On the difference between consecutive prime numbers
by
S. UCHIYAMA (Okayama)

Yuri Vladimirovich Linnik in memoriam

Let $p_n$ denote the $n$th prime number, and define

$$E = \liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n}.$$

The crude estimate $E \leq 1$ follows, as is easily seen, from the fact that $p_n \sim n \log n$ $(n \to \infty)$, which is equivalent to the prime number theorem. The long-standing conjecture that states that $E = 0$, which is obviously the case if there exist infinitely many pairs of primes $p, q$ with a fixed non-zero difference, remains still unproved. The best result on the size of $E$ that is known so far is due to G. Z. Pil'jaf [2], who showed that

$$(1) \quad E \leq \frac{1}{2}(2\sqrt{2} - 1) = 0.457106...$$

improving a previous result of E. Bombieri and H. Davenport [1],

$$E \leq \frac{1}{2}(2 + \sqrt{3}) = 0.466506...$$

The purpose of the present article is to make a further improvement on these results. Indeed, we shall prove the following

Theorem. $(1)$ We have

$$(2) \quad E \leq \frac{9 - \sqrt{8}}{16} = 0.454246...$$

An inspection of Pil'jaf's paper [2] suggests a possibility of ameliorating the estimate (1) for $E$ by an alternative choice of the various parameters therewith concerned. Our proof of (2) is thus a slight modification

$(1)$ After the present paper had been submitted the writer learned from a kind letter of Prof. A. Schinzel that M. N. Huxley obtained, by improving Pil'jaf's argument, the inequality $E \leq (4 + n)/16 = 0.446349...$, which supersedes (2).
Let $k$, $h$ and $r$ be positive integers such that
\[ k = [aM], \quad h = [caM], \]
\[ r = k - h = (1 - c) aM + O(1), \]
where $a = 3(6 - \sqrt{3})/32$ and $c$ is any fixed number satisfying the condition
\[ 1 > c > c_0 = \frac{21 + 2\sqrt{3}}{33} > \frac{2}{3}. \]

Note that $c_0 a = (9 - \sqrt{3})/32$, and $x = (1 - c_0) a$ is the lesser root of the quadratic equation $128x^2 - 48x + 3 = 0$.

We now define the real numbers $c_j (0 < j < k)$ by setting
\[ c_j = \begin{cases} a & \text{for } 0 < j \leq r, \\ 1 & \text{for } r < j \leq h, \\ 1 - \beta(j - h) & \text{for } h < j \leq k, \end{cases} \]
where
\[ a = \frac{3(1 - 8(1 - c) a)}{3 - 256(1 - c)^2 a^2}, \quad \beta = \frac{16a}{M}. \]

It is readily verified that $a > 0$, and that $1 - \beta(j - h) \geq 0$ for $h < j \leq k$, provided $N$ is large enough.

Put
\[ T(x) = \left| \sum_{n=1}^{k} c_j e^{2\pi i x n} \right|^2 = t(0) + 2 \sum_{n=1}^{k} t(n) \cos 4\pi x n. \]

Then we have
\[ t(n) = \begin{cases} c_j c_{j+n} & (0 \leq n < k), \\ 0 & (k < n < 2k), \end{cases} \quad t(k) = 0. \]

It is not difficult to see that the coefficient $t(n)$ of $T(x)$, as a function of $n$, $0 \leq n < k$, can be represented in each of the intervals $0 \leq n < r$, $r \leq n < h - r$, $h - r \leq n < h$, $h \leq n < k$, as a sum of several monomials of degree at most 3 in $n$, and the absolute value of each of the monomials therein involved has for $0 \leq n < k$ the order of magnitude $O(M)$, the $O$-constant being, of course, numerical. We have, in particular,
\[ t(0) = a^2 r + h - \beta r(r+1) + \frac{1}{2} \beta r(r+1)(2r+1), \]
\[ 2 \sum_{n=1}^{k} t(n) = (ar + h - \frac{1}{2} \beta r(r+1))^2 - t(0) \]
and
\[ -\sum_{n=k}^{k/2} t(n) = \frac{1}{2} ar(r+1) - \frac{1}{2} \beta r(r+1)(2r+1), \]
whence we get after some simplifications

$$2 \sum_{n=1}^{h} t(n) = (A(\varepsilon))^2 M^2 + O(M)$$

and

$$32 \sum_{n=h}^{k} t(n) + t(0) M = A(\varepsilon) M^2 + O(M)$$

with

$$A(\varepsilon) = \frac{3a - 4(1 - e)^2 a^2 - 6(1 - e)^3 (7e - 3)a^2}{3 - 256(1 - e)^2 a^2}.$$ 

