Reprinted from

THE QUARTERLY JOURNAL OF

MATHEMATICS

OXFORD SERIES

....

Volume 6 No. 22 June 1935

ON THE DIFFERENCE OF CONSECUTIVE PRIMES

By P. ERDŐS (*Budapest*)

[Received 22 September 1934]

We consider here the question of the intervals between two consecutive prime numbers. Let p_n denote the nth prime. Backlund* proved that, for any positive ϵ and an infinity of n,

$$p_{n+1}-p_n > (2-\epsilon)\log p_n.$$

Brauer and Zeitz† showed that $2-\epsilon$ could be replaced by $4-\epsilon$. Westzynthius‡ proved that for an infinity of n

$$p_{n+1}-p_n > \frac{2\log p_n \operatorname{logloglog} p_n}{\operatorname{loglogloglog} p_n};$$

and Ricci§ has just shown that this can be improved to

$$p_{n+1}-p_n > c\log p_n \operatorname{logloglog} p_n$$

for an infinity of n and with a certain constant c. By increasing the precision of Brauer and Zeitz's method, I shall prove

THEOREM I. *For a certain positive constant c_1 and an infinity of values of n,*

$$p_{n+1}-p_n > \frac{c_1 \log p_n \operatorname{loglog} p_n}{(\operatorname{logloglog} p_n)^2}.$$

We reduce our problem to the proof of the following theorem.

THEOREM II. *For a certain positive constant c_2, we can find $c_2 p_n \log p_n/(\operatorname{loglog} p_n)^2$ consecutive integers so that no one of them is relatively prime to the product $p_1 p_2 \dots p_n$, i.e. each of these integers is divisible by at least one of the primes $p_1, p_2, \dots, p_n$.*

* R. J. Backlund, 'Über die Differenzen zwischen den Zahlen, die zu den n ersten Primzahlen teilerfremd sind': *Commentationes in honorem Ernesti Leonardi Lindelöf*, Helsinki, 1929.

† A. Brauer u. H. Zeitz, 'Über eine zahlentheoretische Behauptung von Legendre': *Sitz. Berliner Math. Ges.* 29 (1930), 116–25; H. Zeitz, *Elementare Betrachtung über eine zahlentheoretische Behauptung von Legendre* (Berlin 1930, Privatdruck).

‡ 'Über die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind', *Comm. Phys.-Math., Helsingfors*, (5) 25 (1931).

§ 'Ricerche aritmetiche sui polinomi II (Intorno a una proposizione non vera di Legendre)': *Rend. Circ. Mat. di Palermo*, 58 (1934).

ON THE DIFFERENCE OF CONSECUTIVE PRIMES

We require some lemmas.

LEMMA 1. *Let m be any positive integer greater than 1, x and y any numbers such that $1 \leqslant x < y < m$, and N the number of primes p less than or equal to m such that $p+1$ is not divisible by any of the primes P, where $x \leqslant P \leqslant y$. Then*

$$N < \frac{c_3 m \log x}{\log m \log y},$$

where c_3 is a constant independent of m, x, and y.

We omit the proof since it is a direct application of the method of Brun.*

LEMMA 2. *If N_0 is the number of those integers not exceeding $p_n \log p_n$, each of whose greatest prime-factors is less than $p_n^{1/(20 \log\log p_n)}$, then $N_0 = o\{p_n/(\log p_n)^2\}$.*

We shall divide the integers we are considering into two classes: (i) those for each of which the number of different prime factors does not exceed $10 \log\log p_n$, and (ii) those for each of which the number of different prime factors exceeds $10 \log\log p_n$. Let the number of integers in these two classes be N_1 and N_2 respectively; then $N_0 = N_1 + N_2$.

If Q is a prime not exceeding $p_n^{1/(20 \log\log p_n)}$, then

$$Q^x > p_n \log p_n \quad \text{if} \quad x > (2 \log p_n)/(\log 2).$$

Hence the number of such primes and powers of such primes less than $p_n \log p_n$ is certainly less than

$$\frac{2 \log p_n}{\log 2} p_n^{1/(20 \log\log p_n)}.$$

But every integer of the class (i) is a product of not more than $10 \log\log p_n$ factors, each being one of these primes or powers. Hence

$$N_1 < \left(\frac{2 \log p_n}{\log 2} p_n^{1/(20 \log\log p_n)}\right)^{10 \log\log p_n}$$
$$= p_n^{\frac{1}{2}} \left(\frac{2 \log p_n}{\log 2}\right)^{10 \log\log p_n} = o\left\{\frac{p_n}{(\log p_n)^2}\right\}.$$

Let $d(k)$ be the number of divisors of k. If k is an integer of the second class, k has more than $10 \log\log p_n$ different prime factors and so

$$d(k) > 2^{10 \log\log p_n} > (\log p_n)^5.$$

* V. Brun, 'Le crible d'Ératosthène et le théorème de Goldbach': *Vidensk. Selsk. Skrifter, Mat.-naturv. Kl. Kristiania*, 3 (1920), and *Comptes Rendus*, 168 (1919). See also 'La série $\frac{1}{5}+\frac{1}{7}+\ldots$ où les dénominateurs sont "nombres premiers jumeaux" est convergente ou finie', *Bull. Soc. Math.* (2) 43 (1919), 1–9.

