Publ. Math. Debrecen
56 / 3-4 (2000), 559–573

On prime factors of integers of the form $ab + 1$

By A. SÁRKÖZY (Budapest) and C. L. STEWART (Waterloo)

To Professor Kálmán Győry
on the occasion of his sixtieth birthday

Abstract. Let N be a positive integer and let A and B be subsets of $\{1,\dots,N\}$. In this article we discuss estimates for the least prime factor and the greatest prime factor of integers of the form $ab+1$ where a is taken from A and b is taken from B.

1. Introduction

If n is a positive integer, p is a prime number and k is a non-negative integer with $p^k \mid n$, $p^{k+1} \nmid n$ then we write $p^k \| n$. For $n > 1$ let $p(n)$ and $P(n)$ denote the least and greatest prime factor of n, respectively.

In the last 15 years many papers have been written on the arithmetical properties of elements of sum sets $A+B$ (defined as the set of the integers of the form $a+b$ with $a \in A$, $b \in B$) where A and B are two "dense" sets of positive integers. In particular, it has been shown that

(i) (Sárközy and Stewart [10]) If $\delta > 0$, N is a positive integer with $N > N_0(\delta)$, $A, B \subset \{1, 2, \dots, N\}$ and

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

Mathematics Subject Classification: 11N36, 11B75.
Key words and phrases: large sieve, greatest prime factor, least prime factor.
Research of the first author partially supported by Hungarian National Foundation for Scientific Research, Grant No. T 029759.
Research of the second author supported in part by Grant A3528 of the Natural Sciences and Engineering Research Council of Canada.

then there are $a \in A$, $b \in B$ with

$$P(a+b) > \frac{c_1(|A|\,|B|)^{1/2}}{\log R \, \log\log R}$$

where $c_1 = c_1(\delta)$ is a positive number and

$$R = \frac{3N}{(|A|\,|B|)^{1/2}}. \tag{1.1}$$

(So that

$$A, B \subset \{1, 2, \dots, N\}, \quad |A|, |B| > \varepsilon N \tag{1.2}$$

and $N > N_1(\varepsilon)$ imply that there are $a \in A$, $b \in B$ with

$$P(a+b) > c_2(\varepsilon)N.\,) \tag{1.3}$$

(ii) (Sárközy and Stewart [11]) If k is a positive integer with $k \geq 2$, $\delta > 0$, N is a positive integer with $N > N_0(\delta, k)$, $A, B \subset \{1, 2, \dots, N\}$ and

$$(|A|\,|B|)^{1/2} > N^{1-\theta_k+\delta}$$

where $\theta_k = (1 + 2k \cdot 4^{k-1})^{-1}$, then there are $a \in A$, $b \in B$ and a prime p with

$$p^k \mid (a+b)$$

and

$$p^k > \frac{c_1(|A|\,|B|)^{1/2}}{\exp(c_2(\log k \, \log R)/\log\log R)}$$

where $c_1 = c_1(\delta, k)$ and $c_2 = c_2(\delta, k)$ are positive numbers and R is defined by (1.1). So that (1.2) and $N > N_1(\varepsilon, k)$ imply that there are $a \in A$, $b \in B$ and a prime p with

$$p^k \mid (a+b), \quad p^k > c_3(\varepsilon, k)N.$$

(iii) (Sárközy and Stewart [12]) If $\beta > 0$, $1/2 < \theta < 1$, N is a positive integer with $N > N_0(\beta, \theta)$, $A, B \subset \{1, 2, \dots, N\}$ and

$$(|A|\,|B|)^{1/2} \geq N^{\theta},$$

then there is a prime number p with

$$\beta < p \le \left(\frac{\log N}{2}\right)^{1/(2\theta-1)}$$

such that every residue class modulo p contains a member of $A+B$. So that having these assumptions, there are $a \in A$, $b \in B$ with

