ON PRIME NUMBERS

EUGENE EHRHART
University Louis Pasteur, Strasbourg, France

(Submitted October 1986)

I. RAREFACTION OF PRIMES IN THE SERIES OF INTEGERS

Let p be a prime number and $\pi(n)$ the number of primes up to n, inclusive. In the capricious succession of primes in the series of integers, the inequality

$$\pi(2n) < 2\pi(n) \qquad (n > 10) \tag{1}$$

shows a certain regularity. It is equivalent to the following proposition.

Theorem 1: The first n integers contain more primes than the n following, for n greater than 10.

I submitted this proposition as a conjecture to Professor G. Robin of the University of Limoges, with two remarks:

—It is true for $n < 10{,}000$, as I have verified it on the computer.

—The inequalities

$$\frac{n}{\log n}\left(1 + \frac{1}{2\log n}\right) < \pi(n) < \frac{n}{\log n}\left(1 + \frac{3}{2\log n}\right) \qquad (n \geqslant 52)$$

established by Rosser and Schoenfeld [1] are not sufficient to prove (1), for they give

$$2\pi(n) - \pi(2n) > \frac{2n}{\log n \log 2n}\left[\log 2 - \frac{3(\log n)^2 - (\log 2n)^2}{2\log n \log 2n}\right],$$

where the expression in brackets is negative.

Robin sent me the following ingenious demonstration of Theorem 1: Suppose that for $n \geqslant n_0$

$$\frac{n}{\log n}\left(1 + \frac{a}{\log n}\right) \leqslant \pi(n) \leqslant \frac{n}{\log n}\left(1 + \frac{b}{\log n}\right). \tag{2}$$

Then for $n \geqslant n_0$

$$\Delta_n = 2\pi(n) - \pi(2n) \geqslant \frac{2n}{\log n \log 2n}\left[\log 2 + \frac{a(\log 2n)^2 - b(\log n)^2}{\log 2n \log n}\right]$$

$$\geqslant \frac{2n}{\log n \log 2n}[\log 2 + (a - b)].$$

Therefore, if $b - a < \log 2$, $\Delta_n > 0$ for $n \geqslant n_0$.

We shall see that we can choose $a = 5/6$ for $n \geqslant 10{,}000$, and verify directly afterward that we can also do this for $n \geqslant 227$. We take $b = 3/2$. We write, as usual, $\theta(n) = \sum_{p \leqslant n} \log p$. Then (see [1], p. 359),

$$\theta(n) = n + R(n),$$

where $R(n) < \dfrac{an}{\log n}$ for $a = \dfrac{5}{6}$ and $n > 10{,}000$.

Lemma: We have

$$\pi(n) \geqslant \frac{n}{\log n} + \frac{5}{6}\frac{n}{(\log n)^2} \qquad (n \geqslant 227).$$

For $n > n_0 = 10{,}000$, with $\log_i(n) = \displaystyle\int_2^n \frac{dt}{\log t}$,

$$\pi(n) - \pi(n_0) = \int_{n_{\bar{0}}}^{n} \frac{d\theta(t)}{\log t}$$

$$= \log_i(n) - \log_i(n_0) + \frac{R(n)}{\log n} - \frac{R(n_{\bar{0}})}{\log n_0} + \int_{n_0}^{n} \frac{R(t)dt}{t(\log t)^2}$$

$$\geqslant \log_i(n) - \log_i(n_0) - \frac{an}{(\log n)^2} - a\int_{n_0}^{n} \frac{dt}{(\log t)^3}$$

since $R(n_{\bar{0}}) < 0$.

Let

$$f(n) = \log_i(n) - \frac{n}{\log n} - \frac{n}{(\log n)^2}.$$

Then, $f'(n) = 2/(\log t)^3$; hence,

$$f(n) - f(n_0) = 2\int_{n_0}^{n} \frac{dt}{(\log t)^3}$$

and

$$\pi(n) - \frac{n}{\log n} - \frac{n}{(\log n)^2} \geqslant \frac{-an}{(\log n)^2} + \left(1 - \frac{a}{2}\right)[f(n) - f(n_0)] + \pi(n_0) - \frac{n_0}{\log n_0} - \frac{n_0}{(\log n_0)^2}.$$

So, for $n_0 = 10{,}000$,

$$\pi(n_0) = 1229 > \frac{n_0}{\log n_0} + \frac{n_0}{(\log n_0)^2}.$$

Since $f(n) > f(n_0)$, we have

$$\pi(n) \geqslant \frac{n}{\log n} + \frac{5}{6}\frac{n}{(\log n)^2} \qquad (n \geqslant 10{,}000).$$

Remark: G. Robin adds: "In [2] the authors claim to have proved (1) and state their demonstration will be published at a later date. As far as I know, it has not yet appeared, but as we saw, the proof is an easy consequence of the results from Rosser and Schoenfeld."

