LONGER THAN AVERAGE INTERVALS CONTAINING NO PRIMES

A. Y. CHEER AND D. A. GOLDSTON

ABSTRACT. We present two methods for proving that there is a positive proportion of intervals which contain no primes and are longer than the average distance between consecutive primes. The first method is based on an argument of Erdös which uses a sieve upper bound for prime twins to bound the density function for gaps between primes. The second method uses known results about the first three moments for the distribution of intervals with a given number of primes. Better results are obtained by assuming that the first n moments are Poisson. The related problem of longer than average gaps between primes is also considered.

I. Introduction. In this paper we examine the occurrence of long intervals containing no prime numbers from a statistical point of view. Previous work on this subject (see [15, Chapter 5]) has been directed toward constructing very long sequences of consecutive composite numbers. However, these sequences occur so infrequently they have no statistical significance. In this paper, we introduce two different methods for finding intervals longer than the average which contain no primes. Both of these methods produce a positive proportion of such intervals.

To state our main results, we introduce the following notation. Let p_n denote the nth prime number, and let

$$(1.1) \qquad d_n = p_{n+1} - p_n$$

denote the nth gap between consecutive primes. By the prime number theorem, as $X \to \infty$

$$(1.2) \qquad \sum_{X \leqslant p_n \leqslant 2X} 1 \sim \frac{X}{\log X}, \qquad \sum_{X \leqslant p_n \leqslant 2X} d_n \sim X;$$

hence the average of the numbers d_n is $\log p_n \sim \log n$. We define

$$(1.3) \qquad a_0(\lambda, X) = \frac{1}{X} \# \{ n \in [X, 2X] : \text{the interval } [n, n + \lambda \log n]$$

$$\text{contains no primes} \}$$

$$= \frac{1}{X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ d_n \geqslant \lambda \log n}} (d_n - \lambda \log n) + o(1);$$

$a_0(\lambda, X)$ is the percentage of intervals of length λ times the average spacing of primes which contain no primes. A simple argument using the prime number theorem shows that for any $\varepsilon > 0$ and X sufficiently large,

$$(1.4) \qquad a_0(\lambda, X) \geqslant 1 - \lambda - \varepsilon.$$

We improve this lower bound by showing

THEOREM 1. *For any $\varepsilon > 0$, and all sufficiently large X, we have*

$$(1.5) \qquad a_0(\lambda, X) \geqslant 2\left(\tfrac{9}{8} - \lambda\right)^2 - \varepsilon, \quad \text{for } \tfrac{7}{8} \leqslant \lambda \leqslant \tfrac{9}{8},$$

and

$$(1.6) \qquad \sum_{X \leqslant p_n \leqslant 2X} d_n^2 \geqslant \left(\frac{193}{192} - \varepsilon\right) X \log X.$$

Equation (1.5) shows that there is a positive proportion of intervals of length up to $\tfrac{9}{8}$ times the average which contain no primes. However, we cannot conclude from this that there is a positive proportion of the gaps d_n which are longer than average, since there could be a large but nonpositive proportion of extremely long gaps which induce $a_0(\lambda, X)$ to be positive. This seems unlikely, but we cannot exclude this possibility without assuming some unproven hypotheses. We shall discuss this further in §9.

The result stated in (1.6) is another measure of the frequency of large gaps d_n. The trivial estimate $\sum_{X \leqslant p_n \leqslant 2X} d_n^2 \geqslant (1 - \varepsilon) X \log X$ follows from (1.2), since by Cauchy's inequality

$$\left(\sum_{X \leqslant p_n \leqslant 2X} d_n\right)^2 \leqslant \left(\sum_{X \leqslant p_n \leqslant 2X} 1\right)\left(\sum_{X \leqslant p_n \leqslant 2X} d_n^2\right).$$

Much work has been done on bounding $\sum_{X \leqslant p_n \leqslant 2X} d_n^2$ from above; the best upper bound known is $X^{23/18+\varepsilon}$ due to Heath-Brown [10], while on assumption of the Riemann Hypothesis Selberg [16] proved the upper bound $X \log^3 X$, and on assumption of the Riemann Hypothesis and a pair correlation hypothesis Heath-Brown [11] has obtained the bound $X \log^2 X$, which falls short of the conjecture of Erdös [4] that the correct upper bound is of order $X \log X$.

Our proof of Theorem 1 is based on the method of Erdös [4]. The constants we obtain in Theorem 1 depend on the constant in the sieve upper bound for prime pairs. We have used here the bound obtained by Bombieri and Davenport [1]. Small improvements in this constant since then lead to slightly better values in Theorem 1. In §3 we prove a more general form of Theorem 1 which explicitly incorporates the sieve upper bound into the theorem. In a recent paper [14], Prachar has also noted that the Erdös method gives results for large gaps.

The main disadvantage with the method used in proving Theorem 1 is that it cannot be expected to ever give the correct size for the lower bounds in (1.5) and (1.6). A complete knowledge of the distribution of primes in intervals of a given length is required to solve these questions. It has been conjectured that the primes are distributed around their average in a Poisson distribution. This conjecture

implies in particular

$$(1.7) \qquad a_0(\lambda, X) \sim e^{-\lambda} \quad \text{as } X \to \infty,$$

and

$$(1.8) \qquad \sum_{X \leqslant p_n \leqslant 2X} d_n^2 \sim 2X \log X.$$

The strongest evidence in support of this conjecture is due to Gallagher [6], who showed that this conjecture is true assuming that the Hardy-Littlewood r-tuple conjecture holds uniformly over tuples of length $\lambda \log X$. We define the kth moment

$$(1.9) \qquad M_k(\lambda, X) = \frac{1}{X} \int_X^{2X} (\pi(x + \lambda \log x) - \pi(x))^k \, dx,$$

where $\pi(x)$ is the number of primes less than or equal to x. Analogous to (1.3) we define, for $n = 0, 1, 2, \ldots,$

$$a_n(\lambda, X) = \frac{1}{X} \text{ measure } \{ x \in [X, 2X] : [x, x + \lambda \log x) \text{ contains}$$

$$\text{exactly } n \text{ primes} \}.$$

It is easy to check that with $n = 0$ this definition agrees with the earlier definition of $a_0(\lambda, X)$ up to $o(1)$. Gallagher proved that the r-tuple conjecture implies

$$(1.10) \qquad M_k(\lambda, X) \sim m_k(\lambda) = \sum_{r=1}^k \sigma(k, r) \frac{\lambda^r}{r!}, \quad \text{as } X \to \infty;$$

here $\sigma(k, r)$ is the Stirling number of the second type, and $m_k(\lambda)$ is the kth moment of the Poisson distribution. Knowing (1.10) for all k completely determines the distribution, and gives

$$(1.11) \qquad a_n(\lambda, X) \sim \frac{e^{-\lambda}\lambda^n}{n!}, \quad \text{as } X \to \infty.$$

In particular, when $n = 0$ we get (1.7).

