Hardy-Ramanujan Journal
Vol.16 (1993) 1-20

On sets of coprime integers in intervals

P. Erdős and Sárközy [1]

1. Throughout this paper we use the following notations : $\mathbb{Z}$ denotes the set of the integers. N denotes the set of the positive integers. For $\mathcal{A} \subset N, m \in N, u \in \mathbb{Z}$ we write $\mathcal{A}_{(m,u)} = \{a : a \in \mathcal{A}, a \equiv u(\text{mod } m)\}$. $\varphi(n)$ denotes Euler's function. p_k denotes the k^{th} prime: $p_1 = 2, p_2 = 3, \cdots$ and we put $P_k = \prod_{i=1}^{k} p_i$. If $k \in N$ and $k \geq 2$, then $\Phi_k(\mathcal{A})$ denotes the number of the k-tuples $(a_1, \cdots, a_k)$ such that $a_1 \in \mathcal{A}, \cdots, a_k \in \mathcal{A}, a_1 < a_2 < \cdots < a_k$ and $(a_i, a_j = 1)$ for $1 \leq i < j \leq k$. If $k \in N, \mathcal{A} \subset N$ and $\Phi_k(\mathcal{A}) = 0$, i.e., $\mathcal{A}$ does not contain a subset $\mathcal{S}$ consisting of k pairwise coprime integers, then $\mathcal{A}$ is said to have property $\mathbf{P}_k$, and Γ_k denotes the family of those subsets of N which have property $\mathbf{P}_k$. We write

$$F_k(n) = \max_{\substack{\mathcal{A} \subset \{1, \cdots, n\} \\ \mathcal{A} \in \Gamma_k}} | \mathcal{A} | .$$

(In other words, $t = F_k(n) + 1$ is the smallest positive integer such that every set $\mathcal{B}$ with $\mathcal{B} \subset \{1, \cdots, n\}, | \mathcal{B} |= t$ contains k pairwise coprime integers). Moreover, for $k, m, n \in N$ we write

$$g_k(m, n) = \max_{\substack{\mathcal{A} \subset \{m, m+1, \cdots, m+n-1\} \\ \mathcal{A} \in \Gamma_k}} | \mathcal{A} | \tag{1}$$

[1] Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. 1901.

and

$$G_k(n) = \max_{m \in N} g_k(m,n)$$

so that, clearly,

$$(g_k(1,n) =) F_k(n) \le G_k(n). \tag{2}$$

For $k,m,n \in N$, let $\psi_k(m,n)$ denote the number of those integers $u \in \{m, m+1, \cdots, m+n-1\}$ which are multiples of at least one of the first k primes, and write

$$\Psi_k(n) = \psi_k(1,n).$$

The set $A \overset{\text{def}}{=} \{a : a \in \{m, m+1, \cdots, m+n-1\}, (a, P_{k-1}) > 1\}$ has property $\mathbf{P}_k$ and thus for this set A we have

$$g_k(m,n) \ge |A| = \psi_{k-1}(m,n), \tag{3}$$

in particular,

$$g_k(1,n) = F_k(n) \ge \psi_{k-1}(1,n) = \Psi_{k-1}(n). \tag{4}$$

Clearly for all $m,n \in N$ we have

$$\psi_k(m,n) = |\{u : m \le u < m+n, (u, P_k) > 1\}| =$$

$$= |\{u : m \le u < m+n\}| - \sum_{d \mid P_k} \mu(d) \, |\{u : m \le u < m+n, d \mid u\}| =$$

(5)

$$= n - \sum_{d \mid P_k} \mu(d) \left(\left[\tfrac{m+n}{d}\right] - \left[\tfrac{m}{d}\right] \right)$$

whence

$$\left| \psi_k(m,n) - \left(1 - \prod_{i=1}^{k} \left(1 - \tfrac{1}{p_i}\right) \right) n \right| =$$

$$= \left| \left(n - \sum_{d \mid P_k} \mu(d) \left(\left[\tfrac{m+n}{d}\right] - \left[\tfrac{m}{d}\right] \right) \right) - \left(n - \sum_{d \mid P_k} \mu(d) \tfrac{n}{d} \right) \right| \le$$

(6)

$$\leq \sum_{d|P_k} |\mu(d)| \, | [\tfrac{m+n}{d}] - [\tfrac{m}{d}] - \tfrac{n}{d} | \leq$$

$$\leq \sum_{d|P_k} (| [\tfrac{m+n}{d}] - \tfrac{m+n}{d} | + | \tfrac{m}{d} - [\tfrac{m}{d}] |) < 2\sum_{d|P_k} 1 = 2^{k+1}.$$

If $P_k \mid n$, then we have $| \{d : m \leq u < m+n, d \mid u\} | = n/d$ thus it follows from (5) that

$$\psi_k(m,n) = n - \sum_{d|P_k} \mu(d)\frac{n}{d} = \left(1 - \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right)\right) n \text{ for } P_k \mid n. \tag{7}$$

In particular, we have

$$\Psi_k(n) = \psi_k(1,n) = | \{u : u \in N, u \leq n, (u, P_k) > 1\} | =$$

$$= - \sum_{d|P_k, d>1} \mu(d) \left[\frac{n}{d}\right] \text{ for all } n \in N \tag{8}$$

and

$$\Psi_k(n) = \left(1 - \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right)\right) n \text{ for } n \in N, P_k \mid n.$$

Finally, for $k, \ell, m, n \in N$, $h_{(k,\ell)}(m,n)$ denotes the maximum of the cardinalities of the sets $\mathcal{A}$ such that $\mathcal{A} \subset \{m, m+1, \cdots, m+n-1\}$, $(a, P_k) = 1$ for all $a \in \mathcal{A}$ and $\mathcal{A} \in \Gamma_\ell$.

