On prime factors of integers which are sums or shifted products

by C.L. Stewart (Waterloo)

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 prime factors of integers of the form $a + b$ and $ab + 1$ where a is from A and b is from B.

1 Introduction

Let A and B be subsets of the first N integers. What information can be gleaned about integers of the form $ab+1$ or $a+b$, with a in A and b in B, from knowledge of the cardinalities of A and B? If A and B are dense subsets of $\{1, \dots, N\}$ then one might expect the integers $a+b$ with a in A and b in B, to have similar arithmetical characteristics to those of the first $2N$ integers and the integers $ab+1$, with a in A and b in B, to have similar arithmetical characteristics to those of the first N^2+1 integers. This phenomenon has been demonstrated in several papers. Even if A and B are not dense subsets of $\{1, \dots, N\}$ it is still possible to deduce some non-trivial estimates for the prime divisors of integers of the form $a+b$ and $ab+1$. In this article we shall survey the estimates which have been obtained for the greatest prime factors of integers $a+b$ and $ab+1$ and for the number of distinct prime factors of the products

$$\prod_{a\in A,\, b\in B} (a+b) \qquad \text{and} \qquad \prod_{a\in A,\, b\in B} (ab+1).$$

2 Results for general sets of integers

For any set X let $|X|$ denote its cardinality and for any integer n with $n \geq 2$ let $P(n)$ denote the greatest prime factor of n and $\omega(n)$ denote the number of distinct prime factors of n. In 1934 in their first joint paper Erdős and Turán [10] proved that if A is a non-empty set of positive integers then

$$\omega\left(\prod_{a,a'\in A} (a+a')\right) \geq \frac{\log|A|}{\log 2}$$

and they asked if a result of this type holds when the summands are taken from different sets.

2000 Mathematics Subject Classification 11N36, 11B75.

Key words and phrases: greatest prime factor, sieve.

This research was supported in part by the Canada Research Chairs Program and by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

In 1986 Győry, Stewart and Tijdeman [15] extended the result of Erdős and Turán to the case where the summands are taken from different sets. They proved, by means of a result on S-unit equations due to Evertse [11] that there is a positive number c_1 such that for any finite sets A and B of positive integers with $|A| \geq |B| \geq 2$,

$$\omega\left(\prod_{a\in A,\, b\in B}(a+b)\right) > c_1 \log |A|. \tag{1}$$

Also in 1986 Stewart and Tijdeman [29] gave an elementary argument to establish a slightly weaker result. They proved that there is a positive number c_2 such that if $|A| = |B| \geq 3$ then

$$\omega\left(\prod_{a\in A,\, b\in B}(a+b)\right) \geq c_2 \frac{\log |A|}{\log\log |A|}.$$

In 1988 Erdős, Stewart and Tijdeman [9] showed that (1) could not be improved by much when they showed that the right hand side of (1) cannot be replaced by $(1/8+\varepsilon)(\log |A|)^2 \log\log |A|$ for any $\varepsilon > 0$. In fact more generally let ε be a real number with $0 < \varepsilon < 1$. They proved that there is a positive number $C(\varepsilon)$, which depends on ε, such that if k and ℓ are integers with k larger than $C(\varepsilon)$ and $2 \leq \ell \leq \log k/\log\log k$ then there exist distinct positive integers $a_1, \ldots, a_k$ and distinct non-negative integers $b_1, \ldots, b_\ell$ such that

$$P\left(\prod_{i=1}^{k}\prod_{j=1}^{\ell}(a_i+b_j)\right) < \left((1+\varepsilon)\frac{\log k}{\ell}\log\left(\frac{\log k}{\ell}\right)\right)^{\ell}. \tag{2}$$

We note that by the Prime Number Theorem it is an immediate consequence of (1) that there is a positive number c_3 such that for any finite sets A and B of positive integers with $|A| \geq |B| \geq 2$, there exist a in A and b in B for which

$$P(a+b) > c_3 \log |A| \log\log |A|.$$

In 1992 Sárközy [22] initiated the study of the multiplicative analogues of results of the above type where one replaces the sums $a+b$ by $ab+1$. In 1996 Győry, Sárközy and Stewart [16] proved the analogue of (1). In particular they proved that if A and B are finite sets of positive integers with $|A| \geq |B| \geq 2$ then

$$\omega\left(\prod_{a\in A,\, b\in B}(ab+1)\right) > c_4 \log |A|, \tag{3}$$

where c_4 is an effectively computable positive number. In fact both (1) and (3) are consequences of the following result established in [16]. Let $n \geq 2$ be an integer and let A and B be finite subsets of $\mathbb{N}^n$ with $|A| \geq |B| \geq 2(n-1)$.

