Acta Math. Hungar.
99 (1–2) (2003), 1–24.

ON SUMS WHICH ARE POWERS

K. GYARMATI, A. SÁRKÖZY* (Budapest) and C. L. STEWART† (Waterloo)

Abstract. The cardinality of sets $\mathcal{A}$ is estimated under the conditions that every element of the sum set $\mathcal{A}+\mathcal{A}$ is a power resp. powerful number (n is said to be powerful if $p \mid n$ implies $p^2 \mid n$). Subset sums with these properties are also studied.

1. Introduction

Let $\mathbf{N}$ denote the set of positive integers and let $\mathbf{Q}$ denote the set of rational numbers. For any subset $\mathcal{A}$ of $\mathbf{N}$ denote the set of subset sums of $\mathcal{A}$ by $\mathcal{P}(\mathcal{A})$. Thus

$$\mathcal{P}(\mathcal{A}) = \left\{ \sum_{a\in\mathcal{A}} \varepsilon_a a : \varepsilon_a = 0 \text{ or } 1, 0 < \sum_{a\in\mathcal{A}} \varepsilon_a < \infty \right\}.$$

Denote the set of the powers by V, so that

$$V = \{x^k : x \in \mathbf{N},\ k \in \mathbf{N},\ k \geqq 2\},$$

and for each integer K, larger than one, put

$$V_K = \{x^k : x \in \mathbf{N},\ k \in \mathbf{N},\ 2 \leqq k \leqq K\}.$$

Let W be the set of the powerful numbers, so that

$$W = \{n : n \in \mathbf{N},\ p \mid n \text{ implies } p^2 \mid n\}.$$

*Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. T029759.

†Research supported in part by the Natural Sciences and Engineering Research Council of Canada, Grant A3528.

Key words and phrases: Gallagher sieve, Linnik's constant, power, powerful, sequences.

2000 Mathematics Subject Classification: 11Bxx, 11N36.

Diophantus initiated the study of sequences $a_1 < a_2 < \cdots$ with the property that $a_i a_j + 1$ is a square for all $1 \leqq i < j$. Later this problem was also studied by Fermat, Euler, Straus and others, and recently Dujella [2] proved that if $\mathcal{A}$ is a set of positive integers with this property, then $\mathcal{A}$ must be finite and, indeed, $|\mathcal{A}| \leqq 9$. Erdős and Moser [3] investigated the additive analogue of the problem. They studied sequences $a_1 < a_2 < \cdots$ with the property that $a_i + a_j$ is a square for $i \neq j$. Rivat, Sárközy and Stewart [12] proved that if $\mathcal{A} = \{a_1, a_2, \dots\}$ is a set with this property and $\mathcal{A} \subset \{1, 2, \dots, N\}$ then, for N large enough,

$$|\mathcal{A}| < 37 \log N. \tag{1.1}$$

Gyarmati [7] has examined generalizations of the problems of Diophantus and Erdős and Moser.

In this paper we shall study sequences $a_1 < a_2 < \cdots$ with the property that $a_i + a_j$ is always a power and sequences for which $a_i + a_j$ is always powerful, see [14] where the latter problem was first proposed. In [8] we treated the case where $a_i a_j + 1$ is always a power. We proved the following. Let N be a positive integer and $\mathcal{A}$ a subset of $\{1, \dots, N\}$. Let K be an integer with $K \geqq 2$. If $aa' + 1$ is in V_K whenever a and a' are distinct integers from $\mathcal{A}$, then

$$|\mathcal{A}| < 160 \frac{K^2}{(\log K)^2} \log\log N, \tag{1.2}$$

for N sufficiently large. Further if $aa' + 1$ is in V whenever a and a' are distinct integers from $\mathcal{A}$, then

$$|A| < 340 (\log N)^2 / \log\log N, \tag{1.3}$$

for N sufficiently large.

We shall establish first the additive analogues of (1.2) and (1.3).

THEOREM 1. *There is an absolute constant c_1 such that if N and K are integers larger than 1, $\mathcal{A}$ is a subset of $\{1, 2, \dots, N\}$ and*

$$a + a' \in V_K \quad \textit{for all} \quad a \in \mathcal{A},\ a' \in \mathcal{A},\ a \neq a', \tag{1.4}$$

then

$$|\mathcal{A}| < c_1 e^{46K} \log N. \tag{1.5}$$

Note that apart from the value of the constant factor on the right hand side of (1.1), this includes (1.1) as a special case.

Replacing V_K by V, we can prove:

THEOREM 2. *There is an absolute constant c_2 such that if N is a positive integer larger than 1, $\mathcal{A}$ is a subset of $\{1, 2, \dots, N\}$ and*

$$a + a' \in V \quad \text{for all} \quad a \in \mathcal{A},\ a' \in \mathcal{A},\ a \neq a', \tag{1.6}$$

then

$$|\mathcal{A}| < c_2 (\log N)^{48}. \tag{1.7}$$

Note that since $V_K \subseteqq V$ for all K, thus (1.7) also holds for sets $\mathcal{A}$ satisfying (1.4). This new upper bound (1.7) is superior to (1.5) in Theorem 1 for $K \gg \log\log N$.

We will also give a lower bound:

THEOREM 3. *There is a number N_0 such that if N is a positive integer larger than N_0, then there is a subset $\mathcal{A}$ of $\{1, 2, \dots, N\}$ with*

$$|\mathcal{A}| \geqq \left[\frac{\log\log N}{4 \log\log\log N} \right]$$

which satisfies (1.6).

In case of infinite sets $\mathcal{A} \subseteqq \mathbf{N}$ the situation is different:

THEOREM 4. *There is no infinite set $\mathcal{A} \subseteqq \mathbf{N}$ satisfying* (1.6).

Considering now the powerful numbers, we will prove the following lower bound.

THEOREM 5. *There is a number N_1 such that if N is a positive integer larger than N_1, then there is a subset $\mathcal{A}$ of $\{1, \dots, N\}$ with*

$$|\mathcal{A}| > \frac{1}{5} \log N \tag{1.8}$$

and for which

$$a + a' \in W \quad \text{for all} \quad a \in \mathcal{A},\ a' \in \mathcal{A}. \tag{1.9}$$

On the other hand we have not been able to give any reasonable upper bound for the cardinality of sets $\mathcal{A} \subset \{1, 2, \dots, N\}$ satisfying (1.9). Thus here we will study the easier problem where the sums are replaced by subset sums. Let $F(N)$ denote the largest cardinality of a subset $\mathcal{A}$ of $\{1, 2, \dots, N\}$ for which $P(\mathcal{A}) \subset W$.

THEOREM 6. *Let N be an integer larger than 3. There are positive absolute constants c_3, c_4 so that*

$$c_3 (\log N)^{1/2} < F(N) < c_4 (\log N)^3 (\log\log N)^{-1/2}. \tag{1.10}$$

2. Lemmas

In order to prove Theorem 1 we need several lemmas.

