PACIFIC JOURNAL OF MATHEMATICS
Vol. 166, No. 2, 1994

ON DIVISORS OF SUMS OF INTEGERS V

A. SÁRKÖZY AND C. L. STEWART

Dedicated to Professor P. Erdős on the occasion of his eightieth birthday.

Let N be a positive integer and let A and B be subsets of $\{1,\dots,N\}$. In this article we shall estimate both the maximum and the average of $\omega(a+b)$, the number of distinct prime factors of $a+b$, where a and b are from A and B respectively.

1. Introduction. For any set X let $|X|$ denote its cardinality and for any integer n larger than one let $\omega(n)$ denote the number of distinct prime factors of n. Let I be an integer larger than one and let ϵ be a positive real number. Let $2 = p_1, p_2, \dots$ be the sequence of prime numbers in increasing order and let m be that positive integer for which $p_1 \cdots p_m \le N \le p_1 \cdots p_{m+1}$. In [**3**], Erdős, Pomerance, Sárközy and Stewart proved that there exist positive numbers C_0 and C_1 which are effectively computable in terms of ϵ, such that if N exceeds C_0 and A and B are subsets of $\{1,\dots,N\}$ with $(|A||B|)^{1/2} > \epsilon N$ then there exist integers a from A and b from B for which

$$\omega(a+b) > m - C_1\sqrt{m}.$$

They also showed that there is a positive real number ϵ, with $\epsilon < 1$, and an effectively computable positive number C_2 such that for each positive integer N there is a subset A of $\{1,\dots,N\}$ with $|A| \ge \epsilon N$ for which

$$\max_{a,a'\in A} \omega(a+a') < m - \frac{C_2\sqrt{m}}{\log m}.$$

Notice by the prime number theorem that

$$m = (1+o(1))(\log N)/(\log\log N).$$

In this article we shall study both the maximum of $\omega(a+b)$ and the average of $\omega(a+b)$ as a and b run over A and B respectively where A and B are subsets of $\{1,\dots,N\}$ for which $(|A||B|)^{1/2}$ is much smaller than ϵN. Our principal tool will be the large sieve inequality.

THEOREM 1. *Let θ be a real number with $1/2 < \theta \leq 1$ and let N be a positive integer. There exists a positive number C_3, which is effectively computable in terms of θ, such thet if A and B are subsets of $\{1,\dots,N\}$ with N greater than C_3 and*

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

then there exists an integer a from A and an integer b from B for which

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

In [**6**] Pomerance, Sárközy and Stewart showed that if A and B are sufficiently dense sets then there is a sum $a+b$ which is divisible by a small prime factor. In particular they proved the following result. Let β be a positive real number. There is a positive number C_4, which is effectively computable in terms of β, such that if A and B are subsets of $\{1,\dots,N\}$ with $(|A||B|)^{1/2} > C_4 N^{1/2}$ then there is a prime number p with $\beta < p < C_4(N/(|A||B|)^{1/2})$, an integer a from A and an integer b from B such that p divides $a+b$. As a byproduct of our proof of Theorem 1 we are able to improve upon this result.

THEOREM 2. *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_5, which is effectively computable in terms of θ and β, such that if A and B are subsets of $\{1,\dots,N\}$ with*

$$(|A||B|)^{1/2} \geq N^{\theta}, \tag{3}$$

and N exceeds C_5 then there is a prime number p with

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

such that every residue class modulo p contains a member of $A+B$.

It follows from the work of Elliott and Sárközy [**1**], see also Erdős, Maier and Sárközy [**2**] and Tenenbaum [**7**], that if A and B are subsets of $\{1,\dots,N\}$ with

$$(|A||B|)^{1/2} = N/\exp(o((\log\log N)^{1/2}\log\log\log N)) \tag{4}$$

and N is sufficiently large then a theorem of Erdős-Kac type holds for $\omega(a+b)$. In particular for A and B satisfying (4) we have

$$\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B}\omega(a+b) \sim \log\log N. \tag{5}$$