Suppose that the n-th coordinate of each vector in A is equal to 1 and any n vectors in $B \cup (0, \dots, 0, 1)$ are linearly independent. There is an effectively computable positive number c_5 such that

$$\omega\left(\prod_{\substack{(a_1,\dots,a_n)\in A\\(b_1,\dots,b_n)\in B}} (a_1b_1 + \cdots + a_nb_n)\right) > c_5 \log|A|. \tag{4}$$

We obtain (1) by taking $n = 2$ and $b_1 = 1$ for all $(b_1, b_2) \in B$ and we obtain (3) by taking $n = 2$ and $b_2 = 1$ for all $(b_1, b_2) \in B$. The proof of (4) depends on work of Evertse and Győry [14] and of Evertse [13] on decomposable form equations and in turn this depends on quantitative versions of Schmidt's Subspace Theorem due to Schmidt [27] and Schlickewei [26].

Győry, Sarközy and Stewart [16] also established a multiplicative analogue of (2). Let ε be a positive real number and let k and ℓ be positive integers with $k \geq 16$ and

$$2 \leq \ell \leq \left(\frac{\log\log k}{\log\log\log k}\right)^{1/2}. \tag{5}$$

They proved that there is a positive number $C_1(\varepsilon)$, which is effectively computable in terms of ε, such that if k exceeds $C_1(\varepsilon)$ then there are sets of positive integers A and B with $|A| = k$ and $|B| = \ell$ for which

$$P\left(\prod_{a\in A}\prod_{b\in B}(ab+1)\right) < (\log k)^{\ell+1+\varepsilon}. \tag{6}$$

They also showed that if (6) is weakened by replacing the exponent $\ell + 1 + \varepsilon$ by 5ℓ then the range (5) for ℓ may be extended to

$$2 \leq \ell \leq c_6 \frac{\log k}{\log\log k},$$

for a positive number c_6.

3 Results for large terms

It follows from (3) and the Prime Number Theorem that if A is a finite set of positive integers with $|A| \geq 2$ then there exist distinct elements a and a' in A for which

$$P(aa' + 1) > c_7 \log|A| \log\log|A|$$

where c_7 is an effectively computable positive number. But what if the size of the integers increases as opposed to the size of the cardinality of A? Győry, Sárközy and Stewart [16] conjectured that if a, b and c denote distinct positive integers then

$$P((ab + 1)(ac + 1)(bc + 1)) \to \infty \tag{7}$$

as $\max(a, b, c) \to \infty$.

Stewart and Tijdeman [30] established the conjecture under the assumption that $\log a / \log(c+1) \to \infty$. Let a, b and c be positive integers with $a \geq b > c$. They proved that there is an effectively computable positive number c_8 for which

$$P((ab+1)(ac+1)(bc+1)) > c_8 \log\left(\frac{\log a}{\log(c+1)}\right). \tag{8}$$

Further, Stewart and Tijdeman [30] also proved that if a, b, c and d are positive integers with $a \geq b > c$ and $a > d$ there exists an effectively computable positive number c_9 such that

$$P((ab+1)(ac+1)(bd+1)(cd+1)) > c_9 \log\log a. \tag{9}$$

The proofs of both (8) and (9) depend on estimates for linear forms in the logarithms of algebraic numbers, see [32].

Győry and Sárközy [17] proved that the conjecture holds in the special case that at least one of the numbers a, b, c, a/b, b/c, a/c has bounded prime factors. This work, later refined by Bugeaud and Luca [3], depends on a result of Evertse [12] on the number of solutions of the S-unit equation and as a consequence does not lead to an effective lower bound in terms of a. Bugeaud [2] was able to give such a bound by applying an estimate of Loxton [19] for simultaneous linear forms in the logarithms of algebraic numbers. Let a, b and c be positive integers with $a \geq b > c$ and let α denote any element of the set $\{a, b, c, a/b, b/c, a/c\}$. Bugeaud proved that there is an effectively computable positive number c_{10} such that

