TSUKUBA J. MATH.
Vol. 11 No. 2 (1987). 257—264

THE DIFFERENCES BETWEEN CONSECUTIVE ALMOST-PRIMES

By

Hiroshi MIKAWA

1. Introduction.

In 1940 P. Erdös [1] proposed the problem to estimate the sum

$$D(x)=\sum_{p_n\leqq x}(p_{n+1}-p_n)^2$$

where p_n denotes the n-th prime. A. Selberg [10] and D. R. Heath-Brown [4] proved that

$$D(x)\ll x(\log x)^3$$

under the Riemann hypothesis, and that, for any $\varepsilon>0$,

$$D(x)\ll x^{7/6+\varepsilon}$$

under the Lindelöf hypothesis, respectively. Furthermore, Heath-Brown [5] showed unconditionaly that, for any $\varepsilon>0$,

$$D(x)\ll x^{23/18+\varepsilon},$$

and he [6] conjectured that

$$D(x)\sim 2x(\log x)\quad \text{as } x\to\infty.$$

U. Meyer considered in his Dissertation the almost-prime analogy of $D(x)$. Let P_2 denote the set of integers with at most two prime factors, multiple factors being counted multiplicity. We replace the primes in $D(x)$ by the almost-primes P_2, and denote the resulting sum $D_2(x)$. In [8] he showed, by the weighted version of a zero density estimate for the Riemann zeta-function, that

$$D_2(x)\ll x^{1.285}(\log x)^{10}.$$

It is the purpose of this paper to make an improvement upon this upper bound.

THEOREM. *We have*

$$D_2(x)\ll x^{1.023}$$

where the implied constant is effectively computable.

In contrast to the Meyer's argument, we appeal to sieve methods, which are the weighted linear sieve of Greaves' type [3] and the prototype of an additive

Received May 15, 1986. Revised October 31, 1986.

large sieve inequality [9]. J. B. Friedlander [2] considered the related proplem from a different point of view. Our argument should be compared with [2], in which also sieve methods were employed.

We use the standard notation in number theory. Especially, for an integer a, $\Omega(a)$ and $\nu(a)$ denote the number of prime factors counted multiplicity and the number of different prime factors of a, respectively. All the implied constants are effectively computable.

I would like to thank Professor G. Greaves for sending me a copy of the preprint of his paper [3].

The present paper is a revised version of part of my Master thesis at Okayama University. I would like to thank Professor S. Uchiyama for encouraging me to publish the paper and careful reading the original manuscript. I would also thank the referee for making the paper easier to read.

2. Reduction of the problem.

In this section we deduce Theorem from Lemma 1 below. We postpone the proof of Lemma 1 until the final section. To simplify the notation, let p_n denote the n-th almost-prime P_2 and write $d_n=p_{n+1}-p_n$. Put $\theta=1.023$.

We will show that

$$\sum_{x<p_n\leqq 2x} d_n^2 \ll x^\theta \tag{1}$$

for all sufficiently large x. The assertion of Theorem immediately follows from this by the routine argument.

LEMMA 1. *We have unifomly for $x\leqq y\leqq 2x$, $(\log x)^3<\Delta\leqq x/2$,*

$$\sum_{\substack{y-\Delta<a\leqq y\\ \Omega(a)\leqq 2}} 1 > C\Delta(\log x)^{-1}+O(\Delta(\log x)^{-3})+E_1(y,\Delta)+E_2(y,\Delta),$$

where the $E_j(y,\Delta)$ $(j=1,2)$ are some quantities depending on y and Δ to be given explicitly in § 4 below and satisfying

$$\int_x^{2x}|E_j(y,\Delta)|^j dy \ll \Delta^{j-1}x^\theta(\log x)^{-j-1},\quad j=1,2. \tag{2}$$

Here the positive constant C is effectively computable.