2. It is easy to see that

$$F_2(n) = \Psi_1(n) = \left[\frac{n}{2}\right] \tag{9}$$

and

$$F_3(n) = \Psi_2(n) = \left[\frac{n}{2}\right] + \left[\frac{n}{3}\right] - \left[\frac{n}{6}\right] \quad \left(= \frac{2}{3} \text{ for } 6 \mid n\right).$$

Erdös, Sárközy and Szemerédi [3],[4] extended and sharpened these statements in various directions, Erdös conjectured long ago (see, e.g., [2]) that

$$F_k(n) = \Psi_{k-1}(n). \tag{11}$$

Szabó and Tóth [5] proved this in the special case $k = 4$:

$$F_4(n) = \Psi_3(n).$$

However, the general case seems to be hopeless at the present.

In this paper, our goal is to study the case of general k and to prove several partial results. In particular, in sections 3 and 4 we will study the connection between the functions $F_k(n)$ and $G_k(n)$. In section 5, the function $h_{(k,\ell)}(m,n)$ will be studied. In section 6, we will give an upper bound for $G_k(n)$. Finally, in section 7, we will generalize several results proved in [3] and [4] by estimating $\Phi_k(\mathcal{A})$.

3. By (2), $F_k(n) \leq G_k(n)$ for all n. First we remark that for $k = 2$ and 3, both $<$ and $=$ occur infinitely often in this inequality.

THEOREM 1. *If* $m, u, v \in N$, *then*

(12) $g_2(m, 2v) = v,$

(13) $g_2(2u, 2v+1) = v+1,$

(14) $g_2(2u-1, 2v+1) = v,$

(15) $g_3(m, 6v) = 4v,$

(16) $g_3(6u-5, 6v-1) = 4v-1,$

(17) $g_3(6u, 6v-1) = 4v.$

It follows trivially from Theorem 1 that

COROLLARY 1. *For all* $v \in N$ *we have*

$$F_2(2v) = G_2(2v) = v,$$

$$G_2(2v+1) = v+1 = F_2(2v+1)+1,$$

$$F_3(6v) = G_3(6v) = 4v$$

and

$$G_3(6v-1) \geq 4v = F_3(6v-1)+1.$$

PROOF OF THEOREM 1. (12),(13) and (14) are near trivial, thus we prove only (15),(16) and (17). By (3) and (5), for all $m, n \in N$ we have

$$g_3(m,n) \geq \psi_2(m,n) = N_1 + N_2 - N_3 \tag{18}$$

where

$$N_1 = |\{a : a \in \{m, m+1, \cdots, m+n-1\}, 2 \mid a\}|,$$

$$N_2 = |\{a : a \in \{m, m+1, \cdots, m+n-1\}, 3 \mid a\}|$$

and

$$N_3 = |\{a : a \in \{m, m+1, \cdots, m+n-1\}, 6 \mid a\}|.$$

It is easy to see that

$$N_1 = 3v, N_2 = 2v \text{ and } N_3 = v \text{ for all } m \text{ and } n = 6v, \tag{19}$$

$$N_1 = 3v-1, N_2 = 2v-1 \text{ and } N_3 = v-1 \text{ for } m = 6u-5, n = 6v-1 \tag{20}$$

and

$$N_1 = 3v, N_2 = 2v \text{ and } N_3 = v \text{ for } m = 6u, n = 6v-1. \tag{21}$$

It follows from (18), (19), (20) and (21) that

$$g_3(m, 6v) \geq 6v, g_3(6u-5, 6v-1) \geq 4v-1, g_3(6u, 6v-1) \geq 4v. \tag{22}$$

Now we will show that

$$g_3(t,6) \leq 4 \quad \text{for all} \quad t \in N. \tag{23}$$

In fact, if $A \subset \{t, t+1, \cdots, t+5\}$ and $|A| \geq 5$, then either A contains three consecutive odd numbers or A contains two consecutive odd numbers and an even number divisible by either 3 or 5. In both cases, these three numbers are pairwise coprime so that $\mathcal{A}$ cannot have property P_3 which proves (19).

Clearly, for all $k, a, b, c \in N$ we have

$$g_k(a, b+c) \le g_k(a, b) + g_k(a+b, c). \tag{24}$$

By (23) and (24) we have

$$g_3(m, 6v) \le \sum_{i=0}^{v-1} g_3(m+6i, 6) \le \sum_{i=0}^{v-1} 4 = 4v. \tag{25}$$

(15) follows from (22) and (25).

Next we will show that

$$g_3(6u-5, 5) \le 3. \tag{26}$$

In fact, if $\mathcal{A} \subset \{6u-5, 6u-4, 6u-3, 6u-2, 6u-1\}$ and $|\mathcal{A}| \ge 4$, then either $\mathcal{A}$ contains three consecutive odd numbers or $\mathcal{A}$ contains $6u-4, 6u-2$ and one of $6u-5, 6u-3$ and $6u-1$. In both cases, the three numbers are pairwise coprime and this proves (26).

It follows from (15), (24) and (26) that

$$g_3(6u-5, 6v-1) \le g_3(6u-5, 5) + g_3(6u, 6(v-1)) \le 3 + 4(v-1) = 4v-1. \tag{27}$$

(16) follows from (22) and (27).

Finally, by (15) we have

$$g_3(6u, 6v-1) \le g_3(6u, 6v) = 4v. \tag{28}$$

(17) follows from (22) and (28).

4. In this section we will show that for all k there is an integer n_k such that $\lim_{k \to +\infty} (G(n_k) - F(n_k)) = +\infty$.