LEMMA 1. *If p is a prime, χ is a multiplicative character* mod p *which is not the principal character, and $\mathcal{A}, \mathcal{B} \subset \{1, 2, \dots, p\}$, then we have*

$$\left| \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \chi(a+b) \right| \leqq \left(p|\mathcal{A}|\,|\mathcal{B}| \right)^{1/2}.$$

PROOF. This is a result of Erdős and Shapiro [4].

In Lemma 2 and its proof, $q, q_1, q_2, \dots$ will denote primes.

LEMMA 2. *Let K be an integer with $K \geqq 2$ and put $Q_K = \prod_{q \leqq K} q$. Assume that p is a prime with*

$$p \equiv 1 \pmod{Q_K}. \tag{2.1}$$

Let $\mathcal{B}$ be a set of integers such that whenever b and b' are distinct elements of $\mathcal{B}$, then $b \not\equiv b' \pmod p$ and there exists an integer k with $2 \leqq k \leqq K$ and an integer x for which

$$b + b' \equiv x^k \pmod p. \tag{2.2}$$

Then there is a positive absolute constant c_5 such that

$$|\mathcal{B}| < c_5 (\log K) 2^{\pi(K)} p^{1/2}. \tag{2.3}$$

PROOF. If (2.2) is solvable, then

$$b + b' \equiv y^q \pmod p \tag{2.4}$$

is also solvable with some prime $q \mid k$ where $q \leqq k \leqq K$. For a prime q and $n \in \mathbf{Z}$, define $f_q(n)$ by $f_q(n) = 1$ if

$$x^q \equiv n \pmod p \tag{2.5}$$

is solvable and $f_q(n) = 0$ if (2.5) is not solvable. Then for $n \in \mathbf{Z}$ we have

$$\prod_{2 \leqq q \leqq K} \left(1 - f_q(n)\right) = \begin{cases} 0 & \text{if (2.5) is solvable for some } q \text{ with } 2 \leqq q \leqq K, \\ 1 & \text{otherwise.} \end{cases}$$

Thus writing

$$F_K(n) = 1 - \prod_{2 \leqq q \leqq K} \big(1 - f_q(n)\big), \tag{2.6}$$

we have

$$F_K(n) = \begin{cases} 1 & \text{if (2.5) is solvable for some } q \text{ with } 2 \leqq q \leqq K, \\ 0 & \text{otherwise.} \end{cases}$$

By our assumption on $\mathcal{B}$, it follows that

$$\sum_{b \in \mathcal{B}} \sum_{b' \in \mathcal{B}} F_K(b + b') \geqq \sum_{\substack{b, b' \in \mathcal{B} \\ b \neq b'}} 1 = |\mathcal{B}|^2 - |\mathcal{B}|. \tag{2.7}$$

If g is a primitive root modulo p and q is a prime with $q \leqq K$, then let $\chi_q(n)$ denote the uniquely determined character modulo p defined by

$$\chi_q(g) = e\left(\frac{1}{q}\right);$$

note that by $q \leqq K$ and (2.1) we have $q \mid p - 1$ thus there is such a character modulo p. This is a character of order q, thus for $(n, p) = 1$ clearly we have

$$f_q(n) = \frac{1}{q} \sum_{j=0}^{q-1} \chi_q^j(n) \qquad \big(\text{for } (n, p) = 1\big). \tag{2.8}$$

It follows from (2.6) and (2.8) that

$$F_K(n) = \sum_{\ell=1}^{\pi(K)} (-1)^{\ell+1} \sum_{q_1 < \cdots < q_\ell \leqq K} \prod_{i=1}^{\ell} f_{q_i}(n) \tag{2.9}$$

$$= \sum_{\ell=1}^{\pi(K)} (-1)^{\ell+1} \sum_{q_1 < \cdots < q_\ell \leqq K} \frac{1}{q_1 \cdots q_\ell} \sum_{j_1=0}^{q_1-1} \cdots \sum_{j_\ell=0}^{q_\ell-1} \chi_{q_1}^{j_1} \cdots \chi_{q_\ell}^{j_\ell}(n)$$

for $(n, p) = 1$. Thus denoting this last expression (for all n) by $F^*(n)$, we have

$$F_K(n) = F_K^*(n) \quad \text{for} \quad (n, p) = 1$$

and clearly

$$F_K(n) = 1 \quad \text{and} \quad F_K^*(n) = 0 \quad \text{for} \quad p \mid n.$$

It follows that

$$\sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} F_K(b+b') = \sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} F_K^*(b+b') + \sum_{\substack{b,b'\in\mathcal{B}\\ p\mid b+b'}} 1. \tag{2.10}$$

For all $b \in \mathcal{B}$ there is at most one b' with $p \mid b+b'$, thus it follows that

$$\sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} F_K(b+b') \leqq \sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} F_K^*(b+b') + |\mathcal{B}|.$$

Now consider the characters $\psi = \chi_{q_1}^{j_1} \cdots \chi_{q_\ell}^{j_\ell}$ in (2.9). We have

$$\psi(g) = \big(\chi_{q_1}(g)\big)^{j_1} \cdots \big(\chi_{q_\ell}(g)\big)^{j_\ell} = e\left(\frac{j_1}{q_1} + \cdots + \frac{j_\ell}{q_\ell}\right),$$

and for $0 \leqq j_1 < q_1, \ldots, 0 \leqq j_\ell < q_\ell$, clearly we have $\frac{j_1}{q_1} + \cdots + \frac{j_\ell}{q_\ell} \in \mathbf{Z}$, hence ψ is the principal character modulo p, if and only if $j_1 = \cdots = j_\ell = 0$. Thus separating out the contribution of the principal character modulo p we obtain

$$\sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} F_K^*(b+b') = \sum_{\ell=1}^{\pi(K)} (-1)^{\ell+1} \sum_{q_1<\cdots<q_\ell\leqq K} \frac{1}{q_1\cdots q_\ell} \sum_{\substack{b,b'\in\mathcal{B}\\ (b+b',p)=1}} 1 \tag{2.11}$$

$$+ \sum_{\ell=1}^{\pi(K)} (-1)^{\ell+1} \sum_{q_1<\cdots<q_\ell\leqq K} \frac{1}{q_1\cdots q_\ell} \sum_{\substack{0\leqq j_1\leqq q_1,\ldots,0\leqq j_\ell\leqq q_\ell\\ (j_1,\ldots,j_\ell)\neq(0,\ldots,0)}} \sum_{b\in\mathcal{B}} \sum_{b'\in\mathcal{B}} \chi_{q_1}^{j_1}\cdots\chi_{q_\ell}^{j_\ell}(b+b')$$

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

Here we have

$$\sum\nolimits_1 = \left(1 - \prod_{q\leqq K}\left(1-\frac{1}{q}\right)\right) \sum_{\substack{b,b'\in\mathcal{B}\\ (b+b',p)=1}} 1 \leqq \left(1 - \prod_{q\leqq K}\left(1-\frac{1}{q}\right)\right)|\mathcal{B}|^2. \tag{2.12}$$

