ANALYSE MATHÉMATIQUE. — *Sur les zéros de la fonction $\zeta(s)$ de Riemann.* Note de M. G.-H. **Hardy.**

1. MM. H. Bohr et E. Landau ont donné tout récemment (¹) la démonstration que la plupart des zéros complexes de $\zeta(s)$ sont situés, quel que soit δ positif, dans le domaine $\frac{1}{2} - \delta < \sigma < \frac{1}{2} + \delta$. Je me propose maintenant de démontrer que, *parmi les zéros de $\zeta(s)$, il y en a une infinité sur la droite* $\sigma = \frac{1}{2}$ (²).

Je pars d'une formule connue de M. Cahen (³), savoir

$$e^{-y} = \frac{1}{2\pi i}\int_{k-i\infty}^{k+i\infty} \Gamma(u)\, y^{-u}\, du \qquad [\Re(y) > 0,\ k > 0];$$

d'où l'on déduit immédiatement

$$1 + \frac{1}{\pi i}\int_{k-i\infty}^{k+i\infty} \Gamma(u)\, y^{-u}\, \zeta(2u)\, du = 1 + 2\sum_{1}^{\infty} e^{-n^2 y} \qquad \left(k > \frac{1}{2}\right).$$

Je prends maintenant pour chemin d'intégration la droite $\sigma = \frac{1}{4}$. En faisant application du théorème de Cauchy et des formules de Riemann

$$\frac{s(s-1)}{2}\Gamma\left(\frac{s}{2}\right)\pi^{-\frac{s}{2}}\zeta(s) = \xi(s) = \xi\left(\frac{1}{2} + ti\right) = \Xi(t),$$

où $\Xi(t)$ est réelle pour t réel, on est conduit à l'équation

$$(1) \qquad 1 + \sqrt{\frac{\pi}{y}} - \frac{2}{\pi}\int_{-\infty}^{\infty}\left(\frac{\pi}{y}\right)^{\frac{1}{4}+ti} \frac{\Xi(2t)}{\frac{1}{4} + 4t^2}\, dt = 1 + 2\sum_{1}^{\infty} e^{-n^2 y}.$$

Dans cette équation, je pose $y = \pi e^{i\alpha}$, où $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$, et

$$e^{-y} = e^{-\pi\cos\alpha - i\pi\sin\alpha} = e^{\pi i\tau} = q = \rho e^{i\Phi};$$

(¹) *Comptes rendus*, 12 janvier 1914.

(²) J'ai communiqué déjà ce résultat à la Société mathématique de Londres (séance du 12 mars 1914).

(³) Thèse (*Annales École Normale supérieure*, 1894, p. 99). Cette formule a été retrouvée par M. Mellin (*Acta Soc. Fennicæ*, t. XX, nº 7, 1895, p. 6) qui en a fait des applications intéressantes.

1914, 1 *Comptes Rendus*, 158, 1012–14.

et j'obtiens la formule

$$(2)\qquad \int_0^\infty \frac{(e^{\alpha t}+e^{-\alpha t})\,\Xi(2t)}{\frac{1}{4}+4t^2}\,dt = \pi\cos\frac{1}{4}\alpha - \frac{1}{2}\pi e^{\frac{1}{i}i\alpha}\,F(q),$$

où

$$F(q) = 1 + 2\sum_1^\infty q^{n^2} = \vartheta_3(0, \tau).$$

Enfin, en différentiant $2p$ fois par rapport à α, on a

$$(3)\qquad \int_0^\infty \frac{(e^{\alpha t}+e^{-\alpha t})\,t^{2p}\,\Xi(2t)}{\frac{1}{4}+4t^2}\,dt = \frac{(-1)^p\pi}{4^{2p}}\cos\frac{1}{4}\alpha - \left(\frac{d}{d\alpha}\right)^{2p}\left[\frac{1}{2}\pi e^{\frac{1}{i}i\alpha}F(q)\right]$$

2. Je vais me servir maintenant d'un lemme tiré de la théorie des fonctions elliptiques. Je suppose que α tende vers $\frac{1}{2}\pi$, de sorte que q tende vers -1 suivant un chemin tangent au rayon $\Phi = \pi$. Cela étant, je dis que le dernier terme de l'équation (3) tend, quel que soit p, vers la limite zéro. Pour cela, il suffit évidemment que toutes les fonctions

$$\left(\frac{d}{dy}\right)^{2p} F(q) = 2\sum_1^\infty n^{4p} q^{n^2}$$

tendent vers zéro. Mais cette dernière proposition se déduit comme corollaire des théorèmes généraux qu'on doit à MM. Bohr et Marcel Riesz, au sujet de la sommabilité des séries de Dirichlet.

La série

$$1^{-s}+0+0-4^{-s}+0+0+0+0+9^{-s}+0+\ldots,$$

convergente pour $\sigma > 0$, représente la fonction

$$(1-2^{1-2s})\,\zeta(2s),$$

régulière dans tout le plan et d'ordre fini dans tout demi-plan $\sigma > \sigma_0$. La série est donc sommable, pour toute valeur de s, par les moyennes de Cesàro d'ordre assez élevé; et pour s entier négatif, elle a la somme

$$(1-2^{1-2s})\,\zeta(2s) = 0.$$

3. Il s'ensuit que, quand α tend vers $\frac{1}{2}\pi$, l'intégrale (3) tend vers la limite $\frac{(-1)^p\pi}{4^{2p}}\cos\frac{1}{8}\pi$. Supposons maintenant que $\Xi(2t)$ garde un signe

pour $t > T > 1$, par exemple le signe positif. Alors on a, par un théorème connu,

$$(4) \qquad \int_0^\infty \frac{\left(e^{\frac{1}{2}\pi t} + e^{-\frac{1}{2}\pi t}\right) t^{2p}\, \Xi(2t)}{\frac{1}{4} + 4t^2}\, dt = \frac{(-1)^p \pi}{4^{2p}} \cos\frac{1}{8}\pi.$$

Soit p impair. On a

$$(5) \qquad \int_T^\infty < -\int_0^T < KT^{2p},$$

où K est indépendant de p. Mais cela est impossible. Il y a en effet, d'après notre hypothèse, un nombre δ positif tel que $\Xi(2t) > \delta$, pour $2T < t < 2T + 1$. Donc

$$(6) \qquad \int_T^\infty > \int_{2T}^{2T+1} > \delta K_1 (2T)^{2p},$$

où K_1, comme K, est positif et ne dépend nullement de p. Enfin, des inégalités (5) et (6) je tire

$$\delta K_1 2^{2p} < K;$$

donc, pour p assez grand, une contradiction.

COMMENTS

See comments following 1921, 2.

NEW PROOFS OF THE PRIME-NUMBER THEOREM AND SIMILAR THEOREMS.

By G. H. HARDY and J. E. LITTLEWOOD.

1. OUR object in writing this paper is to give a short sketch of a new method which we have found for the proof of certain fundamental theorems in the Analytic Theory of Numbers. A fuller account of our researches will be published elsewhere.*

Our method depends upon the use of the formula

$$(1.1) \qquad e^{-y} = \frac{1}{2\pi i} \int_{\kappa - i\infty}^{\kappa + i\infty} \Gamma(s)\, y^{-s}\, ds, \dagger$$

where κ and the real part of y are positive, and y^{-s} has its principal value, in connection with the "Tauberian" theorems proved by us in a recent paper in the *Messenger of Mathematics*.‡ The theorems which we have principally in view are those expressed by the formulæ

$$(1.21) \qquad M(x) = o(x),$$

$$(1.22) \qquad \Sigma \frac{\mu(n)}{n} = 0,$$

$$(1.23) \qquad \psi(x) \sim x,$$

* As part of a memoir, "Contributions to the theory of the Riemann Zeta-function and the theory of the distribution of primes", to appear in the *Acta Mathematica*.

† This formula was first given by Cahen (*Annales de l'École Normale Supérieure*, vol. xi, p. 75). It was found independently by Mellin (*Acta Societatis Fennicae*, vol. xx., No. 7, p. 6), to whom the first rigorous proof is due.

‡ Vol. xliii., p. 134.

1915, 1 (with J. E. Littlewood) *Quarterly Journal of Mathematics*, 46, 215–19.

where $M(x)$, $\mu(n)$, and $\psi(x)$ have their usual meanings. All these theorems are known to be equivalent to* the "Prime-Number Theorem"

$$\Pi(x) \sim \frac{x}{\log x}. \tag{1.24}$$

They will appear here as particular cases of general theorems concerning Dirichlet's series.

2. We begin by stating the following theorem, which is equivalent to Theorems D, E, and F of our paper in the *Messenger of Mathematics* already quoted.

THEOREM A. *Suppose that*

(i) $\lambda_1, \lambda_2, \lambda_3, \ldots$ *is a sequence of numbers satisfying the conditions* $\lambda_1 > 0$, $\lambda_n > \lambda_{n-1}$, $\lambda_n \to \infty$, $\lambda_n/\lambda_{n-1} \to 1$;

(ii) $\alpha \geq 0$;

(iii) a_n *is real and satisfies one or other of the conditions*

$$a_n > -K\lambda_n^{\alpha-1}(\lambda_n - \lambda_{n-1}), \qquad a_n < K\lambda_n^{\alpha-1}(\lambda_n - \lambda_{n-1}),$$

or is complex and of the form

$$O\{\lambda_n^{\alpha-1}(\lambda_n - \lambda_{n-1})\};$$

(iv) *the series* $\qquad f(y) = \Sigma a_n e^{-\lambda_n y}$

is convergent for $y > 0$, *and*

$$f(y) \sim Ay^{-\alpha}$$

as y tends to zero. Then

$$A_n = a_1 + a_2 + \ldots a_n \sim \frac{A\lambda_n^{\alpha}}{\Gamma(1+\alpha)}$$

as n tends to infinity.†

3. We now prove

THEOREM B. *If*

(i) *the series* $\Sigma a_n \lambda_n^{-s}$ *is absolutely convergent for* $\sigma > \sigma_0 > 0$;

(ii) *the function* $F(s)$ *defined by the series is regular for* $\sigma > c$, *where* $0 < c \leq \sigma_0$, *and continuous throughout any finite part of the plane for which* $\sigma \geq c$;

* By this we mean that, from any one of them, the rest can be deduced by elementary reasoning which involves no appeal to the theory of functions of a complex variable.

† When $A = 0$, the last two formulæ must be interpreted as $f(y) = o(y^{-\alpha})$ and $A_n = o(\lambda_n^{\alpha})$ respectively.

(iii) $$F(s) = O(e^{C|t|}),$$
where $C < \frac{1}{2}\pi$, *uniformly for* $\sigma \geq c$:

then the series $$f(y) = \Sigma a_n e^{-\lambda_n y}$$
is convergent for all positive values of y, *and*
$$f(y) = o(y^{-c})$$
as y *tends to zero.*

Suppose that $y > 0$ and $\kappa > \sigma_0$. Then

$$e^{-\lambda_n y} = \frac{1}{2\pi i}\int_{\kappa - i\infty}^{\kappa + i\infty} \Gamma(s)(\lambda_n y)^{-s}\, ds; \tag{3.1}$$

and it is easy to see that we may multiply by a_n and sum with respect to n. We thus obtain

$$f(y) = \frac{1}{2\pi i}\int_{\kappa - i\infty}^{\kappa + i\infty} \Gamma(s) F(s)\, y^{-s}\, ds. \tag{3.2}$$

An application of Cauchy's theorem enables us to replace this equation by

$$\begin{aligned} f(y) &= \frac{1}{2\pi i}\int_{c - i\infty}^{c + i\infty} \Gamma(s) F(s)\, y^{-s}\, ds \\ &= \frac{y^{-c}}{2\pi}\int_{-\infty}^{\infty} \Gamma(c+it) F(c+it)\, y^{-it}\, dt \\ &= \frac{y^{-c}}{2\pi}\int_{-\infty}^{\infty} \Phi(c+it)\, y^{-it}\, dt, \end{aligned} \tag{3.3}$$

say. But the integral

$$\int_{-\infty}^{\infty} |\Phi(c+it)|\, dt$$

is convergent, so that the integrals

$$\int_{-\infty}^{\infty} \Phi(c+it) \begin{matrix}\cos \\ \sin\end{matrix} (t \log y)\, dt,$$

tend to zero as $\log y$ tends, positively or negatively, to infinity.* The theorem now follows from (3.3).

We have supposed λ_n subject to the conditions (i) of Theorem A. The condition

$$\lambda_n / \lambda_{n-1} \to 1$$

is unnecessary here, but is necessary in the next theorem.

* Cf. Landau, *Prace Matematyczno-Fizycznch*, vol. xxi., pp. 173 *et seq.*

Combining Theorems A and B, we obtain

THEOREM C. *If the conditions of Theorem B are satisfied, and a_n is real and satisfies one or other of the inequalities*

$$a_n > -K\lambda_n^{c-1}(\lambda_n - \lambda_{n-1}), \quad a_n < K\lambda_n^{c-1}(\lambda_n - \lambda_{n-1}),$$

or is complex and of the form

$$O\{\lambda_n^{c-1}(\lambda_n - \lambda_{n-1})\},$$

then $$A_n = a_1 + a_1 + \ldots + a_n = o(\lambda_n^c).$$

Suppose in particular that

$$\lambda_n = n, \quad a_n = \mu(n), \quad c = 1.$$

Then $$F(s) = \Sigma \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)},$$

and all the conditions of Theorem C are satisfied. It follows that

$$M(n) = \mu(1) + \mu(2) + \ldots + \mu(n) = o(n);$$

i.e. that (1.21) is true.

4. In order to obtain a direct proof of (1.22) we must modify Theorem C in such a way that it shall apply to the case in which $c = 0$.

THEOREM D. *Suppose that*

(i) *the conditions of Theorem C are satisfied, except that $c = 0$;*

(ii) *$F(s)$ is regular for $s = 0$.*

Then the series Σa_n is convergent, and has the sum $F(0)$.

The proof is very much the same as that of Theorem C. Suppose that $F(s)$ is regular for $|s| \leq \delta$. Then, instead of (3.3), we have

$$(4.1) \qquad f(y) = F(0) + \frac{1}{2\pi i}\int \Gamma(s) F(s) y^{-s} ds,$$

the contour of integration consisting of the parts $(-i\infty, -i\delta)$ and $(i\delta, i\infty)$ of the imaginary axis, and a semicircle γ described on and to the left of the line $(-i\delta, i\delta)$. We show, substantially as in the proof of Theorem C, that the rectilinear parts of the integral tend to zero. Also

$$\begin{aligned}\int_\gamma \Gamma(s) F(s) y^{-s} ds &= \frac{\Gamma(i\delta) F(i\delta) y^{-i\delta} - \Gamma(-i\delta) F(-i\delta) y^{i\delta}}{\log(1/y)} \\ &\quad - \frac{1}{\log(1/y)}\int_\gamma y^{-s} \frac{d}{ds}\{\Gamma(s) F(s)\} ds \\ &= O\left\{\frac{1}{\log(1/y)}\right\} = o(1).\end{aligned}$$

Thus $$f(y) \to F(0)$$
as $y \to 0$, and so $$\Sigma a_n = F(0).$$

The conditions of this theorem are satisfied, for example, when

$$\lambda_n = n, \quad a_n = \frac{\mu(n)}{n}, \quad c = 0,$$

$$F(s) = \frac{1}{\zeta(s+1)}, \quad F(0) = 0;$$

and (1.22) is a corollary.

5. In order to prove (1.23), a slightly different modification of Theorem C is required.

THEOREM E. *Suppose that the conditions of Theorem C are satisfied, except that $F(s)$ is analytic near $s=c$, and has there a simple pole with residue g. Then*

$$A_n = a_1 + a_2 + \ldots + a_n \sim (g/c)\,\lambda_n^{\,c}.$$

The formula (3.3) is in this case replaced by

$$(5.1) \quad f(y) = g\,\Gamma(c)\,y^{-c} + \frac{1}{2\pi i}\int \Gamma(s)\,F(s)\,y^{-s}\,ds,$$

the path of integration being a modification of the line $\sigma = c$ similar to that of the imaginary axis used in the proof of Theorem D. Practically the same argument as was used in the last proof gives the result

$$f(y) \sim g\,\Gamma(c)\,y^{-c},$$

and the theorem follows immediately. If we take

$$\lambda_n = n, \quad a_n = \Lambda(n), \quad c = 1, \quad F(s) = -\zeta'(s)/\zeta(s), \quad g = 1,$$

we obtain (1.23).

We may add in conclusion that the truth of Theorem B does not really depend on the condition $C < \frac{1}{2}\pi$, which may be removed by a modification of the argument. This is, however, immaterial for the applications which we have had in view.

COMMENTS

See 1918, 1 (§ 2.1), and comments thereon.

Prime Numbers

By G. H. HARDY, F.R.S.

(Ordered by the General Committee to be printed *in extenso.*)

THE Theory of Numbers has always been regarded as one of the most obviously useless branches of Pure Mathematics. The accusation is one against which there is no valid defence ; and it is never more just than when directed against the parts of the theory which are more particularly concerned with primes. A science is said to be useful if its development tends to accentuate the existing inequalities in the distribution of wealth, or more directly promotes the destruction of human life. The theory of prime numbers satisfies no such criteria. Those who pursue it will, if they are wise, make no attempt to justify their interest in a subject so trivial and so remote, and will console themselves with the thought that the greatest mathematicians of all ages have found in it a mysterious attraction impossible to resist.

The foundations of the theory were laid by Euclid. Among Euclid's theorems two in particular are of fundamental importance. The first (Euc. vii. 24) is that *if a and b are both prime to c, then ab is also prime to c.* This theorem is the basis of the whole theory of the factorisation of numbers, systematised later by Euler and by Gauss, and in particular of the theorem that *every number can be expressed in one and only one way as a product of primes.* The second theorem (Euc. ix. 20) is that *the number of primes is infinite :* to this theorem I shall return in a moment.

In modern times the theory has developed in two different directions. In the first place there is what may be called roughly the theory of *individual* or *isolated* primes, a theory which it is difficult to define precisely, but of which a general idea may be formed by considering a few of its characteristic problems. How can we determine whether a given number is prime ? what conditions are necessary and what sufficient ? Can we define forms which represent prime numbers only ? Are there infinitely many pairs of primes which differ by 2 ? Is (as Goldbach asserted) every even number the sum of two primes ? This theory I shall dismiss very briefly. We know a number of very beautiful theorems of this character. I need only mention Wilson's theorem, Fermat's theorem, and the extensions of the latter by Lucas. But on the whole the record of research in this direction is a record of failure. The difficulties are too great for the methods of analysis at our command, and the problems remain unsolved.

Very different results are revealed when we turn to the second principal branch of the modern theory, the theory of the *average or asymptotic distribution of primes.* This theory (though one of its most famous problems is still unsolved) is in some ways almost complete, and certainly represents one of the most remarkable triumphs of modern analysis. The theory centres round one theorem, the *Primzahlsatz* or *Prime Number Theorem* ; and it is to the history of this theorem, which may almost be said to embody the history of the whole subject, that I shall devote the remainder of this lecture.*

The problem may be stated crudely as follows : *How many primes*

* A full account of the history of the theorem will be found in Landau's *Handbuch der Lehre von der Verteilung der Primzahlen* (Teubner, 1909).

1915, 10 *British Association Report*, 350–4.

are there less than a given number x? More precisely, let $\pi(x)$ denote the number of primes * not exceeding x: then *what is the order of magnitude of $\pi(x)$?* The Prime Number Theorem provides a complete answer to this last question. It asserts that

$$\pi(x) \backsim \frac{x}{\log x},$$

that is to say, that *$\pi(x)$ and $x/(\log x)$ are asymptotically equivalent,* or that their ratio tends to 1 when x tends to infinity.

The first step towards the proof of this theorem was made by Euclid, when he proved that the number of primes is infinite, or that

$$\pi(x) \rightarrow \infty.$$

Euclid's proof is classical, and can hardly be repeated too often. If the number of primes is finite, let them be 2, 3, 5, . . ., P. The number 2. 3. 5. . . . P + 1 is not divisible by any of 2, 3, 5, ..., P. It is therefore prime itself, or divisible by some prime greater than P; and either alternative contradicts the hypothesis that P is the greatest prime. It is worth remarking that Euclid's reasoning may be used to prove rather more, viz. that the order of $\pi(x)$ is at least as great as that of $\log \log x$.†

The next advances were made by Euler, probably about 1740. It was Euler to whom we owe the introduction into analysis of the Zeta-function, the function on whose properties, as later research has shown, the whole theory depends.

Let $s = \sigma + it$. Then the function $\zeta(s)$ is defined, when $\sigma > 1$, by the equations

$$\zeta(s) = \Sigma n^{-s} = 1^{-s} + 2^{-s} + 3^{-s} + \ldots;$$

and Euler's fundamental contribution to the theory is the formula

$$\zeta(s) = \Pi\left(\frac{1}{1-p^{-s}}\right),$$

where the product extends over all prime values of p. Euler, it is true, considered $\zeta(s)$ as a function of a *real* variable only. But his formula at once indicates the existence of a deep-lying connection between the theory of $\zeta(s)$ and the theory of primes.

Euler deduced from his formula that the series Σp^{-s}, obviously convergent when $s > 1$, is divergent when $s = 1$; and from this it is easy to deduce important consequences as to the order of $\pi(x)$. It is evident that $\pi(x) < x$, so that the order of $\pi(x)$ certainly does not exceed that of x, or, in the notation which is usual now, $\pi(x) = O(x)$.‡ It is an easy corollary of Euler's result that the order of $\pi(x)$ is *not very much less than that of x*; that, for example, $\pi(x) \neq O(x^a)$ for any value of a less than 1; or again, more precisely, that

$$\pi(x) \neq O\left\{\frac{x}{(\log x)^{1+a}}\right\}$$

for any value of a greater than 1.

* It proves most convenient not to count 1 as a prime.

† This was pointed out to me by Prof. H. Bohr of Copenhagen.

‡ $f = O(\phi)$ means that the absolute value of f is less than a constant multiple of ϕ: thus $\sin x = O(1)$, $100x = O(x)$.

It is also easy to prove that the order of $\pi(x)$ is *definitely less than that of* x, or that, as we should express it now, $\pi(x)=o(x)$.* This theorem, when read in conjunction with those which precede, is, I think, enough to suggest the Prime Number Theorem as a very plausible conjecture, or at any rate to suggest that the true order is that of $x/(\log x)$. The theorem was in fact conjectured first by Gauss (1793) and by Legendre (1798); and it is in Legendre's *Essai sur la théorie des nombres* that the conjecture first appears in print.

In this state the problem remained for fifty years, until the publication (1849–1852) of the researches of the Russian mathematician Tschebyschef. I have no time to speak of Tschebyschef's work as fully as it deserves, but his chief results, in so far as they bear directly on the problem now before us, were as follows:—

(1) Tschebyschef showed that the problem is simplified if we take as fundamental not the function $\pi(x)$ itself, but the closely related function

$$\theta(x)=\sum_{p\leq x}\log p$$

(the sum of the logarithms of all primes not exceeding x). He showed that the order of $\theta(x)$ is the same as that of $\pi(x)\log x$, and that the Prime Number Theorem itself is equivalent to the theorem that

$$\theta(x)\sim x.$$

(2) He showed that $\theta(x)$ is actually of order x, and $\pi(x)$ of order $x/(\log x)$, in fact that positive constants A and B exist such that

$$A\frac{x}{\log x}<\pi(x)<B\frac{x}{\log x}.$$

(3) He showed that *if* $\theta(x)/x$ *tends to a limit*, then that limit must be unity.

What Tschebyschef could not prove is that the limit does in fact exist, and, as he failed to prove this, he failed to prove the Prime Number Theorem. And about Tschebyschef's methods (interesting as they are), I shall say nothing; for later research has shown that it was the essential inadequacy of his methods which was responsible for his failure, and that the theorem lies deeper in analysis than any of the ideas on which he relied.

The next great step was taken by Riemann in 1859, and it is in Riemann's famous memoir *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* that we first find the ideas upon which the theory has now been shown really to rest. Riemann did not prove the Prime Number Theorem: it is remarkable, indeed, that he never mentions it. His object was a different one, that of finding an explicit expression for $\pi(x)$, or rather for another closely associated function, as a sum of an infinite series. But it was Riemann who first recognised that, if we are to solve any of these problems, we must study the Zeta-function as a function of the *complex* variable $s=\sigma+it$, and in particular study the distribution of its zeros.

* $f=o(\phi)$ means that $f/\phi\to 0$. Thus $\sin x=o(x)$. This theorem also was stated by Euler, but without satisfactory proof.

Riemann proved

(1) that $\zeta(s)$ is an analytic function of s, regular all over the plane except for a simple pole at the point 1;

(2) that $\zeta(s)$ satisfies the functional equation

$$\zeta(1-s)=2(2\pi)^{-s}\cos\tfrac{1}{2}s\pi\;\Gamma(s)\zeta(s);$$

(3) that $\zeta(s)$ has zeros at the points $-2, -4, -6 \ldots.$, and no other zeros *except possibly complex zeros whose real parts lie between* 0 *and* 1 *inclusive.*

To these propositions he added certain others of which he could produce no satisfactory proof. In particular he asserted that there is in fact an infinity of complex zeros, all naturally situated in the 'critical strip' $0 \leq \sigma < 1$; an assertion now known to be correct. Finally he asserted that it was 'sehr wahrscheinlich' that all these zeros have the real part $\frac{1}{2}$: the notorious 'Riemann hypothesis', unsettled to this day.

We come now to the time when, a hundred years after the conjectures of Gauss and Legendre, the theorem was finally proved. The way was opened by the work of Hadamard on integral transcendental functions. In 1893 Hadamard proved that the complex zeros of Riemann actually exist; and in 1896 he and de la Vallée-Poussin proved independently that *none of them have the real part* 1, and deduced a proof of the Prime Number Theorem.

It is not possible for me now to give an adequate account of the intricate and difficult reasoning by which these theorems are established. But the general ideas which underlie the proofs are, I think, such as should be intelligible to any mathematician.

In the first place Euler's formula shows that $\log \zeta(s)$ behaves, throughout the half-plane $\sigma > 1$, much like the series $\sum p^{-s}$. But $\zeta(s)$ has a simple pole for $s=1$, and so the sum of the series $\sum p^{-1-\delta}$ tends logarithmically to $+\infty$ when $\delta \rightarrow 0$ through positive values. Suppose now that (if possible) $\zeta(1+ti)=0$. Then the real part of $\log \zeta(1+\delta+ti)$, and therefore the real part of the series $\sum p^{-1-\delta-ti}$, tends, also logarithmically, to $-\infty$ when $\delta \rightarrow 0$. It follows that the series

$$\sum p^{-1-\delta},\ -\sum p^{-1-\delta}\cos(t\log p)$$

tend to $+\infty$ *with equal rapidity* when $\delta \rightarrow 0$. As the first series is a series of positive terms, while the signs of the terms in the second series change with a certain regularity, it is natural to suppose that our last conclusion is impossible; and this is in fact not particularly difficult to prove.

I come now to the proof of the Prime Number Theorem itself. If we differentiate Euler's formula logarithmically, we obtain

$$\frac{\zeta'(s)}{\zeta(s)}=\sum\left(\frac{\log p}{p^s}+\frac{\log p}{p^{2s}}+\ldots.\right)=\sum_{p,m}\frac{\log p}{p^{ms}};$$

or (1) $$-\frac{\zeta'(s)}{\zeta(s)}=\sum\frac{\Lambda(n)}{n^s}$$

where p assumes all prime values, m and n all positive integral values, and $\Lambda(n)$ is equal to $\log p$ if n is of the form p^m and to zero otherwise. Let

$$\psi(x) = \sum_{n \leq x} \Lambda(n)$$

Then $\psi(x)$ is, for our present purpose, equivalent to $\theta(x)$: it is easy to show that the difference between the two functions is of order $\sqrt{x}$. We have therefore to prove that $\psi(x) \sim x$.

The series on the right-hand side of the equation (1) is what is called a 'Dirichlet's series'; and the theory of such series resembles the more familiar theory of Taylor's series in one very important respect. *We can express the coefficients by contour integrals in which the function represented by the series appears under the sign of integration.* In particular we can show that

$$(2) \qquad \psi(x) = -\frac{1}{2\pi i}\int \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds,$$

where the path of integration is a line parallel to the imaginary axis and passing to the right of the point $s=1$.

The general idea of the proof is now easy enough to grasp. Every element of the integral (2) is of order x^σ, where $\sigma > 1$: we can therefore draw no *direct* conclusion as to the behaviour of $\psi(x)$ when x is large. But it is at once suggested that we should try to make use of Cauchy's theorem. The subject of integration has a simple pole at the point 1, corresponding to the pole of $\zeta(s)$ itself, and the residue at the pole is precisely x; and there are no other singularities on the line $\sigma = 1$, since $\zeta(s)$, as we have seen, has no poles *or zeros* on that line. Suppose then that we can move the path of integration across to the left of the line, introducing the appropriate correction due to the pole. Plainly we shall then have an expression for $\psi(x) - x$ in the form of an integral *in which every element is of order less than that of* x. And if we can show that the same is true of the integral itself, we shall have proved that $\psi(x) \sim x$, that is to say, we shall have proved the Prime Number Theorem. It will be observed that, if $\zeta(s)$ had zeros whose real part is equal to 1, then the result would be definitely false, since there would be additional residues of order x. It thus becomes clear why the older attempts to prove the theorem, without using the theory of functions of a complex variable, were unsuccessful.

The arguments which I have advanced are not exact: I have merely put forward a chain of reasoning which seems likely to lead to the desired result. The achievement of Hadamard and de la Vallée-Poussin was to replace these plausibilities by rigorous proofs. It might be difficult for me to make clear to you how great this achievement was. Some branches of pure mathematics have the pleasant characteristic that what seems plausible at first sight is generally true. In this theory anyone can make plausible conjectures, and they are almost always false. Nothing short of absolute rigour counts; and it is for this reason that the Analytic Theory of Numbers, while hardly a subject for an amateur, provides the finest possible discipline in accurate reasoning for anyone who will make a real effort to understand its results.

COMMENTS

See 1918, 1 (§ 2.1), and comments thereon.

ON THE DIFFERENCE $\pi(x) - \operatorname{li} x$ (II)

By S. SKEWES

[Received 31 December 1953.—Read 21 January 1954]

Introduction

1. Let $\pi(x)$ denote, as usual, the number of primes less than or equal to x which we suppose always to be not less than 2, and let

$$\operatorname{li} x = \lim_{\epsilon \to 0} \left(\int_0^{1-\epsilon} + \int_{1+\epsilon}^{x} \right) \frac{du}{\log u}.$$

The difference $d(x) = \pi(x) - \operatorname{li} x$ is negative for all values of x up to 10^7, and for all the special values of x for which $\pi(x)$ has been calculated (e.g. $d(x) = -1757$ for $x = 10^9$). Littlewood (**1**) proved in 1914, however, that $d(x)$ changes sign infinitely often, and in particular there exists an X such that $d(x) > 0$ for some $x < X$. This last result is our present subject. Littlewood's method depends on an 'explicit formula', as does all subsequent work, including the present paper.

If θ is the upper bound of the real parts of the zeros $\rho = \beta + i\gamma$ of the Riemann zeta-function $\zeta(s)$, the 'Riemann hypothesis' [(RH) for short] is that $\theta = \frac{1}{2}$; if this is false, then $\frac{1}{2} < \theta \leqslant 1$. In this latter case it had long been known that, for each positive ϵ, $d(x)/x^{\theta-\epsilon}$ oscillates, as x tends to infinity, over a range including ± 1. In proving the mere existence of an X it is therefore permissible to assume (RH), and Littlewood naturally did this.

Littlewood's theorem is a 'pure existence theorem', and does not provide, even when (RH) is assumed, an explicit numerical X.

When we face the problem of a numerical X, free of hypotheses, the argument falls naturally into three stages.

(i) A new method is found which assumes (RH) and provides a numerical $X = X_1$. I gave such a method in 1933 (**3**). In the meantime Ingham (**4**) has developed an alternative method (which he applies to the more general problem of the infinity of changes of sign of $d(x)$). This, adapted to our more special case (one change of sign) and with some further modifications, gives a much better X_1 than my original paper did; the argument is given in full in Part I. One of the advantages of the new method is that we can largely eliminate the ρ's beyond a given point; we operate in fact with the 269 ρ's with $0 < \gamma < 500$, whose position is approximately known.

(ii) (This is easy.) The whole argument in (i) is based on the explicit formula for $\psi_0(x) = \frac{1}{2}\{\psi(x+0)+\psi(x-0)\}$ (in the usual notation of the subject) (**2**). This is

$$\psi_0(x)-x = -\sum_{\rho}\frac{x^{\rho}}{\rho}-\frac{\zeta'(0)}{\zeta(0)}-\tfrac{1}{2}\log\left(1-\frac{1}{x^2}\right) \quad \text{for } x > 1.$$

In the course of the proof the terms of the series $\sum x^{\rho}/\rho$ with $|\gamma| \geqslant G = X_1^3$ can (roughly) be rejected as negligible, (RH) *or not*. It is enough, in other words, to suppose, instead of (RH), only that $\beta = \frac{1}{2}$ for those γ's satisfying $|\gamma| < G$. This hypothesis can in turn be weakened; for $x < X_1$, the $|x^{\beta+i\gamma}|$ concerned differ from $|x^{\frac{1}{2}+i\gamma}|$ by something negligible, provided the β's concerned satisfy

$$b = \beta-\tfrac{1}{2} \leqslant B = X_1^{-3}\log^{-2}X_1.$$

With minor adjustments, then, the proof in (i) can be made to provide an X_1 [actual value $\exp\exp\exp(7{\cdot}703)$] subject only to the double modification of (RH) explained above. This modification, which we will call (H), is, to repeat,

(H) *Every zero $\rho = \beta+i\gamma$ for which $|\gamma| < G = X_1^3$ is such that*

$$b = \beta-\tfrac{1}{2} \leqslant B = X_1^{-3}\log^{-2}X_1.$$

(iii) Since (H) leads to an X_1, it remains only to show that (NH), the negation of (H), leads to an X_2, i.e. that $d(x) > 0$ for some $x < X_2$. Now (NH) asserts the existence of a $\rho = \rho_0 = \beta_0+i\gamma_0$ with

$$0 < \gamma_0 < G = X_1^3, \qquad b_0 = \beta_0-\tfrac{1}{2} > B,$$

where

$$B = X_1^{-3}\log^{-2}X_1;$$

that is, it provides a more or less given ρ to the right of $\sigma = \frac{1}{2}$. In particular, it asserts that $\theta > \frac{1}{2}+B$, in which case an X_2 certainly *exists* in virtue of the old theorem about $d(x) > x^{\theta-\epsilon}$. It is natural to expect further that the proof of that theorem could use the existence of the special ρ_0 to provide a numerical X_2. But this turns out not to be so; the proof in question is another 'pure existence' one. Some further idea is called for, and I am in fact indebted to Professor Littlewood for the sketch of a method for the simpler problem of finding an X such that, for a given $h > 0$ and for some $x < X$, $\psi(x)-x > h\sqrt{x}$.

There is now a last unexpected point. In the past it has always been possible to work with the function $\psi(x)$ and its simpler explicit formula, with only a last minute switch, on established lines, to $\pi(x)$. But with (NH) this is no longer possible, and it is necessary to work, in finding X_2, with $\Pi_0(x) = \frac{1}{2}\{\Pi(x+0)+\Pi(x-0)\}$, where

$$\Pi(x) = \sum_{p^m \leqslant x}\frac{1}{m} = \sum_{m=1}^{M}\frac{1}{m}\pi(x^{1/m}), \qquad M = [\log x/\log 2].$$

The explicit formula for $\Pi_0(x)$ is, for $x > 1$,

$$\Pi_0(x) = \operatorname{li} x - \sum_{\rho} \operatorname{li} x^{\rho} + \int\limits_{x}^{\infty} \frac{du}{(u^2-1)u\log u} - \log 2.$$

In the actual working out of the paper stages (i) and (ii) are telescoped, and (RH) never appears. In Part I we assume (H) from the first, and arrive at an $X_1 = \exp\exp\exp(7{\cdot}703)$. Part II then assumes (NH) and arrives at an X_2 differing negligibly† in expression from e^{X_1}: a (just) permissible X_2 is

$$10^{10^{10^{10^3}}}.$$

I wish in conclusion to express my humble thanks to Professor Littlewood, but for whose patient profanity this paper could never have become fit for publication.

Part I

2. We begin by collecting, in Lemma 1, some results about the zeros ρ. The fundamental theorem underlying all its results is as follows (**5**).

Let $N(T)$ be the number of roots $\rho = \beta + i\gamma$ of the ζ-function satisfying $0 < \beta < 1$, $0 < \gamma \leqslant T$. Then

$$N(T) = \frac{T}{2\pi}\log\frac{T}{2\pi e} + R(T), \tag{1}$$

where $\quad |R(T)| < (0{\cdot}137)\log T + (0{\cdot}443)\log\log T + 4{\cdot}350.$

We make use also of the known values of $\gamma_1, \gamma_2, \ldots, \gamma_{29}$, that is, all the γ's satisfying $0 < \gamma \leqslant 100$. We have now

Lemma 1. *For all $T \geqslant \gamma_1 = 14{\cdot}13\ldots$,*

(i) $\displaystyle\sum_{0<\gamma<T} \frac{1}{\gamma} < \frac{1}{4\pi}\log^2 T,$

(ii) $\displaystyle\sum_{\gamma \geqslant T} \frac{1}{\gamma^2} < \frac{1}{2\pi}\frac{\log T}{T},$

(iii) $\displaystyle\sum_{\gamma>0} \frac{1}{\gamma^2} < 0{\cdot}0233.$

For $|h| < \frac{1}{2}T$,

(iv) $\displaystyle |N(T+h) - N(T)| < \frac{1}{2\pi}(|h| + 1{\cdot}77)\log T + 8{\cdot}7.$

We suppress throughout the details of purely numerical calculations.

† X_1^{100} differs negligibly (in its top index) from X_1. For this and similar reasons some of our approximations can be very crude; only in those bearing on a top index is refinement called for.

We obtain (i) from (1) and the formula†

$$\sum_{0<\gamma_n<T} \frac{1}{\gamma_n} = \sum_{n=1}^{29} \frac{1}{\gamma_n} + \int_{\gamma_{30}}^{T} \frac{N^*(x)}{x^2}\,dx + \frac{N^*(T)}{T},$$

where $N^*(T) = N(T)-29$. Since $\sum_{n=1}^{29} \frac{1}{\gamma_n} < 0{\cdot}5925$, this leads by straightforward calculation to

$$\sum_{0<\gamma<T} \frac{1}{\gamma} = \frac{1}{4\pi}\log^2 T - \frac{\log 2\pi}{2\pi}\log T + R_1(T),$$

where $|R_1(T)| < 0{\cdot}312$. This leads at once to (i).

We obtain (ii) similarly, from (1) and the formula†

$$\sum_{\gamma\geqslant T} \frac{1}{\gamma^2} = \int_T^\infty \frac{2N(x)}{x^3}\,dx - \frac{N(T)}{T^2}.$$

Of the remaining results (iii) is known, and (iv) follows at once from (1).

3. Lemma 2. *Let $\psi_0(x)$ be defined, as usual, by*

$$\psi_0(x) = \tfrac{1}{2}\{\psi(x+0)+\psi(x-0)\},$$

where $\psi(x) = \sum_{n\leqslant x} \Lambda(n)$. For $x > 1$, $\psi_0(x)$ is known to possess the explicit formula‡

$$\psi_0(x)-x = -\sum_\rho \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \tfrac{1}{2}\log\left(1-\frac{1}{x^2}\right), \tag{2}$$

where $\sum_\rho \frac{x^\rho}{\rho}$ is defined as the limit of $\sum_{|\gamma|\leqslant T} \frac{x^\rho}{\rho}$ as $T\to\infty$. If

$$\sum_\rho \frac{x^\rho}{\rho} = \sum_{|\gamma|\leqslant T} \frac{x^\rho}{\rho} + R(x,T),$$

then

(i) $|R(x,T)| < 1000\dfrac{x^3}{x-1}\dfrac{\log^2 T}{T} + 3\log x \quad (x\geqslant e,\ T\geqslant 9)$;

(ii) $|R(x,T)| < (0{\cdot}0001)x^{\frac{1}{2}} \quad (x\geqslant \exp(10^4),\ T\geqslant x^2)$;

(iii) $\left|\sum_\rho \dfrac{x^\rho}{\rho}\right| < 3x\log x \quad (x\geqslant e)$.