THEOREM 2. *There is a positive constant c_1 and a number k_0 such that for all $k \ge k_0$ there is an integer n_k with*

$$G_k(n_k) - F_k(n_k) > c_1 k (\log k)^3 (\log\log k)^{-2}. \tag{29}$$

PROOF. We need the following result of Erdős :

LEMMA 1. *For a certain positive constant c_2 and all $n \in N$ we can find more than $c_2 p_n \log p_n (\log\log p_n)^{-2}$ consecutive integers so that each of them is divisible by at least one of the primes $p_1, p_2, \cdots, p_n$.*

In fact, this is Theorem 2 in [1].

By the prime number theorem, $p_n \sim n \log n$ so that $p_n \log p_n (\log\log p_n)^{-2} \sim n(\log n)^2(\log\log n)^{-2}$. Thus by Lemma 1, there is a positive constant c_3 and for all $k \geq k_1$ there are numbers $n_k, t_k \in N$ so that

$$n_k > c_3 k (\log k)^2 (\log\log k)^{-2} \tag{30}$$

and each of the integers $t_k, t_k+1, \cdots, t_k+n_k-1$ is divisible by at least one of the primes $p_1, p_2, \cdots, p_{k-1}$. Then clearly, the set $\mathcal{A} = \{t_k, t_k+1, \cdots, t_k+n_k-1\}$ has property $\mathbf{P}_k$ whence

$$G_k(n_k) = g_k(t_k, n_k) = n_k. \tag{31}$$

Now we will give an upper bound for $F_k(n_k)$. Assume that $\mathcal{A} \subset \{1, 2, \cdots, n_k\}$ and $\mathcal{A}$ has property $\mathbf{P}_k$. If $q_1 < q_2 < \cdots < q_\ell$ are primes contained in $\mathcal{A}$, then these primes form a subset of $\mathcal{A}$ consisting of pairwise coprime integers. Since $\mathcal{A}$ has property $\mathbf{P}_k$, this implies that $\ell \leq k-1$, i.e., $\mathcal{A}$ contains at most $k-1$ primes. Thus by (30) and the prime number theorem we have

$$\begin{aligned} |\mathcal{A}| &\leq |\{n : n \leq n_k, n \text{ is not prime}\}| + (k-1) = \\ &= n_k - \pi(n_k) + (k-1) < n_k - c_4 n_k (\log n_k)^{-1} + k - 1 < \\ &< n_k - c_5 k (\log k)^3 (\log\log k)^{-2} \end{aligned}$$

whence

$$F_k(n_k) < n_k - c_5 k (\log k)^3 (\log\log k)^{-2}. \tag{32}$$

(29) follows from (31) and (32), and this completes the proof of Theorem 2. On the other hand, we conjecture that

$$\lim_{n \to +\infty} \sup (G_k(n_k) - F_k(n_k)) < +\infty$$

for all k (and, in fact, perhaps the lower bound in (29) is close to the truth). Unfortunately, we have not been able to show this.

Moreover, we conjecture that conjecture (11) can be extended to the more general function $g_k(m,n)$ in the following way : for all $k \geq 2$ and $m,n \in N$, the maximum in (1) is assumed by a set $\mathcal{A} \subset \{m, m+1, \cdots, m+n-1\}$ which consists of the multiples of certain primes $q_1, q_2, \cdots, q_{k-1}$. However, these primes need not be the first $k-1$ primes, as the following example shows : Let $m = 45, n = 10, k = 4$, and consider

$$\mathcal{A} \stackrel{\text{def}}{=} \{a : 45 \leq a \leq 54, (a, 2 \cdot 3 \cdot 7) > 1\} = \{45, 46, 48, 49, 50, 51, 52, 54\}.$$

Then clearly, $\mathcal{A} \in \Gamma_k$ so that

$$g_4(45, 10) \geq |\mathcal{A}| = 8.$$

On the other hand, we have

$$\psi_3(45, 10) = |\{u : 45 \leq u \leq 54, (u, 2 \cdot 3 \cdot 5) > 1\}| = |\{45, 46, 48, 50, 51, 52, 54\}| = 7.$$

5. In this section, we will estimate the function $h_{(k,\ell)}(m,n)$.

THEOREM 3.

(i) *If $k, \ell, m, n \in N$ and $\ell \geq 2$, then we have*

$$h_{(k,\ell)}(m,n) < (\ell - 1)\left(\frac{n}{p_{k+1}P_k} + 2\right)\prod_{i=1}^{k}(p_i - 1). \tag{33}$$

(ii) *If $k, \ell, m, n \in N$ and $\ell \geq 2$, then we have*

$$h_{(k,\ell)}(m,n) \geq n\prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right)\left(1 - \prod_{j=k+1}^{k+\ell-1}\left(1 - \frac{1}{p_j}\right)\right) - P_{k+\ell-1}. \tag{34}$$

Moreover, if $P_{k+\ell-1} \mid n$, then the term $P_{k+\ell-1}$ on the right hand side can be dropped.

PROOF.