$$P(\alpha(ab+1)(ac+1)(bc+1)) > c_{10} \log\log a.$$

The conjecture was finally established independently by Hernández and Luca [18] and Corvaja and Zannier [4] by means of Schmidt's Subspace Theorem. In fact Corvaja and Zannier [4] proved a strengthened version of the conjecture. They proved that if a, b and c are positive integers with $a > b > c$ then

$$P((ab+1)(ac+1)) \to \infty \quad \text{as } a \to \infty.$$

The results of Hernández and Luca and of Corvaja and Zannier are ineffective. Nevertheless Luca [20] was able to make them more explicit. For any prime number p and any integer x let $|x|_p$ denote the p-adic absolute value of x normalized so that $|p|_p = p^{-1}$. For any integer x and set of prime numbers S we put

$$|x|_{\overline{S}} = |x| \prod_{p \in S} |x|_p,$$

so that $|x|_{\overline{S}}$ is the largest divisor of x with no prime factors from S. Luca proved that if S is a finite set of prime numbers there exist positive numbers $C_1(S)$ and $C_2(S)$, which are not effectively computable, such that if a, b and c are positive integers with $a > b > c$ and $a > C_1(S)$ then

$$|(ab+1)(ac+1)|_{\overline{S}} > \exp\left(C_2(S) \frac{\log a}{\log\log a}\right).$$

An additive version of these results was established by Győry, Stewart and Tijdeman [15] in 1986. They proved, by means of a result of Evertse [12], that if a, b and c are distinct positive integers with g.c.d. $(a,b,c)=1$ then

$$P(ab(a+c)(b+c)) \to \infty$$

as $\max(a,b,c) \to \infty$.

4 Results for dense sets of integers

Let $\phi(x)$ denote the distribution function of the normal distribution so that

$$\phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-u^2/2} du.$$

Erdős, Maier and Sárközy [7] proved that an Erdős-Kac theorem applies to the sums $a+b$, counted with multiplicity, when a is from A, b is from B and A and B are dense subsets of $\{1,\ldots,N\}$. In particular they proved that there are positive numbers N_0 and C such that if N exceeds N_0 and ℓ is a positive integer then

$$\left| \frac{1}{|A||B|} \left|\{(a,b) : a \in A,\, b \in B,\, \omega(a+b) \leq \ell\}\right| - \phi\left(\frac{\ell - \log\log N}{(\log\log N)^{1/2}}\right) \right|$$

is at most $CN(|A||B|)^{-1/2}(\log\log N)^{-1/4}$. Tenenbaum [31] subsequently refined this result by replacing the factor $(\log\log N)^{-1/4}$ by $(\log\log N)^{-1/2}$. Elliott and Sárközy [5] obtained another refinement and later [6] they proved a result of similar character for integers of the form $ab+1$.

While the above results show that if A and B are dense subsets of $\{1,\ldots,N\}$ then the typical behaviour of $\omega(a+b)$ and $\omega(ab+1)$ is well understood one may still wonder about extreme values of these functions. For any positive integer N let $m = m(N) = \max\{\omega(k) : 1 \leq k \leq N\}$. One may check that

$$m = (1+o(1))\frac{\log N}{\log\log N} \quad \text{as } N \to \infty.$$

Erdős, Pomerance, Sárközy and Stewart [8] proved in 1993, by means of a combinatorial lemma due to Katona, that for each positive real number ε there are positive numbers $c(\varepsilon)$ and $N_1(\varepsilon)$ such that if N exceeds $N_1(\varepsilon)$ and A and B are subsets of the first N positive integers with $(|A||B|)^{1/2} > \varepsilon N$ then there exist integers a from A and b from B with

$$\omega(a+b) > m - c(\varepsilon)\sqrt{m}. \tag{10}$$

Sárközy [22] extended this result to the case where $A = B$ and $a+b$ in (10) is replaced by $aa'+1$ with $a, a' \in A$. In 1994, Sárközy and Stewart [24] showed that there are sums $a+b$ for which $\omega(a+b)$ is large provided that the weaker requirement

$$(|A||B|)^{1/2} \geq N^{\theta} \tag{11}$$

with $1/2 < \theta \leq 1$, applied. The corresponding result for $ab + 1$ was obtained by Győry, Sárközy and Stewart in [16]. Let θ be a real number with $1/2 < \theta \leq 1$. They proved that there is a positive number $C(\theta)$, which is effectively computable in terms of θ such that if N is a positive integer larger than $C(\theta)$ and A and B are subsets of $\{1, \ldots, N\}$ satisfying (11) then there exists an integer a from A and an integer b from B for which

$$\omega(ab+1) > \frac{1}{6}\left(\theta - \frac{1}{2}\right)^2 \frac{\log N}{\log\log N}. \tag{12}$$

The proof of (12) depends upon multiple applications of the large sieve inequality.