Let δ be a positive real number. If A and B are subsets of $\{1,\dots,N\}$ with $|A| \sim |B| \sim N\exp(-\delta\log\log\log N)$, then (5) need not hold. For instance we may take A and B to be the subset of $\{1,\dots,N\}$ consisting of the multiples of $\prod_{p<\delta\log\log N\log\log\log N} p$. Then for N sufficiently large the average of $\omega(a+b)$ is at least $(1+\delta/2)\log\log N$. On the other hand we conjecture that if A and B are subsets of $\{1,\dots,N\}$ with

$$\min(|A|,|B|) > \exp((\log N)^{1+o(1)}), \tag{6}$$

ϵ is a positive real number and N is sufficiently large in terms of ϵ then

$$\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B}\omega(a+b) > (1-\epsilon)\log\log N. \tag{7}$$

On taking A and B to be positive integers up to $\exp((\log N)^{1-\epsilon})$ we see that condition (6) cannot be weakened substantially. Furthermore, we conjecture that if we let N tend to infinity and A and B run over subsets of $\{1,\dots,N\}$ with

$$\frac{\log(\min(|A|,|B|))}{\log\log N} \to \infty$$

then

$$\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B}\omega(a+b) \to \infty.$$

While we have not been able to establish (7) for all subsets A and B satisfying (6), we have been able to determine the average order for the number of large prime divisors of the sums $a+b$ for sufficiently dense sets A and B. As a consequence we are able to establish (7) for such sets.

THEOREM 3. *There exists an effectively computable positive constant C_6 such that if T and N are positive integers with $T \leq \sqrt{2N}$ and A and B are non-empty subsets of $\{1,\dots,N\}$ then*

$$\left|\frac{1}{|A||B|}\sum_{T<p}\sum_{a\in A, b\in B, p|(a+b)} 1 - (\log\log N - \log\log(3T))\right| < C_6 + \frac{3N}{(|A||B|)^{1/2}T}.$$

We now take $T = N/(|A||B|)^{1/2}$ in Theorem 3 to obtain the following result.

COROLLARY 1. *There exists an effectively computable positive constant C_7 such that if N is a positive integer and A and B are subsets of $\{1,\dots,N\}$ with $|A||B| > N$ then*

$$\left|\frac{1}{|A||B|}\sum_{p>N(|A||B|)^{-1/2}}\ \sum_{a\in A, b\in B, p|(a+b)} 1 - (\log\log N - \log\log N(|A||B|)^{1/2})\right| < C_7.$$

Therefore (7) holds for N sufficiently large provided that A and B are subsets of $\{1,\dots,N\}$ with

$$(|A||B|)^{1/2} = N\exp((\log N)^{o(1)}).$$

2. Preliminary Lemmas. For any real number x let $e(x) = e^{2\pi i}$ and let $||x||$ denote the distance from x to the nearest integer.

Let M and N be integers with N positive and let $a_{M+1}, \ldots, a_{M+N}$ be complex numbers. Define $S(x)$ by

$$S(x) = \sum_{M+1}^{M+N} a_n e(nx). \tag{8}$$

Let X be a set of real numbers which are distinct modulo 1 and define δ by

$$\delta = \min_{x,x' \in X, x \neq x'} \|x - x'\|. \tag{9}$$

The analytical form of the large sieve inequality, (see Theorem 1 of [**5**]), is required for the proof of Theorem 3 and it is given below.

LEMMA 1. *Let $S(x)$ and δ be as in* (8) *and* (9)*, respectively. Then*

$$\sum_{x \in X} |S(x)|^2 \leq (N + \delta^{-1}) \sum_{n=M+1}^{M+N} |a_n|^2.$$

We shall also make use of the following result, see Theorem 1 of [**6**], which was deduced with the aid of the arithmetical form of the large sieve inequality.

LEMMA 2. *Let N be a positive integer and let A and B be non-empty subsets of $\{1, \ldots, N\}$. Let S be a set of prime numbers, let Q be a positive integer and let J denote the number of square-free positive integers up to Q all of whose prime factors are from S. If*

$$J(|A||B|)^{1/2} > N + Q^2, \tag{10}$$

then there is a prime p in S such that each residue class modulo p contains a member of the sum set $A + B$.

Finally, to prove Theorems 1 and 2 we shall require the next result.