Since
$$\sum_{l=1}^{p_n \log p_n} d(l) < 4p_n(\log p_n)^2$$

for sufficiently large n, we have
$$N_2 = o\left\{\frac{p_n}{(\log p_n)^2}\right\}.$$

LEMMA 3. *We can find a constant c_4 so that the number of primes p, less than $c_4 p_n \log p_n/(\log\log p_n)^2$ and such that $p+1$ is not divisible by any prime between $\log p_n$ and $p_n^{1/(20\log\log p_n)}$, is less than $p_n/4\log p_n$.*

We obtain this lemma immediately from Lemma 1 on putting
$$m = \frac{c_4 p_n \log p_n}{(\log\log p_n)^2}, \qquad x = \log p_n, \qquad y = p_n^{1/(20\log\log p_n)}.$$

We return now to Theorem II. We denote by q, r, s, t the primes satisfying the inequalities
$$1 < q \leqslant \log p_n, \qquad \log p_n < r \leqslant p_n^{1/(20\log\log p_n)},$$
$$p_n^{1/(20\log\log p_n)} < s \leqslant \tfrac{1}{2}p_n, \qquad \tfrac{1}{2}p_n < t \leqslant p_n.$$
We denote by $a_1, a_2, \ldots, a_k$ the two sets of integers not greater than $p_n \log p_n$, namely (i) the prime numbers lying between $\frac{1}{2}p_n$ and $c_4 p_n \log p_n/(\log\log p_n)^2$ and not congruent to -1 to any modulus r, (ii) the integers not exceeding $p_n \log p_n$ whose prime factors are included only among the r. Some of the a's may be t's.

LEMMA 4. *The number of the t's is greater than k the number of the a's, if p_n is large enough.*

From Lemmas 2, 3,
$$k < \frac{1}{4}\frac{p_n}{\log p_n} + o\left\{\frac{p_n}{(\log p_n)^2}\right\}.$$
The number of the t's is greater than $\frac{1}{3}p_n/\log p_n$ for large p_n, as is evident from the prime-number theorem, and as can also be proved by elementary methods. This proves the lemma.

We now determine an integer z such that for all q, r, s,
$$0 < z < p_1 p_2 \cdots p_n,$$
$$z \equiv 0 \pmod{q}, \qquad z \equiv 1 \pmod{r}, \qquad z \equiv 0 \pmod{s},$$
$$z + a_i \equiv 0 \pmod{t_i} \qquad (i = 1, 2, \ldots, k).$$

By Lemma 4, the last congruence is always possible, for, as there are more t's than a's, a case such as $z+a_1 \equiv 0 \pmod{t}$, $z+a_2 \equiv 0 \pmod{t}$ cannot occur.

We now show that, if l is any integer such that

$$0 < l < c_2 p_n \log p_n / (\log\log p_n)^2,$$

then no one of the integers

$$z,\ z+1,\ z+2, \ldots,\ z+l$$

is relatively prime to $p_1 p_2 \ldots p_n$.

Now any integer b $(0 < b < l)$ can be placed in one at least of the four following classes:

(i) $b \equiv 0 \pmod{q}$, for some q;

(ii) $b \equiv -1 \pmod{r}$, for some r;

(iii) $b \equiv 0 \pmod{s}$, for some s;

(iv) b is an a_i.

For b cannot be divisible by an r and by a prime greater than $\frac{1}{2}p_n$, since if this were so we should have

$$b > \tfrac{1}{2}p_n r > \tfrac{1}{2}p_n \log p_n > l,$$

for sufficiently large n. Hence, if b does not satisfy (i) or (iii), b is either a product of primes r only, and so satisfies (iv), or b is not divisible by any q, r, s. In the latter case, b must be a prime, for otherwise

$$b > (\tfrac{1}{2}p_n)^2 > l,$$

for sufficiently large n. Since, then, b is a prime between

$$\tfrac{1}{2}p_n \quad \text{and} \quad \frac{c_2 p_n \log p_n}{(\log\log p_n)^2},$$

b is either an a_i, or b satisfies (ii).

It is now clear that $z+b$ is not relatively prime to $p_1 p_2 \ldots p_n$, if

$$b < c_2 p_n \log p_n / (\log\log p_n)^2.$$

Hence also, if $p_1, p_2, \ldots, p_n$ are the primes not exceeding x, say, $z+b$ is not relatively prime to $p_1 p_2 \ldots p_n$, if $b < c_5 x \log x / (\log\log x)^2$, where c_3 is an appropriate constant independent of x. This is clear from the first case on noticing that, by Bertrand's theorem, $p_n \geqslant \frac{1}{2}x$.

We return to the main problem. Take $x = \frac{1}{2}\log p_n$. Then the product of the primes not exceeding x is less than $\frac{1}{2}p_n$ for large p_n by the prime-number theorem, or also by elementary methods. By Theorem II, since now $b < \frac{1}{2}p_n$, we can find K consecutive integers less than p_n, where

$$K = \frac{c_5 \log p_n \log\log p_n}{(\log\log\log p_n)^2},$$

each of which is divisible by a prime less than $\frac{1}{2}\log p_n$. Hence there are at least $K-\frac{1}{2}\log p_n$ $(>\frac{1}{2}K)$ consecutive integers which are not primes.

Thus we have proved that at least one of the intervals between successive primes less than p_n is always of length not less than $c\log p_n \operatorname{loglog} p_n/(\operatorname{logloglog} p_n)^2$ for large p_n and an appropriate constant c. Since this expression is an increasing function of n, it follows immediately that for an infinity of n,

$$p_{n+1}-p_n > \frac{c_1 \log p_n \operatorname{loglog} p_n}{(\operatorname{logloglog} p_n)^2}.$$

I wish to take this opportunity of expressing my gratitude to Professor Mordell for so kindly having helped me in preparing my manuscript.