ON MULTIPLICATIVE BASES OF FINITE SETS

Katalin Fried
Department of Mathematics Teaching and Education Center, Eotvos Lorand University, Institute of Mathematics, Budapest, Hungary
kfried@cs.elte.hu

Katalin Gyarmati[1]
Department of Algebra and Number Theory and MTA–ELTE Geometric and Algebraic Combinatorics Research Group, Eotvos Lorand University, Institute of Mathematics, Budapest, Hungary
gykati@cs.elte.hu

Received: 7/14/17, Revised:1/31/20, Accepted: 6/8/20, Published: 6/18/20

Abstract

We study the density of multiplicative bases of subsets of $\mathbb{Z}$ formed by values of polynomials.

1. Introduction

We will use the following notation. For a set $\mathcal{S} \subseteq \mathbb{Z}$ we denote by $\mathcal{S}(n)$ the cardinality of $\mathcal{S} \cap [1, 2, \dots, n]$. We say that a set $\mathcal{B} \subseteq \mathbb{Z}$ forms a *multiplicative basis of order h of $\mathcal{S}$* if every element of $\mathcal{S}$ can be written as the product of h members of $\mathcal{B}$. While additive bases has been an intensively studied topic in additive number theory, much less attention has been devoted to multiplicative bases. First, multiplicative bases of $[n] \stackrel{\text{def}}{=} [1, 2, \dots, n]$ were studied. It is easy to see that every multiplicative basis of $[n]$ contains the prime numbers up to n. On the other hand, in 2011, Chan [2] proved that there is a multiplicative basis with less than $\pi(n) + c(h+1)^2 \, \frac{n^{2/(h+1)}}{\log^2 n}$ elements (however, he did not use the phrase "multiplicative basis"). This upper bound has been recently sharpened by a factor of h by Pach and Sándor [24]. Namely, if $G_h(n)$ denotes the size of the smallest multiplicative basis of order h of $[n]$, then

$$\pi(n) + 0.5h\frac{n^{2/(h+1)}}{\log^2 n} \leq G_h(n) \leq \pi(n) + 150.4h\frac{n^{2/(h+1)}}{\log^2 n}.$$

[1]Research supported by Hungarian National Research Development and Innovation Funds NK 104183 and K 119528.

Somewhat related problems were studied by Erdős [9]. Next, a few definitions follow.

Definition 1. For a set $\mathcal{S}$ we denote by $G_h(\mathcal{S})$ the size of the smallest multiplicative basis of order h. A basis $\mathcal{B}$ of order h is a *minimal basis of order h of $\mathcal{S}$* if $|\mathcal{B}| = |G_h(\mathcal{S})|$. We call $\mathcal{B}$ a *giant basis of order h of $\mathcal{S}$* if $|\mathcal{B}| \geq |\{1\} \cup \mathcal{S}|$.

In this paper we will study multiplicative bases of order 2 of the set $S(f(x), n) \stackrel{\text{def}}{=} [f(1), f(2), \ldots, f(n)]$ where $f(x) \in \mathbb{Z}[x]$ is a polynomial. Here we remark that related problems were studied by Hajdu and Sárközy in [12], [13] and [14], namely, they studied the multiplicative decomposability of infinite polynomial sets.

Clearly, if $f(x)$ is of the form $f(x) = x^r$ then the following result immediately follows from Chan [2] and Pach and Sándor [24].

Proposition 1. *For every three positive integers r, h and n we have*

$$\pi(n) \leq G_h(S(x^r, n)) \leq \pi(n) + 150.4h\frac{n^{2/(h+1)}}{\log^2 n}.$$

So, for these polynomials $f(x) = x^r$, we know the exact order of magnitude of $G_h(S(f(x), n))$. Now we will study the case of other polynomials. First, we study the simplest case $f(x) = x^2 + 1$. One may conjecture that $S(x^2 + 1, n)$ has only giant bases, but it turns out that this is not the case. There exists a basis with slightly fewer elements than $|\{1\} \cup S(f(x), n)|$. On the other hand we will prove that every multiplicative basis of $S(x^2 + 1, n)$ has at least as many elements as the number of prime numbers of the form $4k + 1$ between n and $2n$. In other words, we will prove the following result.

Theorem 1. *For every $\varepsilon > 0$ there exists a constant $n_0 = n_0(\varepsilon)$ such that for $n > n_0$ we have*

$$\left(\frac{1}{2} - \varepsilon\right)\frac{n}{\log n} \leq G_h(S(x^2 + 1, n)) \leq n - n^{1/2} + (1 + \varepsilon)n^{1/4}.$$

There is a huge gap between the lower and upper bound. An interesting question is the following: which of these bounds is closer to the truth.

Problem 1. Does there exist a constant $\varepsilon_1 > 0$ such that

$$\varepsilon_1 n \leq G_2(S(x^2 + 1, n)) \leq (1 - \varepsilon_1)n$$

is always true?

Next, we study the case of general polynomials $f(x)$. In this case we will be able to prove the following result.