The proof of (i) proceeds by straightforward calculation on the lines of the proof of (2); (ii) follows from (i); and (iii) follows from (2) and the definitions of $\psi_0(x)$ and $\psi(x)$, since

$$\sum_{n\leqslant x} \Lambda(n) \leqslant x\log x \quad \text{and} \quad |\psi_0(x)-\psi(x)| \leqslant \tfrac{1}{2}\log x.$$

† (**2**), 18, Theorem A. ‡ (**2**), 77, Theorem 29.

4. We shall for the present assume the following hypothesis, which we call (H), about the zeros $\rho = \beta+i\gamma$.

(H) *Let* $X_1 = \exp\exp\exp(7{\cdot}703)$, $G = X_1^3$, $B = X_1^{-3}\log^{-2}X_1$. *Then for every zero* ρ *such that* $|\gamma| < G$, β *satisfies*

$$b = \beta-\tfrac{1}{2} \leqslant B.$$

For reference we shall prefix (H) to those results which depend on the hypothesis (H).

(H) LEMMA 3. *Let* $\psi_1(x)$ *be defined, as usual, by*

$$\psi_1(x) = \int\limits_1^x \psi(u)\,du = \sum_{n\leqslant x}(x-n)\Lambda(n).$$

Then, on hypothesis (H),

(i) $|\psi_1(x)-\frac{1}{2}x^2| < 8x \quad (2 < x \leqslant e^8)$;

(ii) $|\psi_1(x)-\frac{1}{2}x^2| < \frac{1}{4}x^{\frac{3}{2}} \quad (e^8 < x < X_1)$.

For $x \geqslant 1$ we have the formula†

$$\psi_1(x)-\tfrac{1}{2}x^2 = -\sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)} - x\frac{\zeta'(0)}{\zeta(0)} + \frac{\zeta'(-1)}{\zeta(-1)} - \sum_{r=1}^{\infty}\frac{x^{1-2r}}{2r(2r-1)}. \tag{3}$$

From Lemma 1 (ii) and (iii), and assuming (H), we have, for $x < X_1$,

$$\begin{aligned} x^{-\frac{3}{2}}\left|\sum_\rho \frac{x^{\rho+1}}{\rho(\rho+1)}\right| &\leqslant \sum_{|\gamma|<G}\left|\frac{x^{\rho-\frac{1}{2}}}{\rho(\rho+1)}\right| + \sum_{|\gamma|\geqslant G}\left|\frac{x^{\rho-\frac{1}{2}}}{\rho(\rho+1)}\right| \\ &\leqslant X_1^B \sum_{|\gamma|<G}\frac{1}{\gamma^2} + X_1^{\frac{1}{2}}\frac{1}{\pi}\frac{\log G}{G} \\ &\leqslant \tfrac{1}{10}. \end{aligned}$$

Substituting in (3) and noting that

$$\zeta'(0)/\zeta(0) = \log 2\pi \quad \text{and} \quad |\zeta'(-1)/\zeta(-1)| < 1,$$

we obtain both (i) and (ii).

5. LEMMA 4.

(i) $\operatorname{li} u < (1{\cdot}0004)u/\log u \quad (u \geqslant \exp(4.10^3))$;

(ii) $\operatorname{li} u < 2u/\log u \quad (u \geqslant 2)$.

The value of li 2 is 1·04... . For $u > u_0 \geqslant e^2$, say,

$$\frac{1}{\log u} < \frac{d}{du}\left(\frac{u}{\log u}\right)\Big/\left(1-\frac{1}{\log u_0}\right).$$

† (2), 73, Theorem 28.

Hence
$$\mathrm{li}\, u = \mathrm{li}\, u_0 + \int\limits_{u_0}^{u} \frac{dv}{\log v}$$
$$< u_0 + \frac{1}{1-1/\log u_0}\left[\frac{v}{\log v}\right]_{u_0}^{u},$$
and the result (i) follows by taking $u_0 = \exp(3.10^3)$. By taking $u_0 = e^2$, we find that (ii) is valid for $u \geqslant 8$, say, and, for $2 \leqslant u < 8$, (ii) is trivial.

6. We define $\Pi(x)$, as usual, by
$$\Pi(x) = \sum_{m=1}^{M} \frac{1}{m}\pi(x^{1/m}), \qquad M = [\log x/\log 2].$$

LEMMA 5. *For $x > \exp(4.10^3)$, either $\pi(\xi) > \mathrm{li}\,\xi$ for some ξ of $2 \leqslant \xi \leqslant x^{\frac{1}{2}}$, or else*
$$0 < \Pi(x)-\pi(x) < (1{\cdot}0005)x^{\frac{1}{2}}/\log x.$$

Supposing the former alternative to be false, we apply Lemma 4 (i) to the first term on the right-hand side of
$$\Pi(x)-\pi(x) = \tfrac{1}{2}\pi(x^{\frac{1}{2}}) + \sum_{m=3}^{M} \frac{1}{m}\pi(x^{1/m}),$$
and Lemma 4 (ii) to the remainder. Then
$$\Pi(x)-\pi(x) \leqslant \tfrac{1}{2}\,\mathrm{li}\, x^{\frac{1}{2}} + 2\sum_{m=3}^{M} x^{1/m}/\log x$$
$$< (1{\cdot}0004)x^{\frac{1}{2}}/\log x + (0{\cdot}0001)x^{\frac{1}{2}}/\log x,$$
and the desired result follows.

7. (H) LEMMA 6. *Let $P(x)$ be defined by*
$$P(x) = (\Pi(x)-\mathrm{li}\, x)-(\psi(x)-x)/\log x.$$
Then, on hypothesis (H),
$$|P(x)| < (0{\cdot}0005)x^{\frac{1}{2}}/\log x \quad (\exp(10^4) \leqslant x < X_1).$$

We have [(**2**), 64]
$$P(x) = \int\limits_{2}^{x} \frac{\psi(u)-u}{u\log^2 u}\,du + \frac{2}{\log 2} - \mathrm{li}\, 2,$$
and therefore, after integrating by parts,
$$|P(x)| \leqslant \left|\frac{\psi_1(x)-\frac{1}{2}x^2}{x\log^2 x} - \frac{\psi_1(2)-2}{2\log^2 2} + \frac{2}{\log 2} - \mathrm{li}\, 2\right| + |J|, \tag{4}$$
where
$$J = \int\limits_{2}^{x} \{\psi_1(u)-\tfrac{1}{2}u^2\}\, d\left\{\frac{1}{u\log^2 u}\right\}.$$

Now $$|J| \leqslant \int_2^x |\psi_1(u)-\tfrac{1}{2}u^2| \frac{4}{u^2\log^2 u}\,du = \int_2^{e^8}+\int_{e^8}^x.$$

Now apply Lemma 3. We have, on (H),

$$|J| < \int_2^{e^8} \frac{32}{u\log^2 u}\,du + \int_{e^8}^x \left(\frac{u^{\frac{1}{4}}}{\log^2 u}\right)\frac{1}{u^{\frac{3}{4}}}\,du.$$

Since $u^{\frac{1}{4}}/\log^2 u$ increases with u for $u > e^8$, it follows that, for

$$\exp(10^4) \leqslant x < X_1,$$

$$|J| < \frac{32}{\log 2}+\frac{x^{\frac{1}{4}}}{\log^2 x}4x^{\frac{1}{4}} < 48+0\cdot0004x^{\frac{1}{2}}/\log x.$$

Substituting this inequality in (4) and applying Lemma 3 (ii) to

$$(\psi_1(x)-\tfrac{1}{2}x^2)/x\log^2 x,$$

we obtain the required inequality.

(H) LEMMA 7. *Assume hypothesis* (H). *Then for any given x satisfying* $\exp(10^4) \leqslant x < X_1$, *either*

(i) $\pi(\xi)-\mathrm{li}\,\xi > 0$ *for some ξ of* $2 \leqslant \xi \leqslant x^{\frac{1}{2}}$,

or else

(ii) '$\psi_0(x)-x > (1\cdot001)x^{\frac{1}{2}}$' *implies* '$\pi(x)-\mathrm{li}\,x > 0$'.

(i) is the first alternative of Lemma 5, and (ii) follows from the second one and Lemma 6, since

$$\begin{aligned}(\pi(x)-\mathrm{li}\,x)\log x &= \{\psi(x)-\psi_0(x)\}+\{\psi_0(x)-x\}-\\ &\quad-\{\psi_0(x)-x-(\Pi(x)-\mathrm{li}\,x)\log x\}-\{(\Pi(x)-\pi(x))\log x\}\\ &> 0+(1\cdot001)x^{\frac{1}{2}}-(0\cdot0005)x^{\frac{1}{2}}-(1\cdot0005)x^{\frac{1}{2}} = 0.\end{aligned}$$

8. (H) LEMMA 8. *On hypothesis* (H),

$$\sum_{|\gamma|<G}\left|\frac{x^{\rho-\frac{1}{2}}}{\rho}-\frac{x^{i\gamma}}{i\gamma}\right| < 0\cdot0234 \quad (\exp(10^4) \leqslant x < X_1).$$

For brevity write the series on the left as $S(x)$, and let, as usual,

$$\rho-\tfrac{1}{2} = \beta-\tfrac{1}{2}+i\gamma = b+i\gamma.$$

Then

$$S(x) = \sum_{|\gamma|<G}\left|\frac{x^{b+i\gamma}}{\beta+i\gamma}-\frac{x^{i\gamma}}{i\gamma}\right| \leqslant \Big(\sum_{|\gamma|<G}\frac{1}{\gamma^2}\Big)\max_{|\gamma|<G}\{|\gamma(x^b-1)|+\beta\}.$$

Applying (H) and Lemma 1 (iii), we have therefore

$$\begin{aligned}S(x) &< (0\cdot0466)[G(2B\log X_1)+\tfrac{1}{2}+B]\\ &< 0\cdot0234.\end{aligned}$$

9. We now develop a modification of Ingham's argument. Consider the formula (see (**4**), 204 (6))

$$\int_a^b \chi(x)(\psi_0(x)-x)\,dx$$
$$= -\sum_\rho \frac{1}{\rho}\int_a^b \chi(x)x^\rho\,dx + \int_a^b \chi(x)\{\tfrac{1}{2}\log(1-x^{-2})^{-1}-\zeta'(0)/\zeta(0)\}\,dx, \qquad (5)$$

where $1 < a < b < \infty$, and $\chi(x)$ is any function integrable in the sense of Lebesgue. Let

$$K(y) = \left(\frac{\sin\frac{1}{2}y}{\frac{1}{2}y}\right)^2,$$

so that, for real α,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} K(y)e^{i\alpha y}\,dy = \begin{cases} 1-|\alpha| & (|\alpha| \leqslant 1), \\ 0 & (|\alpha| > 1). \end{cases} \qquad (6)$$

Let $T = 500$ and ω be any number satisfying $\omega > 2.10^4$. In (5) substitute

$$x = e^u, \qquad \chi(e^u) = e^{-\frac{1}{2}u}TK\{T(u-\omega)\}, \qquad a = e^{\frac{1}{2}\omega}, \qquad b = e^{\frac{3}{2}\omega}.$$

Then, writing for brevity

$$F(u) = \{\psi_0(e^u)-e^u\}e^{-\frac{1}{2}u}, \qquad (7)$$

we have

$$\int_{\frac{1}{2}\omega}^{\frac{3}{2}\omega} TK\{T(u-\omega)\}F(u)\,du = -\sum_\rho \frac{1}{\rho}\int_{\frac{1}{2}\omega}^{\frac{3}{2}\omega} TK\{T(u-\omega)\}e^{(\rho-\frac{1}{2})u}\,du + R, \quad (8)$$

where, if we define $r(u)$ by

$$r(u) = e^{-\frac{1}{2}u}\{\tfrac{1}{2}\log(1-e^{-2u})^{-1}-\zeta'(0)/\zeta(0)\},$$

R is given by

$$R = \int_{\frac{1}{2}\omega}^{\frac{3}{2}\omega} TK\{T(u-\omega)\}r(u)\,du.$$

Since $\frac{1}{2}\omega \leqslant u \leqslant \frac{3}{2}\omega$ and $\omega > 2.10^4$, we have

$$|r(u)| < 2e^{-\frac{1}{4}\omega} < 0{\cdot}00001;$$

hence [in virtue of (6), with $\alpha = 0$]

$$|R| < (0{\cdot}00001)\int_{\frac{1}{2}\omega}^{\frac{3}{2}\omega} TK\{T(u-\omega)\}\,du$$
$$= (0{\cdot}00001)\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)\,dy < (0{\cdot}00001)2\pi. \qquad (9)$$

Substituting $u = \omega+y/T$ in (8), we have from (8) and (9)

$$\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)F(\omega+y/T)\,dy = -\sum_{\rho}\frac{1}{\rho}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)e^{(\rho-\frac{1}{2})(\omega+y/T)}\,dy + R, \quad (10)$$

where $|R| < (0{\cdot}00002)\pi$.

For the infinite series on the right-hand side of (10) we shall substitute the finite series

$$\sum_{|\gamma|<G}\frac{e^{i\gamma\omega}}{i\gamma}\int_{-\infty}^{\infty} K(y)e^{i\gamma y/T}\,dy,$$

where G is the number defined in hypothesis (H), § 4. The total error introduced will be the sum of three errors e_1, e_2, e_3, where e_1 comes from discarding those terms for which $|\gamma| \geqslant G$, e_2 from replacing $(e^{(\rho-\frac{1}{2})(\omega+y/T)})/\rho$ by $(e^{i\gamma(\omega+y/T)})/i\gamma$, and e_3 from extending the limits of integration from $\pm\frac{1}{2}T\omega$ to $\pm\infty$. We shall deal with these errors in separate lemmas.

10. Lemma 9. *For* $2.10^4 \leqslant \omega \leqslant \frac{1}{3}\log G$, *the error*

$$e_1 = \sum_{|\gamma|\geqslant G}\frac{1}{\rho}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)e^{(\rho-\frac{1}{2})(\omega+y/T)}\,dy$$

satisfies $$|e_1| < (0{\cdot}0002)\pi.$$

Since $-\frac{1}{2}T\omega \leqslant y \leqslant \frac{1}{2}T\omega$, we have $\frac{1}{2}\omega \leqslant \omega+y/T \leqslant \frac{3}{2}\omega$. Let

$$M = \sup\left|\sum_{|\gamma|\geqslant G}\frac{1}{\rho}e^{(\rho-\frac{1}{2})m}\right| \quad \text{for } \tfrac{1}{2}\omega \leqslant m \leqslant \tfrac{3}{2}\omega.$$

Then $$|e_1| = \left|\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)\Bigl(\sum_{|\gamma|\geqslant G}\frac{1}{\rho}e^{(\rho-\frac{1}{2})(\omega+y/T)}\Bigr)\,dy\right|$$

$$\leqslant M\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)\,dy \leqslant 2\pi M,$$

and this is less than $(0{\cdot}0001)2\pi$ by Lemma 2 (ii) applied to $e^{-\frac{1}{2}m}R(e^m, G)$, since $(\frac{1}{2}\omega, \frac{3}{2}\omega)$ is included in the appropriate range.

(H) Lemma 10. *On the hypothesis* (H) *and for* ω *subject to*

$$2.10^4 \leqslant \omega \leqslant \tfrac{2}{9}\log G,$$

the error

$$e_2 = \sum_{|\gamma|<G}\frac{1}{\rho}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)e^{(\rho-\frac{1}{2})(\omega+y/T)}\,dy - \sum_{|\gamma|<G}\frac{1}{i\gamma}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)e^{i\gamma(\omega+y/T)}\,dy$$

satisfies $$|e_2| < (0{\cdot}0468)\pi.$$

As in Lemma 9, we have $\frac{1}{2}\omega \leqslant \omega+y/T \leqslant \frac{3}{2}\omega$. Suppose here that m is that value of $\omega+y/T$ for which the value of

$$\sum_{|\gamma|<G} \left| \frac{1}{\rho} e^{(\rho-\frac{1}{2})(\omega+y/T)} - \frac{1}{i\gamma} e^{i\gamma(\omega+y/T)} \right|$$

is greatest. On (H) and for $2.10^4 \leqslant \omega \leqslant \frac{2}{9}\log G$, the error e_2 satisfies

$$|e_2| \leqslant \int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y) \sum_{|\gamma|<G} \left| \frac{1}{\rho} e^{(\rho-\frac{1}{2})(\omega+y/T)} - \frac{1}{i\gamma} e^{i\gamma(\omega+y/T)} \right| dy$$

$$\leqslant \sum_{|\gamma|<G} \left| \frac{1}{\rho} e^{(\rho-\frac{1}{2})m} - \frac{1}{i\gamma} e^{i\gamma m} \right| \int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)\, dy$$

$$< (0{\cdot}0234)2\pi$$

by Lemma 8, since m again lies in the relevant range.

Lemma 11. *For $\omega \geqslant 2.10^4$ and $T = 500$, the error*

$$e_3 = \sum_{|\gamma|<G} \frac{e^{i\gamma\omega}}{i\gamma} \left(\int_{-\infty}^{-\frac{1}{2}T\omega} + \int_{\frac{1}{2}T\omega}^{\infty} \right) K(y) e^{i\gamma y/T}\, dy$$

satisfies $$|e_3| < (0{\cdot}00002)\pi.$$

Since $K(y)$ is an even function of y and the γ's are symmetrically distributed, we have

$$|e_3| \leqslant 4 \sum_{0<\gamma<G} \frac{1}{\gamma} \left| \int_{\frac{1}{2}T\omega}^{\infty} K(y) e^{i\gamma y/T}\, dy \right|$$

$$= 4 \sum_{0<\gamma<T} + 4 \sum_{T\leqslant\gamma<G}.$$

Now we have the two inequalities†

$$\left| \int_{\frac{1}{2}T\omega}^{\infty} K(y) e^{i\gamma y/T}\, dy \right| \begin{cases} \leqslant \displaystyle\int_{\frac{1}{2}T\omega}^{\infty} 4y^{-2}\, dy = \frac{8}{T\omega}, \\ = \left| \displaystyle\int_{\frac{1}{2}T\omega}^{\infty} \frac{T}{i\gamma} \{e^{i\gamma y/T} - e^{\frac{1}{2}i\gamma\omega}\} K'(y)\, dy \right| < \dfrac{2T}{|\gamma|} \dfrac{4}{\frac{1}{2}T\omega}. \end{cases}$$

Using the former inequality in $\sum_{0<\gamma<T}$ and the latter in $\sum_{T\leqslant\gamma<G}$, we have, from Lemma 1 (i) and (ii),

$$|e_3| \leqslant \frac{32}{T\omega} \sum_{0<\gamma<T} \frac{1}{\gamma} + \frac{64}{\omega} \sum_{T\leqslant\gamma<G} \frac{1}{\gamma^2} < \frac{32}{T\omega} \frac{1}{4\pi} \log^2 T + \frac{64}{\omega} \frac{1}{2\pi} \frac{\log T}{T}.$$

Since $\omega > 2.10^4$ and $T = 500$ the required result follows.

† We have $K'(y) = 2\sin y/y^2 - 8\sin^2\frac{1}{2}y/y^3$ and $|K'(y)| < 4/y^2$ in the range concerned.

11. From Lemmas 9, 10, and 11 we may now replace (10), subject to the condition

$$2.10^4 \leqslant \omega < \tfrac{2}{9}\log G, \tag{11}$$

by

$$\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)F(\omega+y/T)\,dy = -\sum_{|\gamma|<G} \frac{e^{i\gamma\omega}}{i\gamma} \int_{-\infty}^{\infty} K(y)e^{i\gamma y/T}\,dy + E, \tag{12}$$

where

$$|E| = |R-e_1-e_2+e_3| < (0{\cdot}0471)\pi.$$

Applying (6) to the series on the right-hand side of (12), we have, still subject to (11),

$$\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)F(\omega+y/T)\,dy = -2\sum_{0<\gamma<T} 2\pi\frac{\sin\gamma\omega}{\gamma}\Big(1-\frac{\gamma}{T}\Big)+E$$

$$> -2\sum_{0<\gamma<T} 2\pi\frac{\sin\gamma\omega}{\gamma}\Big(1-\frac{\gamma}{T}\Big)-(0{\cdot}0471)\pi. \tag{13}$$

Now let $F_M = F_M(\omega)$ be the upper bound of $F(\omega+y/T)$ for the range $-\frac{1}{2}T\omega \leqslant y \leqslant \frac{1}{2}T\omega$. Since $K(y) \geqslant 0$, (13) gives

$$F_M J = F_M\frac{1}{2\pi}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)\,dy \geqslant \frac{1}{2\pi}\int_{-\frac{1}{2}T\omega}^{\frac{1}{2}T\omega} K(y)F(\omega+y/T)\,dy$$

$$> -2\sum_{0<\gamma<T} \frac{\sin\gamma\omega}{\gamma}\Big(1-\frac{\gamma}{T}\Big)-0{\cdot}0236. \tag{14}$$

Now by the definition (7) of F we have

$$F_M = \textit{upper bound of } (\psi_0(x)-x)x^{-\frac{1}{2}} \quad \textit{for } e^{\frac{1}{2}\omega} \leqslant x \leqslant e^{\frac{3}{2}\omega}. \tag{15}$$

We are therefore in a position to establish the following lemma.

(H) Lemma 12. *On the hypothesis* (H) *a sufficient condition that*

$$\pi(x)-\mathrm{li}\,x > 0,$$

for some x satisfying $2 \leqslant x < X_1$,† *is that, for some ω subject to the condition* (11),

$$-\sum_{0<\gamma<500} \frac{\sin\gamma\omega}{\gamma}\Big(1-\frac{\gamma}{500}\Big) > 0{\cdot}5123. \tag{16}$$

When (16) is true we have, by (14)‡ (and the fact that $T = 500$), $F_M J > 1{\cdot}001$, and *a fortiori* $F_M > 1{\cdot}001$ since $J < \frac{1}{2\pi}\int_{-\infty}^{\infty} = 1$. Lemma 12 then follows from Lemma 7.

† We recall that X_1 is the number concerned in (H), § 4, namely $\exp\exp\exp(7{\cdot}703)$.
‡ Which is valid subject to (11).

12. Our problem is now to find a suitable ω. It must be chosen so that the sines in (16) are predominantly negative, and such a choice is made as follows.

The number N of terms in the series on the left-hand side of (16) is equal to the number of γ's satisfying $0 < \gamma < 500$; this is known to be

$$N = 269. \tag{17}$$

Let ω_0 and q be the numbers

$$\omega_0 = 2.10^4+1, \qquad q = 3600. \tag{18}$$

By Dirichlet's theorem there is a number ω' satisfying

$$\omega_0 \leqslant \omega' \leqslant \omega_0 q^N \tag{19}$$

such that
$$\left|\frac{\gamma_n \omega'}{2\pi} - r_n\right| < \frac{1}{q} \quad (n = 1, 2,..., N), \tag{20}$$

where r_n is an integer. Now let

$$\omega = \omega' - k, \tag{21}$$

where
$$k = \tfrac{3}{400}. \tag{22}$$

Then, from (20) and (21),

$$\sin \gamma_n \omega = -\sin(k\gamma_n - \phi_n),$$

where $|\phi_n| < 2\pi/q = 0^\circ\,6'$. Now from (17), (18), (19), and (22),

$$2.10^4 < \omega < \omega_0 q^N = (2.10^4+1)3600^{269} = \exp\exp(7{\cdot}7021...) < \tfrac{2}{9}\log G.$$

The condition (11) is therefore satisfied. Hence, by Lemma 12, we shall have $\pi(x) - \operatorname{li} x > 0$ for some x satisfying

$$2 \leqslant x < X_1$$

provided that

$$S = \sum p(\gamma_n) = \sum_{n=1}^{269} \frac{\sin(k\gamma_n - \phi_n)}{\gamma_n}\left(1 - \frac{\gamma_n}{500}\right) > 0{\cdot}5123. \tag{23}$$

13. The inequality (23) is actually true. The right-hand side is what determines the top index of X_1 and it is here that we try to refine. It will suffice to sketch the numerical considerations involved.

The angles $k\gamma_n - \phi_n$ range from 6° to 215°, and the first 213 sines are positive. In the case of the remaining negative terms, for which

$$180^\circ < k\gamma_n - \phi_n < 215^\circ,$$

the γ's satisfy
$$420 < \gamma_n < 500.$$

Hence, in addition to the fact that $1/\gamma_n$ is small, either the absolute value of the sine or the factor $(1-\gamma_n/500)$ is small, and these negative terms prove to be of little importance. For the rest, sufficient is known about the values of the γ's† to enable us to obtain a lower bound to S by straightforward calculation.

In this the first 29 terms are calculated separately, the remainder are grouped in intervals (of the values of the γ's) of 10. For example, the first group contains the 4 terms for which $100 < \gamma \leqslant 110$, and the last group contains the 7 terms for which $490 < \gamma \leqslant 500$. We obtain a lower bound to the contribution of each term, or group of terms, by making use of the fact that the function $p(\gamma_n)$, defined in (23), satisfies (whatever the particular values of ϕ_n, ϕ_{n+1}) $p(\gamma_n) < p(\gamma_{n+1})$ for $\gamma_n < 457$ (approximately), and thereafter satisfies $p(\gamma_n) > p(\gamma_{n+1})$. We may replace each of the first 29 γ's by the upper bound to the interval in which it is known to lie, and for those groups for which $\gamma \leqslant 450$ we replace each of the γ's in the group by the upper bound of the interval in which it lies. The same replacement applies for the subgroup 450–457. For the subgroup 457–460 and the remaining groups for which $\gamma > 460$, since $p(\gamma_n)$ is now increasing, we replace the γ's in each group by the lower bound of the interval concerned. For example, each of the 4 γ's between 100 and 110 is replaced by 110, while each of the 7 γ's between 470 and 480 is replaced by 470. ϕ_n is replaced by $+6'$ or $-6'$ according as $\gamma_n k < 90°$ or $\gamma_n k > 90°$.

We find that $$S > 0{\cdot}5131 > 0{\cdot}5123.$$

$\pi(x) - \operatorname{li} x$ is therefore positive for some x satisfying

$$2 \leqslant x < X_1 = \exp\exp\exp(7{\cdot}703).$$

Part II

14. Before we can begin developing the consequences of (NH), the negation of (H), we need a number of preliminary results about the function

$$\Pi_0(x) - \operatorname{li} x = \tfrac{1}{2}\{\Pi(x+0)+\Pi(x-0)\} - \operatorname{li} x,$$

where $\Pi(x)$ is defined as in § 6. For $x > 1$ we have ((**2**), 81–82)

$$\Pi_0(x) - \operatorname{li} x = -\sum_\rho \operatorname{li} x^\rho + \int\limits_x^\infty \frac{du}{(u^2-1)u\log u} - \log 2, \tag{24}$$

the series being boundedly convergent in any finite interval $1 < a \leqslant x \leqslant b$. The li function for a complex argument is defined by

$$\operatorname{li} x^\rho = \operatorname{li} e^{\rho \log x}, \tag{25}$$

and, for $w = u+vi$ where $v \neq 0$,

$$\operatorname{li} e^w = \int\limits_{-\infty+vi}^{u+vi} \frac{e^z}{z}\, dz.$$

† (**6**), (**7**), (**8**), (**9**). In addition I have used some calculations performed by Dr. Comrie, kindly lent to me by Professor Titchmarsh.

We define the function $L(t)$ for $t > 0$ by the series

$$L(t) = -e^{-\frac{1}{2}t} \sum_{\rho} \mathrm{li}\, e^{\rho t}. \tag{26}$$

From (25) and (26) we have, for $t > 0$,

$$-L(t) = e^{-\frac{1}{2}t} \sum_{\rho} \int\limits_{-\infty+i\gamma t}^{(\beta+i\gamma)t} \frac{e^z}{z}\, dz = \sum_{\rho} e^{(\rho-\frac{1}{2})t} \int\limits_0^\infty \frac{e^{-v}}{\rho t - v}\, dv$$

$$= \sum e^{(\rho-\frac{1}{2})t}\left[\frac{1}{\rho t} + \int\limits_0^\infty \frac{e^{-v}}{(\rho t - v)^2}\, dv\right]$$

$$= \sum \frac{e^{(\rho-\frac{1}{2})t}}{\rho t} + \sum e^{(\rho-\frac{1}{2})t} \int\limits_0^\infty \frac{e^{-v}\, dv}{(\rho t - v)^2},$$

both series being boundedly convergent in any interval of type

$$0 < a' \leqslant t \leqslant b'$$

since the first is. So

$$-L(t) = \sum_{\rho} u_1(\rho, t) + \sum_{\rho} u_2(\rho, t), \tag{27}$$

$$u_1 = \frac{e^{(\rho-\frac{1}{2})t}}{\rho t}, \qquad u_2 = e^{(\rho-\frac{1}{2})t} \int\limits_0^\infty \frac{e^{-v}\, dv}{(\rho t - v)^2}. \tag{28}$$

15. LEMMA 13. $\quad |L(t)| \leqslant 4e^{\frac{1}{2}t} \quad (t \geqslant 1).$

By Lemma 2 (iii), if $\tau = e^t \geqslant e$,

$$\left|\sum u_1(\rho, t)\right| = \left|\tau^{-\frac{1}{2}}(\log \tau)^{-1} \sum \tau^{\rho}/\rho\right| < 3\tau^{\frac{1}{2}}.$$

In $u_2(\rho, t)$ we have $|\rho t - v|^2 \geqslant |i\gamma t|^2 = \gamma^2 t^2$,

$$|u_2| \leqslant \tau^{\frac{1}{2}} t^{-2} \gamma^{-2}, \qquad \left|\sum u_2\right| \leqslant e^{\frac{1}{2}t} t^{-2}\, 2 \sum_{\gamma > 0} \gamma^{-2} < e^{\frac{1}{2}t},$$

since $\sum \gamma^{-2} < 0{\cdot}05$. The result follows.

16. LEMMA 14. *A sufficient condition that* $\pi(x) - \mathrm{li}\, x > 0$ *for some* x *of* $2 \leqslant x \leqslant X$ *is that, for some* y *satisfying* $10^4 \leqslant y \leqslant \log X$,

$$L(y) \geqslant 1. \tag{29}$$

Suppose the condition of the lemma is satisfied for a certain y, and let $x = e^y$. Then, by Lemma 5, either $\pi(\xi) - \mathrm{li}\, \xi > 0$ for some ξ of $2 \leqslant \xi \leqslant x^{\frac{1}{2}}$, or else

$$\Pi_0(x) - \pi(x) \leqslant \Pi(x) - \pi(x) < (1{\cdot}0005) x^{\frac{1}{2}}/\log x < 2x^{\frac{1}{2}}/\log x.$$

In the first alternative, we have what we want at once, and we have only to consider the second. Now, from (24) and (26), the integral in (24) being positive,

$$\Pi_0(x) - \mathrm{li}\, x > x^{\frac{1}{2}} L(y) - \log 2 \geqslant x^{\frac{1}{2}} - \log 2,$$

and so, from the second alternative,

$$\pi(x)-\operatorname{li} x > x^{\frac{1}{2}}-\log 2-2x^{\frac{1}{2}}/\log x > 0,$$

as desired.

17. Let X_1 and G be, as in § 4,

$$X_1 = \exp\exp\exp(7{\cdot}703), \qquad G = X_1^3.$$

Since we cannot use O's in connexion with numerical bounds, we shall use ϑ's (ϑ', ϑ_1, etc.) to denote numbers, possibly complex, satisfying $|\vartheta| \leqslant 1$. They will in general not be the same from one occurrence to the next, but where more than one occurs in the same expression we distinguish them.

Let $y \geqslant G$, and let λ be any real number satisfying†

$$|\lambda| \leqslant G\ (\leqslant y).$$

Consider‡ the function $F(y,\lambda)$ defined by

$$F(y,\lambda) = \int_{\frac{1}{2}y}^{\infty} [-L(t)]te^{E}\,dt, \tag{30}$$

$$E = E(t,y,\lambda) = -\lambda it-\tfrac{1}{2}(t-y)^2/y. \tag{31}$$

We have the following result.

Lemma 15. *Write $b = \beta-\frac{1}{2}$, $r = \rho-\frac{1}{2}-i\lambda = b+i(\gamma-\lambda)$. Then, subject to $y \geqslant G \geqslant |\lambda|$,*

$$F(y,\lambda) = \sum_{\rho} U(\rho), \tag{32}$$

where

$$U(\rho) = (2\pi y)^{\frac{1}{2}} e^{(r+\frac{1}{2}r^2)y}\left(\frac{1}{\rho}+\frac{400\vartheta}{\gamma^2 y}\right)+\vartheta'\frac{e^{-\frac{1}{10}y}}{\gamma^2}. \tag{33}$$

The proof of this is rather long, and we break it up into two subsidiary lemmas, A and B, and a short final deduction from them. We have among other things to show that the series (27) for L can be integrated term by term in (30): this involves a limit-process $T \to \infty$ for fixed y, λ. The parts of Lemmas A, B dealing with this use O's, which are accordingly uniform in the ρ (or γ), but not in the 'fixed' y, λ; the K's similarly are independent of ρ, γ, but not of y, λ.

Lemma A. *For $u_1(\rho,t)$, defined by* (28), *we have*

$$\int_{T}^{\infty} u_1\, te^{E}\,dt = O\left(\frac{e^{-KT^2}}{\gamma^2}\right), \tag{34}$$

$$\int_{\frac{1}{2}y}^{\infty} u_1\, te^{E}\,dt = \frac{(2\pi y)^{\frac{1}{2}}}{\rho} e^{(r+\frac{1}{2}r^2)y}+\frac{2\vartheta e^{-\frac{1}{8}y}}{\gamma^2}. \tag{35}$$

† These conditions hold throughout the rest of the paper. Note that G is so large that any inequalities like $100y^{10}e^{-y/8} < e^{-y/10}$ that occur in the run of our argument will be true when they are 'true for large y'.

‡ The introduction of $F(y,\lambda)$ is the idea given me by Professor J. E. Littlewood.

LEMMA B. *For $u_2(\rho,t)$, defined by* (28), *we have*

$$\int_T^\infty u_2 te^E\,dt = O\left(\frac{e^{-KT^2}}{\gamma^2}\right), \tag{36}$$

$$\int_{\frac{1}{4}y}^\infty u_2 te^E\,dt = \frac{324\vartheta}{y}(2\pi y)^{\frac{1}{2}}\frac{e^{(r+\frac{1}{2}r^2)y}}{\gamma^2}+\frac{9\vartheta'}{\gamma^2}e^{-\frac{1}{8}y}. \tag{37}$$

18. In Lemmas A, B we may, by symmetry (since λ can take either sign), suppose without loss of generality that $\gamma > 0$.

Proof of Lemma A. We have

$$u_1 te^E = \frac{1}{\rho}e^{\gamma it}f(t), \qquad f(t) = e^{(\beta-\frac{1}{2}-i\lambda)t-\frac{1}{2}(t-y)^2/y}. \tag{38}$$

For $t \geqslant T$,

$$\int_T^t u_1 te^E\,dt = \frac{1}{\rho}\left[\frac{e^{\gamma it}}{\gamma i}f(t)\right]_T^t - \frac{1}{\rho}\int_T^t \frac{e^{\gamma it}}{\gamma i}f'(t)\,dt.$$

As $T\to\infty$ we have, uniformly in $t\geqslant T$, $f(t), f'(t) = O(e^{-KT^2})$. It follows that $\int_T^\infty u_1 te^E\,dt$ exists, and that it is $O(\gamma^{-2}e^{-KT^2})$; and this is (34).

Next,
$$\int_{-\infty}^\infty u_1 te^E\,dt = \frac{1}{\rho}\int_{-\infty}^\infty e^{rt-\frac{1}{2}(t-y)^2/y}\,dt = \frac{(2\pi y)^{\frac{1}{2}}}{\rho}e^{(r+\frac{1}{2}r^2)y}. \tag{39}$$

Again,

$$\int_{-\infty}^{\frac{1}{4}y} u_1 te^E\,dt = \frac{1}{\rho}\left[\frac{e^{\gamma it}}{\gamma i}f(t)\right]_{-\infty}^{\frac{1}{4}y} - \frac{1}{\rho}\int_{-\infty}^{\frac{1}{4}y}\frac{e^{\gamma it}}{\gamma i}f'(t)\,dt = J_1+J_2, \text{ say.} \tag{40}$$

We have $f(-\infty) = 0$ and $|f(\frac{1}{4}y)| \leqslant e^{|b|\frac{1}{4}y-\frac{9}{32}y} < e^{-\frac{1}{8}y}$, so that

$$J_1 = \vartheta\gamma^{-2}e^{-\frac{1}{8}y}.$$

Also, for $-\infty < t \leqslant \frac{1}{4}y$,

$$\begin{aligned}|f'(t)| &= |(\beta-\tfrac{1}{2}-i\lambda)-(t-y)/y|e^{bt-\frac{1}{2}(t-y)^2/y}\\ &\leqslant (\tfrac{1}{2}+|\lambda|+|t-y|)e^{bt-\frac{1}{2}(t-y)^2/y}.\end{aligned}$$

Writing $u = |t-y| = y-t$, and observing that $u \geqslant \frac{3}{4}y$ and

$$\tfrac{1}{2}+|\lambda| < 2y \leqslant \tfrac{8}{3}u,$$

we have
$$|f'| \leqslant 4ue^{b(y-u)-\frac{1}{2}u^2/y} \leqslant 12(by+u)e^{by}e^{-bu-\frac{1}{2}u^2/y},$$

$$\begin{aligned}|J_2| &\leqslant 12\gamma^{-2}e^{by}\int_{\frac{3}{4}y}^\infty (by+u)e^{-bu-\frac{1}{2}u^2/y}\,du = 12\gamma^{-2}ye^{(\frac{1}{4}b-\frac{9}{32})y}\\ &\leqslant 12\gamma^{-2}ye^{-\frac{5}{32}y} < \gamma^{-2}e^{-\frac{1}{8}y}.\end{aligned}$$

So $J_1+J_2 = 2\vartheta\gamma^{-2}e^{-\frac{1}{4}y}$, which, combined with (39) and (40), gives (35) and completes the proof of Lemma A.

19. *Proof of Lemma B.* Let $t \geqslant T$ and $T \to \infty$. We have $|\rho t - v|^2 \geqslant \gamma^2 t^2$, and so, from (28),

$$\left|\int_T^t u_2 t e^E\,dt\right| \leqslant \int_0^\infty e^{-v}\,dv\,.\gamma^{-2}\int_T^t t^{-2}te^{bt-\frac{1}{2}(t-y)^2/y}\,dt$$
$$= O(\gamma^{-2}e^{-KT^2}),$$

and $\int_T^\infty$ exists and satisfies (36).

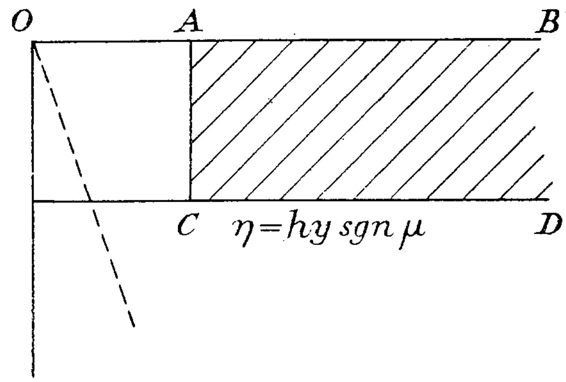


FIG. 1.