Balog and Sárközy [1] were the first to study the greatest prime factor of $a + b$ when A and B are dense subsets of $\{1, \ldots, N\}$. They proved, by means of the large sieve inequality, that there is a positive number N_1 such that if N exceeds N_1 and

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

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

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

In 1986 Sárközy and Stewart [23] refined this result for dense sets A and B by employing the Hardy-Littlewood method. In particular, it follows from their work that if $|A| \gg N$ and $|B| \gg N$ then there exist $\gg N^2/\log N$ pairs (a, b) with a in A and b in B such that

$$P(a+b) \gg N. \tag{13}$$

Put

$$Z = \min\{|A|, |B|\}.$$

In 1992 Ruzsa [21] proved that there exist a in A and b in B for which

$$P(a+b) > c_{11} Z \frac{\log Z}{\log N} \log\left(\frac{\log N}{\log Z}\right),$$

where c_{11} is a positive number. Furthermore he proved that for each positive real number ε there exists a positive number $C(\varepsilon)$ such that if Z exceeds $C(\varepsilon)N^{1/2}$ then there exist a in A and b in B with

$$P(a+b) > \left(\frac{2}{e} - \varepsilon\right) Z. \tag{14}$$

While estimates (13) and (14) are best possible, up to the determination of constants, a different situation applies for the multiplicative case. In this case we have the following conjecture of Sárközy and Stewart [25].

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 $Z > \varepsilon N$ then there are a in A and b in B such that*

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

Sárközy and Stewart [25] were able to give lower bounds for $P(ab+1)$ which are stronger than those for $P(a+b)$, such as (14), for dense sets A and B. In particular they showed that for each positive real number ε there are positive numbers $N_1(\varepsilon)$ and $K(\varepsilon)$, which are effectively computable in terms of ε, such that if N exceeds $N_1(\varepsilon)$ and Z exceeds $K(\varepsilon)N/\log N$ then there are a in A and b in B such that

$$P(ab+1) > (1-\varepsilon)Z\log N. \tag{15}$$

In fact the argument may be modified to give an estimate for $P(ab+1)$ of comparable strength to (15) for Z much smaller as our next result shows.

Theorem 1. *Let θ be a real number with $1/2 < \theta \le 1$. There are numbers $N_0 = N_0(\theta)$ and $C = C(\theta)$, which are effectively computable in terms of θ, such that if $N > N_0$, A, B are subsets of $\{1,\dots,N\}$, $Z = \min\{|A|,|B|\}$ and*

$$Z \ge N^\theta$$

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

$$P(ab+1) > CZ\log Z. \tag{16}$$

Note that, for comparison with (15) as opposed to (14), we may replace $CZ\log Z$ in (16) by $CZ\log N$.

Improvements on (15) and (16) have been obtained for sets which are more dense. For instance Stewart [28] proved that there are effectively computable positive numbers c_1, c_2 and c_3 such that if N exceeds c_1 and

$$Z > c_2\frac{N}{((\log N)/\log\log N)^{1/2}}, \tag{17}$$

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

$$P(ab+1) > N^{1+c_3(Z/N)^2}.$$

The proof employs Weil's estimates for Kloosterman sums. We shall prove the following more explicit version of the above result.

Theorem 2. *Let N be a positive integer, let A and B be subsets of $\{1,\dots,N\}$ and put $Z = \min\{|A|,|B|\}$. Let ε be a real number with $0 < \varepsilon < 1$. There are positive numbers c_1, c_2 and c_3, which are effectively computable in terms of ε, such that if N exceeds c_1 and (17) holds with the new value of c_2 then there are a in A and b in B for which*

$$P(ab+1) > \min(N^{1+(1-\varepsilon)(Z/N)^2}, (c_3(N/\log N)^{4/3}).$$

What happens in the extremal situation where A and B are both equal to $\{1,\ldots,N\}$? Shengli Wu [33] has recently shown, by means of the Bombieri-Vinogradov theorem, that if β is a real number larger than 10 and N is sufficiently large in terms of β then there exist integers a and b from $\{1,\ldots,N\}$ such that

$$P(ab+1) > \frac{N^2}{(\log N)^\beta}.$$

In this special case one has an estimate which approaches that of Conjecture 1.

