ACTA ARITHMETICA
XXXVIII (1980)

A note on large gaps between prime numbers

by

M. N. HUXLEY (Cardiff)

In memory of Professor Paul Turán

1. Introduction. Let p_n denote the nth prime number, $d_n = p_{n+1} - p_n$. As $n \to \infty$, $d_n/\log p_n$ has mean value 1, and numerical evidence is consistent with Cramér's conjecture [2]:

$$d_n \leqslant \big(1+o(1)\big)\log^2 p_n, \tag{1.1}$$

the best bound known is

$$d_n \ll p_n^{7/12+\varepsilon}. \tag{1.2}$$

(We use ε to denote a positive number which may be chosen arbitrarily small, but ε need not denote the same number in each formula. The symbol $\ll$ denotes an inequality in which an unspecified positive constant, depending only on ε, times the right-hand side exceeds the modulus of the left-hand side.) The Riemann hypothesis implies

$$d_n \ll p_n^{1/2}\log p_n. \tag{1.3}$$

Erdős suggested that the proportion of the interval $[1, N]$ covered by long gaps between consecutive prime numbers tended to zero as $N \to \infty$. Wolke [17] proved that

$$\sum_{\substack{p_n < N \\ d_n > N^{1/2}}} d_n \ll N^{\theta}, \tag{1.4}$$

with $\theta = 29/30$. The exponent has been improved by Ivić [10] to $\theta = 23/25$, and by R. J. Cook [1] to $85/98+\varepsilon$. Related results have been given by Moreno [14], Warlimont [16], Heppner [6], Ivić [10] and Heath-Brown [5]. In this note we use more powerful zero-density theorems for $\zeta(s)$. Our results depend on the 'Lindelöf exponent' c for which

$$\zeta(\tfrac{1}{2}+it) \ll |t|^{c}. \tag{1.5}$$

Lindelöf's hypothesis requires that (1.5) be valid for every $c > 0$. The exponent in (1.4) can be taken as

$$\theta = \tfrac{1}{2}+2\max(c, 41/266)+\varepsilon. \tag{1.6}$$

The number $41/266 = 0.1541\ldots$ could be improved slightly, but not below 3/20 by our methods. Note that 41/266 is less than Kolesnik's estimate [12] $c = 173/1067+\varepsilon = 0.1621\ldots$ for the Lindelöf exponent. Assuming the Lindelöf hypothesis we may, however, take

$$\theta = 3/4+\varepsilon. \tag{1.7}$$

If we replace the condition $d_n > N^{1/2}$ in (1.4) by

$$d_n > N^{1/2+\varepsilon_1} \tag{1.8}$$

for any $\varepsilon_1 > 0$, we have (1.4) with

$$\theta = \frac{1}{2}+2\cdot\frac{41}{266}+\varepsilon_2 = \frac{215}{266}+\varepsilon_2, \tag{1.9}$$

and on the Lindelöf hypothesis

$$\theta = \tfrac{1}{2}+\varepsilon_2, \tag{1.10}$$

where ε_2 can be made arbitrarily small. Results can be obtained for longer or shorter gaps. Thus if we sum over gaps for which

$$d_n > N^{7/12}, \tag{1.11}$$

we can take

$$\theta = 3/4+\varepsilon.$$

As in Ivić [10] our method fails for $d_n \leqslant N^{1/6}$.

I would like to thank Cook, Heath-Brown, Ivić and Wolke for making unpublished manuscripts available to me. Heath-Brown also has new bounds for $\zeta(s)$ which may lead to new estimates for the exponent θ of (1.4).