Next we have
$$\int_{\frac{1}{4}y}^\infty u_2 t e^E\,dt = \int_0^\infty e^{-v} H(\rho, v)\,dv,$$

$$H = H(\rho, v) = \int_{\frac{1}{4}y}^\infty e^{rt-\frac{1}{2}(t-y)^2/y}\frac{t\,dt}{(\rho t - v)^2}. \tag{41}$$

We prove (37) of Lemma B by showing that for each v of $(0, \infty)$ H is of the form of the right-hand side of (37) (noting that $\int_0^\infty e^{-v}\,dv = 1$).

We deform the t-contour $\frac{1}{4}y$ to ∞, or AB, in a manner independent of v, as follows. Let $\mu = \gamma - \lambda$, $h = \min(2, |\mu|)$. With $t = \xi + i\eta$ we take a line $\eta = hy\operatorname{sgn}\mu$ $(= \pm hy)$, and replace the original path AB by ACD of the figure (drawn for the worst case, namely $\operatorname{sgn}\mu = -1$). First, the pole $t = v/\rho$ is outside the shaded area, so that† $H = \int_{ACD}$. For the pole is on a line (dotted in the figure) whose slope (tangent) is $-\gamma/\beta$; this is downward‡ and steeper absolutely than $\gamma_1 > 14$, steeper, therefore, than OC in the unfavourable case (of the figure) when C is below A.

† The integrand is uniformly $O(e^{-K\xi^2})$ as $t \to \infty$ in the shaded area.

‡ Recall that in this proof we have $\gamma > 0$.

Taking the integral for H along ACD, then, we have

$$|\rho t-v| \geqslant |\operatorname{im}(\rho t-v)| = |\gamma\xi+\beta\eta|$$
$$\geqslant \gamma\xi-1.2y \geqslant \tfrac{1}{3}\gamma\xi$$

(since $\gamma > 14$, $\xi \geqslant \frac{1}{4}y$). So

$$|H| \leqslant 9\gamma^{-2} \int\limits_{ACD} |e^{E_1}\xi^{-2}t\,dt|, \tag{42}$$

where $$E_1 = rt-\tfrac{1}{2}(t-y)^2/y, \qquad r = b+i\mu, \tag{43}$$

and, as alternative forms,

$$E_1 = -\tfrac{1}{2}y+(r+1)t-\tfrac{1}{2}t^2/y = (r+\tfrac{1}{2}r^2)y-\tfrac{1}{2}[t-(r+1)y]^2/y, \tag{44}$$

$$\operatorname{re} E_1 = -\tfrac{1}{2}y+(b+1)\xi-\mu\eta-\tfrac{1}{2}\xi^2/y+\tfrac{1}{2}\eta^2/y. \tag{45}$$

On AC we have $\xi = \frac{1}{4}y$, $\eta = \sigma hy \operatorname{sgn}\mu$, $0 \leqslant \sigma \leqslant 1$,

$$\operatorname{re} E_1 = [-\tfrac{1}{2}+\tfrac{1}{4}(b+1)]y-\tfrac{1}{32}y-\sigma hy(|\mu|-\tfrac{1}{2}\sigma h),$$

and since the last term is non-positive and $b+1 < \frac{3}{2}$, $\operatorname{re} E_1 < -\frac{1}{7}y$, and

$$\int\limits_{AC} |e^{E_1}\xi^{-2}t\,dt| \leqslant e^{-\frac{1}{7}y}(\tfrac{1}{4}y)^{-2}OC.AC < \tfrac{1}{2}e^{-\frac{1}{8}y}. \tag{46}$$

For CD we have two cases.

Case (i). $|\mu| \leqslant 2$. Here $\eta = \mu y$ ($= \operatorname{im} r.y$),

$$E_1 = (r+\tfrac{1}{2}r^2)y-\tfrac{1}{2}[\xi-(b+1)y]^2/y,$$

$$\int\limits_{CD} |e^{E_1}\xi^{-2}t\,dt| \leqslant |e^{(r+\frac{1}{2}r^2)y}| \int\limits_{\frac{1}{4}y}^{\infty} e^{-\frac{1}{2}[\xi-(b+1)y]^2/y}\{\xi^{-2}(2y+\xi)\}\,d\xi.$$

The curly bracket is greatest for $\xi = \frac{1}{4}y$, and it is then $36y^{-1}$. Taking this outside, and then the integral from $-\infty$ to ∞, we find that $\int\limits_{CD}$ is at most $(2\pi y)^{\frac{1}{2}}|e^{(r+\frac{1}{2}r^2)y}|36y^{-1}$.

Combining this with (42) and (46) we have, in case (i),

$$|H| \leqslant 9\gamma^{-2}(e^{-\frac{1}{8}y}+(2\pi y)^{\frac{1}{2}}|e^{(r+\frac{1}{2}r^2)y}|36y^{-1}). \tag{47}$$

Case (ii). $|\mu| > 2$. Here CD has $\eta = 2y \operatorname{sgn}\mu$. We have from (45)

$$\begin{aligned}\operatorname{re} E_1 &= -\tfrac{1}{2}y+(b+1)\xi-2|\mu|y-\tfrac{1}{2}\xi^2/y+2y \\ &\leqslant -\tfrac{5}{2}y+\tfrac{3}{2}\xi-\tfrac{1}{2}\xi^2/y, \quad \text{since } |\mu| > 2, \\ &= -\tfrac{1}{4}y-\tfrac{1}{4}\xi^2/y-\tfrac{1}{4}(\xi-3y)^2/y \leqslant -\tfrac{1}{4}y-\tfrac{1}{4}\xi^2/y.\end{aligned}$$

As before, $|\xi^{-2}t| \leqslant 36y^{-1}$, and so

$$\int\limits_{CD} |e^{E_1}\xi^{-2}t\,dt| \leqslant 36y^{-1} \int\limits_{-\infty}^{\infty} e^{-\frac{1}{4}y-\frac{1}{4}\xi^2/y}\,d\xi \leqslant \tfrac{1}{2}e^{-\frac{1}{8}y}.$$

From this and (46), $|H| \leqslant 9\gamma^{-2}e^{-\frac{1}{8}y}$, and (47) is true also in case (ii). *A fortiori* H is of the form of the right-hand side of (37), and, as we observed above, this proves (37). This completes the proof of Lemma B.

20. We now have Lemmas A and B (in which γ now is not restricted to be positive), and can take up Lemma 15. By Lemma 13 and (30) we have

$$F = \lim_{T\to\infty} \int\limits_{\frac{1}{4}y}^{T} [-L(t)]te^{E}\,dt,$$

since $\mathrm{re}\,E = -\frac{1}{2}(t-y)^2/y < -Kt^2$ as $t \to \infty$. In $\int\limits_{\frac{1}{4}y}^{T}$ we may substitute $-L(t) = \sum u_1 + \sum u_2$ from (27) and integrate term by term, since the two series are boundedly convergent. If we then replace T by ∞ in each term, the error is

$$\sum \int\limits_{T}^{\infty} u_1\, te^{E}\,dt + \sum \int\limits_{T}^{\infty} u_2\, te^{E}\,dt = O(e^{-KT^2} \textstyle\sum \gamma^{-2}),$$

by (34) from Lemma A and (36) from Lemma B, and this tends to 0 as $T \to \infty$. Hence

$$F(y,\lambda) = \sum_{\rho} U(\rho),$$

where

$$U(\rho) = \int\limits_{\frac{1}{4}y}^{\infty} u_1\, te^{E}\,dt + \int\limits_{\frac{1}{4}y}^{\infty} u_2\, te^{E}\,dt,$$

and when we substitute from (35) and (37) (and make a couple of small adjustments) we arrive at Lemma 15.

21. LEMMA 16.† *For $y \geqslant G \geqslant \lambda \geqslant 0$ we have*

$$F(y,\lambda) = \sum_{|\gamma-\lambda|\leqslant 2} \frac{1}{\rho}(2\pi y)^{\frac{1}{2}}e^{(r+\frac{1}{2}r^2)y}\left(1+\frac{30\vartheta}{y}\right)+\vartheta'.$$

Since $400|\rho|/(\gamma^2 y) < 30/y < 1$, Lemma 15 shows that F is equal to something of the form of the $\sum$ in the lemma, *plus*

$$2\vartheta_1 \sum_{|\gamma-\lambda|>2} \frac{(2\pi y)^{\frac{1}{2}}|e^{r+\frac{1}{2}r^2)y}|}{|\gamma|} + \vartheta_2\, e^{-\frac{1}{16}y} \sum \frac{1}{\gamma^2}. \tag{48}$$

When $|\gamma-\lambda| > 2$ we have

$$\mathrm{re}(r+\tfrac{1}{2}r^2) = b+\tfrac{1}{2}b^2-\tfrac{1}{2}(\gamma-\lambda)^2 < -\tfrac{1}{4}(\gamma-\lambda)^2-\tfrac{3}{8},$$

† (i) From now on λ is non-negative (we normalized in the *proof* above to $\gamma > 0$ and λ of both signs). (ii) The ϑ, of course, varies with the term it occurs in.

and also $|\gamma/(\gamma-\lambda)| \leqslant 1+\lambda < 2y$. The first term in (48) is therefore

$$\vartheta \sum_{\rho} \frac{4y}{\gamma^2}(2\pi y)^{\frac{1}{2}}\{|\gamma-\lambda|e^{-\frac{1}{4}(\gamma-\lambda)^2}\}e^{-\frac{3}{8}y} = \tfrac{1}{2}\vartheta, \tag{49}$$

since the curly bracket is less than (say) 10. Lemma 16 follows.

22. We are now in a position to develop the consequences of (NH), the negation of the hypothesis (H). To assume (NH) is to assume that a zero $\beta_0+i\gamma_0$ exists (with γ_0 positive, by the symmetry) satisfying

$$(\text{NH}) \quad \begin{cases} b_0 = \beta_0-\frac{1}{2} > X_1^{-3}\log^{-2}X_1 = B, \\ 0 < \gamma_0 < X_1^3 = G. \end{cases}$$

We begin by supposing that (for an undetermined Y) the relation '$L(y) \geqslant 1$ for some y' occurring in Lemma 14 is *not* satisfied for the range $G \leqslant y \leqslant 4Y$; that is, we suppose that†

$$L(\eta) < 1 \quad \text{for } G \leqslant \eta \leqslant 4Y. \tag{50}$$

By arguing from the pair of hypotheses (NH) and (50) we find ourselves able to produce a Y_0 (actually G^{10}) such that, if the Y of (50) is Y_0, there is a contradiction. Then (NH) implies (i) that (50) is false for $Y = Y_0$; so (ii) that for *some* y of the range $G \leqslant y \leqslant 4Y_0$ we must have $L(y) \geqslant 1$, when Lemma 14 (with $4Y_0$ for $\log X$) gives $\pi(x)-\operatorname{li} x > 0$ for some x of

$$2 \leqslant x < X = \exp(4Y_0) \ [= \exp(4G^{10})].$$

This, then, is what results from (NH), and since the X is greater than the X_1 derived from (H), it is our final number.

23. LEMMA 17. *If* [*in accordance with* (50)] $L(\eta) < 1$ *for* $G \leqslant \eta \leqslant 4Y$, *then for* $4G \leqslant y \leqslant Y$, $0 \leqslant \lambda \leqslant G$, *we have*

$$|F(y,0)| \leqslant 1, \tag{51}$$

$$|F(y,\lambda)| < 6Y^{\frac{3}{2}}+4. \tag{52}$$

When $\lambda = 0$, the condition $|\gamma-\lambda| \leqslant 2$ is vacuous, and (51) is a case of Lemma 16.

Next, since $L(t)$ is real for $t > 0$, we have, for λ of $0 \leqslant \lambda \leqslant G$, by (30) and (31),

$$\begin{aligned} -F(y,\lambda) &= \int_{\frac{1}{2}y}^{\infty} t\,L(t)(\cos\lambda t - i\sin\lambda t)e^{-\frac{1}{2}(t-y)^2/y}\,dt \\ &= \mathscr{R}-i\mathscr{I}, \text{ say.} \end{aligned} \tag{53}$$

† This means 'for *all* η of the range', and similar interpretations are intended wherever we do not explicitly have 'some'. This being the usual interpretation, we may seem to be labouring the obvious, but the distinctions of 'all' and 'some' are very vital, and complicated by ranges (those in Lemma 17) that 'look' alike, but are not quite so.

Consider the four expressions

$$-F(y,0)\begin{matrix}\pm\mathscr{R}\\ \pm\mathscr{I}\end{matrix} = \left(\int\limits_{\frac{1}{4}y}^{4Y}+\int\limits_{4Y}^{\infty}\right)tL(t)\left\{1\begin{matrix}\pm\cos\lambda t\\ \pm\sin\lambda t\end{matrix}\right\}e^{-\frac{1}{2}(t-y)^2/y}\,dt$$

$$= J_1+J_2, \text{ say.} \tag{54}$$

In J_2 we substitute $|L(t)| \leqslant 4e^{\frac{1}{4}t}$ from Lemma 13, and, remembering that $4G \leqslant y \leqslant Y$, we obtain

$$|J_2| \leqslant \int\limits_{4y}^{\infty} t\,.\,4e^{\frac{1}{4}t}\,.\,2e^{-\frac{1}{2}(t-y)^2/y}\,dt$$

$$= 8e^{-\frac{1}{2}y}\int\limits_{4y}^{\infty} te^{-\frac{1}{2}(t-4y)-\frac{1}{2}(t-2y)^2/y}\,dt < 1. \tag{55}$$

In J_1 we have $G \leqslant t \leqslant 4Y$, and so $L(t) < 1$ by the hypothesis (50); hence, *the curly bracket in* (54) *being in all four cases non-negative*, we have, *algebraically*,

$$J_1 \leqslant \int\limits_{\frac{1}{4}y}^{4Y} 2t\,e^{-\frac{1}{2}(t-y)^2/y}\,dt \leqslant \int\limits_{-\infty}^{\infty}(2y+2|t-y|)e^{-\frac{1}{2}(t-y)^2/y}\,dt$$

$$= 2y(2\pi y)^{\frac{1}{2}}+8y < 6y^{\frac{3}{2}} \leqslant 6Y^{\frac{3}{2}}. \tag{56}$$

Since $|F(y,\lambda)| \leqslant |\mathscr{R}|+|\mathscr{I}|$, from (53), *and since* $|\mathscr{R}|+|\mathscr{I}|$ *is, for each* y, *one (varying with* y*) of the four combinations* $\pm\mathscr{R}\pm\mathscr{I}$, we have, from (54) to (56),

$$|F(y,\lambda)| \leqslant 6Y^{\frac{3}{2}}+1+2|F(y,0)| < 6Y^{\frac{3}{2}}+4,$$

the desired result.

24. We now combine Lemmas 16 and 17, and take $Y = G^{10}$ (Y has this meaning from now on). The upshot is that, subject to (NH), and to the further 'hypothesis'

$$(\mathrm{H}_1) \qquad L(\eta) < 1 \quad (G \leqslant \eta \leqslant 4Y),$$

we have, for λ, y satisfying

$$0 \leqslant \lambda \leqslant G, \tag{57}$$

$$4G \leqslant y \leqslant Y, \tag{58}$$

and for *some* set of ϑ's,

$$\left|\sum_{|\gamma-\lambda|\leqslant 2}\frac{1}{\rho}e^{(r+\frac{1}{2}r^2)y}\left(1+\frac{30\vartheta}{y}\right)\right| < (2\pi y)^{-\frac{1}{2}}(6Y^{\frac{3}{2}}+4+1) < \tfrac{1}{2}Y^{\frac{3}{2}}, \tag{59}$$

where $r = b+i(\gamma-\lambda)$.†

We now take $\lambda = \gamma_0$, where γ_0 is the number in (NH), § 22. [λ duly

† We go on to derive a contradiction from this state of things, as a result of which one of (NH) and (H_1) must be false.

satisfies (57).] So from (59) with $\lambda = \gamma_0$ and so $r = b+i(\gamma-\gamma_0)$,

$$\operatorname{re} \sum_{|\gamma-\gamma_0|\leqslant 2} \frac{i}{\rho}\left(1+\frac{30\vartheta}{y}\right)\exp[(b+\tfrac{1}{2}b^2)y-\tfrac{1}{2}(\gamma-\gamma_0)^2y+i(\gamma-\gamma_0)(1+b)y] < \tfrac{1}{2}Y^{\frac{3}{2}}. \tag{60}$$

We need to know an upper bound for the number N of terms in the sum; Lemma 1 (iv) with $h = 4$, $T = \gamma_0-2$, gives

$$N < \frac{6}{2\pi}\log\gamma_0+8{\cdot}7 < \log G. \tag{61}$$

We proceed to choose, without violating (58), a y $(= y_0)$ for which the real parts of the terms in the sum in (60) are all positive. In the first place, since $\gamma > 14$, the argument of any factor i/ρ lies between $\pm 5°$, and that of any $1+30\vartheta/y$ between $\pm 1°$. Now by Dirichlet's theorem there exists a y_0 satisfying

$$Y^{\frac{1}{2}} \leqslant y_0 \leqslant Y^{\frac{1}{2}}5^N, \tag{62}$$

and such that, for each of the N γ's satisfying $|\gamma-\gamma_0| \leqslant 2$,

$$|(\gamma-\gamma_0)(1+b)y_0-2\pi k| < \tfrac{1}{5}\,.\,2\pi,$$

where k is an integer. Further, with $Y = G^{10}$ and N satisfying (61), $y = y_0$ [satisfying (62)] duly satisfies (58). With $y = y_0$ the arguments of all the terms in the sum in (60) lie between $\pm 80°$; hence the real parts of all the terms are positive, and the sum of them is at least as great as any one term. Choosing the one term to be that with $\gamma = \gamma_0$, we have

$$\frac{\gamma_0}{\gamma_0^2+\beta_0^2}e^{(b_0+\frac{1}{2}b_0^2)y_0}\left\{1-\frac{30}{y_0}\right\} < \tfrac{1}{2}Y^{\frac{3}{2}},$$

$$e^{b_0y_0} < \tfrac{1}{2}(\gamma_0+1/\gamma_0)\{1-30/y_0\}^{-1}Y^{\frac{3}{2}} < \gamma_0 Y^{\frac{3}{2}} \leqslant GY^{\frac{3}{2}},$$

$$e^{By_0} < Y^{\frac{3}{2}}G = G^{16}.$$

With $B = X_1^{-3}\log^{-2}X_1$, $X_1^3 = G$, this contradicts $y_0 \geqslant Y^{\frac{1}{2}} = G^5$ of (62).† So either (NH) is false [and (H) true] or (H_1) is false. In the first case $\pi(x)-\operatorname{li} x > 0$ for an $x < X_1$; in the second it happens for an x of

$$2 \leqslant x < X_2 = \exp(4Y) = \exp(4G^{10}) = \exp(4X_1^{30}).$$

Since $X_2 > X_1 = \exp\exp\exp(7{\cdot}703)$, we conclude finally that $\pi(x)-\operatorname{li} x > 0$ for some $x < X$, where

$$X = \exp\exp\exp\exp(7{\cdot}705) < 10^{10^{10^{10^3}}}.$$

† There is a great deal to spare at this point: see the footnote on p. 50.

REFERENCES

1. J. E. LITTLEWOOD, 'Sur la distribution des nombres premiers', *Comptes Rendus*, 158 (1914), 263–6.
2. A. E. INGHAM, *The Distribution of Prime Numbers*, Cambridge Mathematical Tracts, No. 30 (1932).
3. S. SKEWES, 'On the difference $\pi(x)-\mathrm{li}\, x$ (I)', *J. London Math. Soc.* 8 (1933), 277–83.
4. A. E. INGHAM, 'A note on the distribution of primes', *Acta Arith.* 1 (1936), 201–11.
5. R. J. BACKLUND, 'Über die Nullstellen der Riemannschen Zetafunktion', *Acta Math.* 41 (1918), 345–75.
6. JAHNKE-EMDE, *Funktionentafeln* (Teubner).
7. J. I. HUTCHINSON, 'On the roots of the Riemann zeta function', *Trans. American Math. Soc.* 27 (1925), 49–60.
8. E. C. TITCHMARSH, 'The zeros of the zeta-function', *Proc. Royal Soc.* A, 151 (1935), 234–55.
9. —— 'The zeros of the zeta-function', ibid. 157 (1936), 261–3.

University of Cape Town.

PAIR CORRELATION OF ZEROS AND PRIMES IN SHORT INTERVALS

Daniel A. Goldston and Hugh L. Montgomery*

1. Statement of results.

In 1943, A. Selberg [15] deduced from the Riemann Hypothesis (RH) that

$$\int_1^X \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2 x^{-2}\, dx \ll \delta(\log X)^2 \qquad (1)$$

for $X^{-1} \leqslant \delta \leqslant X^{-1/4}$, $X \geqslant 2$. Selberg was concerned with small values of δ, and the constraint $\delta \leqslant X^{-1/4}$ was imposed more for convenience than out of necessity. For larger δ we have the following result.

Theorem 1. *Assume* RH. *Then*

$$\int_1^X \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2 x^{-2}\, dx \ll \delta(\log X)(\log 2/\delta) \qquad (2)$$

for $0 < \delta \leqslant 1$, $X \geqslant 2$.

In this estimate, the error term for the number of primes in the interval $(x, (1+\delta)x]$ is damped by the factor x^{-2}, and the length of the interval, δx, varies with x. Saffari and Vaughan [14] considered the undamped integral, and derived from RH the estimates

$$\int_1^X \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2\, dx \ll \delta X^2 (\log 2/\delta)^2 \qquad (3)$$

for $0 < \delta \leqslant 1$, and

$$\int_1^X \big(\psi(x+h) - \psi(x) - h\big)^2\, dx \ll hX(\log 2X/h)^2 \qquad (4)$$

---

*Research supported in part by NSF Grant MCS82-01602.

for $0 < h \leqslant X$. It may be similarly shown that RH gives the estimate

$$\int_1^X (\psi(x) - x)^2 \, dx \ll X^2 . \tag{5}$$

Gallagher and Mueller [5] showed that if one assumes not only RH but also the pair correlation conjecture

$$\begin{aligned} &\# \{(\gamma,\gamma') : 0 < \gamma \leqslant T,\ 0 < \gamma - \gamma' \leqslant 2\pi a/\log T\} \\ &\qquad = \Big(\frac{1}{2\pi} \int_0^a 1 - \Big(\frac{\sin \pi u}{\pi u}\Big)^2 du + o(1)\Big) T \log T \end{aligned} \tag{6}$$

then it can be deduced that

$$\int_1^X (\psi((1 + \delta)x) - \psi(x) - \delta x)^2 x^{-2} \, dx \sim \delta(\log 1/\delta)(\log X\sqrt{\delta}) \tag{7}$$

for $X^{-1} \leqslant \delta \leqslant X^{-\varepsilon}$. Here γ denotes the ordinate of a non-trivial zero of the Riemann zeta function. Thus it seems likely that the estimate of Theorem 1 is best possible.

In the course of formulating the conjecture (6), Montgomery [13] also proposed a more precise estimate, namely that

$$F(X,T) \sim \frac{1}{2\pi} T \log T \tag{8}$$

uniformly for $T \leqslant X \leqslant T^A$, for any fixed $A > 1$, where

$$F(X,T) = \sum_{0 < \gamma,\gamma' \leqslant T} X^{i(\gamma-\gamma')} w(\gamma-\gamma') \tag{9}$$

and $w(u) = 4/(4 + u^2)$. We now relate this conjecture to the size of the integral in (3).

Theorem 2. *Assume* RH. *If* $0 < B_1 \leqslant B_2 \leqslant 1$, *then*

$$\int_1^X (\psi((1 + \delta)x) - \psi(x) - \delta x)^2 \, dx \sim \frac{1}{2} \delta X^2 \log 1/\delta \tag{10}$$

uniformly for $X^{-B_2} \leqslant \delta \leqslant X^{-B_1}$, *provided that* (8) *holds uniformly* *for*

$$X^{B_1} (\log X)^{-3} \leqslant T \leqslant X^{B_2} (\log X)^3. \qquad (11)$$

Conversely, if $1 \leqslant A_1 \leqslant A_2 < \infty$, *then* (8) *holds uniformly for* $T^{A_1} \leqslant X \leqslant T^{A_2}$, *provided that* (10) *holds uniformly for*

$$X^{-1/A_1} (\log X)^{-3} \leqslant \delta \leqslant X^{-1/A_2} (\log X)^3 . \qquad (12)$$

Previously Mueller [12] derived (10) from RH and a strong quantitative form of (8). Heath-Brown and Goldston [11] showed that RH and (8) for $T^a \leqslant X \leqslant T^b$, $a < 2 < b$, imply

$$p_{n+1} - p_n = o(p_n^{1/2} (\log p_n)^{1/2}) .$$

This estimate follows easily from Theorem 2 by taking $\delta = \varepsilon X^{-1/2} (\log X)^{1/2}$ in (10). In deriving (10) from (8) we also use the weaker estimate (3). In the case of very small δ, say $\delta \approx (\log X)/X$, we can do better by appealing instead to the bound

$$\int_1^X (\psi((1 + \delta)x) - \psi(x) - \delta x)^2 \, dx \ll \delta X^2 \log X + \delta^2 X^3 \qquad (13)$$

which follows from sieve estimates (see the proof of Lemma 7). In this way we could show that

$$\int_1^X (\pi(x + h) - \pi(x) - h/\log x)^2 \, dx \sim hX/\log X \qquad (14)$$

for $h \approx \log X$, given RH and (8) for $T \leqslant X \leqslant f(T)T \log T$. Here $f(T)$ tends to infinity arbitrarily slowly with T. From this it follows easily that

$$\liminf\, (p_{n+1} - p_n) / \log p_n = 0 .$$

Heath-Brown [10] derived this from a slightly stronger hypothesis.

In assessing the depth of the estimates (8) and (10), we note that (10) is a logarithm sharper than (3), and that (8) is a logarithm sharper than the trivial bound

$$|F(X,T)| \leq F(1,T) \sim \frac{1}{2\pi} T(\log T)^2 \quad . \tag{15}$$

(See Lemma 8.) As in (4), we can relate (10) to primes in intervals of constant length. In summary we have the following

Corollary. *Assume* RH. *Then the following assertions are equivalent*:

(a) *For every fixed* $A \geq 1$, (8) *holds uniformly for* $T \leq X \leq T^A$.

(b) *For every fixed* $\varepsilon > 0$, (10) *holds uniformly for* $X^{-1} \leq \delta \leq X^{-\varepsilon}$.

(c) *For every fixed* $\varepsilon > 0$,

$$\int_0^X (\psi(x + h) - \psi(x) - h)^2 \, dx \sim hX \log X/h \tag{16}$$

holds uniformly for $1 \leq h \leq X^{1-\varepsilon}$.

It is not hard to show that either (b) or (c) implies RH. Gallagher [4] has shown that a weak quantitative form of the prime k-tuple hypothesis gives (16) when $h \approx \log X$.

The path we take between (8) and (10) involves elementary arguments of Abelian and Tauberian character; these are of two sorts. First, we consider the connection between the assertion

$$\int_{-\infty}^{+\infty} e^{-2|y|} f(Y + y) \, dy = 1 + o(1) \tag{17}$$

as $Y \to +\infty$, and the more general assertion

$$\int_a^b R(y) f(Y + y) \, dy = \int_a^b R(y) \, dy + o(1) \tag{18}$$

as $Y \to +\infty$ where R is any Riemann-integrable function. (These two statements are equivalent if f is bounded and non-negative.) This interplay reflects the choice of the weighting function w(u) in the definition (9) of F(X,T). Second, and more intrinsically, we consider a question of Riemann summability (R_2), namely the connection between the two assertions

$$\int_0^\infty \left(\frac{\sin \kappa u}{u}\right)^2 f(u)du = (\pi/2 + o(1))\kappa \log 1/\kappa \tag{19}$$

as $\kappa \to 0^+$, and

$$\int_0^U f(u)du = (1 + o(1))\, U \log U \tag{20}$$

as $U \to +\infty$. Because of the intricacies of the (R_2) method, neither of these assertions implies the other, although they are equivalent for non-negative functions f. The lemmas we formulate below are complicated by the fact that we specify the relation between the parameters κ and U.

2. Lemmas of summability.

Lemma 1. *If*

$$I(Y) = \int_{-\infty}^{+\infty} e^{-2|y|} f(Y + y)dy = 1 + \varepsilon(Y) ,$$

and if $f(y) \geq 0$ *for all* y, *then for any Riemann-integrable function* R(y),

$$\int_a^b R(y)\, f(Y + y)dy = \left(\int_a^b R(y)dy\right) \left(1 + \varepsilon'(y)\right) . \tag{21}$$

If R *is fixed then* $|\varepsilon'(Y)|$ *is small provided that* $|\varepsilon(y)|$ *is small uniformly for* $Y + a - 1 \leq y \leq Y + b + 1$.

In terms of Wiener's general Tauberian theorem, the truth of this lemma hinges on the fact that the Fourier transform of the kernel $k(y) = e^{-2|y|}$, namely the function

$$\hat{k}(t) = \int_{-\infty}^{+\infty} k(y)\, e(-ty)dy = \frac{1}{\pi^2 t^2 + 1} , \quad \left(e(u) = e^{2\pi i u}\right),$$

never vanishes.

Proof. Let $K_c(y) = \max(0, c - |y|)$. By comparing Fourier transforms, or by direct calculation, we see that

$$K_c(y) = \frac{1}{2} e^{-2|y|} - \frac{1}{4} e^{-2|y-c|} - \frac{1}{4} e^{-2|y+c|} + \int_{-c}^{c} (c - |z|)\, e^{-2|y-z|}\, dz \, .$$

Hence

$$\int_{-c}^{c} K_c(y)\, f(Y + y)\, dy = \frac{1}{2} I(Y) - \frac{1}{4} I(Y + c) - \frac{1}{4} I(Y - c) + \int_{-c}^{c} (c - |z|) I(Y + z)\, dz = c^2 + \varepsilon_1(Y)$$

where $|\varepsilon_1|$ is small if $c > 0$ is fixed and if $|\varepsilon(y)|$ is small for $Y - c \leqslant y \leqslant Y + c$. Since

$$\frac{1}{\eta}\big(K_c(y) - K_{c-\eta}(y)\big) \leqslant \chi_{[-c,c]}(y) \leqslant \frac{1}{\eta}\big(K_{c+\eta}(y) - K_c(y)\big) \, ,$$

and since $f \geqslant 0$, we deduce that (21) holds in the case of the step function $R(y) = \chi_{[-c,c]}(y)$. Since the general R can be approximated above and below by step functions, we obtain (21).

Lemma 2. *Suppose that* $f(t)$ *is a continuous non-negative function defined for all* $t \geqslant 0$, *with* $f(t) \ll \log^2 (t + 2)$. *If*

$$J(T) = \int_0^T f(t)dt = \big(1 + \varepsilon(T)\big)T \log T \, ,$$

then

$$\int_0^\infty \Big(\frac{\sin \kappa u}{u}\Big)^2 f(u)du = \big(\pi/2 + \varepsilon'(\kappa)\big)\kappa \log 1/\kappa \qquad (22)$$

where $|\varepsilon'(\kappa)|$ *is small as* $\kappa \to 0^+$ *if* $|\varepsilon(T)|$ *is small uniformly for* $\kappa^{-1} (\log \kappa)^{-2} \leqslant T \leqslant \kappa^{-1}(\log \kappa)^2$.

Proof. We divide the range of integration in (22) into four subintervals: $0 \leqslant u \leqslant \kappa^{-1}(\log \kappa)^{-2} = U_1$, $U_1 \leqslant u \leqslant C\kappa^{-1} = U_2$, $U_2 \leqslant u \leqslant \kappa^{-1}(\log \kappa)^2 = U_3$, and $U_3 \leqslant u < \infty$.
Since $f(t) \ll \log^2(t + 2)$, we see that

$$\int_0^{U_1} < \int_0^{U_1} \kappa^2 \log^2(u+2)\, du < \kappa^2\, U_1 \log^2 U_1 \ll \kappa \quad ,$$

and similarly that

$$\int_{U_3}^{\infty} < \int_{U_3}^{\infty} u^{-2} \log^2 u\, du \ll U_3^{-1} \log^2 U_3 < \kappa \quad .$$

By writing $f(u) = \log 1/\kappa + \log \kappa u + (f(u) - \log u)$, we express the integral from U_1 to U_2 as a sum of three integrals. We note that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 du = \int_0^{\infty} \left(\frac{\sin \kappa u}{u}\right)^2 du + O\left(\kappa(\log \kappa)^{-2}\right)$$

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

and that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 \log \kappa u\, du < \int_0^{\infty} \min(\kappa^2 , u^{-2}) \log \kappa u\, du < \kappa \quad .$$

Put $r(u) = J(u) - u \log u + u$. Then by integrating by parts we see that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 (f(u) - \log u)du < \kappa\Big(1 + (\log \frac{1}{\kappa}) \max_{U_1 \leqslant u \leqslant U_2} |\varepsilon(u)|\Big)\log(C+2) \;.$$

As for the range $U_2 \leqslant u \leqslant U_3$, we see that if $\varepsilon(u) \leqslant 1$ then

$$\int_{U_2}^{U_3} \leqslant \int_{U_2}^{U_3} f(u) u^{-2}\, du \ll U_2^{-1} \log U_2 \ll C^{-1} \kappa \log 1/\kappa \;.$$

We make this small by taking C large. Then the remaining error terms are small if $\varepsilon(u)$ is small.

Lemma 3. *If* K *is even,* K" *continuous,* $\int_{-\infty}^{+\infty} |K| < \infty$, $K(x) \to 0$ *as* $x \to +\infty$, $K' \to 0$ *as* $x \to +\infty$, *and if* $K''(x) \ll x^{-3}$ *as* $x \to +\infty$, *then*

$$\hat{K}(t) = \int_0^{\infty} K''(x)\left(\frac{\sin \pi t x}{\pi t}\right)^2 dx \;. \tag{23}$$

Proof. Integrate by parts twice.

Lemma 4. *If f is a non-negative function defined on* $[0, +\infty)$, $f(t) \ll \log^2(t + 2)$, *and if*

$$I(\kappa) = \int_0^\infty \left(\frac{\sin \kappa t}{t}\right)^2 f(t)dt = (\pi/2 + \varepsilon(\kappa))\kappa \log 1/\kappa$$

then

$$J(T) = \int_0^T f(t)dt = (1 + \varepsilon')T \log T$$

where $|\varepsilon'|$ *is small if* $|\varepsilon(\kappa)| \leqslant \varepsilon$ *uniformly for*

$$T^{-1}(\log T)^{-1} \leqslant \kappa \leqslant T^{-1}(\log T)^2 .$$

Proof. Let K be a kernel with the properties specified in Lemma 3. Replace t by t/T in (23), multiply by $f(t) - \log t$, and integrate over $0 \leqslant t < \infty$. Then we find that

$$\int_0^\infty \left(f(t) - \log t\right) \hat{K}(t/T)dt = \pi^{-2}\, T^2 \int_0^\infty K''(x)\, R(\pi x/T)dx$$

where

$$R(\kappa) = I(\kappa) - \int_0^\infty \left(\frac{\sin \kappa t}{t}\right)^2 \log t \; dt$$

$$= I(\kappa) - \frac{1}{2}\,\pi\kappa \log 1/\kappa + O(\kappa).$$

Since

$$I(\kappa) \ll \int_0^\infty \min(\kappa^2, t^{-2})\log^2(t + 2)dt \ll \kappa \log^2(2 + 1/\kappa)$$

for all $\kappa > 0$, on taking $x_1 = (\log T)^{-1}$ we see that

$$\int_0^{x_1} K''R \ll \int_0^{x_1} xT^{-1} \log^2 T/x \; dx \ll T^{-1}.$$

On taking $x_2 = \frac{1}{4}(\log T)^2$ we find that

$$\int_{x_2}^\infty K''R \ll \int_{x_2}^\infty x^{-3}(x/T) \log^2 T \; dx \ll T^{-1}.$$

Assuming, as we may, that $\varepsilon \geqslant (\log T)^{-1}$, we have $R(\pi x/T) \ll \varepsilon x T^{-1} \log T$ for $x_1 \leqslant x \leqslant x_2$. Hence

$$\int_{x_1}^{x_2} K''R \ll \varepsilon T^{-1}(\log T) \int_0^\infty \min(1, x^{-3})\, x\, dx \ll \varepsilon T^{-1} \log T.$$

For $\eta > 0$ take

$$K(x) = K_\eta(x) = \big(\sin 2\pi x + \sin 2\pi(1+\eta)x\big)\big(2\pi x(1 - 4\eta^2 x^2)\big)^{-1},$$

so that

$$\hat{K}(t) = \begin{cases} 1 & \text{if } |t| \leqslant 1, \\ \cos^2\big(\pi(|t| - 1)/(2\eta)\big) & \text{if } 1 \leqslant |t| \leqslant 1 + \eta, \\ 0 & \text{if } |t| \geqslant 1 + \eta . \end{cases}$$

Thus

$$\int_0^\infty f(t)\hat{K}_\eta(t/T)dt = (1 + O(\eta))T \log T + O_\eta(T) + O_\eta(\varepsilon T \log T) .$$

Since f is non-negative, we see that

$$\int_0^\infty f(t)\hat{K}_\eta\big((1+\eta)t/T\big)dt \leqslant J(T) \leqslant \int_0^\infty f(t)\hat{K}_\eta(t/T)dt ,$$

and we obtain the desired result by taking η small.

In this argument we have made free use of existing treatments of Riemann summability. We note especially Hardy [8, pp. 301, 316, 365] and Hardy and Rogosinski [9, Theorem III].

3. Lemmas of analytic number theory.

As is customary, we write $s = \sigma + it$, and we let $\rho = \beta + i\gamma$ be a typical non-trivial zero of the Riemann zeta function. We first note a simple result of Gallagher [3]:

Lemma 5. *Let* $S(t) = \sum_{\mu \in M} c(\mu)e(\mu t)$ *where* M *is a countable set of real numbers and* $\sum |c(\mu)| < \infty$. *Then*

$$\int_{-T}^{T} |S(t)|^2\, dt < T^2 \int_{-\infty}^{+\infty} \Big| \sum_{\substack{\mu \in M \\ |\mu-u| \leq (4T)^{-1}}} c(\mu)\Big|^2\, du.$$

When a main term is desired, we use the following more elaborate estimate.

Lemma 6. *Let* $S(t)$ *be as above. If* $\delta \geq T^{-1}$ *then*

$$\int_0^T |S(t)|^2\, dt = (T + O(\delta^{-1})) \sum_{\mu \in M} |c(\mu)|^2$$

$$+ O\Big(T \sum_{\substack{\mu,\nu \in M \\ 0 < |\mu-\nu| < \delta}} |c(\mu)c(\nu)|\Big) .$$

Proof. Selberg (see Vaaler[17]) has constructed functions $F_-(t)$ and $F_+(t)$ such that $F_-(T) \leq \chi_{[0,T]}(t) \leq F_+(t)$, $\hat{F}_{\pm}(x) = 0$ for $|x| \geq \delta$, and $\int_{-\infty}^{+\infty} F_{\pm}(t)dt = T \pm \delta^{-1}$. Hence

$$\int_0^T |S|^2 \leq \int_{-\infty}^{+\infty} |S|^2\, F_+ = \sum_{\mu,\nu} c(\mu)\overline{c(\nu)}\, \hat{F}_+(\nu - \mu) .$$

The terms $\mu = \nu$ contribute $(T + \delta^{-1}) \sum_{\mu} |c(\mu)|^2$. Since

$$|\hat{F}_+| \leq \int |F_+| = T + \delta^{-1} \leq 2T ,$$

the terms $\mu \neq \nu$ contribute at most

$$2T \sum_{0 < |\mu-\nu| < \delta} |c(\mu)c(\nu)|.$$

This gives an upper bound, and a corresponding lower bound is derived similarly using F_-.