Moreover, in $\sum_2$ none of the characters $\chi_{q_1}^{j_1}\cdots\chi_{q_\ell}^{j_\ell}$ is the principal character. Thus, by Lemma 1, we have

$$\left|\sum\nolimits_2\right| \tag{2.13}$$

$$\leqq \sum_{\ell=1}^{\pi(K)} \sum_{q_1<\cdots<q_\ell\leqq K} \frac{1}{q_1\cdots q_\ell} \sum_{\substack{0\leqq j_1\leqq q_1,\ldots,0\leqq j_\ell\leqq q_\ell\\(j_1,\ldots,j_\ell)\neq(0,\ldots,0)}} \left|\sum_{b\in\mathcal{B}}\sum_{b'\in\mathcal{B}} \chi_{q_1}^{j_1}\chi_{q_2}^{j_2}\cdots\chi_{q_\ell}^{j_\ell}(b+b')\right|$$

$$\leqq \sum_{\ell=1}^{\pi(K)} \sum_{q_1<\cdots<q_\ell\leqq K} \frac{1}{q_1\cdots q_\ell} \sum_{\substack{0\leqq j_1\leqq q_1,\ldots,0\leqq j_\ell\leqq q_\ell\\(j_1,\ldots,j_\ell)\neq(0,\ldots,0)}} p^{1/2}|\mathcal{B}|$$

$$\leqq p^{1/2}|\mathcal{B}| \sum_{\ell=1}^{\pi(K)} \sum_{q_1<\cdots<q_\ell\leqq K} 1 < p^{1/2}|\mathcal{B}|2^{\pi(K)}.$$

By (2.7), (2.10), (2.11), (2.12) and (2.13) we have

$$|\mathcal{B}|^2 - |\mathcal{B}| \leqq \sum_{b\in\mathcal{B}}\sum_{b'\in\mathcal{B}} F_K(b+b') \leqq \sum\nolimits_1 + \sum\nolimits_2 + |\mathcal{B}|$$

$$< \left(1 - \prod_{q\leqq K}\left(1-\frac{1}{q}\right)\right)|\mathcal{B}|^2 + p^{1/2}|\mathcal{B}|2^{\pi(K)} + |\mathcal{B}|$$

whence

$$\prod_{q\leqq K}\left(1-\frac{1}{q}\right)|\mathcal{B}| \leqq p^{1/2}2^{\pi(K)} + 2.$$

By using Merten's Theorem, (2.3) follows.

LEMMA 3. *Let $\mathcal{A}$ 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 modulo p that contain an element of $\mathcal{A}$. Then for any finite set of primes $\mathcal{P}$ we have*

$$|\mathcal{A}| \leqq \frac{\sum_{p\in\mathcal{P}} \log p - \log N}{\sum_{p\in\mathcal{P}} \frac{\log p}{\nu(p)} - \log N} \tag{2.14}$$

provided that the denominator is positive.

PROOF. This is Gallagher's "larger sieve" [5].

We will also need the fact that in an arithmetic progression of modulus q there are many relatively small primes (below q^c). This fact is closely related to Linnik's theorem on the least prime in an arithmetic progression and, correspondingly, it can be derived from a result whose variants occur in several papers dealing with Linnik's constant. To formulate this result, we need a formula of Turán. Let $k \geqq 2$, and write

$$K(w) = \frac{e^{kw}(e^w - e^{-w})}{2w}, \quad K_1(w) = K(2w \log q)$$

and

$$R(n) = \frac{1}{2\pi i} \int\limits_{\operatorname{Re} w = 2} K_1^2(w) n^{-w}\, dw.$$

Then (see [10]) for all $q \in \mathbf{N}$, $q \geqq 2$ we have

$$R(n) \begin{cases} = 0 & \text{if } \ 1 \leqq n \leqq q^{4k-4} \text{ or } n \geqq q^{4k+4}, \\ < \dfrac{c_6}{\log q} & \text{if } \ q^{4k-4} < n \leqq q^{4k+4}, \end{cases} \tag{2.15}$$

and for $a \in \mathbf{Z}$, $(a, q) = 1$,

$$\sum_{\substack{q^{4k-4} < n < q^{4k+4} \\ n \equiv a \pmod q}} \Lambda(n) R(n) n^{-1} \tag{2.16}$$

$$= \frac{1}{\varphi(q)} \left(1 - \sum_{\chi} \overline{\chi}(a) \sum_{\varrho_\chi} K_1^2(\varrho_\chi - 1) \right) + O(q^{-2})$$

where for each character $\chi \pmod q$, ϱ_χ runs over the non-trivial zeros of $L(s, \chi)$.

LEMMA 4. *There is a positive absolute constant c_7 such that for $k \geqq 19$,*

$$\left| \sum_{\chi} \overline{\chi}(a) \sum_{\varrho_\chi} K_1^2(\varrho_\chi - 1) \right| \leqq \sum_{\chi} \sum_{\varrho_\chi} \left| K_1^2(\varrho_\chi - 1) \right| < 1 - c_7.$$

PROOF. This is proved by Jutila in [10, pp. 59–61] (see also formula (23) in [6]).

LEMMA 5. *There are absolute constants q_0, c_8 such that if $q \geqq q_0$, $a \in \mathbf{Z}$, $(a,q)=1$ and*

$$x \geqq q^{80}, \tag{2.17}$$

then

$$\sum_{\substack{x/q^8<p<x \\ p\equiv a \pmod q}} \frac{\log p}{p} > c_8 \frac{\log q}{\varphi(q)}. \tag{2.18}$$

Note that the exponent in (2.17) could be improved upon by using more recent works [1, 6, 9] on Linnik's constant. However, from our point of view this improvement is not significant.

PROOF. Define k to be the largest integer for which $q^{4k+4} \leqq x$ so that $k \geqq 19$ by (2.17). Thus by Lemma 4, for $q > q_1$ the absolute value of the right hand side of (2.16) is

$$\left| \frac{1}{\varphi(q)} \left(1 - \sum_{\chi} \overline{\chi}(a) \sum_{\varrho_\chi} K_1^2(\varrho_\chi - 1) \right) + O(q^{-2}) \right| \tag{2.19}$$

$$\geqq \frac{1}{\varphi(q)} \left(1 - \left| \sum_{\chi} \sum_{\varrho_\chi} K_1^2(\varrho_\chi - 1) \right| \right) - \frac{c_9}{q^2} > \frac{1}{\varphi(q)} c_7 - \frac{c_9}{q^2} > \frac{c_{10}}{\varphi(q)}.$$

On the other hand, by (2.15) and (2.17) for $q > q_2$ the absolute value of the left hand side of (2.16) is

$$\left| \sum_{\substack{q^{4k-4}<n<q^{4k+4} \\ n\equiv a \pmod q}} \Lambda(n) R(n) n^{-1} \right| < \frac{c_6}{\log q} \sum_{\substack{\frac{x}{q^8}<n<x \\ n\equiv a \pmod q}} \frac{\Lambda(n)}{n} \tag{2.20}$$

$$< \frac{c_6}{\log q} \left(\sum_{\substack{\frac{x}{q^8}<p<x \\ p\equiv a \pmod q}} \frac{\log p}{p} + \sum_{\frac{x}{q^8}<p^k} \frac{\log p}{p^k} \right) < \frac{c_6}{\log q} \sum_{\substack{\frac{x}{q^8}<p<x \\ p\equiv a \pmod q}} \frac{\log p}{p}$$

$$+ \frac{c_{11}}{\log q} \sum_{q^{72}<n} \frac{\log n}{n^2} < \frac{c_6}{\log q} \sum_{\substack{\frac{x}{q^8}<p<x \\ p\equiv a \pmod q}} \frac{\log p}{p} + \frac{1}{q^{71}}.$$

For $q > q_2$, (2.18) follows from (2.16), (2.19) and (2.20), and this completes the proof of Lemma 5.

