Reprinted from Diophantine Analysis, edited by J.Loxton and A.van der Poorten, LMS Lecture Notes vol 109 ©Cambridge University Press 1986. Printed in Great Britain.

ON PRIME FACTORS OF SUMS OF INTEGERS II

C.L. Stewart
University of Waterloo, Ontario N2L 3G1, Canada

R. Tijdeman
University of Leiden, 2300 RA Leiden, The Netherlands

In a part of a paper [3] devoted to his lifelong collaboration with Turán, Erdös wrote (in slightly different notation): "In our first joint paper [4] we proved the following theorem: Let $1 < a_1 < \ldots < a_k$ be any sequence of k integers. Then ($\omega(m)$ is the number of distinct prime factors of m)

$$\omega\Big(\prod_{1\le i<j\le k}(a_i+a_j)\Big) > \frac{\log k}{\log 2}. \tag{1}$$

We always thought that (1) is very far from being the best possible but never could improve it. log k/log 2 can probably be replaced by $k^{1-\varepsilon}$ and perhaps even by k/log k. If $a_i = i$ we see that (1) cannot hold with $(2+\varepsilon)k/\log k$. On the other hand we have no counterexample to

$$\omega\Big(\prod_{1\le i<j\le k}(a_i+a_j)\Big) > (2+o(1))\frac{k}{\log k}. \tag{2}$$

Perhaps such a counterexample will not be difficult to find.

It is annoying that we never could settle our conjecture: To every s there is an f(s) so that if $k > f(s)$ and $1 < a_1 < \ldots < a_k$, $1 < b_1 < \ldots < b_k$ are any two sets of k integers then

$$\omega\Big(\prod_{\substack{1\le i\le k\\ 1\le j\le k}}(a_i+b_j)\Big) > s. \tag{3}$$

I would not be surprised if the proof of (3) is easy and we again overlooked a simple argument." In this paper we give such an elementary argument and a survey of related results.

1. NOTATION

Let $A = \{a_1,\ldots,a_k\}$ be a finite set of positive integers and $B = \{b_1,\ldots,b_\ell\}$ a finite set of non-negative integers. Put

$$\Pi_1 = \prod_{a,b\in A} (a+b), \qquad \Pi_2 = \prod_{\substack{a\in A\\ b\in B}} (a+b).$$

We assume throughout the paper that there is no prime number which divides each sum $a + b$. For the problems which we consider it is no loss of generality to do so.

Let $\omega(n)$ denote the number of distinct prime factors of n and let $P(n)$ denote the greatest prime factor of n. For any set X, let $|X|$ denote the cardinality of X. By $c_1, c_2, \ldots$ we shall denote certain effectively computable positive constants and by $C_1(\varepsilon), C_2(\varepsilon), \ldots$ certain effectively computable positive numbers depending only on ε. We thank A. Schinzel for a useful suggestion.

2. THE CASE $|A| = |B| = 2$

The first non-trivial case is $k = 2$. Obviously $\omega(\Pi_1) > 1$. If $a_1 = 1$ and $a_2 (= 2^p-1)$ is a Mersenne prime, then $\omega(\Pi_1) = 2$. Since it is likely that there are infinitely many Mersenne primes, it is also likely that there are infinitely many sets A with $|A| = 2$ and $\omega(\Pi_1) = 2$. On the other hand, there are only finitely many sets A with $|A| = 2$ and $P(\Pi_1)$ bounded. In fact we have the following more general result.

Theorem 1. (Győry, Stewart and Tijdeman [7]). If a_1, a_2, b_1, b_2 run through positive integers such that g.c.d. $(a_1+b_1, a_1+b_2, a_2+b_1, a_2+b_2) = 1$ and $\max(a_1, a_2, b_1, b_2) \to \infty$ then $P(\Pi_2) \to \infty$.

Note that $P(\Pi_2) = \max_{a\in A, b\in B} P(a+b)$. Theorem 1 is an immediate consequence of a theorem of van der Poorten and Schlickewei [10]. Theorem A is a generalization of this result. We obtain Theorem 1 from Theorem A by taking $x_1 = a_1 + b_1$, $x_2 = -a_1 - b_2$, $x_3 = -a_2 - b_1$, $x_4 = a_2 + b_2$, $\varepsilon = 0$.