(i) Assume that

$$\mathcal{A} \subset \{m, m+1, \cdots, m+n-1\}, \tag{35}$$

(36) $$(a, P_k) = 1 \text{ for all } a \in \mathcal{A}$$

and

(37) $$|\mathcal{A}| \geq (\ell - 1)\left(\frac{n}{p_{k+1}P_K} + 2\right)\prod_{i=1}^{k}(p_i - 1).$$

Let u denote an integer for which $|\mathcal{A}(P_k, u)|$ is maximal :

$$|\mathcal{A}(P_k, u)| \geq |\mathcal{A}(P_k, v)| \quad \text{for all } v \in \mathbb{Z}.$$

By (36) and (37), clearly we have

(38) $$|\mathcal{A}(P_k, u)| \geq \frac{|\mathcal{A}|}{\varphi(P_k)} = |\mathcal{A}|\prod_{i=1}^{k}(p_i - 1)^{-1} \geq (\ell - 1)\left(\frac{n}{p_{k+1}P_k} + 2\right).$$

Define the integers x and y, respectively, by

(39) $$x \equiv u (\text{mod } P_k), x + 1 \leq m \leq x + P_k$$

and

(40) $$x + (y - 1)p_{k+1}P_k < m + n - 1 \leq x + y\, p_{k+1}P_K$$

so that, by (39) and (40),

(41) $$y < \frac{m + n - 1 - x}{p_{k+1}P_k} + 1 \leq \frac{P_k + n - 1}{p_{k+1}P_k} + 1 < \frac{n}{p_{k+1}P_k} + 2.$$

For $i \in N$ write

$$\mathcal{A} = \mathcal{A}_{(P_k,u)} \cap \{x + ((i-1)p_{k+1} + 1)P_k, x + ((i-1)p_{k+1} + 2)P_k, \cdots, x + ip_{k+1}P_k\}.$$

Then we have

$$\mathcal{A}_{(P_k,u)} = \cup_{i=1}^{y}\mathcal{A}_i,$$

thus by (38) and (41) there is an integer z (with $1 \leq z \leq y$) such that

$$|\mathcal{A}_z| \geq \frac{|\mathcal{A}_{(P_k,u)}|}{y} > (\ell - 1)\left(\frac{n}{p_{k+1}P_k} + 2\right)\left(\frac{n}{p_{k+1}P_k} + 2\right)^{-1} = \ell - 1$$

whence

$$(42) \qquad | \mathcal{A}_z | \geq \ell .$$

If $a \in \mathcal{A}_z, a' \in \mathcal{A}_z$ and $a > a'$, then $a - a'$ is of the form $a - a' = jP_k$ where $j \in N, j < p_{k+1}$. The prime factors of jP_k are smaller, than p_{k+1}. It follows that (a, a') has no prime factor greater than p_k. On the other hand, by $a \in \mathcal{A}, a' \in \mathcal{A}$ we have $(a, P_k) = (a', P_k) = 1$, so that (a, a') has no prime factor not exceeding p_k. Thus we have $(a, a') = 1$ so that the elements of $\mathcal{A}_z$ are pairwise coprime. By (42), it follows that $\mathcal{A}_z \subset \mathcal{A}$ contains an ℓ-tuple of pairwise coprime integers so that $\mathcal{A} \notin \Gamma_\ell$. This is so for all $\mathcal{A}$ satisfying (35), (36) and (37) which proves (33).

(ii) Define the set $\mathcal{B}_{(k,\ell)}(m,n)$ by

$$\mathcal{B}_{(k,\ell)}(m,n) = \{b : b \in N, m \leq b < n, (b, P_k) = 1, (b, p_{k+1}p_{k+2} \cdots p_{k+\ell-1}) > 1\}.$$

Then clearly, $\mathcal{B}_{(k,\ell)}(m,n) \in \Gamma_\ell$ and thus

$$(43) \qquad h_{(k,\ell)}(m,n) \geq | \mathcal{B}_{(k,\ell)}(m,n) | .$$

If $P_{k+\ell-1} \mid n$, then by the Chinese remainder theorem we have

(44)

$$| \mathcal{B}_{(k,\ell)}(m,n) | = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right) \left(1 - \prod_{j=k+1}^{k+\ell-1} \left(1 - \frac{1}{p_j}\right)\right) \quad (\text{for } P_{k+\ell-1} \mid n).$$

If $P_{k+\ell-1} \mid n$ is not assumed, then define n' by $P_{k+\ell-1} \mid n', n' \leq n < n' + P_{k+\ell-1}$. Then by (44),

$$| \mathcal{B}_{(k,\ell)}(m,n) | \geq | \mathcal{B}_{(k,\ell)}(m,n') | =$$

$$(45) \quad n' \prod_{i=1}^{k} \left(1 - \tfrac{1}{p_i}\right) \left(1 - \prod_{j=k+1}^{k+\ell-1} \left(1 - \tfrac{1}{p_j}\right)\right) = n \prod_{i=1}^{k} \left(1 - \tfrac{1}{p_i}\right) \left(1 - \prod_{j=k+1}^{k+\ell-1} \left(1 - \tfrac{1}{p_j}\right)\right) -$$

$$-(n - n') \prod_{i=1}^{k} \left(1 - \tfrac{1}{p_i}\right) \left(1 - \prod_{j=k+1}^{k+\ell-1} \left(1 - \tfrac{1}{p_j}\right)\right) \geq$$

$$\geq n\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)\left(1-\prod_{j=k+1}^{k+\ell-1}\left(1-\frac{1}{p_j}\right)\right)-P_{k+\ell-1}$$

The result follows (43), (44) and (45).

6. In this section we will study several consequences of Theorem 3 and, in particular, we will estimate $G_k(n)$. First we consider the important special case $\ell = 2$, when we get a quite sharp estimate for the function $h_{(k,\ell)}(m,n)$. In fact, in this case the error term is bounded for fixed k :

THEOREM 4. *For $k,m,n \in N$ we have*

$$\left| h_{(k,2)}(m,n)-\frac{n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)\right| \leq P_{k+1}.$$

PROOF. By using Theorem 3 with $\ell = 2$ we obtain that

$$h_{(k,2)}(m,n)<\left(\frac{n}{p_{k+1}P_k}+2\right)\prod_{i=1}^{k}(p_i-1)=$$

$$=\frac{n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)+2\prod_{i=1}^{k}(p_i-1)<$$

$$<\frac{n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)+2P_k$$

and

$$h_{(k,2)}(m,n)\geq n\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)\left(1-\left(1-\frac{1}{p_{k+1}}\right)\right)-P_{k+1}=\frac{n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)-P_{k+1}$$

whence the result follows.

