ON THE OCCURRENCE OF LARGE GAPS BETWEEN PRIME NUMBERS

by R. J. COOK

(Received 13 July, 1977)

1. Introduction. Let p_n denote the nth prime number. Erdös asked whether

$$\sum_{\substack{p_n<X \\ p_{n+1}-p_n>X^{1/2+\epsilon}}} (p_{n+1}-p_n) \ll X^c \tag{1}$$

for some constant $c<1$. Moreno [7] obtained a somewhat weaker result and subsequently Wolke [10] proved that

$$\sum_{\substack{p_n<X \\ p_{n+1}-p_n>p_n^{1/2}}} (p_{n+1}-p_n) \ll X^{29/30}.$$

As long ago as 1943 Selberg [8] proved that, if the Riemann Hypothesis is true, then (1) holds with $c=1/2+\eta$ for any $\eta>0$. More recently Warlimont [9] obtained analogous results for sums with a condition $p_{n+1}-p_n>p_n^{\phi}$ where $\phi<1/2$, but his method does not appear to give a computable exponent of X.

THEOREM. *For any* $\varepsilon>0$

$$\sum_{\substack{p_n<X \\ p_{n+1}-p_n>p_n^{1/2}}} (p_{n+1}-p_n) \ll X^{85/98+\varepsilon}. \tag{2}$$

The proof follows similar lines to Wolke's, but uses sharper zero-density estimates for the Riemann ζ-function.

2. Preliminaries. Let

$$\psi(x)=\sum_{p^r\le x} \log p.$$

Then from the explicit formula for $\psi(x)$ (see [1, p. 120] with $T=x^{\alpha}$) we have

$$\psi(x)-x=-\sum_{|\gamma|\le x^{\alpha}} \frac{x^{\rho}}{\rho}+O(x^{1-\alpha}(\log x)^2) \tag{3}$$

for $x\ge 9$ and any constant $\alpha>1/2$, where the summation is over the non-trivial zeros $\rho=\beta+i\gamma$ of the Riemann ζ-function.

We take

$$\alpha=\frac{111}{196}, \qquad U=3x^{1/2}, \qquad e^{\delta}=1+\frac{1}{U}, \tag{4}$$

and, choosing $\sigma_0 > 85/98$ (but close to 85/98), we put

$$\Delta(y) = \psi\left(y + \frac{y}{U}\right) - \psi(y) - \frac{y}{U} + \sum \frac{(e^{\delta\rho} - 1)}{\rho} y^\rho, \tag{5}$$

where $x < y \le 2x$ and the summation is over the zeros of $\zeta(s)$ in the region $|\gamma| \le x^\alpha$, $\sigma_0 \le \beta \le 1$. Then

$$\Delta(y) \ll \left|\sum \frac{(e^{\delta\rho} - 1)}{\rho} y^\rho\right| + x^{1-\alpha} \log^2 x, \tag{6}$$

where the summation is over zeros in the region $|\gamma| \le x^\alpha$, $0 < \beta < \sigma_0$.

Following Moreno [7, Lemma 2], we obtain

$$\int_x^{2x} |\Delta(y)|^2 \, dy \ll \frac{x(\log x)^2}{U^2} \sum x^{2\beta} + x^{3-2\alpha} (\log x)^4 \tag{7}$$

where the summation conditions are the same as in (6).

3. Estimation of the integral. Let $N(\sigma, T)$ denote the number of zeros of $\zeta(s)$ in the rectangle $\sigma \le \beta \le 1$, $|\gamma| \le T$ and write $N(T)$ for $N(0, T)$.

LEMMA 1.

$$N(T) \ll T \log T \tag{8}$$

(See, for example, [2, Chapter 15].)

Integrating we have

$$\begin{aligned}\sum_{\substack{|\gamma| \le x^\alpha \\ 0 \le \beta \le \sigma_0}} (x^{2\beta} - 1) &= 2 \sum_{\substack{|\gamma| \le x^\alpha \\ 0 \le \beta \le \sigma_0}} \int_0^\beta x^{2\sigma} \log x \, d\sigma \\ &= 2 \int_0^{\sigma_0} \sum_{\substack{|\gamma| \le x^\alpha \\ \sigma_0 \ge \beta \ge \sigma}} x^{2\sigma} \log x \, d\sigma \\ &\ll \int_0^{\sigma_0} x^{2\sigma} N(\sigma, x^\alpha) \log x \, d\sigma,\end{aligned}$$

so, from Lemma 1,

$$\sum_{\substack{|\gamma| \le x^\alpha \\ 0 \le \beta \le \sigma_0}} x^{2\beta} \ll x^{1+\alpha} \log x + \int_{1/2}^{\sigma_0} x^{2\sigma} N(\sigma, x^\alpha) \log x \, d\sigma. \tag{9}$$

LEMMA 2.

$$N(\sigma, T) \ll T^{3(1-\sigma)/(2-\sigma)} (\log T)^5 \tag{10}$$

uniformly for $1/2 \le \sigma \le 1$.