Now, let

$$\Pi(\Delta)=\{p_n\in P_2\,;\,x<p_n\leqq 2x,\ 2\Delta<d_n\leqq 4\Delta,\ p_{n+1}\leqq 2x\}.$$

It is well known that $d_n\ll p_n^{1/2}$ for sufficiently large $p_n\in P_2$, so we may assume

(3) $$\Delta \ll x^{1/2}.$$

For any fixed $p_n \in \Pi(\Delta)$, suppose that

$$|E_j(y, \Delta)| \leqq \frac{C}{3} \frac{\Delta}{\log x}, \quad j=1, 2$$

for some $y \in [p_n + d_n/2, p_{n+1})$. Then the right-hand side of (1) is positive for sufficiently large x. But, since

$$y - \Delta > \frac{d_n}{2} + p_n - \frac{d_n}{2} = p_n,$$

the left-hand side of (1) is zero, which is impossible. Thus,

$$|E_1(y, \Delta)| > \frac{C}{3} \frac{\Delta}{\log x} \quad \text{or} \quad |E_2(y, \Delta)| > \frac{C}{3} \frac{\Delta}{\log x},$$

for all $y \in [p_n + d_n/2, p_{n+1})$. Namely,

(4) $$\Pi(\Delta) \subset \bigcup_{j=1}^{2} \Pi_j(\Delta)$$

where

$$\Pi_j(\Delta) = \{p_n \in \Pi(\Delta)\ ;\ |E_j(y, \Delta)| > \frac{C}{3} \frac{\Delta}{\log x} \text{ for all } y \in [p_n + \frac{d_n}{2}, p_{n+1})\}.$$

Now, we have

$$\begin{aligned}\int_x^{2x} |E_j(y, \Delta)|^j dy &\geqq \sum_{p_n \in \Pi_j(\Delta)} \int_{p_n + d_n/2}^{p_{n+1}} |E_j(y, \Delta)|^j dy \\ &> \sum_{p_n \in \Pi_j(\Delta)} \left(\frac{C}{3} \frac{\Delta}{\log x}\right)^j \frac{d_n}{2},\end{aligned}$$

since the intervals $[p_n + d_n/2, p_{n+1}]$ are mutually disjoint. Hence, by (2)

(5) $$\begin{aligned}\sum_{p_n \in \Pi_j(\Delta)} d_n &\ll \left(\frac{\Delta}{\log x}\right)^{-j} \int_x^{2x} |E_j(y, \Delta)|^j dy \\ &\ll \Delta^{-1} x^{\theta} (\log x)^{-1}.\end{aligned}$$

Since there is at most one element $p_n \in P_2$ such that

$$x < p_n \leqq 2x, \quad 2\Delta < d_n \leqq 4\Delta, \quad p_{n+1} > 2x,$$

we have uniformly for $(\log x)^3 < \Delta \leqq x^{1/2}$,

(6) $$\begin{aligned}\sum_{\substack{x < p_n \leqq 2x \\ 2\Delta < d_n \leqq 4\Delta}} d_n &\leqq \sum_{p_n \in \Pi(\Delta)} d_n + 4\Delta \\ &\leqq \sum_{j=1}^{2} \sum_{p_n \in \Pi_j(\Delta)} d_n + 4\Delta \\ &\ll \Delta^{-1} x^{\theta} (\log x)^{-1} + \Delta \\ &\ll \Delta^{-1} x^{\theta} (\log x)^{-1},\end{aligned}$$

by (4), (5) and (3).