LEMMA 3. *Let α and β be real numbers with $\alpha > 1$ and let N be a positive integer. Let T be the set of prime numbers p which satisfy $\beta < p \leq (\log N)^\alpha$ and let S be a subset of T consisting of all but* *at most* $2\log N$ *elements of* T. *Let* R *denote the set of square-free positive integers less than or equal to* N *all of whose prime factors are from* S. *There exists a real number* C_8, *which is effectively computable in terms of* α *and* β, *such that*

$$|R| > 20N^{1-1/\alpha},$$

whenever N *is greater then* C_8.

Proof. C_9, C_{10} and C_{11} will denote positive numbers which are effectively computable in terms of α and β. By the prime number theorem with error term,

$$|S| \geq \pi((\log N)^\alpha) - \pi(\beta) - 2\log N > \frac{(\log N)^\alpha}{\alpha \log\log N}, \tag{11}$$

provided that N is greater than C_9. For any real number x let $[x]$ denote the greatest integer less than or equal to x. We now count the number of distinct ways of choosing $[\log N/(\alpha \log\log N)]$ primes from S. Each choice gives rise to a distinct square-free integer, given by the product of the primes, which does not exceed N and is composed only of primes from S. Then $|R| \geq \omega$ where

$$\omega = \binom{|S|}{\left[\frac{\log N}{\alpha \log\log N}\right]}.$$

Thus

$$\omega \geq \frac{\left(|S| - \left[\frac{\log N}{\alpha\log\log N}\right]\right)^{\frac{\log N}{\alpha \log\log N} - 1}}{\left[\frac{\log N}{\alpha \log\log N}\right]!},$$

and so, by (11) and Stirling's formula,

$$\omega \geq \frac{\left(\frac{(\log N)^\alpha}{\alpha\log\log N}\left(1 - \frac{1}{(\log N)^{\alpha-1}}\right)\right)^{\frac{\log N}{\alpha\log\log N}}}{(\log N)^{\alpha+1}\left(\frac{\log N}{e\alpha \log\log N}\right)^{\frac{\log N}{\alpha\log\log N}}},$$

for $N > C_{10}$. Since $\log(1-x) > -2x$ for $0 < x < 1/2$, we find that, for $N > C_{11}$,

$$\omega \geq N^{1-1/\alpha} e^{\left(\frac{\log N}{\alpha \log\log N} - \frac{2(\log N)^{2-\alpha}}{\alpha \log\log N}\right)} (\log N)^{-\alpha-1},$$

hence

$$\omega > 20N^{1-1/\alpha},$$

as required. □

3. Proof of Theorem 1. Let $\theta_1 = (\theta + 1/2)/2$ and define G and v by

$$G = (\log N)^{1/(2\theta_1 - 1)},$$

and

$$v = \left[\frac{1}{6}\left(\theta - \frac{1}{2}\right)^2 \frac{\log N}{\log\log N}\right] + 1, \tag{12}$$

respectively.

Put $A_0 = A, B_0 = B$ and $W_0 = \emptyset$. We shall construct inductively sets $A_1, \ldots, A_v$, $B_1, \ldots, B_v$ and $W_1, \ldots, W_v$ with the following properties. First, W_i is a set of i primes q satisfying $10 < q \leq G$, $A_i \subseteq A_{i-1}$ and $B_i \subseteq B_{i-1}$ for $i = 1, \ldots, v$. Secondly every element of the sum set $A_i + B_i$ is divisible by each prime in W_i for $i = 1, \ldots, v$. Finally,

$$|A_i| \geq \frac{|A|}{G^{3i}} \quad \text{and} \quad |B_i| \geq \frac{|B|}{G^{3i}}, \tag{13}$$

for $i = 1, \ldots, v$. Note that this suffices to prove our result since A_v and B_v are both non-empty and on taking a from A_v and b from B_v we find that $a + b$ is divisible by the v primes from W_v and so (2) follows from (12).

