On some extremal properties of sequences of integers, II

By P. ERDŐS, A. SÁRKÖZY and E. SZEMERÉDI (Budapest)

1. Let $A=\{a_1<a_2<\dots\}$ be a sequence of positive integers. Put $A(n)=\sum_{a_i\leqq n} 1$. Denote by $f_k(n)$ the smallest integer so that every sequence A satisfying $A(n)\geqq f_k(n)$ contains a subsequence of k terms which are pairwise relatively prime. It is easy to see that

$$f_2(n)=\left[\frac{n}{2}\right]+1,$$

$$f_3(n)=1+\xi_2(n)\left(=\frac{2}{3}n+1 \text{ for } 6/n\right)$$

and it seems likely that

$$f_k(n)=1+\xi_{k-1}(n)$$

where $\xi_{k-1}(n)$ denotes the number of integers not exceeding n which are multiples of at least one of the first $k-1$ primes $2, 3, \dots, p_{k-1}$.

In Part I of this paper (see [3]) we proved in a sharper and more general form several related conjectures stated in [2]. In this paper, we continue this discussion. First we introduce some *notations*. $A_{(m,u)}$ denotes the integers $a_i\in A$, $a_i\equiv u \pmod m$ (and $A_{(m,u)}(n)$ denotes the number of those terms of the sequence $A_{(m,u)}$ which do not exceed n). $\varphi(n)$ denotes Euler's function. We put

$$\varphi_A(u)=\sum_{\substack{a_i\leqq n\\(a_i,u)=1}} 1$$

and

$$\psi_A(u,v)=\sum_{\substack{a_i\leqq n\\(a_i,u)=(a_i,v)=1}} 1.$$

For $k=2,3,\dots,$ $\Phi_k(A)$ denotes the number of the k-tuples $a_{i_1}, a_{i_2}, \dots, a_{i_k}$ such that $a_{i_1}<a_{i_2}<\dots<a_{i_k}\leqq n$ and $(a_{i_x}, a_{i_y})=1$ for $1\leqq x<y\leqq k$. We put

$$F_2(n)=\min_A \max_{a_j\in A} \varphi_A(a_j)$$

and

$$F_3(n)=\min_A \max_{1\leqq x<y\leqq A(n)} \psi_A(a_x,a_y)$$

where the minimum is to be taken over all sequences A satisfying $A(n)\geqq\left[\frac{n}{2}\right]+1$ and $A(n)\geqq\left[\frac{n}{2}\right]+2$, respectively.

$c_1, c_2, \dots, n_0, n_1, \dots$ will denote suitable positive absolute constants.
In Part I of this paper, we proved the following theorems:

Theorem 1. *For* $n > n_0$,

$$F_2(n) > c_1 n/\text{loglog}\, n.$$

Theorem 2. *There exists constants* c_2, c_3, c_4, n_1 *such that*

$$A_{(2,1)}(n) = s, \quad 1 \leqq s < c_2 n$$

and

$$A(n) > \frac{n}{2}$$

imply that for $n > n_1$,

$$\max_{a_i \in A} \varphi_A(a_i) > c_3 n/\text{loglog}\, \frac{n}{s}$$

and

$$\Phi_2(A) > c_4 sn/\text{loglog}\, \frac{n}{s}.$$

Theorem 3. *To every* $0 < \varepsilon (< 1/2)$, *there exist constants* $c_5 = c_5(\varepsilon)$ *and* $n_2 = n_2(\varepsilon)$ *such that if* $n > n_2$,

$$A_{(2,1)}(n) = s \geqq \varepsilon n$$

and

$$A(n) > \frac{n}{2},$$

then

$$\Phi_2(A) > c_5 n^2.$$

(Note that Theorem 1 is a consequence of Theorems 2 and 3.)

2. Throughout this section, we will assume for simplicity that n is even; all our results could be extended easily for odd n.
P. Erdős conjectured in [2] that if

$$A(n) \geqq \frac{n}{2} + 2$$

then there exists a 4-tuple a_x, a_y, a_u, a_v such that

$$(a_x, a_u) = (a_x, a_v) = (a_y, a_u) = (a_y, a_v) = 1.$$

In this section, we are going to prove the following sharper form of this conjecture:

Theorem 4. *For* $n > n_3$,

$$F_3(n) > c_6 n/(\log\log n)^2.$$

We first prove two other theorems which will easily imply Theorem 4.