Theorem A. (Evertse [6]). Let $S = \{p_1,\ldots,p_s\}$ be a finite set of prime numbers and let n be a positive integer. Let ε be a real number with

$0 < \varepsilon < 1$. There are only finitely many n-tuples $(x_1,\ldots,x_n)$ of rational integers composed of prime numbers from S such that

$$x_1 + \ldots + x_n = 0,$$

$$\text{g.c.d. } (x_1,\ldots,x_n) = 1,$$

$$x_{i_1} + \ldots + x_{i_t} \neq 0$$

for each non-empty, proper subset $\{i_1,\ldots,i_t\}$ of $\{1,\ldots,n\}$, and

$$0 < \prod_{j=0}^{s} |x_1 \ldots x_n|_{p_j} < \Big(\max_{i=1,\ldots,n} |x_i|\Big)^{\varepsilon}$$

where $|\ |_{p_0}$ denotes the ordinary absolute value.

Since this theorem is proved by the p-adic analogue of the ineffective Thue-Siegel-Roth-Schmidt method, no non-trivial lower bound for $P(\Pi_2)$ can be derived from the proof of Theorem 1.

It would be interesting to have an upper bound for the number of possible n-tuples in Theorem A, even only for some fixed ε. For $n = 3$ such results are known. Mahler [9] proved that there are at most c_1^s triples x_1,x_2,x_3 as in Theorem A when $\varepsilon = 0$. This result was generalized by Evertse as follows.

Theorem B. (Evertse [5]). Let K be an algebraic number field of degree m, let λ,μ be non-zero elements of K and let S be a finite set of primes of K of cardinality s containing the infinite primes. Then the equation

$$\lambda x + \mu y = 1 \tag{4}$$

in S-units x,y has at most

$$3 \times 7^{m+2s}$$

solutions. (An element a of K is called an S-unit if $|a|_p = 1$ for all primes p outside S and its equivalence classes.)

Lewis and Mahler [8] showed that there are at most c_2^s triples x_1,x_2,x_3 satisfying the conditions of Theorem A with $n = 3$ and $\varepsilon = 1/19$. Using an idea of Evertse, we can state their result as follows (cf. [12]).

Theorem C. Let $p_1,\ldots,p_s$ be distinct prime numbers. There exists an effectively computable positive constant c_2 such that there are at most c_2^s pairs of coprime integers x,y satisfying

$$0 < \prod_{j=0}^{s} |xy(x-y)|_{p_j} < (\max(|x|,|y|))^{1/19}.$$

We shall give applications of these results in section 4.

3. LOWER BOUNDS FOR $\omega(\Pi_1)$ AND $\omega(\Pi_2)$: ELEMENTARY PROOFS

We recall a result mentioned in the introduction.

Theorem 2. (Erdös and Turán [4]). Let A be a set of positive integers with $|A| = k > 2$. Then

$$\omega(\Pi_1) > \frac{\log k}{\log 2}.$$

The elementary and elegant proof is based on a lemma which states that, for any odd prime p, every set of 2n positive integers contains a subset V of n integers satisfying

$$|v_1+v_2|_p = \max(|v_1|_p, |v_2|_p) \qquad \text{for all } v_1,v_2 \in V.$$

Erdös and Turán stressed the elementary character of their proof.

The following result answers Erdös' question mentioned in the introduction. The proof is elementary, but entirely different from the proof of Erdös and Turán.

Theorem 3. Let A and B be sets of positive integers with $|A| = |B| = k > 3$. Then

$$\omega(\Pi_2) > c_3 \frac{\log k}{\log\log k}.$$

The proof is based on the following application of the box principle.