3. Completion of the proof of Theorem 1

Assume that $\mathcal{A}$ satisfies (1.4). Thus if a and a' are distinct elements of $\mathcal{A}$, there are positive integers x and k with $2 \leqq k \leqq K$ and

$$a + a' = x^k. \tag{3.1}$$

Let p be any prime satisfying (2.1), and define $\mathcal{B}_p$ by

$$\mathcal{B}_p = \{b : b \in \{0, 1, \dots, p-1\},\ \exists\, a \in \mathcal{A} \text{ such that } a \equiv b \pmod{p}\}.$$

Since (3.1) holds for all $a, a' \in \mathcal{A}$, $a \neq a'$ with some $x \in \mathbf{N}$ and $2 \leqq k \leqq K$, thus $\mathcal{B}_p$ satisfies the assumptions in Lemma 2 for each of these primes p. By Lemma 2 it follows that (2.3) holds. Put $|\mathcal{B}_p| = \nu(p)$ so that

$$\nu(p) < c_5 (\log K) 2^{\pi(K)} p^{1/2} \tag{3.2}$$

for all p satisfying (2.1).

As in the statement of Lemma 2, we put

$$Q_K = \prod_{q \leqq K} q,$$

where the product is taken over primes q with $q \leqq K$.

Next write $Z = CQ_K^{80}(\log N)^2$ where C is a positive number larger than 3 which will be fixed later. Clearly we have

$$Z \geqq Q_K^{80}. \tag{3.3}$$

Write

$$\mathcal{P} = \{p : p \text{ prime}, p \leqq Z,\ p \equiv 1 \pmod{Q_K}\},$$

and observe that by (3.2) we have

$$\begin{aligned} \sum_{p \in \mathcal{P}} \frac{\log p}{\nu(p)} &> c_{12} \frac{1}{(\log K) 2^{\pi(K)}} \sum_{p \in \mathcal{P}} \frac{\log p}{p^{1/2}} \\ &= c_{12} \frac{1}{(\log K) 2^{\pi(K)}} \sum_{\substack{p \leqq Z \\ p \equiv 1 \pmod{Q_K}}} \frac{\log p}{p^{1/2}} \\ &> c_{12} \frac{1}{(\log K) 2^{\pi(K)}} \sum_{\substack{\frac{Z}{Q_K^8} < p < Z \\ p \equiv 1 \pmod{Q_K}}} \frac{\log p}{p} \left(\frac{Z}{Q_K^8}\right)^{1/2}. \end{aligned} \tag{3.4}$$

By (3.3), (2.17) in Lemma 5 holds with Q_K and Z in place of q and x, respectively, so that we may use Lemma 5 to estimate the last sum in (3.4). We obtain

$$\sum_{p\in\mathcal{P}} \frac{\log p}{\nu(p)} > c_{12}\frac{1}{(\log K)2^{\pi(K)}} c_8 \frac{Z^{1/2}\log Q_K}{Q_K^4\varphi(Q_K)}$$

$$> c_{13}C^{1/2}\frac{\log Q_K \log\log Q_K}{(\log K)2^{\pi(K)}} Q_K^{35}\log N.$$

Since for $K\to\infty$ we have $Q_K = e^{(1+o(1))K}$,

$$\frac{\log Q_K \log\log Q_K}{(\log K)2^{\pi(K)}} Q_K^{35} = e^{(35+o(1))K}.$$

It follows that if C is large enough, then uniformly for $K \geqq 2$ and $N \geqq 2$ we have

$$\sum_{p\in\mathcal{P}} \frac{\log p}{\nu(p)} > 2e^{34K}\log N$$

whence

$$\sum_{p\in\mathcal{P}} \frac{\log p}{\nu(p)} - \log N > e^{34K}\log N > 0. \tag{3.5}$$

We are now in a position to apply Lemma 3. By the Brun–Titchmarsh Theorem, the numerator in (2.14) is

$$\sum_{p\in\mathcal{P}} \log p - \log N < \sum_{\substack{p\leqq Z \\ p\equiv 1 \pmod{Q_K}}} \log Z < 2\frac{Z}{\varphi(Q_K)\log Z}\log Z \tag{3.6}$$

$$< c_{14}Z\frac{\log\log Q_K}{Q_K} < c_{15}Q_K^{79}(\log N)^2\log\log Q_K < c_{16}e^{80K}(\log N)^2.$$

(1.5) follows from (2.14), (3.5) and (3.6), and this completes the proof of Theorem 1.

4. Proof of Theorem 2

Let $\mathcal{A} = \{a_1, a_2, \dots, a_t\}$, and write each a_i in the (unique) form

$$a_i = 2^{u_i}(4w_i + e_i) \tag{4.1}$$

where u_i, w_i are non-negative integers and $e_i \in \{-1, +1\}$. Set $L = \left[\frac{\log N}{\log 2}\right]$ so that in (4.1) we have $0 \leqq u_i \leqq L$. We will group the a_i's according to the values of u_i and e_i, i.e., for $0 \leqq u < L$, $e \in \{-1, +1\}$ we write

$$\mathcal{A}_{(u,e)} = \{a_i : 1 \leqq i \leqq t,\ u_i = u,\ e_i = e\}.$$

Then we have

$$\mathcal{A} = \bigcup_{u=0}^{L} \bigcup_{e \in \{-1,+1\}} \mathcal{A}_{(u,e)} \tag{4.2}$$

so that it remains to give an upper estimate for $\left|\mathcal{A}_{(u,e)}\right|$ (for all u, e).

Fix u and e, and assume that $a, a' \in \mathcal{A}_{(u,e)}$. Then by $\mathcal{A}_{(u,e)} \subseteqq \mathcal{A}$ we have $a + a' \in V$ so that there are integers x, k with $a + a' = x^k$. If q is a prime with $q \mid k$ (say, q is the smallest prime divisor of k), then, writing $y = x^{k/q}$, we have

$$a + a' = y^q. \tag{4.3}$$

Moreover, for all $a, a' \in \mathcal{A}_{(u,e)}$, clearly $a + a'$ can be written in the form

$$a + a' = 2^{u+1}(2z + 1). \tag{4.4}$$

It follows from (4.3) and (4.4) that

$$q \mid (u + 1). \tag{4.5}$$

Write

$$Q(u) = \prod_{q \mid (u+1)} q. \tag{4.6}$$

Now we will use an argument which is a variant of the proof of Theorem 1; thus we will leave some details to the reader.