(This is due to Ingham [5].)

https://doi.org/10.1017/S0017089500003700 Published online by Cambridge University Press

Lemma 2 shows that the contribution to the integral in (9) coming from the interval $1/2 \leq \sigma \leq 3/4$ is

$$\begin{aligned} &\ll (\log x)^6 \max_{1/2 \leq \sigma \leq 3/4} x^{2\sigma + 3\alpha(1-\sigma)/(2-\sigma)} \\ &\ll (\log x)^6 x^{3/2+3\alpha/5} \ll x^{3-2\alpha}. \end{aligned} \tag{11}$$

LEMMA 3.

$$N(\sigma, T) \ll T^{3(1-\sigma)/(3\sigma-1)} (\log T)^{44} \tag{12}$$

uniformly in $3/4 \leq \sigma \leq 1$.

(This is due to Huxley [**3**].)

We apply Lemma 3 on the interval $[\frac{3}{4}, \frac{5}{6}]$ to give a contribution to the integral which is

$$\begin{aligned} &\ll (\log x)^{45} \max_{3/4 \leq \sigma \leq 5/6} x^{2\sigma + 3\alpha(1-\sigma)/(3\sigma-1)} \\ &\ll x^{5/3+\alpha/3} (\log x)^{45} \ll x^{3-2\alpha}. \end{aligned} \tag{13}$$

LEMMA 4. *For any* $\eta > 0$,

$$N(\sigma, T) \ll T^{48(1-\sigma)/37(2\sigma-1)+\eta} \tag{14}$$

uniformly in $61/74 \leq \sigma \leq 37/42$, *where the implicit constant depends on* η.

(This is due to Huxley [**4**].)

Applying Lemma 4 on the interval $[\frac{5}{6}, \sigma_0]$, at $\sigma = 85/98$ we have

$$2\sigma + 48\alpha(1-\sigma)/37(2\sigma-1) = 3 - 2\alpha = 85/98 + 1,$$

so, choosing σ_0 close enough to 85/98, the contribution to the integral coming from the interval $[\frac{5}{6}, \sigma_0]$ is

$$\ll x^{3-2\alpha+\varepsilon}. \tag{15}$$

Combining these estimates, for any $\varepsilon > 0$ and σ_0 satisfying $85/98 < \sigma_0 \leq 85/98 + \delta(\varepsilon)$,

$$\int_{1/2}^{\sigma_0} x^{2\sigma} N(\sigma, x^\alpha) \log x \, d\sigma \ll x^{3-2\alpha+\varepsilon}, \tag{16}$$

and so, from (7) and (9), and recalling that $U = 3x^{1/2}$,

$$\int_x^{2x} |\Delta(y)|^2 \, dy \ll x^{3-2\alpha+\varepsilon}. \tag{17}$$

4. Estimation of a sum. We break the sum

$$\sum \frac{(e^{\delta\rho} - 1)}{\rho} y^\rho$$

https://doi.org/10.1017/S0017089500003700 Published online by Cambridge University Press

occurring in the definition of $\Delta(y)$ into 4 parts, and observe that

$$\frac{(e^{\delta\rho}-1)}{\rho}y^{\rho} \ll \frac{y^{\beta}}{U}.$$

Using Lemma 13 and

$$\sum_{\substack{|\gamma|\le x^{\alpha} \\ \sigma_0\le\beta\le 1}} x^{\beta} \ll \int_{\sigma_0}^{1} x^{\sigma}N(\sigma, x^{\alpha})\log x\, d\sigma,$$

we see that those zeros in the strip $\sigma_0\le\beta\le 37/42$ make a contribution

$$\begin{aligned} &\ll U^{-1}\log x \max x^{\beta+48\alpha(1-\beta)/37(2\sigma_0-1)+\eta} \\ &\ll U^{-1}x^{1-\theta} \end{aligned} \tag{18}$$

for some $\theta>0$, as η can be taken arbitrarily small.

LEMMA 5. *For any* $\eta>0$,

$$N(\sigma, T) \ll x^{3(1-\sigma)/2\sigma+\eta} \tag{19}$$

uniformly in $37/42\le\sigma\le 1$.

