

Complete prime subsets of consecutive integers

P. Erdős and J.L. Selfridge

Denote by $\pi(x)$ the number of primes not exceeding x .

A well known conjecture states that for $1 < y \leq x$,

$$(1) \quad \pi(x+y) \leq \pi(x) + \pi(y).$$

The proof of (1) unfortunately seems to be hopeless at present.

No methods seem to be available to prove it. Conceivably a counter-example could be found by using computers, but we believe that the conjecture is true. It has been proved for small values of y . Also for $x = y$ it is a classical result of Landau when x is large, and recently was proved for all x by Schoenfeld.

Hardy and Littlewood observed that Brun's sieve method gives

$$\pi(x+y) - \pi(x) < \frac{c_1 y}{\log y},$$

and Selberg proved by his improvement of Brun's method that

$$\pi(x+y) - \pi(x) < (2 + o(1)) \frac{y}{\log y}.$$

This result was recently strengthened by Montgomery to

$$\pi(x+y) - \pi(x) < \frac{2y}{\log y}.$$

It would be very interesting if 2 could be replaced by a smaller constant in the Montgomery-Selberg result.

Hardy and Littlewood put

$$\rho(y) = \limsup_{x \rightarrow \infty} (\pi(x+y) - \pi(x)) .$$

Presumably

$$\rho(y) - \pi(y) \rightarrow -\infty ,$$

but it seems to be very hard to get any result on $\rho(y)$. In view of this unsatisfactory situation P. Erdős [3] introduced the following related functions which probably are easier to investigate. Let $\{a_i\}$ be a sequence of integers, prime in pairs, with

$$(2) \quad 0 \leq n < a_1 < \dots < a_t \leq n+k , \quad (a_i, a_j) = 1 , \quad 1 \leq i < j \leq t .$$

The sequence is called *complete* if for every s in $n < s \leq n+k$, $(s, a_i) > 1$ for some i . Put

$$F(n,k) = \max t , \quad f(n,k) = \min t ,$$

where the maximum and minimum are taken with respect to all complete sequences satisfying (2). Clearly $F(n,k)$ and $f(n,k)$ are periodic functions of n (the period is a divisor of the product of primes less than k).

Now we shall consider the four functions

$$(3) \quad \max_n F(n,k) , \quad \min_n F(n,k) , \quad \max_n f(n,k) , \quad \min_n f(n,k) ,$$

and try to obtain some non-trivial results on these functions.

(In what follows max and min are always taken over all n).

Trivially

$$(4) \quad \min f(n, k) = 2 ,$$

for it suffices to put $n = P - k$, where P is the product of all primes less than k , and take $a_1 = n + k - 1$, $a_2 = a_1 + 1$.

Erdos thought that perhaps $\max F(n, k) \leq \pi(k) + 1$, but this conjecture is clearly wrong. We show first of all that for infinitely many k ,

$$(5) \quad \max F(n, k) > \pi(k) + \left(\frac{1}{2} + o(1)\right) \frac{k^{\frac{1}{2}}}{\log k} .$$

It immediately follows from the prime number theorem that for infinitely many k ,

$$(6) \quad \pi(k + k^{\frac{1}{2}}) > \pi(k) + \left(\frac{1}{2} + o(1)\right) \frac{k^{\frac{1}{2}}}{\log k} .$$

This inequality almost certainly holds for all k but even $\pi(k + k^{\frac{1}{2}}) - \pi(k) > 1$ seems hopeless at present, though it seems certain to hold for all $k \geq 117$.

Let k be an integer for which (6) holds, and put $n = [k^{\frac{1}{2}}]$. The a 's are the primes in $(k^{\frac{1}{2}}, k + k^{\frac{1}{2}})$, together with one power (in this interval) of each prime $p \leq k^{\frac{1}{2}}$. Thus for every k ,

$$\pi(k + k^{\frac{1}{2}}) \leq \max F(n, k) ,$$

and (6) implies (5).

We now prove a sharper result.

Theorem 1.

$$\max_n F(n, k) > \pi(k) + (\log 2 - \frac{1}{2} - o(1)) \left(\frac{k}{(\log k)^2} \right) .$$

Proof: Let $b = n + [\frac{1}{2}k] + 1$. We take $b \equiv 0 \pmod{p}$ for all $p < p_r$, where p_r is the least prime for which $2p_r + r > \pi(\frac{1}{2}k) + \pi(\frac{1}{2}k - \frac{1}{2}) + 3$. Now the set of a 's will consist of $b - 1$ and $b + 1$, together with every $b - p$ for which $p \leq \frac{1}{2}k$, and every $b + p$ for which $p_r \leq p \leq \frac{1}{2}k - \frac{1}{2}$. The total number of such a 's is $t = 2 + \pi(\frac{1}{2}k) + \pi(\frac{1}{2}k - \frac{1}{2}) - (r - 1)$. So if $p \geq p_r$, then $2p > t$ (by choice of p_r), and there must be some residue class \pmod{p} which contains at most one of the a 's. Thus we can choose $b \pmod{p}$ so that p will divide at most one of the a 's. Then these a 's are clearly pairwise relatively prime, and their number is greater than $2\pi(\frac{1}{2}k) - r$.