Theorem 5. *There exist constants c_7, c_8, c_9 and n_4 such that if* $n > n_4$,

(1) $$A_{(2,1)}(n) = s, \quad 2 \leqq s < c_7 n$$

and

(2) $$A(n) > \frac{n}{2}$$

then there exist at least $c_8 s^2$ pairs a_x, a_y ($a_x \in A, a_y \in A$) satisfying $1 \leqq a_x < a_y \leqq n$ *and*

(3) $$\psi_A(a_x, a_y) > c_9 n \Big/ \left(\log\log \frac{n}{s}\right)^2.$$

PROOF. We need the following known lemma (see [1]).

Lemma 1. *The number of integers $1 \leqq k \leqq n$ satisfying $\varphi(k)/k < 1/t$ is less than $n \exp(-\exp c_{10} t)$ (where $\exp z = e^z$), uniformly in $t > 2$.*

Let us apply Lemma 1 with

$$t = \frac{1}{c_{10}} \log\log \frac{2n}{s}.$$

($t > 2$ holds for small enough c_7.) We obtain that the number of integers $1 \leqq k \leqq n$ which satisfy $\varphi(k)/k < 1/t$ (where t is defined by (3)) is less than $s/2 (\geqq 1)$. Denote now by $b_1 < \ldots < b_r \leqq n$, $r > s/2 (\geqq 1)$ the integers in $A_{(2,1)}$ satisfying $\varphi(b_i)/b_i > 1/t$. We are going to show that for $1 \leqq x < y \leqq r$,

(4) $$\psi_A(b_x, b_y) > c_9 n / \log\log \frac{n}{s}$$

provided that c_7 and c_9 are sufficiently small (and n is large).

Clearly, the number of integers $2u \leqq n$ satisfying $(2u, b_x) = (2u, b_y) = 1$ is

$$\left[\frac{n}{2}\right] + \sum_{p_{i_1} p_{i_2} \ldots p_{i_k} / [b_x, b_y]} (-1)^k \left[\frac{n}{2 p_{i_1} p_{i_2} \ldots p_{i_k}}\right].$$

Here for n large, the number of terms is

$$2^{\nu([b_x, b_y])} < 2^{4 \log n / \log\log n}$$

(where $\nu(m)$ denotes the number of the distinct prime factors of m) since it is well-known (and follows from the prime number theorem or a more elementary theorem) that for $m < N$,

(5) $$\nu(m) < 2 \log N / \mathrm{loglog}\, N,$$

hence

$$\nu([b_x, b_y]) < 2 \log n^2 / \mathrm{loglog}\, n^2 < 4 \log n / \mathrm{loglog}\, n.$$

Thus

$$\sum_{\substack{u\leqq n/2\\(2u,\, b_x)=(2u,\, b_y)=1}} 1 \geqq \frac{n}{2} \prod_{p/[b_x, b_y]} \left(1-\frac{1}{p}\right) - 2^{4\log n/\log\log n} \geqq$$

$$\geqq \frac{n}{2} \prod_{p/b_x} \left(1-\frac{1}{p}\right) \prod_{p/b_y} \left(1-\frac{1}{p}\right) - 2^{4\log n/\log\log n} =$$

$$= \frac{n}{2} \frac{\varphi(b_x)}{b_x} \frac{\varphi(b_y)}{b_y} - 2^{4\log n/\log\log n} >$$

$$> \frac{n}{2t^2} - 2^{4\log n/\log\log n} > \frac{n}{3t^2}$$

for sufficiently large n (with respect to (3)). Hence, we obtain by a simple computation (with respect to (1) and (2)) that for sufficiently small c_7 and c_9,

$$\psi_A(b_x, b_y) \geqq \sum_{\substack{u\leqq n/2\\(2u,\, b_x)=(2u,\, b_y)=1}} 1 - \sum_{\substack{u\leqq n/2\\2u\notin A}} 1 >$$

$$> \frac{n}{3t^2} - \left(\frac{n}{2} - A_{(2,0)}(n)\right) > \frac{n}{3t^2} - \frac{n}{2} + (A(n) - A_{(2,1)}(n)) >$$

$$> \frac{n}{3t^2} - A_{(2,1)}(n) = \frac{n}{3t^2} - s > c_9 n \bigg/ \left(\log\log\frac{n}{s}\right)^2,$$

provided that n is large enough which proves (4).