(This is due to Huxley [**4**].)

Applying Lemma 5 on the interval $37/42\le\beta\le 23/24$, we obtain a contribution

$$\begin{aligned} &\ll U^{-1}\log x \max x^{\beta+63\alpha(1-\beta)/37+\eta} \\ &\ll U^{-1}x^{1-\theta} \end{aligned} \tag{20}$$

for some $\theta>0$.

LEMMA 6.

$$N(\sigma, T) \ll T^{84(1-\sigma)/55}(\log T)^{50} \tag{21}$$

uniformly in the interval $[23/24, 1]$.

(This follows immediately from Corollary 12.4 of Montgomery [**6**].)

LEMMA 7. *For some positive constant c, we have* $\zeta(s)\ne 0$ *in the region*

$$\sigma = \operatorname{re} s > 1-\frac{c}{(\log\tau)^{2/3}(\log\log\tau)^{1/3}}, \qquad \tau = |\operatorname{im} s|+2. \tag{22}$$

(See [**6**, Corollary 11, 4].)

Taking, for some suitable c,

$$\sigma_1 = 1-\frac{c}{(\log x)^{2/3}(\log\log x)^{1/3}},$$

https://doi.org/10.1017/S0017089500003700 Published online by Cambridge University Press

and using Lemma 6 on the interval $[23/24, \sigma_1]$ we have a contribution

$$\ll U^{-1}\log x \max x^{\beta+84\alpha(1-\beta)/55}(\log x)^{50}$$
$$\ll x/U \log x \tag{23}$$

from the remaining zeros.

5. Proof of the theorem. Let x be large and let p_n, p_{n+1} be two consecutive primes satisfying

$$x < p_n < p_{n+1} \le 2x \quad \text{and} \quad p_{n+1} - p_n > p_n^{1/2}.$$

Choosing y so that

$$p_n < y \le p_n + \tfrac{1}{3}(p_{n+1} - p_n) = q_n,$$

say, we have $y + y/U < p_{n+1}$ and therefore

$$|\Delta(y)| = \left|\psi\left(y + \frac{y}{U}\right) - \psi(y) - \frac{y}{U} + \sum \frac{(e^{\delta\rho} - 1)}{\rho} y^\rho\right|$$
$$= \left|\frac{y}{U} + O\left(\frac{y}{U \log x}\right)\right| \ge \frac{x}{2U} \gg x^{1/2}, \tag{24}$$

if x is sufficiently large. Then

$$\int_{p_n}^{q_n} |\Delta(y)|^2\, dy \gg x(q_n - p_n) \gg x(p_{n+1} - p_n).$$

Hence

$$\sum_{\substack{x < p_n < p_{n+1} \le 2x \\ p_{n+1} - p_n > p_n^{1/2}}} (p_{n+1} - p_n) \ll x^{-1} \int_x^{2x} |\Delta(y)|^2\, dy$$
$$\ll x^{2-2\alpha+\varepsilon} = x^{85/98+\varepsilon}. \tag{25}$$

The theorem now follows on summing estimates from the intervals $[X/2^{\nu+1}, X/2^\nu]$ for $\nu = 0, 1, \ldots$.

REFERENCES

1. K. Chandrasekharan, *Arithmetical functions* (Springer-Verlag, 1970).

2. H. Davenport, *Multiplicative number theory* (Markham, 1967).

3. M. N. Huxley, On the difference between consecutive primes, *Invent. Math.* **15** (1972), 164–170.

4. M. N. Huxley, Large values of Dirichlet polynomials II, *Acta Arith.* **27** (1975), 159–169.

5. A. E. Ingham, On the estimation of $N(\sigma, T)$, *Quart. J. Math. Oxford Ser*(1) **11** (1940), 291–292.

6. H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics No. 227 (Springer-Verlag, 1971).

https://doi.org/10.1017/S0017089500003700 Published online by Cambridge University Press

7. C. J. Moreno, The average size of gaps between primes, *Mathematika* **21** (1974), 96–100.

8. A. Selberg, On the normal density of primes in small intervals and the difference between consecutive primes, *Arch. Math. Naturvid.* **47** (1943), 1–19.

9. R. Warlimont, Uber die Haufigkeit grosser Differenzen konsekutiver Primzahlen, *Monatsh. Math.* **83** (1977), 59–63.

10. D. Wolke, Grosse Differenzen zwischen aufeinanderfolgenden Primzahlen, *Math. Ann.* **218** (1975), 269–271.

UNIVERSITY OF SHEFFIELD
SHEFFIELD S10 2TN