Using

$$\pi(x) = \frac{x}{\log x} + (1 + o(1)) \frac{x}{(\log x)^2} ,$$

we obtain

$$\max F(n, k) > 2\pi\left(\frac{k}{2}\right) - r > \frac{k}{\log k} + \frac{(\frac{1}{2} + \log 2)k}{(\log k)^2} - o\left(\frac{k}{(\log k)^2}\right)$$

$$\pi(k) + \frac{(\log 2 - \frac{1}{2})k}{(\log k)^2} - o\left(\frac{k}{(\log k)^2}\right),$$

which proves theorem 1.

By Selberg's sieve we easily obtain

$$\max F(n,k) < (2 + o(1)) \frac{k}{\log k}.$$

It appears likely that

$$(7) \quad \max F(n,k) = (1 + o(1)) \frac{k}{\log k};$$

a proof of this seems a very difficult problem.

Theorem 2. For every $\epsilon > 0$ and $k > k_0(\epsilon)$,

$$k^{1-\epsilon} < \min F(n,k) < \frac{c k (\log \log k)^2}{(\log k)^2 \log \log \log k}.$$

First we prove the upper bound. By a well known theorem of Rankin, for every k there is an n so that for every i , $1 \leq i \leq k$, $n+i$ has a prime factor not exceeding

$$\frac{c k (\log \log k)^2}{\log k \log \log \log k} = \ell.$$

For this n clearly $F(n,k) \leq \ell$, so the prime number theorem immediately implies the upper bound in Theorem 2.

An unpublished result of Rosser (see the forthcoming book

on sieve methods by Halberstam and Richert) implies that for every $\epsilon' > 0$, there is a constant $c = c(\epsilon')$ such that at least $c \frac{k}{\log k}$ integers between n and $n+k$ have all their prime factors greater than $k^{\frac{1}{2}-\epsilon'}$. Let these integers be $u_1 < \dots < u_s$, where $s \geq c \frac{k}{\log k}$. Then, if $\epsilon' < \frac{1}{6}$, u_i is divisible by at most two primes not exceeding k and these are both greater than $k^{\frac{1}{2}-\epsilon'}$. Hence for fixed u_i there are at most $2k^{\frac{1}{2}+\epsilon'} + 2$ integers j for which $(u_i, u_j) > 1$. Thus clearly

$$F(n, k) > \frac{s}{2k^{\frac{1}{2}+\epsilon'} + 2} > k^{\frac{1}{2}-\epsilon'}$$

which completes the proof of Theorem 2.

We would guess that the true order of magnitude of $\min F(n, k)$ is $\frac{k}{(\log k)^\alpha}$ for some α but this seems hopeless at present. It would be very interesting to prove

$$\min F(n, k) > k^{\frac{1}{2}+\epsilon}$$

Clearly both functions $\min F(n, k)$ and $\max F(n, k)$ are monotonic in k but $\max f(n, k)$ is certainly not monotonic, since

$$\max f(n, 6) = 3, \quad \max f(n, 5) = 4.$$

Possibly this is the only case where $\max f(n, k)$ is not monotonic. We computed $\max f(n, k)$ for $k \leq 45$. Perhaps

$$\max f(n,k) - \pi(k) \rightarrow -\infty,$$

but we have no nontrivial upper bound for $\max f(n,k)$. Denote by $\phi(n,k)$ the number of integers s in $n < s \leq n+k$ with no prime factor less than k . Clearly

$$f(n,k) \geq \phi(n,k)$$

and

$$F(n,k) \leq \pi(k) + \phi(n,k).$$

More generally for every $l \leq k$,

$$F(n,k) \leq \pi(l) + \phi(n,l).$$

A simple averaging argument gives

$$\lim_{k \rightarrow \infty} \frac{\log k}{k} \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n=1}^x f(n,k) = \lim_{k \rightarrow \infty} \frac{\log k}{k} \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n=1}^x F(n,k) = e^{-\gamma}.$$