$$p(a+b) \le \left(\frac{\log N}{2}\right)^{1/(2\theta-1)}.$$

2. The results

One might like to study the multiplicative analogues of sum set results. One way of doing this, proposed by Sárközy [7], is to replace the sums $a+b$ by the numbers $ab+1$ (see also [4] and [8]). However, it should be noted that the first result on the arithmetic properties of numbers $ab+1$ is due, probably, to Vinogradov (see Chapter V of [14]). Let p be a prime number and k be an integer coprime with p. Let $\left(\frac{n}{p}\right)$ denote the Legendre symbol of n over p. Vinogradov established the estimate

$$\left|\sum_{a\in A}\sum_{b\in B}\left(\frac{ab+k}{p}\right)\right| < (2|A||B|p)^{1/2}.$$

This result can be considered as the multiplicative analogue of the recent results of Friedlander and Iwaniec [2] on sums of the form $\sum_{a\in A}\sum_{b\in B}\chi(a+b)$ where χ is a non-principal character modulo a prime p.

In this paper, first we will prove the multiplicative analogue of result (iii).

Theorem 1. *Let N be a positive integer and let θ and β be real numbers with $1/2 < \theta < 1$. There is a positive number c, which is effectively computable in terms of θ and β, such that if A and B are subsets of $\{1,\dots,N\}$ with*

$$(|A|\,|B|)^{1/2} \ge N^{\theta},$$

and N exceeds c then there is a prime number p with

$$\beta < p \le \left(\frac{\log N}{2}\right)^{1/(2\theta-1)},$$

and integers a in A and b in B such that p divides $ab+1$.

Next, we will study the multiplicative analogue of result (i). Almost certainly the following conjectures are true.

Conjecture 1. *For each positive real number ε there are positive real numbers $N_0(\varepsilon)$ and $c(\varepsilon)$ such that if N exceeds $N_0(\varepsilon)$ and (1.2) holds, then there are a in A and b in B with*

$$P(ab+1) > c(\varepsilon)N^2.$$

Conjecture 2. *For each positive real number ε and each integer k, with $k \geq 2$, there are positive real numbers $N_0(\varepsilon, k)$ and $c(\varepsilon, k)$ such that if N exceeds $N_0(\varepsilon, k)$ and (1.2) holds, then there are a in A and b in B and a prime p with*

$$p^k \mid ab+1 \quad \text{and} \quad p^k > c(\varepsilon, k)N^2.$$

However, these conjectures seem to be hopelessly difficult.

For the additive case of these conjectures we have applied the Hardy–Littlewood method [10], [11]. Since the multiplicative problems are of a binary nature, the Hardy–Littlewood method fails completely here. Another approach used in several related papers [1], [12] and [13] is based on the application of the large sieve. In the multiplicative case this approach works too, however one gets only relatively weak partial results. In case of Conjecture 1, the natural limit of this approach is to show that assuming (1.2), there are a, b with

$$P(ab+1) \gg N.$$

By an elementary argument, reminiscent both of Gallagher's larger sieve [3] and of Ruzsa's argument in [6], we shall show that

$$\max_{a\in A, b\in B} \frac{P(ab+1)}{N} \to +\infty.$$

Theorem 2. *For each $\varepsilon > 0$ there are numbers $N_0 = N_0(\varepsilon)$ and $C = C(\varepsilon)$, which are effectively computable in terms of ε, such that if $N > N_0$, $A, B \subset \{1, 2, \ldots, N\}$ and*

$$\min(|A|, |B|) > C\frac{N}{\log N}, \tag{2.1}$$

then there are a in A and b in B such that

$$P(ab+1) > (1-\varepsilon)\min(|A|, |B|)\log N. \tag{2.2}$$

In the case of Conjecture 2, again the Hardy–Littlewood method fails for the same reasons. The method of the proof of Theorem 2 fails as well. Again, the application of the large sieve gives a partial result. By a straightforward application of the prime power moduli large sieve in [9], one gets, assuming (1.2), that there are $a \in A$, $b \in B$ and a prime p with

