NNTDM 2 (1996) 3, 33-41

THE TWIN PRIME PROBLEM, ACCORDING TO HARDY-LITTLEWOOD

by Aldo Peretti

Abstract:

The ideas suggested by Hardy and Littlewood in their paper (1) concerning the subject of the title, are explained at full length. A further contribution is made by the introduction of Poisson's integral formula in the solution of the problem.

1. In pages 41 and 42 of their celebrated paper on P. N. III (ref. (1)), Hardy and Littlewood indicate, in a too condensed form, how could be proved the existence of infinite twin primes. Furthermore, they conjecture a formula for $P_k(n)$, the pairs of primes p_1, $p_1 + k = p_2$, less than x, expressed in their conjecture B, given below.

The purpose of this paper is to develop in full detail the central ideas of these two pages, as dense as fruitful.

2. Obtention of a generating function for the twin primes.

Consider the function

$$[1] \qquad f(x) = \sum_{p} \log p . \; x^p$$

where p denotes the prime numbers. It is well known that $f(x)$ has a natural boundary on $|x| = 1$. Next, we choose

$$[2] \qquad x = R \, e^{i\Psi}$$

where, of course $0 < R < 1$.

Afterwards, we form

$$[3] \qquad J(R) = \frac{1}{2\pi} \int_{0}^{2\pi} f(R\, e^{i\Psi}) \; f(R\, e^{-i\Psi}) \; (R\, e^{i\Psi})^k \; d\Psi$$

where k denotes a fixed natural number. We have then

[4] $$J(R) = \frac{1}{2\pi} \int_0^{2\pi} \{\Sigma \log p_1 \; R^{p_1} \; e^{p_1 i\Psi}\} \; \{\Sigma \log p_2 \; R^{p_2} \; e^{-p_2 i\Psi}\} \; R^k \; e^{ik\Psi} \; d\Psi$$

Inside the unit circle the series are uniformly convergent; their product is equally well uniformly convergent, and so term by term integration is permissible, over a finite interval, even after the multiplication by $e^{ik\Psi}$.

We obtain thus:

[5] $$J(R) = \frac{1}{2\pi} \sum_{p_1} \sum_{p_2} \log p_1 \; \log p_2 \; R^{p_1+p_2+k} \int_0^{2\pi} e^{i\Psi(p_1-p_2+k)} \; d\Psi$$

Now, it is an obvious fact that

[6] $$\int_0^{2\pi} e^{i\Psi(p_1-p_2+k)} \; d\Psi = \begin{cases} 0 \text{ if } p_1 - p_2 + k \neq 0 \\ 2\pi \text{ if } p_1 - p_2 + k = 0 \end{cases}$$

i.e.:

[7] $$\int_0^{2\pi} e^{i\Psi(p_1-p_2+k)} \; d\Psi = \begin{cases} 2\pi \text{ if } p_1 \text{ and } p_2 \text{ are twin primes of difference } k \\ 0 \text{ in any other case} \end{cases}$$

Hence [5] reduces to:

[8] $$J(R) = \Sigma \; \Sigma \; \log p_1 \; \log p_2 \; R^{p_1+p_2+k}$$

where now the double sum is no longer extended to all the primes, but only to the twin primes of difference k.

Evidently, formula [8] can be regarded as the generating function for the logarithms of the twin primes. (This does not exclude, of course, the fact that the double sum could be a finite series).

In order to draw some conclusion about the order of magnitude of $J(R)$ as R approximates to the unit circle, we need to find some alternative way for evaluating $J(R)$, and this is what we perform in the following paragraph, through an alternative representation of $f(x)$.

3. The Farey dissection of $f(x)$.

The formula

$$[9] \qquad f(z) = \sum_p \log p.\ z^p = \sum_{q=1}^{[\sqrt{n}]} \sum_{\substack{h=o \\ (h,q)=1}}^{q} \frac{\mu(q)}{\varphi(q)\left(1 - \frac{z}{\rho}\right)} + A\, n^{\vartheta + 1/4 + \varepsilon}$$

is to be found in ref. (2).