Theorem 2. *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial of degree $r \geq 2$ and write $f(x)$ as a product of irreducible polynomials over $\mathbb{Z}[x]$, say*

$$f(x) = f_1(x)f_2(x)\cdots f_s(x), \tag{1}$$

where s denotes the number of irreducible factors in (1)*. Then*

$$\frac{n}{(\log n)^{s\log r/\log 2}} \ll G_2(S(f(x),n)).$$

We remark that from Theorem 2 we immediately get the following result.

Corollary 1. *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial of degree $r \geq 2$. Then*

$$\frac{n}{(\log n)^{r\log r/\log 2}} \ll G_2(S(f(x),n)).$$

In the case of the polynomial $f(x) = x^2 + 1$, Theorem 2 gives the same lower bound as the one in Theorem 1.

As a general upper bound one can give the trivial bound $|\{1\} \cup S(f(x),n)| \leq n+1$. Regarding the upper bound we ask the following questions.

Problem 2. Is there any polynomial $f(x)$ such that for every n the set $S(f(x),n)$ has only giant bases of order 2, in other words do we have

$$|\mathcal{B}| \geq |\{1\} \cup S(f(x),n)|$$

for every basis $\mathcal{B}$ of order 2? Or, is there a general non-trivial upper bound for $G_2(S(f(x),n))$?

Perhaps the lower bound in Theorem 2 can be sharpened. We also ask the following question.

Problem 3. Is it possible to give a better general lower bound for $G_2(S(f(x),n))$ than the bound $\frac{n}{(\log n)^{s\log r/log2}}$ in Theorem 2?

So far we have been considering multiplicative bases of $S(f(x),n) = \{f(1), f(2), f(3), \ldots, f(n)\}$. Next, we study the multiplicative bases of its subsets, i.e., sets of the form

$$\mathcal{W} \stackrel{\text{def}}{=} \{f(a_1), f(a_2), f(a_3), \ldots, f(a_k)\}, \tag{2}$$

where $1 \leq a_1 < a_2 < \cdots < a_k \leq n$ are integers. If $\mathcal{B}$ is a multiplicative basis of order 2 of $\mathcal{W}$, then each element of $\mathcal{W}$ can be written in the form b_ib_j with $b_i, b_j \in \mathcal{B}$, and thus

$$|\mathcal{W}| \leq |\mathcal{B}|^2,$$

and so

$$|\mathcal{W}|^{1/2} \leq |\mathcal{B}|\,. \tag{3}$$

In the case of polynomials $f(x)$ of degree 2, this problem is somewhat related to the study of Diophantine tuples (see e.g. [1], [4], [5], [6], [7], [8], [15]).

Our goal is to study whether (3) is the best possible general lower bound. Under some not too restrictive conditions on the a_i's in $\mathcal{W}$ we will prove $|\mathcal{W}|^{2/3} \ll |\mathcal{B}|$.

Theorem 3. *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial of degree $\deg f \geq 2$ and $u, a_1, a_2, \ldots, a_k$ be positive integers such that*

$$u \leq a_1 < a_2 < \cdots < a_k < 2u. \tag{4}$$

We define $\mathcal{W}$ by (2). If $\mathcal{B}$ is a multiplicative basis of order 2 of $\mathcal{W}$, then

$$|\mathcal{W}|^{2/3} \ll |\mathcal{B}|\,, \tag{5}$$

where the implied constant factor depends only on the polynomial $f(x)$.

Remark 1. If $f(x)$ is of the form $f(x) = x^r + a_{r-3}x^{r-3} + \cdots + a_{r-4}x^{r-4} + \cdots + a_0$ (so the coefficients of the terms x^{r-1} and x^{r-2} are 0) and if instead of (4) only $u \leq a_1 < a_2 < \cdots < a_k < u^2$ is assumed, then $|\mathcal{W}|^{2/3} \ll |\mathcal{B}|$ can be proved similarly to Theorem 3.

Regarding Theorem 3 we ask the following question.

Problem 4. Is it true that the lower bound (5) holds for arbitrary a_i's, i.e., is condition (4) necessary in Theorem 3? In this general case what lower bound can be given for $|\mathcal{B}|$?

Remark 2. Let $\mathcal{B}$ be a multiplicative basis of order 2 of the set $\mathcal{W}$ defined in Theorem 3. Probably, the lower bound (5) in the case of certain special polynomials can be sharpened to $|\mathcal{W}|^{3/4} \ll |\mathcal{B}|$. For more details see the end of the proof of Theorem 3.

Finally, we will present a problem concerning sets having only giant bases. Clearly the set $I = \{a^2, a^2+1, a^2+2, \ldots, a^2+a\}$ has only giant bases. To see this, let $\mathcal{B}$ be a multiplicative basis of I of order 2. We split $\mathcal{B}$ into two disjoint subsets, so $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ where

$$\mathcal{B}_1 \stackrel{\text{def}}{=} \{b \in \mathcal{B} :\ b \leq a\},$$
$$\mathcal{B}_2 \stackrel{\text{def}}{=} \{b \in \mathcal{B} :\ b \geq a+1\}.$$

If $b_i b_j \in I$ and $b_i < b_j$, then $b_i \leq a$ and $b_j \geq a+1$. Thus for $b_i b_j \in I$ and $b_i < b_j$, we have $b_i \in \mathcal{B}_1$ and $b_j \in \mathcal{B}_2$.