Our second method for examining larger than average intervals containing no primes uses information about a finite number of moments $M_k(\lambda, X)$. A brief summary of the moment method is presented in [2]. Our present knowledge about these moments is quite limited, and the results we obtain unconditionally are much weaker than what is obtained in Theorem 1. In §5 we show there is a constant $c(\lambda) > 0$ such that

$$(1.12) \qquad a_0(\lambda, X) > c(\lambda) \text{ for } \lambda = 1.004 \text{ and } X \text{ sufficiently large.}$$

However, if we knew the second and third moments, we could do substantially better than Theorem 1.

THEOREM 2. *Assume* (1.10) *holds for* $k = 2$ *and* 3. *Then, for* $\varepsilon > 0$ *and all sufficiently large* X, *we have*

$$(1.13) \qquad a_0(\lambda, X) \geqslant 1 - \lambda + \tfrac{5}{12}\lambda^2 - \tfrac{1}{12}\lambda^3 - \varepsilon, \quad \text{for } 0 \leqslant \lambda \leqslant 2,$$

and

$$(1.14) \qquad \sum_{X \leqslant p_n \leqslant 2X} d_n^2 \geqslant \left(\frac{14}{9} - \varepsilon \right) X \log X.$$

We also prove in §6 a more general theorem which gives lower bounds for $a_0(\lambda, X)$ in terms of the total number of moments assumed to be Poisson. As might be expected, this lower bound converges to $e^{-\lambda}$ as more moments are used. We show that the most that can be proven for distributions with the first three moments Poisson is

$$a_0(\lambda, X) \geqslant c(\lambda) > 0 \quad \text{for } \lambda < 2.$$

The general problem for n moments is a linear programming problem with no simple optimal solution. We examine this more closely in §6.

Finally in §9, we show that our results hold not only for intervals but also for a positive proportion of gaps d_n, provided the second moment is Poisson.

ACKNOWLEDGMENT. We would like to thank A. Odlyzko and C. Pomerance who (independently) pointed out that the method of Erdös for small gaps could be adapted to large gaps, and who both gave proofs of results similar to Theorem 1. The result stated in (1.12) was obtained jointly with D. R. Heath-Brown, whom we would also like to thank.

2. The method of Erdös. In 1940 Erdös [4] proved that there are infinitely many n such that

$$(2.1) \qquad d_n < (1 - A)\log p_n,$$

where A is an absolute positive constant. We now give a description of this method similar to the one due to Huxley [12]. We define

$$(2.2) \qquad G(\lambda, X) = \frac{1}{X/\log X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ d_n \geqslant \lambda \log p_n}} 1.$$

This function gives the percentage of gaps d_n which are longer than λ times the average; $F(\lambda, X) = 1 - G(\lambda, X)$ is the distribution function for these gaps. We may write $G(\lambda, X)$ formally in terms of a density function $\rho(u, X)$ as

$$(2.3) \qquad G(\lambda, X) = \int_\lambda^\infty \rho(u, X) \, du,$$

where $\rho(u, X)$ is determined by (2.2) only up to integration on fixed intervals. With this notation, equation (1.2) becomes

$$(2.4) \qquad \int_0^\infty \rho(u, X) \, du \sim 1, \qquad \int_0^\infty u\rho(u, X) \, du \sim 1, \quad \text{as } X \to \infty.$$

The second relation above is most easily seen by expressing

$$\frac{1}{X} \sum_{X \leqslant p_n \leqslant 2X} d_n \sim -\int_0^\infty u \, dG(u, X),$$

where the integral is a Stieltjes integral, and then applying (2.3). On combining the results in (2.4), we obtain

$$\int_0^1 (1 - u)\rho(u, X)\,du \sim \int_1^\infty (u - 1)\rho(u, X)\,du.$$

Hence we may interpret $\rho(u, X)$ as a density function with total mass 1 and center of mass at $u = 1$. Erdös's insight was that by using a sieve upper bound for prime pairs one can (in effect) upper bound $\rho(u, X)$, which forces $\rho(u, X)$ to spread out on either side of $u = 1$.

The sieve result needed is the following: for p and p' primes

$$(2.5) \qquad \sum_{\substack{X \leqslant p,\, p' \leqslant 2X \\ p' - p = 2k}} 1 < (\mathscr{C} + \varepsilon) 2H(k) \frac{X}{\log^2 X},$$

for any $\varepsilon > 0$ and X sufficiently large, and uniformly for all k. Here

$$(2.6) \qquad H(k) = H \prod_{\substack{p \mid k \\ p > 2}} \frac{p - 1}{p - 2}, \qquad H = \prod_{p > 2} \left(1 - (p - 1)^{-2}\right),$$

and $\mathscr{C}$ is an absolute constant obtained by the sieve method. The upper bound (2.5) for $\mathscr{C} = 4$ was proved in [1] (see also [9]). The best value of $\mathscr{C}$ presently known is $\mathscr{C} = 3.9171$, due to Chen [3]. It has been conjectured that (2.5) holds with asymptotic equality when $\mathscr{C} = 1$. The arithmetic function $H(k)$ has the important property that it is on average equal to 1:

$$(2.7) \qquad \sum_{k=1}^n H(k) \sim n$$

(see [1]).

We now are ready to obtain the bound needed for $\rho(u, X)$:

LEMMA 1. *For any Riemann integrable function $f(u) \geqslant 0$, $0 \leqslant a < b < \infty$, we have for $\varepsilon > 0$ and X sufficiently large*

$$(2.8) \qquad 0 \leqslant \int_a^b f(u)\rho(u, X)\,du < (\mathscr{C} + \varepsilon) \int_a^b f(u)\,du.$$

PROOF. By (2.3), for $0 \leqslant \alpha < \beta$,

$$\int_\alpha^\beta \rho(u, X)\,du = G(\alpha, X) - G(\beta, X)$$

$$= \frac{1}{X/\log X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ \alpha \log p_n \leqslant d_n < \beta \log p_n}} 1.$$

Hence $0 \leqslant \int_\alpha^\beta \rho(u, X)\, du$, and also

$$\int_\alpha^\beta \rho(u, X)\, du \leqslant \frac{1}{X/\log X} \sum_{\substack{X \leqslant p \leqslant 2X \\ \alpha \log X \leqslant p' - p \leqslant \beta \log(2X)}} 1$$

$$\leqslant \frac{1}{X/\log X} \sum_{\alpha \log X \leqslant 2k \leqslant \beta \log(2X)} \left(\sum_{\substack{X \leqslant p,\, p' \leqslant 2(X+k) \\ p' - p = 2k}} 1 \right)$$

$$\leqslant \frac{2(\mathscr{C} + \varepsilon)}{\log X} \left(\sum_{(\alpha \log X)/2 \leqslant k \leqslant (\beta \log(2X))/2} H(k) \right)$$

$$\leqslant (\mathscr{C} + \varepsilon)(\beta - \alpha) = (\mathscr{C} + \varepsilon) \int_\alpha^\beta 1\, du,$$

where we have used (2.5) and (2.7), and ε differs at each step as necessary. We conclude (2.8) holds for nonnegative step functions, and thus for Riemann integrable $f(u) \geqslant 0$ by using upper and lower step function approximations.