Here $\mu(q)$ is the Moebius function; $\varphi(q)$ is Euler's function: quantity of naturals $\leq q$ and prime with q. ρ denotes a primitive root of unity: $\rho = e^{2\pi i h/q}$, with $(h,q) = 1$ and ϑ is the upper bound of the real part of the zeros of the L-series involved in the problem.

The approximation is valid as far as $z = R\, e^{i\Psi}$, with $R = e^{-1/n}$

$$[10] \qquad x = e^{-y} \qquad y = \frac{1}{n} - \vartheta i \qquad \vartheta_1 < \vartheta \leq \vartheta_2 \qquad \frac{\pi}{2\sqrt{n}} < \vartheta_1, \vartheta_2 < \frac{2\pi}{\sqrt{n}}$$

Now, $J(R)$ is nothing but

$$[11] \qquad J(R) = \frac{1}{2\pi} \int_o^{2\pi} f(z)\, f(\bar{z})\, z^{k-1}\, dz$$

($\bar{z}$ = conjugate of z), or, what is the same:

$$[12] \qquad J(R) = \frac{1}{2\pi i} \int_C f(z)\, f(\bar{z})\, z^{k-1}\, dz$$

where C can be any circle of radius $R < 1$ surrounding the origin.

4. We form now:

$$[13] \qquad f(z)\, f(\bar{z}) = \sum_{q_1} \sum_{q_2} \sum_{h_1} \sum_{h_2} \frac{\mu(q_1)}{\varphi(q_1)} \frac{\mu(q_2)}{\varphi(q_2)} \frac{1}{\left(1 - \frac{z}{\rho_1}\right)\left(1 - \frac{\bar{z}}{\rho_2}\right)} +$$

$$+ A\, n^{\vartheta+1/4+\varepsilon} \sum_{q} \sum_{h} \frac{\mu(q)}{\varphi(q)\left(1-\frac{z}{\rho_1}\right)} + A^2\, n^{2\vartheta+1/2+\varepsilon} +$$

$$+ A\, n^{\vartheta+1/4+\varepsilon} \sum_{q} \sum_{h} \frac{\mu(q)}{\varphi(q)\left(1-\frac{\bar{z}}{\rho_1}\right)}$$

The arcs along which we can perform the integration of [13] are those limited by the inequalities [10], that represent only a small part of the whole circle C . But here Hardy and Littlewood make explicitly the following hypothesis:

<u>Hypothesis</u>

When we perform the integration of [13] along C , we obtain an asymptotic value for J (R). That is to say that they assume

$$[14] \qquad J(R) \sim \frac{1}{2\pi i} \int_{C} \sum_{q_1} \frac{\mu(q_1)}{\varphi(q_1)} \sum_{q_2} \frac{\mu(q_2)}{\varphi(q_2)} \sum_{h_1} \sum_{h_2} \frac{z^{k-1}\, dz}{\left(1-\frac{z}{\rho_1}\right)\left(1-\frac{\bar{z}}{\rho_2}\right)}$$

$$+ \frac{A}{2\pi i}\, n^{\vartheta+1/4+\varepsilon} \int_{C} \sum_{q} \sum_{h} \frac{\mu(q)\, z^{k-1}}{\varphi(q)\left(1-\frac{z}{\rho}\right)} + \frac{1}{2\pi i} \int_{C} A^2\, n^{2\vartheta+1/2+\varepsilon}\, z^{k-1}\, dz$$

$$+ \frac{A}{2\pi i}\, n^{\vartheta+1/4+\varepsilon} \int_{C} \sum_{q} \sum_{h} \frac{\mu(q)\, z^{k-1}}{\varphi(q)\left(1-\frac{\bar{z}}{\rho}\right)}$$

The third term at right vanishes, because it is the integral along C of an analytic function inside C . At the remaining terms of the right handside we can interchange the signs $\int$ and Σ because we are dealing with finite sums, so that

$$[15] \qquad J(R) = \sum_{q_1} \frac{\mu(q_1)}{\varphi(q_1)} \sum_{q_2} \frac{\mu(q_2)}{\varphi(q_2)} \sum_{h_1} \sum_{h_2} \frac{1}{2\pi i} \int_{C} \frac{z^{k-1}\, dz}{\left(1-\frac{z}{\rho_1}\right)\left(1-\frac{\bar{z}}{\rho_2}\right)}$$

$$+ \frac{A}{2\pi i} n^{\vartheta+1/4+\varepsilon} \left\{ \sum_q \sum_h \frac{\mu(q)}{\varphi(q)} \left(\int_c \frac{z^{k-1} \, dz}{1 - \frac{z}{\rho}} + \int_c \frac{z^{k-1} \, dz}{1 - \frac{\bar{z}}{\rho}} \right) \right\}$$

