A basis of finite and infinite sets with small representation function

Artūras Dubickas

Department of Mathematics and Informatics, Vilnius University,
Naugarduko 24, Vilnius LT-03225, Lithuania

arturas.dubickas@mif.vu.lt

Submitted: Nov 7, 2011; Accepted: Dec 19, 2011; Published: Jan 6, 2012
Mathematics Subject Classifications: 11B13, 11B34, 11B83, 11R06

Abstract

Let A be a subset of the set of nonnegative integers $\mathbb{N}\cup\{0\}$, and let $r_A(n)$ be the number of representations of $n \geqslant 0$ by the sum $a+b$ with $a,b \in A$. Then $\left(\sum_{a\in A} x^a\right)^2 = \sum_{n=0}^{\infty} r_A(n)x^n$. We show that an old result of Erdős asserting that there is a basis A of $\mathbb{N}\cup\{0\}$, i.e., $r_A(n) \geqslant 1$ for $n \geqslant 0$, whose representation function $r_A(n)$ satisfies $r_A(n) < (2e+\varepsilon)\log n$ for each sufficiently large integer n. Towards a polynomial version of the Erdős-Turán conjecture we prove that for each $\varepsilon > 0$ and each sufficiently large integer n there is a set $A \subseteq \{0,1,\dots,n\}$ such that the square of the corresponding Newman polynomial $f(x) := \sum_{a\in A} x^a$ of degree n has all of its $2n+1$ coefficients in the interval $[1,(1+\varepsilon)(4/\pi)(\log n)^2]$. Finally, it is shown that the correct order of growth for $H(f^2)$ of those reciprocal Newman polynomials f of degree n whose squares f^2 have all their $2n+1$ coefficients positive is $\sqrt{n}$. More precisely, if the Newman polynomial $f(x) = \sum_{a\in A} x^a$ of degree n is reciprocal, i.e., $A = n - A$, then $A + A = \{0,1,\dots,2n\}$ implies that the coefficient for x^n in $f(x)^2$ is at least $2\sqrt{n}-3$. In the opposite direction, we explicitly construct a reciprocal Newman polynomial $f(x)$ of degree n such that the coefficients of its square $f(x)^2$ all belong to the interval $[1, 2\sqrt{2n}+4]$.

1 Introduction

Throughout the paper, let

$$\mathbb{N} := \{1,2,3,\dots\},$$

and let A be a finite or infinite subset of the set $\mathbb{N}\cup\{0\}$. The square of the power series

$$f_A(x) = \sum_{a\in A} x^a$$

associated with A is given by the formulae

$$f_A(x)^2 = \sum_{n=0}^{\infty} r_A(n)x^n,$$

where $r_A(n)$ stands for the number of representations of the integer $n \geqslant 0$ by the sum $a+b$ with $a, b \in A$, namely,

$$r_A(n) := |\{(a,b) \in A^2 \ : \ a+b=n\}|.$$

One of the unsolved conjectures of Erdős and Turán [9] (which is a 500 USD problem in [7]) asserts that $\limsup_{n\to\infty} r_A(n) = \infty$ in case when A is an asymptotic basis of $\mathbb{N}$, i.e., $r_A(n) \geqslant 1$ for each sufficiently large integer n. In general, $A \subseteq \mathbb{N} \cup \{0\}$ is called a *basis* of $B \subseteq \mathbb{N} \cup \{0\}$ if every element of B belongs to the sumset

$$A + A = \{a + a' \ : \ a, a' \in A\},$$

i.e., $r_A(n) \geqslant 1$ for each $n \in B$. It is known that for any $A \subseteq \mathbb{N} \cup \{0\}$ the values of $r_A(n)$, where $n \geqslant 0$, cannot all lie in the interval $[1,5]$ (see [11]), and in $[1,7]$ (see [2]). By an entirely different method, Sándor [17] showed that the values of $r_A(n)$, where n runs through all sufficiently large integers, cannot all lie in the interval $[u,v]$, where $u > (\sqrt{v}-1)^2$. See also [1], [12], [13] for some further work on the Erdős-Turán conjecture.