We will replace Lemma 2 by

LEMMA 2′. *Let $u \in \mathbf{N}$ and define $Q(u)$ by (4.6). Assume that p is a prime with*

$$p \equiv 1 \pmod{Q(u)}. \tag{4.7}$$

Let $\mathcal{B}$ be a set of integers such that whenever b and b' are distinct elements of $\mathcal{B}$, then $b \not\equiv b' \pmod{p}$ and there exists a prime q with $q \mid (u+1)$ and an integer y for which

$$b + b' \equiv y^q \pmod{p}.$$

Then there is a positive absolute constant c_{17} such that

$$|\mathcal{B}| < c_{17} 2^{\omega(u+1)} \prod_{q \mid (u+1)} \left(1 - \frac{1}{q}\right)^{-1} p^{1/2}. \tag{4.8}$$

PROOF. Replacing Q_K by $Q(u)$ in the proof of Lemma 2′, in the same way we obtain

$$\prod_{q \mid (u+1)} \left(1 - \frac{1}{q}\right) |\mathcal{B}| \leqq p^{1/2} 2^{\omega(u+1)} + 2. \tag{4.9}$$

(4.8) follows from (4.9), and this completes the proof of the lemma.

Now let p be any prime satisfying (4.7), and define B_p by

$$\mathcal{B}_p = \left\{ b : \, b \in \{0, 1, \ldots, p-1\}, \; \exists \, a \in \mathcal{A}_{(u,e)} \text{ such that } a \equiv b \pmod{p} \right\}.$$

Since (4.3) holds for all $a, a' \in \mathcal{A}_{(u,e)}$, $a \neq a'$ with some $y \in \mathbf{N}$ and a prime q satisfying (4.5), thus $\mathcal{B}_p$ satisfies the assumptions in Lemma 2′ for each of these primes p. By Lemma 2′ it follows that, writing $|\mathcal{B}_p| = \nu(p)$ we have

$$\nu(p) = |\mathcal{B}_p| < c_{17} 2^{\omega(u+1)} \prod_{q \mid (u+1)} \left(1 - \frac{1}{q}\right)^{-1} p^{1/2}. \tag{4.10}$$

Next we write $Z(u) = C\big(Q(u)\big)^{80} (\log N)^2$ where C is a positive number larger than 3 which will be fixed later. Clearly we have

$$Z(u) \geqq \big(Q(u)\big)^{80}. \tag{4.11}$$

Write

$$\mathcal{P}(u) = \left\{ p : \, p \text{ prime}, \, p \leqq Z(u), \; p \equiv 1 \pmod{Q(u)} \right\}.$$

By (4.10) we have

$$\sum_{p\in\mathcal{P}(u)}\frac{\log p}{\nu(p)} > c_{18}2^{-\omega(u+1)}\prod_{q|(u+1)}\left(1-\frac{1}{q}\right)\sum_{p\in\mathcal{P}(u)}\frac{\log p}{p^{1/2}}$$

$$> c_{18}2^{-\omega(u+1)}\prod_{q|(u+1)}\left(1-\frac{1}{q}\right)\sum_{\substack{Z(u)/\left(Q(u)\right)^8<p\leqq Z(u)\\ p\equiv 1 \pmod{Q(u)}}}\frac{\log p}{p}\left(\frac{Z(u)}{\left(Q(u)\right)^8}\right)^{1/2}.$$

By (4.11), (2.17) in Lemma 5 holds with $Q(u)$ and $Z(u)$ in place of q and x, respectively, so that we may use Lemma 5 to estimate the last sum. We obtain

$$\sum_{p\in\mathcal{P}(u)}\frac{\log p}{\nu(p)} > c_{18}2^{-\omega(u+1)}\prod_{q|(u+1)}\left(1-\frac{1}{q}\right)\frac{\left(Z(u)\right)^{1/2}}{\left(Q(u)\right)^4}c_8\frac{\log Q(u)}{\varphi\left(Q(u)\right)}$$

$$> c_8c_{18}2^{-\omega(u+1)}C^{1/2}\left(Q(u)\right)^{35}\log Q(u)\log N.$$

Since we have

$$2^{\omega(u+1)} = \prod_{q|(u+1)}2 \leqq \prod_{q|(u+1)}q = Q(u),$$

thus it follows that

$$\sum_{p\in\mathcal{P}(u)}\frac{\log p}{\nu(p)} > c_8c_{18}\ C^{1/2}\left(Q(u)\right)^{34}\log N. \tag{4.12}$$

Now we fix C: we choose it so large (in terms of c_8 and c_{18}) that the constant factor in this lower bound is greater than 2. Then we obtain

$$\sum_{p\in\mathcal{P}(u)}\frac{\log p}{\nu(p)} > 2\left(Q(u)\right)^{34}\log N,$$

so that, by (4.12),

$$\sum_{p\in\mathcal{P}(u)}\frac{\log p}{\nu(p)} - \log N > \left(Q(u)\right)^{34}\log N > 0. \tag{4.13}$$

Thus we may apply Lemma 3 with $\mathcal{A}_{(u,e)}$ and $\mathcal{P}(u)$ in place of $\mathcal{A}$ and $\mathcal{P}$, respectively. Then by the Brun–Titchmarsh Theorem, the numerator in (2.14) is

$$\sum_{p\in\mathcal{P}(u)} \log p - \log N < \sum_{\substack{p\leqq Z(u)\\ p\equiv 1 \pmod{Q(u)}}} \log Z(u) \tag{4.14}$$

$$< 2\frac{Z(u)}{\varphi\big(Q(u)\big)\log Z(u)}\log Z(u) \leqq 2Z(u) = 2C\big(Q(u)\big)^{80}(\log N)^2.$$

It follows from (2.14), (4.13) and (4.14) that

$$\big|\mathcal{A}_{(u,e)}\big| < c_{19}\big(Q(u)\big)^{46}\log N. \tag{4.15}$$

For all $u \leqq L$ we have

$$Q(u) = \prod_{q|(u+1)} q \leqq u+1 \leqq L+1. \tag{4.16}$$

Thus it follows from (4.15) that

$$\big|\mathcal{A}_{(u,e)}\big| < c_{19}(L+1)^{46}\log N. \tag{4.17}$$

By (4.2) and (4.17) we have

$$|\mathcal{A}| = \sum_{u=0}^{L}\sum_{e\in\{-1,+1\}}\big|\mathcal{A}_{(u,e)}\big|$$

$$\leqq \max_{u,e}\big|\mathcal{A}_{(u,e)}\big|\sum_{u=0}^{L}\sum_{e\in\{-1,+1\}} 1 < c_{19}(L+1)^{46}(\log N)2(L+1)$$

$$= 2c_{19}(L+1)^{47}\log N = 2c_{19}\left(\left[\frac{\log N}{\log 2}\right]+1\right)^{47}\log N < c_{20}(\log N)^{48}$$

which completes the proof of Theorem 2.