2. Wolke's method. For simplicity we consider the range

$$N \leqslant p_n < p_{n+1} \leqslant 2N, \tag{2.1}$$

where N is sufficiently large. A result of the type (1.4) may be deduced by letting N run through the powers of 2 and summing over the various ranges. Let as usual

$$\psi(x) = \sum_{a\geqslant 1}\sum_{\substack{p \text{ prime}\\ p^a\leqslant x}}\log p. \tag{2.2}$$

Our starting point is the approximate formula ([15], Satz 4.5), valid for $N \leqslant x \leqslant 2N$ and for $T \leqslant N$,

$$\psi(x) = x-\sum_{\varrho}\frac{x^{\varrho}}{\varrho}+O\left(\frac{Nl^2}{T}\right). \tag{2.3}$$

where $l = \log 2N$, and the sum is over zeros $\varrho = \beta+i\gamma$ of $\zeta(s)$ in the rectangle $0 \leqslant \beta \leqslant 1$, $|\gamma| \leqslant T$. To discuss gaps with

$$d_n = p_{n+1}-p_n \geqslant h = 2\delta N \tag{2.4}$$

we let

$$E(x) = \psi(x+\delta x)-\psi(x)-\delta x \doteq -\sum_{\varrho}x^{\varrho}\delta(\varrho)+O(Nl^2/T), \tag{2.5}$$

where

$$\delta(\varrho) = \int_1^{1+\delta}u^{\varrho-1}\,du, \tag{2.6}$$

$$|\delta(\varrho)| \leqslant \delta. \tag{2.7}$$

By (1.2) $d_n < N^{2/3}$, so that for $a \geqslant 3$ there is at most one ath power between p_n and p_{n+1}. The number of squares in an interval of length d contained in the interval (2.1) is

$$\leqslant 1+d/2N^{1/2}, \tag{2.8}$$

so that

$$\psi(x+\delta x)-\psi(x) \leqslant \delta xl/2N^{1/2}+2l^2 \tag{2.9}$$

when

$$N \leqslant p_n \leqslant x < x+\delta x \leqslant p_{n+1} \leqslant 2N. \tag{2.10}$$

Hence if N is sufficiently large and if $h \geqslant 8l^2$,

$$E(x) \leqslant -\delta x/2 \tag{2.11}$$

on a subinterval of $[p_n, p_{n+1}]$ of length at least $d_n/2$. We now write

$$E(x) = E_1(x)+E_2(x)+E_3(x), \tag{2.12}$$

where for some α to be chosen, $E_1(x)$ is the sum over zeros $\varrho = \beta+i\gamma$ with $\beta < \alpha$, $E_2(x)$ over zeros with $\beta \geqslant \alpha$ and $E_3(x)$ is the error term in (2.5). If

$$T = BNl^2/h \tag{2.13}$$

with some sufficiently large constant B, then

$$|E_3(x)| \leqslant \delta N/6 \leqslant \delta x/6 \tag{2.14}$$

for $N \leqslant x \leqslant 2N$. To estimate $E_1(x)$ and $E_2(x)$ we write as usual $N(\sigma, T)$

for the number of zeros of $\zeta(s)$ with $\beta \geq \sigma$, $|\gamma| \leq T$. By dividing the interval $\alpha \leq \sigma \leq 1$ into subintervals of width $1/l$, we find

$$|E_2(x)| \ll 2\delta l \max_{\alpha \leq \sigma \leq 1} N^{\sigma} N(\sigma, T). \tag{2.15}$$

The known bounds for $N(\sigma, T)$ can be thrown into the form

$$N(\sigma, T) \ll T^{A(\sigma)(1-\sigma)}, \tag{2.16}$$

where $A(\sigma)$ increases from $\sigma = 1/2$ to $\sigma = 3/4$, then decreases for $\sigma > 3/4$. Also $N(\sigma, T)$ is zero for

$$\sigma \geq \sigma_0 = 1 - \frac{D}{l^{2/3}\log^{1/3} l}, \tag{2.17}$$

where D is a constant. Hence if for some $\varepsilon > 0$

$$T^{A(\alpha)} \ll N^{1-\varepsilon}, \tag{2.18}$$

then

$$(T^{A(\sigma)}/N)^{1-\sigma} < 1/12l \tag{2.19}$$

for $\alpha \leq \sigma \leq \sigma_0$, and

$$|E_2(x)| \leq \delta N/6. \tag{2.20}$$

We now have

$$|E_1(x)| \geq \delta N/6 \tag{2.21}$$

on a subinterval of $[p_n, p_{n+1}]$ of length at least $d_n/2$.