In the opposite direction, Erdős in [6] answered a question of Sidon and showed that there exists a basis A of $\mathbb{N} \cup \{0\}$ such that

$$r_A(n) \leqslant c_1 \log n \tag{1}$$

for some positive constant c_1 and each $n \geqslant 2$. Representations by the sums of k terms have been considered in [8]. In [16] Ruzsa proved that there is a basis A of $\mathbb{N} \cup \{0\}$ whose representation function $r_A(n)$ is bounded on average, namely,

$$\frac{1}{n} \sum_{k=0}^{n-1} r_A(k)^2 \leqslant c_2 \tag{2}$$

for each $n \geqslant 1$. Recently, Tang [18] showed that there is an asymptotic basis A of $\mathbb{N} \cup \{0\}$ for which (2) holds with the constant $c_2 = 1449757928$ for each sufficiently large n. He then refined his construction based on an earlier paper [20] and derived the same result with the smaller constant 1069693154 (see [19]). Finally, during the Paul Turán memorial conference in Budapest Yong-Gao Chen and Quan-Hui Yang (see `http://www.renyi.hu/~turan100/abstracts.pdf`) announced that there is a basis A of $\mathbb{N} \cup \{0\}$ for which (2) holds with the constant 3000 for each $n \geqslant 1$.

The original paper of Erdős [6] is based on some combinatorial construction with probabilistic flavor. From there one can get some explicit but quite large constant c_1 in (1). Our first theorem gives a small constant $2e = 5.4365\ldots$:

Theorem 1 *For each ε satisfying $0 < \varepsilon < 1/2$ there is a positive constant $c(\varepsilon)$ and a basis A of $\mathbb{N} \cup \{0\}$ such that*

$$0.1\varepsilon^2 \log n \leqslant r_A(n) \leqslant (2e + \varepsilon) \log n + c(\varepsilon) \tag{3}$$

for every $n \geqslant 2$.

In [4] the author raised a polynomial version of the Erdős-Turán problem. Suppose that $f(x)$ is a polynomial of degree n with coefficients in $\{0, 1\}$ (often called a *Newman* polynomial after [15]) such that $f(x)^2$ has positive coefficients for x^j, $j = 0, 1, \ldots, 2n$. *What is the smallest possible maximal coefficient of $f(x)^2$? Is it bounded or unbounded in terms of n?* Equivalently, we ask for the smallest possible value of

$$\max_{0 \leqslant k \leqslant 2n} r_A(k),$$

where $A \subseteq \{0, 1, \ldots, n\}$ satisfies $A + A = \{0, 1, \ldots, 2n\}$.

Exactly the same question, although without interpretation in terms of sets and sumsets, can be asked for the polynomial $f(x)$ of degree n with nonnegative coefficients. Under additional assumption of f being a *reciprocal* polynomial, namely, $f(x) = x^n f(1/x)$, it was proved in [5] that if the coefficients of $f(x)^2$ are all at least 1 then the largest coefficient of $f(x)^2$ must be at least $\kappa_{\rm rec}(n)$, where $\kappa_{\rm rec}(n) \sim \frac{2}{\pi} \log n$ as $n \to \infty$. The extremal reciprocal polynomial with nonnegative coefficients was found explicitly in [5]:

$$\sum_{k=0}^{\lfloor n/2 \rfloor} 2^{-2k} \binom{2k}{k} x^k + \sum_{k=0}^{n-\lfloor n/2 \rfloor - 1} 2^{-2k} \binom{2k}{k} x^{n-k}. \tag{4}$$

In fact, the first $\lfloor n/2 \rfloor + 1$ and the last $\lfloor n/2 \rfloor + 1$ coefficients of its square are all equal to 1 (see [5]). We conjectured in [5] that the extremal polynomial (with nonnegative coefficients) in the general case should be the same reciprocal polynomial (4). However, there are no results in this direction so far (neither for general polynomials with real nonnegative coefficients nor for Newman polynomials). Below, we shall give three results of this type for Newman polynomials.

For a general Newman polynomial we prove that

Theorem 2 *For each $\varepsilon > 0$ and each integer $n \geqslant n_0(\varepsilon)$ there is Newman polynomial of degree n whose square has all of its coefficients in the interval $[1, (1 + \varepsilon)(4/\pi)(\log n)^2]$.*

