ACTA ARITHMETICA
VIII (1963)

On the difference of consecutive numbers prime to n

by

C. Hooley (Bristol)

1. Introduction. In 1940 Erdös [1] made the following conjecture. Let n be any positive integer greater than 1, and let $a_1, a_2, \ldots, a_{\varphi(n)}$, where $a_i < a_{i+1}$, be the $\varphi(n)$ integers not exceeding n that are relatively prime to n. Then

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1}-a_i)^2 = O\left(\frac{n^2}{\varphi(n)}\right).$$

In this paper we prove that for any fixed real number a such that $1 \leqslant a < 2$, we have

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1}-a_i)^a = O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\}.$$

A proof of Erdös' conjecture would involve the extension of this result to the case $a = 2$. Here, however, the method just fails, but we are able to modify the argument in order to prove the weaker result

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1}-a_i)^2 = O(n\log\log n).$$

Throughout this paper the constants implied by the O notation depend at most on a.

2. Estimation of $G(n, h)$. Let

$$f(m) = \begin{cases} 1, & \text{if} \quad (m, n) = 1, \\ 0, & \text{if} \quad (m, n) > 1, \end{cases}$$

and let $F(m, h)$ be defined for any positive integer h by

$$F(m, h) = \sum_{r=m}^{m+h-1} f(r).$$

We first obtain an upper bound for the sum

$$G(n,h)=\sum_{m=1}^{n}\left(F(m,h)-h\frac{\varphi(n)}{n}\right)^2.$$

Erdös [2] has recently given an upper bound for a similar sum in connection with another problem related to the Euler φ function, and has outlined a proof. In this paper, however, the upper bound needs to be sharper and to hold uniformly with respect to h, so it has been considered advisable to give the estimations in some detail.

Since

$$\sum_{m=1}^{n}F(m,h)=h\sum_{r=1}^{n}f(r)=h\varphi(n),$$

we have

(1) $$G(n,h)=\sum_{m=1}^{n}F^2(m,h)-h^2\frac{\varphi^2(n)}{n}.$$

Now

$$\sum_{m=1}^{n}F^2(m,h)=\sum_{m=1}^{n}\sum_{r,s=m}^{m+h-1}f(r)f(s)$$
$$=h\sum_{r=1}^{n}f^2(r)+2\sum_{k=1}^{h-1}(h-k)\sum_{r=1}^{n}f(r)f(r+k),$$

since $f(l)=f(l+n)$ for any integer l. Hence

(2) $$\sum_{m=1}^{n}F^2(m,h)=h\varphi(n)+2\sum_{k=1}^{h-1}(h-k)\sum_{r=1}^{n}f(r)f(r+k)$$
$$=h\varphi(n)+S(n,h),\quad \text{say}.$$

We now estimate $S(n,h)$. The proof is given for odd values of n only; the proof for even values of n is similar in principle, but different in minor details. A straightforward application of the sieve method gives for n of either parity

$$\sum_{r=1}^{n}f(r)f(r+k)=n\prod_{\substack{p|n\\ p|k}}\left(1-\frac{1}{p}\right)\prod_{\substack{p|n\\ p\nmid k}}\left(1-\frac{2}{p}\right).$$

Therefore, if n be odd,

$$\sum_{r=1}^{n}f(r)f(r+k)=n\prod_{p|n}\left(1-\frac{2}{p}\right)\prod_{\substack{p|n\\ p|k}}\left(1+\frac{1}{p-2}\right)$$
$$=n\psi(n)\,\theta[(n,k)],\quad \text{say},$$

where (n,k) is the highest common factor of n and k. Now for any square free number $m=p_1p_2\dots p_t$ define $\varrho(m)$ to be $(p_1-2)(p_2-2)\dots(p_t-2)$. Then for any odd integer N we have

$$\theta(N)=\sum_{d|N}\frac{|\mu(d)|}{\varrho(d)}.$$

Hence

$$\sum_{k\leqslant x}\theta[(n,k)]=\sum_{k\leqslant x}\sum_{\substack{d|n\\ d|k}}\frac{|\mu(d)|}{\varrho(d)}$$
$$=\sum_{d|n}\frac{|\mu(d)|}{\varrho(d)}\sum_{k'\leqslant x/d}1$$
$$=\sum_{\substack{d|n\\ (\mathrm{A})}}\frac{|\mu(d)|}{\varrho(d)}\left(\frac{x}{d}+O(1)\right),$$

where (A) denotes the condition that d have no prime factors exceeding x. Thus

(3) $$\sum_{k\leqslant x}\theta[(n,k)]=x\sum_{d|n}\frac{|\mu(d)|}{d\varrho(d)}+O\left(x\sum_{d\geqslant x}\frac{|\mu(d)|}{d\varrho(d)}\right)+O\left(\sum_{\substack{d|n\\ (\mathrm{A})}}\frac{|\mu(d)|}{\varrho(d)}\right)$$
$$=x\prod_{p|n}\left(1+\frac{1}{p(p-2)}\right)+O(1)+O\left(\log(x+2)\right),$$

since

$$\frac{|\mu(d)|}{\varrho(d)}=O\left(\frac{\sigma_{-1}^2(d)}{d}\right)\quad\text{and}\quad\prod_{2<p\leqslant x}\left(1+\frac{1}{p-2}\right)=O\left(\log(x+2)\right).$$