This follows from the well known theorem of Mertens,

$$\prod_{p < k} \left(1 - \frac{1}{p}\right) = (1 + o(1)) \frac{e^{-\gamma}}{\log k}.$$

By the method used in [4, Thm. 3] we can prove

$$\lim_{k \rightarrow \infty} \left(\frac{\log k}{k}\right)^2 \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n=1}^x f(n,k)^2 = \lim_{k \rightarrow \infty} \left(\frac{\log k}{k}\right)^2 \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{n=1}^x F(n,k)^2 = e^{-2\gamma}.$$

Thus as $k \rightarrow \infty$, for almost all n

$$f(n,k) = (1 + o(1))F(n,k) = (1 + o(1)) \frac{e^{-\gamma}k}{\log k}.$$

We thus have

$$\max_n f(n, k) \geq (1 - o(1)) \frac{e^{-\gamma k}}{\log k},$$

and we have not been able to improve this.

Finally we remark that there are several problems connected with $\min f(n, k)$. The first is to determine or estimate the smallest positive integer n_k for which $f(n_k, k) = 2$. Trivially $n_k \leq P - k$, where P is the product of all primes less than k . We have found that $k = 99925854$ gives strict inequality. For this value of k , take m to be the smallest positive integer satisfying the congruences

$$m + k \equiv 0 \pmod{p_1 p_2 p_3},$$

$$m + 1 \equiv 0 \pmod{P/p_1 p_2 p_3}.$$

Take $p_1 = 10061$, $p_2 = 20123$, $p_3 = 35281$ and notice that $k - 1$ is prime. Since $k < p_1^2$, the only numbers between $m + 1$ and $m + k$ which have no prime factor in common with $m + 1$ are $m + 2$, $m + p_1 + 1$, $m + p_2 + 1$, and $m + p_3 + 1$. But $m + 2 \equiv m + p_2 + 1 \equiv 0 \pmod{p_1}$, $m + p_1 + 1 \equiv 0 \pmod{p_3}$, and $m + p_3 + 1 \equiv 0 \pmod{p_2}$. Then obviously $n_k \leq m < P - k$.

Now let n_k' be the smallest integer for which there are two integers a and b , $n_k' < a < b \leq n_k' + k$, so that $(n + j, ab) > 1$ for $1 \leq j \leq k$. The difference between n_k'

and n_k is that we do not require $(a,b) = 1$.

Theorem 3. For all sufficiently large k ,

$$(8) \quad n_k' < \frac{1}{2^p} < P - k,$$

where P is the product of all primes less than k .

A recent theorem of Motohashi [5] states that there are infinitely many primes p such that the smallest prime $q \equiv 1 \pmod{p}$ is less than $p^{1.64}$. This immediately gives (8) for every k , $p + q < k \leq p^2$. To see this, let m be the smallest positive integer satisfying the congruences

$$m + p + q + 1 \equiv 0 \pmod{pq},$$

$$m + 1 \equiv 0 \pmod{P/pq}.$$

If $a = m + 1$ and $b = m + p + q + 1$, then clearly $(m + j, ab) > 1$ for $1 \leq j < p^2$. Then obviously $n_k' \leq m$. Let $m^* = P - m - k - 1$. From the construction of m it is easy to see that $m^* > 0$ and that $n_k' \leq m^*$. Since either m or m^* is less than $\frac{1}{2}(P-k)$, the theorem is proved.

A simple modification of this proof shows that the intervals $(p + q, p^2)$ cover every integer for $k > k_0$. In fact probably Theorem 3 holds for every $k > 34$ ($34 = 11 + 23$). For the proof of this one would have to obtain explicitly given constants in Bombieri's theorem [1].

The smallest value of k for which (8) holds is given by $k = 17$, $n_k' = 2183$, $a = 2184$, $b = 2200$. Perhaps $n_k' = \prod_{p < k} p - k$ for $25 < k \leq 34$ but we did not investigate this.

A well known theorem of Pillai and Szekeres [3] states that for $k \leq 16$, every set of k consecutive integers contains one which is relatively prime to the others. (A theorem of A. Brauer and Pillai [2] states that this theorem fails for every $k \geq 17$.) The theorem of Pillai and Szekeres easily implies that if $1 < b - a < 16$ there is always a c between a and b such that $(a, c) = (b, c) = 1$. This easily implies that $n_k' = \prod_{p < k} p - k$ for $k \leq 16$. We leave the simple details to the reader.