Finally,

$$\sum_{x<p_n\leqq 2x} d_n^2 = \sum_{\substack{x<p_n\leqq 2x \\ d_n\leqq x^{\theta-1}}} d_n^2 + \sum_{\substack{x<p_n\leqq 2x \\ d_n>x^{\theta-1}}} d_n^2$$

$$\leqq x^{\theta-1} \sum_{x<p_n\leqq 2x} d_n + \sum_{x^{\theta-1}\ll \Delta\ll x^{1/2}} \sum_{\substack{x<p_n\leqq 2x \\ 2\Delta<d_n\leqq 4\Delta}} d_n^2$$

where Δ's run through the powers of 2. By (6) we get

$$\sum_{x<p_n\leqq 2x} d_n^2 \ll x^{\theta} + \sum_{\Delta} \Delta \sum_{\substack{x<p_n\leqq 2x \\ 2\Delta<d_n\leqq 4\Delta}} d_n$$

$$\ll x^{\theta} + \sum_{\Delta} x^{\theta}(\log x)^{-1}$$

$$\ll x^{\theta},$$

as required.

3. Lemmas.

Firstly we state the results of [3] merely in as simple a way as is sufficient for our application.

LEMMA 2. *If* $\Lambda=(\log 2x)(\log D)^{-1}<1.95544$, *then we have*

$$\sum_{\substack{y-\Delta<a\leqq y \\ \Omega(a)\leqq 2}} 1 > C(\Lambda)\,\Delta(\log x)^{-1} + \sum_{d<D} \lambda_d r_d(y, \Delta) - \sum_{\substack{y-\Delta<n\leqq y \\ p^2|n \\ D^v\leqq p<D^u}} 1,$$

uniformly for

$$x\leqq y\leqq 2x,\quad 2<\Delta\leqq x/2$$

where the positive constant $C(\Lambda)$ *is effective and depends on* Λ *only*,

$$|\lambda_d|\leqq \mu(d)^2 3^{\nu(d)},$$

$$r_d(y, \Delta)=\left[\frac{y}{d}\right]-\left[\frac{y-\Delta}{d}\right]-\frac{\Delta}{d},$$

and u *and* v *are the absolute constants such that* $0.01<v<u<1$.

LEMMA 3. *For any real* t, *and* $H>2$, *we have*

$$t-[t]-\frac{1}{2}=-\frac{1}{2\pi i}\sum_{0<|h|<H}\frac{1}{h}e(ht)+0\left(\min\left(1, \frac{1}{H||t||}\right)\right)$$

where

$$e(t)=e^{2\pi it},$$

and

$$||t||=\min_{n\in Z}|t-n|.$$

PROOF. See [7], for example.

LEMMA 4. *Let* $1 < A \ll 1$. *For any different real numbers* (b_n) *and any complex numbers* (c_n), *we have*

$$\int_x^{Ax} |\sum_n c_n e(b_n u)|^2 du \ll (x+\delta^{-1}) \sum_n |c_n|^2$$

where

$$\delta = \min_{m \neq n} |b_m - b_n|.$$

PROOF. This is the corollary 2 in [9].

LEMMA 5. *Let* $1 < A \ll 1$. *For any* $H > 2$, *we have*

$$\int_x^{Ax} \min\left(H, \frac{1}{||u||}\right) du \ll x(\log H).$$

PROOF.

$$\sum_{\substack{x-1<n\leqq Ax \\ n\in \boldsymbol{Z}}} \int_n^{n+1} \min\left(H, \frac{1}{||u||}\right) du \ll \sum_n \int_0^{1/2} \min\left(H, \frac{1}{u}\right) du$$

$$\ll x(\log H).$$

4. Proof of Lemma 1.

We begin with considering

$$\sum_{\substack{y-\Delta<n\leqq y \\ p^2|n \\ D^v\leqq p<D^u}} 1 = \sum_{\substack{y-\Delta<n\leqq y \\ p^2|n \\ D^v\leqq p<x^{1/3}}} 1 + \sum_{\substack{y-\Delta<n\leqq y \\ p^2|n \\ x^{1/3}\leqq p<D^u}} 1 = R_1 + R_2, \text{ say.}$$