Hence, integrating (3) in the range $0\leqslant x\leqslant h$, we have

$$2\sum_{k=1}^{h-1}(h-k)\,\theta[(n,k)]=h^2\prod_{p|n}\left(1+\frac{1}{p(p-2)}\right)+O(h\log 2h),$$

and so

(4) $$S(n,h)=h^2n\psi(n)\prod_{p|n}\left(1+\frac{1}{p(p-2)}\right)+O\left(n\psi(n)h\log 2h\right)$$
$$=h^2n\prod_{p|n}\left(1-\frac{1}{p}\right)^2+O\left(n\psi(n)h\log 2h\right)$$
$$=h^2\frac{\varphi^2(n)}{n}+O\left(h\log 2h\frac{\varphi^2(n)}{n}\right).$$

This result also holds for n even.

We deduce from (1), (2), and (4),

(5) $$G(n,h)=h\varphi(n)+O\left(h\log 2h\cdot\frac{\varphi^2(n)}{n}\right).$$

3. The final inequalities. Let $N_r \equiv N_r(n)$ be the number of intervals $a_{i+1}-a_i$ of length r, and let

$$S_r^{(t)} \equiv S_r^{(t)}(n) = N_r + 2^t N_{r+1} + 3^t N_{r+2} + \dots$$

Setting $h = r-1$ in (5) we have

$$(r-1)^2 \frac{\varphi^2(n)}{n^2} S_r^{(1)} \leqslant (r-1)\varphi(n) + O\left(r \log 2r \cdot \frac{\varphi^2(n)}{n}\right),$$

and therefore

$$S_r^{(1)} = O\left(\frac{1}{r}\cdot\frac{n^2}{\varphi(n)}\right) + O\left(\frac{\log 2r}{r}\cdot n\right). \tag{6}$$

Now

$$\sum_{l \geqslant r} N_l l^a = \sum_{l \geqslant r} (S_l^{(0)} - S_{l+1}^{(0)}) l^a \tag{7}$$

$$= O\left(r^a S_r^{(0)} + \sum_{l \geqslant r} S_l^{(0)} l^{a-1}\right)$$

$$= O\left(r^a S_r^{(0)} + r^{a-1} S_r^{(1)} + \sum_{l \geqslant r} S_l^{(1)} l^{a-2}\right).$$

Let $r = \left[\frac{n}{\varphi(n)}\right] + 1$. Then, since $S_r^{(0)} = O(\varphi(n))$ and $S_r^{(1)} = O(n)$, we have, by (6) and (7), when $1 \leqslant a < 2$,

$$\sum_{l \geqslant r} N_l l^a = O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\} + O\left\{\sum_{l \geqslant r} \frac{n^2}{\varphi(n)} l^{a-3}\right\} + O\left\{\sum_{l \geqslant r} n \log 2l \cdot l^{a-3}\right\} \tag{8}$$

$$= O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\} + O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-2} \log\left(\frac{2n}{\varphi(n)}\right)\right\} = O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\}.$$

Also

$$\sum_{l<r} N_l l^a \leqslant r^{a-1} \sum_{l<r} N_l l \leqslant r^{a-1} n = O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\}. \tag{9}$$

Combining (8) and (9) we obtain

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1}-a_i)^a = \sum_{l} N_l l^a = \sum_{l<r} N_l l^a + \sum_{l \geqslant r} N_l l^a = O\left\{n\left(\frac{n}{\varphi(n)}\right)^{a-1}\right\}.$$

We only sketch the procedure for the case $a = 2$. A standard application of Brun's method shows that there is a positive absolute constant C such that

$$F(m, h) > 0 \quad \text{for} \quad h > \log^C 2n = R, \quad \text{say.}$$

Hence $N_l = 0$ for $l > R+1$, and we obtain easily

$$\sum_{i=1}^{\varphi(n)-1} (a_{i+1}-a_i)^2 = \sum_{l} N_l l^2$$

$$= O\left(\frac{n^2}{\varphi(n)}\right) + O\left(\sum_{l \leqslant R+1} \frac{n^2}{\varphi(n)} l^{-1}\right) + O\left\{\sum_{l \leqslant R+1} n \log 2l \cdot l^{-1}\right\}$$

$$= O\left(\frac{n^2}{\varphi(n)} \log\log n\right) + O\left(n(\log\log n)^2\right)$$

$$= O\left(n(\log\log n)^2\right).$$

References

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

[2] — *On the integers relatively prime to n and on a number-theoretic function considered by Jacobsthal*, Math. Scand. 10 (1962), pp. 163-170.

BRISTOL UNIVERSITY, ENGLAND

Reçu par la Rédaction le 26. 1. 1963