We remark that while we cannot improve on Theorem 1 for "small" K (say, $K = o(\log\log N)$) and on the upper bound following from Theorem 2 under assumption (1.4) with "large" K (say, $K \gg \log\log N$), a small improvement could be made on the exponent of $\log N$ in Theorem 2 assuming

(1.4) with "medium size" K. This could be done replacing $Q(u)$ in the proof Theorem 2 by

$$Q_K(u) = \prod_{\substack{q|(u+1)\\ q\leqq K}} q,$$

and then instead of using a uniform (in u) upper bound of type (4.16) for $Q_K(u)$, utilizing the fact that $Q_K(u)$ is "small" for a "random" $u \leqq L$. However, this would make the proof more complicated and the improvement obtained in this way is not very significant (more could be achieved by improving on Lemma 5), thus we decided to present here only this simpler version.

5. Proof of Theorem 3

Let us put $\mathcal{A} = \{x, 2x, 3x, \dots, nx\}$ where $n = \left[\frac{\log\log N}{4\log\log N}\right]$. We will fix the value of x later. Then for all $a, a' \in \mathcal{A}$ we have $a + a' \in \{2x, 3x, \dots, 2nx\}$. Let p_i denote the i-th prime number. Define the positive integer ℓ by $p_\ell \leqq 2n < p_{\ell+1}$. For $1 \leqq t \leqq 2n-1$, write

$$t+1 = p_1^{\beta_{1,t}} p_2^{\beta_{2,t}} \cdots p_\ell^{\beta_{\ell,t}}$$

and suppose that x is also of the form $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\ell^{\alpha_\ell}$. Then $(t+1)x$ is a p_t-th power for all $1 \leqq t \leqq 2n-1$ if and only if

$$\alpha_i + \beta_{i,t} \equiv 0 \pmod{p_t} \quad \text{for} \quad 1 \leqq i \leqq \ell,\ 1 \leqq t \leqq 2n-1. \tag{5.1}$$

It follows from the Chinese Remainder Theorem that (5.1) has a solution in $(\alpha_1, \alpha_2, \dots, \alpha_\ell)$ which is unique modulo $p_1 p_2 \cdots p_{2n-1}$. So we may assume that $\alpha_1, \alpha_2, \dots, \alpha_\ell$ satisfy (5.1) and

$$0 \leqq \alpha_i < p_1 p_2 \cdots p_{2n-1} \quad \text{for} \quad 1 \leqq i \leqq \ell;$$

this defines x. Then the sum of any two elements of $\mathcal{A}$ is a power and

$$0 \leqq \alpha_i < p_1 p_2 \cdots p_{2n-1} < 3^{p_{2n-1}} < e^{3n\log n},$$

$$nx = np_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\ell^{\alpha_\ell} < n(p_1 p_2 \cdots p_\ell)^{\exp(3n\log n)} < n3^{p_\ell \exp(3n\log n)}$$

$$\leqq n3^{2n\exp(3n\log n)} < \exp\big(\exp(4n\log n)\big) < N$$

for N large enough.

6. Proof of Theorem 4

We will prove by contradiction: assume that, contrary to the assertion, there is an infinite set $\mathcal{A} = \{a_1, a_2, \dots\} \subseteqq \mathbf{N}$ satisfying (1.6). Write each a_i in the form

$$a_i = 2^{u_i}(2v_i + 1) \quad (\text{for } i = 1, 2, \dots) \tag{6.1}$$

where u_i, v_i are non-negative integers. We have to distinguish two cases.

Case 1: the sequence $\{u_1, u_2, \dots\}$ is bounded. Then there is an integer u which occurs infinitely many times in this sequence

$$u_i = u \tag{6.2}$$

for infinitely many $i \in \mathbf{N}$. The set of the a_i's satisfying (6.2) contains an infinite subset so the v_i's in the representation (6.1) are of the same parity. Denote this subset of $\mathcal{A}$ by $\mathcal{B} = \{b_1, b_2, \dots\}$ so that

$$b_i = 2^u(4w_i + e) \quad (\text{for } i = 1, 2, \dots)$$

where $e \in \{-1, +1\}$. Then for all $1 \leqq i < j$, the sum $b_i + b_j$ can be written in the form

$$b_i + b_j = 2^{(u+1)}(2z + 1).$$

By $\mathcal{B} \subset \mathcal{A}$ and (1.6) this sum must be a power; clearly, the exponent k in this power must satisfy $k \mid (u + 1)$ so that k may assume only finitely many values. Thus by the infinite Ramsey theorem [11], we may select an infinite subset $\mathcal{D} = \{d_1, d_2, \dots\} \subseteqq \mathcal{B}$ (with $d_1 < d_2 < \dots$) so that each sum $d_i + d_j$ is a k-th power with the same exponent k. Write $d_i + d_1 = x_i^k$, $d_i + d_2 = y_i^k$ (for $i = 3, 4, \dots$). Then we have $y_i^k - x_i^k = d_2 - d_1$, so that the equation

$$y_i^k - x_i^k = d_2 - d_1 \quad (> 0) \tag{6.3}$$

has infinitely many solutions. But this is clearly impossible (since the difference between the consecutive powers tends to infinity), and this contradiction completes the proof in Case 1.

Case 2: the sequence $\{u_1, u_2, \dots\}$ (defined by (6.1)) is not bounded. Then we can select an infinite subset $\{a_{i_1}, a_{i_2}, \dots\}$ of $\mathcal{A}$ so that $u_{i_1} < u_{i_2} < \cdots$. Write $a_{i_j} = b_j$, $\{b_1, b_2, \dots\} = \mathcal{B}$ ($\subseteqq \mathcal{A}$), $u_{i_j} = q_j$ so that we have

$$b_j = 2^{q_j}(2r_j + 1) \quad (\text{for } j = 1, 2, \dots) \tag{6.4}$$

and

$$q_1 < q_2 < \cdots . \tag{6.5}$$

Consider now the sums $b_j + b_1$ and $b_j + b_2$ with $j = 3, 4, \dots$. By $\mathcal{B} \subseteq \mathcal{A}$ and (1.6), these sums must be powers; by (6.4) and (6.5) the exponent of 2 in the canonical form of these sums is q_1, resp. q_2, so that the exponent in both powers $b_j + b_1$ and $b_j + b_2$ is at most q_2. Since these exponents are bounded, by the pigeon hole principle there is a pair k, ℓ (with $k, \ell \geqq 2$) and a subset $\mathcal{D} = \{d_1, d_2, \dots\} \subseteq \{b_3, b_4, \dots\}$ so that $d_i + b_1$ and $d_i + b_2$ can be written in the form $d_i + b_1 = x_i^k$, resp. $d_i + b_2 = y_i^\ell$ whence $y_i^\ell - x_i^k = b_2 - b_1$, so that the equation

$$y^\ell - x^k = b_2 - b_1 \quad (> 0) \tag{6.6}$$

has infinitely many solutions. But this is impossible for $\ell = k$ trivially (as in the case of (6.3)). For $\ell \neq k$, since (6.4) is an irreducible Thue equation with a degree $\max\{k, \ell\} \geqq 3$, thus by a well-known theorem [15], (6.6) has only finitely many solutions. This contradiction completes the proof of theorem.