Lemma 1. Let x,y,g and n be integers greater than one with $g < y$ and let $d_1,\dots,d_m$ be distinct integers with $y < d_i < x$ for $i = 1,\dots,m$. Put $s = \omega(d_1 \dots d_m)$. If

$$m > n((3e \log x)/\log(y/g))^s \tag{5}$$

then there exist n distinct integers $d_{\ell_1},\dots,d_{\ell_n}$ from $(d_1,\dots,d_m)$ for which

$$\text{g.c.d.}\ (d_{\ell_1},\dots,d_{\ell_n}) > g.$$

Proof. Let $p_1,\dots,p_s$ be the prime numbers which divide $d_1 \dots d_m$ and write

$$d_i = p_1^{\ell_{1,i}} \dots p_s^{\ell_{s,i}},$$

for $i = 1,\dots,m$. To d_i we associate the vector $v_i = (\ell_{1,i} \log p_1,\dots,\ell_{s,i} \log p_s)$ in $\mathbf{R}^s$. Notice that all the vectors $v_1,\dots,v_m$ lie in

$$D = \{(y_1,\dots,y_s) \mid y_1 + \dots + y_s < \log x \text{ and } 0 < y_i \text{ for } i=1,\dots,s\}.$$

Put $w = [(s \log x)/\log(y/g)] + 1$ and observe that

$$(s \log x)/\log(y/g) < w < (2s \log x)/\log(y/g). \tag{6}$$

For any positive integers $j_1,\dots,j_s$ define the box

$$D_{j_1,\dots,j_s} = \{(y_1,\dots,y_s) \mid ((j_i-1)\log x)/w < y_i < (j_i \log x)/w,\ i=1,\dots,s\}.$$

Let M be the number of such boxes for which

$$\sum_{i=1}^{s} (j_i-1)/w < 1.$$

Notice that these M boxes cover D and are contained in

$$D_0 = \{(y_1,\dots,y_s) \mid y_1+\dots+y_s < (1+(s/w))\log x \text{ and } 0 < y_i \text{ for } i=1,\dots,s\}.$$

Each box has volume $((\log x)/w)^s$. The volume of D_0 is $((1+(s/w))\log x)^s/s!$. Therefore

$$M < (s+w)^s/s!,$$

and so, by (6), and the inequality $s! > (s/e)^s$,

$$M < ((3e \log x)/\log(y/g))^s.$$

If (5) holds m/M is at least n and so there is one box $D_{j_1,\ldots,j_s}$ which contains at least n vectors. Let $d_{\ell_1},\ldots,d_{\ell_n}$ be the integers associated with these vectors and put

$$g_1 = \text{g.c.d.}\ (d_{\ell_1},\ldots,d_{\ell_n}).$$

Since $d_{\ell_1},\ldots,d_{\ell_n}$ are in the same box,

$$\log|d_{\ell_i}|_{p_j}^{-1} > \log|d_{\ell_1}|_{p_j}^{-1} - (\log x)/w,$$

for $i = 1,\ldots,n$ and $j = 1,\ldots,s$. Thus, by the product formula for valuations,

$$\log g_1 = \sum_{j=1}^{s} \log|g_1|_{p_j}^{-1} > \sum_{j=1}^{s} (\log|d_{\ell_1}|_{p_j}^{-1} - (\log x)/w),$$

hence

$$\log g_1 > \log d_{\ell_1} - (s \log x)/w.$$

But $d_{\ell_1} > y$ and thus

$$g_1 > yx^{-s/w},$$

hence, by (6), $g_1 > g$ as required. □

Lemma 1 is used twice in the proof of Lemma 2.

Lemma 2. Let c,k and s be integers with $c > 6$, $k > 2$, $s > 2$. Let $S = \{p_1,\ldots,p_s\}$ be a finite set of prime numbers. Let $A = \{a_1,\ldots,a_k\}$ and $B = \{b_1,\ldots,b_k\}$ be sets of positive integers such that $a_i + b_j$ is composed of prime numbers from S for $i,j = 1,\ldots,k$. Suppose $a_1 < \ldots < a_k$ and $b_1 < \ldots < b_k$ and put $N = a_k + b_k$. If $k > 2(10\ cs)^{2s}$ and $b_k > N^{1-1/cs}$, then $a_k > N^{1-2/cs}$ and there exist integers g and i_1, with $g > N^{1-6/c}$ and $1 < i_1 < k$ such that