Suppose that i is an integer with $0 \leq i < v$ and that A_i, B_i and W_i have been constructed with the above properties. We shall now show how to construct A_{i+1}, B_{i+1} and W_{i+1}. First, for each prime p with $10 < p \leq G$ let $a_1, \ldots, a_{j(p)}$ be representatives for those residue classes modulo p which are occupied by fewer than $|A_i|/p^3$ terms of A_i. For each prime p with $10 < p \leq G$ we remove from A_i those terms of A_i which are congruent to one of $a_1, \dots, a_{j(p)}$ modulo p. We are left with a subset A_i' of A_i with

$$(14)\ |A_i'| \geq |A_i| \left(1 - \sum_{10<p\leq G} \frac{j(p)}{p^3}\right) \geq |A_i| \left(1 - \sum_{10<p} \frac{1}{p^2}\right) \geq \frac{|A_i|}{10}$$

and such that for each prime p with $10 < p \leq G$ and each a' in A_i', the number of terms of A_i which are congruent to a' modulo p is at least $|A_i|/p^3$. Similarly, we produce a subset B_i' of B_i with

$$|B_i'| \geq \frac{|B_i|}{10} \tag{15}$$

and such that for each prime p with $10 < p \leq G$ and each residue class modulo p which contains an element of B_i' the number of terms of B_i in the residue class is at least $|B_i|/p^3$.

The number of terms in W_i is i which is less than v and, by (12), is at most $\log N$. Thus we may apply Lemma 3 with $\beta = 10$ and $\alpha = 1/(2\theta_1 - 1)$ to conclude that there is a real number C_{12}, which is effectively computable in terms of θ, such that if N exceeds C_{12} then the number of square-free positive integers less than or equal to $N^{1/2}$ all of whose prime factors p satisfy $10 < p \leq G$ and $p \notin W_i$ is greater than

$$20\, N^{\frac{1}{2}(1-(2\theta_1-1))} = 20\, N^{1-\theta_1}. \tag{16}$$

By our inductive assumption (13) and by (1) and (12), we obtain

$$(|A_i||B_i|)^{1/2} \geq (|A||B|)^{1/2} G^{-3i} \geq N^{\theta_1}. \tag{17}$$

Thus, by (14), (15) and (17),

$$(|A_i'||B_i'|)^{1/2} \geq \frac{N^{\theta_1}}{10}. \tag{18}$$

We now apply Lemma 2 with $A = A_i'$, $B = B_i'$, $Q = N^{1/2}$ and S the set of primes p with $10 < p \leq G$ and $p \notin W_i$. Then J, the number of square-free integers up to Q divisible only by primes from S, is greater than $20N^{1-\theta_1}$ by (16), for $N > C_{12}$ and so, by (18), inequality (10) holds. Thus there is a prime q_{i+1} in S, an element

a' in A_i' and an element b' in B_i' such that q_{i+1} divides $a' + b'$. We put

$$A_{i+1} = \{a \in A_i \ : \ a \equiv a' \pmod{q_{i+1}}\},$$
$$B_{i+1} = \{b \in B_i \ : \ b \equiv b' \pmod{q_{i+1}}\},$$

and

$$W_{i+1} = W_i \cup \{q_{i+1}\}.$$

By our construction every element of $A_{i+1} + B_{i+1}$ is divisible by each prime in W_{i+1}. Further, we have, by (13),

$$|A_{i+1}| \geq \frac{|A_i|}{q_{i+1}^3} \geq \frac{|A_i|}{G^3} \geq \frac{|A|}{G^{3(i+1)}},$$

and

$$|B_{i+1}| \geq \frac{|B|}{G^{3(i+1)}},$$

as required. Our result now follows.

4. Proof of Theorem 2. Let S be the set of primes p which satisfy $\beta < p \leq (\log(N^{1/2}))^{1/(2\theta-1)}$. Put $\alpha = 1/(2\theta - 1)$ and observe that α is a real number greater than one since $1/2 < \theta < 1$. Next let J denote the number of square-free positive integer less than or equal to $N^{1/2}$ all of whose prime factors are from S. By Lemma 3 there exists a positive number C_{13}, which is effectively computable in terms of θ, such that if N exceeds C_{13}, then

$$J > 20(N^{1/2})^{1-(2\theta-1)} = 20N^{1-\theta}. \tag{19}$$

We now apply Lemma 2 with $Q = N^{1/2}$ and with J and S as above. From (3) and (19) we obtain (10) and so our result follows from Lemma 2.