5 Preliminary lemmas

For positive integers N and t we put

$$V_t(N) = \{(m,n) \in \mathbb{Z}\times\mathbb{Z} \mid 1 \le m \le N,\ 1 \le n \le N,\ t \mid mn+1\}$$

and denote by $d(t)$ the number of positive integers which divide t. In [28] Stewart deduced from Weil's estimates for Kloosterman sums the following result.

Lemma 1.

$$|V_t(N)| = \frac{\varphi(t)}{t^2}N^2 + O\left(t^{1/2}d(t)^{3/2}(\log t)^2 + \frac{Nd(t)\log t}{t}\right).$$

For the proofs of Theorems 1 and 2 we shall also require a minor variation on Lemma 4 of [25]. Let U be a subset of $\{1,\ldots,N\}$, m be a positive integer and h be an integer. We put

$$r(U,h,m) = |\{n : n \in U,\, n \equiv h (\operatorname{mod} m)\}|. \tag{18}$$

Lemma 2. *Let N and M be integers with $1 \le M \le N$ and let U be a subset of $\{1,\ldots,N\}$. Then*

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

Proof. We shall follow the proof of Lemma 4 of [25] which treats the case $M = N$. Put

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

We have

$$\sum_{p\le N} \log p \operatorname{ord}_p D(U) = \log D(U) \le \log\left(\prod_{\substack{n,n'\in U\\ n'<n}} N\right) = \binom{|U|}{2}\log N, \tag{19}$$

where ord_p denotes the p-adic order. Furthermore

$$\begin{aligned}
\mathrm{ord}_p D(U) &= \sum_{\substack{n,n' \in U \\ n' < n}} \mathrm{ord}_p(n - n') \\
&= \sum_{\substack{n,n' \in U \\ n' < n}} \left| \left\{ k : k \le \frac{\log N}{\log p},\, p^k \mid n - n' \right\} \right| \\
&= \sum_{k \le \frac{\log N}{\log p}} |\{(n, n') : n, n' \in U,\, n' < n,\, p^k \mid n - n'\}| \\
&= \sum_{k \le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} |\{(n, n') : n, n' \in U,\ n' < n,\ n \equiv n' \equiv h \pmod{p^k}\}| \\
&= \sum_{k \le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} \binom{r(U, h, p^k)}{2} \\
&= \sum_{k \le \frac{\log N}{\log p}} \left(\frac{1}{2} \sum_{h=1}^{p^k} (r(U, h, p^k))^2 - \frac{1}{2} \sum_{h=1}^{p^k} r(U, h, p^k) \right) \\
&= \frac{1}{2} \sum_{k \le \frac{\log N}{\log p}} \left(\sum_{h=1}^{p^k} (r(U, h, p^k))^2 - |U| \right). \qquad (20)
\end{aligned}$$

Therefore, by (19) and (20),

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

Since $\sum_{h=1}^{p^k} r(U, h, p^k)^2 - |U| \ge 0$ we see that

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

Therefore

$$\sum_{p \le M} \log p \sum_{k \le \frac{\log N}{\log p}} \sum_{h=1}^{p^k} r(U, h, p^k)^2 \le 2 \binom{|U|}{2} \log N + |U| \pi(M) \log N$$

as required. $\square$

6 An estimate from below

For the proofs of Theorems 1 and 2 we may assume, by removing terms from either A or B, if necessary, that

$$Z = \min(|A|, |B|) = |A| = |B|. \tag{21}$$

Define E by

$$E = \prod_{a \in A,\, b \in B} (ab + 1). \tag{22}$$

Let ε be a real number with $0 < \varepsilon < 1$. Then, by (21),

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

Therefore provided that $Z \geq N^{1/2}$, as in the hypotheses for Theorems 1 and 2, and that N is sufficiently large in terms of ε,

$$\frac{\varepsilon Z}{10} \geq Z^{1 - \frac{\varepsilon}{10}}$$

and so

$$\log E \geq 2 \left(1 - \frac{\varepsilon}{10} \right)^3 Z^2 \log Z. \tag{23}$$

For brevity we write

$$P = P \left(\prod_{a \in A,\ b \in B} (ab + 1) \right)$$

and we put

$$E_1 = \prod_{p \leq N} p^{\operatorname{ord}_p E}, \tag{24}$$

where the product is taken over primes p up to N. We shall require an upper bound for E_1 for the proof of Theorem 2.