5. We must evaluate now:

$$[16] \qquad I_{1,2} = \frac{1}{2\pi i} \int_c \frac{z^{k-1} \, dz}{\left(1 - \frac{z}{\rho_1}\right)\left(1 - \frac{\bar{z}}{\rho_2}\right)} = \frac{\rho_1 \rho_2}{2\pi i} \int_c \frac{z^{k-1} \, dz}{(z - \rho_1)(\bar{z} - \rho_2)}$$

The integrand is an analytic function of z, with poles at $z = \rho_1$ and $\bar{z} = \rho_2$, i.e. at $z = \bar{\rho}_2 = \rho_2^{-1}$. These poles are exactly in the boundary $R = 1$, but as by hypothesis, $R < 1$, the integrand is analytic inside C. Hence, by Cauchy's integral theorem:

$$[17] \qquad I_{1,2} = 0 \qquad \text{whatever be the radius } R < 1$$

There is, however, a unique exception to the former statement, which occurs when $\rho_1 = \rho_2^{-1} = \rho$. Then we have:

$$[18] \qquad I_{1,2} = I^*_{1,2} = \frac{1}{2\pi i} \int_c \frac{z^{k-1} \, dz}{\left(1 - \frac{z}{\rho}\right)(1 - z\rho)}$$

Putting $z = R\, e^{i\alpha}$ $\rho = e^{2\pi i h/q}$, we obtain:

$$[19] \qquad I^*_{1,2} = \frac{R^k}{2\pi} \int_0^{2\pi} \frac{e^{ik\alpha} \, d\alpha}{1 - 2R \cos(\alpha - 2\pi h/q) + R^2}$$

This is nothing but a special case of Poisson's classical integral formula (ref. (3)):

$$[20] \qquad u(r_1, \beta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{r_2^2 - r_1^2}{r_2^2 - 2 r_1 r_2 \cos(\alpha - \beta) + r_1^2} \, u(r_2, \alpha) \, d\alpha$$

(where $0 \leq r_1 < r_2$), when $r_2 = 1$, $r_1 = R < 1$, $u(R, \alpha) = (R\, e^{i\alpha})^k$.

We deduce thus

[21] $$I^*_{1,2} = \frac{u(R_1\ 2\pi h/q)}{1 - R^2} = \frac{R^k}{1 - R^2}\, e^{2\pi ihk/q}$$

Summaryzing, we have:

[22] $$I_{1,2} = \begin{cases} 0 & \text{if } \rho_1 \neq \rho_2^{-1} \text{ whatever be R} \\ \dfrac{R^k}{1 - R^2}\, e^{2\pi ihk/q} & \text{if } \rho_1 = \rho_2^{-1} \end{cases}$$

6. We consider now

[23] $$I\rho = \frac{1}{2\pi i} \int_c \left\{ \frac{z^{k-1}\, dz}{\left(1 - \frac{z}{\rho}\right)} + \frac{z^{k-1}\, dz}{\left(1 - \frac{\bar{z}}{\rho}\right)} \right\}$$

The integrand has only two simple poles at $z = \rho$ and $\bar{z} = \rho$, that lie outside the contour C defined by $R < 1$; hence; by Cauchy's theorem.