Lemma 7. *Let* $C(x) > 0$ *be a continuous function such that* $C(x) \approx C(y)$ *whenever* $x \approx y$. If $|c(p)| \leqslant C(p)$ *for all primes* p, *and if* $\delta \geqslant T^{-1}$, *then*

$$\int_0^T \left| \sum_p c(p)p^{it}\right|^2 dt = (T + O(\delta^{-1})) \sum_p |c(p)|^2 + O\Big(\delta T \int_{\delta^{-1}}^{\infty} C(u)^2\, u(\log u)^{-2}\, du\Big)$$

Proof. We appeal to the previous lemma. In the second error term, the primes $p \in (X,2X]$ contribute

$$T\, C(X)^2 \sum_{X<p\leqslant 2X}\ \sum_{p<p'\leqslant(1+2\delta)p} 1 \ll T\, C(X)^2 \sum_{1\leqslant k\leqslant 4\delta X} \pi_2(2X,k)$$

where $\pi_2(x,k)$ denotes the number of primes $p \leqslant x$ for which $p + k$ is also prime. It is well-known (see Halberstam and Richert [7, p.117]) that

$$\pi_2(x,k) \ll \big(k/\phi(k)\big)x(\log x)^{-2}$$

uniformly for $x \geqslant 2$, $k \neq 0$. Since $\sum_{k\leqslant K} k/\phi(k) \ll K$, it follows that our upper bound is

$$\ll T\, C(X)^2\, \delta X^2\, (\log X)^{-2} \ll \delta T \int_X^{2X} C(u)^2\, u(\log u)^{-2}\, du\ .$$

We put $X = \delta^{-1}\, 2^r$ and sum over $r \geqslant 0$ to obtain the desired result.

We now present the main known properties of $F(X,T)$.

Lemma 8. *Assume* RH, *and let* $F(X,T)$ *be as in* (9). *Then* $F(X,T) \geqslant 0$, $F(X,T) = F(1/X,T)$, *and*

$$F(X,T) = T\big(X^{-2}(\log T)^2 + \log X\big)\Big(\frac{1}{2\pi} + O\big((\log T)^{-1/2}\, (\log\log T)^{1/2}\big)\Big) \tag{24}$$

uniformly for $1 \leqslant X \leqslant T$.

Proof. The first assertion is an immediate consequence of either of the two identities

$$F(X,T) = 2\pi \int_{-\infty}^{+\infty} e^{-4\pi|u|} \Big| \sum_{0<\gamma\leq T} X^{i\gamma} e(\gamma u) \Big|^2 du, \tag{25}$$

or

$$F(X,T) = \frac{2}{\pi} \int_{-\infty}^{+\infty} \Big| \sum_{0<\gamma\leq T} \frac{X^{i\gamma}}{1+(t-\gamma)^2} \Big|^2 dt.$$

The observation that F is non-negative has also been made by Mueller (unpublished). The second assertion is obvious from the definition of F. The estimate (24) is substantially due to Goldston [6, Lemma B], and may be proved by substituting an appeal to Lemma 7 in the argument of Montgomery [13].

Lemma 9. *If* $0 \leq h \leq T$ *then*

$$\#\{(\gamma,\gamma') : 0 \leq \gamma \leq T, |\gamma - \gamma'| \leq h\} \ll (1 + h\log T)T\log T. \tag{27}$$

Proof. We argue unconditionally, although if RH is assumed then the above follows easily from Lemma 8 (see (6) of Montgomery [13]). Let $N(T) = \#\{\gamma : 0 < \gamma \leq T\}$. Following Selberg, Fujii [2] showed that

$$\int_0^T \Big(N(t+h) - N(t) - \frac{1}{2\pi} h\log t\Big)^2 dt \ll T\log(2 + h\log T)$$

for $0 \leq h \leq 1$. Hence

$$\int_0^T \big(N(t+h) - N(t)\big)^2 dt \ll h^2 T(\log T)^2$$

for $(\log T)^{-1} \leq h \leq 1$. This gives (27) in this case. To derive (27) when $0 \leq h \leq (\log T)^{-1}$, it suffices to consider $h = (\log T)^{-1}$. As for the range $1 \leq h \leq T$, it suffices to use the bound

$$N(T+1) - N(T) \ll \log T \tag{28}$$

(see Titchmarch [16, p. 178]).

Lemma 10. *For* $0 < \delta \leq 1$ *let*

$$a(s) = ((1+\delta)^s - 1)/s \; . \tag{29}$$

If $|c(\gamma)| \leq 1$ *for all* γ *then*

$$\int_{-\infty}^{+\infty} |a(it)|^2 \, \Big| \sum_{\gamma} \frac{c(\gamma)}{1+(t-\gamma)^2} \Big|^2 \, dt = \int_{-\infty}^{+\infty} \Big| \sum_{|\gamma| \leq Z} \frac{a(1/2+i\gamma)c(\gamma)}{1+(t-\gamma)^2} \Big|^2 \, dt$$

$$+ O(\delta^2(\log 2/\delta)^3) + O(Z^{-1}(\log Z)^3) \tag{30}$$

provide that $Z \geq 1/\delta$.

Proof. By (28), the sum that occurs in the integral on the left is $\ll \log(2+|t|)$. Since

$$a(s) \ll \min(\delta, |s|^{-1}) \tag{31}$$

in the strip $|\sigma| \leq 1/\delta$, it follows by Cauchy's formula or by direct calculation that

$$a'(s) \ll \min(\delta^2 \, , \, \delta/|s|) \tag{32}$$

for $|\sigma| \leq (2\delta)^{-1}$. Hence in particular,

$$a(it) - a(\tfrac{1}{2} + it) \ll \min(\delta^2 \, , \, \delta/|t|) \; ,$$

and consequently

$$|a(it)|^2 - |a(\tfrac{1}{2} + it)|^2 \ll \min(\delta^3, \, \delta/t^2) \; .$$

Let I denote the integral on the left in (30), and J the corresponding integral with $a(it)$ replaced by $a(\frac{1}{2} + it)$. Then

$$I - J \ll \int \min(\delta^3, \, \delta/t^2)(\log(2+|t|))^2 \, dt \ll \delta^2 (\log 2/\delta)^2 \; .$$

Write J in the form $J = \int |A|^2$. From (28) and (31) we see that

$$A \ll \min(\delta, |t|^{-1}) \log(2 + |t|) \tag{33}$$

Now let K be the integral with $a(\tfrac{1}{2} + it)$ replaced by $a(\tfrac{1}{2} + i\gamma)$, and write $K = \int |B|^2$. Then B also satisfies the estimate (33). From (31) and (32) we see that

$$a(\tfrac{1}{2} + i\gamma) - a(\tfrac{1}{2} + it) \ll |t - \gamma| \min (\delta^2, \delta/|t|) .$$

Thus

$$A - B \ll \min(\delta^2, \delta/|t|)(\log(2/\delta + |t|))^2,$$

so that

$$|A|^2 - |B|^2 \ll \min(\delta^3, \delta/t^2)(\log (2/\delta + |t|))^3 ,$$

and hence

$$J - K \ll \delta^2(\log 2/\delta)^3 .$$

Finally, let $L = \int |C|^2$ be the integral on the right in (30). We note that C also satisfies the estimate (33). Since

$$B - C \ll \min(Z^{-1}, |t|^{-1}) \log(2Z + |t|),$$

we find that

$$|B|^2 - |C|^2 \ll \min(Z^{-1}(1 + |t|)^{-1}, t^{-2})(\log(2Z + |t|))^2 .$$

Thus

$$K - L \ll Z^{-1} (\log 2Z)^3 ,$$

and the proof is complete.

4. Proof of Theorem 1.

Although we arrange the technical details differently, the ideas are entirely the same as in Selberg's paper. If $\delta X \leq 1$ then there is at most one prime power in the interval $(x, (1 + \delta)x]$, so that our integral is

$$\ll \delta \sum_{n\leq X} \Lambda(n)^2 /n + \delta^2 X \ll \delta(\log X)^2 \quad ,$$

which suffices. We now suppose that $\delta X > 1$. By the above argument we see that

$$\int_0^{1/\delta} \ldots \ll \delta(\log 2/\delta)^2.$$

Thus it suffices to consider the range $1/\delta \leq x \leq X$. Here we apply the explicit formula for $\psi(x)$ (see Davenport [1, 17]), which gives

$$\begin{aligned}\psi\big((1+\delta)x\big) - \psi(x) - \delta x = &- \sum_{|\rho|\leq Z} a(\rho)x^{\rho} \\ &+ O\Big((\log x)\min(1, \frac{x}{Z\,\|\, x \,\|})\Big) \\ &+ O\Big((\log x)\min(1, \frac{x}{Z\,\|\, (1+\delta)x \,\|})\Big) \\ &+ O(x\, Z^{-1}(\log xZ)^2)\end{aligned} \tag{34}$$

where $a(s)$ is given in (29), and $\|\,\theta\,\| = \min_n \|\theta - n\|$ is the distance from θ to the nearest integer. The error terms contribute a negligible amount if we take $Z = X(\log X)^2$. Writing $\rho = \frac{1}{2} + i\gamma$, $x = e^y$, $Y = \log X$, we see that it remains to show that

$$\int_{\log 1/\delta}^{Y} \Big| \sum_{|\gamma|\leq Z} a(\rho)\, e^{i\gamma y}\Big|^2 dy \ll \delta Y \log 2/\delta. \tag{35}$$

By Lemma 5 we see that this integral is

$$\ll Y^2 \int_{-\infty}^{\infty} \Big(\sum_{\substack{|\gamma|\leq Z \\ |\gamma-2\pi u|\leq 2/Y}} a(\rho)^2\Big)\, du \ll Y \sum_{\substack{|\gamma|\leq Z \\ |\gamma'|\leq Z \\ |\gamma-\gamma'|\leq 4/Y}} |a(\rho)a(\rho')| \;.$$

By (31) and Lemma 9 this gives (35), and the proof is complete.

5. Proof of Theorem 2.

We first assume (8) as needed, and derive (10). Let

$$J(T) = J(X,T) = 4 \int_0^T \left| \sum_\gamma \frac{X^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt \quad .$$

Montgomery [13] (see his (26), but beware of the changes in notation) used (28) to show that

$$J(X,T) = 2\pi\ F(X,T) + O\big((\log T)^3\big) \ .$$

Thus (8) is equivalent to

$$J(X,T) = \big(1 + o(1)\big)T \log T \quad . \tag{36}$$

With a(s) defined in (29), we note that

$$|a(it)|^2 = 4\left(\frac{\sin \kappa t}{t}\right)^2$$

where $\kappa = \frac{1}{2}\log (1+\delta)$. Then by Lemma 2 we deduce that

$$\int_0^\infty |a(it)|^2 \left|\sum_\gamma \frac{X^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt = (\pi/2 + o(1))\kappa \log 1/\kappa$$
$$= (\pi/4 + o(1))\delta \log 1/\delta \ . \tag{37}$$

The values of T for which we have used (8) lie in the range

$$\delta^{-1}(\log 1/\delta)^{-2} \leqslant T \leqslant 3\delta^{-1}(\log 1/\delta)^2 \ . \tag{38}$$

The integrand is even, so that the value is doubled if we integrate over negative values of t as well. Then by Lemma 10

$$\int_{-\infty}^{+\infty} \left| \sum_{|\gamma| \leqslant Z} \frac{a(\rho)\ X^{i\gamma}}{1+(t-\gamma)^2}\right|^2 dt = \big(\pi/2 + o(1)\big)\delta \log 1/\delta$$

provided that $Z \geqslant \delta^{-1}(\log 1/\delta)^3$. Let S(t) denote the above sum over γ. Its Fourier transform is

$$\hat{S}(u) = \int_{-\infty}^{+\infty} S(t)\ e(-tu)dt = \pi \sum_{|\gamma| \leqslant Z} a(\rho)\ X^{i\gamma}\ e(-\gamma u)e^{-2\pi|u|} \ .$$

Hence by Plancherel's identity the integral above is

$$= \pi^2 \int_{-\infty}^{+\infty} \left| \sum_{|\gamma| \leqslant Z} a(\rho)X^{i\gamma}\ e(-\gamma u)\right|^2 e^{-4\pi|u|}\ du \ .$$

On writing $Y = \log X$, $-2\pi u = y$, we find that

$$\int_{-\infty}^{+\infty} \Big| \sum_{|\gamma| \leq Z} a(\rho)\, e^{i\gamma(Y+y)} \Big|^2 e^{-2|y|}\, dy = (1 + o(1))\delta \log 1/\delta \,. \tag{39}$$

In Lemma 1 we take

$$R(y) = \begin{cases} e^{2y} & 0 \leq y \leq \log 2, \\ 0 & \text{otherwise} . \end{cases}$$

On making the change of variable $x = e^{Y+y}$ we deduce that

$$\int_X^{2X} \Big| \sum_{|\gamma| \leq Z} a(\rho) x^{\rho} \Big|^2 dx = \big(3/2 + o(1) \big)\, \delta X^2 \log 1/\delta \,.$$

We replace X by $X2^{-k}$, sum over k, $1 \leq k \leq K$, and use the explicit formula (34) with $Z = X(\log X)^3$ to see that

$$\int_{X2^{-K}}^{X} \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2\, dx = \tfrac{1}{2} \big(1-2^{-2K} + o(1)\big)\, \delta X^2 \log 1/\delta.$$

We take $K = [\log\log X]$, and note that it suffices to have (8) in the range (11). To bound the contribution of the range $1 \leq x \leq X2^{-K}$, we appeal to (3) with X replaced by $X2^{-K}$. Thus we have (10) .

We now deduce (8) from (10). By integrating (10) by parts from X_1 to $X_2 = X_1(\log X_1)^{2/3}$, we find that

$$\int_{X_1}^{X_2} \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2\, x^{-4} dx = \big(\tfrac{1}{2} + o(1)\big)\delta(\log 1/\delta) X_1^{-2}.$$

From (3) we similarly deduce that

$$\int_{X_2}^{\infty} \big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2 x^{-4}\, dx \ll \delta\, (\log 1/\delta)^2\, X_2^{-2}$$

$$= O\big(\delta\, (\log 1/\delta)\, X_1^{-2} \big) \,.$$

We add these relations, and multiply through by X_1^2 . By making a further appeal to (10) with $X = X_1$ we deduce that

$$\int_0^{\infty} \min(x^2/X_1^2,\, X_1^2/x^2)\big(\psi((1+\delta)x) - \psi(x) - \delta x\big)^2\, x^{-2}\, dx$$

$$= (1 + o(1))\delta \log 1/\delta \,.$$

We write X for X_1, put $Y = \log X$, $x = e^{Y+y}$, and appeal to the explicit formula (34) with $Z = X(\log X)^3$, and we find that we have (39). Retracing our steps, we find that we have (37). Then by Lemma 4 we obtain (36), and hence (8). The values of δ and X for which we have used (10) also satisfy (12).

6. Proof of the Corollary.

We note that Lemma 8 gives (8) when

$$X(\log X)^{-3} \leqslant T \leqslant X,$$

and that (10) is trivial when

$$X^{-1}(\log X)^{-3} \leqslant \delta \leqslant X^{-1}.$$

Thus the equivalence of (a) and (b) follows immediately from Theorem 2.

We now show that (b) implies (c). We suppress the converse argument, which is similar. The method here is that of Saffari and Vaughan [14]. Our first goal is to deduce from (b) that

$$\int_0^H \int_0^X (\psi(x+h) - \psi(x) - h)^2 \, dxdh \sim \frac{1}{2} H^2 X \log X/H \tag{40}$$

uniformly for $1 \leqslant H \leqslant X^{1-\varepsilon}$. To this end it suffices to show that

$$\int_{\frac{1}{2}X}^X \int_0^H (\psi(x+h) - \psi(x) - h)^2 \, dh \, dx \sim \frac{1}{4} H^2 X \log X/H \tag{41}$$

In this integral we replace h by $\delta = h/x$, and invert the order of integration. Thus the left hand side above is

$$\int_0^{H/X} \int_{\frac{1}{2}X}^X f(x,\delta x)^2 \, x \, dx \, d\delta + \int_{H/X}^{2H/X} \int_{\frac{1}{2}X}^{H/\delta} f(x,\delta x)^2 \, x \, dx \, d\delta$$

where $f(x,y) = \psi(x+y) - \psi(x) - y$. By integrating by parts, we see from (b) that if $A \approx B \approx X$ then

$$\int_A^B f(x,\delta x)^2 \, x \, dx = \frac{1}{3} (B^3 - A^3) \, \delta \log 1/\delta + O(X^3 \delta \log 1/\delta).$$

This yields (41) . Then (40) follows by replacing X by $X2^{-k}$ in (41), summing over $0 \leq k \leq K = [2 \log\log X]$, and by appealing to (4) with X replaced by $X2^{-K-1}$.

We now deduce (c) from (40). Suppose that $0 < \eta < 1$. By differencing in (40) we see that

$$\int_H^{(1+\eta)H} \int_0^X f(x,h)^2 \, dx \, dh = (\eta + \tfrac{1}{2}\eta^2 + o(1)) XH^2 \log X/H \quad .$$

Let $g(x,h) = f(x,H)$. From the identity

$$f^2 - g^2 = 2f(f-g) - (f-g)^2$$

and the Cauchy-Schwartz inequality we find that

$$\iint f^2 - g^2 \ll \left(\iint f^2\right)^{1/2} \left(\iint (f-g)^2\right)^{1/2} + \iint (f-g)^2 \quad .$$

But $f(x,h) - g(x,h) = f(x+H,h-H)$, so that

$$\begin{aligned} \iint (f-g)^2 &= \int_0^{\eta H} \int_H^{X+H} f(x,h)^2 \, dx \, dh \\ &\ll \eta^2 H^2 X \log X/H \end{aligned}$$

by (40). Hence we see that

$$\begin{aligned} \eta H \int_0^X (\psi(x+H) - \psi(x) - H)^2 \, dx &= \iint g^2 \\ &= \iint f^2 + O(\eta^{3/2} XH^2 \log X/H) \\ &= \left(\eta + O(\eta^{3/2}) + o(1)\right) XH^2 \log X/H \ . \end{aligned}$$

We now divide both sides by ηH, and obtain the desired result by letting $\eta \to 0^+$ sufficiently slowly.

References.

1. H. Davenport, Multiplicative Number Theory, Second Edition, Springer-Verlag, 1980.

2. A. Fujii, On the zeros of Dirichlet L-functions, I, *Trans. Amer. Math. Soc.* **196** (1974), 225-235. (Corrections to this paper are noted in *Trans. Amer. Math. Soc.* **267** (1981), pp 38-39, and in [5; pp. 219-220].)

3. P.X. Gallagher, A large sieve density estimate near $\sigma = 1$, *Invent. Math.* **11** (1970), 329-339.

4. P.X. Gallagher, On the distribution of primes in short intervals, *Mathematika* **23** (1976), 4-9.

5. P.X. Gallagher and Julia H. Mueller, Primes and zeros in short intervals, *J. Reine Agnew. Math.* **303/304** (1978), 205-220.

6. Daniel A. Goldston, Large differences between consecutive prime numbers, Thesis, University of California Berkeley, 1981.

7. H.Halberstam and H.-E. Richert, Sieve Methods. Academic Press, London, 1974.

8. G.H. Hardy, Divergent Series, Oxford University Press, 1963.

9. G.H. Hardy and W.W. Rogosinski, Notes on Fourier series (I): On sine series with positive coefficients, *J. London Math. Soc.* **18** (1943), 50-57.

10. D.R. Heath-Brown, Gaps between primes, and the pair correlation of zeros of the zeta-function, *Acta Arith.* **41** (1982), 85-99.

11. D.R. Heath-Brown and D.A. Goldston, A note on the difference between consecutive primes, *Math. Ann.* **266** (1984), 317-320.

12. Julia Huang (=J.H. Mueller), Primes and zeros in short intervals, Thesis, Columbia University, 1976.

13. H.L Montgomery, The pair correlation of zeros of the zeta function, *Proc. Sympos. Pure Math.* **24** (1973), 181-193.

14. B. Saffari and R.C. Vaughan, On the fractional Parts of x/n and related sequences II, *Ann. Inst. Fourier* (Grenoble) **27**, no. 2, (1977), 1-30.

15. A. Selberg, On the normal density of primes in small intervals, and the difference between consecutive primes, *Arch. Math. Naturvid.* **47**, no. 6, (1943), 87-105.

16. E.C. Titchmarsh, The theory of the Riemann zeta-function, Oxford University Press, 1951.

17. J.D. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.*, **12**, No. 2, (1985), 183-216.

D. A. Goldston
San Jose State University,
San Jose, CA 95192,
U.S.A.

H. L. Montgomery
University of Michigan,
Ann Arbor, MI 48109,
U.S.A.

Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée.

§ 1. Legendre, dans sa Théorie des nombres *), propose une formule pour déterminer combien il y a de nombres premiers depuis 1 jusqu'à une limite donnée. Il commence par comparer sa formule avec l'énumération immédiate des nombres premiers faite dans les tables les plus étendues, nommément depuis 10000 jusqu'à 1000000, et l'applique ensuite à la solution de plusieurs questions. Malgré la concordance prononcée de la formule de Legendre avec les tables des nombres premiers, nous nous permettons néanmoins d'élever quelques doutes sur son exactitude, et par conséquent aussi sur les résultats qu'on en a tirés. Nous fondons notre assertion sur un théorème, relatif aux propriétés de la fonction qui détermine combien il y a de nombres premiers inférieurs à une limite donnée, théorème dont on peut déduire plusieurs conséquences curieuses. Nous allons d'abord donner la démonstration du théorème en question, et nous en présenterons ensuite quelques applications.

I-er Théorème.

§ 2. *Si l'on représente par $\varphi(x)$ la totalité des nombres premiers inférieurs à x, par n un entier quelconque, enfin par ρ une quantité > 0, la somme*

$$\sum_{x=2}^{x=\infty}\left[\varphi(x+1)-\varphi(x)-\frac{1}{\log x}\right]\frac{\log^n x}{x^{1+\rho}}$$

jouira de la propriété de s'approcher d'une limite finie, à mesure que ρ converge vers zéro.

*) Tome 2, page 65 (Troisième édition).

Démonstration. Commençons par démontrer que la propriété en question a lieu pour les fonctions que l'on obtient par la différentiation successive des trois expressions

$$\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}, \quad \log \rho - \Sigma \log \left(1 - \frac{1}{\mu^{1+\rho}}\right),$$
$$\Sigma \log \left(1 - \frac{1}{\mu^{1+\rho}}\right) + \sum \frac{1}{\mu^{1+\rho}}$$

par rapport à ρ; ici, comme par la suite, la sommation par rapport à m s'étend à tous les entiers depuis $m = 2$ jusqu'à $m = \infty$, et par rapport à μ seulement aux nombres premiers, également depuis $\mu = 2$ jusqu'à $\mu = \infty$.

Considérons la première expression. Il est facile de voir que l'on a

$$\int_0^\infty \frac{e^{-x}}{e^x - 1} x^\rho \, dx = \sum \frac{1}{m^{1+\rho}} \cdot \int_0^\infty e^{-x} x^\rho \, dx,$$
$$\int_0^\infty e^{-x} x^{-1+\rho} \, dx = \frac{1}{\rho} \int_0^\infty e^{-x} x^\rho \, dx,$$

et par conséquent

$$\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho} = \frac{\int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x}\right) e^{-x} x^\rho \, dx}{\int_0^\infty e^{-x} x^\rho \, dx}.$$

En vertu de cette équation la dérivée d'un ordre quelconque n de $\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}$ par rapport à ρ sera égale à une fraction, dont le dénominateur est $\left[\int_0^\infty e^{-x} x^\rho \, dx\right]^{n+1}$ et le numerateur une fonction entière des expressions

$$\int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x}\right) e^{-x} x^\rho \, dx, \quad \int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x}\right) e^{-x} x^\rho \log x \, dx,$$
$$\int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x}\right) e^{-x} x^\rho \log^2 x \, dx, \ldots \int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x}\right) e^{-x} x^\rho \log^n x \, dx,$$
$$\int_0^\infty e^{-x} x^\rho \, dx, \int_0^\infty e^{-x} x^\rho \log x \, dx, \int_0^\infty e^{-x} x^\rho \log^2 x \, dx, \ldots \int_0^\infty e^{-x} x^\rho \log^n x \, dx.$$

Or, une telle fraction, pour $n = 0$ aussi bien que pour $n > 0$, s'approche d'une limite finie à mesure que ρ converge vers zéro; car la limite de l'intégrale $\int_0^\infty e^{-x}\, x^\rho\, dx$ pour $\rho = 0$ est 1, et les intégrales

$$\int_0^\infty \left(\frac{1}{e^x-1} - \frac{1}{x}\right) e^{-x}\, x^\rho\, dx, \quad \int_0^\infty \left(\frac{1}{e^x-1} - \frac{1}{x}\right) e^{-x}\, x^\rho \log x\, dx,$$

$$\int_0^\infty \left(\frac{1}{e^x-1} - \frac{1}{x}\right) e^{-x}\, x^\rho \log^2 x\, dx, \ldots \int_0^\infty \left(\frac{1}{e^x-1} - \frac{1}{x}\right) e^{-x}\, x^\rho \log^n x\, dx,$$

$$\int_0^\infty e^{-x}\, x^\rho \log x\, dx, \quad \int_0^\infty e^{-x}\, x^\rho \log^2 x\, dx, \ldots \int_0^\infty e^{-x}\, x^\rho \log^n x\, dx$$

pour $\rho = 0$ conservent évidemment des valeurs finies.

Ainsi, il est certain que la fonction $\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}$, aussi bien que ses dérivées successives, resteront finies à mesure que ρ convergera vers la limite zéro.

Considérons actuellement la fonction

$$\log \rho - \Sigma \log \left(1 - \frac{1}{\mu^{1+\rho}}\right).$$

On sait que

$$\left[\left(1 - \frac{1}{2^{1+\rho}}\right)\left(1 - \frac{1}{3^{1+\rho}}\right)\left(1 - \frac{1}{5^{1+\rho}}\right) \ldots\ldots\right]^{-1}$$
$$= 1 + \frac{1}{2^{1+\rho}} + \frac{1}{3^{1+\rho}} + \frac{1}{4^{1+\rho}} + \ldots\ldots;$$

d'où l'on tire

$$-\log\left(1 - \frac{1}{2^{1+\rho}}\right) - \log\left(1 - \frac{1}{3^{1+\rho}}\right) - \log\left(1 - \frac{1}{5^{1+\rho}}\right) \ldots$$
$$= \log\left(1 + \frac{1}{2^{1+\rho}} + \frac{1}{3^{1+\rho}} + \frac{1}{4^{1+\rho}} + \ldots\right),$$

équation qui, d'après la notation admise plus haut, peut être écrite de cette manière

$$-\Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right) = \log\left(1 + \sum \frac{1}{m^{1+\rho}}\right);$$

donc

$$\log \rho - \Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right) = \log\left(1 + \sum \frac{1}{m^{1+\rho}}\right)\rho;$$

ou bien

$$\log \rho - \Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right) = \log\left[1 + \rho + \left(\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}\right)\rho\right].$$

Cette équation fait voir que toutes les dérivées de

$$\log \rho - \Sigma \log \left(1 - \frac{1}{\mu^{1+\rho}}\right)$$

suivant ρ, s'exprimeront au moyen d'un nombre fini de fractions, dont les dénominateurs seront des puissances entières et positives de

$$1 + \rho + \left(\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}\right)\rho,$$

et les numérateurs des fonctions entières de ρ, de l'expression $\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}$ et de ses dérivées par rapport à ρ. Or, de telles fractions s'approcheront d'une limite finie à mesure que ρ convergera vers zéro; car l'expression $1+\rho+\left(\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}\right)\rho$, qui entre dans les dénominateurs de ces fractions, tendra vers la limite 1 à mesure que ρ s'approchera de zéro, et cela parceque la différence $\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}$, dans cette hypothèse, reste finie comme nous l'avons démontré plus haut. Quant à ce qui concerne les numérateurs, comme ils ne contiennent la différence $\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}$ et ses dérivées que sous forme entière, et que ces fonctions tendent vers une limite finie quand ρ converge vers zéro, il en sera de même pour ses numérateurs.

Il nous reste encore à démontrer que la même propriété à lieu relativement aux dérivées de la fonction

$$\Sigma \log \left(1 - \frac{1}{\mu^{1+\rho}}\right) + \sum \frac{1}{\mu^{1+\rho}}.$$

Nous remarquerons d'abord que sa première dérivée sera

$$\sum \frac{1}{\mu^{2+2\rho}} \cdot \frac{\log \mu}{1 - \frac{1}{\mu^{1+\rho}}}.$$

Il est facile de voir par la forme de cette fonction que les dérivées des ordres supérieurs s'exprimeront également au moyen d'un nombre fini des termes tels que

$$\sum \frac{1}{\mu^{2+2\rho}} \cdot \frac{\log^p \mu}{1 - \frac{1}{\mu^{1+\rho}}} \cdot \frac{1}{\mu^s \left(1 - \frac{1}{\mu^{1+\rho}}\right)^r},$$

p, q, r n'étant par inférieurs à zéro. Mais chaque terme de cette nature, pour des valeurs de ρ non-inférieures à zéro, a une valeur finie; en effet, pour $\rho = 0$ et $\rho > 0$, la fonction sous le signe Σ sera une quantité d'un ordre supérieur au premier par rapport à $\frac{1}{\mu}$.

Après nous être convaincu que les dérivées des trois expressions

$$\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}, \quad \log \rho - \Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right),$$
$$\Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right) + \sum \frac{1}{\mu^{1+\rho}},$$

pour des valeurs de ρ convergentes vers zéro, tendent vers des limites finies, nous concluons que la même propriété aura également lieu par rapport à l'expression

$$\frac{d^n\left[\Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right) + \sum \frac{1}{\mu^{1+\rho}}\right]}{d\rho^n} - \frac{d^n\left[\log \rho - \Sigma \log\left(1 - \frac{1}{\mu^{1+\rho}}\right)\right]}{d\rho^n} - \frac{d^{n-1}\left(\sum \frac{1}{m^{1+\rho}} - \frac{1}{\rho}\right)}{d\rho^{n-1}},$$

laquelle, après les différentiations effectuées, ce réduira à

$$\pm\left(\sum \frac{\log^n \mu}{\mu^{1+\rho}} - \sum \frac{\log^{n-1} m}{m^{1+\rho}}\right).$$

Ce qui vient d'être dit renferme le théorême énoncé plus haut, car il est facile de remarquer que, d'après notre notation, la différence

$$\sum \frac{\log^n \mu}{\mu^{1+\rho}} - \sum \frac{\log^{n-1} m}{m^{1+\rho}}$$

est identique avec l'expression

$$\sum_{x=2}^{x=\infty} \left[\varphi(x+1) - \varphi(x) - \frac{1}{\log x}\right] \frac{\log^n x}{x^{1+\rho}},$$

ou bien, ce qui revient au même, avec

$$\sum_{x=2}^{x=\infty} \left[\varphi(x+1) - \varphi(x)\right] \frac{\log^n x}{x^{1+\rho}} - \sum_{x=2}^{x=\infty} \frac{\log^{n-1} x}{x^{1+\rho}},$$

Pour le faire voir il n'y a qu'à observer que le premier terme de cette différence est simplement égal à $\sum \frac{\log^n \mu}{\mu^{1+\rho}}$, parceque le facteur $\varphi(x+1) - \varphi(x)$ de $\frac{\log^n x}{x^{1+\rho}}$ se réduit, par la définition même de la fonction φ, à 1 ou à 0 suivant que x est un nombre premier ou un nombre composé. Quant au second terme $\sum_{x=2}^{x=\infty} \frac{\log^{n-1} x}{x^{1+\rho}}$, il se transforme évidemment en $\sum \frac{\log^{n-1} m}{m^{1+\rho}}$ par le changement de x en m.

De cette manière la proposition que nous avions en vue de démontrer, se trouve complètement établie.

§ 3. Le théorème dont on vient de donner la démonstration conduit à plusieurs propriétés curieuses de la fonction qui détermine combien il y a de nombres premiers inférieurs à une limite donnée. Et d'abord observons que la différence

$$\frac{1}{\log x} - \int_x^{x+1} \frac{dx}{\log x},$$

pour x très grand, est une quantité infiniment petite du premier ordre par rapport à $\frac{1}{x}$; par conséquent l'expression

$$\left(\frac{1}{\log x} - \int_x^{x+1} \frac{dx}{\log x}\right) \frac{\log^n x}{x^{1+\rho}},$$

pour x très grand, sera de l'ordre $2 + \rho$ relativement à $\frac{1}{x}$; d'après cela, la somme

$$\sum_{x=2}^{x=\infty} \left(\frac{1}{\log x} - \int_x^{x+1} \frac{dx}{\log x}\right) \frac{\log^n x}{x^{1+\rho}}$$

pour des valeurs de ρ non-inférieures à zéro, restera finie. Ajoutant cette somme à l'expression

$$\sum_{x=2}^{x=\infty} \left[\varphi(x+1) - \varphi(x) - \frac{1}{\log x}\right] \frac{\log^n x}{x^{1+\rho}},$$

pour laquelle le théorème I-er a lieu, nous concluons que la valeur de

$$\sum_{x=2}^{x=\infty} \left[\varphi(x+1) - \varphi(x) - \int_x^{x+1} \frac{dx}{\log x}\right] \frac{\log^n x}{x^{1+\rho}}\, dx$$

restera finie à mesure que ρ convergera vers la limite zéro. De là on tire le théorème suivant:

II-ème Théorème.

La fonction $\varphi(x)$, qui désigne combien il y a de nombres premiers inférieurs à x, satisfera, entre les limites $x = 2$ et $x = \infty$, une infinité de fois aux deux inégalités

$$\varphi(x) > \int_2^x \frac{dx}{\log x} - \frac{\alpha x}{\log^n x} \quad \text{et} \quad \varphi(x) < \int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x},$$

quelque petite que soit la valeur de α, supposée positive, et quelque grand que soit en même temps le nombre n.

Démonstration. Nous nous contenterons de démontrer l'une de ces deux inégalités, parceque l'autre s'établira tout-à-fait de la même manière. Choisissons, par exemple, la suivante:

$$\varphi(x) < \int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}. \tag{1}$$

Pour prouver que cette inégalité est satisfaite une infinité de fois, admettons d'abord que le contraire ait lieu, et voyons quelles seront les conséquences de cette hypothèse. Soit a un entier supérieur à e^n et supérieur en même temps au plus grand nombre qui satisfait à l'inégalité (1). Dans cette supposition on aura pour $x > a$ l'inégalité

$$\varphi(x) \geqq \int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}, \quad \log x > n,$$

et par conséquent

$$\varphi(x) - \int_2^x \frac{dx}{\log x} \geqq \frac{\alpha x}{\log^n x}, \quad \frac{n}{\log x} < 1. \tag{2}$$

Or, si l'on admettait les inégalités (2), il en résulterait, contrairement à ce qui a été démontré plus haut, que l'expression

$$\sum_{x=2}^{x=\infty} \left[\varphi(x+1) - \varphi(x) - \int_x^{x+1} \frac{dx}{\log x} \right] \frac{\log^n x}{x^{1+\rho}},$$

au lieu de converger vers une limite finie pour des valeurs très petites de ρ, s'approcherait de la limite $+\infty$. En effet, nous pouvons considérer cette expression comme la limite de

$$\sum_{x=2}^{x=s} \left[\varphi(x+1) - \varphi(x) - \int_x^{x+1} \frac{dx}{\log x} \right] \frac{\log^n x}{x^{1+\rho}}$$

pour $s = \infty$. Supposant donc $s > a$, cette quantité peut être mise sous la forme

$$C + \sum_{x=a+1}^{x=s} \left[\varphi(x+1) - \varphi(x) - \int_x^{x+1} \frac{dx}{\log x} \right] \frac{\log^n x}{x^{1+\rho}}, \tag{3}$$

en faisant pour abréger

$$C = \sum_{x=2}^{x=a} \left[\varphi(x+1) - \varphi(x) - \int_x^{x+1} \frac{dx}{\log x} \right] \frac{\log^n x}{x^{1+\rho}},$$

et observant que C désignera une quantité finie pour $\rho = 0$ et $\rho > 0$.

Or, l'expression (3), en vertu de la formule connue

$$\sum_{a+1}^{s} u_x (v_{x+1} - v_x) = u_s v_{s+1} - u_a v_{a+1} - \sum_{a+1}^{s} v_x (u_x - u_{x-1}),$$

et après avoir fait

$$v_x = \varphi(x) - \int_2^x \frac{dx}{\log x}, \quad u_x = \frac{\log^n x}{x^{1+\rho}},$$

se transformera dans la suivante:

$$C - \left[\varphi(a+1) - \int_2^{a+1} \frac{dx}{\log x}\right] \frac{\log^n a}{a^{1+\rho}} + \left[\varphi(s+1) - \int_2^{s+1} \frac{dx}{\log x}\right] \frac{\log^n s}{s^{1+\rho}}$$

$$- \sum_{x=a+1}^{x=s} \left[\varphi(x) - \int_2^x \frac{dx}{\log x}\right] \left[\frac{\log^n x}{x^{1+\rho}} - \frac{\log^n (x-1)}{(x-1)^{1+\rho}}\right],$$

qui, à son tour, en faisant $\theta > 0$ et < 1, pourra s'écrire comme il suit:

$$C - \left[\varphi(a+1) - \int_2^{a+1} \frac{dx}{\log x}\right] \frac{\log^n a}{a^{1+\rho}} + \left[\varphi(s+1) - \int_2^{s+1} \frac{dx}{\log x}\right] \frac{\log^n s}{s^{1+\rho}} +$$

$$+ \sum_{x=a+1}^{x=s} \left[\varphi(x) - \int_2^x \frac{dx}{\log x}\right] \left[1 + \rho - \frac{n}{\log(x-\theta)}\right] \frac{\log^n (x-\theta)}{(x-\theta)^{2+\rho}}.$$

Si l'on représente par F la somme des deux premiers termes de cette expression, et si l'on observe de plus que le troisième est positif en vertu de la condition (2), on sera en droit de conclure que l'expression précédente a une valeur supérieure à

$$F + \sum_{x=a+1}^{x=s} \left[\varphi(x) - \int_2^x \frac{dx}{\log x}\right] \left[1 + \rho - \frac{n}{\log(x-\theta)}\right] \frac{\log^n (x-\theta)}{(x-\theta)^{2+\rho}}.$$

Les mêmes conditions (2) font voir que dans cette expression la fonction sous le signe Σ conservera une valeur positive entre les limites. En outre, on aura entre les limites de la sommation 1°) $1 + \rho - \frac{n}{\log(x-\theta)} > 1 - \frac{n}{\log a}$; car $\rho > 0$, $x \geqq a+1$, $\theta < 1$; 2°) $\varphi(x) - \int_2^x \frac{dx}{\log x} > \frac{\alpha(x-\theta)}{\log^n (x-\theta)}$; car $\varphi(x) - \int_2^x \frac{dx}{\log x} \geqq \frac{\alpha x}{\log^n x}$ en vertu de la première des inégalités (2), et en vertu de la seconde la dérivée de $\frac{\alpha x}{\log^n x}$, égale à $\frac{\alpha}{\log^n x}\left(1-\frac{n}{\log x}\right)$, est positive, ce qui donne $\frac{\alpha x}{\log^n x} > \frac{\alpha(x-\theta)}{\log^n (x-\theta)}$. Donc l'expression précédente surpasse la somme

$$F+\sum_{x=a+1}^{x=s} \frac{\alpha(x-\theta)}{\log^n(x-\theta)}\left(1-\frac{n}{\log a}\right)\frac{\log^n(x-\theta)}{(x-\theta)^{2+\rho}},$$

qui, après les réductions, devient

$$F+\alpha\left(1-\frac{n}{\log a}\right)\sum_{x=a+1}^{x=s}\frac{1}{(x-\theta)^{1+\rho}};$$

or, cette dernière expression est évidemment supérieure à celle-ci

$$F+\alpha\left(1-\frac{n}{\log a}\right)\sum_{x=a+1}^{x=s}\frac{1}{x^{1+\rho}},$$

laquelle, pour $s=\infty$, se réduit à

$$F+\alpha\left(1-\frac{n}{\log a}\right)\sum_{x=a+1}^{x=\infty}\frac{1}{x^{1+\rho}},$$

ou à

$$F+\alpha\left(1-\frac{n}{\log a}\right)\frac{\int_0^\infty \frac{e^{-ax}}{e^x-1}\,x^\rho\,dx}{\int_0^\infty e^{-x}x^\rho\,dx}.$$

Il est facile de faire voir que la quantité à laquelle nous sommes parvenus converge vers la limite $+\infty$ pour $\rho=0$. En effet, on a d'abord $\int_0^\infty \frac{e^{-ax}}{e^x-1}\,dx=+\infty$, $\int_0^\infty e^{-x}\,dx=1$, de plus α et $1-\frac{n}{\log a}$ sont toutes deux des quantités positives, la première par hypothèse, et la seconde en vertu de la dernière inégalité (2).