$$p^k \mid (ab+1)$$

and

$$p^k > c(\varepsilon, k)\left(\frac{N}{\log N}\right)^{k/(2k-1)}. \tag{2.3}$$

However, this lower bound is not quite satisfactory. Namely, the natural limit of the sieve approach seems to be that the sifting moduli can be as large as cN; this would correspond to

$$p^k > c(\varepsilon, k)N \tag{2.4}$$

in place of (2.3). We shall treat this problem in a subsequent paper by means of an improved version of the prime power moduli large sieve.

3. Proof of Theorem 1

Lemma 1 (Large sieve)**.** *Let M and N be integers with N positive. Let A be a set of integers in the interval $[M+1, M+N]$. For each prime p let $\nu(p)$ denote the number of residue classes* (mod p) *which contain a member of A. Then for any positive number Q we have*

$$|A| \le \frac{N+Q^2}{L}$$

where

$$L = \sum_{q \le Q}{}' \prod_{p|q} (p - \nu(p))/\nu(p),$$

the dash indicating the sum is over square-free positive integers q.

Proof. See Theorem 7.1 of [5]. □

Lemma 2. *Let p be an integer with $p \ge 3$ and let*

$$D = \left\{ (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1 + x_2 \le 1 + \frac{1}{p} \quad \text{and} \quad \frac{1}{p} \le x_i \le \frac{p-1}{p} \text{ for } i = 1, 2 \right\}.$$

Then

$$\min_D \left(\frac{1}{x_1} - 1 \right) \left(\frac{1}{x_2} - 1 \right) = \frac{1}{2} \left(\frac{p-2}{p-1} \right).$$

Proof. Put $f(x_1, x_2) = \left(\frac{1}{x_1} - 1 \right) \left(\frac{1}{x_2} - 1 \right)$. We readily check that there are no local maxima or minima of f in D and so the minimum occurs on the boundary. Next note that on that part of the boundary of D with either $x_1 = \frac{1}{p}$ or $x_2 = \frac{1}{p}$ one has $f(x_1, x_2) \ge 1$. Further for that part of the boundary where either $x_1 = \frac{p-1}{p}$ or $x_2 = \frac{p-1}{p}$, $f(x_1, x_2) \ge \frac{1}{2} \left(\frac{p-2}{p-1} \right)$ with equality holding when (x_1, x_2) is $\left(\frac{p-1}{p}, \frac{2}{p} \right)$ or $\left(\frac{2}{p}, \frac{p-1}{p} \right)$. Finally on the line segment from $\left(\frac{2}{p}, \frac{p-1}{p} \right)$ to $\left(\frac{p-1}{p}, \frac{2}{p} \right)$ we find that the minimum value is attained at the endpoints and so our result follows. □

Lemma 3. *Let N be a positive integer and let A and B be non-empty subsets of $\{1, \dots, N\}$. Let α and β be real numbers with $\alpha > 1$. Let T be the set of primes p which satisfy $\beta < p \le \left(\frac{\log N}{2} \right)^\alpha$ and let S be a subset of T consisting of all but at most $2 \log N$ elements of T. There is a real number C which is effectively computable in terms of α and β such that if N exceeds C and*

$$(|A|\,|B|)^{1/2} \ge \frac{N^{\frac{1+1/\alpha}{2}}}{10} \tag{3.1}$$

then there is a prime p from S and integers a in A and b in B such that p divides $ab + 1$.