If $\ell > 2$ but ℓ is not much greater, than k, then still we have a satisfactory estimate for $h_{(k,\ell)}(m,n)$:

THEOREM 5. *For all $\varepsilon > 0$ there exist numbers $\delta > 0, \eta > 0$ and $k_0(\varepsilon)$ such that for all $k \in N, k > k_0(\varepsilon)$ we have*

$$(1-\varepsilon)n\left(\sum_{i=k+1}^{k+\ell-1}\frac{1}{p_i}\right)\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)<h_{(k,\ell)}(m,n)<(1+\varepsilon)\frac{(\ell-1)n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)$$

(46) *for* $k > k_0(\varepsilon), 2 \le \ell < k^{1+\delta}$, *all* $m \in N$ *and* $n > n_0(k, \varepsilon)$,

$$(1-\varepsilon)\frac{(\ell-1)n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right) < h_{(k,\ell)}(m,n) < (1+\varepsilon)\frac{(\ell-1)n}{p_{k+1}}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)$$

(47) *for* $k > k_0(\varepsilon), \ell < \eta k$, *all* $m \in N$ *and* $n > n_0(k, \varepsilon, \eta)$

and

(48)

$$(1-\varepsilon)\log 2\,\frac{n}{\log k}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right) < h_{(k,k+1)}(m,n) < (1+\varepsilon)\frac{n}{\log k}\prod_{i=1}^{k}\left(1-\frac{1}{p_i}\right)$$

for $k > k_0(\varepsilon)$, *all* $m \in N$ *and* $n > n_0(k, \varepsilon)$.

PROOF. The upper bound in (46) follows trivially from (33) in Theorem 3. To derive the lower bound in (46) from (34) in Theorem 4, observe that $1 - x = e^{-x+O(x^2)}$ and $e^{-x} = 1 - x + O(x^2)$ for $x \to 0$, moreover, it follows from

$$\sum_{p<x}\frac{1}{p} = \textit{loglog } x + C + o(1) \tag{49}$$

and the prime number theorem that if $\rho > 0, \delta$ is sufficiently small in terms of ρ and $k > k_0(\rho)$, then

$$\sum_{i=k+1}^{[k^{1+\delta}]}\frac{1}{p_i} < \rho.$$

(47) follows trivially from (46) and the prime number theorem.

(48) follows from (46), (49) and the prime number theorem and this completes the proof of Theorem 5.

One may use (48) to estimate $G_k(n)$:

THEOREM 6. *For all* $\varepsilon > 0$ *there is a number* $k_0 = k_0(\varepsilon)$ *such that for* $k > k_0(\varepsilon)$, *all* $m \in N$ *and* $n > n_0(k, \varepsilon)$ *we have*

$$-2^k \le g_k(m,n) - \left(1 - \prod_{i=1}^{k-1}\left(1-\frac{1}{p_i}\right)\right)n < (1+\varepsilon)\frac{n}{\log k}\prod_{i=1}^{k-1}\left(1-\frac{1}{p_i}\right). \tag{50}$$

Moreover, for $P_{k-1} \mid n$ the lower bound can be replaced by 0.

Note that since except for $g_k(m,n)$, all the terms in (50) are independent of m, it follows that

COROLLARY 2. *For $\varepsilon > 0, k > k_0(\varepsilon)$ and $n > n_0(k,\varepsilon)$ we have*

$$-2^k < G_k(n) - \left(1 - \prod_{i=1}^{k-1}\left(1 - \frac{1}{p_i}\right)\right) n < (1+\varepsilon)\frac{n}{\log k}\prod_{i=1}^{k-1}\left(1 - \frac{1}{p_i}\right)$$

where the lower bound can be replaced by 0 *for $P_k \mid n$.*

PROOF OF THEOREM 6. It follows from (3) and (6) that

$$g_k(m,n) - \left(1 - \prod_{i=1}^{k-1}\left(1 - \frac{1}{p_i}\right)\right) n \ge \psi_{k-1}(m,n) - \tag{51}$$

$$- \left(1 - \prod_{i=1}^{k-1}\left(1 - \frac{1}{p_i}\right)\right) n \ge -2^k$$

and, by (7), this lower bound can be replaced by 0 for $P_{k-1} \mid n$.

On the other hand, assume that $\mathcal{A} \subset \{m, m+1, \cdots, m+n-1\}$ and $\mathcal{A} \in \Gamma_k$. Then by (6) and (48), for $k > k_0(\varepsilon)$, all $m \in N$ and $n > n_0(k,\varepsilon)$ we have

$$|\mathcal{A}| = |\{a : a \in \mathcal{A}, (a, P_{k-1}) > 1\}| + |\{a : a \in \mathcal{A}, (a, P_{k-1}) = 1\}| \le$$

$$\le |\{a : m \le a < m+n, (a, P_{k-1}) > 1\}| + h_{(k-1,k)}(m,n) =$$

$$= \psi_{k-1}(m,n) + h_{(k-1,k)}(m,n) < \tag{52}$$

$$< \left(\left(1 - \prod_{i=1}^{k-1}\left(1 - \tfrac{1}{p_i}\right)\right) n + 2^k\right) + \left(1 + \tfrac{\varepsilon}{2}\right)\tfrac{n}{\log(k-1)}\prod_{i=1}^{k-1}\left(1 - \tfrac{1}{p_i}\right) <$$

$$< \left(1 - \prod_{i=1}^{k-1}\left(1 - \tfrac{1}{p_i}\right)\right) n + (1+\varepsilon)\tfrac{n}{\log k}\prod_{i=1}^{k-1}\left(1 - \tfrac{1}{p_i}\right).$$

The result follows from (51) and (52).