For each $b \in \mathcal{B}_2$ there exists at most one element i of I for which $b \mid i$ since $|I| = a + 1 \leq b$. Thus

$$a + 1 = |I| \leq |\mathcal{B}_2| < |\mathcal{B}|\,,$$

from which the statement follows.

Our final problem is the following problem.

Problem 5. Let $I = \{m + 1, m + 2, \ldots, m + n\}$, where m and n are integers. For what m and n's does I have only giant bases?

2. Proofs of Theorems 1, 2 and 3

2.1. Proof of Theorem 1

First, we prove that for $n > n_0(\varepsilon)$ we have

$$\left(\frac{1}{2} - \varepsilon\right) \frac{n}{\log n} \leq G_h(S(x^2 + 1, n)). \tag{6}$$

Let $\mathcal{B}$ be a multiplicative basis of order h of $S(x^2 + 1, n)$. Let $\mathcal{P}$ denote the following set

$$\mathcal{P} \stackrel{\text{def}}{=} \{p : \ p \text{ is a prime of the form } 4k + 1 \text{ and } n < p < 2n\}. \tag{7}$$

To every prime $p \in \mathcal{P}$ we assign the smallest positive integer $g = g(p)$ with

$$p \mid g(p)^2 + 1.$$

If $p \in \mathcal{P}$, then p is a prime number of the form $4k + 1$, thus the congruence

$$x^2 \equiv -1 \pmod{p}$$

has two different solutions and one of them is between 1 and $(p - 1)/2$, so that

$$1 \leq g(p) \leq \frac{p - 1}{2} < n. \tag{8}$$

Now $\mathcal{B}$ is a multiplicative basis of $S(x^2 + 1, n)$, thus it is also a multiplicative basis of its subsets, so that $\mathcal{B}$ is a multiplicative basis of

$$S_1 \stackrel{\text{def}}{=} \{g(p)^2 + 1 : \ p \in \mathcal{P}\}$$

since $S_1 \subset S(x^2 + 1, n)$ by (8).

For every $p \in \mathcal{P}$, S_1 contains a multiple of p since $p \mid g(p)^2 + 1$. Thus $\mathcal{B}$ contains a multiple of p, which we denote by $h(p)$. Then $h(p) \in \mathcal{B}$ and $p \mid h(p)$.

We will prove that for $p, q \in \mathcal{P}$, $p \neq q$,

$$h(p) = h(q)$$

is not possible. Assume that contrary to this statement we have $p \neq q$ and $h(p) = h(q)$. Then

$$p \mid h(p),\ q \mid h(q).$$

Thus

$$pq \mid h(p) = h(q).$$

By $p, q \in \mathcal{P}$ we have $n + 1 \leq p, q$, whence

$$(n+1)^2 \leq pq \leq h(p) = h(q). \tag{9}$$

But $\mathcal{B}$ is a multiplicative basis of $S(x^2+1, n)$ so that its elements are less or equal than $n^2 + 1$, thus

$$h(p) = h(q) \leq n^2 + 1,$$

which contradicts (9).

Thus the function $h:\ \mathcal{P} \to \mathcal{B}$ is injective, so that

$$|\mathcal{P}| \leq |\mathcal{B}|,$$

which proves (6).

In order to prove

$$G_h(S(x^2+1, n)) \leq n - n^{1/2} + (1+\varepsilon)n^{1/4},$$

we will prove a slightly stronger upper bound, namely, $G_h(S(x^2+1, n)) \leq n - n^{1/2} + n^{1/4} + 2$. It is enough to construct a multiplicative basis $\mathcal{B}$ of order h of $S(x^2+1, n)$ with

$$|\mathcal{B}| \leq n - n^{1/2} + n^{1/4} + 2.$$

First, observe that

$$\left(a^2+1\right)\left((a+1)^2+1\right) = \left(a^2+a+1\right)^2 + 1. \tag{10}$$

Let

$$\mathcal{B} \stackrel{\text{def}}{=} \{x^2+1:\ 0 \leq x \leq n\} \setminus \{\left(a^2+a+1\right)^2 + 1:\ n^{1/2} + 0.5 \leq a^2+a+1 \leq n\}.$$

In order to prove that $\mathcal{B}$ is a multiplicative basis of order h it is enough to prove that for $1 \leq x \leq n$ the integer x^2+1 can be written as a product of h elements of $\mathcal{B}$. If x is not of the form a^2+a+1 where $n^{1/2} + 0.5 \leq a^2+a+1 \leq n$, then it is clear that

$$x^2 + 1 = b_1 b_2 b_3 \cdots b_h \tag{11}$$

where $b_1 = x^2 + 1 \in \mathcal{B}$ and $b_2 = b_3 = \cdots = b_h = 1 \in \mathcal{B}$.