To complete the proof of Theorem 5, observe that $b_x \in A$ and $b_y \in A$ in (4), furthermore, (4) holds for any pair x, y such that $1 \leqq x < y \leqq r$, and here $r > s/2 (\geqq 1)$.

Theorem 6. *To every* $0 < \varepsilon (< 1/2)$, *there exist constants* $c_{11} = c_{11}(\varepsilon)$ and $n_5 = n_5(\varepsilon)$ *such that if* $n > n_5$,

$$A_{(2,1)}(n) = s > \varepsilon n$$

and

$$A(n) > n/2$$

then there exist at least $c_{10} n^2$ *pairs* a_x, a_y $(a_x \in A,\, a_y \in A)$ *satisfying* $1 \leqq a_x < a_y \leqq n$ *and*

$$\psi_A(a_x, a_y) > c_{11} n.$$

(Note that for $\varepsilon n < s < c_7 n$, Theorem 6 would follow from Theorem 5, but for the large values of s, we need a separate proof.)

PROOF. We are going to show that Theorem 3 implies Theorem 6.

By Theorem 3 and Cauchy's inequality,

(6)
$$\sum_{1\leqq x<y\leqq A(n)} \psi_A(a_x, a_y) = \sum_{1\leqq x<y\leqq A(n)} \sum_{\substack{a_i\leqq n\\(a_i,a_x)=(a_i,a_y)=1}} 1 =$$
$$= \sum_{a_i\leqq n}\Big(\sum_{\substack{1\leqq x<y\leqq A(n)\\(a_i,a_x)=(a_i,a_y)=1}} 1\Big) = \sum_{a_i\leqq n}\binom{\varphi_A(a_i)}{2} = \frac{1}{2}\sum_{a_i\leqq n}(\varphi_A(a_i))^2 - \frac{1}{2}\sum_{a_i\leqq n}\varphi_A(a_i) \geqq$$
$$\geqq \frac{1}{2}\frac{\big(\sum_{a_i\leqq n}\varphi_A(a_i)\big)^2}{n} - \frac{1}{2}\sum_{a_i\leqq n} n \geqq \frac{1}{2n}\Big(\sum_{a_i\leqq n}\Big(\sum_{\substack{a_j\leqq n\\(a_j,a_i)=1}} 1\Big)\Big)^2 - \frac{1}{2}n^2 \geqq$$
$$\geqq \frac{1}{2n}(2\Phi_2(A))^2 - \frac{1}{2}n^2 > \frac{1}{2n}(2c_5(\varepsilon)n^2)^2 - \frac{1}{2}n^2 > c_{12}(\varepsilon)n^3.$$

On the other hand, we have

(7)
$$\sum_{1\leqq x<y\leqq A(n)} \psi_A(a_x, a_y) =$$
$$= \sum_{\substack{1\leqq x<y\leqq A(n)\\ \psi_A(a_x,a_y)>c_{11}n}} \psi_A(a_x, a_y) + \sum_{\substack{1\leqq x<y\leqq A(n)\\ \psi_A(a_x,a_y)<c_{11}n}} \psi_A(a_x, a_y) \leqq$$
$$\leqq \sum_{1\leqq x<y\leqq n} c_{11}n + \sum_{\substack{1\leqq x<y\leqq A(n)\\ \psi_A(a_x,a_y)>c_{11}n}} n \leqq \frac{c_{11}}{2}n^3 + n\sum_{\substack{1\leqq x<y\leqq A(n)\\ \psi_A(a_x,a_y)>c_{11}n}} 1.$$

If c_{11} is sufficiently small (depending on ε) then (6) and (7) yield the statement of Theorem 6.