Lemma 3. *Let $\varepsilon > 0$ and suppose that Z exceeds $N/(\log N)^{1/2}$. There exists a positive number $N_0(\varepsilon)$, which is effectively computable in terms of ε, such that for $N > N_0(\varepsilon)$,*

$$\log E_1 < (1 + \varepsilon) Z^2 \log N.$$

Proof. This follows from the proof of Theorem 2 of [25], see 4.14 of [25]. $\square$

7 Proof of Theorem 1

Our proof proceeds by a comparison of estimates for E, recall (22). Put

$$\delta = \theta - \frac{1}{2}.$$

By (23), for N sufficiently large in terms of δ,

$$\log E \geq (2-\delta) Z^2 \log Z. \tag{25}$$

We now observe that we may suppose that $P \leq N$. For if $P > N$ and $Z \leq K(1/2)N/\log N$, recall the definition of $K(1/2)$ from (15), then for N sufficiently large,

$$Z \log Z \leq K\left(\frac{1}{2}\right) \frac{N}{\log N} \log N = K\left(\frac{1}{2}\right) N < K\left(\frac{1}{2}\right) P.$$

Thus $P > (K(1/2))^{-1} Z \log Z$ as required. On the other hand if $Z > K(1/2)N/\log N$ then, by (15) with $\varepsilon = 1/2$,

$$P > \frac{1}{2} Z \log N \geq \frac{1}{2} Z \log Z,$$

for $N > N_1(1/2)$ as required. Therefore we may suppose that $P \leq N$.

We have

$$\begin{aligned}
\log E &= \sum_{p \leq P} \operatorname{ord}_p \left(\prod_{\substack{a \in A \\ b \in B}} (ab+1) \right) \log p \\
&= \sum_{p \leq P} \log p \sum_{k \leq \frac{\log(N^2+1)}{\log p}} \left| \{(a,b) : a \in A,\, b \in B,\, ab \equiv -1 \pmod{p^k}\} \right| \\
&= \sum\nolimits_1 + \sum\nolimits_2
\end{aligned} \tag{26}$$

where in $\sum_1$ we sum over $p \leq P$, $k \leq \log N/\log p$ while in $\sum_2$ we have $p \leq P$ and $\log(N+1)/\log p \leq k \leq \log(N^2+1)/\log p$.

We have

$$\sum\nolimits_1 = \sum_{p \leq P} \log p \left(\sum_{k \leq \frac{\log N}{\log p}} \sum_{\substack{1 \leq h \leq p^k \\ (h,p^k)=1}} \left| \{a \in A,\, a \equiv h \pmod{p^k}\} \right| \left| \{b \in B,\, b \equiv \overline{h} \pmod{p^k}\} \right| \right),$$

where for each integer h coprime with p^k we let $\overline{h}$ denote the unique integer with $1 \leq \overline{h} \leq p^k$ for which $h\overline{h} \equiv -1 \pmod{p^k}$. Therefore, by (18),

$$\sum\nolimits_1 = \sum_{p \leq P} \log p \sum_{k \leq \frac{\log N}{\log p}} \sum_{\substack{1 \leq h \leq p^k \\ (h,p^k)=1}} r(A,h,p^k) r(B,\overline{h},p^k).$$

Since $xy \le (1/2)(x^2+y^2)$ for any non-negative real numbers x and y we see that

$$\begin{aligned}\sum\nolimits_1 &\le \frac{1}{2}\sum_{p\le P}\log p \sum_{k\le\frac{\log N}{\log p}}\ \sum_{\substack{1\le h\le p^k\\(h,p)=1}} (r^2(A,h,p^k)+r^2(B,\overline{h},p^k))\\ &\le \frac{1}{2}\sum_{p\le P}\log p \sum_{k\le\frac{\log N}{\log p}}\ \sum_{\substack{1\le h\le p^k\\(h,p)=1}} (r^2(A,h,p^k)+r^2(B,h,p^k)).\end{aligned}$$

Therefore since $P\le N$, by Lemma 2, and (21),

$$\sum\nolimits_1 \le Z(Z-1+\pi(P))\log N. \tag{27}$$

We shall now estimate $\sum_2$. Notice that if p^k exceeds N then for each a in A there is at most one b in B for which $ab\equiv -1 \pmod{p^k}$. Therefore

$$\begin{aligned}\sum\nolimits_2 &\le \sum_{p\le P}\log p \sum_{\frac{\log N}{\log p}\le k\le\frac{\log(N^2+1)}{\log p}} |A|\\ &\le |A|\sum_{p\le P}\log p\frac{\log(N^2+1)}{\log p}\\ &\le 3Z\pi(P)\log N. \end{aligned}\tag{28}$$