In terms of sumsets Theorem 2 asserts that for each $\varepsilon > 0$ and each sufficiently large n there is subset A of the set $\{0, 1, \ldots, n\}$ such that $A + A = \{0, 1, \ldots, 2n\}$ and the number of representations of each given $k \in \{0, 1, \ldots, 2n\}$ by the sum $a + a'$, with $a, a' \in A$, is at most $(1 + \varepsilon)(4/\pi)(\log n)^2$, i.e.,

$$1 \leqslant r_A(k) \leqslant (1 + \varepsilon)(4/\pi)(\log n)^2$$

for every $k = 0, 1, \ldots, 2n$. We remark that under a slightly weaker assumption

$$\{0, 1, \ldots, \lfloor (2 - \varepsilon)n \rfloor\} \subseteq A + A$$

Theorem 1 gives a stronger bound with $(\log n)^2$ replaced by $\log n$:

Corollary 3 *For each $\varepsilon > 0$ there is a positive constant $C = C(\varepsilon)$ such that for every integer $n \geqslant 2$ there is a set $A \subseteq \{0, 1, \ldots, n\}$ for which the sumset $A + A$ contains the set $\{0, 1, \ldots, \lfloor (2-\varepsilon)n \rfloor\}$ and*

$$r_A(k) \leqslant C \log n$$

for every $k = 0, 1, \ldots, 2n$.

Finally, the next theorem asserts that for a reciprocal Newman polynomial the correct growth is of the order $\sqrt{n}$:

Theorem 4 *For each reciprocal Newman polynomial $f(x)$ of degree n whose square has all of its $2n+1$ coefficients at least 1, the middle coefficient for x^n in $f(x)^2$ must be at least $2\sqrt{n}-3$. On the other hand, for each $n \in \mathbb{N}$ there is a reciprocal Newman polynomial of degree n such that the coefficients of its square are all in the interval $[1, 2\sqrt{2n}+4]$.*

The first part of Theorem 4 will be derived by a simple counting argument, while to prove the second part we shall use the following explicit example

$$\sum_{i=0}^{t-1}(x^i + x^{n-i}) + \sum_{j=1}^{s}(x^{jt} + x^{n-jt}) + \delta(x), \tag{5}$$

where

$$t := \lfloor \sqrt{n/2} \rfloor, \quad s := \lceil n/2t \rceil - 1,$$

$$\delta(x) := \begin{cases} 0, & \text{if } \{n/2t\} \leqslant 1/2, \\ x^{n/2}, & \text{if } \{n/2t\} > 1/2 \text{ and } n \text{ is even}, \\ x^{(n-1)/2} + x^{(n+1)/2}, & \text{if } \{n/2t\} > 1/2 \text{ and } n \text{ is odd}. \end{cases}$$

The constants -3 and 4 in Theorem 4 can be easily improved. However, we do not know for which constant in the interval $[2, 2\sqrt{2}]$ both parts of Theorem 4 hold, so we ask for the best possible constant κ for $\sqrt{n}$ in the sense that for each $\varepsilon > 0$ and each sufficiently large $n \in \mathbb{N}$ the first statement of Theorem 4 holds with $(\kappa - \varepsilon)\sqrt{n}$ instead of $2\sqrt{n} - 3$ while the second holds with $(\kappa + \varepsilon)\sqrt{n}$ instead of $2\sqrt{2n} + 4$.

In the next section we shall prove Theorem 4. Its proof is independent from the other parts of the paper. Sections 3 and 4 contain probabilistic and analytic preparation for the proofs of Theorems 1 and 2. Their proofs will be completed in Sections 5 and 6, respectively. In Section 5 we shall also prove Corollary 3.

2 Proof of Theorem 4

Let $f(x)$ be a Newman polynomial of degree n whose square $f(x)^2$ has its coefficients (from the constant coefficient to the leading coefficient) at least 1. Let A be the subset of $B := \{0, 1, \ldots, \lfloor n/2 \rfloor\}$ consisting of those indices j whose coefficients for x^j in $f(x)$ are equal to 1. Evidently, the sumset $A + A$ must contain the set B. Since $A + A$ has at most

$$|A|(|A| - 1)/2 + |A| = |A|(|A| + 1|)/2$$

distinct elements, we must have

$$n/2 < \lfloor n/2 \rfloor + 1 = |B| \leqslant |A|(|A|+1)/2 < (|A|+1)^2/2.$$

Hence

$$|A| > \sqrt{n} - 1. \tag{6}$$

On the other hand, since $f(x)$ is reciprocal, the middle coefficient of the polynomial $f(x)^2$ equals $2|A|-1$ if n is even and $n/2 \in A$, and equals $2|A|$ otherwise. Using (6) we deduce that the coefficient for x^n in $f(x)^2$ is at least

$$2|A| - 1 = 2(|A|+1) - 3 > 2\sqrt{n} - 3,$$

as claimed.