Theorem 4 follows easily from Theorems 5 and 6. Namely, if
$$2 \leqq s = A_{(2,1)}(n) < c_7 n$$
then Theorem 5 yields that
$$\max_{1\leqq x<y\leqq A(n)} \psi_A(a_x, a_y) > c_9 n/(\log\log n)^2,$$
while if
$$s = A_{(2,1)}(n) \geqq c_7 n$$
then applying Theorem 6 with c_7 in place of ε, we obtain the much sharper
$$\max_{1\leqq x<y\leqq A(n)} \psi_A(a_x, a_y) > c_{11}(c_7)n$$
which completes the proof of Theorem 4.

Finally, we remark that using the same method, also the following theorem could be proved:

Theorem 7. *If* $n>n_6$,
$$A_{(2,1)}(n) = s(>0), \quad A(n) > \frac{n}{2}$$

and

$$r = \min\left\{s, \left[\frac{1}{10}\log\log n\right]\right\} \tag{8}$$

then there exist integers $b_1 < b_2 < \ldots < b_r$ *and* $d_1 < d_2 < \ldots < d_r$ *such that* $b_i, d_i \in A$ *for* $i = 1, 2, \ldots, r$ *and*

$$(b_i, d_j) = 1 \quad \text{for} \quad 1 \leqq i, j \leqq r.$$

(The statement of this theorem is, perhaps, true even with $\min\{s, (1/4-\varepsilon)n/\log n\}$ on the right of (8) but this can not be proved by our method.)

3. Starting out from an other conjecture of P. Erdős, we will prove the following analogue of Theorem 3 for triplets a_x, a_y, a_z instead of pairs a_x, a_y:

Theorem 8. *To every* $0 < \varepsilon (< 1/2)$, *there exist constants* $c_{12} = c_{12}(\varepsilon)$ *and* $n_7 = n_7(\varepsilon)$ *such that if* $n > n_7$ *and*

$$A(n) > \left(\frac{2}{3} + \varepsilon\right) n \tag{9}$$

then

$$\Phi_3(A) > c_{12} n^3.$$

PROOF. Denote by P_r the product of the primes not exceeding r. We need

Lemma 2. *To every* $\varrho > 0$ *and* $\delta > 0$ *there is an* $r_0 = r_0(\varrho, \delta)$ *so that if* $r \geqq r_0$, $n > n_8(\varrho, \delta, r)$ *and* $u = 1, 2, \ldots, P_r$ *then for all but* $\varrho \frac{n}{P_r}$ *integers k satisfying*

$$1 \leqq k \leqq n, \quad k \equiv u \pmod{P_r},$$

we have

$$\alpha(k) = \prod_{\substack{p|k \\ p > r}} \left(1 - \frac{1}{p}\right) > 1 - \delta.$$

This lemma is identical with Lemma 2 in [3].

Now we prove Theorem 8. Let r denote a positive integer for which

$$r \geqq r_0\left(\frac{\varepsilon}{4}, \frac{\varepsilon}{4}\right) \quad \text{and} \quad r \geqq 3 \tag{10}$$

hold.

By (9),

$$\frac{P_r}{6} \max_{0 \leqq k \leqq P_r/6-1} \sum_{i=1}^{6} A_{(P_r, 6k+i)}(n) \geqq$$

$$\geqq \sum_{k=0}^{P_r/6-1} \left(\sum_{i=1}^{6} A_{(P_r, 6k+i)}(n)\right) = \sum_{j=1}^{P_r} A_{(P_r, j)}(n) = A(n) > \left(\frac{2}{3} + \varepsilon\right) n.$$

This implies the existence of an integer k such that $0 \leqq k \leqq P_r/6-1$ and

$$\sum_{i=1}^{6} A_{(P_r, 6k+i)}(n) > \frac{6}{P_r}\left(\frac{2}{3} + \varepsilon\right) n = (4 + 6\varepsilon)\frac{n}{P_r}. \tag{11}$$

Clearly, for every u,

$$A_{(P_r,u)}(n) < \frac{n}{P_r}+1. \tag{12}$$

(11) and (12) imply that there exist integers $i_1, \dots, i_5$ such that

$$1 \leqq i_1 < \dots < i_5 \leqq 6 \tag{13}$$

and

$$A_{(P_r,6k+i_j)}(n) > 2\varepsilon\frac{n}{P_r} \quad \text{for} \quad j=1, \dots, 5, \tag{14}$$

since otherwise

$$\sum_{i=1}^{n} A_{(P_r,6k+i)}(n) \leqq 4\left(\frac{n}{P_r}+1\right)+2\left(2\varepsilon\frac{n}{P_r}\right) = (4+4\varepsilon)\frac{n}{P_r}+4 < (4+6\varepsilon)\frac{n}{P_r}$$

would hold, in contradiction with (11).