Accordingly, by (26), (27), and (28),

$$\log E \le (Z^2+4Z\pi(P))\log N.$$

Thus, by (25),

$$(2-\delta)Z^2\log Z \le (Z^2+4Z\pi(P))\log N$$

hence

$$(2-\delta)Z\left(\frac{\log Z}{\log N}\right) \le Z+4\pi(P)$$

so

$$\frac{Z}{4}\left((2-\delta)\left(\frac{\log Z}{\log N}\right)-1\right)\le \pi(P).$$

By hypothesis $Z\ge N^\theta$ and so

$$\frac{Z}{4}((2-\delta)\theta-1)<\pi(P).$$

Since $\theta = 1/2+\delta$ we see that $(2-\delta)\theta-1=(3/2)\delta-\delta^2$ and $(3/2)\delta-\delta^2\ge (3/2)\delta-(1/2)\delta=\delta$. Therefore

$$\frac{\delta Z}{4}<\pi(P)$$

and our result now follows from the Prime Number Theorem.

8 Proof of Theorem 2

Let ε be a real number with $0 < \varepsilon < 1$ and let $N_0, N_1, \ldots$ denote positive numbers which are effectively computable in terms of ε. We shall suppose that (21) holds and that E and E_1 are defined as in (22) and (24) respectively. Then by (17) and (23) for $N > N_1$,

$$\log E > (2-\varepsilon) Z^2 \log N. \tag{29}$$

Further, by Lemma 3, for $N > N_2$,

$$\log E_1 < (1+\varepsilon) Z^2 \log N. \tag{30}$$

Put $E_2 = E/E_1$ and note that by (29) and (30)

$$\log E_2 > (1-2\varepsilon) Z^2 \log N. \tag{31}$$

Certainly

$$E_2 \leq \prod_{N \leq p \leq P} p^{\operatorname{ord}_p G} \tag{32}$$

where

$$G = \prod_{1 \leq m,\, n \leq N} (mn+1).$$

Put $P = NY$ and note that if p exceeds N then p^2 exceeds N^2+1 and so

$$\sum_{N < p \leq NY} \log p \operatorname{ord}_p G = \sum_{N < p \leq NY} \log p \sum_{\substack{1 \leq m,\, n \leq N \\ p \mid mn+1}} 1. \tag{33}$$

But, by Lemma 1,

$$\sum_{\substack{1 \leq m,\, n \leq N \\ p \mid mn+1}} 1 = \frac{p-1}{p^2} N^2 + O\left(p^{1/2} (\log p)^2 + \frac{N \log p}{p}\right). \tag{34}$$

Suppose that $P \leq (\varepsilon N / \log N)^{4/3}$ since otherwise our result holds. Thus by (34), for each prime p with $N < p \leq NY$ we have, for $N > N_3$,

$$\sum_{\substack{1 \leq m,\, n \leq N \\ p \mid mn+1}} 1 < (1+\varepsilon) \frac{N^2}{p}.$$

Therefore by (33), for $N > N_3$,

$$\sum_{N < p \leq NY} \log p \operatorname{ord}_p G < (1+\varepsilon) N^2 \sum_{N < p \leq NY} \frac{\log p}{p}.$$

By (17) and Theorem 1 we see that

$$Y > (\log N)^{1/2} \tag{35}$$

for $N > N_4$.

Since
$$\sum_{p \le x} \frac{\log p}{p} = \log x + O(1)$$
we have, by (35), that for $N > N_5$,
$$\sum_{N < p \le NY} \log p \operatorname{ord}_p G < (1 + 2\varepsilon) N^2 \log Y. \tag{36}$$
It follows from (31), (32) and (36) that
$$(1 - 2\varepsilon) Z^2 \log N < (1 + 2\varepsilon) N^2 \log Y$$
hence
$$N^{\left(\frac{1-2\varepsilon}{1+2\varepsilon}\right)\left(\frac{Z}{N}\right)^2} < Y$$
as required.

References

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

[2] Y. Bugeaud, On the greatest prime factor of $(ab+1)(bc+1)(ca+1)$, *Acta Arith.* **86** (1998), 45–49.