Professor Robin also sent me a demonstration of a more general proposition, which I submitted to him again as a conjecture.

Theorem 2: For all integers k and n greater than 2,

$$\pi(kn) < k\pi(n). \tag{2}$$

Suppose (3) is verified for $n \geqslant n_0$. Then,

$$k\pi(n) - \pi(kn) \geqslant \frac{kn}{\log n \log kn}\left[\log k + \frac{a(\log kn)^2 - b(\log n)^2}{\log kn \log n}\right]$$

$$\geqslant \frac{kn}{\log n \log kn}[\log k + (a - b)].$$

We have seen that with $n_0 \geqslant 10{,}000$, we can take $a = 5/6$, $b = 3/2$. So

$$\log k + (a - b) > 0 \quad \text{for } k \geqslant 2.$$

For $n \geqslant 59$, we can choose $a = \frac{1}{2}$ and $b = \frac{3}{2}$; consequently, $\log k + (a - b) > 0$, for $k \geqslant 3$. We verify directly afterward that (3) holds for $k = 3$ and $n < 59$. For $n \geqslant 17$, we can choose $a = 0$ and $b = \frac{3}{3}$; consequently, $\log k + (a - b) > 0$, for $k \geqslant 5$.

The case $k = 4$ being treated as $k = 2$, it remains to study (3) for $n \leqslant 16$. We did this successively for

$$13 \leqslant n \leqslant 16,\ 11 \leqslant n \leqslant 12,\ 7 \leqslant n \leqslant 10,\ 5 \leqslant n \leqslant 6,\ 3 \leqslant n \leqslant 4,$$

using for $\pi(km)$ the majoration

$$\pi(n) < \frac{5}{4}\frac{n}{\log n} \quad (n \geqslant 2).$$

II. RAREFACTION OF TWINS IN THE SERIES OF PRIMES

Twins are two primes with a difference of 2.

Theorem 3: In the infinite series of great primes, twins are extremely rare.

Let p_n be the n^{th} prime number. The probability that 7, for instance, does not divide $p_n + 2$ is 5/6, for the equiprobable remainders in the division of p_n by 7 are 2, 3, 4, 5, and 6. The probability P_n that $p_n + 2$ will be a prime is also

$$P_n = \prod_{2<p<\sqrt{p_n}} \frac{p - 2}{p - 1}.$$

Therefore, if n tends to infinity, P_n tends to zero, like

$$\prod_{p<\sqrt{p_n}} \frac{p - 1}{p},$$

which is greater.

Remark: Our reasoning is not quite rigorous, because we utilize implicitly the independence of prime numbers. A conjecture in [3] substitutes an approached value of P_n,

$$\prod_{2<p<p_n\cdot 0,5615\ldots} \frac{p-2}{p-1} \quad \text{to} \quad \prod_{2<p<p_n\cdot 0,5} \frac{p-2}{p-1},$$

where $0,5615\ldots = e^{-\gamma}$, with Euler's constant $\gamma = 0,5772\ldots$. Mertens proved that

$$\log P_m \prod_{i \leqslant m} \frac{p_i - 1}{p_i}$$

tends to $e^{-\gamma}$, if m tends to infinity.

Yet our reasoning carries away, I think, any doubt that the probability P_n tends to zero. If you disagree, consider Theorems 3 and 4 as conjectures.

Theorem 4: The series of primes presents arbitrarily great intervals without twins.

Indeed, if every interval of k consecutive primes presented at least one pair of twins, the probability P_n would be at least $1/k$ and could not tend to zero.

Remarks (see [4]):

1. The table below gives an idea of the rarefaction of twins; the 150,000 first integers after 10^n present t pairs of twins:

n	8	9	10	11	12	13	14	15
t	584	461	314	309	259	211	191	166

2. One estimates empirically that the number of pairs of twins up to n has the order of

$$132032 \frac{n}{(\log n)^2}.$$

The estimate leads one to believe there exist an infinity of twin primes.

REFERENCES

1. L. Schoenfeld. "Shaper Bounds for the Chebychev Functions $\theta(x)$ and $\psi(x)$." *Mathematics of Computation* **30** (1976):337-360.
2. J. Rosser & L. Schoenfeld. "Approximate Formulas for Some Functions of Prime Numbers." *Illinois Journal of Mathematics* **6** (1962):64-94.
3. H. Riesel. *Prime Numbers and Computer Methods for Factorisation* (Chap. 3). Basel: Birkhäuser, 1935.
4. P. Davis & R. Hersch. *The Mathematical Experiences*, pp. 215-216. Boston, 1982.

◆◇◆◇◆