Next, we consider the polynomial $f(x)$ given in (5) for $n \geqslant 18$. Observe first that

$$st = t\lceil n/2t - 1 \rceil < t(n/2t) = n/2$$

and

$$n/2 - st \leqslant n/2 - (n/2t - 1)t = t.$$

The inequality $\{n/2t\} > 1/2$ is equivalent to $\lceil n/2t \rceil < 1/2 + n/2t$ which is equivalent to $n - st > st + t$, i.e., the gap between st and $n - st$ is greater than t. So if $\{n/2t\} \leqslant 1/2$ then the gap between st and $n - st$ is at most t. Also, $t = \lfloor \sqrt{n/2} \rfloor \geqslant 3$ provided that $n \geqslant 18$. It follows that the terms (zero, one or two) of the polynomial $\delta(x)$ are between x^{st} and x^{n-st}. Consequently, f is a reciprocal Newman polynomial for each $n \geqslant 18$.

Since $f(x)^2$ is reciprocal, to prove that all the coefficients of $f(x)^2$ are at least 1 it suffices to show that its coefficients for x^j, where $j = 0, 1, \ldots, n$, are nonzero. This time, let A be the subset of $\{0, 1, \ldots, n\}$ consisting of those indices j whose coefficients for x^j in $f(x)$ are equal to 1. By (5), we see that

$$A = \{0, 1, \ldots, t, 2t, \ldots, st, u_n, v_n, n - st, n - (s-1)t, \ldots, n - t, n - t + 1, \ldots, n\}.$$

Here, $u_n = v_n = 0$ if $n - st \leqslant st + t$. If $n - st > st + t$, then $u_n = v_n = n/2$ for n even and $u_n = (n-1)/2$, $v_n = (n+1)/2$ for n odd. The gaps between consecutive elements $t, 2t, \ldots, st, u_n, v_n, n-st, \ldots, n-2t, n-t$ of A are at most t. (Here, u_n, v_n are only present if $n - st > st + t$.) Thus each integer $k \in [0, n]$ belongs to the sumset $A + A$. As $f(x)^2$ is reciprocal, it follows that all of the coefficients of $f(x)^2$ are at least 1.

We next show that the coefficients of $f(x)^2$ are at most $2\sqrt{2n} + 5$ for each $n \geqslant 18$. Since $f(x)^2$ is reciprocal, it suffices to prove this for the coefficients a_m of x^m, where $0 \leqslant m \leqslant n$. Clearly, $a_0 = 1$, $a_1 = 2$ and for $m \geqslant 2$

$$a_m = 2|A \cap \{0, 1, \ldots, \lceil m/2 \rceil - 1\}| + \delta_{m/2},$$

where $\delta_{m/2} = 1$ for m even and $m/2 \in A$ and $\delta_{m/2} = 0$ otherwise. Hence for each $m = 2, \ldots, n$ we have

$$a_m = 2|A \cap \{0, 1, \ldots, \lceil m/2 \rceil - 1\}| + \delta_{m/2} \leqslant a_n = 2|A \cap \{0, 1, \ldots, \lceil n/2 \rceil - 1\}| + \delta_{n/2}.$$

It is easy to see that $a_n = 2(s+t)$ if $u_n = v_n = 0$, $a_n = 2(s+t)+1$ if $u_n = n/2$, and $a_n = 2(s+t+1)$ if $u_n = (n-1)/2$. In the third case, in view of $s = \lceil n/2t - 1 \rceil < n/2t - 1/2$ we find that

$$a_n = 2(s+t+1) = 2t + 2\lceil n/2t - 1 \rceil + 2 < 2t + 2(n/2t - 1/2) + 2 = 2t + n/t + 1.$$

In the first two cases, using $s = \lceil n/2t - 1 \rceil < n/2t$ we obtain the same bound, because then

$$a_n \leqslant 2(s+t) + 1 = 2t + 2\lceil n/2t - 1 \rceil + 1 < 2t + 2(n/2t) + 1 = 2t + n/t + 1.$$

Using $t = \lfloor \sqrt{n/2} \rfloor$, we see that $2t \leqslant 2\sqrt{n/2} = \sqrt{2n}$ and $t > \sqrt{n/2} - 1$. Hence

$$\frac{n}{t} < \frac{n}{\sqrt{n/2} - 1} \leqslant \sqrt{2n} + 3$$

for $n \geqslant 18$. Consequently, for $m = 2, \ldots, n$

$$a_m \leqslant a_n \leqslant 2t + n/t + 1 < \sqrt{2n} + \sqrt{2n} + 3 + 1 = 2\sqrt{2n} + 4.$$

This proves the second part of Theorem 4 for $n \geqslant 18$.