Proof. Suppose the contrary. For each prime p let $\nu_1(p)$ denote the number of residue classes modulo p which contain an element of A and let $\nu_2(p)$ denote the number of those which contain an element of B. It follows from the large sieve inequality, Lemma 1 (applying it to estimate both $|A|$ and $|B|$), that, for each $Q \geq 1$,

$$(|A|\,|B|)^{1/2} \leq (N + Q^2)/H, \tag{3.2}$$

where

$$H = \left(\prod_{i=1}^{2} \sum_{q \leq Q}{}' \prod_{p|q} \left(\frac{p}{\nu_i(p)} - 1 \right) \right)^{1/2}$$

and where the dash indicates the sum is over square-free positive integers q. From the Cauchy–Schwarz inequality we have

$$H \geq \sum_{q \leq Q}{}' \left(\prod_{p|q} \left(\frac{p}{\nu_1(p)} - 1 \right) \left(\frac{p}{\nu_2(p)} - 1 \right) \right)^{1/2}.$$

Let $Q = N^{1/2}$ and let R be the set of integers from $\{1, \ldots, [N^{1/2}]\}$ composed of $\left[\frac{\log Q}{\alpha \log\log Q}\right]$ distinct primes p from S with $p \geq 11$. Then

$$H \geq \sum_{r \in R} \left(\prod_{p|r} \left(\frac{p}{\nu_1(p)} - 1 \right) \left(\frac{p}{\nu_2(p)} - 1 \right) \right)^{1/2}. \tag{3.3}$$

By assumption, for each prime p from S the congruence $ab \equiv -1 \pmod{p}$ has no solution with $a \in A$, $b \in B$. Let $\nu_1^*(p)$ denote the number of residue classes different from the residue class of 0 that contain an element of A and let $\nu_2^*(p)$ be the number for B. We have $\nu_1^*(p) + \nu_2^*(p) \leq p - 1$ and thus $\nu_1(p) + \nu_2(p) \leq p + 1$. By Lemma 2

$$\left(\frac{p}{\nu_1(p)} - 1 \right) \left(\frac{p}{\nu_2(p)} - 1 \right) \geq \frac{1}{2} - \frac{1}{2(p-1)},$$

for $p \in S$. Thus for $p \in S$ with $p \geq 11$ we plainly have

$$\left(\frac{p}{\nu_1(p)} - 1 \right) \left(\frac{p}{\nu_2(p)} - 1 \right) \geq \frac{9}{20}. \tag{3.4}$$

Therefore from (3.2), (3.3) and (3.4) we see that

$$(|A|\,|B|)^{1/2} \leq \frac{2N}{H'} \tag{3.5}$$

where

$$H' \geq \left(\frac{9}{20}\right)^{\frac{1}{2}\left[\frac{\log Q}{\alpha \log\log Q}\right]} |R|. \tag{3.6}$$

It remains to estimate $|R|$. Let S' be the subset of S of primes bigger than 10. Then by the prime number theorem with error term,

$$|S'| \geq \pi((\log Q)^{\alpha}) - \pi(\beta) - \pi(10) - 2\log N > \frac{(\log Q)^{\alpha}}{\alpha \log\log Q} \tag{3.7}$$

provided that $N > c_1$, where $c_1, c_2, \dots$ are positive numbers which are effectively computable in terms of α and β. We now count the number of distinct ways of choosing $[\log Q/\alpha \log\log Q]$ primes from S'. Each choice gives rise to a distinct square-free integer, given by the product of the primes, which does not exceed Q. Then

$$|R| \geq \binom{|S'|}{\left[\frac{\log Q}{\alpha \log\log Q}\right]} \geq \frac{\left(|S'| - \left[\frac{\log Q}{\alpha \log\log Q}\right]\right)^{\frac{\log Q}{\alpha \log\log Q}-1}}{\left[\frac{\log Q}{\alpha \log\log Q}\right]!}.$$