7. Proof of Theorem 5

Write $P_x = \prod_{p \leqq x} p$. Let x denote the greatest positive integer with

$$\frac{1}{2} x P_x^2 \leqq N, \tag{7.1}$$

and let

$$\mathcal{A} = \left\{ i P_x^2 : \, i \leqq \frac{x}{2} \right\}. \tag{7.2}$$

Then $\mathcal{A} \subset \{1, 2, \dots N\}$. Moreover, by the prime number theorem $P_x = \exp\big(\big(1 + o(1)\big)\, x\big)$, and thus it follows that $\exp\big(\big(2 + o(1)\big)\, x\big) = N$ whence

$$x = \left(\frac{1}{2} + o(1) \right) \log N. \tag{7.3}$$

By (7.2) we have

$$|\mathcal{A}| = \left(\frac{1}{2} + o(1) \right) x. \tag{7.4}$$

(1.8) follows from (7.3) and (7.4). Finally, if $a \in \mathcal{A}$, $a' \in \mathcal{A}$, then $a + a' = \ell P_x^2$ for an integer ℓ with $1 \leqq \ell \leqq x$. Thus if p is a prime divisor of $a + a'$, it is at most x and thus p^2 divides $a + a'$, which proves (1.9).

8. Proof of Theorem 6

The lower bound can be proved by a construction similar to the one employed in the proof of Theorem 5. Define P_x as in Section 7, let x denote the greatest positive integer with $x^{1/2}P_x^2 \leqq N$, so that

$$x = \big(1/2 + o(1)\big)\log N, \tag{8.1}$$

and let $\mathcal{A} = \{iP_x^2 : i \leqq x^{1/2}\}$. Then we have

$$|\mathcal{A}| = \left(\left(\frac{1}{2} + o(1)\right)\log N\right)^{1/2} = \big(2^{-1/2} + o(1)\big)(\log N)^{1/2}$$

and every element of $\mathcal{P}(\mathcal{A})$ is powerful which proves the lower bound in (1.10).

In order to prove the upper bound in (1.10) we need two lemmas.

LEMMA 6. *If N is a positive integer, $\mathcal{A} \subset \{1, 2, \dots, N\}$ and*

$$|\mathcal{A}| > c_{21}(N\log N)^{1/2}, \tag{8.2}$$

then there are positive integers d, y, z such that

$$d < c_{22}N|\mathcal{A}|^{-1}, \quad z > c_{23}|\mathcal{A}|^2, \quad y < c_{24}Nz|\mathcal{A}|^{-2}$$

and

$$\big\{yd, (y+1)d, \dots, zd\big\} \subset \mathcal{P}(\mathcal{A}).$$

PROOF. This is a theorem of Sárközy [13]. (Note that in [13] the numerical values of the constants c_{21}, c_{22}, c_{23} and c_{27} are computed. However, we do not need these values here.)

LEMMA 7. *There is an absolute constant c_{25} such that if p is a prime,*

$$\mathcal{B} \subset \{1, 2, \dots, p^2 - 1\}$$

and

$$|\mathcal{B}| > c_{25}p(\log p)^{1/2}, \tag{8.3}$$

then there is a subset sum

$$s = \sum_{b\in\mathcal{B}} \varepsilon_b b \in \mathcal{P}(\mathcal{B}) \tag{8.4}$$

with

$$p \mid s, \quad p^2 \nmid s. \tag{8.5}$$

PROOF. If c_{25} in (8.3) is large enough, then (8.2) holds with p^2 and $\mathcal{B}$ in place of N and $\mathcal{A}$, respectively, so that we may use Lemma 6. We obtain that there are d, y and z with

$$d < c_{22}p^2|\mathcal{B}|^{-1} < c_{26}p(\log p)^{-1/2}, \tag{8.6}$$

$$z > c_{23}|\mathcal{B}|^2 > c_{27}p^2 \log p, \tag{8.7}$$

$$y < c_{24}p^2 z|\mathcal{B}|^{-2} < c_{28}z(\log p)^{-1} \tag{8.8}$$

and

$$\{yd, (y+1)d, \dots, zd\} \subset \mathcal{P}(\mathcal{B}). \tag{8.9}$$

If p is large enough (which, clearly, can be assumed), then it follows from (8.6) that

$$(d, p) = 1. \tag{8.10}$$

Consider the first two multiples of p in the arithmetic progression yd, $(y+1)d, (y+2)d, \dots$. By (8.10) one of them, say, ud satisfies

$$p \mid ud, \quad p^2 \nmid ud, \tag{8.11}$$

and we also have

$$y \leqq u < y + 2p. \tag{8.12}$$

By (8.7) and (8.8) for large enough p we have

$$z - y > z - c_{28}z(\log p)^{-1} > \frac{z}{2} > c_{29}p^2 \log p. \tag{8.13}$$

It follows from (8.12) and (8.13) that $y \leqq u < z$ so that, by (8.9),

$$ud \in \mathcal{P}(\mathcal{A}). \tag{8.14}$$

By (8.11) and (8.14), (8.4) and (8.5) hold with $s = ud$ and this completes the proof of the lemma.

By using Lemma 7, we may complete the proof of the upper bound in (8.15) in the following way. It suffices to show that if

$$\mathcal{A} \subset \{1, 2, \dots, N\} \tag{8.15}$$

and

every element of $\mathcal{P}(\mathcal{A})$ is powerful, (8.16)

then we must have

$$|\mathcal{A}| < c_4 (\log N)^3 (\log\log N)^{-1/2}. \tag{8.17}$$

Again we write $P_x = \prod_{p \leqq x} p$, and let x denote the smallest positive integer with

$$P_x^2 > N \tag{8.18}$$

so that, by the prime number theorem,

$$x = \left(\frac{1}{2} + o(1)\right) \log N. \tag{8.19}$$

For all $p \leqq x$, write

$$\mathcal{A}_p = \{a : a \in \mathcal{A},\ p^2 \nmid a\}. \tag{8.20}$$

If there is an $\overline{a}$ with

$$\overline{a} \in \mathcal{A} \setminus \bigcup_{p \leqq x} \mathcal{A}_p \tag{8.21}$$

then $\overline{a} \notin \mathcal{A}_p$ for all $p \leqq x$, whence $p^2 \mid \overline{a}$ for all $p \leqq x$ so that

$$P_x^2 \mid \overline{a}. \tag{8.22}$$

However, by (8.15), (8.18) and (8.21) this is not possible. Thus there is no $\overline{a}$ satisfying (8.21) so that

$$\mathcal{A} = \bigcup_{p \leqq x} \mathcal{A}_p. \tag{8.23}$$

Thus in order to prove (8.17), we have to give an upper bound for $|\mathcal{A}_p|$ for all $p \leqq x$.