$$g \mid a_{i_1} + b_j \qquad \text{for} \quad j = 1,\ldots,k.$$

Proof. We apply Lemma 1 with $m = k$, $x = N$, $y = N^{1-1/cs}$ and $g = N^{1-2/cs}$ to the numbers a_1+b_k, $a_2+b_k,\ldots,a_k+b_k$. We obtain that there exist integers $i_1,\ldots,i_n$ with $1 < i_1 < i_2 < \ldots < i_n < k$ and $n > 2(10\ cs)^s$ such that

$$g_1 > N^{1-2/cs}$$

where g_1 = g.c.d. $(a_{i_1}+b_k,\ldots,a_{i_n}+b_k)$. Since g_1 divides $a_{i_2} - a_{i_1}$, we have $a_{i_2} > g_1$. This gives the first assertion. Thus, for any j from $\{1,\ldots,k\}$, the numbers $a_{i_2}+b_j$, $a_{i_3}+b_j$, ..., $a_{i_n}+b_j$ lie between $N^{1-2/cs}$ and N. Therefore, by Lemma 1 applied with $m = n-1$, $x = N$, $y = N^{1-2/cs}$ and $g = N^{1-3/cs}$, there exist two integers $h_1(j)$ and $h_2(j)$ from $\{i_2,i_3,\ldots,i_n\}$ with $h_1(j) < h_2(j)$ for which

$$g_2(j) > N^{1-3/cs}$$

where $g_2(j)$ = g.c.d. $(a_{h_1(j)} + b_j,\ a_{h_2(j)} + b_j)$. Put $g_3(j) =$ g.c.d. $(g_1,g_2(j))$ and notice that both g_1 and $g_2(j)$ divide $a_{h_2(j)} - a_{h_1(j)}$. Since $a_{h_2(j)} - a_{h_1(j)} < a_{h_2(j)} < N$, we have

$$g_3(j) = \frac{g_1 g_2(j)}{\text{l.c.m.}(g_1,g_2(j))} > \frac{N^{1-2/cs}N^{1-3/cs}}{N} = N^{1-5/cs}.$$

Notice that $g_3(j)$ divides $b_k - b_j$ and is composed of primes from $\{p_1,\ldots,p_s\}$. Furthermore

$$\left|\frac{g_1}{g_3(j)}\right|_{p_\sigma}^{-1} \leq \left|\frac{g_1}{g_3(j)}\right| \leq N^{5/cs}$$

for $j = 1,\ldots,k$ and $\sigma = 1,\ldots,s$. Hence

$$\prod_{\sigma=1}^{s} \max_{j=1,\ldots,k} \left|\frac{g_1}{g_3(j)}\right|_{p_\sigma}^{-1} \leq N^{5/c}.$$

This implies that there exists an integer g_4 with $g_4 | g_3(j)$ for $j = 1,\ldots,k$ and

$$g_4 \geq g_1 N^{-5/c} \geq N^{1-2/cs-5/c} \geq N^{1-6/c}.$$

It follows that $g_4 | g_1$ and $g_4 | b_k - b_j$ for $j = 1,\ldots,k$. Hence

$$g_4 | a_i + b_j$$

for $i = i_1$ and $j = 1,\ldots,k$. □

<u>Proof of Theorem 3</u>. Let $a_1,\ldots,a_k$ denote the elements of A and $b_1,\ldots,b_k$ the elements of B. We shall suppose $0 < a_1 < \ldots < a_k$ and $0 < b_1 < \ldots < b_k$. Further we may assume, without loss of generality, that there is no prime number p such that

$$p | a_i + b_j \qquad \text{for} \quad i,j = 1,\ldots,k. \tag{7}$$

Finally we may assume that $a_k < b_k$. Put $N = a_k + b_k$.

We shall prove that if

$$k > 10^{6s} s^{2s} \tag{8}$$

then

$$\omega(\Pi_2) > s.$$

This suffices to establish our result. For $s = 1$ it is obvious. We shall therefore assume that

$$\omega(\Pi_2) = s > 1$$

and that k satisfies (8) and we shall show that this implies (7) for some prime p. Let $p_1,\ldots,p_s$ be the primes which divide Π_2.