7. In this section, we will generalize Theorem 3 in [3] and Theorem 8 in [4]. For $k \in N, k \geq 2$ we write

$$\alpha_k = \inf_{n \in N} \frac{G_k(n)}{n}.$$

(It is easy to see that $\alpha_2 = 1/2, \alpha_3 = 2/3$).

THEOREM 7. *Let $k \in N, k \geq 2$ and $\varepsilon > 0$. There are numbers $n_0 = n_0(k, \varepsilon)$ and $c_6 = c_6(k, \varepsilon)$ such that if $n > n_0, m \in N, \mathcal{A} \subset \{m, m+1, \cdots, m+n-1\}$ and*

(53) $$| \mathcal{A} | > (\alpha_k + \varepsilon)n,$$

then

(54) $$\Phi_k(\mathcal{A}) > c_6 n^k.$$

Note that it follows trivially from Theorem 7 that for all $k \geq 2$, $\lim_{n \to +\infty} \frac{G_k(n)}{n}$ exists (and this limit is equal to α_k).

PROOF. The proof will be based on the following lemma :

LEMMA 2. *To every $\rho > 0$ and $\delta > 0$ there is an $r_0 = r_0(\rho, \delta)$ so that if $r \geq r_0, n > n_1(\rho, \delta, r), m \in N$ and $u = 1, 2, \cdots, P_r$, then for all but $\rho \frac{n}{P_r}$ integers b satisfying*

(55) $$m \leq b < m + n, \quad b \equiv u \ (mod\ P_r),$$

we have

$$\gamma(b) \overset{\text{def}}{=} \sum_{p | b, p_r < p < n} \frac{1}{p} < \delta.$$

PROOF. Denote the set of the integers satisfying (55) by $\mathcal{B}$ and write $\mathcal{B}_1 = \{b : b \in \mathcal{B}, \gamma(b) \geq \delta\}$. By the Chinese remainder theorem, $\sum_{i=1}^{+\infty} \frac{1}{p_i^2} < +\infty$ and $\sum_{p<n} \frac{1}{p} \sim \log \log n$, for $r > r_0(\rho, \delta)$ and $n > n_1(\rho, \delta, r)$ we have

$$\sum_{b\in B}\gamma(b)=\sum_{\substack{m\le b<m+n\\ b\equiv u(\bmod P_r)}}\ \sum_{\substack{p\mid b\\ p_r<p<n}}\frac{1}{p}=$$

$$(56)\quad =\sum_{p_r<p<n}\frac{1}{p}\sum_{\substack{m\le b<m+n\\ b\equiv u(\bmod P_r),\,p\mid b}}1<$$

$$<\sum_{p_r<p<n}\frac{1}{p}\left(\frac{n}{pP_r}+1\right)<\frac{n}{P_r}\sum_{p_r<p}\frac{1}{p^2}+\sum_{p<n}\frac{1}{p}<$$

$$<\frac{\rho\delta}{2}\cdot\frac{n}{P_r}+2\log\log\ n<\rho\delta\frac{n}{P_r}.$$

On the other hand, clearly we have

$$(57)\qquad \sum_{b\in B}\gamma(b)\ge\sum_{b\in B_1}\gamma(b)\ge\sum_{b\in B}\delta=\delta\mid B_1\mid.$$

It follows from (56) and (57) that

$$\mid B_1\mid<\rho\frac{n}{P_r}$$

and this completes the proof of Lemma 2.

Now we prove Theorem 7. By the definition of α_k, there is a positive integer n_1 such that

$$G_k(n_1)<(\alpha_k+\frac{\varepsilon}{4})n_1,$$

so that

$$(58)\qquad g_k(m,n_1)<(\alpha_k+\frac{\varepsilon}{4})n_1\ \text{ for all } m\in N.$$

Let r be a positive integer such that

$$(59)\qquad r>\max\left(n_1,r_0\left(\frac{\varepsilon}{8n_1},\frac{\varepsilon}{32k}\right)\right)$$

and write $n_1P_r=M$.

By (53) we have

$$P_r\max_{0\le i<P_r}\sum_{j=1}^{n_1}\mid A_{(M,in_1+j)}\mid\ge$$

$$\geq \sum_{i=0}^{P_r-1}\sum_{j=1}^{n_1} | \mathcal{A}_{(M,in_1+j)} | = \sum_{u=1}^{M} | \mathcal{A}_{(M,u)} | = | \mathcal{A} | > (\alpha_k+\varepsilon)n.$$

It follows that there is an integer i such that $0 \leq i < P_r$ and

$$\sum_{j=1}^{n_1} | \mathcal{A}_{(M,in_1+j)} | > (\alpha_k+\varepsilon)\frac{n}{P_r}. \tag{60}$$

Clearly, for all $u \in \mathbb{Z}$ we have

$$| \mathcal{A}_{(M,u)} | \leq | \{a : m \leq a < m+n, a \equiv u (\mathrm{mod}\ M)\} | < \frac{n}{M}+1. \tag{61}$$