REMARKS. The sieve constant $\mathscr{C}$ defined in equation (2.5) has very recently been improved by Fouvry and Grupp [5] for restricted values of k. They have obtained $\mathscr{C} = 3.454$ when $k = 1$, and have indicated that their proof can be generalized to hold for $k \leqslant (\log X)^A$, for any fixed A. This is sufficient for our purposes since we only needed $k \ll \log X$ in the proof of Lemma 1. If this value of $\mathscr{C}$ is used in Theorem 3 (see next section) we obtain gaps of length $8/7$ instead of $9/8$ in (1.5) and the constant $193/192$ may be replaced by $145/144$ in (1.6).

3. Proof of Theorem 1. Our goal is to find numbers $\lambda > 0$ and $c(\lambda) > 0$, independent of X, such that, given any $\varepsilon > 0$,

(3.1) $a_0(\lambda, X) \geqslant c(\lambda) - \varepsilon,$ for all $X \geqslant X_0(\varepsilon)$.

For any given X, $a_0(\lambda, X)$ is a nonincreasing function of λ, and hence $a_0(\mu) \geqslant c(\lambda) - \varepsilon$ for all $\mu \leqslant \lambda$. Therefore it makes sense to try to find the largest λ for which (3.1) holds. We define

(3.2) $\Lambda = \sup\{\lambda \colon (3.1) \text{ holds}\}.$

In other words $a_0(\lambda) > 0$ for all sufficiently large X if λ is any given number less than Λ. We now prove a generalization of Theorem 1.

THEOREM 3. *We have $\Lambda \geqslant 1 + 1/2\mathscr{C}$. Furthermore, for $\varepsilon > 0$ and X sufficiently large,*

(3.3) $a_0(\lambda, X) \geqslant \dfrac{\mathscr{C}}{2}\left(1 + \dfrac{1}{2\mathscr{C}} - \lambda\right)^2 - \varepsilon$ *for* $1 - \dfrac{1}{2\mathscr{C}} \leqslant \lambda \leqslant 1 + \dfrac{1}{2\mathscr{C}},$

and

(3.4) $\displaystyle\sum_{X \leqslant p_n \leqslant 2X} d_n^2 \geqslant \left(1 + \frac{1}{12\mathscr{C}^2} - \varepsilon\right) X \log X.$

$\mathscr{C}$ is any number for which (2.5) holds ($\mathscr{C} = 4$ is Theorem 1). We make the comment that these lower bounds are what would be obtained if $\rho(u, X) = \mathscr{C}$ for $1 - 1/2\mathscr{C} \leqslant u \leqslant 1 + 1/2\mathscr{C}$ and zero everywhere else. This may be viewed in some sense as the worst case for the theorem. For convenience of notation we drop all reference to X and ε. This has no effect on the argument; each equation would only be modified by a $o(1)$ term or an ε term, and X would be taken sufficiently large. Therefore, we rewrite (2.4) and (2.8) as

$$(3.5) \qquad \int_0^\infty \rho(u)\,du = 1, \qquad \int_0^\infty u\rho(u)\,du = 1,$$

and, for $f(u) \geqslant 0$ and Riemann integrable,

$$(3.6) \qquad 0 \leqslant \int_a^b \rho(u)f(u)\,du \leqslant \mathscr{C}\int_a^b f(u)\,du.$$

By (1.3), (2.2), and (2.3) we have

$$a_0(\lambda, X) \sim \frac{1}{X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ d_n \geqslant \lambda \log p_n}} (d_n - \lambda \log p_n)$$

$$\sim -\int_\lambda^\infty (u - \lambda)\,dG(u, X)$$

$$\sim \int_\lambda^\infty (u - \lambda)\rho(u, X)\,du,$$

and therefore we write this as

$$(3.7) \qquad a_0(\lambda) = \int_\lambda^\infty (u - \lambda)\rho(u)\,du.$$

We now prove $\Lambda \geqslant 1 + 1/2\mathscr{C}$ by the same argument used by Erdös, but with our present notation; for $0 < a \leqslant b$,

$$1 = \int_0^\infty u\rho(u)\,du = \int_0^a u\rho(u)\,du + \int_a^b u\rho(u)\,du + \int_b^\infty u\rho(u)\,du$$

$$\leqslant a\int_0^a \rho(u)\,du + \int_a^b u\rho(u)\,du + \int_b^\infty u\rho(u)\,du.$$

We note that we have not assumed here $\rho(u)$ is positive (although we may think of it as positive); rather we have used the lower bound in (3.6) to show

$$\int_0^a u\rho(u)\,du = a\int_0^a \rho(u)\,du - \int_0^a (a - u)\rho(u)\,du \leqslant a\int_0^a \rho(u)\,du.$$

This comment will also apply at later points in the argument. Continuing, we have

$$1 \leqslant a\left(1 - \int_a^\infty \rho(u)\,du\right) + \int_a^b u\rho(u)\,du + \int_b^\infty u\rho(u)\,du,$$

$$\leqslant a + \int_a^b (u - a)\rho(u)\,du + \int_b^\infty (u - a)\rho(u)\,du$$

$$\leqslant a + \mathscr{C}\int_a^b (u - a)\,du + \int_b^\infty (u - a)\rho(u)\,du,$$

and hence

$$(3.8) \qquad 1 - a - \frac{\mathscr{C}}{2}(b - a)^2 \leqslant \int_b^\infty (u - a)\rho(u)\,du.$$

For fixed b the left side is maximized with $a = b - 1/\mathscr{C}$, which gives

$$1 - b + \frac{1}{2\mathscr{C}} \leqslant \int_b^\infty \left(u - b + \frac{1}{\mathscr{C}} \right) \rho(u) \, du.$$

The left side is positive for any $b < 1 + 1/2\mathscr{C}$. The right-hand side by (3.7) is equal to

$$a_0(b) + \frac{1}{\mathscr{C}} \int_b^\infty \rho(u) \, du \leqslant a_0(b) + \frac{1}{\mathscr{C}} \int_{b-\delta}^\infty \frac{u - (b - \delta)}{\delta} \rho(u) \, du$$

$$= a_0(b) + \frac{1}{\delta\mathscr{C}} a_0(b - \delta)$$

$$\leqslant a_0(b - \delta)\left[1 + \frac{1}{\delta\mathscr{C}} \right].$$

Hence, since $a_0(b - \delta) > 0$ for any b less than $1 + 1/2\mathscr{C}$ and taking $\delta > 0$ sufficiently small, we conclude $\Lambda \geqslant 1 + 1/2\mathscr{C}$.