If $x = a_1^2 + a_1 + 1$ for some integer a_1 and $n^{1/2} + 0.5 \leq a_1^2 + a_1 + 1 \leq n$, then by (10) we have

$$x^2 + 1 = \left(a_1^2 + a_1 + 1\right)^2 + 1 = \left(a_1^2 + 1\right)\left((a_1 + 1)^2 + 1\right).$$

Thus

$$x^2 + 1 = b_1 b_2 b_3 \cdots b_h,$$

with $b_1 = a_1^2 + 1,\ b_2 = (a_1 + 1)^2 + 1,\ b_3 = \cdots = b_h = 1$. It is easy to see that since $a_1^2 + a_1 + 1 \leq n$, it follows that

$$a_1 < a_1 + 1 < n^{1/2} + 0.5.$$

Then

$$b_1, b_2 \notin \{y^2 + 1 : n^{1/2} + 0.5 \leq y \leq n\},$$

therefore

$$b_1, b_2 \notin \{\left(a^2 + a + 1\right)^2 + 1 :\ n^{1/2} + 0.5 \leq a^2 + a + 1 \leq n\}.$$

Thus by the definition of $\mathcal{B}$ we have $b_1, b_2 \in \mathcal{B}$ and we also have $b_3 = b_4 = \cdots = b_h = 1 \in \mathcal{B}$. Computing the number of elements of $\mathcal{B}$ we get

$$|\mathcal{B}| \leq n - n^{1/2} + (1 + \varepsilon)n^{1/4},$$

which was to be proved. □

2.2. Proof of Theorem 2

Throughout the proof $c_1, c_2, c_3, \ldots$ will denote constants depending only on the polynomial $f(x)$. We may also suppose that the leading coefficient of $f(x)$ is positive.

Let $\tau(a)$ denote the number of positive divisors of a positive integer a. It is well-known that

$$\sum_{a=1}^{n} \tau(a) = n \log n + O(n).$$

In 1952 Erdős [10] extended this result to polynomials, namely, he proved the following result.

Lemma 1 (Erdős). *Let $f(x) \in \mathbb{Z}[x]$ be an irreducible polynomial. There exist positive integers c_1 and c_2 depending on $f(x)$ such that for $n \geq 2$ we have*

$$c_1 n \log n < \sum_{a=1}^{n} \tau(f(a)) < c_2 n \log n. \tag{12}$$

Erdős gave an existence proof, but he could not give bounds on the order of magnitude of the constants c_1 and c_2 in Lemma 1. Recently Lapkova [19] achieved some good bounds in the case of polynomials of degree 2. Related results can be found in [3].

In order to prove Theorem 2 we will need only the upper bound in (12). Let s denote the number of irreducible factors $f_j(x)$ in (1). Using Erdős's lemma we will prove the following result.

Lemma 2. *There exists a constant c_3 depending only on the polynomial $f(x)$ such that for every integer n large enough we have that the set*

$$E(f(x),n) \stackrel{\text{def}}{=} \{a : n/4 \le a \le n \text{ and } \tau(f(a)) < c_3(\log n)^s\} \tag{13}$$

has at least $n/4$ different elements.

Proof of Lemma 2. Let s denote the number of irreducible factors $f_j(x)$ in (1). By Erdős's lemma, for $1 \le j \le s$ we have

$$\sum_{a=1}^{n} \tau(f_j(a)) < c_2 n \log n.$$

Thus

$$\sum_{a=1}^{n} (\tau(f_1(a)) + \tau(f_2(a)) + \cdots + \tau(f_s(a))) < s c_2 n \log n = c_4 n \log n. \tag{14}$$

Let

$$\mathcal{A}_1 \stackrel{\text{def}}{=} \{1 \le a \le n : \ \tau(f_1(a)) + \tau(f_2(a)) + \cdots + \tau(f_s(a)) \ge 2c_4 \log n\},$$
$$\mathcal{A}_2 \stackrel{\text{def}}{=} \{1 \le a \le n : \ \tau(f_1(a)) + \tau(f_2(a)) + \cdots + \tau(f_s(a)) < 2c_4 \log n\}.$$

Clearly $\mathcal{A}_1$ and $\mathcal{A}_2$ are disjoint and

$$|\mathcal{A}_1| + |\mathcal{A}_2| = n. \tag{15}$$

By (14)

$$\begin{aligned} |\mathcal{A}_1| \cdot 2c_4 \log n &\le \sum_{a \in \mathcal{A}_1} (\tau(f_1(a)) + \tau(f_2(a)) + \cdots + \tau(f_s(a))) \\ &\le \sum_{a=1}^{n} (\tau(f_1(a)) + \tau(f_2(a)) + \cdots + \tau(f_s(a))) \\ &< c_4 n \log n. \end{aligned}$$