Thus by (3.7) and Stirling's formula

$$|R| \geq \frac{\left(\frac{(\log Q)^{\alpha}}{\alpha \log\log Q}\left(1 - \frac{1}{(\log Q)^{\alpha-1}}\right)\right)^{\frac{\log Q}{\alpha \log\log Q}}}{(\log Q)^{\alpha+1}\left(\frac{\log Q}{e\alpha \log\log Q}\right)^{\frac{\log Q}{\alpha \log\log Q}}},$$

for $N > c_2$. Since $\log(1-x) > -2x$ for $0 < x < 1/2$,

$$|R| \geq Q^{1-1/\alpha} e^{\left(\frac{\log Q}{\alpha \log\log Q} - \frac{2(\log Q)^{2-\alpha}}{\alpha \log\log Q}\right)} (\log Q)^{-\alpha-1}, \tag{3.8}$$

for $N > c_3$. Further, since $\left(\frac{20}{9}\right)^{1/2} < e$, it follows from (3.6) and (3.8) that

$$H' > 20 Q^{1-1/\alpha},$$

for $N > c_4$. Therefore, by (3.5),

$$(|A|\,|B|)^{1/2} < \frac{N}{10}^{1-(1/2)(1-1/\alpha)} = \frac{N}{10}^{\frac{1+1/\alpha}{2}},$$

for $N > c_5$ which contradicts (3.1). The result now follows.

Proof of Theorem 1. Let S be the set of primes p which satisfy $\beta < p \le (\log(N^{1/2}))^{1/(2\theta-1)}$. Put $\alpha = 1/(2\theta - 1)$ and note that α is a real number greater than one since $\frac{1}{2} < \theta < 1$. Theorem 1 now follows from Lemma 3 on noting that $(1+1/\alpha)/2 = \theta$. □

4. Proof of Theorem 2

First note that we may assume $|A| = |B|$. Put

$$|A| = |B| = Z. \tag{4.1}$$

Let

$$E = \prod_{a\in A}\prod_{b\in B}(ab+1).$$

Then clearly we have

$$\begin{aligned} E &\ge \prod_{\substack{a\in A\\ a\ge \varepsilon Z/10}} \prod_{\substack{b\in B\\ b\ge \varepsilon Z/10}} \left(\left(\frac{\varepsilon Z}{10}\right)^2 + 1\right) \\ &> \left(\frac{\varepsilon Z}{10}\right)^{2(|A|-\varepsilon Z/10)(|B|-\varepsilon Z/10)} = \left(\frac{\varepsilon Z}{10}\right)^{2(1-\varepsilon/10)^2 Z^2}. \end{aligned} \tag{4.2}$$

If C and N are large enough in terms of ε, then it follows from (2.1), (4.1) and (4.2) that

$$\log E > 2\left(1-\frac{\varepsilon}{10}\right)^2 Z^2 \log\left(\frac{\varepsilon Z}{10}\right) > 2\left(1-\frac{\varepsilon}{5}\right) Z^2 \log N. \tag{4.3}$$

If p is a prime with $p \le N^2+1$, then define $u(p)$ by $p^{u(p)} \| E$, and for each positive integer k write $\alpha(p,k) = |\{(a,b) : a\in A,\ b \in B,\ p^k \mid ab+1\}|$ so that

$$E = \prod_{p\le N^2+1} p^{u(p)}$$

where

$$u(p) = \sum_{k \le \frac{\log(N^2+1)}{\log p}} \alpha(p,k). \tag{4.4}$$

Write

$$T = (1-\varepsilon)\min(|A|,|B|)\log N,$$

and let P_1, P_2 and P_3 denote the set of the primes p with $p \le N$, $N < p \le T$ and $T < p \le N^2+1$, respectively, and write

$$E = E_1 E_2 E_3 \tag{4.5}$$

where

$$E_i = \prod_{p \in P_i} p^{u(p)} \quad \text{for } i = 1,2,3. \tag{4.6}$$

Then it suffices to prove that $E_3 > 1$, or, equivalently, that

$$\log E_3 > 0. \tag{4.7}$$

Next we will give an upper bound for E_1. If U is a subset of $\{1,\dots,N\}$, m is a positive integer and h is an integer, then write

$$r(U,h,m) = |\{n : n \in U,\ n \equiv h \pmod{m}\}|.$$

When $(h,m)=1$, let $\overline{h}(m)$ denote the integer from $\{1,\dots,m\}$ with

$$h\overline{h}(m) \equiv -1 \pmod{m}.$$

We shall need the following lemma.