We now prove (3.3). The trivial estimate is obtained as follows:

$$1 = \int_0^\infty u\rho(u) \, du = \int_0^\lambda u\rho(u) \, du + \int_\lambda^\infty u\rho(u) \, du$$

$$\leqslant \lambda \int_0^\lambda \rho(u) \, du + \int_\lambda^\infty u\rho(u) \, du$$

$$= \lambda + \int_\lambda^\infty (u - \lambda)\rho(u) \, du$$

$$= \lambda + a_0(\lambda).$$

Hence,

(3.9) $$1 - \lambda \leqslant a_0(\lambda).$$

This is the result (1.4), due to Gallagher and mentioned in [11]. To obtain a better estimate, we start from equation (3.8); for $0 < a \leqslant b \leqslant c$,

$$1 - a - \frac{\mathscr{C}}{2}(b - a)^2 \leqslant \int_b^\infty (u - a)\rho(u) \, du$$

$$= \int_b^\infty (u - b)\rho(u) \, du + \int_b^c (b - a)\rho(u) \, du$$

$$+ \int_c^\infty (b - a)\rho(u) \, du$$

$$\leqslant a_0(b) + \int_b^c (b - a)\rho(u) \, du + (b - a)\int_c^\infty \frac{u - b}{c - b}\rho(u) \, du$$

$$\leqslant a_0(b) + \int_b^c (b - a)\left(1 - \frac{u - b}{c - b}\right)\rho(u) \, du$$

$$+ \frac{b - a}{c - b} \int_b^\infty (u - b)\rho(u) \, du$$

$$\leqslant a_0(b)\left(1 + \frac{b - a}{c - b}\right) + \frac{\mathscr{C}}{2}(b - a)(c - b).$$

We conclude

$$a_0(b) \geqslant \left(1 - a - \frac{\mathscr{C}}{2}(b-a)^2 - \frac{\mathscr{C}}{2}(b-a)(c-b)\right)\left(1 + \frac{b-a}{c-b}\right)^{-1}.$$

Now, given b, we wish to pick a and c to maximize the right side. Making a change of variable $h = b - a$, $k = (c-b)/h$ we have $h \geqslant 0$, $k \geqslant 0$, and

$$a_0(b) \geqslant \left(1 + h - \frac{\mathscr{C}}{2}(1+k)h^2 - b\right)\left(1 + \frac{1}{k}\right)^{-1}.$$

A simple calculation shows that the expression on the right is maximized when

$$h = \frac{1}{\mathscr{C}(k+1)} \quad \text{and} \quad k = \frac{1 - 2\mathscr{C}(b-1)}{1 + 2\mathscr{C}(b-1)},$$

subject to the condition $1 - 1/2\mathscr{C} \leqslant b \leqslant 1 + 1/2\mathscr{C}$. Using these choices of h and k and taking $b = \lambda$ gives (3.3). For $0 \leqslant \lambda \leqslant 1 - 1/2\mathscr{C}$ the trivial estimate (3.9) is better.

Finally we prove (3.4). First, we note

(3.10)
$$\sum_{X \leqslant p_n \leqslant 2X} d_n^2 \sim 2X \log X \int_0^\infty a_0(\lambda, X) \, d\lambda,$$

since by (1.3)

$$\int_0^\infty a_0(\lambda, X) \, d\lambda \sim \frac{1}{X} \sum_{X \leqslant p_n \leqslant 2X} \int_0^{d_n/\log X} (d_n - \lambda \log X) \, d\lambda$$

$$\sim \frac{1}{2X \log X} \sum_{X \leqslant p_n \leqslant 2X} d_n^2.$$

Now, by (3.3) and (3.9)

$$\sum_{X \leqslant p_n \leqslant 2X} d_n^2 \geqslant 2(1-\varepsilon)X \log X \int_0^{1+1/2c} a_0(\lambda, X) \, d\lambda$$

$$\geqslant 2(1-\varepsilon)X \log X$$

$$\times \left\{ \int_0^{1-1/2\mathscr{C}}(1-\lambda) \, d\lambda + \int_{1-1/2\mathscr{C}}^{1+1/2\mathscr{C}} \frac{\mathscr{C}}{2}\left(1 + \frac{1}{2\mathscr{C}} - \lambda\right)^2 d\lambda \right\}$$

$$\geqslant 2(1-\varepsilon)X \log X \left\{ \frac{1}{2} + \frac{1}{24\mathscr{C}^2} \right\}$$

$$\geqslant \left(1 + \frac{1}{12\mathscr{C}^2} - \varepsilon\right)X \log X.$$

This proves (3.4). We may express $\sum_{X \leqslant p_n \leqslant 2X} d_n^2$ directly in terms of $\rho(u, X)$ by the formula

(3.11)
$$\sum_{X \leqslant p_n \leqslant 2X} d_n^2 \sim X \log X \int_0^\infty u^2 \rho(u, X) \, du,$$

which is obtained by substituting (3.7) into (3.10) and inverting the order of integration.

4. The moment method. We recall from §1 that

$$(4.1) \qquad M_k(\lambda, X) = \frac{1}{X} \int_X^{2X} \left(\pi(x + \lambda \log x) - \pi(x) \right)^k dx,$$

and

$$(4.2) \qquad a_n(\lambda, X) = \frac{1}{X} \text{ measure } \{ x \in [X, 2X] : [x, x + \lambda \log x) \text{ contains}$$

$$\text{exactly } n \text{ primes} \}.$$

Since $(\pi(x + \lambda \log x) - \pi(x))^k$ is a step function, we may integrate in (4.1) and obtain, for $k \geq 1$,

$$(4.3) \qquad M_k(\lambda, X) = \sum_{n=1}^{\infty} n^k a_n(\lambda, X).$$

The sum above is actually a finite sum since $a_n(\lambda, X) = 0$ for $n \geq X$. We also have the obvious relation

$$(4.4) \qquad a_0(\lambda, X) + \sum_{n=1}^{\infty} a_n(\lambda, X) = 1.$$

We are now ready to describe our method. Suppose $a_0(\lambda, X) = 0$ for some given λ and X large. Moreover, suppose we have upper and/or lower bounds for $M_k(\lambda, X)$, $k = 1, 2, \ldots, N$. Equations (4.3) and (4.4) now form a system of $N + 1$ linear equations in the unknowns $a_1(\lambda, X)$, $a_2(\lambda, X), \ldots,$. We seek solutions such that $a_1(\lambda, X) \geq 0$, $a_2(\lambda, X) \geq 0, \ldots$. This is a linear programming problem, and we know that if for a given λ the system is solvable, then there is a solution with at most $N + 1$ nonzero a's. On the other hand, if there is no solution for a given λ, then the assumption $a_0(\lambda, X) = 0$ is untenable, and we conclude $\Lambda \geq \lambda$ (Λ is defined in (3.2)). For a given N (not too large) a computer may be used to solve the system as accurately as desired. In general, certain simple arguments may be used to show a system is unsolvable, although usually there is some loss in doing this.