Thus

$$|\mathcal{A}_1| < n/2.$$

From this and (15) we have

$$|\mathcal{A}_2| > n/2. \tag{16}$$

Next, we will use the inequality $\tau(xy) \le \tau(x)\tau(y)$ and the inequality between the arithmetic and geometric means. For $a \in \mathcal{A}_2$ we have

$$\begin{aligned}\tau(f(a)) &= \tau\left(f_1(a)f_2(a)\cdots\tau f_s(a)\right)\\ &\le \tau\left(f_1(a)\right)\tau\left(f_2(a)\right)\cdots\tau\left(f_s(a)\right)\\ &\le \left(\frac{\tau\left(f_1(a)\right)+\tau\left(f_2(a)\right)+\cdots+\tau\left(f_s(a)\right)}{s}\right)^s\\ &< \left(\frac{2c_4\log n}{s}\right)^s\\ &= c_5(\log n)^s. \end{aligned}\tag{17}$$

Define $\mathcal{C}$ by

$$\mathcal{C} \stackrel{\text{def}}{=} \{a:\ n/4 \le a < n \text{ and } a \in \mathcal{A}_2\}.$$

Clearly by (16) we have

$$|\mathcal{C}| \ge |\mathcal{A}_2| - n/4 > n/4. \tag{18}$$

By $\mathcal{C} \subseteq \mathcal{A}_2$ and (17) we have for $a \in \mathcal{C}$

$$\tau(f(a)) < c_5(\log n)^s.$$

Thus, if we define $E(f(x),n)$ by (13) with c_5 in place of c_3, we have $\mathcal{C} \subseteq E(f(x),n)$. By this and (18) we have

$$n/4 < |\mathcal{C}| \le |E(f(x),n)|,$$

which proves Lemma 2. □

Define $F(f(x),n)$ by

$$F(f(x),n) \stackrel{\text{def}}{=} \{f(a) : n/4 \le a \le n \text{ and } \tau(f(a)) < c_3(\log n)^s\}. \tag{19}$$

For a fixed number c the equation $f(x) = c$ has at most $r = \deg f$ solutions. Thus we have

$$|F(f(x),n)| \ge \frac{1}{r}\,|E(f(x),n)| > \frac{n}{4r} = c_6 n. \tag{20}$$

Next, we prove the following lemma.

Lemma 3. *Let $\mathcal{B}$ be a multiplicative basis of $F(f(x),n)$ of order 2. Then*

$$|\mathcal{B}| \gg \frac{n}{(\log n)^{s\log r/\log 2}}.$$

From Lemma 3 we immediately get Theorem 2. If $\mathcal{B}$ is a multiplicative basis of $S(f(x), n)$, then it is also a multiplicative basis of $F(f(x), n)$ since $F(f(x), n) \subseteq S(f(x), n)$.

Proof of Lemma 3. Define a graph $\mathcal{G}$ as follows: Its vertices are the elements of $\mathcal{B}$. Two vertices v_1, v_2 are joined by an edge $\{v_1, v_2\}$ if and only if

$$v_1 v_2 \in F(f(x), n).$$

In other words, there exists $a \in E(f(x), n)$ (so $n/4 \leq a < n$ and $\tau(f(a)) < c_3(\log n)^s$) such that

$$v_1 v_2 = f(a). \tag{21}$$

By the definition of $F(f(x), n)$ we have

$$\max\{\tau(v_1), \tau(v_2)\} \leq \tau(v_1 v_2) < c_3(\log n)^s. \tag{22}$$

Then for the number of vertices and edges of $\mathcal{G}$ we have

$$|V(\mathcal{G})| = |\mathcal{B}|, \tag{23}$$

$$|E(\mathcal{G})| \geq |F(f(x), n)| > c_6 n. \tag{24}$$

Let $f(x)$ be of the form $f(x) = a_r x^r + a_{r-1} x^{r-1} + \cdots + a_1 x + a_0$. In (21) we have $a \geq n/4$, thus for n large enough we have

$$v_1 v_2 = f(a) > \frac{a_r}{2} a^r > \frac{a_r}{2} (n/4)^r = c_7^2 n^r \geq c_7^2 n^2.$$

It follows that for an arbitrary edge $e = \{v_1, v_2\}$ of $\mathcal{G}$ we have

$$\text{either } v_1 > c_7 n \text{ or } v_2 > c_7 n. \tag{25}$$

We split the set of vertices $\mathcal{B}$ into two disjoint sets:

$$\mathcal{B}_1 = \{v \in \mathcal{B} : \ v > c_7 n\},$$
$$\mathcal{B}_2 = \{v \in \mathcal{B} : \ v \leq c_7 n\}.$$

By (25), clearly for every edge $e = \{v_1, v_2\}$ of $\mathcal{G}$ we have $v_1 \in \mathcal{B}_1$ or $v_2 \in \mathcal{B}_1$. Thus if we denote the degree of a vertex $v \in \mathcal{B}$ in $\mathcal{G}$ by $d(v)$, then

$$|E(\mathcal{G})| \leq \sum_{v \in \mathcal{B}_1} d(v). \tag{26}$$

In Lemma 4 we give an upper bound for the degree of a vertex of $\mathcal{B}_1$.