Lemma 4. *If N is a positive integer and $U \subset \{1,2,\dots,N\}$ then we have*

$$\sum_{p \le N} \log p \sum_{k \le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} \left(r(U,h,p^k)\right)^2 \le |U|\left(|U|-1+\pi(N)\right)\log N. \tag{4.8}$$

Proof of Lemma 4. Write

$$D(U) = \prod_{\substack{n,n' \in U \\ n' < n}} (n - n'),$$

and for $p \le N$ define the integer $v(U,p)$ by $p^{v(U,p)} \| D(U)$. Then we have

$$
\begin{aligned}
\sum_{p\le N} v(U,p)\log p &= \log \prod_{p\le N} p^{v(U,p)} = \log D(U) \le \log \prod_{\substack{n,n'\in U\\ n'<n}} N \\
&= |\{(n,n') : n,n'\in U,\ n'<n\}| \log N \qquad (4.9)\\
&= \binom{|U|}{2} \log N.
\end{aligned}
$$

Moreover, defining $\beta(m,p)$ by $p^{\beta(m,p)} \| m$, clearly we have

$$
\begin{aligned}
v(U,p) &= \sum_{\substack{n,n'\in U\\ n'<n}} \beta(n-n',p) = \sum_{\substack{n,n'\in U\\ n'<n}} \left|\left\{k : k \le \frac{\log N}{\log p},\ p^k \mid n-n'\right\}\right| \\
&= \sum_{k\le \frac{\log N}{\log p}} \left|\left\{(n,n') : n,n'\in U,\ n'<n,\ p^k \mid n-n'\right\}\right| \\
&= \sum_{k\le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} \left|\left\{(n,n') : n,n'\in U,\ n'<n,\ n\equiv n'\equiv h \pmod{p^k}\right\}\right| \\
&= \sum_{k\le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} \binom{|\{n : n\in U,\ n\equiv h \pmod{p^k}\}|}{2} \\
&= \sum_{k\le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} \binom{r(U,h,p^k)}{2} \\
&= \sum_{k\le \frac{\log N}{\log p}} \left(\frac12 \sum_{h=1}^{p^k} \big(r(U,h,p^k)\big)^2 - \frac12 \sum_{h=1}^{p^k} r(U,h,p^k)\right) \\
&= \frac12 \sum_{k\le \frac{\log N}{\log p}} \left(\sum_{h=1}^{p^k} \big(r(U,h,p^k)\big)^2 - |U|\right)
\end{aligned}
$$

whence

$$\sum_{p\le N} v(U,p)\log p = \frac{1}{2}\sum_{p\le N}\log p \sum_{k\le \frac{\log N}{\log p}}\left(\sum_{h=1}^{p^k}(r(U,h,p^k))^2 - |U|\right)$$

$$\geq \frac{1}{2}\sum_{p\le N}\left(\log p\sum_{k\le \frac{\log N}{\log p}}\sum_{h=1}^{p^k}(r(U,h,p^k))^2 - |U|\log N\right) \tag{4.10}$$

$$= \frac{1}{2}\left(\sum_{p\le N}\log p\sum_{k\le \frac{\log n}{\log p}}\sum_{h=1}^{p^k}(r(U,h,p^k))^2 - |U|\pi(N)\log N\right).$$