By Lemma 3, we have

$$\begin{aligned} R_1 &= \sum_{D^v\leqq p<x^{1/3}} \sum_{(y-\Delta)/p^2<m\leqq y/p^2} 1 \\ &= \sum_{D^v\leqq p<x^{1/3}} \frac{\Delta}{p^2} + \sum_p r_{p2}(y, \Delta) \\ &= 0(\Delta(\log x)^{-3}) + \sum_p \sum_{0<|h|<p} \frac{1}{2\pi i h}\left\{1-e\left(-\frac{h}{p^2}\Delta\right)\right\} e\left(\frac{h}{p^2}y\right) + \\ &\qquad + \sum_p 0\left(\min\left((1, \frac{1}{p||y/p^2||}\right) + \min\left(1, \frac{1}{p||(y-\Delta)/p^2||}\right)\right) \\ &= 0(\Delta(\log x)^{-3}) + R_{12} + R_{11}, \text{ say.} \end{aligned}$$

We have then

$$\int_x^{2x} |R_{12}|^2 dy \ll (\log x) \max_{D^v\leqq P<x^{1/3}} \int_x^{2x} |\sum_{\substack{0<h<p \\ P<p\leqq 2P}} \frac{1}{h}\left\{1-e\left(-\frac{h}{p^2}\Delta\right)\right\} e\left(\frac{h}{p^2}y\right)|^2 dy$$

$$\ll (\log x) \max_P (x+P^4) \sum_{h,p} \frac{1}{h^2} \sin^2\Big(\frac{\pi h \Delta}{p^2}\Big)$$
$$\ll (\log x) \max_P (x+P^4) \frac{\Delta}{P}$$
$$\ll \Delta x (\log x).$$

by Lemma 4. Moreover we have, by Lemma 5,

$$\int_x^{2x} |R_{11}| dy \ll \sum_{D^v \leqq p < x^{1/3}} \int_x^{2x} \min\Big(1, \frac{1}{p\|y/p^2\|}\Big) dy$$
$$= \sum_p p \int_{x/p^2}^{2x/p^2} \min\Big(p, \frac{1}{\|u\|}\Big) du$$
$$\ll \sum_p p \frac{x}{p^2} (\log x)$$
$$\ll x (\log x).$$

We next deal with R_2.

$$R_2 = \sum_{xD^{-2u} < m \leqq 2x^{1/3}} \sum_{(y-\Delta)/m < p^2 \leqq y/m} 1$$
$$\leqq \sum_m \sum_{\substack{(y-\Delta)/m < k^2 \leqq y/m \\ k \in \boldsymbol{Z}}} 1$$
$$\leqq \sum_m \sum_{\substack{\sqrt{y/m} - \Delta/\sqrt{mx} < k \leqq \sqrt{y/m} \\ k \in \boldsymbol{Z}}} 1$$
$$= \sum_m \frac{\Delta}{\sqrt{mx}} + \sum_m r_{\sqrt{m}}(\sqrt{y}, \Delta x^{-1/2})$$
$$= O(\Delta (\log x)^{-3}) + \sum_m \sum_{0 < |h| < m} \frac{1}{2\pi i h} \Big\{1 - e\Big(-\frac{h}{\sqrt{m}} \Delta x^{-1/2}\Big)\Big\} e\Big(\frac{h}{\sqrt{m}} y\Big) +$$
$$+ \sum_m \Big(\min\Big(1, \frac{1}{m\|\sqrt{y/m}\|}\Big) + \min\Big(1, \frac{1}{m\|\sqrt{y/m} - \Delta/\sqrt{mx}\|}\Big)\Big)$$
$$= O(\Delta (\log x)^{-3}) + R_{22} + R_{21}, \text{ say.}$$