[24] $$I\rho = 0$$

7. With the results [22] and [24] at hand, we return to [15] deducing that:

[25] $$J(R) \cong \sum_{q=1}^{[\sqrt{n}]} \frac{\mu^2(q)}{\varphi^2(q)} \sum_{\substack{h=1 \\ (h,q)=1}}^{q} e^{2\pi ihk/q}\, \frac{R^k}{1 - R^2}$$

Next, by definition, we have that

[26] $$\sum_{\substack{h \\ (h,q)=1}} e^{2\pi ihk/q} = C_q(k)$$

where $C_q(n)$ is Ramanujan's function. Hence

[27] $$\sum_{q=1}^{[\sqrt{n}]} \frac{\mu^2(q)}{\varphi^2(q)} C_q(k)$$

is the singular series of Hardy-Littlewood truncated at $q = [\sqrt{n}]$.
We denote it by $S_{\sqrt{n}}(k)$. It is bounded as $n \to \infty$, so that we can write:

[28] $$S_{\sqrt{n}}(k) = S_\infty(k) \left\{ 1 + \left(\frac{1}{n} \right) \right\}$$

or, otherwise expressed:

[29] $$S_{\sqrt{n}}(k) \sim S_\infty(k)$$

The multiplicative properties of the functions that appear in the singular series enables us to transform it into a product, the result being (ref. (1) p. 42):

[30] $$S_\infty.(k) = 2 \prod_{p=3}^{\infty} \left(1 - \frac{1}{(p-1)^2} \right) \prod_{p/k} \frac{p-1}{p-2}$$

(from here follows the result [29])
Hence, from [26] follows:

[31] $$J(R) \sim \frac{S_\infty(k)\ R^k}{1 - R^2}$$

But the series development [8] of $J(R)$ can be written as:

[32] $$J(R) = \sum \sum \log p_1 \ \log p_2 \ R^{2p_2}$$

due to the fact that $p_1 = p_2 - k$.
Hence, if we put $R^2 = r$, we can write [32] taking account of [31] as:

[33] $$J(r) = \sum \sum \log p_1 \ \log p_2 \ r^{p_2} \sim \frac{r^{k/2}}{1 - r} S_\infty(k) \sim \frac{S_\infty(k)}{1 - r}$$

The partial sum $s(n)$ of the coefficients of $J(r)$ up to $p_2 = n$ is:

[34] $$s(n) = \sum_{\text{twins} \leq n} \sum \log p_1 \ \log p_2$$

We apply now to the series development [33] of $J(r)$ the tauberian theorem of ref. (3) p. 225, that states that if

$$f(x) \sim \frac{C}{1-x}$$

as $x \to 1$, and the coefficients a_i in the power series expansion of $f(x)$ are positive, then $s(n) = a_1 + a_2 + \ldots + a_n \sim C.n.$

(This is a very particular case of a much more general theorem of Hardy and Littlewood in ref. (4)).

On account of [33] we obtain:

[35] $$\sum_{\text{twins} \leq n} \sum \log p_1 \ \log p_2 \sim n\, S_\infty(k)$$

that is justly hypothesis B of P. N. III.

This implies, as Hardy and Littlewood point out, that $P_k(n)$, the quantity of twin primes (with difference k) $< n$ would be

$$P_k(n) \sim \frac{n}{\log^2 n} S_\infty(k)$$

By the way, we can recall that

$$S_\infty(k) = 2 \prod_{p=3}^{\infty} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p/k} \frac{p-1}{p-2}$$

where the p are odd primes and

$$\prod_{p=3}^{\infty} \left(1 - \frac{1}{(p-1)^2}\right) = 1{,}3023\ldots$$

8. The question that presents itself in the preceding lines is if the hypothesis made in paragraph 4 above can be proved by some way.
My personal opinion is that following the model that appears in Landau's "Vorlesungen", V Teil, Kap. 6 Zweiter und Dritter Schritt formulas (224) - (230), p. 220 the hypothesis concerning the sums that appear in [13] could be ultimately proved.

* * * * * * * * * * *

REFERENCES

(1) G. Hardy and J. Littlewood: Some problems in "Partitio Numerorum" III - On the expression of a number as a sum of primes. Acta Math. (1922) p. 1-70.

(2) E. Landau - Vorlesungen Uber Zahlentheorie - Chelsea Edition. p. 218

(3) E. Titchmarsh. The Theory of Functions. Oxford Univ. Press. 2nd. edition. Chap. III, § 3.6, p. 124.

(4) G. Hardy and J. Littlewood: Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive. Proc. London Math. Soc. (2) 13 (1914) p. 174-91.

Aldo Peretti
Facultad de Ciencia y Tecnología - Universidad del Salvador
Rodríguez Peña 640
(1020) Buenos Aires, Argentina