We now estimate

$$\int_N^{2N} |E_1(x)|^2\, dx \tag{2.22}$$
$$= \int_{u=1}^{1+\delta} \int_{v=1}^{1+\delta} \sum_{\varrho} \sum_{\varrho'} \delta(\varrho)\delta(\varrho') \frac{(2N)^{\varrho+\varrho'+1} - N^{\varrho+\varrho'+1}}{\varrho+\varrho'+1} u^{\varrho-1} v^{\varrho'-1}\, dv\, du$$
$$\leq 9\delta^2 \sum_{\varrho} \sum_{\varrho'} \frac{N^{\beta+\beta'+1}}{|\varrho+\varrho'+1|} \ll \delta^2 l^2 \sum_{\varrho} N^{2\beta+1} \ll \delta^2 l^3 \max_{1/2 \leq \sigma \leq \alpha} N(\sigma, T) N^{2\sigma+1},$$

where we have used the fact that

$$N(0, t+1) - N(0, t) \ll l \tag{2.23}$$

for $0 \leq t \leq T$. The maximum in (2.22) need not occur at the endpoints of the range $1/2 \leq \sigma \leq \alpha$.

3. Calculations. With $h = N^{1/2}$ we may choose any α for which $A(\alpha) < 2$. By Jutila [11] we have $A(\alpha) < 2+\varepsilon$ for $\alpha \geq 11/14$. If c is an exponent for

which (1.5) holds, and $c \geq 1/7$, so that $1/2+2c \geq 11/14$, then $A(\alpha) < 2$ for $\alpha > 1/2+2c$. This was shown in formula (10.19) of [8] using Haneke's estimate [4], which gives $c = 6/37+\varepsilon$ for any $\varepsilon > 0$. The later estimate of Kolesnik [12] permits $c = 173/1067+\varepsilon$, and we may sharpen (10.19) of [8] to

$$N(\sigma, T) \ll T^{8c(1-\sigma)/(2\sigma-1)+\varepsilon} \tag{3.1}$$

for an interval of values of σ about the point $1/2+2c$. We must now compute the maximum in (2.22). Since by (2.13)

$$T = BN^{1/2} l^2, \tag{3.2}$$

we want the maximum of

$$f(\sigma) = 2\sigma + A(\sigma)(1-\sigma)/2. \tag{3.3}$$

For $1/2 \leq \sigma \leq 3/4$ we use the classical bound of Ingham [9] with

$$A(\sigma) = 3/(2-\sigma)+\varepsilon. \tag{3.4}$$

Since $A(\sigma) < 4$ and $A(\sigma)$ is increasing, the maximum on this range is $9/5+\varepsilon$ at $\sigma = 3/4$. For $3/4 \leq \sigma \leq 11/14$ we use the estimate

$$A(\sigma) = 3/(3\sigma-1)+\varepsilon \tag{3.5}$$

from [7], which can, however, be improved slightly by the methods of the author and Jutila (see for example [8] and [11]). The maximum of $f(\sigma)$ on this range is again at the right-hand end point, $481/266+\varepsilon$ at $\sigma = 11/14$. As remarked above, we may take $A(\sigma) = 2+\varepsilon$ for $\sigma \geq 11/14$. The best estimates obtainable for $A(\sigma)$ form a continuous function of σ, whilst those we are quoting are discontinuous at 11/14. The maximum of $f(\sigma)$ for $11/14 \leq \sigma \leq \alpha$ is now $1+\alpha+\varepsilon$ at $\sigma = \alpha$. Hence we have

$$\frac{\delta^2 N^2}{36} \sum_{d_n > N^{1/2}} \frac{d_n}{2} \leq \int_N^{2N} |E_1(x)|^2\, dx \ll \delta^2 N^{2+\varepsilon} l^3 \max(N^{\alpha}, N^{215/266}), \tag{3.6}$$

the sum being over n for which (2.1) holds. Taking $\alpha = 1/2+2c+\varepsilon$, we deduce (1.4) with the exponent θ given by (1.6).