It follows from (13) that the sequence $\{i_1, \dots, i_5\}$ contains a subsequence $\{j_1, j_2, j_3\}$ of 3 terms which are pairwise relatively prime. Let us put $6k+j_i=u_i$ for $i=1, 2, 3$; then we have

$$(u_1, u_2) = (u_1, u_3) = (u_2, u_3) = 1, \quad |u_\mu - u_\nu| \leqq 5 \quad \text{for} \quad 1 \leqq \mu, \nu \leqq 3, \tag{15}$$

and by (14),

$$A_{(P_r,u_i)}(n) > 2\varepsilon\frac{n}{P_r}. \tag{16}$$

Let $b_1 < \dots < b_t$ denote the sequence of those integers b for which

$$b \in A_{(P_r,u_1)} \quad \text{and} \quad \prod_{\substack{p/b \\ p>r}}\left(1-\frac{1}{p}\right) > 1-\frac{\varepsilon}{4}. \tag{17}$$

Lemma 2 yields with respect to (10) and (14) that

$$t > A_{(P_r,u_1)}(n) - \frac{\varepsilon}{4}\frac{n}{P_r} > 2\varepsilon\frac{n}{P_r} - \frac{\varepsilon}{4}\frac{n}{P_r} > \varepsilon\frac{n}{P_r}. \tag{18}$$

We are going to estimate from below the number of solutions

$$(b_i, a_x) = 1, \quad a_x \in A_{(P_r,u_2)} \tag{19}$$

(for i fixed).

Assume that $p/(b_i, d)$, $d \equiv u_2 \pmod{P_r}$. By (10), (15) and (17), these imply $p>r$. Denote by $D_i(P_r, u_2)$ the number of those integers d for which $d \leqq n$, $d \equiv u_2$ (mod P_r) and $(b_i, d)=1$. We have by a simple argument

$$\left| D_i(P_r, u_2) - \frac{n}{P_r}\prod_{\substack{p/b_i \\ p>r}}\left(1-\frac{1}{p}\right)\right| \leqq 2^{\nu(b_i)} < 2^{2\log n/\log\log n} \tag{20}$$

(with respect to (5)). Thus in view of (17),

(21) $$D_i(P_r, u_2) > \frac{n}{P_r} \prod_{\substack{p/b_i \\ p>r}} \left(1-\frac{1}{p}\right) - 2^{2\log n/\log\log n} >$$

$$> \left(1-\frac{\varepsilon}{4}\right)\frac{n}{P_r} - 2^{2\log n/\log\log n} > \left(1-\frac{\varepsilon}{2}\right)\frac{n}{P_r}$$

(for n large).

Denoting the number of solutions of (19) by v_i, we have by (16) and (21)

(22) $$v_i \geqq A_{(P_r,u_2)}(n) - \sum_{\substack{d \leqq n \\ d\equiv u_2 \,(\mathrm{mod}\, P_r) \\ (b_i,d)>1}} 1 =$$

$$= A_{(P_r,u_2)}(n) - \Big(\sum_{\substack{d\leqq n \\ d\equiv u_2\,(\mathrm{mod}\, P_r)}} 1 - D_i(P_r, u_2)\Big) <$$

$$> 2\varepsilon\frac{n}{P_r} - \left(\frac{n}{P_r}+1\right) + \left(1-\frac{\varepsilon}{2}\right)\frac{n}{P_r} = \frac{3\varepsilon}{2}\frac{n}{P_r} - 1 > \varepsilon\frac{n}{P_r}.$$

Let $d_1^{(i)} < \ldots < d_{w_i}^{(i)}$ denote the sequence of those integers d for which