5. Proof of (1.12). We first illustrate the moment method by applying it to prove equation (1.12). The following is known about $M_k(\lambda, X)$: For fixed λ, $\varepsilon > 0$, and X sufficiently large,

$$(5.1) \qquad M_1(\lambda, X) \sim \lambda, \quad \text{as } X \to \infty,$$

$$(5.2) \qquad M_2(\lambda, X) \geq \lambda/2 + \lambda^2 - \varepsilon,$$

$$(5.3) \qquad M_2(\lambda, X) \leq \lambda + \mathscr{C}\lambda^2 + \varepsilon,$$

and

$$(5.4) \qquad M_3(\lambda, X) \leq \lambda + 3\mathscr{C}\lambda^2 + \mathscr{D}\lambda^3.$$

Equation (5.1) is the prime number theorem, equation (5.2) is from [7], and equations (5.3) and (5.4) follow from sieve upper bounds applied to the argument on page 5 of [6]. The number $\mathscr{C}$ is the same constant as in (2.5), and $\mathscr{D}$ is the constant corresponding to the upper bound in the 3-dimensional sieve for prime triples. The best known value for $\mathscr{D}$ is 48 (see [9]).

We now turn to the proof of (1.12). We may ignore all ε and X dependences, since these have no effect on the proof. We fix λ, and write $a_n(\lambda, X)$ as a_n. Assume $a_0 = 0$. The equations corresponding to (4.3), (4.4), (5.1), (5.2), and (5.4) with $\mathscr{C} = 4$ and $\mathscr{D} = 48$ are

$$(5.5) \qquad \sum_{n=1}^{\infty} a_n = 1, \qquad \sum_{n=1}^{\infty} n a_n = \lambda, \qquad \sum_{n=1}^{\infty} n^2 a_n \geqslant \frac{\lambda}{2} + \lambda^2,$$

$$\sum_{n=1}^{\infty} n^3 a_n \leqslant \lambda + 12\lambda^2 + 48\lambda^3.$$

By recombining we have

$$(5.6) \qquad \sum_{n=1}^{\infty} (n-1) a_n = \lambda - 1, \qquad \sum_{n=1}^{\infty} (n-1)^2 a_n \geqslant \lambda^2 - \frac{3}{2}\lambda + 1,$$

$$\sum_{n=1}^{\infty} (n-1)^3 a_n \leqslant 48\lambda^3 + 9\lambda^2 + \frac{5}{2}\lambda - 1.$$

By Cauchy's inequality,

$$(5.7) \qquad \left(\sum_{n=1}^{\infty} (n-1)^2 a_n \right)^2 \leqslant \left(\sum_{n=1}^{\infty} (n-1) a_n \right)\left(\sum_{n=1}^{\infty} (n-1)^3 a_n \right),$$

which gives from (5.6) after simplifying

$$0 \leqslant 47\lambda^3 - 36\lambda^2 - \tfrac{43}{4}\lambda - \tfrac{1}{2}.$$

This is false unless λ is larger than the root $\lambda_0 = 1.004259085\ldots$, which proves (1.12). The above argument is very close to optimal. Using a linear programming computation we find (5.6) is unsolvable for $\lambda = 1.0042591123$, while for $\lambda = 1.0042591124$ we have the solution $a_1 = .99996$, $a_{117} = 1.3041 \times 10^{-9}$, $a_{118} = 3.6518 \times 10^{-6}$, and $a_{119} = 3.2472 \times 10^{-5}$ (rounded to 4 decimal places).

6. N Poisson moments. We now turn to the question of obtaining information about Λ and $a_0(\lambda, X)$ given that the first N moments $M_k(\lambda, X)$ are Poisson. Specifically, consider the system of $N + 1$ equations given by (4.3) with $M_k(\lambda, X) = m_k(\lambda)$ for $k = 1, 2, \ldots, N$, $(m_k(\lambda)$ is the kth Poisson moment defined in (1.10)) and (4.4) with $a_0(\lambda, X) = 0$. We define Λ_N^* to be the supremum of the set of all λ's such that these $N + 1$ equations have no solution in nonnegative numbers $a_n(\lambda)$. Clearly

$$(6.1) \qquad \Lambda \geqslant \Lambda_N^*, \qquad \Lambda_{N+1}^* \geqslant \Lambda_N^*.$$

We will prove in the next section the following:

THEOREM 4. *We have*

$$(6.2) \qquad \Lambda_1^* = 1, \qquad \Lambda_2^* = 1, \qquad \Lambda_3^* = 2,$$

and in general

$$(6.3) \qquad \Lambda_{2N+1}^* \geqslant .55\left(N + \tfrac{3}{2}\right) \quad \text{for } N \geqslant 2.$$

Furthermore, assuming equation (1.10) *holds for* $k = 1, 2, \ldots, 2N + 1$, *then for* $\varepsilon > 0$ *and* X *sufficiently large*,

$$(6.4) \qquad a_0(\lambda, X) \geqslant \sum_{n=0}^{2N+1} \frac{(-\lambda)^n}{n!} - \varepsilon.$$

The method used to prove the theorem does not actually need the asymptotic results for the $M_k(\lambda, X)$'s, but only upper bounds for the odd moments and lower bounds for the even moments.

The lower bound in (6.4) approaches the correct asymptotic expression $e^{-\lambda}$ for a Poisson distribution as $N \to \infty$. However, neither (6.3) nor (6.4) are sharp. We give in §8 a different method that may be used to refine (6.3) and (6.4). In particular, we use this method to prove Theorem 2 in the introduction.

It seems likely that $\Lambda_{2N}^* = \Lambda_{2N-1}^*$. Computationally one finds that this is the case. The solution of the system with $k = 2N$ has $2N + 1$ nonzero a_n's. Numerically, as $\lambda \to \Lambda_{2N-1}^*$, the first $2N$ nonzero a_n's will solve the $k = 2N - 1$ system, and the remaining nonzero a_n moves off to infinity ($n \to \infty$) and its value tends to zero. We have computed Λ_N^* for $N \leqslant 11$. Table 1 summarizes the results. The even N are not included since they agree with the previous odd N.