Since $N > k > 2^{20s}$, we have $b_k > N/2 > N^{1-1/20s}$. Therefore, by Lemma 2 with c = 20, we have $a_k > N^{1-1/10s}$ and there exist integers g_5 and i_1 with $g_5 > N^{7/10}$ and $1 < i_1 < k$ such that

$$g_5 | a_{i_1} + b_j \qquad \text{for} \quad j = 1,\ldots,k.$$

By applying Lemma 2 with A and B interchanged and c = 10, we obtain that there exist integers g_6 and j_1 with $g_6 > N^{2/5}$ and $1 < j_1 < k$ such that

$$g_6 | a_i + b_{j_1} \qquad \text{for} \quad i = 1,\ldots,k.$$

Put g = g.c.d. (g_5,g_6) and notice that both g_5 and g_6 divide $a_{i_1} + b_{j_1}$.

Hence

$$g = \frac{g_5 g_6}{\text{l.c.m.}(g_5,g_6)} > \frac{N^{7/10}N^{2/5}}{N} = N^{1/10} > 1.$$

Since $a_i + b_j = (a_i+b_{j_1}) - (a_{i_1}+b_{j_1}) + (a_{i_1}+b_j)$, we have

$$g | a_i+b_j \qquad \text{for} \quad i,j = 1,\ldots,k.$$

Hence there exists a prime p such that (7) holds. □

4. LOWER BOUNDS FOR $\omega(\Pi_1)$ AND $\omega(\Pi_2)$: NON-ELEMENTARY PROOFS

It is easy to derive the following refinement of Theorem 2 from Mahler's result mentioned in section 2 which was published just one year earlier than the paper of Erdös and Turán appeared. Mahler's result has a non-elementary proof, however.

Theorem 4. Let $A = \{a_1,\ldots,a_k\}$ be a set of positive integers with $k > 2$. Put $B = \{0,a_1\}$. Then

$$\omega\Big(\prod_{i=1}^{k}((a_i+a_1)(a_i+a_i))\Big) \geq \omega(\Pi_2) > c_4 \log k. \tag{9}$$

Proof. Let $\prod_{i=1}^{k}(a_i(a_i+a_1))$ be composed of the prime numbers $p_1,\ldots,p_s$. Then each of $a_1,a_2,\ldots,a_k,a_1+a_1,a_2+a_1,\ldots,a_k+a_1$ is composed of these primes. Put $d_i = \text{g.c.d.}(a_i,a_i+a_1)$, $x_i = (a_i+a_1)/d_i$ and $y_i = a_i/d_i$ for $i = 1,\ldots,k$. Then $\text{g.c.d.}(x_i,y_i) = 1$ for $i = 1,\ldots,k$ and all the pairs (x_i,y_i) are distinct. Moreover, each pair x_i,y_i satisfies

$$\prod_{j=0}^{s}|x_iy_i(x_i-y_i)|_{p_j} = 1.$$

Hence $k < c_1^s$ by Mahler's result (or by Theorem B or C). □

It is possible to refine Theorem 3 in a similar way. The essential difference with the above proof is that we no longer know that $x_i - y_i$ is composed of the prime numbers $p_1,\ldots,p_s$ and we can therefore not apply Mahler's result. However, for $i = 1,\ldots,k$, each pair $a_i + b_1$, $a_i + b_2$ provides a solution of the equation $(b_1-b_2)^{-1}x + (b_2-b_1)^{-1}y = 1$. Hence a straightforward application of Theorem B yields the following result.

Theorem 5. (Győry, Stewart and Tijdeman [7]). Let A and B be sets of positive integers with $k = |A| \geq |B| \geq 2$. Then

$$\omega(\Pi_2) > c_5 \log k. \tag{10}$$

We note that a slightly weaker result, namely

$$\omega(\Pi_2) > c_6 \log k/\log\log k$$

can be derived by using Lemma 1 and Theorem C. See Stewart and Tijdeman [12].