To complete the proof observe that for $n = 1$ and $n = 2$ one can take the reciprocal Newman polynomial $1 + x$ and $1 + x + x^2$, respectively. For $3 \leqslant n \leqslant 17$ we may consider the reciprocal Newman polynomial $f(x) = \sum_{a \in A} x^a$, where

$$A := \{0, 1, 2, \ldots \lceil n/3 \rceil\} \cup \{n - \lceil n/3 \rceil, \ldots, n-1, n\}.$$

Then the coefficients of $f(x)^2$ are all at least 1 while the largest coefficient of $f(x)^2$ equals $2\lceil n/3 \rceil + 2$. One can easily verify that this is less than $2\sqrt{2n} + 4$ in the range $3 \leqslant n \leqslant 17$.

3 A bit of probability theory

Let $X_1, \ldots, X_s$ be s independent Bernoulli trials, where

$$\mathbb{P}(X_i = 1) = p_i \in [0, 1] \quad \text{and} \quad \mathbb{P}(X_i = 0) = 1 - p_i$$

for $i = 1, \ldots, s$. Set $X := X_1 + \cdots + X_s$ and

$$\mathbb{E}(X) := \sum_{i=1}^{s} p_i$$

for the expectation of the random variable X. Then Chernoff's inequality (named after [3], see, e.g., [14]) asserts that

Lemma 5 *For any $\delta > 0$ we have*

$$\mathbb{P}(X > (1+\delta)\mathbb{E}(X)) \leqslant e^{-((1+\delta)\log(1+\delta) - \delta)\mathbb{E}(X)} \tag{7}$$

and

$$\mathbb{P}(X < (1-\delta)\mathbb{E}(X)) \leqslant e^{-(\delta + (1-\delta)\log(1-\delta))\mathbb{E}(X)}. \tag{8}$$

Since $(1+\delta)\log(1+\delta)-\delta > \delta^2/3$ and $\delta+(1-\delta)\log(1-\delta) > \delta^2/2$ for $0<\delta<1$, inequalities (7) and (8) imply the following symmetric form of Lemma 5

$$\mathbb{P}(|X-\mathbb{E}(X)| > \delta\mathbb{E}(X)) \leqslant e^{-\delta^2\mathbb{E}(X)/2} + e^{-\delta^2\mathbb{E}(X)/3} \leqslant 2e^{-\delta^2\mathbb{E}(X)/3} \tag{9}$$

for every δ satisfying $0<\delta<1$.

For the proof of Theorems 1 and 2 we define mutually independent random variables Y_k and Y_k^* taking only values 0 and 1, by

$$\mathbb{P}(Y_0=1)=\mathbb{P}(Y_0^*=1)=\mathbb{P}(Y_1=1)=\mathbb{P}(Y_1^*=1)=\mathbb{P}(Y_2=1)=\mathbb{P}(Y_2^*=1)=1$$

and

$$\mathbb{P}(Y_k=1)=\mathbb{P}(Y_k^*=1)=p_k := \lambda\sqrt{\frac{2\log k}{\pi k}} \tag{10}$$

for each integer $k \geqslant 3$. Here, λ will be chosen in the interval

$$1<\lambda<2, \tag{11}$$

so that $0<p_k \leqslant p_3 < 2\sqrt{\frac{2\log 3}{3\pi}} < 0.97 < 1$ for $k \geqslant 3$, by (10) and (11). For convenience, we shall also use the notation $p_0=p_1=p_2=1$, so, by (10),

$$\mathbb{P}(Y_k=0)=\mathbb{P}(Y_k^*=0)=1-p_k$$

for every nonnegative integer k, and

$$p_0=p_1=p_2>p_3>p_4>p_5>p_6>\dots. \tag{12}$$

To prove Theorem 1 we consider the random series

$$f(x) := \sum_{k=0}^{\infty} Y_k x^k \tag{13}$$

with coefficients $0, 1$. The square of $f(x)$ is given by

$$f(x)^2 = \sum_{m=0}^{\infty} Z_m x^m, \tag{14}$$

where

$$Z_m := 2\sum_{0\leqslant k<m/2} Y_k Y_{m-k} + Y_{m/2} \tag{15}$$

and throughout the convention $Y_{m/2}=p_{m/2}=0$ is adopted if m is odd. In the sum