[3] Y. Bugeaud and F. Luca, A quantitative lower bound for the greatest prime factor of $(ab+1)(bc+1)(ca+1)$, *Acta Arith.* **114** (2004), 275–294.

[4] P. Corvaja and U. Zannier, On the greatest prime factor of $(ab+1)(ac+1)$, *Proc. Amer. Math. Soc.* **131** (2003), 1705–1709.

[5] P.D.T.A. Elliott and A. Sárközy, The distribution of the number of prime divisors of sums $a+b$, *J. Number Theory* **29** (1988), 94–99.

[6] P.D.T.A. Elliott and A. Sárközy, The distribution of the number of prime divisors of form $ab+1$, *New trends in probability and statistics,* Vol. 4 (Palanga, 1996), 313–321.

[7] P. Erdős, H. Maier and A. Sárközy, On the distribution of the number of prime factors of sums $a+b$, *Trans. Amer. Math. Soc.* **302** (1987), 269–280.

[8] P. Erdős, C. Pomerance, A. Sárközy and C.L. Stewart, On elements of sumsets with many prime factors, *J. Number Theory* **44** (1993), 93–104.

[9] P. Erdős, C.L. Stewart and R. Tijdeman, Some diophantine equations with many solutions, *Compositio Math.* **66** (1988), 37–56.

[10] P. Erdős and P. Turán, On a problem in the elementary theory of numbers, *Amer. Math. Monthly* **41** (1934), 608–611.

[11] J.-H. Evertse, On equations in S-units and the Thue-Mahler equation, *Invent. Math.* **75** (1984), 561–584.

[12] J.-H. Evertse, On sums of S-units and linear recurrences, *Compositio Math.* **53** (1984), 225–244.

[13] J-H. Evertse, The number of solutions of decomposable form equations, *Invent. Math.* **122** (1995), 559-601.

[14] J.-H. Evertse and K. Győry, Finiteness criteria for decomposable form equations, *Acta Arith.* **50** (1988), 357–379.

[15] K. Győry, C.L. Stewart and R. Tijdeman, On prime factors of sums of integers I, *Compositio Math.* **59** (1986), 81–88.

[16] 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.

[17] K. Győry and A. Sárközy, On prime factors of integers of the form $(ab+1)(bc+1)(ca+1)$, *Acta Arith.* **79** (1997), 163–171.

[18] S. Hernández and F. Luca, On the largest prime factor of $(ab+1)(ac+1)(bc+1)$, *Bol. Soc. Mat. Mexicana* **9** (2003), 235–244.

[19] J.H. Loxton, Some problems involving powers of integers, *Acta Arith.* **46** (1986), 113–123.

[20] F. Luca, On the greatest common divisor of $u-1$ and $v-1$ with u and v near S-units, *Monatshefte Math.* **146** (2005), 239–256.

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

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

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

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

[25] A. Sárközy and C.L. Stewart, On prime factors of integers of the form $ab+1$, *Publicationes Math. Debrecen* **56** (2000), 559–573.

[26] H.P. Schlickewei, The quantitative Subspace Theorem for number fields, *Compositio Math.* **82** (1992), 245–273.

[27] W.M. Schmidt, The subspace theorem in diophantine approximations, *Compositio Math.* **69** (1989), 121–173.

[28] C.L. Stewart, On the greatest prime factor of integers of the form $ab+1$, *Periodica Math. Hungarica* **43** (2001), 81–91.

[29] C.L. Stewart and R. Tijdeman, On prime factors of sums of integers II, *Diophantine Analysis, LMS Lecture Notes* **109**, Cambridge University Press (1986), 83–98.

[30] C.L. Stewart and R. Tijdeman, On the greatest prime factor of $(ab+1)(ac+1)(bc+1)$, *Acta Arith.* **79** (1997), 93–101.

[31] G. Tenenbaum, Facteurs premiers de sommes d'entiers, *Proc. Amer. Math. Soc.* **106** (1989), 287–296.

[32] M. Waldschmidt, Minorations de combinaisons linéaires de logarithmes de nombres algébriques, *Canad. J. Math.* **45** (1993), 176–224.

[33] S. Wu, Higher-dimensional Kloosterman sums and the greatest prime factor of integers of the form $a_1a_2\cdots a_{k+1}+1$, Ph.D. thesis, University of Waterloo, 2007.

Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario
Canada N2L 3G1
email: cstewart@uwaterloo.ca