5. UPPER BOUNDS FOR $\omega(\Pi_1)$ AND $\omega(\Pi_2)$

It is likely that Theorems 2 and 3 are far from best possible results, but we have nothing to add to the words of Erdös cited in the introduction. However, we shall show that Theorems 4 and 5 are not far from being best possible.

Theorem 6. Let ε be a positive real number. For every $k > C_1(\varepsilon)$ there exist positive integers $a_1,\ldots,a_k$ such that

$$\omega\Big(\prod_{i=1}^{k}(a_i(a_i+a_1))\Big) < 2(\log k)^{2+\varepsilon}.$$

This implies that the right-hand sides of (9) and (10) cannot be replaced by $(\log k)^{2+\varepsilon}$ for sufficiently large k.

The proof of Theorem 6 is based on the following counterpart of Mahler's estimate.

Theorem D. (Stewart and Tijdeman). Let $\varepsilon > 0$. Let $p_1,\ldots,p_s$ be the first s prime numbers. Then for $s > C_2(\varepsilon)$ there exists a positive integer $a < e^s$ such that the equation

$$x - y = a$$

has at least $\exp(s^{1/2-\varepsilon})$ solutions in relatively prime positive integers x and y composed of $p_1,\ldots,p_s$. Further there are arbitrarily large finite sets S of prime numbers such that the equation

$$x - y = z$$

with $(x,y,z) = 1$ has at least $\exp(|S|^{1/2-\varepsilon})$ solutions in positive integers x,y and z composed of primes from S.

The proof of Theorem D is based on an estimate of de Bruijn [2] for the number of positive integers $< x$ composed of prime numbers $< y$.

Proof of Theorem 6. We may assume $0 < \varepsilon < 1$. Let $p_1,\ldots,p_s$ be the first s prime numbers where $s > C_2(\varepsilon/8)$. Apply Theorem D. Let a_1 be an integer

with $0 < a_1 < e^s$ such that the equation $x - y = a_1$ has at least $\exp(s^{1/2-\varepsilon/8})$ solutions in integers x,y composed of $p_1,\ldots,p_s$. Denote these solutions (x,y) with $y \neq a_1$ by

$$(a_2+a_1,a_2),\ (a_3+a_1,a_3),\ \ldots,\ (a_k+a_1,a_k).$$

Then

$$\omega\Big(\prod_{i=1}^{k}(a_i(a_i+a_1))\Big) < \omega(2a_1) + s.$$

Since $a_1 < e^s$, we have $\omega(2a_1) < s$. Further $k > \exp(s^{1/2-\varepsilon/8})$, hence $s < (\log k)^{2/(1-\varepsilon/4)} < (\log k)^{2+\varepsilon}$. Combining these inequalities we obtain

$$\omega\Big(\prod_{i=1}^{k}(a_i(a_i+a_1))\Big) < 2s < 2(\log k)^{2+\varepsilon}. \qquad \square$$

6. LOWER BOUNDS FOR $P(\Pi_1)$ AND $P(\Pi_2)$

Lower bounds for $\omega(\Pi_1)$ and $\omega(\Pi_2)$ give immediately lower bounds for $P(\Pi_1)$ and $P(\Pi_2)$. In fact by applying the prime number theorem we deduce the following result from Theorem 5.

Theorem 7. Let A and B be sets of positive integers with $k = |A| > |B| > 2$ and $k > 3$. Then

$$P(\Pi_2) > c_7 \log k \log\log k. \tag{11}$$

By Theorem 11 the right hand side of inequality (11) can not be replaced by $(\log k)^{2+\varepsilon}$ for any $\varepsilon > 0$. If, however, at least one of the terms of A or B is large and the terms have no common factor we are able to improve upon (11).