It follows from (4.9) and (4.10) that the left hand side of (4.8) is

$$\leq 2\binom{|U|}{2}\log N + |U|\pi(N)\log N = |U|\left(|U| - 1 + \pi(N)\right)\log N$$

and this completes the proof of the lemma. □

By (4.4) and (4.6), we may estimate $\log E_1$ in the following way:

$$\log E_1 = \sum_{p\in P_1} u(p)\log p = \sum_{p\le N}\left(\sum_{k\le\frac{\log(N^2+1)}{\log p}}\alpha(p,k)\right)\log p \tag{4.11}$$

$$= \sum_{p\le N}\log p\sum_{k\le\frac{\log(N^2+1)}{\log p}}|\{(a,b) : a\in A,\ b\in B,\ ab\equiv -1 \pmod{p^k}\}|$$

$$= \sum\nolimits_1 + \sum\nolimits_2$$

where in $\sum_1$ we sum over $p\le N$, $k\le\frac{\log N}{\log p}$, while in $\sum_2$ we have $p\le N$, $\frac{\log(N+1)}{\log p}\le k\le\frac{\log(N^2+1)}{\log p}$. Using the inequality $|xy|\le\frac{1}{2}(x^2+y^2)$ we obtain

$$\sum\nolimits_1 = \sum_{p\le N}\log p\Bigg(\sum_{k\le\frac{\log N}{\log p}}\sum_{\substack{1\le h\le p^k\\(h,p^k)=1}}|\{a : a\in A,\ a\equiv h \pmod{p^k}\}|$$

$$\cdot\,|\{b : b\in B,\ b\equiv \overline{h}(p^k) \pmod{p^k}\}|\Bigg)$$

$$= \sum_{p \le N} \log p \sum_{k \le \frac{\log N}{\log p}} \sum_{\substack{1 \le h \le p^k \\ (h,p^k)=1}} r(A,h,p^k) r(B,\overline{h}(p^k),p^k)$$

$$= \frac{1}{2} \sum_{p \le N} \log p \sum_{k \le \frac{\log N}{\log p}} \left(\sum_{\substack{1 \le h \le p^k \\ (h,p^k)=1}} r^2(A,h,p^k) + \sum_{\substack{1 \le h \le p^k \\ (h,p^k)=1}} r^2(B,\overline{h}(p^k),p^k) \right)$$

$$= \frac{1}{2} \sum_{p \le N} \log p \sum_{k \le \frac{\log N}{\log p}} \sum_{\substack{1 \le h \le p^k \\ (h,p^k)=1}} (r^2(A,h,p^k) + r^2(B,h,p^k)).$$

Using Lemma 4 with A, respectively B, in place of U, by (4.1) we obtain

$$\begin{aligned} \textstyle\sum_1 &\le \frac{1}{2}(|A|(|A|-1+\pi(N)) + |B|(|B|-1+\pi(N))) \log N \\ &= Z(Z-1+\pi(N)) \log N \le (Z^2 + Z\pi(N)) \log N. \end{aligned} \tag{4.12}$$

To estimate $\sum_2$ observe that if $N < p^k$ and a is fixed (which can be done in $|A|$ ways), then, by $B \subset \{1,2,\dots,N\}$, there is at most one $b \in B$ with

$$ab \equiv -1 \pmod{p^k}.$$

It follows that

$$\begin{aligned} \textstyle\sum_2 &\le \sum_{p \le N} \log p \sum_{k \le \frac{\log(N^2+1)}{\log p}} |A| \\ &\le |A| \sum_{p \le N} \log p \frac{\log(N^2+1)}{\log p} < 3Z \log N \pi(N). \end{aligned} \tag{4.13}$$

By (2.1), (4.11), (4.12) and (4.13), for large enough C we have

$$\log E_1 < (Z^2 + 4Z\pi(N)) \log N < \left(1 + \frac{\varepsilon}{5}\right) Z^2 \log N. \tag{4.14}$$

To estimate $\log E_2$, observe that for $p > N$, $k \ge 2$, $A, B \subset \{1,2,\dots,N\}$ we have $\alpha(p,k) = 0$. Moreover, for $p > N$ and fixed $a \le N$ there is at most one $b \le N$ with $p \mid ab+1$. Thus by (4.1) and (4.4), for $p \in P_2$ we have

$$\begin{aligned} u(p) &= \sum_{k \le \frac{\log(N^2+1)}{\log p}} \alpha(p,k) = \alpha(p,1) \\ &= |\{(a,b) : a \in A,\ b \in B,\ p \mid ab+1\}| \qquad (4.15) \\ &= \sum_{a \in A} |\{b : b \in B,\ p \mid ab+1\}| \le \sum_{a \in A} 1 = |A| = Z. \end{aligned}$$