(60) and (61) imply that for sufficiently large n there exist integers $j_1, j_2, \cdots, j_t$ such that

$$1 \leq j_1 \leq j_2 \leq \cdots < j_t \leq n_1, \tag{62}$$

$$t \geq (\alpha_k + \frac{\varepsilon}{4})n_1 \tag{63}$$

and

$$| \mathcal{A}_{(M,in_1+j_u)} | > \frac{\varepsilon}{4}\ \frac{n}{M} \quad \text{for } u = 1,2,\cdots,t, \tag{64}$$

since otherwise, writing $\mathcal{J}_1 = \{j : 1 \leq j \leq n_1, | \mathcal{A}_{(M,in_1+j)} | > \frac{\varepsilon}{4}\ \frac{n}{M}\}$ and $\mathcal{J}_2 = \{j : 1 \leq j \leq n_1, | \mathcal{A}_{(M,in_1+j)} | \leq \frac{\varepsilon}{4}\ \frac{n}{M}\}$ by (61) for large n we had

$$\sum_{j=1}^{n_1} | \mathcal{A}_{(M,in_1+j)} | = \sum_{j\in\mathcal{J}_1} | \mathcal{A}_{(M,in_1+j)} | + \sum_{j\in\mathcal{J}_2} | \mathcal{A}_{(M,in_1+j)} | \leq$$

$$\leq \sum_{j\in\mathcal{J}_1}(\tfrac{n}{M}+1) + \sum_{j\in\mathcal{J}_2}\tfrac{\varepsilon}{4}\ \tfrac{n}{M} = (\tfrac{n}{M}+1) | \mathcal{J}_1 | + \tfrac{\varepsilon}{4}\tfrac{n}{M} | \mathcal{J}_2 | <$$

$$(\tfrac{n}{M}+1)(\alpha_k+\tfrac{\varepsilon}{4})n_1 + \tfrac{\varepsilon}{4}\ \tfrac{n}{M}\ n_1 =$$

$$= (\alpha_k+\tfrac{\varepsilon}{2})\tfrac{n}{P_r} + (\alpha_k+\tfrac{\varepsilon}{4})n_1 < (\alpha_k+\varepsilon)\tfrac{n}{P_r}$$

and this contradicts (60).

By (58), (62) and (63), the set $\{in_1 + j_1, in_1 + j_2, \cdots, in_1 + j_t\}$ contains a subset $\{v_1, v_2, \cdots, v_k\}$ consisting of k pairwise coprime integers. Then we have

(65) $$v_x \not\equiv v_y \pmod M \quad \text{for} \quad 1 \le x < y \le k,$$

(66) $$(v_x, v_y) = 1 \quad \text{for} \quad 1 \le x < y \le k$$

and, by (64),

(67) $$|\mathcal{A}_{(M,v_x)}| > \frac{\varepsilon}{4}\frac{n}{M} \quad \text{for} \quad x = 1, 2, \cdots, k.$$

Now we will show that it suffices to prove

LEMMA 3. *Using the notations above and writing* $\mathcal{D} = \{d : m \le d < m + n, \gamma(d) = \sum_{p|d, p_r < p < n} \frac{1}{p} < \frac{\varepsilon}{32k}\}$,

(i) *there are more than* $\frac{\varepsilon}{8}\frac{n}{m}$ *integers* d_1, *satisfying* $d_1 \in \mathcal{D} \cap \mathcal{A}_{(M,v_1)}$,

(ii) *if* $j \in \{2, 3, \cdots, k\}$ *and* $d_1, \cdots, d_{j-1}$ *are integers with* $d_1 \in \mathcal{D} \cap \mathcal{A}_{(M,V_1)}, \cdots, d_{j-1} \in \mathcal{D} \cap \mathcal{A}_{(Mv_{j-1})}$ *and (for* $j \ge 3$*)*

(68) $$(d_x, d_y) = 1 \quad \textit{for} \quad 1 \le x < y \le j - 1,$$

then there are more than $\frac{\varepsilon}{16}\frac{n}{M}$ *integers* d_j *satisfying* $d_j \in \mathcal{D} \cap \mathcal{A}_{(M,V_j)}$ *and*

$$(d_x, d_j) = 1 \quad \textit{for} \quad 1 \le x \le j - 1.$$

Assume namely that Lemma 3 has been proved. Select a d_1 in the way described in (i). Then select $d_2, \cdots, d_k$ successively in the way described in (ii). In this way, we obtain distinct k-tuples $(d_1, d_2, \cdots, d_k)$ all whose elements belong to $\mathcal{A}$ and whose elements are pairwise coprime. Thus $\Psi_k(\mathcal{A})$ is greater than or equal to the number of these k-tuples $(d_1, \cdots, d_k)$. To give a lower bound for the number of these k-tuples, observe that by (i), d_1 can be chosen in more than $\frac{\varepsilon}{8}\frac{n}{M}$ ways; if d_1 is given, then by (ii), d_2 can be chosen in more, than $\frac{\varepsilon}{16}\frac{n}{M}$ ways (independently of d_1), etc., finally, if $d_1,\cdots,d_{k-1}$ are given, then d_k can be chosen in more, than $\frac{\varepsilon}{16}\frac{n}{m}$ ways. Thus the total number of these k-tuples $(d_1,\cdots,d_k)$ is greater, than $\frac{\varepsilon}{8}\frac{n}{M}\cdot\frac{\varepsilon}{16}\frac{n}{M}\cdot\cdots\cdot\frac{\varepsilon}{16}\frac{n}{M}$ so that

$$\phi_k(\mathcal{A}) > \frac{2\varepsilon^k}{16^k M^k}n^k > c_7(k,\varepsilon)n^k$$

and this proves (54).

It remains to prove Lemma 3.