TABLE 1

N	Λ_N^*	λ solution	$a_n(\lambda)$ (values rounded)
3	2	2	$a_1 = \frac{2}{3}$, $a_4 = \frac{1}{3}$
5	$3.11713\ldots$	3.11714	$a_1 = .31863$, $a_2 = 4.2756 \times 10^{-6}$, $a_3 = .1775$, $a_4 = .42868$, $a_7 = .050237$, $a_8 = .024955$.
7	$4.143770\ldots$	4.143771	$a_1 = .14772$, $a_3 = .19838$, $a_4 = .39981$, $a_6 = .02365$, $a_7 = .21883$, $a_8 = 5.1699 \times 10^{-8}$, $a_{11} = .011365$, $a_{12} = 2.4362 \times 10^{-4}$.
9	$5.238078\ldots$	5.238079	$a_1 = .061289$, $a_3 = .15968$, $a_4 = .28304$, $a_5 = 6.4803 \times 10^{-7}$, $a_6 = .16019$, $a_7 = .25862$, $a_{10} = .061939$, $a_{11} = .01358$, $a_{14} = 6.2149 \times 10^{-4}$, $a_{15} = .0011129$.
11	$6.291643\ldots$	6.2916435	$a_1 = .025323$, $a_3 = .10443$, $a_4 = .18339$, $a_5 = 8.761 \times 10^{-8}$, $a_6 = .23838$, $a_7 = .23595$, $a_9 = .05441$, $a_{10} = .13743$, $a_{13} = .010835$, $a_{14} = .0089696$, $a_{18} = 1.7866 \times 10^{-4}$, $a_{19} = 7.9083 \times 10^{-5}$.

7. Proof of Theorem 4. In this section and the next we ignore all X and ε dependences, since these have no effect on the proofs. We write $a_n(\lambda, X)$ as a_n (or $a_n(\lambda)$ if we wish to emphasize the λ dependence).

We first prove (6.2). The result $\Lambda_1^* = 1$ is trivial, and $\Lambda_2^* = 1$ follows from noting the system may be solved for $\lambda = 1 + \varepsilon$, $\varepsilon > 0$, with solution a_1, a_2, and a_r, where $r = r(\varepsilon)$ is any sufficiently large number. Now suppose the first three moments are

Poisson. In place of (5.5) we have

(7.1) $$\sum_{n=1}^{\infty} a_n = 1, \qquad \sum_{n=1}^{\infty} na_n = \lambda, \qquad \sum_{n=1}^{\infty} n^2 a_n = \lambda + \lambda^2,$$

$$\sum_{n=1}^{\infty} n^3 a_n = \lambda + 3\lambda^2 + \lambda^3.$$

On recombining we have

$$\sum_{n=1}^{\infty} (n-1) a_n = \lambda - 1, \qquad \sum_{n=1}^{\infty} (n-1)^2 a_n = \lambda^2 - \lambda + 1,$$

$$\sum_{n=1}^{\infty} (n-1)^3 a_n = \lambda^3 + \lambda - 1;$$

substituting into (5.7) and simplifying we obtain

$$0 \leqslant \lambda^2 (\lambda - 2),$$

whence $\lambda \geqslant 2$. Since $a_1 = \frac{2}{3}$, $a_4 = \frac{1}{3}$ satisfy (7.1) when $\lambda = 2$, we conclude $\Lambda_3^* = 2$.

To prove (6.3) and (6.4), we make use of the well-known relation for the Poisson distribution

(7.2) $$\sum_{n=0}^{\infty} n(n-1) \cdots (n-k+1) a_n(\lambda) = \lambda^k, \qquad k \geqslant 1.$$

Assuming (1.10) holds for $k = 1, 2, \ldots, m$, it is clear that (4.3) can be used to prove (7.2) for $k = 1, 2, \ldots, m$, and conversely (7.2) can be used to obtain (1.10). Hence, we may use (7.2) in place of (1.10) and (4.3). Now define

(7.3) $$f(z) = \sum_{n=0}^{\infty} a_n z^n;$$

here the sum is actually finite since $a_n = 0$ for $n \geqslant^s X$. Assume (7.2) holds for $k = 1, 2, \ldots m$.

By Taylor's theorem with remainder we have

$$f(0) = \sum_{n=0}^{m} \frac{f^{(n)}(1)(-1)^n}{n!} + \frac{f^{(m+1)}(\xi)}{(m+1)!}(-1)^{m+1},$$

where ξ is some number $0 \leqslant \xi \leqslant 1$. Now $f(0) = a_0$, $f^{(n)}(1) = \lambda^n$ for $n \leqslant m$, and hence

$$a_0(\lambda) = \sum_{n=0}^{m} \frac{(-\lambda)^n}{n!} + \frac{(-1)^{m+1} f^{(m+1)}(\xi)}{(m+1)!}.$$

If we had assumed (7.2) holds for all fixed k, then $a_0(\lambda) \sim e^{-\lambda}$ on taking $k = k(X) \to \infty$ as $X \to \infty$. To prove (6.4), we take $M = 2N + 1$, and note trivially $f^{(2N+2)}(\xi) \geqslant 0$. Hence

$$a_0(\lambda) \geqslant \sum_{n=0}^{2N+1} \frac{(-\lambda)^n}{n!}.$$

This proves (6.4).

To prove (6.3), we show the polynomial

$$(7.4) \qquad P_k(\lambda) = \sum_{n=0}^{k} \frac{(-\lambda)^n}{n!}$$

is positive for $\lambda < .278(k + 2)$. By Taylor's theorem

$$e^{-\lambda} = P_k(\lambda) + \frac{(-\lambda)^{k+1} e^{-\xi}}{(k+1)!} \qquad \text{for some } 0 \leqslant \xi \leqslant \lambda.$$

Hence for k even $P_k(\lambda) > e^{-\lambda} > 0$ for all $\lambda \geqslant 0$, while for k odd we see $P_k(\lambda)$ is strictly decreasing since $P'_k(\lambda) = -P_{k-1}(\lambda)$. Hence for odd k, $P_k(\lambda)$ will have exactly one root. To locate the root, note $P_k(\lambda) > 0$ by the Taylor expansion if $e^{-\lambda} > (\lambda^{k+1})/(k+1)!$ or $(k+1)\log\lambda + \lambda < \log(k+1)!$. Now, by Stirling's formula [13], for $m \geqslant 1$, we have

$$(7.5) \quad \log \Gamma(m) = \left(m - \tfrac{1}{2}\right)\log m - m + \tfrac{1}{2}\log 2\pi + w(m), \qquad |w(m)| \leqslant 1/12m.$$

Using $\Gamma(m) = (m - 1)!$, and taking $k = m - 2$, $\lambda = \theta m$, we find the above condition is satisfied if

$$m(1 + \theta + \log\theta) < \log(\sqrt{2\pi m}\,\theta) - 1/12m.$$

Now, $1 + \theta + \log\theta = 0$ has the solution $\theta = .27846\ldots$, hence the left side is negative for $\theta \leqslant .27846$. However, the right side is positive if $m \geqslant 5$ and $\theta \geqslant \tfrac{1}{5}$. Hence

$$P_k(\lambda) > 0 \quad \text{if } \lambda < .27846(k + 2) \text{ and } k = 3, 5, 7, \ldots.$$

This proves (6.3). A similar argument can be used to show $P_k(\lambda)$ is negative if $\lambda > ((k + 1)!)^{1/(k+1)} \sim k/e$. The first few roots $P_k(\lambda_k) = 0$ are $\lambda_3 = 1.596\ldots$, $\lambda_5 = 2.180\ldots$, $\lambda_7 = 2.759\ldots$.