5. Proof of Theorem 3. Put $R = [\sqrt{2N}]$. We have

$$\left|\sum_{a\in A}\sum_{b\in B}\sum_{T<p,p|a+b} 1 - \sum_{a\in A}\sum_{b\in B}\sum_{T<p\leq R,p|a+b} 1\right|$$
$$= \left|\sum_{a\in A}\sum_{b\in B}\sum_{R<p\leq 2N,p|a+b} 1\right| \leq \left|\sum_{a\in A}\sum_{b\in B} 1\right| = |A||B|.$$

We define, for each real number α,

$$F(\alpha) = \sum_{a\in A} e(a\alpha) \quad \text{and} \quad G(\alpha) = \sum_{b\in B} e(b\alpha).$$

Then

$$\sum_{a\in A}\sum_{b\in B}\sum_{T<p\le R, p|a+b} 1 = \sum_{T<p\le R} \frac{1}{p}\sum_{h=0}^{p-1} F\left(\frac{h}{p}\right) G\left(\frac{h}{p}\right) \tag{21}$$

$$= \sum_{T<p\le R} \frac{1}{p}\left(|A||B| + \sum_{h=0}^{p-1} F\left(\frac{h}{p}\right) G\left(\frac{h}{p}\right)\right).$$

Further there is an effectively computable positive constant C_{14} such that

$$\left|\sum_{T<p\le R} \frac{1}{p} - (\log\log R - \log\log(3T))\right| < C_{14}, \tag{22}$$

see Theorem 427 of [**4**]. Put

$$H = \left|\sum_{a\in A}\sum_{b\in B}\sum_{T<p, p|a+b} 1 - |A||B|(\log\log N - \log\log(3T))\right|.$$

By (20), (21) and (22),

$$H \le C_{15}|A||B| + \sum_{T<p\le R} \frac{1}{p}\sum_{h=1}^{p-1}\left|F\left(\frac{h}{p}\right) G\left(\frac{h}{p}\right)\right|.$$

For all real numbers u and v, $|u||v| \le (|u|^2 + |v|^2)/2$ and thus

$$H \le C_{15}|A||B| + \frac{1}{2}\sum_{T<p\le R} \frac{1}{p}\sum_{h=1}^{p-1}\left(\left(\frac{|B|}{|A|}\right)^{1/2}\left|F\left(\frac{h}{p}\right)\right|^2 + \left(\frac{|A|}{|B|}\right)^{1/2}\left|G\left(\frac{h}{p}\right)\right|^2\right). \tag{23}$$

Put

$$S(n) = \sum_{p<n}\sum_{h=1}^{p-1}\left|F\left(\frac{h}{p}\right)\right|^2.$$

Then by Lemma 1, for $n \le R$,

$$S(n) \le (N + n^2)|A| \le 3N|A|.$$

Thus we obtain

(24)
$$\begin{aligned}\sum_{T<p\le R} \frac{1}{p} \sum_{h=1}^{p-1} \left| F\left(\frac{h}{p}\right)\right|^2 &= \sum_{n=T+1}^{R} \frac{S(n) - S(n-1)}{n} \\ &= \sum_{n=T+1}^{R} S(n)\left(\frac{1}{n} - \frac{1}{n+1}\right) - \frac{S(T)}{T+1} + \frac{S(R)}{R+1} \\ &= \sum_{n=T+1}^{R} 3N|A|\left(\frac{1}{n} - \frac{1}{n+1}\right) + \frac{3N|A|}{R+1} = \frac{3N|A|}{T+1},\end{aligned}$$

and similarly

$$\sum_{T<p\le R} \frac{1}{p} \sum_{h=1}^{p-1} \left| G\left(\frac{h}{p}\right)\right|^2 \le \frac{3N|B|}{T+1}. \tag{25}$$

Our result follows from (23), (24) and (25).

References

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

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

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

[4] G.H. Hardy, E.M. Wright, *An introduction to the theory of numbers*, 5-th ed., Oxford, 1979.

[5] H.L. Montgomery, R.C. Vaughan, *The large seive*, Mathematika, **20** (1973), 119-134.

[6] C. Pomerance, A. Sárközy, C.L. Stewart, *On divisors of sums of integers*, III, Pacific J. Math., **133** (1988), 363-379.

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

Received September 9, 1991, and accepted for publication July 29, 1993.
The research of the second author was supported in part by a Killam Research Fellowship and by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

THE UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO, CANADA N2L 3G1
E-mail address: cstewart@watserv1.uwaterloo.ca

Permanent address of A. Sárközy:
MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13-15,
BUDAPEST, HUNGARY, H-1053