PROOF OF LEMMA 3. By Lemma 2 (with $\rho=\frac{\varepsilon}{8n_1}$ and $\delta=\frac{\varepsilon}{32k}$) (59) and (67) for $1\le x\le k$ we have

$$|\mathcal{D}\cap\mathcal{A}_{(M,V_x)}|=|\mathcal{A}_{(M,V_x)}|-|\{a:a\in\mathcal{A}_{(M,V_x)},a\notin\mathcal{D}\}|>$$

$$>\frac{\varepsilon}{4}\frac{n}{M}-|\{a:a\equiv v_x(\text{mod }M),m\le a<m+n,\gamma(a)\ge\tfrac{\varepsilon}{32k}\}|\ge$$

$$\geq\frac{\varepsilon}{4}\frac{n}{M}-|\{a:a\equiv v_x(\text{mod }P_r),m\le a<m+n,\gamma(a)\ge\tfrac{\varepsilon}{32k}\}|\ge \tag{69}$$

$$\ge\frac{\varepsilon}{4}\frac{n}{M}-\frac{\varepsilon}{8n_1}\frac{n}{P_r}=\frac{\varepsilon}{8}\frac{n}{M}$$

which, with $x=1$, proves (i).

Assume now that $2\le j\le k$ and $d_1,\cdots,d_{j-1}$ are given as described in (ii). Then clearly,

$$|\{d:d\in\mathcal{D}\cap\mathcal{A}_{(M,V_j)},(d_1,d)=\cdots=(d_{j-1},d)=1\}|\ge$$

$$\ge|\mathcal{D}\cap\mathcal{A}_{(M,V_j)}|-\sum_{i=1}^{j-1}|\{d:d\in\mathcal{D}\cap\mathcal{A}_{(M,V_j)},(d_i,d)>1\}|. \tag{70}$$

Assume that $1\le i\le j-1, d\in\mathcal{D}\cap\mathcal{A}_{(M,V_j)}$ and p is a prime with

$$p\mid(d_i,d). \tag{71}$$

By $m\le d_i,d<m+n$, we have

$$|d_i-d|<n. \tag{72}$$

Moreover, by $d_i \in \mathcal{A}_{(M,V_i)}, d \in \mathcal{A}_{(M,V_j)}, i < j$ and (65) we have

$$d_i \neq d. \tag{73}$$

It follows from (71), (72) and (73) that $p < n$.

By $d_i \in \mathcal{A}_{(M,V_i)}, d \in \mathcal{A}_{(M,V_j)}$ we have

$$d \equiv v_i \pmod M, \quad d \equiv v_j \pmod M. \tag{74}$$

If $p \le p_r$, then it follows from (71), (74) and $p \mid M = n_1 P_r$ that

$$v_i \equiv v_j \equiv 0 \pmod p$$

so that $p \mid (v_i, v_j)$ with $i < j$ which contradicts (66).

Thus (71) implies that $p_r < p < n$ and thus, by (59), $(p, M) = 1$ so that, by $d_i \in \mathcal{D}$, for $1 \le i \le j-1$ we have

$$\begin{aligned} &|\{d : d \in \mathcal{D} \cap \mathcal{A}_{(M,V_j)}, (d_i, d) > 1\}| \le \\ &\le \sum_{p \mid d_i, p_r < p < n} |\{d : m \le d < m+n, d \equiv v_j (\mathrm{mod}\ M)\}, p \mid d\}| < \\ &< \sum_{p \mid d_i, p_r < p < n} \left(\tfrac{n}{pM} + 1\right) = \\ &= \tfrac{n}{M} \sum_{p \mid d_i, p_r < p < n} \tfrac{1}{p} + |\{p : p \mid d_i, p_r < p < n\}| < \\ &< \tfrac{\varepsilon}{32k} \tfrac{n}{M} + |\{p : p \mid d_i, p_r < p < n\}|. \end{aligned} \tag{75}$$

By (68), it follows (69) (with $x = j$), (70) and (75) that for large n we have

$$\begin{aligned} &|\{d : d \in \mathcal{D} \cap \mathcal{A}_{(M,V_j)}, (d_1, d) = \dots = (d_{j-1}, d) = 1\}| \ge \\ &\ge \tfrac{\varepsilon}{8} \tfrac{n}{M} - \sum_{i=1}^{j-1} \left(\tfrac{\varepsilon}{32k} \tfrac{n}{M} + |\{p : p \mid d_i, p_r < p < n\}|\right) = \\ &= \tfrac{\varepsilon}{8} \tfrac{n}{M} - (j-1) \tfrac{\varepsilon}{32k} \tfrac{n}{M} - \sum_{i=1}^{j-1} |\{p : p \mid d_i, p_r < p < n\}| > \\ &> \tfrac{\varepsilon}{8} \tfrac{n}{M} - \tfrac{\varepsilon}{32} \tfrac{n}{M} - \pi(n) > \tfrac{\varepsilon}{16} \tfrac{n}{M} \end{aligned}$$

which proves (ii) and this completes the proof of Lemma 3.

References

[1] P. Erdös, On the difference of consecutive primes, Quart. J. Oxford 6 (1935), 124-128.

[2] P. Erdös, Remarks in number theory, IV (in Hungarian), Mat. Lapok 13 (1962), 228-255.

[3] P. Erdös, A. Sárközy and E. Szemerédi, On some extremal properties of sequences of integers, Annales Univ. Sci. Budapest, Eötvös 12 (1969), 131-135.

[4] P. Erdös, A. Sárközy and E. Szemerédi, On some extremal properties of sequences of integers, II, Publicationes Math. Debrecen 27 (1980), 117-125.

[5] C. Szabó and G. Tóth, Maximal sequences not containing 4 pairwise coprime integers (in Hungarian), Mat. Lapok 32 (1985), 253-257.