8. Proof of Theorem 2. We now give a different method for finding lower bounds for $a_0(\lambda, X)$, and prove Theorem 2. Starting from (4.4), we have

$$(8.1) \qquad a_0(\lambda) = 1 - \sum_{n=1}^{\infty} a_n(\lambda) \geqslant 1 - \sum_{n=1}^{\infty} p(n) a_n(\lambda),$$

where $p(n)$ is any function such that $p(n) \geqslant 1$ for $n = 1, 2, 3, \ldots$. Suppose $p(n)$ is a polynomial with $p(0) = 0$, say $p(n) = \sum_{j=1}^{N} b_j n^j$. It is clear that in this case we can evaluate the right-hand side of (8.1) by using (1.10) and (4.3) for $k = 1, 2, 3, \ldots, N$. Specifically, suppose

$$(8.2) \quad p(n) = \sum_{k=1}^{N} C_k n(n - 1) \cdots (n - k + 1), \quad p(n) \geqslant 1 \text{ for } n = 1, 2, 3, \ldots.$$

Then, assuming (1.10) holds for $k = 1, 2, \ldots, N$, by (7.2) we have

$$(8.3) \qquad a_0(\lambda) \geqslant 1 - \sum_{k=1}^{N} C_k \lambda^k.$$

We now prove Theorem 2. Let
$$p(n) = \tfrac{1}{12}(n-1)(n-3)(n-4) + 1.$$
Clearly $p(n) \geq 1$ for $n = 1, 2, 3, \ldots$. Putting $p(n)$ into the form (8.2) we have
$$p(n) = n - \tfrac{5}{12}n(n-1) + \tfrac{1}{12}n(n-1)(n-2),$$
which by (8.3) proves (1.13). Since our $p(n)$ is of degree 3, we have used (1.10) for $k = 1, 2,$ and 3 (note: $k = 1$ is the prime number theorem (see equation (5.1))). Equation (1.14) follows from (1.13) by using (3.10).

We mention that either of the polynomials
$$p(n) = \tfrac{1}{20}(n-1)(n-4)(n-5) + 1 \quad \text{or} \quad p(n) = \tfrac{1}{16}(n-1)(n-4)^2 + 1$$
gives $\Lambda_3^* \geq 2$, but are not as large as the $p(n)$ used to prove Theorem 2.

In the general case, the C_k's in (8.2) may be readily computed from the values of $p(n)$ at $n = 0, 1, 2, \ldots, N$ by the Newton forward divided-difference formula. In this case

(8.4)
$$C_k = \frac{1}{k!}\Delta^k p(0),$$

where $\Delta^k p(0)$ is determined by the formula $\Delta^1 p(a) = p(a+1) - p(a)$, $\Delta^2 p(a) = \Delta^1 p(a+1) - \Delta^1 p(a)$, and in general $\Delta^k p(a) = \Delta^{k-1}p(a+1) - \Delta^{k-1}p(a)$. The simplest choice for $p(n)$ is

(8.5)
$$p(n) = \frac{(n-1)(n-2)\cdots(n-2N-1)}{(2N+1)!} + 1, \qquad N \geq 1.$$

In this case $p(0) = 0$, $p(1) = p(2) = \cdots = p(2N+1) = 1$, and hence we have $\Delta^1 p(0) = 1$, $\Delta^1 p(j) = 0$ for $j = 1, 2, \ldots, N+1$; $\Delta^2 p(0) = -1$, $\Delta^2 p(j) = 0$ for $j = 1, 2, \ldots, 2N+1$; and so on. Thus $\Delta^k p(0) = (-1)^{k-1}$ and by (8.4) $C_k = (-1)^{k-1}/k!$ which by (8.3) gives
$$a_0 \geq 1 - \sum_{k=1}^{2N+1}(-1)^{k-1}\lambda^k/k! = \sum_{k=0}^{2N+1}(-\lambda)^k/k!$$
This gives another proof of (6.4).

The question of finding optimal polynomials appears to be quite complicated. To begin with, we can replace (8.5) by the polynomial
(8.6)
$$p(n) = \frac{(n-1)(n-n_1)(n-n_1-1)(n-n_2)(n-n_2-1)\cdots(n-n_N)(n-n_N-1)}{(n_1)(n_1+1)(n_2)(n_2+1)\cdots(n_N)(n_N+1)} + 1,$$

where $n_1 < n_2 < n_3 \cdots < n_N$ are arbitrary positive integers. This polynomial of degree $2N + 1$ is easily seen to satisfy $p(n) \geq 1$ for all $n \geq 1$. It might also be noted that the solutions in Table 1 tend to have the a's occurring in consecutive pairs, which agrees with how $p(n)$ is constructed. As an example, consider

(8.7)
$$p(n) = \frac{(n-1)(n-3)(n-4)(n-6)(n-7)}{3 \cdot 4 \cdot 6 \cdot 7} + 1,$$

which gives

(8.8)
$$a_0(\lambda) \geq 1 - \lambda + \frac{29}{63}\lambda^2 - \frac{8}{63}\lambda^3 + \frac{11}{504}\lambda^4 - \frac{1}{504}\lambda^5.$$

This polynomial has its root at $\lambda = 3$, proving

$$(8.9) \qquad\qquad\qquad\qquad \Lambda_5^* \geqslant 3,$$

and gives a lower bound of $(1.782\ldots)X \log X$ in (1.14), on the assumption of now having 5 Poisson moments.

9. Gaps between consecutive primes. Up to this point, we have only considered the problem of finding a positive proportion of intervals of length λ with no primes contained in them; in other words showing $a_0(\lambda, X)$ is positive. In terms of consecutive primes, we have by equation (1.3):

$$(9.1) \qquad a_0(\lambda, X) \sim \frac{1}{X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ d_n \geqslant \lambda \log p_n}} (d_n - \lambda \log n), \quad \text{as } X \to \infty.$$

Consequently, we see that, for $\varepsilon > 0$,

$$(9.2) \qquad\qquad\qquad d_n > (\Lambda - \varepsilon) \log p_n,$$

for infinitely many n. However, it cannot be shown that a positive proportion of the gaps d_n have this property. Recall $G(\lambda, X)$ from (2.2):

$$(9.3) \qquad\qquad G(\lambda, X) = \frac{1}{X/\log X} \sum_{\substack{X \leqslant p_n \leqslant 2X \\ d_n \geqslant \lambda \log p_n}} 1.$$

We wish to find λ such that $G(\lambda, X)$ is positive for all sufficiently large X. The reason a lower bound for $a_0(\lambda, X)$ in (9.1) does not lead to a lower bound in (9.3) is that there could be $CX/d(X)\log X$ gaps d_n of length $\geqslant d(X)\log X$, where $d(X) \to \infty$ as $X \to \infty$, which will make $a_0(\lambda, X)$ positive (for *any* fixed λ), yet only contribute $o(1)$ to $G(\lambda, X)$. One cannot rule out the possibility that these are the only larger than average gaps d_n that occur.