Also, we have

$$\sum_{n=1}^{k} t(n) H(n) = \sum_{n=1}^{k} t(n) + O(M(\log M)^2)$$

for $l = 1$ and $h$. This is an immediate consequence of the following lemma which has been proved by Rankin [3].

**Lemma 3.** If $s$ is a non-negative real number, then for arbitrary integers $a, b$ with $0 \leq a < b \leq M$ we have

$$\sum_{n=1}^{b} n^s H(n) = \sum_{n=1}^{b} n^s + O(M^2(\log M)^2),$$

where the $O$-constant may depend only on $s$.

3. Let the integer $N$ be sufficiently large. We define as in Pilt'jul [2]

$$W(N; 2n) = Z(N; 2n) - Z(N,M^{-2}; 2n),$$

where we have set $M = \log N$, as before.

It follows from Lemmas 1 and 2 that one has for any fixed $\varepsilon > 0$

$$\sum_{n=1}^{k} t(n) W(N; 2n) > 2N \sum_{n=1}^{k} t(n) H(n) - (1 + 2\varepsilon) t(0) N M,$$

since

$$\sum_{n=1}^{k} t(n) Z(N,M^{-2}; 2n) = O(N).$$

Assume now that $W(N; 2n) = 0$ for all integers $n$ lying in the interval $1 \leq n < h$. Then, by Lemma 2 we have for any fixed $\delta > 0$

$$\sum_{n=1}^{h} t(n) W(N; 2n) \leq \sum_{n=1}^{h} t(n) Z(N; 2n) < (8 + \delta) N \sum_{n=1}^{h} t(n) H(n),$$

and we arrive at the inequality

$$(8 + \delta) \sum_{n=1}^{h} t(n) H(n) + (1 + 2\varepsilon) t(0) M > 2 \sum_{n=1}^{h} t(n) H(n)$$

or

$$32 \sum_{n=h}^{k} t(n) + t(0) M > 8 \sum_{n=1}^{h} t(n) - 3M^2,$$

on appealing to the results obtained in \S 2.

It follows from this that

$$A(\varepsilon) > \frac{4}{3} (A(\varepsilon))^2 - 4\varepsilon$$

or

$$1 > 4A(\varepsilon) - \varepsilon,'$$

where $\varepsilon'$ is a positive real number that can be made as small as we please, if we let $N$ tend to infinity. However, this is obviously impossible, since we have taken $1 > \varepsilon > \varepsilon_0$ and this implies that

$$4A(\varepsilon) > 4A(\varepsilon_0) = 1;$$

in fact, it will suffice to observe that the quantity

$$(1 - \varepsilon)^2 (28\varepsilon_0 - 12\varepsilon - 1)$$

is, as a function of $\varepsilon$, monotonically decreasing in the interval

$$\frac{44 + \sqrt{3}}{63} < \varepsilon < 1,$$

where $\varepsilon_1 < \varepsilon_0$.

We thus have proved that, if $N$ is sufficiently large, then we must always have

$$W(N; 2n) > 0$$

for at least one $n$ with $1 \leq n < h = \lceil e \log N \rceil$, which implies that $E \leq 2e\varepsilon$. Since $\varepsilon$ may be taken arbitrarily close to $\varepsilon_0$, this concludes the proof of our assertion (2).

**References**


**DEPARTMENT OF MATHEMATICS**

**OKAYAMA UNIVERSITY**

**Okayama, Japan**

Received on 20.6.1973 (425)