Lemma 4. *For $v \in \mathcal{B}_1$ we have*

$$d(v) \ll (\log n)^{s \log r / \log 2}.$$

Before proving Lemma 4 we show that from Lemma 4 we immediately get Lemma 3. By Lemma 4, (24), and (26), we have

$$c_6 n < |E(\mathcal{G})| \le \sum_{v\in\mathcal{B}_1} d(v) \ll \sum_{v\in\mathcal{B}_1} (\log n)^{s\log r/\log 2} \ll |\mathcal{B}_1|\,(\log n)^{s\log r/\log 2}$$
$$\ll |\mathcal{B}|\,(\log n)^{s\log r/\log 2},$$

whence

$$\frac{n}{(\log n)^{s\log r/\log 2}} < |\mathcal{B}|\,,$$

which proves Lemma 3. □

Thus in order to prove Theorem 2 it is enough to prove Lemma 4.

Proof of Lemma 4. If $d(v) = 0$, then the statement of the lemma is trivial. Now assume that there exists $v' \in \mathcal{B}$ such that $e = \{v, v'\}$ is an edge of $\mathcal{G}$, so that there exists $n/4 \le a < n$ for which $\tau(f(a)) < c_3(\log n)^s$ and

$$vv' = f(a).$$

Then

$$\tau(v) \le \tau(vv') = \tau(f(a)) < c_3(\log n)^s. \tag{27}$$

Next, we introduce some notation. Let $D(f)$ denote the discriminant of the polynomial $f(x)$. For a prime p let $\ell(p)$ denote the largest integer for which

$$p^{\ell(p)} \mid D(f)$$

(thus $p^{\ell(p)+1} \nmid D(f)$). For $m \in \mathbb{N}$ denote by $N(f(x), m)$ the number of solutions of the congruence

$$f(x) \equiv 0 \pmod{m}.$$

In 1921, Nagel [20] and Ore [21] proved that if p is a prime and $k \in \mathbb{N}$, then

$$N(f(x), p^k) \le r p^{2\ell(p)}. \tag{28}$$

This was considerably improved by Sándor [22], Huxley [16] and Stewart [23], but for our purpose (28) is sufficient. Let m be a composite number. By the Chinese Remainder Theorem we have

$$N(f(x), m) = \prod_{p^k \| m} N(f(x), p^k).$$

Using (28) we have

$$\begin{aligned} N(f(x),m) &\le \prod_{p\mid m} rp^{2\ell(p)} = r^{\omega(m)} \prod_{p\mid m} p^{2\ell(p)} \\ &= r^{\omega(m)} \prod_{p\mid m,\ \ell(p)\neq 0} p^{2\ell(p)} \le r^{\omega(m)} \prod_{p,\ \ell(p)\neq 0} p^{2\ell(p)} \\ &\le r^{\omega(m)} \prod_{p\mid D(f)} p^{2\ell(p)} = r^{\omega(m)} D(f)^2 \\ &= c_8 r^{\omega(m)}. \end{aligned} \tag{29}$$

Now we are ready to give an upper bound for $d(v)$ if $v \in \mathcal{B}_1$. We get

$$\begin{aligned} d(v) &= |\{v' \in \mathcal{B} :\ vv' = f(a) \text{ with } a \in E(f(x),n)\}| \\ &\le |\{a \in E(f(x),n) :\ f(a) \equiv 0 \pmod v\}| \\ &\le |\{1 \le a \le n :\ f(a) \equiv 0 \pmod v\}|. \end{aligned}$$

Since $v \in \mathcal{B}_1$, we have $v > c_7 n$. Let $c_9 = \lceil \frac{1}{c_7} \rceil$. Then

$$\begin{aligned} d(v) &\le \left| \{1 \le a \le \frac{1}{c_7} v :\ f(a) \equiv 0 \pmod v\} \right| \\ &\le |\{1 \le a \le c_9 v :\ f(a) \equiv 0 \pmod v\}| \\ &\le c_9 \, |\{1 \le a \le v :\ f(a) \equiv 0 \pmod v\}| \\ &= c_9 N(f(x), v). \end{aligned}$$

By (29) we have

$$\begin{aligned} d(v) &\le c_9 c_8 r^{\omega(v)} = c_{10} r^{\omega(v)} \\ & c_{10} \left(2^{\omega(v)}\right)^{\log r/\log 2} = c_{10} \tau(v)^{\log r/\log 2}. \end{aligned}$$

By (27) we have

$$d(v) < c_{10} \left(c_3 (\log n)^s\right)^{\log r/\log 2} = c_{11} (\log n)^{s \log r/\log 2},$$

which completes the proof of Lemma 4 and thus also of Theorem 2. □

2.3. Proof of Theorem 3

Throughout the proof we may suppose that u is large enough, depending on the polynomial $f(x)$. Let $f(x)$ be a polynomial of the form

$$f(x) = g_r x^r + g_{r-1} x^{r-1} + \cdots + g_1 x + g_0.$$

Define β by $\beta \stackrel{\text{def}}{=} \frac{g_{r-1}}{rg_r}$ and the polynomial $p(x)$ by

$$\begin{aligned} p(x) &\stackrel{\text{def}}{=} f(x-\beta) \\ &= g_r\left(x-\frac{g_{r-1}}{rg_r}\right)^r + g_{r-1}\left(x-\frac{g_{r-1}}{rg_r}\right)^{r-1} + \dots g_1\left(x-\frac{g_{r-1}}{rg_r}\right) + g_0. \end{aligned}$$

Clearly, $p(x)$ can be written in the form

$$p(x) = q_r x^r + q_m x^m + q_{m-1}x^{m-1} + q_{m-2}x^{m-2} + \cdots + q_1 x + q_0, \tag{30}$$

where $q_m \neq 0$ and $m \leq r-2$ (in other words, the coefficients of $x^{r-1}, x^{r-2}, \dots, x^{m+1}$ in $p(x)$ are 0). We remark that if $g_{r-1} = 0$, then $f(x) = p(x)$.