Fix a prime $p \leqq x$, and for $n \in \mathbf{N}$ let $r(n)$ denote the least non-negative residue of n modulo p^2, so that

$$r(n) \equiv n \pmod{p^2}, \quad 0 \leqq r(n) < p^2.$$

Write

$$\mathcal{B} = \left\{ b : \exists\, a \text{ with } a \in \mathcal{A}_p,\ r(a) = b \right\}.$$

By (8.20) we have

$$0 \notin \mathcal{B}. \tag{8.24}$$

Assume now that there is an

$$s = b_1 + \cdots + b_t \in \mathcal{P}(\mathcal{B}) \quad (\text{where } b_1, \ldots, b_t \in \mathcal{B},\ b_1 < \cdots < b_t) \tag{8.25}$$

with

$$p \mid s, \quad p^2 \nmid s. \tag{8.26}$$

Then by the definition of $\mathcal{B}$, there are

$$a_1, \ldots, a_t \in \mathcal{A} \tag{8.27}$$

with $r(a_1) = b_1, \ldots, r(a_t) = b_t$ whence

$$a_1 + \cdots + a_t \equiv b_1 + \cdots + b_t \equiv s \pmod{p^2},$$

so that, by (8.26), $p \mid a_1 + \cdots + a_t$, $p^2 \nmid a_1 + \cdots + a_t$, and thus

$$a_1 + \cdots + a_t \text{ is not powerful.} \tag{8.28}$$

Moreover, the numbers $a_1, \ldots, a_t$ are distinct and so by (8.27) we have

$$a_1 + \cdots + a_t \in \mathcal{P}(\mathcal{A}). \tag{8.29}$$

But (8.28) and (8.29) contradict (8.16), and this shows that there is no s satisfying (8.25) and (8.26). In particular

$$\text{if} \quad s \in \mathcal{P}(\mathcal{B}) \quad \text{then} \quad p \mid s,\ p^2 \nmid s \text{ cannot hold.} \tag{8.30}$$

By Lemma 7, it follows that

$$|\mathcal{B}| \leqq c_{25} p (\log p)^{1/2}. \tag{8.31}$$

Now we will show that for all $b \in \mathcal{B}$ we have

$$\left|\left\{a : a \in \mathcal{A}_p,\ a \equiv b \pmod{p^2}\right\}\right| < p. \tag{8.32}$$

By (8.24) and (8.30), and since clearly $\mathcal{B} \subset \mathcal{P}(\mathcal{B})$, we have

$$(b, p) = 1 \quad \text{for all } b \in \mathcal{B}. \tag{8.33}$$

Assume now that, contrary to (8.32), for some $b \in \mathcal{B}$ there are $a_1, \dots, a_p \in \mathcal{A}$ with $a_1 < \dots < a_p$ and

$$a_1 \equiv \dots \equiv a_p \equiv b \pmod{p^2}. \tag{8.34}$$

Then we have

$$a_1 + \dots + a_p \in \mathcal{P}(\mathcal{A}) \tag{8.35}$$

and, by (8.34), $a_1 + \dots + a_p \equiv pb \pmod{p^2}$. By (8.33), it follows that $p \mid a_1 + \dots + a_p$ and $p^2 \nmid a_1 + \dots + a_p$ and thus

$$a_1 + \dots + a_p \text{ is not powerful.} \tag{8.36}$$

(8.35) and (8.36) contradict (8.16), and this completes the proof of (8.32).

It follows from (8.31) and (8.32) that

$$|\mathcal{A}_p| = \sum_{b \in \mathcal{B}} \left|\left\{a : a \in \mathcal{A}_p,\ a \equiv b \pmod{p^2}\right\}\right| < \sum_{b \in \mathcal{B}} p = |\mathcal{B}| p \leqq c_{25} p^2 (\log p)^{1/2}.$$

Thus, by (8.19) and (8.23), we have

$$|\mathcal{A}| \leqq \sum_{p \leqq x} |\mathcal{A}_p| \leqq c_{20} \sum_{p \leqq x} p^2 (\log p)^{1/2} < c_{30} (\log N)^3 (\log\log N)^{-1/2}$$

which proves (8.17) and completes the proof of Theorem 6.

Acknowledgement. We would like to thank the referee, Professor I. Z. Ruzsa for his valuable remarks and, in particular, for his suggestion which enabled us to improve considerably on our Theorems 2, 3 and 4.

References

[1] J. R. Chen, On the least prime in an arithmetical progression and theorems concerning the zeros of Dirichlet's L-functions, II, *Sci. Sinica,* **22** (1979), 859–889.
[2] A. Dujella, Absolute bound for the size of Diophantine m-tuples, *Acta Arith.,* to appear.
[3] P. Erdős, Quelques problemes de la theorie des nombres, *Monographie de l'enseignement mathématique,* Genève (1963), 81–135.
[4] P. Erdős and H. N. Shapiro, On the least primitive root of a prime, *Pacific J. Math.,* **7** (1957), 861–865.
[5] P. X. Gallagher, A larger sieve, *Acta Arith.,* **18** (1971), 77–81.
[6] S. Graham, On Linnik's constant, *Acta Arith.,* **39** (1981), 163–179.
[7] K. Gyarmati, On a problem of Diophantus, *Acta Arith.,* **97** (2001), 53–65.
[8] K. Gyarmati, A. Sárközy and C. L. Stewart, On shifted products which are powers, *Mathematika,* to appear.
[9] R. Heath-Brown, Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression, *Proc. London Math. Soc.,* **64** (1992), 265–338.
[10] M. Jutila, On Linnik's constant, *Math. Scand.,* **41** (1977), 45–62.
[11] F. P. Ramsey, On a problem for formal logic, *Proc. London. Math. Soc.,* **30** (1930), 264–286.
[12] J. Rivat, A. Sárközy and C. L. Stewart, Congruence properties of the Ω-function on sumsets, *Illinois J. Math.,* **43** (1999), 1–18.
[13] A. Sárközy, Finite addition theorems, II, *J. Number Theory,* **48** (1994), 197–218.
[14] A. Sárközy, Unsolved problems in number theory, *Periodica Math. Hungar.,* **42** (2001), 17–35.
[15] A. Thue, Über Annäherungswerte algebraischer Zahlen, *J. reine angew. Math.,* **135** (1909), 284–305.

(Received March 9, 2001; revised August 21, 2002)

DEPARTMENT OF ALGEBRA AND NUMBER THEORY
EÖTVÖS LORÁND UNIVERSITY
H-1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C
HUNGARY
E-MAIL: GYKATI@CS.ELTE.HU
SARKOZY@CS.ELTE.HU

DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA N2L 3G1
E-MAIL: CSTEWART@WATSERV1.UWATERLOO.CA