If we assume the Lindelöf hypothesis, after Turán and Halász [3] we may take $\alpha = 3/4+\varepsilon$, and $A(\sigma) = 2+\varepsilon$ for $1/2 \leq \sigma \leq \alpha$. The maximum of $f(\sigma)$ for $1/2 \leq \sigma \leq \alpha$ is $1+\alpha+\varepsilon$ at $\sigma = \alpha$, which gives (1.7).

If we sharpen the condition $d_n \geq h$ by taking $h = N^{1/2+\varepsilon_1}$ in place of $h = N^{1/2}$, we may take any α for which $A(\alpha) < 2+\varepsilon_3$, in particular, $\alpha = 11/14$, and, if the Lindelöf hypothesis be assumed, even $\alpha = 1/2+\varepsilon_4$. This gives the bounds (1.9) and (1.10). Finally with $h = N^{7/12}$ we want

$A(\alpha) < 12/5$, and we can choose $\alpha = 3/4+\varepsilon$. In place of (3.3) we want the maximum of

$$f(\sigma) = 2\sigma + \frac{5}{12} A(\sigma)(1-\sigma) \tag{3.7}$$

on the range $1/2 \leqslant \sigma \leqslant \alpha$ with $A(\sigma)$ given by (3.4), and this is $1+\alpha+\varepsilon$ at $\sigma = \alpha$. As before, this leads to $\theta = \alpha+\varepsilon$, and the exponent (1.12).

References

[1] R. J. Cook, *On the occurrence of large gaps between prime numbers*, Glasgow Math. J., 20 (1979), pp. 43–48.

[2] H. Cramér, *On the order of magnitude of the difference between consecutive prime numbers*, Acta Arith. 2 (1937), pp. 23–46.

[3] G. Halász and P. Turán, *On the distribution of the roots of Riemann zeta and allied functions I*, J. Number Theory 1 (1969), pp. 121–137.

[4] W. Haneke, *Verschärfung der Abschätzung von* $\zeta(\frac{1}{2}+it)$, Acta Arith. 8 (1963), pp. 357–430.

[5] D. R. Heath-Brown, Ph. D. thesis, Cambridge 1977.

[6] E. Heppner, *Bemerkungen über Lücken zwischen Primzahlen*, Arch. Math. 27 (1976), pp. 492–494.

[7] M. N. Huxley, *On the difference between consecutive primes*, Inventiones Math. 15 (1972), pp. 164–170.

[8] — *Large values of Dirichlet polynomials*, Acta Arith. 24 (1973), pp. 329–346, corrigendum ibid. 26 (1974), p. 444.

[9] A. E. Ingham, *On the estimation of* $N(\sigma, T)$, Quarterly J. Math. (Oxford), 11 (1940), pp. 291–292.

[10] A. Ivić, *On sums of large differences between consecutive primes*, to be revised for Math. Ann.

[11] M. Jutila, *Zero-density estimates for L-functions*, Acta Arith. 32 (1976), pp. 55–62.

[12] G. A. Kolesnik, *On the estimation of certain trigonometric sums*, ibid. 25 (1973), pp. 7–30.

[13] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture notes in Math. 227, Springer-Verlag 1971.

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

[15] K. Prachar, *Primzahlverteilung*, Springer-Verlag 1957.

[16] R. Warlimont, *Über die Häufigkeit grosser Differenzen konsekutiver Primzahlen*, Monatshefte Math. 83 (1977), pp. 59–63.

[17] D. Wolke, *Grosse Differenzen zwischen aufeinanderfolgenden Primzahlen*, Math. Ann. 218 (1975), pp. 269–271.

Received on 27. 1. 1978 (1036)

Note added in proof by the editor. In the revised version of [10] that has already appeared in Math. Ann. 241 (1979), pp. 1–9, Ivić proves (1.4) with $\theta = 215/266+\varepsilon$ and announces the value $\theta = 3/4+\varepsilon$ due to Heath-Brown.