$$V_m := \sum_{0\leqslant k<m/2} Y_k Y_{m-k} \tag{16}$$

the summands Y_kY_{m-k}, where $0 \leqslant k < m/2$, are mutually independent random variables taking only values 0 and 1, so that Lemma 5 is applicable to $X = V_m$. By (10), we have

$$\mathbb{P}(Y_kY_{m-k} = 1) = \mathbb{P}(Y_k = 1)\mathbb{P}(Y_{m-k} = 1) = p_kp_{m-k}.$$

Thus the expectation of V_m is

$$\mathbb{E}(V_m) = S_m := \sum_{0 \leqslant k < m/2} p_kp_{m-k}. \tag{17}$$

In a similar fashion in the proof of Theorem 2 we will consider the random Newman polynomial

$$f(x) := \sum_{k=0}^{n} U_kx^k = \sum_{0 \leqslant k < n/2} (Y_kx^k + Y_k^*x^{n-k}) + Y_{n/2}x^{n/2}, \tag{18}$$

where $Y_{m/2} = Y_{m/2}^* = p_{m/2} = 0$ if m is odd. Note that $U_k = Y_k$ for $k \leqslant n/2$ and $U_k = Y_{n-k}^*$ for $n/2 < k \leqslant n$. The square of f is given by

$$f(x)^2 = \sum_{m=0}^{2n} Z_mx^m, \tag{19}$$

where

$$Z_m := 2 \sum_{0 \leqslant k < m/2} U_kU_{m-k} + U_{m/2} \tag{20}$$

for $m \leqslant n$. By symmetry (see (10), (18) and (19)), for each interval $I \subseteq \mathbb{R}$ we must have

$$\mathbb{P}(Z_m \in I) = \mathbb{P}(Z_{2n-m} \in I) \tag{21}$$

for $n < m \leqslant 2n$.

In the sum

$$V_m := \sum_{0 \leqslant k < m/2} U_kU_{m-k} \tag{22}$$

U_kU_{m-k}, where $0 \leqslant k < m/2$, are mutually independent random variables taking only values 0 and 1, so we will be able to apply Lemma 5 to V_m. This time, for $k \leqslant n/2$ we have $U_{m-k} = Y_{m-k}$ if $m - k \leqslant n/2$ and $U_{m-k} = Y_{n-m+k}^*$ if $m - k > n/2$. Hence, by (10),

$$\mathbb{P}(U_kU_{m-k} = 1) = \mathbb{P}(Y_kU_{m-k} = 1) = \mathbb{P}(Y_k = 1)\mathbb{P}(U_{m-k} = 1) = p_kp_{\min\{m-k,n-m+k\}}.$$

Thus the expectation of V_m for $m \leqslant n$ is

$$\mathbb{E}(V_m) = T_m := \sum_{0 \leqslant k < m/2} p_kp_{\min\{m-k,n-m+k\}}. \tag{23}$$

In addition to this we shall also use the Borel-Cantelli lemma (see, e.g., [10]).

Lemma 6 *Let $E_0, E_1, E_2, \ldots$ be a sequence of events in some probability space. If*

$$\sum_{j=0}^{\infty} \mathbb{P}(E_j) < \infty$$

then the probability of the event consisting in the occurrence of only a finite number out of the events E_j, $j = 0, 1, \ldots$, is equal to 1.