Nous étant assurés de cette manière que, dans l'hypothèse admise, non seulement la somme

$$\sum_{x=a}^{x=\infty}\left[\varphi(x+1)-\varphi(x)-\int_x^{x+1}\frac{dx}{\log x}\right]\frac{\log^n x}{x^{1+\rho}},$$

mais aussi une quantité plus petite qu'elle se réduit à $+\infty$, nous sommes en droit de conclure que l'hypothèse en question est inadmissible, d'où découle de suite la légitimité du théorème II.

§ 4. Il sera facile actuellement, en vertu de la proposition précédente, de démontrer le théorème qui suit:

III-éme Théorème.

L'expression $\frac{x}{\varphi(x)} - \log x$, *pour* $x = \infty$, *ne peut avoir une limite différente de* -1.

Démonstration. Soit L la limite de la différence $\frac{x}{\varphi(x)} - \log x$ pour $x = \infty$. Dans cette hypothése on pourra toujours trouver un nombre N tellement grand que pour $x > N$ la valeur de $\frac{x}{\varphi(x)} - \log x$ sera comprise entre les limites $L - \varepsilon$ et $L + \varepsilon$, ε étant aussi petite qu'on voudra. Ainsi, pour de semblables valeurs de x, et lorsque $\varepsilon > 0$, on aura

$$\frac{x}{\varphi(x)} - \log x > L - \varepsilon, \quad \frac{x}{\varphi(x)} - \log x < L + \varepsilon. \tag{4}$$

Mais, en vertu du théorème précédent, les inégalités

$$\varphi(x) > \int_2^x \frac{dx}{\log x} - \frac{\alpha x}{\log^n x}, \quad \varphi(x) < \int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}$$

sont satisfaites par une infinité de valeurs de x, et par conséquent aussi par des valeurs de x supérieures à N, pour lesquelles les inégalités (4) ont lieu. Or, ces inégalités, combinées avec celles que nous venons d'écrire, conduisent à

$$\frac{x}{\int_2^x \frac{dx}{\log x} - \frac{\alpha x}{\log^n x}} - \log x > L - \varepsilon, \quad \frac{x}{\int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}} - \log x < L + \varepsilon;$$

d'où l'on tire

$$L + 1 < \frac{x - (\log x - 1)\left(\int_2^x \frac{dx}{\log x} - \frac{\alpha x}{\log^n x}\right)}{\int_2^x \frac{dx}{\log x} - \frac{\alpha x}{\log^n x}} + \varepsilon,$$

$$L + 1 > \frac{x - (\log x - 1)\left(\int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}\right)}{\int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}} - \varepsilon.$$

On voit par ces inégalités que la valeur numérique de $L+1$ ne surpasse pas celle de l'une des expressions qui en forment les seconds membres. De plus ε peut devenir aussi petite qu'on voudra dans l'hypothèse de N très grand, et on peut en dire autant de chacune des quantités

$$\frac{x-(\log x-1)\left(\int_2^x \frac{dx}{\log x} \mp \frac{\alpha x}{\log^n x}\right)}{\int_2^x \frac{dx}{\log x} \mp \frac{\alpha x}{\log^n x}};$$

car, pour $x=\infty$, on trouve par les principes du calcul différentiel que leur limite commune est zéro.

Nous étant ainsi convaincus que les limites

$$\frac{x-(\log x-1)\left(\int_2^x \frac{dx}{\log x} \mp \frac{\alpha x}{\log^n x}\right)}{\int_2^x \frac{dx}{\log x} \mp \frac{\alpha x}{\log^n x}} \pm \varepsilon,$$

de la valeur numérique de $L+1$ peuvent être diminuées à volonté, nous sommes en droit de conclure que $L+1=0$, et par conséquent $L=-1$, ce qu'il s'agissait de démontrer.

Ce que nous venons de prouver relativement à la limite de la valeur de $\frac{x}{\varphi(x)}-\log x$, pour $x=\infty$, ne s'accorde pas avec une formule donnée par Legendre pour déterminer approximativement combien il y a de nombres premiers inférieurs à une limite donnée. D'après lui la fonction $\varphi(x)$, pour x très grand, est exprimée avec une approximation suffisante par la formule

$$\varphi(x)=\frac{x}{\log x-1{,}08366},$$

qui donne pour la limite de $\frac{x}{\varphi(x)}-\log x$ le nombre $-1{,}08366$ au lieu de -1.

§ 5. En partant du théorême II on peut déterminer la limite supérieure du dégre de précision avec lequel la fonction, désignée par $\varphi(x)$, peut être remplacée par toute autre fonction donnée $f(x)$. Dans ce qui va suivre nous comparerons la difference $f(x)-\varphi(x)$ avec les expressions

$$\frac{x}{\log x},\quad \frac{x}{\log^2 x},\quad \frac{x}{\log^3 x},\ldots$$

et, pour abréger le discours, nous dirons que A est *une quantité de l'ordre* $\frac{x}{\log^n x}$, quand le rapport de A à $\frac{x}{\log^m x}$ pour $x = \infty$, sera infini pour $m > n$ et zéro pour $m < n$. Cela posé, nous allons démontrer le théorême suivant:

IV-éme Théorème.

Quand l'expression

$$\frac{\log^n x}{x}\left(f(x) - \int_2^x \frac{dx}{\log x}\right),$$

pour $x = \infty$, a pour limite une quantité finie ou infinie, la fonction $f(x)$ ne peut représenter $\varphi(x)$ exactement en quantités de l'ordre $\frac{x}{\log^n x}$ inclusivement.

Démonstration. Soit L la limite vers laquelle converge l'expression

$$\frac{\log^n x}{x}\left(f(x) - \int_2^x \frac{dx}{\log x}\right)$$

à mesure que x s'approche de l'infini. Comme L, par hypothèse, est différente de zéro, elle ne pourra être égale qu'à une quantité positive ou négative. Supposons la positive; notre raisonnement s'appliquera sans difficulté au cas de $L < 0$.

Si la limite L de l'expression que nous considérons, pour $x = \infty$, est supérieure à zéro, nous pourrons trouver un nombre N assez grand et tel que, pour $x > N$, la valeur de l'expression

$$\frac{\log^n x}{x}\left(f(x) - \int_2^x \frac{dx}{\log x}\right)$$

reste constamment supérieure à une certaine quantité positive l.

Nous aurons donc pour $x > N$

$$\frac{\log^n x}{x}\left(f(x) - \int_2^x \frac{dx}{\log x}\right) > l. \tag{5}$$

Mais, en vertu du théorème II, quelque petit que soit $\alpha = \frac{l}{2}$, nous aurons pour un nombre infini de valeurs de x l'inégalité

$$\varphi(x) < \int_2^x \frac{dx}{\log x} + \frac{\alpha x}{\log^n x}, \tag{6}$$

qui donne

$$f(x) - \int_2^x \frac{dx}{\log x} < f(x) - \varphi(x) + \frac{\alpha x}{\log^n x};$$

en la multipliant par $\frac{\log^n x}{x}$, et observant que $\alpha = \frac{l}{2}$, on trouve

$$\frac{\log^n x}{x}\left[f(x) - \int_2^x \frac{dx}{\log x}\right] < \frac{\log^n x}{x}\left[f(x) - \varphi(x)\right] + \frac{l}{2}$$

ou bien, en vertu de l'négalité (5),

$$\frac{\log^n x}{x}\left[f(x) - \varphi(x)\right] > \frac{l}{2}.$$

Or, cette inégalité ayant lieu en même temps que celles marquées par les numéros (5) et (6) pour une infinité de valeurs de x prouve, à cause de $\frac{l}{2} > 0$, que la limite de

$$\frac{\log^n x}{x}\left[f(x) - \varphi(x)\right],$$

pour $x = \infty$, ne peut pas être égale à zéro. Si donc cette limite est différente de zéro, la différence $f(x) - \varphi(x)$, d'après la convention établie plus haut, est une quantité de l'ordre $\frac{x}{\log^n x}$ ou d'un ordre inférieur; par conséquent $f(x)$ diffère de $\varphi(x)$ d'une quantité de l'ordre $\frac{x}{\log^n x}$, ou bien d'un ordre inférieur, ce qu'il s'agissait de démontrer.

En nous basant sur ce théorème, nous pouvons faire voir que la formule de Legendre $\frac{x}{\log x - 1{,}08366}$, pour laquelle la limite de l'expression

$$\frac{\log^2 x}{x}\left(\frac{x}{\log x - 1{,}08366} - \int_2^x \frac{dx}{\log x}\right),$$

quand $x = \infty$, est égale à 0,08366, ne peut exprimer $\varphi(x)$ avec un degré de précision allant jusqu'aux quantités de l'ordre $\frac{x}{\log^2 x}$ inclusivement.

On trouve avec la même facilité les valeurs des constantes A et B telles que la fonction $\frac{x}{A\log x+B}$ puisse représenter $\varphi(x)$ avec une précision poussée aux quantités de l'ordre $\frac{x}{\log^2 x}$ inclusivement. En vertu du théorême précédent de telles valeurs de A et B doivent satisfaire à l'équation

$$\lim.\left[\frac{\log^2 x}{x}\left(\frac{x}{A\log x+B}-\int_2^x\frac{dx}{\log x}\right)\right]_{x=\infty}=0.$$

Le dévelloppement de $\frac{x}{A\log x+B}$ donne

$$\frac{x}{A\log x+B}=\frac{1}{A}\cdot\frac{x}{\log x}-\frac{B}{A^2}\cdot\frac{x}{\log^2 x}+\frac{B^2}{A^3}\cdot\frac{x}{\log^3 x}-\ldots$$

De plus, intégrant $\int_2^x\frac{dx}{\log x}$ par parties, on trouve

$$\int_2^x\frac{dx}{\log x}=\frac{x}{\log x}+\frac{x}{\log^2 x}+2\int_2^x\frac{dx}{\log^3 x}+C.$$

En vertu de ce qui vient d'être trouvé l'équation précédente se réduit à

$$\lim.\left\{\frac{\log^2 x}{x}\left(\begin{array}{l}\frac{1}{A}\cdot\frac{x}{\log x}-\frac{B}{A^2}\cdot\frac{x}{\log^2 x}+\frac{B^2}{A^3}\cdot\frac{x}{\log^3 x}-\ldots\\ \ldots-\frac{x}{\log x}-\frac{x}{\log^2 x}-2\int_2^x\frac{dx}{\log^3 x}-C\end{array}\right)\right\}_{x=\infty}=0,$$

ou bien

$$\lim.\left\{\begin{array}{l}\left(\frac{1}{A}-1\right)\log x-\left(\frac{B}{A^2}+1\right)+\frac{B^2}{A^3}\,\frac{1}{\log x}-\ldots\\ \ldots-2\,\frac{\log^2 x}{x}\int_2^x\frac{dx}{\log^3 x}-C\,\frac{\log^2 x}{x}\end{array}\right\}_{x=\infty}=0.$$

Or, si l'on observe que tous les termes à partir du troisième convergent vers zéro pour des valeurs croissantes de x, on verra immédiatement qu'on ne peut satisfaire à l'équation précédente qu'en faisant $\frac{1}{A}-1=0$, $\frac{B}{A^2}+1=0$. D'où $A=1$, $B=-1$.

Ainsi, de toutes les fonctions de la forme $\frac{x}{A\log x+B}$ la seule $\frac{x}{\log x-1}$ peut exprimer $\varphi(x)$ avec une précision poussée aux quantités de l'ordre $\frac{x}{\log^2 x}$ inclusivement.

§ 6. Démontrons actuellement un théorème concernant le choix de la fonction qui détermine, avec un degré de précision requis, la fonction que nous avons représentée par $\varphi(x)$.

V-éme Théorème.

Si la fonction $\varphi(x)$ *qui désigne combien il y a de nombres premiers inférieurs à* x, *peut être représentée algébriquement avec une précision poussée aux quantités de l'ordre* $\frac{x}{\log^n x}$ *inclusivement au moyen des fonctions* x, $\log x$, e^x, *alors elle s'exprimera par la formule*

$$\frac{x}{\log x}+\frac{1.x}{\log^2 x}+\frac{1.2.x}{\log^3 x}+\ldots+\frac{1.2.3\ldots(n-1)x}{\log^n x}$$

Démonstration. Soit $f(x)$ la fonction qui, contenant sous forme algébrique x, $\log x$, e^x, exprime $\varphi(x)$ exactement jusqu'aux quantités de l'ordre $\frac{x}{\log^n x}$ inclusivement; l'expression

$$\frac{\log^n x}{x}\left[f(x)-\frac{x}{\log x}-\frac{1.x}{\log^2 x}-\frac{1.2.x}{\log^3 x}-\ldots-\frac{1.2.3\ldots(n-1)x}{\log^n x}\right]$$

pour des valeurs croissantes de x, devra converger soit vers zéro, soit vers une limite finie ou infiniment grande. En effet, s'il n'en était pas ainsi, la première dérivée de cette expression changerait de signe une infinité de fois pour des valeurs de x croissantes jusqu'à $+\infty$, ce qui ne peut arriver, comme il est facile de s'en assurer, avec une fonction algébrique de x, $\log x$, e^x *).

Ainsi, on aura nécessairement pour $f(x)$

$$(7)\quad \lim.\left\{\frac{\log^n x}{x}\left(f(x)-\frac{x}{\log x}-\frac{1.x}{\log^2 x}-\frac{1.2x}{\log^3 x}-\ldots-\frac{1.2.3\ldots(n-1)x}{\log^n x}\right)\right\}_{x=\infty}=L.$$

---

*) Il est très facile de s'assurer qu'une fonction algébrique de x, $\log x$, e^x cesse de changer de signe pour une valeur de x surpassant une certaine limite. Si la fonction dont il s'agit est entière, alors son signe dépendra d'un terme de la forme $Kx^{m'}.\log^{m''} x.e^{m'''x}$, pour des valeurs de x plus ou moins considérables, ce terme ne changeant pas de signe pour $x>1$. Pour toute autre fonction algébrique de x, $\log x$, e^x, que nous représenterons par y, on démontrera la même proposition de la manière suivante: observons d'abord que la fonction y sera la racine de l'équation $u_0y^m+u_1y^{m-1}+u_2y^{m-2}+\ldots+u_{m-1}y+u_m=0$, u_0, u_1, u_2,... u_{m-1}, u_m étant des fonctions entières de x, $\log x$, e^x; si l'on représente par v la fonction qui résulte de l'élimination de y entre l'équation précédente et sa dérivée, alors les fonctions u_0, u_m et v, comme entières, finiront par ne plus se réduire à zéro ou à changer de signe pour des valeurs de x surpassant une certaine limite; il arrivera donc que y conservera également son signe, car, pour des valeurs de x qui ne réduisent pas v à zéro, l'équation $u_0y^m+u_1y^{m-1}+\ldots+u_{m-1}y+u_m=0$ ne peut avoir de racines égales, et quand les racines sont inégales, le signe de l'une d'elles ne peut changer qu'en supposant que le signe de u_0 ou u_m change. — Cette propriété peut être étendue à beaucoup d'autres fonctions, pour lesquelles, par cette raison, le théorème V ainsi que les conséquences qui s'en déduisent, auront également lieu.

Mais, d'un autre côté, il est facile de s'assurer que

$$\lim. \left[\frac{\log^n x}{x}\left(\frac{x}{\log x}+\frac{1.x}{\log^2 x}+\frac{1.2.x}{\log^3 x}+\ldots+\frac{1.2\ldots(n-1)x}{\log^n x}-\int_2^x\frac{dx}{\log x}\right)\right]_{x=\infty}=0;$$

cette équation ajoutée à la précédente donne

$$\lim. \left[\frac{\log^n x}{x}\left(f(x)-\int_2^x\frac{dx}{\log x}\right)\right]_{x=\infty}=L.$$

Or, comme par hypothèse $f(x)$ représente $\varphi(x)$ exactement jusqu'aux quantités de l'ordre $\frac{x}{\log^n x}$ inclusivement, et que d'après le théorème IV cela ne peut avoir lieu, si la limite de

$$\frac{\log^n x}{x}\left[f(x)-\int_2^x\frac{dx}{\log x}\right],$$

pour $x=\infty$, n'est pas zéro, on aura $L=0$; cela posé, l'équation (7), pour $L=0$, se réduit à

$$\lim. \left\{\frac{\log^n x}{x}\left[f(x)-\frac{x}{\log x}-\frac{1.x}{\log^2 x}-\frac{1.2.x}{\log^3 x}-\ldots-\frac{1.2.3\ldots(n-1)x}{\log^n x}\right]\right\}_{x=\infty}=0,$$

ce qui prouve que la fonction

$$\frac{x}{\log x}+\frac{1.x}{\log^2 x}+\frac{1.2x}{\log^3 x}+\ldots+\frac{1.2.3\ldots(n-1)x}{\log^n x}$$

ne diffère pas de $f(x)$ de quantités de l'ordre $\frac{x}{\log^n x}$ et d'ordres inférieurs, et que par conséquent elle peut, aussi bien que $f(x)$, représenter $\varphi(x)$ avec une précision poussée jusqu'aux quantités de l'ordre $\frac{x}{\log^n x}$ inclusivement, ce qu'il s'agissait de démontrer.

D'après le théoréme que nous venons d'établir, nous concluons que si la fonction $\varphi(x)$, qui représente combien il y a de nombres premiers inférieurs à x, peut être exprimée algébriquement au moyen de x, $\log x$, e^x jusqu'aux quantités des ordres $\frac{x}{\log x}, \frac{x}{\log^2 x}, \frac{x}{\log^3 x}, \ldots$ inclusivement, elle devra s'exprimer par

$$\frac{x}{\log x},\ \frac{x}{\log x}+\frac{1.x}{\log^2 x},\ \frac{x}{\log x}+\frac{1.x}{\log^2 x}+\frac{1.2.x}{\log^3 x},\ldots$$

De plus, comme ces sommes ne sont autre chose que les valeurs successives de l'integrale $\int_2^x\frac{dx}{\log x}$, poussées aux quantités des ordres $\frac{x}{\log x}, \frac{x}{\log^2 x}, \frac{x}{\log^3 x}, \ldots$

nous sommes en droit de conclure que dans toutes ces hypothèses l'intégrale $\int_2^x \frac{dx}{\log x}$ exprimera $\varphi(x)$ avec exactitude jusqu'aux quantités de l'ordre pour lequel elle peut encore s'exprimer algébriquement au moyen de x, $\log x$, e^x.

Il est facile de se convaincre par les tables des nombres premiers que l'intégrale $\int_2^x \frac{dx}{\log x}$, pour x très grand, exprime avec assez de précision combien il y a de nombres premiers inférieurs à x. Mais ces tables sont trop peu étendues pour pouvoir décider de la supériorité de la formule $\int_2^x \frac{dx}{\log x}$ sur la formule de Legendre $\frac{x}{\log x - 1{,}08366}$ ou sur toute autre analogue. Dans les limites de ces tables les deux fonctions $\int_2^x \frac{dx}{\log x}$ et $\frac{x}{\log x - 1{,}08366}$ différent peu entr'elles; mais leur différence $\frac{x}{\log x - 1{,}08366} - \int_2^x \frac{dx}{\log x}$, ayant un *minimum* pour $x = e^{\frac{(1{,}08366)^2}{0{,}08366}} = 1247689$, croitra constamment jusqu'à l'infini après cette valeur, et déjà pour $x > 10000000$, aura une valeur considérable. C'est pour des nombres de cette grandeur que l'avantage de l'une des deux formules $\int_2^x \frac{dx}{\log x}$, $\frac{x}{\log x - 1{,}08366}$ devra se manifester. Mais pour effectuer cette vérification il faudrait avoir des tables de nombres premiers beaucoup plus étendues que celle que l'on possède.

§ 7. En adoptant l'intégrale $\int_2^x \frac{dx}{\log x}$ pour la valeur approchée de $\varphi(x)$ nous serons obligés de changer toutes les formules auxquelles Legendre est parvenu en partant de l'hypothsée $\varphi(x) = \frac{x}{\log x - 1{,}08366}$; nos formules ne seront pas plus compliquées que les siennes, et auront sur elles l'avantage d'être plus approchées d'après les théorèmes qui ont été demontrés plus haut.

Appliquons notre formule à la détermination de la somme des deux séries

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \ldots + \frac{1}{X},$$

$$\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\ldots\ldots\left(1 - \frac{1}{X}\right)$$

pour X très grand.

Pour déterminer la somme de la première de ces deux séries, nous observerons que, d'après la notation admise plus haut, l'on a

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\sum_{x=2}^{x=X}\frac{\varphi(x+1)-\varphi(x)}{x},$$

car, $\varphi(x)$ représentant la totalité des nombres premiers inférieurs à x, la différence $\varphi(x+1)-\varphi(x)$ se réduira à zéro toutes les fois que x sera un nombre composé, et à 1, quand x sera premier.

Supposons X très grand, et désignons par λ un nombre inférieur à X, mais assez grand cependant pour que la fonction $\varphi(x)$, entre les limites $x=\lambda$ et $x=X$, puisse être représentée avec une exactitude suffisante par l'intégrale $\int_2^x \frac{dx}{\log x}$. Dans cette hypothèse l'équation précédente pourra s'écrire de cette manière:

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\sum_{x=2}^{x=\lambda-1}\frac{\varphi(x+1)-\varphi(x)}{x}+\sum_{x=\lambda}^{x=X}\frac{\varphi(x+1)-\varphi(x)}{x}.$$

Remplaçant dans la dernière somme $\varphi(x)$ par $\int_2^x \frac{dx}{\log x}$, on aura

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\sum_{x=2}^{x=\lambda-1}\frac{\varphi(x+1)-\varphi(x)}{x}+\sum_{x=\lambda}^{x=X}\frac{\int_x^{x+1}\frac{dx}{\log x}}{x}.$$

Or, l'intégrale $\int_x^{x+1}\frac{dx}{\log x}$ peut être représentée par $\frac{1}{\log x}$ avec exactitude jusqu'aux quantités de l'ordre $\frac{1}{x}$; de plus, la somme $\sum_{x=\lambda}^{x=X}\frac{1}{x\log x}$ peut être remplacée par l'intégrale $\int_\lambda^X \frac{dx}{x\log x}$ avec le même degré de précision.

Sous ces conditions l'équation précédente deviendra

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\sum_{x=2}^{x=\lambda-1}\frac{\varphi(x+1)-\varphi(x)}{x}+\int_\lambda^x\frac{dx}{x\log x},$$

ou bien, effectuant l'intégration,

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\sum_{x=2}^{x=\lambda-1}\frac{\varphi(x+1)-\varphi(x)}{x}-\log\log\lambda+\log\log X$$

Enfin, si l'on remplace par C la quantité

$$\sum_{x=2}^{x=\lambda-1} \frac{\varphi(x+1)-\varphi(x)}{x} - \log\log\lambda,$$

indépendante de x, on trouvera

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=C+\log\log X. \tag{8}$$

Lorsque l'on aura déterminé la valeur de la constante C, cette équation pourra servir à trouver par approximation la somme de la série

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}$$

quand X sera très grand.

La formule que nous venons de trouver est plus simple que celle de Legendre, d'aprés laquelle on a

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}=\log(\log X-0{,}08366)+C.$$

Passons maintenant à la détermination du produit

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\ldots\left(1-\frac{1}{X}\right)=P.$$

Prenant le logarithme des deux membres de cette équation, on aura la formule

$$\log P=\log\left(1-\frac{1}{2}\right)+\log\left(1-\frac{1}{3}\right)+\log\left(1-\frac{1}{5}\right)+\ldots+\log\left(1-\frac{1}{X}\right)$$

qui peut encore s'écrire de la manière suivante:

$$\log P=-\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}\right)+\frac{1}{2}+\log\left(1-\frac{1}{2}\right)+$$
$$\frac{1}{3}+\log\left(1-\frac{1}{3}\right)+\frac{1}{5}+\log\left(1-\frac{1}{5}\right)+\ldots+\frac{1}{X}+\log\left(1-\frac{1}{X}\right).$$

Observons actuellement que la série finie

$$\frac{1}{2}+\log\left(1-\frac{1}{2}\right)+\frac{1}{3}+\log\left(1-\frac{1}{3}\right)+\frac{1}{5}+\log\left(1-\frac{1}{5}\right)+\ldots+\frac{1}{X}+\log\left(1-\frac{1}{X}\right),$$

aux quantités de l'ordre $\frac{1}{X}$ près, peut être remplacée par la série infinie

$$\frac{1}{2}+\log\left(1-\frac{1}{2}\right)+\frac{1}{3}+\log\left(1-\frac{1}{3}\right)+\frac{1}{5}+\log\left(1-\frac{1}{5}\right)+\ldots$$

Or, la différence entre ces deux séries est évidemment inférieure à la somme

$$-\frac{1}{X+1}-\log\left(1-\frac{1}{X+1}\right)-\frac{1}{X+2}-\log\left(1-\frac{1}{X+2}\right)-\ldots$$

qui, elle même, est inférieure à l'intégrale

$$\int_X^\infty \left[-\frac{1}{x}-\log\left(1-\frac{1}{x}\right)\right]dx = 1+(X-1)\log\left(1-\frac{1}{X}\right);$$

de plus, comme la valeur de $1+(X-1)\log\left(1-\frac{1}{X}\right)$ pour X très grand, est une quantité infiniment petite du premier ordre par rapport à $\frac{1}{X}$, nous en concluons que la substitution qui vient d'être indiquée est permise.

D'après ce qui vient d'être dit, si l'on représente par C' la somme de la série infinie

$$\frac{1}{2}+\log\left(1-\frac{1}{2}\right)+\frac{1}{3}+\log\left(1-\frac{1}{3}\right)+\frac{1}{5}+\log\left(1-\frac{1}{5}\right)+\ldots$$

la valeur de $\log P$ s'exprimera, aux quantités de l'ordre de $\frac{1}{X}$ près, par la formule

$$\log P = -\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}\right)+C'.$$

Substituant pour

$$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\ldots+\frac{1}{X}$$

la valeur (8) trouvée plus haut, on obtiendra

$$\log P = -C-\log\log X + C',$$

d'où

$$P=\frac{e^{C'-C}}{\log X}.$$

Enfin, faisant pour abréger $e^{C'-C}=C_0$, et remplaçant P par le produit

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\ldots\left(1-\frac{1}{X}\right),$$

nous aurons

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\ldots\left(1-\frac{1}{X}\right)=\frac{C_0}{\log X}.$$

Legendre a trouvé, pour la valeur du même produit, la formule suivante:

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\ldots\left(1-\frac{1}{X}\right)=\frac{C_0}{\log X-0,08366}.$$

PAIR CORRELATION OF ZEROS AND PRIMES IN SHORT INTERVALS

Daniel A. Goldston and Hugh L. Montgomery*

1. Statement of results.

In 1943, A. Selberg [15] deduced from the Riemann Hypothesis (RH) that

$$\int_1^X \left(\psi((1+\delta)x) - \psi(x) - \delta x\right)^2 x^{-2}\, dx \ll \delta(\log X)^2 \tag{1}$$

for $X^{-1} \leqslant \delta \leqslant X^{-1/4}$, $X \geqslant 2$. Selberg was concerned with small values of δ, and the constraint $\delta \leqslant X^{-1/4}$ was imposed more for convenience than out of necessity. For larger δ we have the following result.

Theorem 1. *Assume* RH. *Then*

$$\int_1^X \left(\psi((1+\delta)x) - \psi(x) - \delta x\right)^2 x^{-2}\, dx \ll \delta(\log X)(\log 2/\delta) \tag{2}$$

for $0 < \delta \leqslant 1$, $X \geqslant 2$.

In this estimate, the error term for the number of primes in the interval $(x, (1+\delta)x]$ is damped by the factor x^{-2}, and the length of the interval, δx, varies with x. Saffari and Vaughan [14] considered the undamped integral, and derived from RH the estimates

$$\int_1^X \left(\psi((1+\delta)x) - \psi(x) - \delta x\right)^2 dx \ll \delta X^2 (\log 2/\delta)^2 \tag{3}$$

for $0 < \delta \leqslant 1$, and

$$\int_1^X \left(\psi(x+h) - \psi(x) - h\right)^2 dx \ll hX(\log 2X/h)^2 \tag{4}$$

*Research supported in part by NSF Grant MCS82-01602.

for $0 < h \leqslant X$. It may be similarly shown that RH gives the estimate

$$\int_1^X (\psi(x) - x)^2 \, dx \ll X^2 . \tag{5}$$

Gallagher and Mueller [5] showed that if one assumes not only RH but also the pair correlation conjecture

$$\begin{aligned} &\# \{(\gamma,\gamma') : 0 < \gamma \leqslant T,\ 0 < \gamma - \gamma' \leqslant 2\pi a/\log T\} \\ &= \Big(\frac{1}{2\pi} \int_0^a 1 - \Big(\frac{\sin \pi u}{\pi u}\Big)^2 du + o(1)\Big) T \log T \end{aligned} \tag{6}$$

then it can be deduced that

$$\int_1^X (\psi((1+\delta)x) - \psi(x) - \delta x)^2 x^{-2} \, dx \sim \delta(\log 1/\delta)(\log X\sqrt{\delta}) \tag{7}$$

for $X^{-1} \leqslant \delta \leqslant X^{-\varepsilon}$. Here γ denotes the ordinate of a non-trivial zero of the Riemann zeta function. Thus it seems likely that the estimate of Theorem 1 is best possible.

In the course of formulating the conjecture (6), Montgomery [13] also proposed a more precise estimate, namely that

$$F(X,T) \sim \frac{1}{2\pi} T \log T \tag{8}$$

uniformly for $T \leqslant X \leqslant T^A$, for any fixed $A > 1$, where

$$F(X,T) = \sum_{0 < \gamma,\gamma' \leqslant T} X^{i(\gamma-\gamma')} w(\gamma-\gamma') \tag{9}$$

and $w(u) = 4/(4 + u^2)$. We now relate this conjecture to the size of the integral in (3).

Theorem 2. *Assume* RH. *If* $0 < B_1 \leqslant B_2 \leqslant 1$, *then*

$$\int_1^X (\psi((1+\delta)x) - \psi(x) - \delta x)^2 \, dx \sim \frac{1}{2} \delta X^2 \log 1/\delta \tag{10}$$

uniformly for $X^{-B_2} \leqslant \delta \leqslant X^{-B_1}$, *provided that* (8) *holds uniformly* *for*

$$X^{B_1} (\log X)^{-3} \leq T \leq X^{B_2} (\log X)^3. \tag{11}$$

Conversely, if $1 \leq A_1 \leq A_2 < \infty$, *then* (8) *holds uniformly for* $T^{A_1} \leq X \leq T^{A_2}$, *provided that* (10) *holds uniformly for*

$$X^{-1/A_1} (\log X)^{-3} \leq \delta \leq X^{-1/A_2} (\log X)^3 . \tag{12}$$

Previously Mueller [12] derived (10) from RH and a strong quantitative form of (8). Heath-Brown and Goldston [11] showed that RH and (8) for $T^a \leq X \leq T^b$, $a < 2 < b$, imply

$$p_{n+1} - p_n = o(p_n^{1/2} (\log p_n)^{1/2}) .$$

This estimate follows easily from Theorem 2 by taking $\delta = \varepsilon X^{-1/2} (\log X)^{1/2}$ in (10). In deriving (10) from (8) we also use the weaker estimate (3). In the case of very small δ, say $\delta \approx (\log X)/X$, we can do better by appealing instead to the bound

$$\int_1^X (\psi((1 + \delta)x) - \psi(x) - \delta x)^2 \, dx \ll \delta X^2 \log X + \delta^2 X^3 \tag{13}$$

which follows from sieve estimates (see the proof of Lemma 7). In this way we could show that

$$\int_1^X (\pi(x + h) - \pi(x) - h/\log x)^2 \, dx \sim hX/\log X \tag{14}$$

for $h \approx \log X$, given RH and (8) for $T \leq X \leq f(T)T \log T$. Here $f(T)$ tends to infinity arbitrarily slowly with T. From this it follows easily that

$$\liminf \, (p_{n+1} - p_n) / \log p_n = 0 .$$

Heath-Brown [10] derived this from a slightly stronger hypothesis.

In assessing the depth of the estimates (8) and (10), we note that (10) is a logarithm sharper than (3), and that (8) is a logarithm sharper than the trivial bound

$$|F(X,T)| \leq F(1,T) \sim \frac{1}{2\pi} T(\log T)^2 \quad . \tag{15}$$

(See Lemma 8.) As in (4), we can relate (10) to primes in intervals of constant length. In summary we have the following

Corollary. *Assume* RH. *Then the following assertions are equivalent*:

(a) *For every fixed* $A \geq 1$, (8) *holds uniformly for* $T \leq X \leq T^A$.

(b) *For every fixed* $\varepsilon > 0$, (10) *holds uniformly for* $X^{-1} \leq \delta \leq X^{-\varepsilon}$.

(c) *For every fixed* $\varepsilon > 0$,

$$\int_0^X \big(\psi(x + h) - \psi(x) - h\big)^2 \, dx \sim hX \log X/h \tag{16}$$

holds uniformly for $1 \leq h \leq X^{1-\varepsilon}$.

It is not hard to show that either (b) or (c) implies RH. Gallagher [4] has shown that a weak quantitative form of the prime k-tuple hypothesis gives (16) when $h \approx \log X$.

The path we take between (8) and (10) involves elementary arguments of Abelian and Tauberian character; these are of two sorts. First, we consider the connection between the assertion

$$\int_{-\infty}^{+\infty} e^{-2|y|} f(Y + y) \, dy = 1 + o(1) \tag{17}$$

as $Y \to +\infty$, and the more general assertion

$$\int_a^b R(y) f(Y + y) \, dy = \int_a^b R(y) \, dy + o(1) \tag{18}$$

as $Y \to +\infty$ where R is any Riemann-integrable function. (These two statements are equivalent if f is bounded and non-negative.) This interplay reflects the choice of the weighting function w(u) in the definition (9) of F(X,T). Second, and more intrinsically, we consider a question of Riemann summability (R_2), namely the connection between the two assertions

$$\int_0^{\infty} \left(\frac{\sin \kappa u}{u}\right)^2 f(u)du = (\pi/2 + o(1))\kappa \log 1/\kappa \tag{19}$$

as $\kappa \to 0^+$, and

$$\int_0^{U} f(u)du = (1 + o(1)) U \log U \tag{20}$$

as $U \to +\infty$. Because of the intricacies of the (R_2) method, neither of these assertions implies the other, although they are equivalent for non-negative functions f. The lemmas we formulate below are complicated by the fact that we specify the relation between the parameters κ and U.

2. Lemmas of summability.

Lemma 1. *If*

$$I(Y) = \int_{-\infty}^{+\infty} e^{-2|y|} f(Y + y)dy = 1 + \varepsilon(Y) ,$$

and if $f(y) \geq 0$ *for all* y, *then for any Riemann-integrable function* $R(y)$,

$$\int_a^b R(y) f(Y + y)dy = \left(\int_a^b R(y)dy\right) \left(1 + \varepsilon'(y)\right) . \tag{21}$$

If R *is fixed then* $|\varepsilon'(Y)|$ *is small provided that* $|\varepsilon(y)|$ *is small uniformly for* $Y + a - 1 \leq y \leq Y + b + 1$.

In terms of Wiener's general Tauberian theorem, the truth of this lemma hinges on the fact that the Fourier transform of the kernel $k(y) = e^{-2|y|}$, namely the function

$$\hat{k}(t) = \int_{-\infty}^{+\infty} k(y) e(-ty)dy = \frac{1}{\pi^2 t^2 + 1} , \quad \left(e(u) = e^{2\pi i u}\right),$$

never vanishes.

Proof. Let $K_c(y) = \max(0, c - |y|)$. By comparing Fourier transforms, or by direct calculation, we see that

$$K_c(y) = \frac{1}{2} e^{-2|y|} - \frac{1}{4} e^{-2|y-c|} - \frac{1}{4} e^{-2|y+c|}$$

$$+ \int_{-c}^{c} (c - |z|) e^{-2|y-z|} dz .$$

Hence

$$\int_{-c}^{c} K_c(y) f(Y + y) dy = \frac{1}{2} I(Y) - \frac{1}{4} I(Y + c) - \frac{1}{4} I(Y - c)$$

$$+ \int_{-c}^{c} (c - |z|)I(Y + z) dz$$

$$= c^2 + \varepsilon_1(Y)$$

where $|\varepsilon_1|$ is small if $c > 0$ is fixed and if $|\varepsilon(y)|$ is small for $Y - c \leqslant y \leqslant Y + c$. Since

$$\frac{1}{\eta}\left(K_c(y) - K_{c-\eta}(y)\right) \leqslant \chi_{[-c,c]}(y) \leqslant \frac{1}{\eta}\left(K_{c+\eta}(y) - K_c(y)\right) ,$$

and since $f \geqslant 0$, we deduce that (21) holds in the case of the step function $R(y) = \chi_{[-c,c]}(y)$. Since the general R can be approximated above and below by step functions, we obtain (21).

Lemma 2. *Suppose that* $f(t)$ *is a continuous non-negative function defined for all* $t \geqslant 0$, *with* $f(t) \ll \log^2 (t + 2)$. *If*

$$J(T) = \int_0^T f(t)dt = \left(1 + \varepsilon(T)\right)T \log T ,$$

then

$$\int_0^\infty \left(\frac{\sin \kappa u}{u}\right)^2 f(u)du = \left(\pi/2 + \varepsilon'(\kappa)\right)\kappa \log 1/\kappa \qquad (22)$$

where $|\varepsilon'(\kappa)|$ *is small as* $\kappa \to 0^+$ *if* $|\varepsilon(T)|$ *is small uniformly for* $\kappa^{-1} (\log \kappa)^{-2} \leqslant T \leqslant \kappa^{-1}(\log \kappa)^2$.

Proof. We divide the range of integration in (22) into four subintervals: $0 \leqslant u \leqslant \kappa^{-1}(\log \kappa)^{-2} = U_1$, $U_1 \leqslant u \leqslant C\kappa^{-1} = U_2$, $U_2 \leqslant u \leqslant \kappa^{-1}(\log \kappa)^2 = U_3$, and $U_3 \leqslant u < \infty$.
Since $f(t) \ll \log^2(t + 2)$, we see that

$$\int_0^{U_1} < \int_0^{U_1} \kappa^2 \log^2(u+2)\,du < \kappa^2 U_1 \log^2 U_1 \ll \kappa \quad ,$$

and similarly that

$$\int_{U_3}^{\infty} < \int_{U_3}^{\infty} u^{-2} \log^2 u\,du \ll U_3^{-1} \log^2 U_3 < \kappa \quad .$$

By writing $f(u) = \log 1/\kappa + \log \kappa u + (f(u) - \log u)$, we express the integral from U_1 to U_2 as a sum of three integrals. We note that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 du = \int_0^{\infty} \left(\frac{\sin \kappa u}{u}\right)^2 du + O(\kappa(\log \kappa)^{-2})$$
$$= \frac{\pi}{2}\kappa(1 + O(\log \kappa)^{-2}),$$

and that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 \log \kappa u\,du < \int_0^{\infty} \min(\kappa^2 , u^{-2}) \log \kappa u\,du < \kappa \quad .$$

Put $r(u) = J(u) - u \log u + u$. Then by integrating by parts we see that

$$\int_{U_1}^{U_2} \left(\frac{\sin \kappa u}{u}\right)^2 (f(u)-\log u)du < \kappa\left(1+ \left(\log \frac{1}{\kappa}\right) \max_{U_1 \leqslant u \leqslant U_2} |\varepsilon(u)|\right)\log(C+2) \quad .$$

As for the range $U_2 \leqslant u \leqslant U_3$, we see that if $\varepsilon(u) \leqslant 1$ then

$$\int_{U_2}^{U_3} \leqslant \int_{U_2}^{U_3} f(u)u^{-2}\,du \ll U_2^{-1} \log U_2 \ll C^{-1}\kappa \log 1/\kappa \ .$$

We make this small by taking C large. Then the remaining error terms are small if $\varepsilon(u)$ is small.