We can, however, conditionally handle this question. We prove

THEOREM 5. *Assume equation* (1.10) *holds for* $k = 2$ ($m_2(\lambda, X) \sim \lambda + \lambda^2$). *Then, if* $a_0(\lambda, X) \geqslant c(\lambda) - \varepsilon > 0$ *for any* $\varepsilon > 0$ *and all X sufficiently large,*

$$(9.4) \qquad\qquad\qquad G(\lambda, X) \gg c(\lambda).$$

Hence $G(\lambda, X)$ *is positive for all sufficiently large X if* $\lambda < \Lambda$.

PROOF. Returning to the notation of §3, we have for any $0 < \lambda_1 < \lambda_2$,

$$a_0(\lambda_1) - a_0(\lambda_2) = \int_{\lambda_1}^\infty (u - \lambda_1)\rho(u)\, du - \int_{\lambda_2}^\infty (u - \lambda_2)\rho(u)\, du$$

$$= \int_{\lambda_1}^{\lambda_2} (u - \lambda_1)\rho(u)\, du + (\lambda_2 - \lambda_1)\int_{\lambda_2}^\infty \rho(u)\, du$$

$$= \int_{\lambda_1}^{\lambda_2} (u - \lambda_1)\rho(u)\, du + (\lambda_2 - \lambda_1)G(\lambda_2).$$

Hence

$$G(\lambda_2) = \frac{a_0(\lambda_1) - a_0(\lambda_2)}{\lambda_2 - \lambda_1} - \frac{1}{\lambda_2 - \lambda_1}\int_{\lambda_1}^{\lambda_2} (u - \lambda_1)\rho(u)\, du.$$

Similarly,

$$G(\lambda_1) = \frac{a_0(\lambda_1) - a_0(\lambda_2)}{\lambda_2 - \lambda_1} - \frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} (u - \lambda_2)\rho(u)\,du.$$

We conclude, for $0 < \lambda_1 < \lambda_2$,

$$(9.5) \qquad\qquad G(\lambda_2) \leqslant \frac{a_0(\lambda_1) - a_0(\lambda_2)}{\lambda_2 - \lambda_1} \leqslant G(\lambda_1).$$

We now show that, assuming $M_2(\lambda, X) \sim \lambda + \lambda^2$, $a_0(\lambda) \to 0$ as $\lambda \to \infty$. Hence we may pick λ_2 large enough so that $a_0(\lambda_2) < \frac{1}{2}c(\lambda_1) \leqslant \frac{1}{2}a_0(\lambda_1)$ which by (9.5) proves (9.4) and the theorem. Now, by (4.3) and (4.4) we have

$$\sum_{n=1}^{\infty} a_n = 1 - a_0, \qquad \sum_{n=1}^{\infty} na_n \sim \lambda, \qquad \sum_{n=1}^{\infty} n^2 a_n \sim \lambda + \lambda^2.$$

By Cauchy's inequality

$$\left(\sum_{n=1}^{\infty} na_n \right)^2 \leqslant \left(\sum_{n=1}^{\infty} a_n \right) \left(\sum_{n=1}^{\infty} n^2 a_n \right),$$

and hence we obtain, for $\varepsilon > 0$ and X sufficiently large,

$$(9.6) \qquad\qquad a_0(\lambda, X) \leqslant \frac{1}{1 + \lambda} + \varepsilon.$$

This shows $a_0(\lambda) \to 0$ as $\lambda \to \infty$, and completes the proof.

REMARKS. $a_0(\lambda) \to 0$ as $\lambda \to \infty$ was proven by Heath-Brown [11] assuming the Riemann Hypothesis and a form of the pair correlation conjecture for zeros of the zeta function. Under the same conjectures, it has been proven in [8] that $M_2(\lambda, X) \sim \lambda + \lambda^2$ as $X \to \infty$, so (9.6) also follows from these conjectures. In Theorems 2 and 4 we have *assumed* (1.10) for $k = 2$ is true. Hence, this gives a positive proportion of gaps $d_n \geqslant (\Lambda - \varepsilon) \log p_n$ for the Λ obtained there. At present, this cannot be shown unconditionally for any $\Lambda > 1$.

REFERENCES

1. E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. London Ser. A **293** (1966), 1–18.

2. A. Y. Cheer and D. A. Goldston, *A moment method for primes in short intervals*, C. R. Math. Rep. Acad. Sci. Canada, **11**, (1987) No. 2, 101–106.

3. J. R. Chen, *On the Goldbach's problem and the sieve methods*, Sci. Sinica **21** (1978), 701–739.

4. P. Erdös, *The difference of consecutive primes*, Duke Math. J. **6** (1940), 438–441.

5. E. Fouvry and F. Grupp, *On the switching principle in sieve theory*, J. Reine Angew. Math. **370** (1986), 101–126.

6. P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematica **23** (1976), 4–9.

7. D. A. Goldston, *The second moment for prime numbers*, Quart. J. Math. Oxford (2), **35** (1984), 153–163.

8. D. A. Goldston and H. L. Montgomery, *Pair correlation of zeros and primes in short intervals*, Analytic Number Theory and Diophantine Problem, Birkhäuser, Boston, Mass., 1987.

9. H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, 1974.

10. D. R. Heath-Brown, *The differences between consecutive primes*. III, J. London Math. Soc. [2] **20** (1979), 177–178.

11. _____, *Gaps between primes, and the pair correlation of zeros of the zeta-functions*, Acta Arith. **41** (1982), 85–99.

12. M. Huxley, *Estimating* $\lim\inf(p_{n+r} - p_n)/\log p_n$, Séminaire de Théorie des Nombres, Anneé 1980–1981, no. 18, 1–10.

13. N. N. Lebedev, *Special functions and their applications*, Dover, New York, 1972.

14. K. Prachar, *Bemerkungen über Primzahlen in kurzen Reihen*, Acta Arith. **44** (1984), 175–180.

15. _____, *Primzahlverteilung*, Springer-Verlag, Berlin, 1957.

16. A. Selberg, *On the normal density of primes in small intervals and the difference between consecutive primes*, Arch. Math. Naturvid. **47** (1943), No. 6, 87–105.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS, CALIFORNIA 95616

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, SAN JOSE STATE UNIVERSITY, SAN JOSE, CALIFORNIA 95192