Theorem 8. (Györy, Stewart and Tijdeman [7]). Let ε be a positive real number and let $a_1,\ldots,a_k$ and b be positive integers. Put $A = (a_1,\ldots,a_k)$ and $B = (0,b)$. If $k > C_3(\varepsilon)$ and

$$\text{g.c.d.}\ (a_1,\ldots,a_k,b) = 1$$

then

$$P(\Pi_2) > \min((1-\varepsilon)k \log k,\ c_8 \operatorname{loglog}(a_k+b)).$$

For the proof of Theorem 8 we employ estimates for linear forms in logarithms due to Baker and, in the p-adic case, to van der Poorten.

Let N be a positive integer. If A and B are dense subsets of $\{1,\dots,N\}$ then it is possible to improve on the above estimates. For example Balog and Sárközy used the large sieve inequality to prove the following theorem.

Theorem 9. (Balog and Sárközy [1]). Let N be a positive integer and let A and B be subsets of $\{1,\dots,N\}$. If

$$(|A||B|)^{1/2} > 10N^{1/2}\log N$$

and $N > c_9$, then

$$P(\Pi_2) > \frac{(|A||B|)^{1/2}}{16 \log N}.$$

Further, Sárközy and Stewart used the Hardy-Littlewood circle method to establish the following result.

Theorem 10. (Sárközy and Stewart [11]). Let ε be a positive real number, let N be a positive integer and let A and B be subsets of $\{1,\dots,N\}$. Put $R = 3N/(|A||B|)^{1/2}$. If

$$(|A||B|)^{1/2} > N^{5/6+\varepsilon}$$

and $N > C_4(\varepsilon)$, then

$$P(\Pi_2) > c_{10} \frac{(|A||B|)^{1/2}}{\log R \operatorname{loglog} R}.$$

7. UPPER BOUNDS FOR $P(\Pi_1)$ AND $P(\Pi_2)$

The following theorem shows that the right hand side of inequality (11) cannot be replaced by $(\log k)^{2+\varepsilon}$ for any $\varepsilon > 0$.

Theorem 11 Let ε be a positive real number. For every $k > C_5(\varepsilon)$ there exist positive integers $a_1,\ldots,a_k$ and a positive integer b such that, with $A = \{a_1,\ldots,a_k\}$ and $B = \{0,b\}$,

$$P(\Pi_2) < (\log k)^{2+\varepsilon}. \tag{12}$$

Proof. We may assume $\varepsilon < 1$. Let s be the number of primes not exceeding $(\log k)^{2+\varepsilon}$. Then, by the prime number theorem,

$$s > \frac{(\log k)^{2+\varepsilon}}{3\ \mathrm{loglog}\ k} > (\log k)^{2/(1-\varepsilon/3)} \tag{13}$$

if $k > C_6(\varepsilon)$. Let $p_1,\ldots,p_s$ be the first s prime numbers. If k is so large that $s > C_2(\varepsilon/6)$ then, according to Theorem D, there exists a positive integer b such that the equation $x - y = b$ has at least $\exp(s^{1/2-\varepsilon/6})$ solutions in positive integers x,y. Note that, by (13),

$$\exp(s^{1/2-\varepsilon/6}) > k.$$

Let $(x,y) = (a_1+b,a_1)$, (a_2+b,a_2), ..., (a_k+b,a_k) be a set of solutions. Then $a_1,\ldots,a_k$ and b yield (12). □

The proof of Theorem 11 does not guarantee that P(b) is small. Such a result is given by the following theorem which should be compared with Theorems 6 and 11. The proof of Theorem 12 is self-contained.

Theorem 12. Let ε be a positive real number. For every $k > C_7(\varepsilon)$ there exist positive integers $a_1,a_2,\ldots,a_k$ such that

$$P\Big(\prod_{i=1}^{k}(a_i(a_i+a_1))\Big) < \exp(((2+\varepsilon)\log k\ \mathrm{loglog}\ k)^{1/2}). \tag{14}$$

Proof. We may assume $\varepsilon < 1$. Let p be the smallest prime number such that

$$k < \exp((1-\varepsilon/8)(\log p)^2/2\mathrm{loglog}\ p). \tag{15}$$

Let $p_1,p_2,\ldots$ denote the prime numbers in increasing order and let $p = p_s$. By the prime number theorem we have

$$s > \frac{p}{2 \log p}, \tag{16}$$

for $k > c_{11}$. Put

$$t = \frac{(1-\varepsilon/8)\log p}{2 \operatorname{loglog} p} \quad \text{and} \quad u = [t] + 3. \tag{17}$$