Lemma 3. *If* K *is even,* K" *continuous,* $\int_{-\infty}^{+\infty} |K| < \infty$,
$K(x) \to 0$ *as* $x \to +\infty$, $K' \to 0$ *as* $x \to +\infty$, *and if* $K''(x) \ll x^{-3}$ *as* $x \to +\infty$, *then*

$$\hat{K}(t) = \int_0^{\infty} K''(x)\left(\frac{\sin \pi t x}{\pi t}\right)^2 dx \ . \qquad (23)$$

Proof. Integrate by parts twice.

Lemma 4. *If f is a non-negative function defined on* $[0, +\infty)$, $f(t) \ll \log^2(t + 2)$, *and if*

$$I(\kappa) = \int_0^\infty \left(\frac{\sin \kappa t}{t}\right)^2 f(t)dt = (\pi/2 + \varepsilon(\kappa))\kappa \log 1/\kappa$$

then

$$J(T) = \int_0^T f(t)dt = (1 + \varepsilon')T \log T$$

where $|\varepsilon'|$ *is small if* $|\varepsilon(\kappa)| \leqslant \varepsilon$ *uniformly for*

$$T^{-1}(\log T)^{-1} \leqslant \kappa \leqslant T^{-1}(\log T)^2 .$$

Proof. Let K be a kernel with the properties specified in Lemma 3. Replace t by t/T in (23), multiply by $f(t) - \log t$, and integrate over $0 \leqslant t < \infty$. Then we find that

$$\int_0^\infty \left(f(t) - \log t\right) \hat{K}(t/T)dt = \pi^{-2} T^2 \int_0^\infty K''(x) R(\pi x/T)dx$$

where

$$R(\kappa) = I(\kappa) - \int_0^\infty \left(\frac{\sin \kappa t}{t}\right)^2 \log t \, dt$$

$$= I(\kappa) - \frac{1}{2} \pi\kappa \log 1/\kappa + O(\kappa).$$

Since

$$I(\kappa) \ll \int_0^\infty \min(\kappa^2, t^{-2})\log^2(t + 2)dt \ll \kappa \log^2(2 + 1/\kappa)$$

for all $\kappa > 0$, on taking $x_1 = (\log T)^{-1}$ we see that

$$\int_0^{x_1} K''R \ll \int_0^{x_1} xT^{-1} \log^2 T/x \, dx \ll T^{-1}.$$

On taking $x_2 = \frac{1}{4}(\log T)^2$ we find that

$$\int_{x_2}^\infty K''R \ll \int_{x_2}^\infty x^{-3}(x/T) \log^2 T \, dx \ll T^{-1}.$$

Assuming, as we may, that $\varepsilon \geq (\log T)^{-1}$, we have $R(\pi x/T) \ll \varepsilon\, xT^{-1}\log T$ for $x_1 \leq x \leq x_2$. Hence

$$\int_{x_1}^{x_2} K''R \ll \varepsilon\, T^{-1}(\log T) \int_0^\infty \min(1, x^{-3})\, x\, dx \ll \varepsilon\, T^{-1}\log T.$$

For $\eta > 0$ take

$$K(x) = K_\eta(x) = \left(\sin 2\pi x + \sin 2\pi(1+\eta)x\right)\left(2\pi x(1 - 4\eta^2 x^2)\right)^{-1} ,$$

so that

$$\hat{K}(t) = \begin{cases} 1 & \text{if } |t| \leq 1, \\ \cos^2\left(\pi(|t| - 1)/(2\eta)\right) & \text{if } 1 \leq |t| \leq 1+\eta, \\ 0 & \text{if } |t| \geq 1+\eta . \end{cases}$$

Thus

$$\int_0^\infty f(t)\hat{K}_\eta(t/T)dt = (1 + O(\eta))T \log T + O_\eta(T) + O_\eta(\varepsilon\, T \log T) .$$

Since f is non-negative, we see that

$$\int_0^\infty f(t)\hat{K}_\eta\left((1+\eta)t/T\right)dt \leq J(T) \leq \int_0^\infty f(t)\hat{K}_\eta(t/T)dt ,$$

and we obtain the desired result by taking η small.

In this argument we have made free use of existing treatments of Riemann summability. We note especially Hardy [8, pp. 301, 316, 365] and Hardy and Rogosinski [9, Theorem III].

3. Lemmas of analytic number theory.

As is customary, we write $s = \sigma + it$, and we let $\rho = \beta + i\gamma$ be a typical non-trivial zero of the Riemann zeta function. We first note a simple result of Gallagher [3]:

Lemma 5. *Let* $S(t) = \sum_{\mu \in M} c(\mu)e(\mu t)$ *where* M *is a countable set of real numbers and* $\sum |c(\mu)| < \infty$. *Then*

$$\int_{-T}^{T} |S(t)|^2 \, dt < T^2 \int_{-\infty}^{+\infty} \Big| \sum_{\substack{\mu \in M \\ |\mu - u| \leqslant (4T)^{-1}}} c(\mu) \Big|^2 \, du.$$

When a main term is desired, we use the following more elaborate estimate.

Lemma 6. *Let* S(t) *be as above. If* $\delta \geqslant T^{-1}$ *then*

$$\int_0^T |S(t)|^2 \, dt = (T + O(\delta^{-1})) \sum_{\mu \in M} |c(\mu)|^2$$

$$+ O\Big(T \sum_{\substack{\mu,\nu \in M \\ 0 < |\mu - \nu| < \delta}} |c(\mu)c(\nu)| \Big) .$$

Proof. Selberg (see Vaaler[17]) has constructed functions $F_-(t)$ and $F_+(t)$ such that $F_-(T) \leqslant \chi_{[0,T]}(t) \leqslant F_+(t)$, $\hat{F}_\pm(x) = 0$ for $|x| \geqslant \delta$, and $\int_{-\infty}^{+\infty} F_\pm(t)dt = T \pm \delta^{-1}$. Hence

$$\int_0^T |S|^2 \leqslant \int_{-\infty}^{+\infty} |S|^2 F_+ = \sum_{\mu,\nu} c(\mu)\overline{c(\nu)} \, \hat{F}_+(\nu - \mu) .$$

The terms $\mu = \nu$ contribute $(T + \delta^{-1}) \sum_\mu |c(\mu)|^2$. Since

$$|\hat{F}_+| \leqslant \int |F_+| = T + \delta^{-1} \leqslant 2T ,$$

the terms $\mu \neq \nu$ contribute at most

$$2T \sum_{0 < |\mu - \nu| < \delta} |c(\mu)c(\nu)|.$$

This gives an upper bound, and a corresponding lower bound is derived similarly using F_-.

Lemma 7. *Let* $C(x) > 0$ *be a continuous function such that* $C(x) \approx C(y)$ *whenever* $x \approx y$. If $|c(p)| \leqslant C(p)$ *for all primes* p, *and if* $\delta \geqslant T^{-1}$, *then*

$$\int_0^T \Big| \sum_p c(p)p^{it}\Big|^2\, dt = (T + O(\delta^{-1})) \sum_p |c(p)|^2 + O\Big(\delta T \int_{\delta^{-1}}^{\infty} C(u)^2\, u(\log u)^{-2}\, du\Big)$$

Proof. We appeal to the previous lemma. In the second error term, the primes $p \in (X,2X]$ contribute

$$T\, C(X)^2 \sum_{X<p\leqslant 2X}\ \sum_{p<p'\leqslant(1+2\delta)p} 1 \ll T\, C(X)^2 \sum_{1\leqslant k\leqslant 4\delta X} \pi_2(2X,k)$$

where $\pi_2(x,k)$ denotes the number of primes $p \leqslant x$ for which $p + k$ is also prime. It is well-known (see Halberstam and Richert [7, p.117]) that

$$\pi_2(x,k) \ll \big(k/\phi(k)\big)x(\log x)^{-2}$$

uniformly for $x \geqslant 2$, $k \neq 0$. Since $\sum_{k \leqslant K} k/\phi(k) \ll K$, it follows that our upper bound is

$$\ll T\, C(X)^2\, \delta X^2\, (\log X)^{-2} \ll \delta T \int_X^{2X} C(u)^2\, u(\log u)^{-2}\, du\ .$$

We put $X = \delta^{-1}\, 2^r$ and sum over $r \geqslant 0$ to obtain the desired result.

We now present the main known properties of $F(X,T)$.

Lemma 8. *Assume* RH, *and let* $F(X,T)$ *be as in* (9). *Then* $F(X,T) \geqslant 0$, $F(X,T) = F(1/X,T)$, *and*

$$F(X,T) = T\big(X^{-2}(\log T)^2 + \log X\big)\Big(\frac{1}{2\pi} + O\big((\log T)^{-1/2}\, (\log\log T)^{1/2}\big)\Big) \tag{24}$$

uniformly for $1 \leqslant X \leqslant T$.

Proof. The first assertion is an immediate consequence of either of the two identities

$$F(X,T) = 2\pi \int_{-\infty}^{+\infty} e^{-4\pi|u|} \left| \sum_{0<\gamma\leq T} X^{i\gamma} e(\gamma u) \right|^2 du, \qquad (25)$$

or

$$F(X,T) = \frac{2}{\pi} \int_{-\infty}^{+\infty} \left| \sum_{0<\gamma\leq T} \frac{X^{i\gamma}}{1+(t-\gamma)^2} \right|^2 dt.$$

The observation that F is non-negative has also been made by Mueller (unpublished). The second assertion is obvious from the definition of F. The estimate (24) is substantially due to Goldston [6, Lemma B], and may be proved by substituting an appeal to Lemma 7 in the argument of Montgomery [13].

Lemma 9. *If* $0 \leq h \leq T$ *then*

$$\#\{(\gamma,\gamma') : 0 \leq \gamma \leq T,\ |\gamma - \gamma'| \leq h\} \ll (1 + h \log T) T \log T. \qquad (27)$$

Proof. We argue unconditionally, although if RH is assumed then the above follows easily from Lemma 8 (see (6) of Montgomery [13]). Let $N(T) = \#\{\gamma : 0 < \gamma \leq T\}$. Following Selberg, Fujii [2] showed that

$$\int_0^T \left(N(t+h) - N(t) - \frac{1}{2\pi} h \log t \right)^2 dt \ll T \log(2 + h \log T)$$

for $0 \leq h \leq 1$. Hence

$$\int_0^T \left(N(t+h) - N(t) \right)^2 dt \ll h^2 T (\log T)^2$$

for $(\log T)^{-1} \leq h \leq 1$. This gives (27) in this case. To derive (27) when $0 \leq h \leq (\log T)^{-1}$, it suffices to consider $h = (\log T)^{-1}$. As for the range $1 \leq h \leq T$, it suffices to use the bound

$$N(T+1) - N(T) \ll \log T \qquad (28)$$

(see Titchmarch [16,p. 178]).

Lemma 10. *For* $0 < \delta \leqslant 1$ *let*

$$a(s) = ((1+\delta)^s - 1)/s \; . \tag{29}$$

If $|c(\gamma)| \leqslant 1$ *for all* γ *then*

$$\int_{-\infty}^{+\infty} |a(it)|^2 \; \Big|\sum_{\gamma} \frac{c(\gamma)}{1+(t-\gamma)^2}\Big|^2 \, dt = \int_{-\infty}^{+\infty} \Big|\sum_{|\gamma|\leqslant Z} \frac{a(1/2+i\gamma)c(\gamma)}{1+(t-\gamma)^2}\Big|^2 \, dt$$

$$+ \, O(\delta^2(\log 2/\delta)^3) + O(Z^{-1}(\log Z)^3) \tag{30}$$

provide that $Z \geqslant 1/\delta$.

Proof. By (28), the sum that occurs in the integral on the left is $\ll \log(2+|t|)$. Since

$$a(s) \ll \min(\delta, \; |s|^{-1}) \tag{31}$$

in the strip $|\sigma| \leqslant 1/\delta$, it follows by Cauchy's formula or by direct calculation that

$$a'(s) \ll \min(\delta^2 \; , \; \delta/|s|) \tag{32}$$

for $|\sigma| \leqslant (2\delta)^{-1}$. Hence in particular,

$$a(it) - a(\tfrac{1}{2} + it) \ll \min(\delta^2 \; , \; \delta/|t|) \; ,$$

and consequently

$$|a(it)|^2 - |a(\tfrac{1}{2} + it)|^2 \ll \min(\delta^3, \; \delta/t^2) \; .$$

Let I denote the integral on the left in (30), and J the corresponding integral with a(it) replaced by a($\tfrac{1}{2}$ + it). Then

$$I - J \ll \int \min(\delta^3, \; \delta/t^2)(\log(2+|t|))^2 \, dt \ll \delta^2 (\log 2/\delta)^2 \; .$$

Write J in the form $J = \int |A|^2$. From (28) and (31) we see that

$$A \ll \min(\delta,\ |t|^{-1}) \log(2 + |t|) \tag{33}$$

Now let K be the integral with $a(1/2 + it)$ replaced by $a(1/2 + i\gamma)$, and write $K = \int |B|^2$. Then B also satisfies the estimate (33). From (31) and (32) we see that

$$a(1/2 + i\gamma) - a(1/2 + it) \ll |t - \gamma| \min(\delta^2,\ \delta/|t|) .$$

Thus

$$A - B \ll \min(\delta^2,\ \delta/|t|)(\log(2/\delta + |t|))^2,$$

so that

$$|A|^2 - |B|^2 \ll \min(\delta^3,\ \delta/t^2)(\log(2/\delta + |t|))^3 ,$$

and hence

$$J - K \ll \delta^2(\log 2/\delta)^3 .$$

Finally, let $L = \int |C|^2$ be the integral on the right in (30). We note that C also satisfies the estimate (33). Since

$$B - C \ll \min(Z^{-1},\ |t|^{-1}) \log(2Z + |t|),$$

we find that

$$|B|^2 - |C|^2 \ll \min(Z^{-1}(1 + |t|)^{-1},\ t^{-2})(\log(2Z + |t|))^2 .$$

Thus

$$K - L \ll Z^{-1} (\log 2Z)^3 ,$$

and the proof is complete.

4. Proof of Theorem 1.

Although we arrange the technical details differently, the ideas are entirely the same as in Selberg's paper. If $\delta X \le 1$ then there is at most one prime power in the interval $(x, (1 + \delta)x]$, so that our integral is

$$\ll \delta \sum_{n\leqslant X} \Lambda(n)^2/n + \delta^2 X \ll \delta(\log X)^2 ,$$

which suffices. We now suppose that $\delta X > 1$. By the above argument we see that

$$\int_0^{1/\delta} \dots \ll \delta(\log 2/\delta)^2.$$

Thus it suffices to consider the range $1/\delta \leqslant x \leqslant X$. Here we apply the explicit formula for $\psi(x)$ (see Davenport [1, 17]), which gives

$$\begin{aligned}\psi\big((1+\delta)x\big) - \psi(x) - \delta x = &- \sum_{|\rho|\leqslant Z} a(\rho)x^{\rho} \\ &+ O\Big((\log x)\min(1, \frac{x}{Z\,\|x\|})\Big) \\ &+ O\Big((\log x)\min(1, \frac{x}{Z\,\|(1+\delta)x\|})\Big) \\ &+ O(x\,Z^{-1}(\log xZ)^2)\end{aligned} \tag{34}$$

where $a(s)$ is given in (29), and $\|\theta\| = \min_n \|\theta - n\|$ is the distance from θ to the nearest integer. The error terms contribute a negligible amount if we take $Z = X(\log X)^2$. Writing $\rho = \frac{1}{2} + i\gamma$, $x = e^y$, $Y = \log X$, we see that it remains to show that

$$\int_{\log 1/\delta}^{Y} \Big| \sum_{|\gamma|\leqslant Z} a(\rho)\, e^{i\gamma y}\Big|^2 dy \ll \delta Y \log 2/\delta. \tag{35}$$

By Lemma 5 we see that this integral is

$$\ll Y^2 \int_{-\infty}^{\infty} \Big(\sum_{\substack{|\gamma|\leqslant Z \\ |\gamma - 2\pi u|\leqslant 2/Y}} a(\rho)^2\Big)\, du \ll Y \sum_{\substack{|\gamma|\leqslant Z \\ |\gamma'|\leqslant Z \\ |\gamma-\gamma'|\leqslant 4/Y}} |a(\rho)a(\rho')| .$$

By (31) and Lemma 9 this gives (35), and the proof is complete.

5. Proof of Theorem 2.

We first assume (8) as needed, and derive (10). Let

$$J(T) = J(X,T) = 4 \int_0^T \Big| \sum_\gamma \frac{X^{i\gamma}}{1 + (t-\gamma)^2} \Big|^2 dt \quad .$$

Montgomery [13] (see his (26), but beware of the changes in notation) used (28) to show that

$$J(X,T) = 2\pi \, F(X,T) + O\big((\log T)^3\big) \ .$$

Thus (8) is equivalent to

$$J(X,T) = \big(1 + o(1)\big) T \log T \quad . \tag{36}$$

With a(s) defined in (29), we note that

$$|a(it)|^2 = 4\Big(\frac{\sin \kappa t}{t} \Big)^2$$

where $\kappa = \frac{1}{2} \log (1 + \delta)$. Then by Lemma 2 we deduce that

$$\begin{aligned} \int_0^\infty |a(it)|^2 \Big| \sum_\gamma \frac{X^{i\gamma}}{1 + (t-\gamma)^2} \Big|^2 dt &= (\pi/2 + o(1))\kappa \log 1/\kappa \\ &= (\pi/4 + o(1))\delta \log 1/\delta \ . \end{aligned} \tag{37}$$

The values of T for which we have used (8) lie in the range

$$\delta^{-1}(\log 1/\delta)^{-2} \leqslant T \leqslant 3\delta^{-1}(\log 1/\delta)^2 \ . \tag{38}$$

The integrand is even, so that the value is doubled if we integrate over negative values of t as well. Then by Lemma 10

$$\int_{-\infty}^{+\infty} \Big| \sum_{|\gamma| \leqslant Z} \frac{a(\rho)\, X^{i\gamma}}{1 + (t-\gamma)^2} \Big|^2 dt = \big(\pi/2 + o(1)\big)\delta \log 1/\delta$$

provided that $Z \geqslant \delta^{-1}(\log 1/\delta)^3$. Let S(t) denote the above sum over γ. Its Fourier transform is

$$\hat{S}(u) = \int_{-\infty}^{+\infty} S(t)\, e(-tu) dt = \pi \sum_{|\gamma| \leqslant Z} a(\rho)\, X^{i\gamma}\, e(-\gamma u) e^{-2\pi|u|} \ .$$

Hence by Plancherel's identity the integral above is

$$= \pi^2 \int_{-\infty}^{+\infty} \Big| \sum_{|\gamma| \leqslant Z} a(\rho) X^{i\gamma}\, e(-\gamma u) \Big|^2 e^{-4\pi|u|}\, du \ .$$

On writing $Y = \log X$, $-2\pi u = y$, we find that

$$\int_{-\infty}^{+\infty} \Big| \sum_{|\gamma| \leq Z} a(\rho)\, e^{i\gamma(Y+y)} \Big|^2 e^{-2|y|}\, dy = (1 + o(1))\delta \log 1/\delta \ . \tag{39}$$

In Lemma 1 we take

$$R(y) = \begin{cases} e^{2y} & 0 \leq y \leq \log 2, \\ 0 & \text{otherwise} . \end{cases}$$

On making the change of variable $x = e^{Y+y}$ we deduce that

$$\int_X^{2X} \Big| \sum_{|\gamma| \leq Z} a(\rho) x^{\rho} \Big|^2 dx = \big(3/2 + o(1) \big)\ \delta X^2 \log 1/\delta \ .$$

We replace X by $X2^{-k}$, sum over k, $1 \leq k \leq K$, and use the explicit formula (34) with $Z = X(\log X)^3$ to see that

$$\int_{X2^{-K}}^{X} \big(\psi\big((1+\delta)x\big) - \psi(x) - \delta x\big)^2\, dx = \frac{1}{2} \big(1 - 2^{-2K} + o(1)\big)\ \delta X^2 \log 1/\delta .$$

We take $K = [\log\log X]$, and note that it suffices to have (8) in the range (11). To bound the contribution of the range $1 \leq x \leq X2^{-K}$, we appeal to (3) with X replaced by $X2^{-K}$. Thus we have (10) .

We now deduce (8) from (10). By integrating (10) by parts from X_1 to $X_2 = X_1(\log X_1)^{2/3}$, we find that

$$\int_{X_1}^{X_2} \big(\psi\big((1+\delta)x\big) - \psi(x) - \delta x\big)^2\, x^{-4} dx = \big(\tfrac{1}{2} + o(1)\big)\delta(\log 1/\delta) X_1^{-2} .$$

From (3) we similarly deduce that

$$\int_{X_2}^{\infty} \big(\psi\big((1+\delta)x\big) - \psi(x) - \delta x\big)^2 x^{-4}\, dx \ll \delta\, (\log 1/\delta)^2\, X_2^{-2}$$

$$= O\big(\delta\, (\log 1/\delta)\, X_1^{-2} \big) \ .$$

We add these relations, and multiply through by X_1^2 . By making a further appeal to (10) with $X = X_1$ we deduce that

$$\int_0^{\infty} \min(x^2/X_1^2,\ X_1^2/x^2)\big(\psi\big((1+\delta)x\big) - \psi(x) - \delta x\big)^2\, x^{-2}\, dx$$

$$= (1 + o(1))\delta \log 1/\delta \ .$$

We write X for X_1, put $Y = \log X$, $x = e^{Y+y}$, and appeal to the explicit formula (34) with $Z = X(\log X)^3$, and we find that we have (39). Retracing our steps, we find that we have (37). Then by Lemma 4 we obtain (36), and hence (8) . The values of δ and X for which we have used (10) also satisfy (12) .

6. Proof of the Corollary.

We note that Lemma 8 gives (8) when

$$X(\log X)^{-3} \leqslant T \leqslant X,$$

and that (10) is trivial when

$$X^{-1}(\log X)^{-3} \leqslant \delta \leqslant X^{-1}.$$

Thus the equivalence of (a) and (b) follows immediately from Theorem 2.

We now show that (b) implies (c). We suppress the converse argument, which is similar. The method here is that of Saffari and Vaughan [14]. Our first goal is to deduce from (b) that

$$\int_0^H \int_0^X (\psi(x+h) - \psi(x) - h)^2 \, dxdh \sim \frac{1}{2} H^2 X \log X/H \tag{40}$$

uniformly for $1 \leqslant H \leqslant X^{1-\varepsilon}$. To this end it suffices to show that

$$\int_{\frac{1}{2}X}^{X} \int_0^H (\psi(x+h) - \psi(x) - h)^2 \, dh \, dx \sim \frac{1}{4} H^2 X \log X/H \tag{41}$$

In this integral we replace h by $\delta = h/x$, and invert the order of integration. Thus the left hand side above is

$$\int_0^{H/X} \int_{\frac{1}{2}X}^{X} f(x,\delta x)^2 \, x \, dx \, d\delta + \int_{H/X}^{2H/X} \int_{\frac{1}{2}X}^{H/\delta} f(x,\delta x)^2 \, x \, dx \, d\delta$$

where $f(x,y) = \psi(x+y) - \psi(x) - y$. By integrating by parts, we see from (b) that if $A \approx B \approx X$ then

$$\int_A^B f(x,\delta x)^2 \, x \, dx = \frac{1}{3} (B^3 - A^3) \, \delta \log 1/\delta + O(X^3 \delta \log 1/\delta).$$

This yields (41) . Then (40) follows by replacing X by $X2^{-k}$ in (41), summing over $0 \leqslant k \leqslant K = [2 \log\log X]$, and by appealing to (4) with X replaced by $X2^{-K-1}$.

We now deduce (c) from (40). Suppose that $0 < \eta < 1$. By differencing in (40) we see that

$$\int_H^{(1+\eta)H} \int_0^X f(x,h)^2 \, dx \, dh = (\eta + \tfrac{1}{2}\eta^2 + o(1))XH^2 \log X/H \quad .$$

Let $g(x,h) = f(x,H)$. From the identity

$$f^2 - g^2 = 2f(f-g) - (f-g)^2$$

and the Cauchy-Schwartz inequality we find that

$$\iint f^2 - g^2 \ll \left(\iint f^2\right)^{1/2} \left(\iint (f-g)^2\right)^{1/2} + \iint (f-g)^2 \quad .$$

But $f(x,h) - g(x,h) = f(x+H,h-H)$, so that

$$\begin{aligned} \iint (f-g)^2 &= \int_0^{\eta H} \int_H^{X+H} f(x,h)^2 \, dx \, dh \\ &\ll \eta^2 H^2 X \log X/H \end{aligned}$$

by (40). Hence we see that

$$\begin{aligned} \eta H \int_0^X (\psi(x+H) - \psi(x) - H)^2 \, dx &= \iint g^2 \\ &= \iint f^2 + O(\eta^{3/2} XH^2 \log X/H) \\ &= \left(\eta + O(\eta^{3/2}) + o(1)\right) XH^2 \log X/H \ . \end{aligned}$$

We now divide both sides by ηH, and obtain the desired result by letting $\eta \to 0^+$ sufficiently slowly.

References.

1. H. Davenport, Multiplicative Number Theory, Second Edition, Springer-Verlag, 1980.

2. A. Fujii, On the zeros of Dirichlet L-functions, I, *Trans. Amer. Math. Soc.* **196** (1974), 225-235. (Corrections to this paper are noted in *Trans. Amer. Math. Soc.* **267** (1981), pp 38-39, and in [5; pp. 219-220].)

3. P.X. Gallagher, A large sieve density estimate near $\sigma = 1$, *Invent. Math.* **11** (1970), 329-339.

4. P.X. Gallagher, On the distribution of primes in short intervals, *Mathematika* **23** (1976), 4-9.

5. P.X. Gallagher and Julia H. Mueller, Primes and zeros in short intervals, *J. Reine Agnew. Math.* **303/304** (1978), 205-220.

6. Daniel A. Goldston, Large differences between consecutive prime numbers, Thesis, University of California Berkeley, 1981.

7. H.Halberstam and H.-E. Richert, Sieve Methods. Academic Press, London, 1974.

8. G.H. Hardy, Divergent Series, Oxford University Press, 1963.

9. G.H. Hardy and W.W. Rogosinski, Notes on Fourier series (I): On sine series with positive coefficients, *J. London Math. Soc.* **18** (1943), 50-57.

10. D.R. Heath-Brown, Gaps between primes, and the pair correlation of zeros of the zeta-function, *Acta Arith.* **41** (1982), 85-99.

11. D.R. Heath-Brown and D.A. Goldston, A note on the difference between consecutive primes, *Math. Ann.* **266** (1984), 317-320.

12. Julia Huang (=J.H. Mueller), Primes and zeros in short intervals, Thesis, Columbia University, 1976.

13. H.L Montgomery, The pair correlation of zeros of the zeta function, *Proc. Sympos. Pure Math.* **24** (1973), 181-193.

14. B. Saffari and R.C. Vaughan, On the fractional Parts of x/n and related sequences II, *Ann. Inst. Fourier* (Grenoble) **27**, no. 2, (1977), 1-30.

15. A. Selberg, On the normal density of primes in small intervals, and the difference between consecutive primes, *Arch. Math. Naturvid.* **47**, no. 6, (1943), 87-105.

16. E.C. Titchmarsh, The theory of the Riemann zeta-function, Oxford University Press, 1951.

17. J.D. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.*, **12**, No. 2, (1985), 183-216.

D. A. Goldston
San Jose State University,
San Jose, CA 95192,
U.S.A.

H. L. Montgomery
University of Michigan,
Ann Arbor, MI 48109,
U.S.A.

THÉORIE DES NOMBRES. — Sur le Théorème des Nombres Premiers. Note de **Hédi Daboussi,** présentée par Jean-Pierre Kahane.

Remise le 19 décembre 1983.

Nous donnons une nouvelle démonstration du théorème des nombres premiers n'utilisant pas l'inégalité de Selberg.

NUMBER THEORY. — On the Prime Number Theorem.

We give a new elementary proof of the prime number theorem which does not use Selberg's inequality.

A H. Delange et P. Erdős à l'occasion de leur 70^e anniversaire.

I. 1. Soit $y \geqq 2$ et v_y, u_y deux fonctions complètement multiplicatives définies par :

$$v_y(p) = \begin{cases} 1 & \text{si } p \leqq y \\ 0 & \text{si } p > y' \end{cases} \qquad u_y(p) = \begin{cases} 1 & \text{si } p > y \\ 0 & \text{si } p \leqq y' \end{cases}$$

la lettre p désignant des nombres premiers.

Λ désigne la fonction de Von Mangoldt, $\mathbf{1}$ la fonction constante égale à 1, μ la fonction de Möbius; ainsi, par exemple $\log n = (\Lambda * \mathbf{1})(n)$, où le signe $*$ désigne la convolution de Dirichlet. On notera $V_y(t) = \sum_{n \leqq t} v_y(n) \mu(n)$, $V_y^*(t) = \sum_{n \leqq t} v_y(n)$ et $M(t) = \sum_{n \leqq t} \mu(n)$. Nous montrerons que $\lim_{x \to \infty} |M(x)/x| = 0$.

I. 2. *Aperçu de la méthode.* — Nous démontrerons que pour tout $y \geqq 2$:

$$\overline{\lim_{x \to \infty}} |M(x)/x| \leqq \{ \prod_{p \leqq y} (1-(1/p)) \} \int_1^\infty (|V_y(t)|/t^2)\, dt. \tag{1}$$

Soit $\alpha = \overline{\lim} |M(x)|/x$; évidemment $\alpha \leqq 1$.

Nous établirons qu'il existe $\delta > 1$ tel que pour tout β, $\alpha < \beta < 2$, on a :

$$\int_1^y (|V_y(t)|/t^2)\, dt \leqq \beta/\delta \log y + O(1), \tag{2}$$

et que l'on a :

$$\int_y^\infty (|V_y(t)|/t^2)\, dt \leqq \beta(C-1) \log y + o(\log y), \tag{3}$$

où $C = \lim_{y \to \infty} (\log y)^{-1} \prod_{p \leqq y} (1-(1/p))^{-1}$ (Il est connu que $C = e^\gamma$, où γ est la constante d'Euler, nous n'en ferons pas usage).

(3) entraînera que $\alpha \leqq \beta(1 - C^{-1}(1-(1/\delta))$, le facteur de β étant < 1, il en résulte, en faisant tendre β vers α, que $\alpha = 0$).

II. 1. Les séries $\Sigma\, v_y(n)/n$ et $\Sigma\, (v_y(n)/n) \mu(n)$ sont absolument convergentes avec pour sommes $\prod_{p \leqq y} (1-(1/p))^{-1}$ et $\prod_{p \leqq y} (1-(1/p))$.

On en déduit (en partant de $u_y = v_y \mu * \mathbf{1}$) que :

$$\lim_{x \to \infty} (1/x) \sum_{n \leqq x} u_y(n) \quad \text{existe et est égale à} \quad \prod_{p \leqq y} (1-(1/p)).$$

D'après les définitions de v_y et u_y, il est clair que $\mu(n) = (\mu u_y * \mu v_y)(n)$ pour tout entier

n. Ainsi :

$$\mathrm{M}(x)=\sum_{n\leqq x}\mu(n)\,u_y(n)\,\mathrm{V}_y(x/n).$$

Notons $d_1=1<d_2\ldots<d_q$ la suite finie des entiers sans facteur carré ayant tous leurs diviseurs premiers $\leqq y$, et remarquons que, si n vérifie $x/d_{j+1}<n\leqq x/d_j$, alors $\mathrm{V}_y(x/n)=\mathrm{V}_y(d_j)$. On obtient ainsi :

$$\mathrm{M}(x)=\sum_{j=1}^{q-1}\mathrm{V}_y(d_j)\sum_{x/d_{j+1}<n\leqq x/d_j}u_y(n)\,\mu(n)+\mathrm{V}_y(d_q)\sum_{n\leqq x/d_q}u_y(n)\,\mu(n),$$

$$\overline{\lim_{x\to\infty}}\,|\mathrm{M}(x)/x|\leqq\sum_{j=1}^{q-1}|\mathrm{V}_y(d_j)|\lim_{x\to\infty}(1/x)\sum_{x/d_{j+1}<n\leqq x/d_j}u_y(n)+|\mathrm{V}_y(d_q)|\lim_{x\to\infty}(1/x)\sum_{n\leqq x/d_q}u_y(n).$$

Ce qui fournit (1).

II. 2. On sait qu'il existe $\mathrm{M}>0$ tel que, pour tous nombres a et b positifs :

$$\left|\int_a^b(\mathrm{M}(t)/t^2)\,dt\right|\leqq\mathrm{M}.$$

Prenons $\alpha<\beta<2$ et x_β tel que pour $x\geqq x_\beta$, $|\mathrm{M}(x)|\leqq\beta x$.

Soit $\delta=\min(2,\ 1+(\alpha^2/4\,\mathrm{M}))$.

Puisque $v_y(n)=1$ si $n\leqq y$ et donc $\mathrm{V}_y(t)=\mathrm{M}(t)$ si $t\leqq y$, l'inégalité (2) s'écrit :

$$(2)'\qquad\int_1^y(|\mathrm{M}(t)|/t^2)\,dt\leqq\beta/\delta\log y+O(1).$$

Une telle inégalité intervient dans la méthode de Selberg ([1], [3]). L'inégalité (2)′ s'établit par la méthode utilisée en [2] pour prouver le lemme 5.8.

III. Quelques lemmes.

III. 1. Lemme 1. — *Soit h une fonction définie sur* $[y,\ +\infty[$, *positive, décroissante et possédant une dérivée continue. On a :*

Pour tout $t\geqq y$: $\displaystyle\sum_{p\leqq y}(\log p/p)\,h(pt)=\int_t^{yt}(h(v)/v)\,dv+O(h(y)).$

Pour tout $t\geqq 1$: $\displaystyle\sum_{y/t<p\leqq y}(\log p/p)\,h(pt)=\int_y^{yt}(h(v)/v)\,dv+O(h(y)).$

Ce lemme s'obtient par intégration par parties grâce à la relation :

$$\sum_{p\leqq t}\log p/p=\log t+O(1).$$

III. 2. Lemme 2. — *Posons, pour $s>0$:*

$$k(s)=\int_0^\infty e^{-sx}\,e^{f(x)}\,dx,\qquad où\quad f(x)=\int_0^x((1-e^{-u})/u)\,du.$$

Alors la fonction k est positive, décroissante et indéfiniment dérivable. De plus :

$$(4)\qquad sk(s)-\int_s^{s+1}k(u)\,du=1\quad\text{pour tout }s>0.$$

$\Big[$Il est immédiat que $\displaystyle\int_s^{s+1}k(u)\,du=\int_0^\infty e^{f(x)}\,e^{-sx}f'(x)\,dx$. En intégrant par parties on obtient (4)$\Big]$.

III.3. LEMME 3. — *Soit k la fonction définie au lemme 2, on a :*

$$\int_1^2 k(u)\,(2-u)\,du = C-1. \tag{5}$$

Nous établirons ce lemme par une méthode purement arithmétique, une méthode analogue nous fournira (3). De la relation $\log = \Lambda * \mathbf{1}$, nous déduisons $v_y \log = v_y \Lambda * v_y$, et donc :

$$\sum_{n \leqq t} v_y(n) \log n = \sum_{n \leqq t} v_y(n)\,\Lambda(n)\,V_y^*(t/n);$$

ou encore, $V_y^*(t) \log t = \sum_{n \leqq t} v_y(n)\,\Lambda(n)\,V_y^*(t/n) + \sum_{n \leqq t} v_y(n) \log(t/n)$.

Par définition de Λ et v_y, on a :

$$V_y^*(t) \log t = \sum_{\substack{p \leqq t \\ p \leqq y}} \log p\, V_y^*(t/p) + \sum_{\substack{p \leqq y \\ p^r \leqq t,\, r \geqq 2}} \log p\, V_y^*(t/p^r) + \sum_{n \leqq t} v_y(n) \log(t/n).$$

Posons pour $t > y$: $h(t) = (1/\log y)\, k(\log t/\log y)$. Alors :

$$\int_y^\infty (V_y^*(t)/t^2) \log t \,.\, h(t)\,dt = \int_y^\infty \sum_{\substack{p \leqq t \\ p \leqq y}} \log p\, V_y^*(t/p)\,(h(t)/t^2)\,dt + E_1 + E_2, \tag{6}$$

où

$$E_1 = \int_y^\infty \sum_{\substack{p \leqq y \\ p^r \leqq t,\, r \geqq 2}} \log p\, V_y^*(t/p^r)\,(h(t)/t^2)\,dt, \qquad E_2 = \int_y^\infty \sum_{n \leqq t} v_y(n) \log(t/n)\,(h(t)/t^2)\,dt.$$

La décroissance de h entraîne que :

$$E_1 \leqq h(y)\,.\,\Big(\sum_{\substack{r \geqq 2 \\ p}} \log p/p^r\Big)\,.\int_1^\infty (V_y^*(u)/u^2)\,du, \qquad E_2 \leqq h(y)\,.\,\Big(\sum_n v_y(n)/n\Big)\,.\int_1^\infty (\log t/t^2)\,dt.$$

Par ailleurs $\sum v_y(n)/n = \int_1^\infty (V_y^*(u)/u^2)\,du = O(\log y)$, ce qui implique que $E_1 = O(1)$ et $E_2 = O(1)$. L'intégrale à droite de (6) s'écrit :

$$\int_y^\infty \sum_{\substack{p \leqq t \\ p \leqq y}} \log p\, V_y^*(t/p)\,(h(t)/t^2)\,dt = \sum_{p \leqq y} \log p/p \int_{y/p}^y (V_y^*(t)/t^2)\,h(pt)\,dt$$

$$+ \sum_{p \leqq y} \log p/p \int_y^\infty (V_y^*(t)/t^2)\,h(pt)\,dt$$

$$= \int_1^y V_y^*(t)/t^2 \sum_{y/t < p \leqq y} (\log p/p)\,h(pt)\,dt + \int_y^\infty V_y^*(t)/t^2 \sum_{p \leqq y} (\log p/p)\,h(pt)\,dt,$$

ce qui, grâce au lemme 1, donne :

$$\int_y^\infty V_y^*(t)/t^2 \left\{ \log t\,.\,h(t) - \int_t^{yt} (h(v)/v)\,dv \right\} dt = \int_1^y V_y^*(t)/t^2 \left(\int_y^{yt} (h(v)/v)\,dv \right) dt + O(1).$$

Il découle du lemme 2 et de la définition de h que :

$$\log t\,.\,h(t) - \int_t^{yt} (h(v)/v)\,dv = 1 \quad \text{pour tout } t \geqq 1.$$

En utilisant également le fait que $V_y^*(t) = [t] = t + O(1)$ pour tout $t \leqq y$, on obtient par un calcul simple que :

$$\int_y^\infty (V_y^*(t)/t^2)\,dt = \left(\int_1^2 k(u)\,(2-u)\,du\right)\log y + O(1).$$

Par ailleurs :

$$\int_y^\infty (V_y^*(t)/t^2)\,dt = \sum_n v_y(n)/n - \sum_{n\leqq y} v_y(n)/n = (\mathrm{C}+o(1))\log y - \log y + O(1).$$

Ces deux formes de l'intégrale fournissent l'égalité (5).

IV. PREUVE DE L'INÉGALITÉ (3). — De la relation : $-\mu\log=\mu*\Lambda$, nous déduisons que $-v_y\,\mu\log = v_y\,\mu * v_y\,\Lambda$.