It follows from (4.6) and (4.15) that

$$\log E_2 = \sum_{p \in P_2} u(p) \log p \le Z \sum_{N < p \le T} \log p.$$

By (2.1), (4.1) and the definition of T, and using the prime number theorem, for C large we obtain that

$$\log E_2 \le Z(1+o(1))(T-N) < \left(1 - \frac{4\varepsilon}{5}\right) Z^2 \log N. \qquad (4.16)$$

It follows from (4.3), (4.5), (4.14) and (4.16) that

$$\begin{aligned} \log E_3 = \log E - \log E_1 - \log E_2 &> 2\left(1 - \frac{\varepsilon}{5}\right) Z^2 \log N \\ &- \left(1 + \frac{\varepsilon}{5}\right) Z^2 \log N - \left(1 - \frac{4\varepsilon}{5}\right) Z^2 \log N = \frac{\varepsilon}{5} Z^2 \log N > 0 \end{aligned}$$

which proves (4.7).

References

[1] A. Balog and A. Sárközy, On sums of sequences of integers, II, *Acta Math. Hung.* **44** (1984), 169–179.

[2] J. Friedlander and H. Iwaniec, Estimates for character sums, *Proc. Amer. Math. Soc.* **119** (1993), 365–372.

[3] P. X. Gallagher, A larger sieve, *Acta Arith.* **18** (1971), 77–81.

[4] K. Győry, A. Sárközy and C. L. Stewart, On the number of prime factors of integers of the form $ab+1$, *Acta Arith.* **74** (1996), 365–385.

[5] H. E. Richert, Lectures on sieve methods, *Tata Institute of Fundamental Research*, *Bombay*, 1976.

[6] I. Ruzsa, Large prime factors of sums, *Studia Sci. Math. Hungar.* **27** (1992), 463–470.

[7] A. Sárközy, On sums $a+b$ and numbers of the form $ab+1$ with many prime factors, *Grazer Math. Ber.* **318** (1992), 141–154.

[8] A. Sárközy, On the average value for the number of divisors of numbers of form $ab+1$, *Acta Math. Hungar.* **66** (1995), 223–245.

[9] A. Sárközy, A note on the arithmetic form of the large sieve, *Studia Sci. Math. Hungar.* **27** (1992), 83–95.

[10] A. Sárközy and C. L. Stewart, On divisors of sums of integers, II, *J. Reine Angew. Math.* **365** (1986), 171–191.

[11] A. Sárközy and C. L. Stewart, On divisors of sums of integers, IV, *Canadian J. Math.* **40** (1988), 788–816.

[12] A. Sárközy and C. L. Stewart, On divisors of sums of integers, V, *Pacific J. Math.* **166** (1994), 373–384.

[13] G. N. Sárközy, On a problem of P. Erdős, *Acta Math. Hungar.* **60** (1992), 271–282.

[14] I. M. Vinogradov, An introduction to the theory of numbers, *Pergamon Press, London and New York*, 1955.

A. SÁRKÖZY
DEPARTMENT OF ALGEBRA AND NUMBER THEORY
EÖTVÖS LORÁND UNIVERSITY
H–1088 BUDAPEST, RÁKÓCZI ÚT 5.
HUNGARY

E-mail: sarkozy@cs.elte.hu

C. L. STEWART
DEPARTMENT OF PURE MATHEMATICS
THE UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO N2L 3G1
CANADA

E-mail: cstewart@watserv1.uwaterloo.ca

(Received August 23, 1999; revised April 4, 2000)