Let S denote the set of positive integers at most p^u which are composed of $p_1,\ldots,p_s$. Then, by (16),

$$|S| > \frac{s^u}{u!} > (\frac{s}{u})^u > (\frac{p}{2u \log p})^u.$$

Thus, by (17),

$$|S| > 2p^{u-1}$$

for $k > C_8(\varepsilon)$. Consider the gaps between consecutive elements of S. Since an interval of length p^u can contain no more than p^{u-1} distinct intervals of length p, there are at least p^{u-1} pairs of integers x,y in S with $0 < x-y < p$. By (17) and (15) we have

$$p^{u-1} > p^{t+1} > pk.$$

Hence for some integer a_1 with $0 < a_1 < m$, there are at least k pairs x,y from S such that $x - y = a_1$ and so k distinct elements $y = a_1,\ldots,a_k$ in S for which the equation $x - y = a_1$ has a solution x in S. Since

$$k > \exp((1-\varepsilon/8)(\log p_{s-1})^2/2 \operatorname{loglog} p_{s-1})$$

we have

$$(1-\varepsilon/8)(\log p_{s-1})^2/2 \operatorname{loglog} p_{s-1} < \log k \quad \text{and} \quad p_{s-1} < k$$

for $k > C_9(\varepsilon)$. Since $p_s < 2p_{s-1}$ by Bertrand's postulate, we obtain

$$P(\prod_{i=1}^{k} (a_i(a_i+a_1))) < p_s < 2p_{s-1} < 2\exp(((2+\varepsilon/2)\log k \operatorname{loglog} k)^{1/2})$$

$$< \exp(((2+\varepsilon)\log k \operatorname{loglog} k)^{1/2})$$

for $k > C_{10}(\varepsilon)$. □

We remark that it is possible to replace the factor $2 + \varepsilon$ on the right hand side of inequality (14) by $1 + \varepsilon$ by using estimates of de Bruijn [2] as in the proof of Theorem D.

REFERENCES

1. A. Balog and A. Sárközy, On sums of sequences of integers, II, Acta Math. Hungar., to appear.
2. N.G. de Bruijn, On the number of positive integers $\leq x$ and free of prime factors $> y$, II, Nederl. Akad. Wetensch. Proc. Ser. A 69 = Indag. Math. 28 (1966), 239-247.
3. P. Erdös, Problems in number theory and combinatorics, Proc. Sixth Manitoba Conf. Numerical Math. (Univ. Manitoba, Winnipeg, Man., 1976), pp. 35-58, Congress Numer., 18. Utilitas Math., Winnipeg, Man., 1977.
4. P. Erdös, and P. Turán, On a problem in the elementary theory of numbers, Amer. Math. Monthly 41 (1934), 608-611.
5. J.-H. Evertse, On equations in S-units and the Thue-Mahler equation, Invent. Math. 75 (1984), 561-584.
6. J.-H. Evertse, On sums of S-units and linear recurrences, Compositio Math. 53 (1984), 225-244.
7. K. Györy, C.L. Stewart and R. Tijdeman, On prime factors of sums of integers I, Compositio Math., to appear.
8. D.J. Lewis and K. Mahler, On the representation of integers by binary forms, Acta Arith. 6 (1961), 333-363.
9. K. Mahler, Zur Approximation algebraischer Zahlen I,II, Math. Ann. 107 (1933), 691-730 and 108 (1933), 37-55.
10. A.J. van der Poorten and H.P. Schlickewei, The growth conditions for recurrence sequences, Macquarie Math. Report 82-0041 (1982).
11. A. Sárközy and C.L. Stewart, On divisors of sums of integers II, J. Reine Angew. Math., to appear.
12. C.L. Stewart and R. Tijdeman, On prime factors of sums of integers, Univ. Leiden Math. Inst. Report 11 (1985).

The research of the first author was supported in part by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.