En raisonnant comme au paragraphe précédent, nous avons successivement :

$$|\mathrm{V}_y(t)|\log t \leqq \sum_{n\leqq t} v_y(n)\,\Lambda(n)\,|\mathrm{V}_y(t/n)| + \sum_{n\leqq t} v_y(n)\log(t/n),$$

et, avec la fonction h définie plus haut,

$$\int_y^\infty |\mathrm{V}_y(t)|/t^2 \left\{\log t\,.\,h(t) - \int_t^{yt}(h(v)/v)\,dv\right\}dt \leqq \int_1^y |\mathrm{V}_y(t)|/t^2\left(\int_y^{yt}(h(v)/v)\,dv\right)dt + O(1),$$

et finalement :

$$\int_y^\infty (|\mathrm{V}_y(t)|/t^2)\,dt \leqq \int_1^y |\mathrm{M}(t)|/t^2\left(\int_y^{yt}(h(v)/v)\,dv\right)dt + O(1).$$

En majorant $|\mathrm{M}(t)|$ par βt pour $t\geqq x_\beta$ et en effectuant l'intégration à droite, on a :

$$\int_y^\infty (|\mathrm{V}_y(t)|/t^2)\,dt \leqq \beta\left(\int_1^2 k(u)\,(2-u)\,du\right)\log y + O(1),$$

et donc l'inégalité (3) grâce au lemme 3.

RÉFÉRENCES BIBLIOGRAPHIQUES

[1] P. ERDŐS, On a new method in elementary number theory which leads to an elementary proof of the prime Number Theorem, *Proc. Nat. Acad. Sc. U.S.A.*, 35, 1949, p. 374-384.

[2] N. LEVINSON, A motivated account of an elementary proof of the prime number theorem, *Amer. Math. Monthly*, 76, 1969, p. 225-245.

[3] A. SELBERG, An elementary proof of the prime number theorem, *Ann. of Math.*, (2), 50, 1949, p. 305-313.

H. D. : *Université Paris-Sud, Département de Mathématiques*, bât. 425, 91405 *Orsay*.

On Newman's Quick Way to the Prime Number Theorem

J. Korevaar

1. Introduction and Overview

There are several interesting functions in number theory whose tables look quite irregular, but which exhibit surprising asymptotic regularity as $x \to \infty$. A notable example is the function $\pi(x)$ which counts the number of primes p not exceeding x.

1.1. The Famous Prime Number Theorem

$$\pi(x) = \sum_{p \leqslant x} 1 \sim \frac{x}{\log x} \quad \text{as} \quad x \to \infty, \tag{1.1}$$

was surmised already by Legendre and Gauss. However, it took a hundred years before the first proofs appeared, one by Hadamard and one by de la Vallée Poussin (1896). Their and all but one of the subsequent proofs make heavy use of the Riemann zeta function. (The one exception is the long so-called elementary proof by Selberg [11] and Erdös [4].)

For Re $s > 1$ the zeta function is given by the Dirichlet series

D. J. Newman

$$\zeta(s) = \sum_{1}^{\infty} \frac{1}{n^s}. \tag{1.2a}$$

By the unique representation of positive integers n as products of prime powers, the series may be converted to the Euler product (cf. [5])

$$\begin{aligned}\zeta(s) &= \left(1 + \frac{1}{p_1^s} + \frac{1}{p_1^{2s}} + \dots\right)\left(1 + \frac{1}{p_2^s} + \frac{1}{p_2^{2s}} + \dots\right)\dots \\ &= \prod_p \frac{1}{1 - p^{-s}}.\end{aligned} \tag{1.2b}$$

The above function element is analytic for Re $s > 1$ and can be continued across the line Re $s = 1$ (Fig. 1). More precisely, the difference

$$\zeta(s) - \frac{1}{s-1}$$

can be continued analytically to the half-plane Re $s > 0$ (cf. § B.1 in the box on p. 111) and in fact to all of $\mathbb{C}$. The essential property of $\zeta(s)$ in the proofs of the prime number theorem is its non-vanishing on the line Re $s = 1$

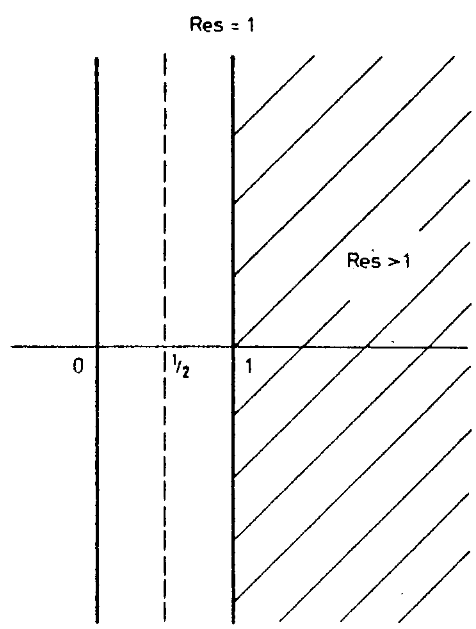


Figure 1

(cf. § B.2). [The zeta function has many zeros in the strip $0 < \text{Re } s < 1$. Riemann's conjecture (1859) that they all lie on the central line $\text{Re } s = \frac{1}{2}$ remains unproven to this day.]

For about fifty years now, the standard proofs of the prime number theorem have involved some form of Wiener's Tauberian theory for Fourier integrals, usually the Ikehara-Wiener theorem of § 1.2 (see Wiener [14] and cf. various books, for example Doetsch [2], Chandrasekharan [1], Heins [6]). Thus the proof of the prime number theorem has remained quite difficult until the recent breakthrough by D. J. Newman [10].

In 1980, he succeeded in replacing the Wiener theory in the proof by an ingenious application of complex integration theory, involving nothing more difficult than Cauchy's integral formula, together with suitable estimates. We present Newman's method in § 2 (applying it to Laplace integrals instead of Dirichlet series). In this method, the Ikehara-Wiener Tauberian theorem is replaced by a poor man's version which also readily leads to the prime number theorem.

Excellent accounts of the history of the prime number theorem and the zeta function may be found in the books of Landau [9], Ingham [7], Titchmarsh [13] and Edwards [3].

1.2. A Gem from Ingham's Work

Newman's method leads directly to the following pretty theorem which is already contained in work of Ingham [8]. However, Ingham used Wiener's method to prove his (more general) results.

Auxiliary Tauberian theorem. *Let $F(t)$ be bounded on $(0, \infty)$ and integrable over every finite subinterval, so that the Laplace transform*

$$G(z) = \int_0^\infty F(t)e^{-zt}dt \tag{1.3}$$

is well-defined and analytic throughout the open half-plane $\text{Re } z > 0$. *Suppose that $G(z)$ can be continued analytically to a neighborhood of every point on the imaginary axis. Then*

$$\int_0^\infty F(t)dt \text{ exists} \tag{1.4}$$

as an improper integral [*and is equal to* $G(0)$].

Under the given hypothesis, the Laplace integral (1.3) will converge everywhere on the imaginary axis. For the conclusion (1.4), it is actually sufficient that $G(z)$ have a continuous extension to the closed half-plane $\text{Re } z \geqslant 0$ which is smooth at $z = 0$: see § 2.

At first glance, the above theorem looks quite different from the

Ikehara-Wiener theorem [14]: Let $f(x)$ be nonnegative and nondecreasing on $[1, \infty)$ and such that the Mellin transform

$$g_0(s) = \int_1^\infty x^{-s}df(x) = -f(1) + s\int_1^\infty f(x)x^{-s-1}dx$$

exists for $\text{Re } s > 1$. Suppose that for some constant c, the function

$$g_0(s) - \frac{c}{s-1}$$

has a continuous extension to the closed half-plane $\text{Re } s \geqslant 1$. Then

$$f(x)/x \to c \quad \text{as} \quad x \to \infty.$$

This is an extremely useful theorem, but what could we do with the auxiliary theorem in the same direction? We will show that the latter has a corollary which is just as good for the application that we want to make.

1.3. A Poor Man's Ikehara-Wiener Theorem

We will establish the following

Corollary to the auxiliary theorem. *Let $f(x)$ be nonnegative, nondecreasing and $O(x)$ on $[1, \infty)$, so that its Mellin transform*

$$g(s) = s\int_1^\infty f(x)x^{-s-1}dx \tag{1.5}$$

is well-defined and analytic throughout the half-plane $\text{Re } s > 1$. *Suppose that for some constant c, the function*

$$g(s) - \frac{c}{s-1} \tag{1.6}$$

can be continued analytically to a neighborhood of every point on the line $\text{Re } s = 1$. *Then*

$$f(x)/x \to c \quad \textit{as} \quad x \to \infty. \tag{1.7}$$

Derivation from the auxiliary theorem. Let $f(x)$ and $g(s)$ satisfy the hypotheses of the corollary. We set $x = e^t$ and define

$$e^{-t}f(e^t) - c = F(t),$$

so that $F(t)$ is bounded on $(0, \infty)$. Its Laplace transform will be

$$G(z) = \int_0^\infty \{e^{-t}f(e^t) - c\}\, e^{-zt}dt$$

$$= \int_1^\infty f(x)x^{-z-2}dx - \frac{c}{z} = \frac{1}{z+1}\left\{g(z+1) - \frac{c}{z} - c\right\}.$$

Thus by the hypothesis of the corollary, $G(z)$ can be continued analytically to a neighborhood of every point on the imaginary axis. We may now apply the auxiliary theorem from § 1.2.

What does its conclusion tell us? Setting $t = \log x$ we find that the improper integrals

$$\int_0^\infty \{e^{-t}f(e^t) - c\}\, dt = \int_1^\infty \frac{f(x) - cx}{x^2}\, dx \tag{1.8}$$

exist. Using the fact that $f(x)$ is an increasing function, one readily derives that $f(x) \sim cx$ in the sense of (1.7).

Indeed, suppose for a moment that $\limsup f(x)/x > c$ $(\geqslant 0)$. Then there would be a positive constant δ such that for certain arbitrarily large numbers y

$$f(y) > (c + 2\delta)y.$$

It would follow that

$$f(x) > (c + 2\delta)y > (c + \delta)x \quad \text{for} \quad y < x < \rho y$$

where $\rho = (c + 2\delta)/(c + \delta)$. But then

$$\int_y^{\rho y} \frac{f(x) - cx}{x^2}\, dx > \int_y^{\rho y} \frac{\delta}{x}\, dx = \delta \log \rho$$

for those same numbers y, contradicting the existence of (1.8).

One similarly disposes of the contingency $\liminf f(x)/x < c$ (in this case c would have to be positive and one would consider intervals $\theta y < x < y$ with $\theta < 1$ where $f(x) < (c - \delta)x$). Thus $f(x)/x \to c$.

1.4. Corollary ⇒ Prime Number Theorem

This step is routine to number theorists. One takes $f(x) = \psi(x)$, where $\psi(x)$ is that well-known function from prime number theory,

$$\psi(x) = \sum_{p^m \leqslant x} \log p \tag{1.9}$$

(the summation is over all prime powers not exceeding x). It is a simple fact (first noticed by Chebyshev) that $\pi(x) = O(x/\log x)$ or equivalently, $\psi(x) = O(x)$ (cf. § B.4 in the box for more details). Thus $f(x)$ is as the corollary wants it.

What about its Mellin transform $g(s)$? A standard calculation based on the Euler product in (1.2) shows that

$$g(s) = -\frac{\zeta'(s)}{\zeta(s)}, \quad \operatorname{Re} s > 1$$

(cf. § B.3). Since $\zeta(s)$ behaves like $1/(s-1)$ around $s = 1$, the same is true for $g(s)$. The analyticity of $\zeta(s)$ at the points of the line $\operatorname{Re} s = 1$ (different from $s = 1$) and its non-vanishing there imply that $g(s)$ can be continued analytically to a neighborhood of every one of those points (cf. §§ B.1, B.2). Thus

$$g(s) - \frac{1}{s-1}$$

has an analytic continuation to a neighborhood of the closed half-plane $\operatorname{Re} s \geqslant 1$.

The conclusion of the corollary now tells us that

$$\psi(x)/x \to 1 \quad \text{as} \quad x \to \infty,$$

and this is equivalent to the prime number theorem (1.1) (cf. § B.4).

2. Newman's Beautiful Method

2.1. Proof of the Auxiliary Tauberian Theorem

Let $F(t)$ be bounded on $(0, \infty)$ and such that its Laplace transform $G(z)$ can be continued to a function (still called $G(z)$) which is analytic in a neighborhood of the closed half-plane $\operatorname{Re} z \geqslant 0$. We may and will assume that

$$|F(t)| \leqslant 1, \quad t > 0.$$

For $0 < \lambda < \infty$ we write

$$G_\lambda(z) = \int_0^\lambda F(t)e^{-zt}dt. \tag{2.1}$$

Observe that $G_\lambda(z)$ is analytic for all z. We will show that

$$G_\lambda(0) = \int_0^\lambda F(t)dt \to G(0) \quad \text{as} \quad \lambda \to \infty.$$

Some details left out in 1.4

We begin with the necessary facts about the zeta function.

B.1. Analytic continuation of $\zeta(s)$. Simple transformations show that for Re $s > 2$

$$\zeta(s) = \sum_1^\infty \frac{n}{n^s} - \sum_1^\infty \frac{n-1}{n^s} = \sum_1^\infty \frac{n}{n^s} - \sum_1^\infty \frac{n}{(n+1)^s} = \sum_1^\infty n\left\{\frac{1}{n^s} - \frac{1}{(n+1)^s}\right\} = \sum_1^\infty ns\int_n^{n+1} x^{-s-1}dx = s\sum_1^\infty \int_n^{n+1} [x]x^{-s-1}dx =$$

$$= s\int_1^\infty [x]x^{-s-1}dx, \tag{B.1}$$

where $[x]$ denotes the largest integer $\leqslant x$. Since first and final member are analytic for Re $s > 1$, the integral formula holds throughout that half-plane.

It is reasonable to compare the integral with

$$s\int_1^\infty x\cdot x^{-s-1}dx = \frac{s}{s-1} = 1 + \frac{1}{s-1}. \tag{B.2}$$

Combination of (B.1) and (B.2) gives

$$\zeta(s) - \frac{1}{s-1} = 1 + s\int_1^\infty ([x]-x)x^{-s-1}dx. \tag{B.3}$$

The new integral converges and represents an analytic function throughout the half-plane Re $s > 0$. Thus (B.3) provides an analytic continuation of the left-hand side to that half-plane.

B.2. Non-vanishing of $\zeta(s)$ *for* Re $s \geqslant 1$. The Euler product in (1.2) shows that $\zeta(s) \neq 0$ for Re $s > 1$. For Re $s = 1$ we will use Mertens's clever proof of 1898. The key fact is the inequality

$$3 + 4\cos\theta + \cos 2\theta = 2(1+\cos\theta)^2 \geqslant 0, \quad \theta \text{ real.} \tag{B.4}$$

Suppose that $\zeta(1+ib)$ would be equal to 0, where b is real and $\neq 0$. Then the auxiliary analytic function

$$\varphi(s) = \zeta^3(s)\zeta^4(s+ib)\zeta(s+2ib)$$

would have a zero for $s = 1$: the pole of $\zeta^3(s)$ could not cancel the zero of $\zeta^4(s+ib)$. It would follow that

$$\log|\varphi(s)| \to -\infty \quad \text{as} \quad s \to 1. \tag{B.5}$$

We now take s *real* and > 1. By the Euler product,

$$\log|\zeta(s+it)| = -\text{Re}\sum_p \log(1-p^{-s-it}) = \text{Re}\sum_p \{p^{-s-it} + \tfrac{1}{2}(p^2)^{-s-it} + \tfrac{1}{3}(p^3)^{-s-it} + \ldots\} = \text{Re}\sum_1^\infty a_n n^{-s-it} \quad \text{with} \quad a_n \geqslant 0.$$

Thus

$$\log|\varphi(s)| = \text{Re}\sum_1^\infty a_n n^{-s}(3 + 4n^{-ib} + n^{-2ib}) = \sum_1^\infty a_n n^{-s}\{3 + 4\cos(b\log n) + \cos(2b\log n)\} \geqslant 0$$

because of (B.4), contradicting (B.5).

B.3. Representations for $\zeta'(s)/\zeta(s)$. Logarithmic differentiation of the Euler product in (1.2) gives

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{p^{-s}}{1-p^{-s}}\log p = \sum_p (p^{-s} + p^{-2s} + \ldots)\log p = \sum_1^\infty \Lambda(n)n^{-s}, \tag{B.6}$$

where $\Lambda(n)$ is the von Mangoldt function,

$$\Lambda(n) = \begin{cases} \log p \text{ if } n = p^m, \\ 0 \quad \text{ if } n \text{ is not a prime power.} \end{cases}$$

The corresponding partial sum function is equal to $\psi(x)$:

$$\psi(x) = \sum_{p^m \leqslant x} \log p = \sum_{n \leqslant x} \Lambda(n). \tag{B.7}$$

Proceeding as in (B.1), the series (B.6) leads to the integral representation

$$-\frac{\zeta'(s)}{\zeta(s)} = s \int_1^\infty \psi(x) x^{-s-1} dx, \quad \text{Re } s > 1. \tag{B.8}$$

The integral converges and is analytic for Re $s > 1$ since by (B.7), $\psi(x) \leqslant x \log x$.

B.4. Relation between $\psi(x)$ and $\pi(x)$. By (B.7), $\psi(x)$ counts $\log p$ (for fixed p) as many times as there are powers $p^m \leqslant x$, hence

$$\psi(x) = \sum_{p \leqslant x} \left[\frac{\log x}{\log p}\right] \log p \leqslant \log x \sum_{p \leqslant x} 1 = \pi(x) \log x. \tag{B.9}$$

On the other hand, when $1 < y < x$,

$$\pi(x) = \pi(y) + \sum_{y < p \leqslant x} 1 \leqslant \pi(y) + \sum_{y < p \leqslant x} \frac{\log p}{\log y} < y + \frac{\psi(x)}{\log y}.$$

Taking $y = x/\log^2 x$ one thus finds that

$$\pi(x) \frac{\log x}{x} < \frac{1}{\log x} + \frac{\psi(x)}{x} \frac{\log x}{\log x - 2 \log \log x}. \tag{B.10}$$

Combination of (B.9) and (B.10) shows that

$$\lim \pi(x) \frac{\log x}{x} = 1 \quad \text{if and only if} \quad \lim \frac{\psi(x)}{x} = 1. \tag{B.11}$$

We finally indicate a standard proof of the estimate

$$\psi(x) = 0(x). \tag{B.12}$$

For positive integral n, the binomial coefficient $\binom{2n}{n}$ must be divisible by all primes p on $(n, 2n]$. Hence

$$\prod_{n < p \leqslant 2n} p \leqslant \binom{2n}{n} < 2^{2n},$$

so that

$$\sum_{2^{k-1} < p \leqslant 2^k} \log p \leqslant 2^k \log 2.$$

It follows that

$$\sum_{p \leqslant 2^k} \log p \leqslant (2^k + 2^{k-1} + \ldots + 1) \log 2 < 2^{k+1} \log 2$$

and hence there is a constant C such that

$$\sum_{p \leqslant x} \log p \leqslant Cx.$$

Since the prime powers higher than the first contribute at most a term $0(x^{1/2+\epsilon})$ to $\psi(x)$, inequality (B.12) follows.

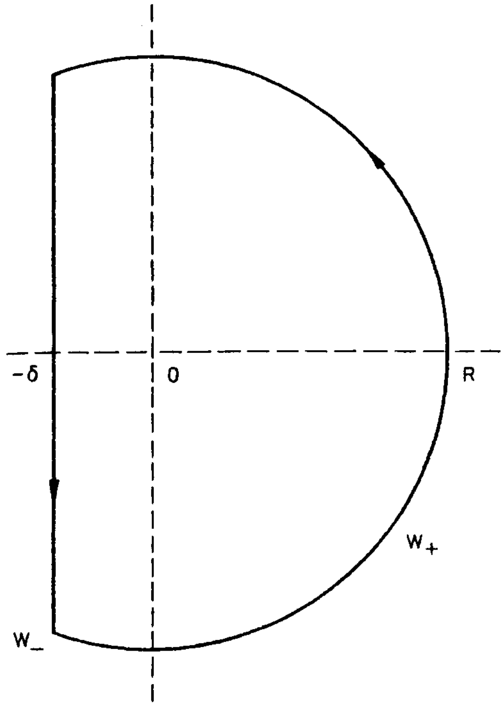


Figure 2

First idea. We try to estimate $G(0) - G_\lambda(0)$ with the aid of Cauchy's formula. Thus we look for a suitable path of integration W around 0. The simplest choice would be a circle, but we can not go too far into the left half-plane because we know nothing about $G(z)$ there. So for given $R > 0$, the positively oriented path W will consist of an arc of the circle $|z| = R$ and a segment of the vertical line $\operatorname{Re} z = -\delta$ (Fig. 2). Here the number $\delta = \delta(R) > 0$ is chosen so small that $G(z)$ is analytic on and inside W. We denote the part of W in $\operatorname{Re} z > 0$ by W_+, the part in $\operatorname{Re} z < 0$ by W_-. By Cauchy's formula,

$$G(0) - G_\lambda(0) = \frac{1}{2\pi i} \int_W \{G(z) - G_\lambda(z)\} \frac{1}{z} dz. \tag{2.2}$$

We have the following simple estimates:

for $x = \operatorname{Re} z > 0$,

$$|G(z) - G_\lambda(z)| = |\int_\lambda^\infty F(t) e^{-zt} dt| \leqslant \int_\lambda^\infty e^{-xt} dt = \frac{1}{x} e^{-\lambda x}; \tag{2.3}$$

for $x = \operatorname{Re} z < 0$,

$$|G_\lambda(z)| = |\int_0^\lambda F(t) e^{-zt} dt| \leqslant \int_0^\lambda e^{-xt} dt < \frac{1}{|x|} e^{-\lambda x}. \tag{2.4}$$

Second idea. Observe the similarity between the bounds obtained in (2.3) and (2.4)! It will be advantageous to multiply $G(z)$ and $G_\lambda(z)$ in (2.2) by $e^{\lambda z}$. This will not affect the left-hand side, but in estimating on W, the exponential $e^{-\lambda x}$ (large on W_-) will disappear from (2.3) and (2.4).

Third idea. Could we also get rid of the troublesome factor $1/|x|$ in the estimates which is bad near the imaginary axis? Yes, this can be done by adding the term z/R^2 to $1/z$ in (2.2), again without affecting the left-hand side. (For the experts: this trick is used also in Carleman's formula for the zeros of an analytic function in a half-plane, cf. [12].) The resulting modified formula is

$$G(0) - G_\lambda(0) = \frac{1}{2\pi i} \int_W \{G(z) - G_\lambda(z)\} e^{\lambda z} \left(\frac{1}{z} + \frac{z}{R^2}\right) dz. \tag{2.5}$$

Let us start harvesting. On the circle $|z| = R$,

$$\frac{1}{z} + \frac{z}{R^2} = \frac{2x}{R^2}. \tag{2.6}$$

Thus on W_+ the integrand $I(z)$ in (2.5) may be estimated as follows (see (2.3)):

$$|I(z)| \leqslant \frac{1}{x} e^{-\lambda x} e^{\lambda x} \frac{2x}{R^2} = \frac{2}{R^2}.$$

The corresponding integral is harmless:

$$\left|\frac{1}{2\pi i} \int_{W_+} I(z) dz\right| \leqslant \frac{1}{2\pi} \frac{2}{R^2} \pi R = \frac{1}{R}! \tag{2.7}$$

Fourth idea. We now turn to the part of (2.5) due to W_-. Since $G_\lambda(z)$ is analytic for all z, we may replace the integral over W_- involving $G_\lambda(z)$ by the corresponding integral over the semi-circle

$$W_-^* : \{|z| = R\} \cap \{\operatorname{Re} z < 0\}$$

(Fig. 3). Cauchy's theorem and inequality (2.4) readily give

$$\left|\frac{1}{2\pi i} \int_{W_-} G_\lambda(z) e^{\lambda z} \left(\frac{1}{z} + \frac{z}{R^2}\right) dz\right| = \left|\frac{1}{2\pi i} \int_{W_-^*} \ldots dz\right| < \frac{1}{R}. \tag{2.8}$$

We finally tackle the remaining integral

$$\frac{1}{2\pi i} \int_{W_-} G(z) e^{\lambda z} \left(\frac{1}{z} + \frac{z}{R^2}\right) dz. \tag{2.9}$$

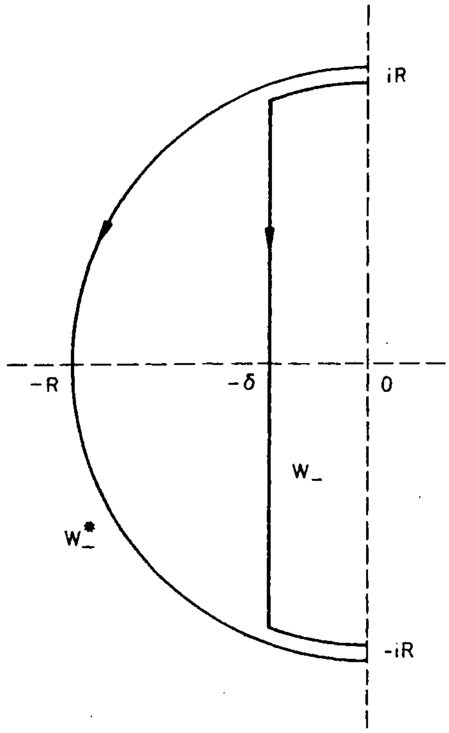


Figure 3

By the analyticity of $G(z)$ on W_- there will be a constant $B = B(R, \delta)$ such that

$$|G(z)\left(\frac{1}{z} + \frac{z}{R^2}\right)| \leqslant B \quad \text{on} \quad W_-.$$

It follows that

$$|G(z)e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right)| \leqslant Be^{\lambda x}.$$

Hence on the part of W_- where $x \leqslant -\delta_1 < 0$, the integrand in (2.9) tends to zero uniformly as $\lambda \to \infty$. On the remaining small part of W_- (we take $\delta_1 < \delta$ small), the integrand is bounded by B. Thus for fixed W, the integral in (2.9) tends to zero as $\lambda \to \infty$.

Conclusion. For given $\epsilon > 0$ one may choose $R = 1/\epsilon$. One next chooses δ so small that $G(z)$ is analytic on and inside W. One finally determines λ_0 so large that (2.9) is bounded by ϵ for all $\lambda > \lambda_0$. Then by (2.5) and (2.7)–(2.9),

$$|G(0) - G_\lambda(0)| < 3\epsilon \quad \text{for} \quad \lambda > \lambda_0.$$

In other words, $G_\lambda(0) \to G(0)$ as $\lambda \to \infty$.

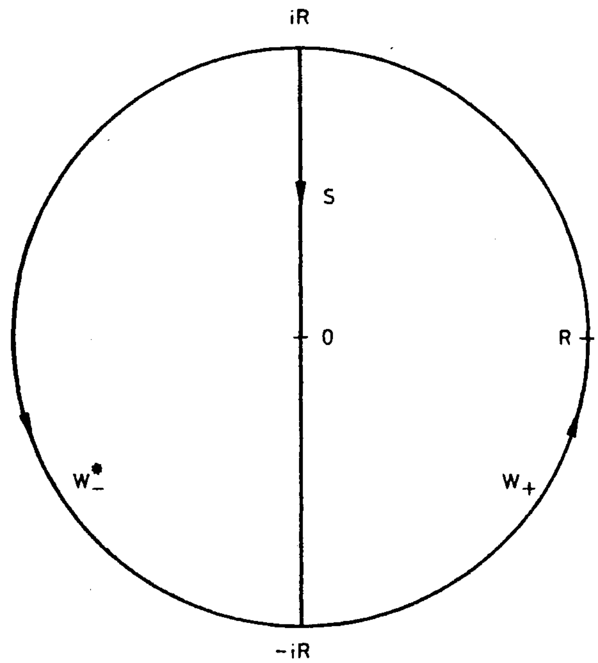


Figure 4

2.2. Relaxing the Conditions on $G(z)$

In the above proof, it is not really necessary to take $G(z)$ into the left half-plane. By modifying $F(t)$ on some finite interval one may assume that $G(0) = 0$. Then $G(z)/z$ will be analytic for Re $z \geqslant 0$ and thus

$$G(0) = 0 = \frac{1}{2\pi i} \int_{W_+ \cup S} G(z)e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right)dz,$$

where S is the segment $[iR, -iR]$ of the imaginary axis (Fig. 4). For $G_\lambda(0)$ we integrate over the circle $|z| = R$. Subtracting, one obtains

$$G(0) - G_\lambda(0) = \frac{1}{2\pi i} \int_{W_+} \{G(z) - G_\lambda(z)\}e^{\lambda z}\left(\frac{1}{z} + \frac{z}{R^2}\right)dz$$
$$+ \frac{1}{2\pi i} \int_{iR}^{-iR} G(z)e^{\lambda z} \ldots dz - \frac{1}{2\pi i} \int_{W_-^*} G_\lambda(z) \ldots dz. \tag{2.10}$$

The first and third integral are just as before. To deal with the second integral one may apply integration by parts or the Riemann-Lebesgue lemma.

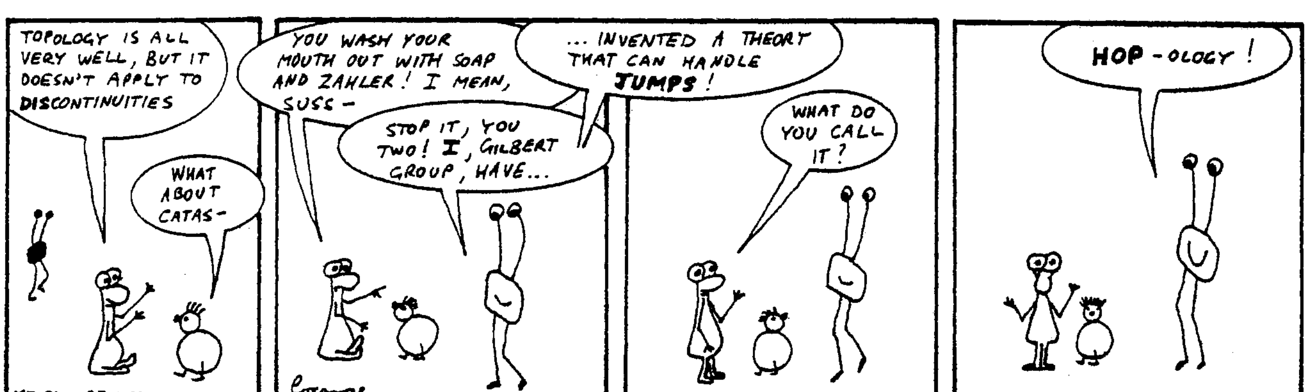
In order to arrive at (2.10), we have not used any analyticity of $G(z)$ at points of the imaginary axis. It would be more than enough to know that $G(z)/z$ can be extended continuously to Re $z \geqslant 0$. The Riemann-Lebesgue lemma will then handle the second integral.

Conclusion. In the auxiliary Tauberian theorem, it is sufficient to require that $\{G(z) - G(0)\}/z$ can be extended continuously to the closed half-plane $\operatorname{Re} z \geqslant 0$.

References

1. K. Chandrasekharan: Introduction to analytic number theory. Springer, Berlin, 1968
2. G. Doetsch: Theorie und Anwendung der Laplace-Transformation. Springer, Berlin, 1937
3. H. M. Edwards: Riemann's zeta function. Academic Press, New York, 1974
4. P. Erdös: On a new method in elementary number theory. Proc. Nat. Acad. Sci. U. S. A. *35* (1949) 374–384
5. G. H. Hardy, E. M. Wright: An introduction to the theory of numbers. Oxford Univ. Press, London, 1954
6. M. Heins: Complex function theory. Academic Press, New York, 1968
7. A. E. Ingham: The distribution of prime numbers. Cambridge Univ. Press, 1932; reprinted by Hafner, New York 1971
8. A. E. Ingham: On Wiener's method in Tauberian theorems. Proc. London Math. Soc. (2) *38* (1935) 458–480
9. E. Landau: Handbuch der Lehre von der Verteilung der Primzahlen (2 volumes). Teubner, Leipzig, 1909; reprinted by Chelsea, New York, 1953
10. D. J. Newman: Simple analytic proof of the prime number theorem. American Math. Monthly *87* (1980) 693–696
11. A. Selberg: An elementary proof of the prime number theorem. Ann. of Math. (2) *50* (1949) 305–313
12. E. C. Titchmarsh: The theory of functions. Oxford Univ. Press, London, 1939
13. E. C. Titchmarsh: The theory of the Riemann zeta-function. Oxford Univ. Press, London, 1951
14. N. Wiener: Tauberian theorems. Ann. of Math. *33* (1932) 1–100

J. Korevaar
Mathematisch Instituut
Universiteit van Amsterdam
Roeterstraat 15
Amsterdam



A DISPROOF OF A CONJECTURE OF PÓLYA

C. B. Haselgrove

Let $\lambda(n)$ be Liouville's function defined by

$$\lambda(n) = (-1)^{\nu},$$

where ν is the number of prime factors of n, repeated factors being counted according to their multiplicity. Alternatively, $\lambda(n)$ may be defined by the relation

$$\zeta(2s)/\zeta(s) = \sum_{n=1}^{\infty} \lambda(n)\, n^{-s},$$

where $\zeta(s)$ is the zeta function of Riemann.

Pólya [5] conjectured that the sum

$$L(x) = \sum_{n \leqslant x} \lambda(n)$$

is negative or zero for all $x \geqslant 2$. The author has verified that this conjecture is true for $x \leqslant 250{,}000$ (Royal Society Depository for Unpublished Mathematical Tables, No. 65)†.

A number of results have been deduced on the assumption of the truth of Pólya's conjecture, in particular that all the complex zeros of $\zeta(s)$ lie on the line $s = \frac{1}{2} + it$, with t real (the Riemann hypothesis), and that all these zeros are simple. These and other results are described by Ingham [2] in a paper in which he proves a further consequence, namely that the imaginary parts of some of the zeros of the zeta function above the real axis are linearly dependent (with rational integral multipliers).

In the course of his paper Ingham, assuming the Riemann hypothesis and the simplicity of the zeros, defines two functions

$$A(u) = e^{-\frac{1}{2}u}\, L(e^u)$$

and

$$A_T^*(u) = \alpha_0 + 2\Re \sum_{0<\gamma_n<T} \left(1 - \frac{\gamma_n}{T}\right) \alpha_n\, e^{i\gamma_n u},$$

where $n = 1, 2, \ldots,$ $T > 0$, $\alpha_0 = 1/\zeta(\frac{1}{2})$, γ_n runs through the imaginary parts of the zeros $\rho_n = \frac{1}{2} + i\gamma_n$ of $\zeta(s)$, and $\alpha_n = \zeta(2\rho_n)/\rho_n\, \zeta'(\rho_n)$. He then proves that

$$\varliminf A(u) \leqslant \varliminf A_T^*(u) \leqslant \varlimsup A_T^*(u) \leqslant \varlimsup A(u),$$

where the limits are taken as $u \to \infty$ with T fixed.

This suggests a method for investigating the truth of Pólya's conjecture. It follows from Dirichlet's theorem on Diophantine approximation, or from the fact that the function $A_T^*(u)$ is almost periodic, that $\varliminf A_T^*(u) \leqslant A_T^*(u) \leqslant \varlimsup A_T^*(u)$. Thus if we can find T, u such that

† D. H. Lehmer has verified the conjecture for $x \leqslant 600{,}000$ (private communication).

[Mathematika 5 (1958), 141–145]

$A_T^*(u) > 0$, it will follow that $\overline{\lim}\, A_T^*(u) > 0$ and hence that $L(e^u) > 0$ for some u, *i.e.* that Pólya's conjecture is false. Such values of T and u have in fact been found. Since the failure of the Riemann hypothesis would in any case imply the falsehood of Pólya's conjecture, it is immaterial that his argument rests on the assumption of the Riemann hypothesis.

Now that electronic computers are available it is possible to calculate $\zeta(s)$ over large ranges with considerable accuracy. Methods of calculating have been described by Lehmer [**3**], and by Haselgrove and Miller [**1**], who give tables of $\zeta(s)$ computed on the EDSAC I at the University Mathematical Laboratory, Cambridge, and the Mark I computer at Manchester University. In the course of computing the tables the first 1500 numbers γ_n were evaluated with an error of at most 3×10^{-8}. These values were used to find $\zeta(2\rho_n)$ and $\zeta'(\rho_n)$ and hence $\zeta(2\rho_n)/\rho_n \zeta'(\rho_n)$. Table I gives $\left|\dfrac{\zeta(2\rho_n)}{\rho_n \zeta'(\rho_n)}\right|$ and $\dfrac{1}{\pi}\,\mathrm{ph}\left(\dfrac{\zeta(2\rho_n)}{\rho_n \zeta'(\rho_n)}\right)$ for the first 50 zeros above the real axis. These quantities should be accurate to within 1 in the last decimal given. Here $\mathrm{ph}(z)$ denotes the argument (or phase) of the complex number z.

In order to find a value of u such that $A_T^*(u) > 0$ we observe that the largest contributions come from the first, second and seventh zeros. We therefore select for trial such positive or negative values of u as make the contributions of these three zeros positive and as large as possible. It is desirable to find a value of u which is not too large so that errors in γ_n do not cause large errors in $\gamma_n u$. (On the other hand smaller values of u require larger values of T which increases the amount of computation necessary.) I am indebted to Mr. J. Leech for finding the set of values

$$u = 28{\cdot}148 + l \times 139{\cdot}5794 + m \times 33{\cdot}7836 + n \times 17{\cdot}3363,$$

where l, m and n are integers satisfying $|l| \leqslant 50$, $|m| \leqslant 3$, $|n| \leqslant 1$. It is found that for $u = 831{\cdot}847$ (corresponding nearly to $l = 6$, $m = -1$ and $n = 0$) and $T = 1000$ the sum $A_T^*(u)$ is positive. This value of u is not claimed to be the smallest for which this occurs. For $T = 1000$ the sum involves 649 zeros. In Table II $A_T^*(u)$ is given for $T = 1000$ and $u = 831{\cdot}8\ (0{\cdot}001)\ 831{\cdot}859$.

This result would lead us to suspect that $L(e^u)$ becomes positive in the neighbourhood of $u = 831{\cdot}847$, although there is no proof that this is where the change of sign occurs. Some idea of the similarity of behaviour of $A_T^*(u)$ and $e^{-\frac{1}{2}u} L(e^u)$ is given by Fig. 1, for the range $10{\cdot}70 \leqslant u \leqslant 10{\cdot}85$ with $T = 200$. It is well known that $L(x)$ attains the value of -2 at $x = e^u = 48{,}512$.