Let $\mathcal{B} = \{b_1, b_2, \dots, b_t\}$ be a multiplicative basis of $\mathcal{W}$ of order 2. We will use the following lemma.

Lemma 5. *There exist constants c_1 and $c_2 > 1$ depending only on the polynomial $f(x)$ $(= p(x-\beta))$ such that if b_1, b_2, b_3, b_4 are integers greater than c_1 for which*

$$\begin{aligned} b_1 b_3 &= f(x_1) = p(x_1 - \beta) \\ b_1 b_4 &= f(x_2) = p(x_2 - \beta) \\ b_2 b_3 &= f(x_3) = p(x_3 - \beta) \\ b_2 b_4 &= f(x_4) = p(x_4 - \beta) \end{aligned}$$

hold for some integers x_1, x_2, x_3, x_4, then

$$\begin{aligned} c_2 b_1 b_3 < b_2 b_4 \text{ if } m = r-2 \text{ in (30) and} \\ c_2 \left(b_1 b_3\right)^2 < b_2 b_4 \text{ if } m \leq r-3 \text{ in (30).} \end{aligned}$$

Proof of Lemma 5. This is a combination of Lemma 1 and Lemma 2 in [17].

Next, we define a graph $\mathcal{G}$ by the following: Its vertices are the elements of $\mathcal{B}$, so that $V(\mathcal{G}) = \mathcal{B}$. There is an edge between the vertices $b_1 \in \mathcal{B}$ and $b_2 \in \mathcal{B}$ if and only if there exists an $1 \leq j \leq k$ such that

$$b_1 b_2 = f(a_j) = p(a_j - \beta).$$

We will denote this edge by $\{b_1, b_2\}$.

Now $\mathcal{B}$ is a multiplicative basis of order 2 of $\mathcal{W}$, thus the number $|E(\mathcal{G})|$ of the edges of $\mathcal{G}$ satisfies

$$|E(\mathcal{G})| \geq |\mathcal{W}|. \tag{31}$$

Next, we will use the constants c_1 and c_2 defined in Lemma 5. We will color the edges of $\mathcal{G}$ by different colors. We color an edge $\{b_1, b_2\}$ of $\mathcal{G}$ by the first color

if $b_1 \le c_1$ or $b_2 \le c_1$. Clearly, the number of edges colored by the first color is $\le 2c_1 |\mathcal{B}|$. For $i \ge 2$ we color the edge $\{b_1, b_2\}$ of $\mathcal{G}$ by the i-th color if

$$\frac{1}{2} g_r c_2^{i-2} u^r \le b_1 b_2 < \frac{1}{2} g_r c_2^{i-2} u^r. \tag{32}$$

Here $b_1 b_2 = f(a_j)$ for some $1 \le j \le k$. Since the leading coefficient of $f(x)$ is positive ($g_r > 0$), and using (4), we have

$$\frac{1}{2} g_r u^r \le f(a_1), \dots, f(a_k) \le 2g_r (2u)^r$$

if u is large enough depending on the polynomial $f(x)$. Thus we have

$$\frac{1}{2} g_r u^r \le b_1 b_2 \le 2^{r+1} g_r u^r.$$

Using this and (32), we see that the number of different colors is less than a constant c_4 depending on the polynomial $f(x)$.

By Lemma 5 the graph $\mathcal{G}$ does not contain a cycle of length 4 such that the edges of the cycle are colored by the same i-th color for an $i \ge 2$. By the Kövári–Sós–Turán theorem [18], if a graph $\mathcal{G}$ has n vertices and does not contain a cycle of length 4, then it has at most

$$1 + n + \left[\frac{1}{2} n^{3/2}\right] \tag{33}$$

edges. (Here we remark that in [18] the authors studied matrices containing 0's and 1's and not graphs, but considering the adjacency matrix of $\mathcal{G}$ one may get the upper bound in (33).) We have at most c_4 different colors, thus

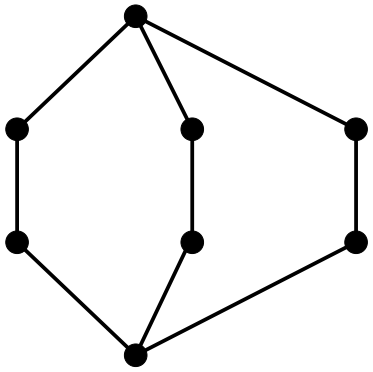
$$|E(\mathcal{G})| \ll |V(\mathcal{G})|^{3/2} = |\mathcal{B}|^{3/2}$$

where the implied constant depends on the polynomial $f(x)$. Using (31) we get

$$|\mathcal{W}| \ll |\mathcal{B}|^{3/2},$$

from which the theorem follows.