(23) $$(b_i, d) = 1, \quad d \in A_{(P_r,u_2)} \quad \text{and} \quad \prod_{\substack{p/d \\ p>r}} \left(1-\frac{1}{p}\right) > 1-\frac{\varepsilon}{4}.$$

Lemma 2 yields by (10) and (22) that

(24) $$w_i \geqq v_i - \frac{\varepsilon}{4}\frac{n}{P_r} > \varepsilon\frac{n}{P_r} - \frac{\varepsilon}{4}\frac{n}{P_r} > \frac{\varepsilon}{2}\frac{n}{P_r}.$$

Let us denote the number of solutions of

(25) $$(b_i, a_y) = (d_j^{(i)}, a_y) = 1, \quad a_y \in A_{(P_r,u_3)}$$

(for i, j fixed) by $z_j^{(i)}$.

By (15), (17) and (23), if $d\equiv u_3 \,(\mathrm{mod}\, P_r)$ and $p/(b_i, e)$ or $p/(d_j^{(i)}, e)$ then $p>r$. Denote by $E_j^{(i)}(P_r, u_3)$ the number of those integers e for which $e\leqq n$, $e\equiv u_3 \,(\mathrm{mod}\, P_r)$ and $(b_i, e) = (d_j^{(i)}, e) = 1$. With respect to (5), we have

(26) $$\left| E_j^{(i)}(P_r, u_3) - \frac{n}{P_r} \prod_{\substack{p/b_i d_j^{(i)} \\ p>r}} \left(1-\frac{1}{p}\right)\right| < 2^{\nu(b_i d_j^{(i)})} <$$

$$< 2^{2\log n^2/\log\log n^2} < 2^{4\log n/\log\log n}.$$

We obtain from (17), (23) and (26) for sufficiently large n that

$$E_j^{(i)}(P_r, u_3) > \frac{n}{P_r} \prod_{\substack{p \mid b_i d_j^{(i)} \\ p > r}} \left(1 - \frac{1}{p}\right) - 2^{4 \log n / \log \log n} = \tag{27}$$

$$= \frac{n}{P_r} \prod_{\substack{p \mid b_i \\ p > r}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \mid d_j^{(i)} \\ p > r}} \left(1 - \frac{1}{p}\right) - 2^{4 \log n / \log \log n} >$$

$$> \frac{n}{P_r} \left(1 - \frac{\varepsilon}{4}\right) \left(1 - \frac{\varepsilon}{4}\right) - 2^{4 \log n / \log \log n} > \left(1 - \frac{\varepsilon}{2}\right) \frac{n}{P_r}.$$

(16) and (27) yield that

$$z_j^{(i)} \geqq A_{(P_r, u_3)}(n) - \Big(\sum_{\substack{e \leqq n \\ e \equiv u_3 \pmod{P_r}}} 1 - E_j^{(i)}(P_r, u_3) \Big) > \tag{28}$$

$$> 2\varepsilon \frac{n}{P_r} - \left(\frac{n}{P_r} + 1\right) + \left(1 - \frac{\varepsilon}{2}\right) \frac{n}{P_r} = \frac{3\varepsilon}{2} \frac{n}{P_r} - 1 > \varepsilon \frac{n}{P_r}.$$

By (17), (23) and (25), the triplets b_i, $d_j^{(i)}$, a_y satisfy

$$(b_i, d_j^{(i)}) = (b_i, a_y) = (d_j^{(i)}, a_y) = 1, \quad b_i, d_j^{(i)}, a_y \in A,$$

and by (18), (24) and (28), their number is greater than

$$\varepsilon \frac{n}{P_r} \cdot \frac{\varepsilon}{2} \frac{n}{P_r} \cdot \varepsilon \frac{n}{P_r} = c_{12}(\varepsilon) n^3$$

which completes the proof of Theorem 8.

References

[1] P. Erdős, Some remarks about additive and multiplicative functions, *Bull. Amer. Math. Soc.* (1946), 527—537.

[2] P. Erdős, Remarks on number theory, V. Extremal problems in number theory, II. (in Hungarian), *Mat. Lapok* **17** (1966), 135—155.

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

(Received April 23, 1977.)