The numbers γ_n and $|\zeta'(\rho_n)|$ were calculated and checked on the EDSAC I. The numbers $|\alpha_n|$ and $(1/\pi)\,\mathrm{ph}\,\alpha_n$ were calculated on both the EDSAC I and the Mark I; slight discrepancies were attributable to rounding errors in the EDSAC I results [these errors were known to occur only in

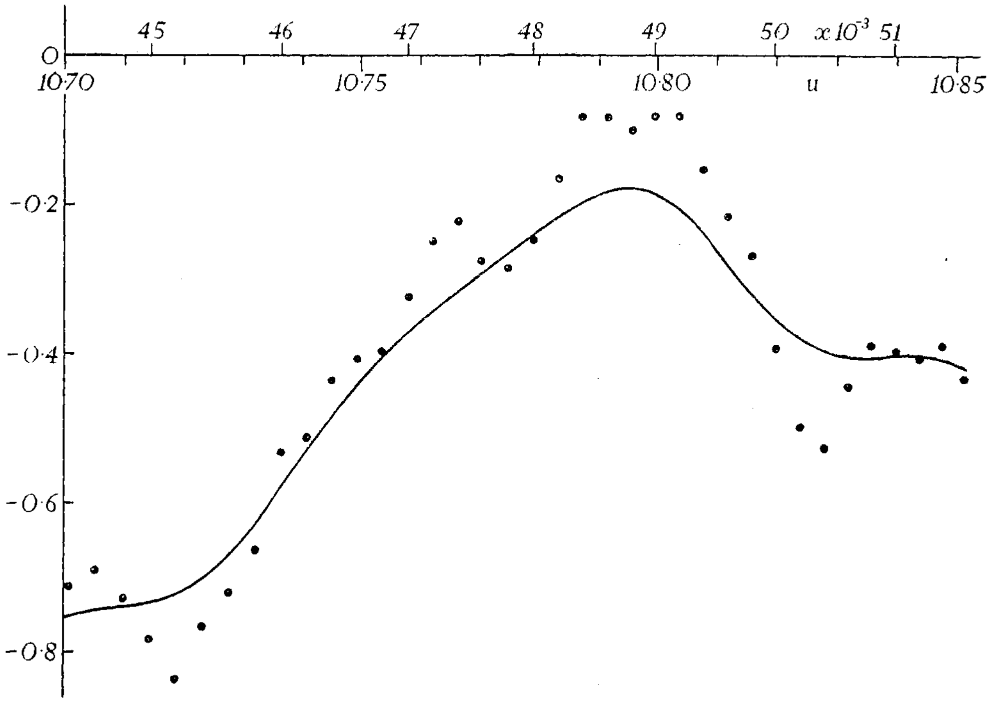


Fig. 1. The continuous line represents $A_T(u)$ with $T = 200$ and the points are values of $e^{-\frac{1}{2}u}L(e^u) = x^{-\frac{1}{2}}L(x)$ for $x = 44{,}400\ (200)\ 51{,}600$.

the calculation of $\zeta(2\rho_n)$]. The function $A_T^*(u)$ was also calculated on both machines for $u = 831{\cdot}845\ (0{\cdot}001)\ 831{\cdot}859$ with

$$T = 10{,}000\pi/32 = 981{\cdot}487\ldots.$$

The agreement of the results was within the limits set by the same rounding errors. We are thus led to the conclusion that Pólya's conjecture is false.

The methods described above may be applied to several similar problems. In particular we mention Mertens' hypothesis (Mertens [4], Ingham [2]) that

$$|M(x)| < x^{\frac{1}{2}},$$

where $M(x) = \sum_{n \leqslant x} \mu(n)$ and $\mu(n)$ is the function of Möbius, and Turán's suggestion [6] that

$$\sum_{n \leqslant x} \frac{\lambda(n)}{n} \geqslant 0 \quad (x \geqslant 1).$$

The functions $A_T^*(u)$ for these two problems would be

$$A_T^*(u) = 2\Re \sum_{0<\gamma_n<T} \left(1 - \frac{\gamma_n}{T}\right) \frac{1}{\rho_n\,\zeta'(\rho_n)}\, e^{i\gamma_n u}$$

and $$A_T^*(u) = -\frac{1}{\zeta(\frac{1}{2})} + 2\Re \sum_{0<\gamma_n<T} \left(1 - \frac{\gamma_n}{T}\right) \frac{\zeta(2\rho_n)}{(\rho_n - 1)\,\zeta'(\rho_n)}\, e^{i\gamma_n u}$$

respectively. In order to disprove Mertens' hypothesis it would be sufficient to find T, u such that $|A_T^*(u)| > 1$. In the case of Turán's sum it would be sufficient to find T, u for which the corresponding $A_T^*(u) < 0$.

TABLE I

$\alpha_0 = -0\cdot68476524$

n	γ_n	$\lvert\alpha_n\rvert$	$\frac{1}{\pi}$ ph α_n
1	14·13472513	0·17371523	−0·6474507
2	21·02203961	0·03476036	−0·3551160
3	25·01085756	0·01556708	−0·4134004
4	30·42487610	0·01297375	−0·0833945
5	32·93506159	0·01785754	−0·7548282
6	37·58617814	0·01288878	−0·7272660
7	40·91871901	0·03150963	−0·6028458
8	43·32707326	0·01230750	−0·8272898
9	48·00515087	0·00720805	0·0184893
10	49·77383246	0·02021980	−0·5793702
11	52·97032147	0·00535633	−0·4192135
12	56·44624768	0·00755698	−0·2768936
13	59·34704400	0·01237921	−0·4192967
14	60·83177851	0·00599826	−0·9329956
15	65·11254403	0·00477803	−0·2803717
16	67·07981051	0·00581876	−0·4098320
17	69·54640170	0·00478356	−0·7596008
18	72·06715766	0·00974345	−0·4975657
19	75·70469068	0·00548818	0·0765724
20	77·14484005	0·02312381	−0·5730905
21	79·33737500	0·00252000	−0·6244279
22	82·91038084	0·00387834	−0·3756371
23	84·73549297	0·00235891	−0·3687321
24	87·42527461	0·00384655	−0·4377069
25	88·80911120	0·00525947	−0·8231806
26	92·49189925	0·00138235	−0·4700115
27	94·65134403	0·00848399	−0·2819908
28	95·87063423	0·00570993	−1·0702284
29	98·83119420	0·00198488	−0·3237511
30	101·31785098	0·00179165	−0·4550882
31	103·72553802	0·00634339	−0·4263908
32	105·44662306	0·00704923	−0·7109605
33	107·16861114	0·00167392	−0·7086325
34	111·02953568	0·00541237	0·1826984
35	111·87465895	0·00526590	−0·7576648
36	114·32022105	0·00474366	−0·6894735
37	116·22668017	0·00179214	−0·3943627
38	118·79078298	0·00155757	−0·5521109
39	121·37012475	0·00260999	−0·2066128
40	122·94682964	0·01392294	−0·6669428
41	124·25681830	0·00197417	−0·8638090
42	127·51668401	0·00197639	−0·5442688
43	129·57870397	0·00195627	−0·3872602
44	131·08768873	0·00863392	−0·4815026
45	133·49773700	0·00180997	−0·3578010
46	134·75650989	0·00311216	−0·8288625
47	138·11604194	0·00104793	−0·2091791
48	139·73620909	0·00426514	−0·0393277
49	141·12370727	0·00307670	−1·0484089
50	143·11184585	0·00167694	−0·7713534

TABLE II

$T = 1000$

u	$A^*_T(u)$	u	$A^*_T(u)$	u	$A^*_T(u)$
831·800	−0·43329	831·820	−0·30119	831·840	−0·13583
·801	−0·42140	·821	−0·27534	·841	−0·12063
·802	−0·41040	·822	−0·25347	·842	−0·09590
·803	−0·40181	·823	−0·23640	·843	−0·06610
·804	−0·39640	·824	−0·22333	·844	−0·03705
·805	−0·39382	·825	−0·21269	·845	−0·01395
·806	−0·39287	·826	−0·20305	·846	0·00014
·807	−0·39220	·827	−0·19370	·847	0·00495
·808	−0·39097	·828	−0·18445	·848	0·00265
·809	−0·38918	·829	−0·17512	·849	−0·00328
831·810	−0·38762	831·830	−0·16518	831·850	−0·00950
·811	−0·38723	·831	−0·15397	·851	−0·01404
·812	−0·38853	·832	−0·14152	·852	−0·01693
·813	−0·39107	·833	−0·12920	·853	−0·01981
·814	−0·39325	·834	−0·11960	·854	−0·02493
·815	−0·39265	·835	−0·11547	·855	−0·03390
·816	−0·38674	·836	−0·11807	·856	−0·04698
·817	−0·37380	·837	−0·12600	·857	−1·56321
·818	−0·35378	·838	−0·13514	·858	−0·08124
·819	−0·32850	·839	−0·13999	·859	−0·10024

Note added September, 1958. Since submitting this paper the author has demonstrated the failure of Turán's inequality. The corresponding function $A^*_T(u)$ was shown to be negative, for $T = 1000$, at $u = 853{\cdot}853$ and $u = 996{\cdot}980$. He has not been able to disprove Mertens' conjecture, but it may well be disproved when much faster machines now being planned (about 1000 times faster than the EDSAC I and the Mark I) are available.

References.

1. C. B. Haselgrove and J. C. P. Miller, *Tables of the Riemann zeta function* (Royal Society Mathematical Tables, Vol. 6) (in the press).
2. A. E. Ingham, " On two conjectures in the theory of numbers ", *American J. of Math.*, 64 (1942), 313–319.
3. D. H. Lehmer, " Extended computation of the Riemann zeta function ", *Mathematika*, 3 (1956), 102–108.
4. F. Mertens, " Uber eine zahlentheoretische Funktion ", *Sitzungsberichte Akad. Wien.*, 106, Abt. 2a (1897), 761–830.
5. G. Pólya, " Verschiedene Bemerkungen zur Zahlentheorie ", *Jahresbericht der deutschen Math.-Vereinigung*, 28 (1919), 31–40.
6. P. Turán, " On some approximative Dirichlet-polynomials in the theory of the zeta-function of Riemann ", *Danske Vid. Selsk. Mat.-Fys. Medd.*, 24, No. 17 (1948).

The University,
Manchester 13.

(*Received 29th August*, 1958.)

THE PAIR CORRELATION OF ZEROS OF THE ZETA FUNCTION

H. L. MONTGOMERY

1. **Statement of results.** We assume the Riemann Hypothesis (RH) throughout this paper; $\varrho = \frac{1}{2} + i\gamma$ denotes a nontrivial zero of the Riemann zeta function. Our object is to investigate the distribution of the differences $\gamma - \gamma'$ between the zeros. It would thus be desirable to know the Fourier transform of the distribution function of the numbers $\gamma - \gamma'$; with this in mind we take

$$F(\alpha) = F(\alpha, T) = \left(\frac{T}{2\pi} \log T\right)^{-1} \sum_{0 < \gamma \leqq T;\, 0 < \gamma' \leqq T} T^{i\alpha(\gamma - \gamma')} w(\gamma - \gamma'), \tag{1}$$

where α and $T \geqq 2$ are real. Here $w(u)$ is a suitable weighting function, $w(u) = 4/(4+u^2)$, so $w(0) = 1$. Our results concerning $F(\alpha)$ are stated in the following

THEOREM. *(Assume RH.) For real α, $T \geqq 2$, let $F(\alpha)$ be defined by* (1). *Then $F(\alpha)$ is real, and $F(\alpha) = F(-\alpha)$. If $T > T_0(\varepsilon)$ then $F(\alpha) \geqq -\varepsilon$ for all α. For fixed α satisfying $0 \leqq \alpha < 1$ we have*

$$F(\alpha) = (1 + o(1))\, T^{-2\alpha} \log T + \alpha + o(1) \tag{2}$$

as T tends to infinity; this holds uniformly for $0 \leqq \alpha \leqq 1 - \varepsilon$.

The first term on the right-hand side of the above behaves in the limit as a Dirac δ-function; it reflects the fact that if $\alpha = 0$ then all the terms in (1) are positive. With more effort we could show that (2) holds uniformly throughout $0 \leqq \alpha \leqq 1$.

To investigate sums involving $\gamma - \gamma'$ we have only to convolve $F(\alpha)$ with an

AMS 1970 *subject classifications.* Primary 10H05.

appropriate kernel $\hat{r}(\alpha)$; from (1) alone it is immediate that

$$(3) \qquad \sum_{0<\gamma\leqq T;\, 0<\gamma'\leqq T} r\left((\gamma-\gamma')\frac{\log T}{2\pi}\right) w(\gamma-\gamma') = \left(\frac{T}{2\pi}\log T\right)\int_{-\infty}^{+\infty} F(\alpha)\,\hat{r}(\alpha)\,d\alpha .$$

Here $\hat{r}$ is the Fourier transform of r,

$$(4) \qquad \hat{r}(\alpha) = \int_{-\infty}^{+\infty} r(u)\, e(-\alpha u)\, du \qquad (e(\theta) = e^{2\pi i\theta}).$$

Our theorem gives us little information about $F(\alpha)$ for $\alpha \geqq 1$, so for the most part we restrict our attention to kernels $\hat{r}$ which vanish outside $[-1+\delta, 1-\delta]$. Particular choices of $\hat{r}(\alpha)$ give us

COROLLARY 1. *(Assume RH.) If* $0<\alpha<1$ *is fixed then*

$$(5) \qquad \sum_{0<\gamma\leqq T;\, 0<\gamma'\leqq T} \left(\frac{\sin\alpha(\gamma-\gamma')\log T}{\alpha(\gamma-\gamma')\log T}\right) w(\gamma-\gamma') \sim \left(\frac{1}{2\alpha}+\frac{\alpha}{2}\right)\frac{T}{2\pi}\log T,$$

and

$$(6) \qquad \sum_{0<\gamma\leqq T;\, 0<\gamma'\leqq T} \left(\frac{\sin(\alpha/2)(\gamma-\gamma')\log T}{(\alpha/2)(\gamma-\gamma')\log T}\right)^2 w(\gamma-\gamma') \sim \left(\frac{1}{\alpha}+\frac{\alpha}{3}\right)\frac{T}{2\pi}\log T.$$

In the latter assertion one can delete the factor $w(\gamma-\gamma')$ if one wishes. We use (6) to derive

COROLLARY 2. *(Assume RH.) As T tends to infinity*

$$(7) \qquad \sum_{0<\gamma\leqq T;\, \varrho \text{ simple}} 1 \geqq \left(\tfrac{2}{3}+o(1)\right)\frac{T}{2\pi}\log T.$$

The number of zeros of $\zeta(s)$ with $0<\gamma\leqq T$ is $\sim(T/2\pi)\log T$, so the above asserts that at least $\frac{2}{3}$ of the zeros are simple. It is known (see [**6**]) that the first 3,500,000 zeros are simple and lie on the critical line $\sigma=\frac{1}{2}$. Although one expects that all the zeros of $\zeta(s)$ are simple, the only other result in this direction is due to A. Selberg [**7**]. His result holds unconditionally; it states that a positive density of the zeros of $\zeta(s)$ are of odd order and lie on the critical line.

Let $0<\gamma_1\leqq\gamma_2\leqq\ldots$ denote the imaginary parts of the zeros of $\zeta(s)$ in the upper half-plane. The average of $\gamma_{n+1}-\gamma_n$ is $2\pi/\log\gamma_n$; our Theorem enables us to show that $\gamma_{n+1}-\gamma_n$ is not always near its average.

COROLLARY 3. *(Assume RH.) We can compute a constant λ so that*

$$\liminf_n (\gamma_{n+1}-\gamma_n)(\log\gamma_n/2\pi) \leqq \lambda < 1. \tag{8}$$

A complicated argument would permit one to show that in fact $\gamma_{n+1}-\gamma_n \leqq 2\pi\lambda/\log\gamma_n$ for a positive density of n. This, with the fact that the average value is $2\pi/\log\gamma_n$, enables one to assert that

$$\limsup_n (\gamma_{n+1}-\gamma_n)(\log\gamma_n/2\pi) \geqq \lambda' > 1. \tag{9}$$

We note that if $\zeta(s)$ has infinitely many multiple zeros then we may take $\lambda=0$ in (8). Our proof allows us to take $\lambda=0.68$. It would be of interest to have $\lambda<\frac{1}{4}$, as P. J. Weinberger and I have established the following: Let $d>0$ be square-free, and put $K=\boldsymbol{Q}((-d)^{1/2})$. Let $h(-d)$ be the class number of K, and let $\zeta_K(s)=\zeta(s)\cdot L(s,\chi)$ be the Dedekind zeta function of K. For each positive A, ε there is an effectively computable constant $d_0=d_0(A,\varepsilon)$ such that if $h(-d)\leqq A$, $d>d_0$, then all zeros of $\zeta_K(s)$ which are in the rectangle $0<\sigma<1$, $0\leqq t\leqq d^{1/2-\varepsilon}$ lie on the line $\sigma=\frac{1}{2}$; if $\frac{1}{2}+i\gamma_n$, $\frac{1}{2}+i\gamma_{n+1}$ are consecutive zeros of $\zeta_K(s)$ in this range then

$$(1-\varepsilon)\frac{2\pi}{\log d(\gamma_n+2)^2} \leqq \gamma_{n+1}-\gamma_n \leqq (1+\varepsilon)\frac{2\pi}{\log d(\gamma_n+2)^2}. \tag{10}$$

One may inquire about the behaviour of $F(\alpha)$ for $\alpha\geqq 1$. Our first observation is that (2) cannot hold uniformly for $0\leqq\alpha\leqq C$ if C is large. For if it did then (6) would hold for $0<\alpha\leqq C$. Write (6) as $G(\alpha)\sim H(\alpha)$. On one hand $|\sin 2x|\leqq 2|\sin x|$, so $G(2\alpha)\leqq G(\alpha)$ for all α. On the other hand $H(2\alpha)>\frac{3}{2}H(\alpha)$ for $\alpha\geqq 2$. This suggests that $F(\alpha)$ makes some change in its behaviour for $\alpha\geqq 1$. Further considerations of the above sort lead one to believe that certain averages of $F(\alpha)$ over large α are close to 1. At the end of § 3 we describe two heuristic arguments which suggest that

$$F(\alpha)=1+o(1) \tag{11}$$

for $\alpha\geqq 1$, uniformly in bounded intervals. This, with the Theorem, completely determines F, so an appropriate use of (3) leads immediately to a

CONJECTURE. *For fixed $\alpha<\beta$,*

$$(12)\qquad \sum_{\substack{0<\gamma\leqq T\\ 0<\gamma'\leqq T\\ 2\pi\alpha/\log T\leqq\gamma-\gamma'\leqq 2\pi\beta/\log T}} 1 \sim \left(\int_\alpha^\beta 1-\left(\frac{\sin\pi u}{\pi u}\right)^2 du + \delta(\alpha,\beta)\right)\frac{T}{2\pi}\log T$$

as T tends to infinity. Here $\delta(\alpha,\beta)=1$ if $0\in[\alpha,\beta]$, $\delta(\alpha,\beta)=0$ otherwise.

The Dirac δ-function occurs naturally in the above, for if $0\in[\alpha,\beta]$ then the sum includes terms $\gamma=\gamma'$.

The assertions (11) and (12) are essentially equivalent. From either it immediately follows that almost all zeros are simple. From (11) it is easy to see how Corollary 1 ought to be extended: If (11) is true then for $\alpha\geqq 1$,

$$(13)\qquad \sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T}\left(\frac{\sin\alpha(\gamma-\gamma')\log T}{\alpha(\gamma-\gamma')\log T}\right)w(\gamma-\gamma')\sim\frac{T}{2\pi}\log T,$$

and

$$(14)\qquad \sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T}\left(\frac{\sin(\alpha/2)(\gamma-\gamma')\log T}{(\alpha/2)(\gamma-\gamma')\log T}\right)^2 w(\gamma-\gamma')\sim\left(1+\frac{1}{3\alpha^2}\right)\frac{T}{2\pi}\log T.$$

In a certain standard terminology the Conjecture may be formulated as the assertion that $1-((\sin\pi u)/\pi u)^2$ is the pair correlation function of the zeros of the zeta function. F. J. Dyson has drawn my attention to the fact that the eigenvalues of a random complex Hermitian or unitary matrix of large order have precisely the same pair correlation function (see [**3**, equations (6.13), (9.61)]). This means that the Conjecture fits well with the view that there is a linear operator (not yet discovered) whose eigenvalues characterize the zeros of the zeta function. The eigenvalues of a random real symmetric matrix of large order have a different pair correlation, and the eigenvalues of a random symplectic matrix of large order have yet another pair correlation. In fact the "form factors" $F_r(\alpha)$, $F_s(\alpha)$ of these latter pair correlations are nonlinear for $0<\alpha<1$, so our Theorem enables us to distinguish the behaviour of the zeros of $\zeta(s)$ from the eigenvalues of such matrices. Hence, if there is a linear operator whose eigenvalues characterize the zeros of the zeta function, we might expect that it is complex Hermitian or unitary.

One might extend the present work to investigate the k-tuple correlation of the zeros of the zeta function. If the analogy with random complex Hermitian matrices appears to continue, then one might conjecture that the k-tuple correlation function $\hat{F}(u_1, u_2, \ldots, u_k)$ is given by

$$(15)\qquad \hat{F}(u_1, u_2,\ldots, u_k)=\det A,$$

where $A=[a_{ij}]$ is the $k\times k$ matrix with entries $a_{ii}=1$, $a_{ij}=(\sin\pi(u_i-u_j))/\pi(u_i-u_j)$ for $i\neq j$. Here the normalization is the same as in the Conjecture, which is the case $k=2$ of the above.

If one continues to draw on the analogy with random complex Hermitian matrices then one may formulate a conjecture concerning the distribution of the numbers $\gamma_{n+1}-\gamma_n$. The precise conjecture involves a complicated (but calculable) spheroidal function. Thus, or otherwise, one may conjecture that

$$\liminf_n(\gamma_{n+1}-\gamma_n)\log\gamma_n=0, \tag{16}$$

and

$$\limsup_n(\gamma_{n+1}-\gamma_n)\log\gamma_n=+\infty; \tag{17}$$

so Corollary 3 is probably far from the truth.

It would be interesting to see how numerical evidence compares with the above conjectures. The first several thousand zeros have been computed, so it would not be difficult to assemble relevant statistics. However, data on the failures of "Gram's law" indicate that the asymptotic behaviour is approached very slowly. Thus the numerical evidence may not be particularly illuminating.

2. **An explicit formula.** In proving our Theorem we require the following formula, which relates zeros of $\zeta(s)$ to prime numbers.

LEMMA. *If $1<\sigma<2$ and $x\geqq 1$ then*

$$\begin{aligned}(2\sigma-1)\sum_\gamma\frac{x^{i\gamma}}{(\sigma-\frac{1}{2})^2+(t-\gamma)^2}&=-x^{-1/2}\left(\sum_{n\leqq x}\Lambda(n)\left(\frac{x}{n}\right)^{1-\sigma+it}+\sum_{n>x}\Lambda(n)\left(\frac{x}{n}\right)^{\sigma+it}\right)\\&\quad+x^{1/2-\sigma+it}(\log\tau+O_\sigma(1))+O_\sigma(x^{1/2}\tau^{-1}),\end{aligned}\tag{18}$$

where $\tau=|t|+2$. The implicit constants depend only on σ.

PROOF. It is well known (see [**2**, p. 353]) that if $x>1$, $x\neq p^n$, then

$$\sum_{n\leqq x}\Lambda(n)\,n^{-s}=-\frac{\zeta'}{\zeta}(s)+\frac{x^{1-s}}{1-s}-\sum_\varrho\frac{x^{\varrho-s}}{\varrho-s}+\sum_{n=1}^\infty\frac{x^{-2n-s}}{2n+s}$$

provided $s\neq 1$, $s\neq\varrho$, $s\neq -2n$. This does not depend on RH, but if we assume RH then the above may be expressed as

$$\sum_{\varrho} \frac{x^{i\gamma - it}}{\sigma - \frac{1}{2} + it - i\gamma} = x^{\sigma - 1/2} \left(\frac{\zeta'}{\zeta}(s) + \sum_{n \leqq x} \Lambda(n)\, n^{-s} - \frac{x^{1-s}}{1-s} - \sum_{n=1}^{\infty} \frac{x^{-2n-s}}{2n+s} \right). \tag{19}$$

If we replace s by $1-\sigma+it$ in the above then we have

$$\begin{aligned} \sum_{\varrho} \frac{x^{i\gamma - it}}{\frac{1}{2} - \sigma + it - i\gamma} = x^{1/2 - \sigma} \Bigg(& \frac{\zeta'}{\zeta}(1-\sigma+it) + \sum_{n \leqq x} \Lambda(n)\, n^{\sigma - 1 - it} \\ & - \frac{x^{\sigma - it}}{\sigma - it} - \sum_{n=1}^{\infty} \frac{x^{-2n-1+\sigma-it}}{2n+1-\sigma+it} \Bigg). \end{aligned} \tag{20}$$

We subtract respective sides of (20) from (19), and use the relation

$$\frac{\zeta'}{\zeta}(s) = -\sum_{n=1}^{\infty} \Lambda(n)\, n^{-s}, \tag{21}$$

which holds for $\sigma > 1$. We find that

$$\begin{aligned} (2\sigma - 1) \sum_{\varrho} \frac{x^{i\gamma}}{(\sigma - \frac{1}{2})^2 + (t-\gamma)^2} = & -x^{-1/2} \left(\sum_{n \leqq x} \Lambda(n) \left(\frac{x}{n}\right)^{1-\sigma+it} + \sum_{n > x} \Lambda(n) \left(\frac{x}{n}\right)^{\sigma+it} \right) \\ & - \frac{\zeta'}{\zeta}(1-\sigma+it)\, x^{1/2-\sigma+it} + \frac{x^{1/2}(2\sigma-1)}{(\sigma-1+it)(\sigma-it)} \\ & - x^{-1/2} \sum_{n=1}^{\infty} \frac{(2\sigma-1)\, x^{-2n}}{(\sigma-1-it-2n)(\sigma+it+2n)}. \end{aligned} \tag{22}$$

Both sides of the above are continuous for all $x \geqq 1$, so we no longer exclude the values $x=1$, $x=p^n$. If $1<\sigma<2$, then from the logarithmic derivative of the functional equation of the zeta function (see [**1**, pp. 75, 82–83]) we have

$$\frac{\zeta'}{\zeta}(1-\sigma+it) = -\frac{\zeta'}{\zeta}(\sigma - it) - \log\tau + O_\sigma(1);$$

from (21) we see that this is $= -\log\tau + O_\sigma(1)$. Hence the right-hand side of (22) is

$$\begin{aligned} = -x^{-1/2} & \left(\sum_{n \leqq x} \Lambda(n) \left(\frac{x}{n}\right)^{1-\sigma+it} + \sum_{n > x} \Lambda(n) \left(\frac{x}{n}\right)^{\sigma+it} \right) \\ & + x^{1/2-\sigma+it}(\log\tau + O_\sigma(1)) + O_\sigma(x^{1/2}\tau^{-2}) + O_\sigma(x^{-2}\tau^{-1}), \end{aligned}$$

which gives the result.

3. **Proof of the Theorem.** The first assertion of the Theorem follows from the observation that we may interchange γ and γ' in (1). To prove the remaining assertions, take $\sigma = \frac{3}{2}$ in the Lemma, and write (18) briefly as $L(x, t) = R(x, t)$. We evaluate the integrals $\int_0^T |L(x, t)|^2\, dt$, $\int_0^T |R(x, t)|^2\, dt$.

We treat the left-hand side first. We have

$$(23)\qquad \int_0^T |L(x, t)|^2\, dt = 4 \sum_{\gamma, \gamma'} x^{i(\gamma - \gamma')} \int_0^T \frac{dt}{(1+(t-\gamma)^2)\,(1+(t-\gamma')^2)}.$$

We wish to exclude those numbers $\gamma \notin [0, T]$. It suffices to show that

$$(24)\qquad \sum_{\gamma, \gamma';\, \gamma \notin [0, T]} \int_0^T \frac{dt}{(1+(t-\gamma)^2)\,(1+(t-\gamma')^2)} \ll \log^3 T,$$

for then (23) is

$$(25)\qquad = 4 \sum_{0<\gamma \leqq T;\, 0<\gamma' \leqq T} x^{i(\gamma-\gamma')} \int_0^T \frac{dt}{(1+(t-\gamma)^2)\,(1+(t-\gamma')^2)} + O(\log^3 T).$$

To prove (24) we use the fact (Theorem 9.2 of [**8**]) that if $T \geqq 2$ then there are $\ll \log T$ zeros for which $T \leqq \gamma \leqq T+1$. From this it is immediate that if $0 \leqq t \leqq T$ then

$$\sum_{\gamma;\, \gamma \notin [0, T]} \frac{1}{1+(t-\gamma)^2} \ll \left(\frac{1}{t+1} + \frac{1}{T-t+1}\right) \log T,$$

and

$$\sum_{\gamma'} \frac{1}{1+(t-\gamma')^2} \ll \log T.$$

On the left-hand side of (24) we take the sums inside and use the above estimates. The integration is then trivial, and we obtain (24).

Arguing similarly we may also show that

$$\sum_{0<\gamma \leqq T;\, 0<\gamma' \leqq T} \int_T^\infty \frac{dt}{(1+(t-\gamma)^2)\,(1+(t-\gamma')^2)} \ll \log^2 T \int_T^\infty \frac{dt}{(t-T+1)^2} \ll \log^2 T.$$

The estimation of $\sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T}\int_{-\infty}^{0}\dots$ is the same, so we see that (25) is

$$=4\sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T} x^{i(\gamma-\gamma')}\int_{-\infty}^{+\infty}\frac{dt}{(1+(t-\gamma)^2)\,(1+(t-\gamma')^2)}+O(\log^3 T).$$

From the calculus of residues we deduce that the definite integral is $=(\pi/2)\,w(\gamma-\gamma')$, so the above is

$$=2\pi\sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T} x^{i(\gamma-\gamma')}w(\gamma-\gamma')+O(\log^3 T).$$

If we put $x=T^\alpha$ then we have

$$\int_0^T |L(T^\alpha, t)|^2\,dt = F(\alpha)\,T\log T+O(\log^3 T). \tag{26}$$

Here the left-hand side is clearly nonnegative, so we have the second assertion of the Theorem.

To complete the proof of the Theorem we prove (2); to this end we evaluate $\int_0^T |R(x,t)|^2\,dt$. In the first place

$$\int_0^T |x^{-1+it}\log\tau|^2\,dt = \frac{T}{x^2}(\log^2 T+O(\log T)) \tag{27}$$

for all $x\geqq 1$, $T\geqq 2$. To compute the mean square of the Dirichlet series on the right-hand side of (18) we use the following quantitative form (see [5]) of Parseval's identity for Dirichlet series:

$$\int_0^T \Big|\sum_n a_n n^{-it}\Big|^2\,dt = \sum_n |a_n|^2(T+O(n)). \tag{28}$$

We could instead use the weaker relation

$$\int_0^T \Big|\sum_{n\leqq N} a_n n^{-it}\Big|^2\,dt = (T+O(N))\sum_{n\leqq N}|a_n|^2\,;$$

this is Theorem 1.6 of [4]. However, the latter is restricted to Dirichlet polynomials, so we simplify our treatment by arguing from (28). We have

$$\frac{1}{x}\int_0^T \left| \sum_{n\leqq x} \Lambda(n)\left(\frac{x}{n}\right)^{-1/2+it} + \sum_{n>x} \Lambda(n)\left(\frac{x}{n}\right)^{3/2+it} \right|^2 dt$$
$$= \frac{1}{x}\sum_{n\leqq x} \Lambda(n)^2\left(\frac{x}{n}\right)^{-1}(T+O(n)) + \frac{1}{x}\sum_{n>x} \Lambda(n)^2\left(\frac{x}{n}\right)^{3}(T+O(n)).$$

By the prime number theorem with error term this is

$$= T(\log x + O(1)) + O(x\log x). \tag{29}$$

As for the error terms in (18), we see that

$$\int_0^T |x^{-1+it}|^2\, dt = \frac{T}{x^2}, \tag{30}$$

and

$$\int_0^T x\tau^{-2}\, dt \ll x. \tag{31}$$

We now combine our estimates (27), (29), (30), (31); we employ the following consequence of the Cauchy-Schwarz inequality: If $M_k = \int_0^T |A_k(t)|^2\, dt$ and $M_1 \geqq M_2 \geqq M_3 \geqq M_4$, then

$$\int_0^T \left| \sum_{k=1}^{4} A_k(t) \right|^2 dt = M_1 + O((M_1 M_2)^{1/2}).$$

We consider three cases.

Case 1. $1 \leqq x \leqq (\log T)^{3/4}$. Then our M_1 term is given by (27). Our other terms are uniformly $o(M_1)$, so our expression is $=(1+o(1))\,(T/x^2)\log^2 T$.

Case 2. $(\log T)^{3/4} < x \leqq (\log T)^{3/2}$. In this case all our estimates are uniformly $o(T\log T)$.

Case 3. $(\log T)^{3/2} < x \leqq T/\log T$. Then our M_1 term is given by (29). All our other terms are uniformly $o(M_1)$, so our expression is $=(1+o(1))\,T\log x$.

If we put $x=T^\alpha$ then we may express our result by saying that

$$\int_0^T |R(T^\alpha, t)|^2\,dt = ((1+o(1))\,T^{-2\alpha}\log T+\alpha+o(1))\,T\log T,$$

uniformly for $0\leqq\alpha\leqq 1-\varepsilon$. This and (26) give (2), so the proof is complete.

If $\alpha>1$ in the above then $x>T$, so the second error term in (29) is no longer smaller than the main term. The error term (31) also gives problems; a little consideration reveals that what we require is to know the size of

$$\int_0^T \left|\frac{1}{x}\sum_{n\leqq x}\Lambda(n)\,n^{1/2-it}+x\sum_{n>x}\Lambda(n)\,n^{-3/2-it}-\frac{2x^{1/2-it}}{(\frac{1}{2}+it)(\frac{3}{2}-it)}\right|^2 dt. \tag{32}$$

If we multiply out the integrand and integrate terms individually, we find that there are too many nondiagonal terms to be ignored. We may, however, collect terms so that the above is expressed in terms of sums of the sort $\sum_{n\leqq y}\Lambda(n)\,\Lambda(n+h)$. There are various indications that this sum is approximately $c(h)\,y$, where $c(h)$ is a certain arithmetic constant. If we replace these sums by their conjectured approximations $c(h)\,y$, then our new expression is $\sim T\log T$. Moreover, there is a reasonable hypothesis as to the size of the differences

$$\sum_{n\leqq y}\Lambda(n)\,\Lambda(n+h)-c(h)\,y \tag{33}$$

which if true would allow us to carry out our program for $1\leqq\alpha<2$. If the differences (33) are not only reasonably small but also behave independently for different h then (32) is $\sim T\log T$ for all $\alpha\geqq 1$.

Another indication of the behaviour of the expression (32) can be obtained by considering its "q-analogue." The expression

$$\sum_{q\leqq Q}\frac{1}{\varphi(q)}\sum_{\chi\neq\chi_0}\left|\frac{1}{x}\sum_{n\leqq x}\Lambda(n)\,\chi(n)\,n^{1/2}+\sum_{n>x}\Lambda(n)\,\chi(n)\,n^{-3/2}\right|^2 \tag{34}$$

may be shown to be $\sim Q\log x$ for $Q\geqq x$, in analogy with (29). If $x(\log x)^{-A}\leqq Q\leqq x$ then we may use an established technique [**4**, Chapter 17] to show that (34) is $\sim Q\log Q$. If GRH is true then this latter asymptotic relationship holds for $x^{3/4+\varepsilon}<Q\leqq x$. This corresponds to $1\leqq\alpha<\frac{4}{3}$. One does not expect a change in the behaviour for larger α, but a more delicate error-term analysis is needed if the result is to be extended.

4. **The corollaries.** To prove Corollary 1 we use our Theorem in conjunction with (3). To obtain (5) we take $r(u)=(\sin 2\pi\alpha u)/2\pi\alpha u$. The Theorem makes it a simple task to compute

$$\int_{-\infty}^{+\infty} F(\beta)\,\hat{r}(\beta)\,d\beta=\frac{1}{2\alpha}\int_{-\alpha}^{\alpha} F(\beta)\,d\beta.$$

To obtain (6) we take $r(u)=((\sin \pi\alpha u)/\pi\alpha u)^2$. Again from the Theorem it is easy to compute

$$\int_{-\infty}^{+\infty} F(\beta)\,\hat{r}(\beta)\,d\beta=\frac{1}{\alpha^2}\int_{-\alpha}^{+\alpha} (\alpha-\beta)\,F(\beta)\,d\beta.$$

We now prove Corollary 2. Let m_ϱ be the multiplicity of the zero ϱ. In a sum over $0<\gamma\leqq T$, our convention concerning multiple zeros is that zeros are counted according to their multiplicities. This is accomplished by allowing γ to take on the same value m_ϱ times. In particular,

$$\sum_{0<\gamma\leqq T} m_\varrho=\sum_{\substack{0<\gamma\leqq T\\ 0<\gamma'\leqq T\\ \gamma=\gamma'}} 1$$

for on both sides a given zero ϱ is counted with weight m_ϱ^2. But

$$\sum_{\substack{0<\gamma\leqq T\\ 0<\gamma'\leqq T\\ \gamma=\gamma'}} 1\leqq \sum_{0<\gamma\leqq T;\,0<\gamma'\leqq T}\left(\frac{\sin(\alpha/2)\,(\gamma-\gamma')\log T}{(\alpha/2)\,(\gamma-\gamma')\log T}\right)^2 w(\gamma-\gamma'),$$

and if we take $\alpha=1-\delta$ then from (6) the above is

$$\leqq(\tfrac{4}{3}+\varepsilon)\,(T/2\pi)\log T.$$

Hence we have demonstrated that

$$\sum_{0<\gamma\leqq T} m_\varrho\leqq(\tfrac{4}{3}+o(1))\,(T/2\pi)\log T.$$

Now

$$\sum_{0<\gamma\leqq T;\,\varrho \text{ simple}} 1 \geqq \sum_{0<\gamma\leqq T} (2-m_\varrho) \geqq (2-\tfrac{4}{3}+o(1)) \frac{T}{2\pi}\log T,$$

so we have Corollary 2. The kernel $\hat{r}(u)$ which we have used does not appear to be optimal for our purpose, so presumably one can improve slightly on the constant $\frac{2}{3}$.

We now turn to the first assertion of Corollary 3. We take $r(u) = \max(1-(|u|/\lambda), 0)$ in (3), and choose λ later. Now $\hat{r}(\alpha)$ is nonnegative, and $\int_0^\infty \hat{r}(\alpha)\, d\alpha < \infty$, so our Theorem permits us to calculate a lower bound for the right-hand side of (3). We see that

$$\int_{-\infty}^{+\infty} F(\alpha)\,\hat{r}(\alpha)\,d\alpha \geqq (1+o(1))\left(\lambda + 2\lambda\int_0^1 \alpha\left(\frac{\sin \pi\lambda\alpha}{\pi\lambda\alpha}\right) d\alpha\right)\frac{T}{2\pi}\log T.$$

We may assume that all but finitely many zeros are simple, so the terms $\gamma=\gamma'$ in (3) contribute an amount $\sim (T/2\pi)\log T$. Hence

$$\sum_{\substack{0<\gamma\leqq T\\ 0<\gamma'\leqq T\\ 0<\gamma-\gamma'<2\pi\lambda/\log T}} 1 \geqq (\tfrac{1}{2}+o(1))\, C(\lambda)\frac{T}{2\pi}\log T$$

where

$$C(\lambda) = \lambda + (1/\pi^2\lambda)\operatorname{Cin}(2\pi\lambda) - 1.$$

Here $\operatorname{Cin}(x)$ is the "cosine integral,"

$$\operatorname{Cin} x = \int_0^x \frac{1-\cos u}{u}\, du.$$

Note that the integrand is nonnegative, so that $\operatorname{Cin}(x) > 0$ for $x>0$. To obtain (8) we show that $C(\lambda)>0$ for some $\lambda<1$. This is easy, because $C(1) = (1/\pi^2) \cdot \operatorname{Cin}(2\pi) > 0$, and $C(\lambda)$ is continuous. In fact a little calculation reveals that $C(0.68) > 0$. We have not determined the optimal kernel $\hat{r}(\alpha)$, so one should be able to improve on the constant 0.68.

References

1. H. Davenport, *Multiplicative number theory*, Lectures in Advanced Math., no. 1, Markham, Chicago, Ill., 1967. MR **36** #117.

2. Edmund Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Berlin, 1909.

3. M. L. Mehta, *Random matrices and the statistical theory of energy levels*, Academic Press, New York, 1967. MR **36** #3554.

4. Hugh L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math., vol. 227, Springer-Verlag, Berlin and New York, 1971.

5. Hugh L. Montgomery and R. C. Vaughan, *Hilbert's inequality* (to appear).

6. J. B. Rosser, J. M. Yohe and L. Schoenfeld, *Rigorous computation and the zeros of the Riemann zeta-function* (*with discussion*), Information Processing 68 (Proc. IFIP Congress, Edinburgh, 1968), vol. 1, Math., Software, North-Holland, Amsterdam, 1969, pp. 70–76. MR **41** #2892.

7. Atle Selberg, *On the zeros of Riemann's zeta-function*, Skr. Norske Vid. Akad. Oslo I (1942), no. 10, 59 pp. MR **6**, 58.

8. E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Clarendon Press, Oxford, 1951. MR **13**, 741.

Trinity College
Cambridge, England