Finally we state a few more unsolved problems:

1. For which k is it true that if $(a, b) = 1$, $1 < b - a = k$, then there always is a c between a and b such that $(ab, c) = 1$? Perhaps for every sufficiently large k one can find integers $b - a = k$ so that for every $a < c < b$, $(ab, c) > 1$, but we do not know if this holds for every sufficiently large k .

It is easy to see that if 1. is true for a given k , then

$$n_k = \prod_{p < k} p - k.$$

It seems certain that

$$n_k' = o\left(\prod_{p < k} p\right).$$

On the other hand n_k' certainly increases very fast. From Rosser's result used in the proof of Theorem 2 it follows that

$n_k' > \exp(k^{\frac{1}{2}-\epsilon})$. Very likely

$$(9) \quad \phi(n, n+k) > c_k \frac{k}{\log k}$$

holds for every $n < k^a$. This is probably very deep.

2. Is there a k so that for some set of k consecutive integers $n+1, \dots, n+k$,

$$(n+i, \prod_{\substack{j=1 \\ j \neq i}}^k (n+j)) = A(n,i)$$

is composite for all $i, 1 \leq i \leq k$? Perhaps every sufficiently large k has this property.

The following problem is probably much more difficult: Is there a k so that $A(n,i)$ has more than r distinct prime factors for all $i, 1 \leq i \leq k$? For $r = 0$ this is the Brauer-Pillai-Szekeres result, but for $r = 1$ it is probably quite difficult and the answer may very well be 'yes' for $r = 1$ and 'no' for $r > 1$.

We attach a small table of values of $\max f$, $\min F$ and $\max F$

and some indication of values of $n + 1$ for which these are attained. In the table of $\min F$, one choice for $n + 1$ is $P - \lfloor \frac{1}{2}k \rfloor$, except when $k = 38$ or 39 , for which $n + 1 = 2162$ may be chosen. Notice the overlap with the Pillai-Szekeres example [2] of 17 consecutive numbers, none relatively prime to the product of the others. In the table of $\max F$, the $n + 1$ column is left blank if one choice for $n + 1$ is $P - 1$.

We wish to acknowledge the assistance of E.F. Ecklund,
R.B. Eggleton and R.K. Guy.

k	$\max f$	$n+1$	$\min F$	$\max F$	$n+1$	k	$\max F$	$n+1$
4	3	1	3	3		46	16	
5	4	1	3	4		47	16	
6	3	1	4	4		48	16	
7	4	1	4	5		49	17	
8	4	1	4	5		50	17	
9	5	5	4	6		51	18	11
10	5	4	5	6		52	18	10
11	5	1	5	6		53	18	9
12	5	1	5	6		54	18	8
13	6	1	5	7		55	18	
14	6	1	6	7		56	18	
15	7	95	6	8		57	19	11
16	7	94	6	8		58	19	10
17	7	67	6	8		59	19	9
18	7	92	6	8		60	19	8
19	8	95	6	9		61	20	11
20	8	94	6	9		62	20	10
21	8	11	6	10		63	21	11
22	8	10	7	10		64	21	10
23	8	1	7	10		65	21	9
24	8	1	7	10		66	21	8
25	9	7	7	11		67	21	7
26	9	6	8	11		68	21	6
27	9	5	8	11		69	22	11
28	9	4	8	11		70	22	10
29	9	1	8	11		71	22	9
30	9	1	8	11		72	22	8
31	10	1	8	12		73	23	11
32	10	1	8	12		74	23	10
33	11	11	8	13		75	23	
34	11	10	9	13		76	23	
35	11	9	9	13		77	23	
36	11	8	9	13		78	23	
37	12	11	9	14	7	79	24	11
38	12	10	9	14	6	80	24	10
39	13	$> 10^4$	9	15	11	81	24	
40	13	$> 10^4$	10	15	10	82	24	
41	13		10	15		83	24	
42	13		10	15		84	24	
43	14		10	16	11	85	25	
44	14		10	16	10	86	25	
45	14		10	16		87	26	1217

References

- [1] E. Bombieri, On the large sieve, *Mathematika* 12 (1965), 201-225.
- [2] See e.g. R. J. Evans, On blocks of N consecutive integers, *Amer. Math. Monthly*, 76 (1969), 48-49.
- [3] P. Erdős, ["]Some problems in number theory, in A.O.L. Atkin and B.J. Birch (eds.), Computers in number theory, Academic Press, London, 1971, 405-414.
- [4] P. Erdős, ["]The coefficients of the cyclotomic polynomial, *Bull. Amer. Math. Soc.* 52 (1946), 179-184.
- [5] Motohashi Yoichi, A note on the least prime in an arithmetic progression with a prime difference, *Acta Arith.* 17 (1970), 283-285.