Probably, it can be proved that if m in (30) is significantly smaller than the degree r of the polynomial, then the subgraphs $\mathcal{G}_i$ of $\mathcal{G}$ formed by the edges of $\mathcal{G}$ colored by the i-th color (where $i \ge 2$) do not contain the following graph $\theta_{3,3}$:



From this, using the theorem of Faudree and Simonovits [11] in extremal graph theory, one may obtain the bound

$$|\mathcal{W}| \le \sum_i E(\mathcal{G}_i) \ll c_1 |\mathcal{B}| + \sum_{i\ge 2} |V(\mathcal{G}_i)|^{1+1/3} \ll |\mathcal{B}|^{4/3},$$

from which

$$|\mathcal{B}| \gg |\mathcal{W}|^{3/4} \tag{34}$$

follows. However, the proof of that these subgraphs of $\mathcal{G}$ do not contain $\theta_{3,3}$ would be rather lengthly and complicated, and the desired lower bound (34) is just slightly stronger than the one in Theorem 3, and it is also far from the truth. Thus we do not present the details of the proof here.

Acknowledgments. The authors would like to thank the referee for his careful reading and valuable comments.

References

[1] A. Baker and H. Davenport, The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$, *Quart J. Math. Oxford* (2) **20** (1969), 129–137.

[2] T. H. Chan, On sets of integers, none of which divides the product of k others, *Eur. J. Comb.* (3) **32** (2011), 443–447.

[3] F. Delmer, Probléme de diviseurs, *Séminaire de théorie des nombres de Bordeaux 1970–1971*, Talk no. 22, pp. 1–20.

[4] L. E. Dickson, *History of the Theory of Numbers* Vol. 2, Chelsea, New York 1966, pp. 514–519.

[5] Diophantus of Alexandria, *Arithmetics and the Book of Polygonal Numbers*, (I. G. Bashmakova, Ed.), Nauka, Moscow, 1974 (Russian), pp. 103–104, 232.

[6] A. Dujella, An absolute bound for the size of Diophantine m-tuples, *J. Number Theory* **89** (2001), 126–150.

[7] A. Dujella, There are only finitely many Diophantine quintuples, *J. Reine Angew. Math.* **566** (2004), 183–214.

[8] A. Dujella and A. Pethő, A generalization of a theorem of Baker and Davenport, *Quart. J. Math. Oxford Ser. (2)* **49** (1998), 291–306.

[9] P. Erdős, On sequences of integers no one of which divides the product of two others and some related problems, *Mitt. Forsch.-Inst. Math. Mech. Univ. Tomsk* **2** (1938), 74–82.

[10] P. Erdős, On the sum $\sum_{k=1}^{x} d(f(k))$, *J. London Math. Soc.* **27** (1952), 7–15.

[11] R. J. Faudree and M. Simonovits, On a class of degenerate extremal graph problems, *Combinatorica* **3** (1) (1983), 83–93.

[12] L. Hajdu and A. Sárközy, On multiplicative decompositions of polynomial sequences, I, *Acta Arith.* **184** (2018), 139–150.

[13] L. Hajdu and A. Sárközy, On multiplicative decompositions of polynomial sequences, II, *Acta Arith.* **186** (2018), 191–200.

[14] L. Hajdu and A. Sárközy, On multiplicative decompositions of polynomial sequences, III, *Acta Arith.* **193** (2020), 193-216.

[15] Bo He, A. Togbé and V. Ziegler, There is no Diophantine quintuple, *Trans. Am. Math. Soc.* **371** (2019), 6665–6709.

[16] M. N. Huxley, A note on polynomial congruence, in: *Recent Progress in Analytic Number Theory, Volume I*, eds.: H. Halberstam, C. Hooley, pp. 193–196, 1981, London, Academic Press.

[17] K. Gyarmati, A polynomial extension of a problem of Diophantus, *Publ. Math. Debrecen* **66** (2005), 389–405.

[18] T. Kövári, V. T. Sós, P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.* **3** (1954) 50–57.

[19] K. Lapkova, Explicit upper bound for an average number of divisors of irreducible quadratic polynomials, *Monatsh. Math.* **186** (4) (2018), 663-673.

[20] T. Nagel, Généralisation d'un théorème de Tchebicheff, *Journ. de Math.* **8** (1921) 343–356.

[21] O. Ore, Anzahl der Wurseln höherer Kongruenzen, *Norsk Matematisk Tidsskrift (3 Aagang)*, (1921), 63–66.

[22] G. Sándor, Uber die Anzahl der Lösungen einer Kongruenz, *Acta Math.* **87** (1952), 13–17.

[23] C. L. Stewart, On the number of solutions of polynomial congruences, *C. R. Math. Rep. Acad. Sci. Canada* **13** (6) (1991), 271–273.

[24] P. P. Pach, C. Sándor, Multiplicative bases and an Erdős Problem, *Combinatorica*, **38** (2018), 1175–1203.