We have as before

$$\int_x^{2x} |R_{22}|^2 dy$$
$$\ll \int_x^{2x} \Big| \sum_{\substack{0 < h < m \\ xD^{-2u} < m \leqq 2x^{1/3}}}' \Big(\sum_{k < x^{1/3}} \frac{1}{k}\Big) \frac{1}{h} \Big\{1 - e\Big(-\frac{h}{\sqrt{m}} \Delta x^{-1/2}\Big)\Big\} e\Big(\frac{h}{\sqrt{m}} \sqrt{y}\Big) \Big|^2 dy$$
$$\ll \sqrt{x} \int_{\sqrt{x}}^{\sqrt{2x}} \Big| \sum_{h,m}' \Big(\sum_k \frac{1}{k}\Big) \frac{1}{h} \Big\{1 - e\Big(-\frac{h}{\sqrt{m}} \Delta x^{-1/2}\Big)\Big\} e\Big(\frac{h}{\sqrt{m}} u\Big) \Big|^2 du$$

where $\sum'$ indicates that $h/\sqrt{m}$'s are different from each other. Since

$$\Big(\sqrt{m_1} + \sqrt{m_2}\Big) \Big| \frac{h_1}{\sqrt{m_1}} - \frac{h_2}{\sqrt{m_2}} \Big| > \Big(\frac{h_1}{\sqrt{m_1}} + \frac{h_2}{\sqrt{m_2}}\Big) \Big| \frac{h_1}{\sqrt{m_2}} - \frac{h_2}{\sqrt{m_2}} \Big|$$
$$= \Big| \frac{h_1^2}{m_1} - \frac{h_2^2}{m_2} \Big|$$
$$\geqq \frac{1}{m_1 m_2},$$

we see that

$$\min_{\neq}\left|\frac{h_1}{\sqrt{m_1}}-\frac{h_2}{\sqrt{m_2}}\right|\gg(x^{1/3})^{-5/2}.$$

Hence, by Lemma 4, (5) contributes at most

$$\begin{aligned}&\sqrt{x}\,(\sqrt{x}+(x^{1/3})^{5/2})\sum_{\substack{0<h<m\\ xD^{-2u}<m\leqq 2x^{1/3}}}\Big(\sum_{k<x^{1/3}}\frac{1}{k}\Big)^2\frac{1}{h^2}\sin^2\Big(\frac{\pi h\Delta}{\sqrt{mx}}\Big)\\ &\ll\sqrt{x}\,(\sqrt{x}+(x^{1/3})^{5/2})(\log x)^2\sum_{m\leqq 2x^{1/3}}\frac{\Delta}{\sqrt{mx}}\\ &\ll\Delta(\sqrt{x}+(x^{1/3})^{5/2})(x^{1/3})^{1/2}(\log x)^2\\ &\ll\Delta x(\log x)^2.\end{aligned}$$

Moreover we have

$$\begin{aligned}\int_x^{2x}|R_{21}|dy&\ll\sum_{xD^{-2u}<m\leqq 2x^{1/3}}\int_x^{2x}\min\Big(1,\frac{1}{m\|\sqrt{y/m}\|}\Big)dy\\ &\ll\sum_m\frac{1}{m}m\sqrt{\frac{x}{m}}\int_{\sqrt{x/m}}^{\sqrt{2x/m}}\min\Big(m,\frac{1}{\|u\|}\Big)du\\ &\ll\sum_m\frac{x}{m}(\log x)\\ &\ll x(\log x)^2,\end{aligned}$$

by Lemma 5.

Now we proceed to the remainder terms with the sieve estimate.

$$\begin{aligned}\sum_{d<D}\lambda_d r_d(y,\Delta)&=\sum_{d<D}\lambda_d\frac{1}{2\pi i}\sum_{0<|h|<d}\frac{1}{h}\Big\{1-e\Big(-\frac{h}{d}\Delta\Big)\Big\}e\Big(\frac{h}{d}y\Big)+\\ &\quad+\sum_{d<D}\lambda_d\,0\Big(\min\Big(1,\frac{1}{d\|y/d\|}\Big)+\min\Big(1,\frac{1}{d\|(y-\Delta)/d\|}\Big)\Big)\\ &=R_{32}+R_{31},\ \text{say.}\end{aligned}$$

By Lemma 4, we have

$$\begin{aligned}\int_x^{2x}|R_{32}|^2dy&\ll\int_x^{2x}\Big|\sum_{\substack{0<h<d<D\\ (h,d)=1}}\Big(\sum_m\frac{\lambda_{dm}}{m}\Big)\frac{1}{h}\Big\{1-e\Big(-\frac{h}{d}\Delta\Big)\Big\}e\Big(\frac{h}{d}y\Big)\Big|^2dy\\ &\ll(x+D^2)\sum_{h,d}\Big(\sum\frac{\lambda_{dm}}{m}\Big)^2\frac{1}{h^2}\sin^2\Big(\frac{\pi h\Delta}{d}\Big)\\ &\ll\Delta(x+D^2)\sum_{d<D}\frac{9^{\nu(d)}}{d}\Big(\sum_{m<D}\frac{3^{\nu(m)}}{m}\Big)^2\\ &\ll\Delta(x+D^2)(\log x)^{15},\end{aligned}$$

since $\mu(dm)^2=1$. Moreover, Lemma 5 yields

$$\begin{aligned}\int_x^{2x}|R_{31}|dy&\ll\sum_{d<D}|\lambda_d|\int_x^{2x}\min\Big(1,\frac{1}{d\|y/d\|}\Big)dy\\ &\ll\sum_{d<D}|\lambda_d|\int_{x/d}^{2x/d}\min\Big(d,\frac{1}{\|u\|}\Big)du\end{aligned}$$

$$\ll \sum_{d<D} 3^{\nu(d)} \frac{x}{d} (\log x)$$
$$\ll x(\log x)^4.$$

Put

$$E_j(y, \Delta) = R_{1j} + R_{2j} + R_{3j}, \quad j=1, 2.$$

Then, by the above argument, we see

$$\int_x^{2x} |E_1(y, \Delta)| dy \ll x(\log x)^4$$

and

$$\int_x^{2x} |E_2(y, \Delta)|^2 dy \ll \Delta(x+D^2)(\log x)^{15}.$$

Taking

$$D = (2x)^{0.5115} (\log x)^{-9},$$

so that

$$\Delta = \frac{\log 2x}{0.5115(\log 2x) - 9(\log\log x)} < \frac{1}{0.5114} < 1.95544,$$

we get Lemma 1.

This completes our proof.

References

[1] Erdös, P., The difference of consecutive primes., Duke Math. J. **6** (1940), 438-441.

[2] Friedlander, J. B., Moments of sifted sequences., Math. Ann. **267** (1984), 101-106.

[3] Greaves, G., The weighted linear sieve and Selberg's λ^2-method., Acta Arith. **47** (1986), 71-96.

[4] Heath-Brown, D. R., The differences between consecutive primes II., J. London Math. Soc. (2) **19** (1979), 207-220.

[5] Heath-Brown, D. R., The differences between consecutive primes III., ibid (2) **20** (1979), 177-178.

[6] Heath-Brown, D. R., Gaps between primes, and the pair correlation of zeros of the zeta-function., Acta Arith. **41** (1983), 85-99.

[7] Huxley, M. N., The distribution of prime numbers., Oxford, 1972.

[8] Meyer, U., Eine Summe über Differenzen aufeinanderfolgender Fastprimzahlen P_2., Arch. Math. **42** (1984), 448-454.

[9] Montgomery, M. H., and Vaughan, R. C., Hilbert's inequality., J. London Math. Soc. (2) **8** (1974), 73-82.

[10] Selberg, A., On the normal density of primes in short intervals and the difference between consecutive primes., Arch. Math. Naturvid. **47** (1943), 83-105.

Institute of Mathematics
University of Tsukuba
Ibaraki Pref.
305 Japan