Finally, using the notation $E^c = \Omega \setminus E$ for an event E in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, from $\left(\cup_{k=1}^{\ell} E_k\right) \cup \left(\cap_{k=1}^{\ell} E_k^c\right) = \Omega$ we obtain the next standard estimate

$$\mathbb{P}(\cap_{k=1}^{\ell} E_k^c) = 1 - \mathbb{P}(\cup_{k=1}^{\ell} E_k) \geqslant 1 - \sum_{k=1}^{\ell} \mathbb{P}(E_k). \tag{24}$$

4 ...and analysis

Lemma 7 *For S_m given in (17) and p_k given in (10) we have*

$$S_m \sim \lambda^2 \log m \quad as \quad m \to \infty.$$

Proof: Writing

$$\sqrt{\frac{\log k \log(m-k)}{k(m-k)}} = \frac{1}{m} \sqrt{\frac{(\log m + \log(k/m))(\log m + \log(1-k/m))}{(k/m)(1-k/m)}}$$

for $3 \leqslant k \leqslant m-3$ and replacing the sum by the corresponding Riemann integral in view of (10) we find that

$$S_m = \sum_{0 \leqslant k < m/2} p_k p_{m-k} \sim \frac{2\lambda^2 \log m}{\pi} \int_0^{1/2} \frac{dz}{\sqrt{z(1-z)}}$$

$$= \frac{\lambda^2 \log m}{\pi} \int_0^1 \frac{dz}{\sqrt{z(1-z)}} = \frac{\lambda^2 \log m}{\pi} \frac{\Gamma(1/2)^2}{\Gamma(1)} = \lambda^2 \log m$$

as $m \to \infty$. (Here, we used the values of the Gamma function $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(1) = 1$.) Thus $S_m \sim \lambda^2 \log m$ as $m \to \infty$, as claimed. ■

We next evaluate T_m defined in (23) for $m \leqslant n$. (Recall that the probabilities p_k are defined in (10) and S_m is given in (17).)

Lemma 8 *We have*

$$T_m = S_m \tag{25}$$

for $m \leqslant \lfloor n/2 \rfloor$,

$$S_m \leqslant T_m \leqslant T_n \tag{26}$$

for $\lfloor n/2 \rfloor < m \leqslant n$, and

$$T_n \sim (\lambda^2/\pi)(\log n)^2 \quad as \quad n \to \infty. \tag{27}$$

Proof: Observe that $p_{\min\{m-k,n-m+k\}} = p_{m-k}$ for $m \leqslant \lfloor n/2 \rfloor$. So (17) and (23) yield (25). Assume next that $\lfloor n/2 \rfloor < m \leqslant n$. Then $p_{\min\{m-k,n-m+k\}} = p_{n-m+k}$ for $k < m - n/2$ and $p_{\min\{m-k,n-m+k\}} = p_{m-k}$ for $k \geqslant m - n/2$. It follows that

$$T_m = \sum_{0 \leqslant k < m-n/2} p_k p_{n-m+k} + \sum_{m-n/2 \leqslant k < m/2} p_k p_{m-k}.$$

Note that, by (12), the sequence $p_0, p_1, p_2, p_3, \ldots$ is nonincreasing. By replacing each p_{n-m+k} is the first sum by p_{m-k} and using $p_{n-m+k} \geqslant p_{m-k}$ we obtain

$$T_m \geqslant \sum_{0 \leqslant k < m-n/2} p_k p_{m-k} + \sum_{m-n/2 \leqslant k < m/2} p_k p_{m-k} = \sum_{0 \leqslant k < m/2} p_k p_{m-k} = S_m.$$

Similarly, by replacing each p_{n-m+k} by p_k in the first sum (so that $p_{n-m+k} \leqslant p_k$) of T_m and each p_{m-k} by p_k is the second sum (so that $p_{m-k} \leqslant p_k$) of T_m, we get

$$T_m \leqslant \sum_{0 \leqslant k < m-n/2} p_k^2 + \sum_{m-n/2 \leqslant k < m/2} p_k^2 = \sum_{0 \leqslant k < m/2} p_k^2 \leqslant \sum_{0 \leqslant k < n/2} p_k^2 = T_n.$$

This completes the proof of (26).

To prove (27) observe that, by (10), we have

$$T_n = \sum_{0 \leqslant k < n/2} p_k^2 = 3 + \frac{2\lambda^2}{\pi} \sum_{3 \leqslant k < n/2} \frac{\log k}{k} \sim \frac{\lambda^2}{\pi} (\log n)^2$$

as $n \to \infty$. ■

Recall that if $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial in $\mathbb{R}[x]$ then its *height* and *length* are defined by

$$H(f) := \max_{0 \leqslant j \leqslant n} |a_j| \quad \text{and} \quad L(f) := \sum_{j=0}^{n} |a_j|,$$

respectively. For the infinite series $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots$, we define its height by the formula

$$H(f) := \sup_{j \geqslant 0} |a_j|.$$

Lemma 9 *Let f be a polynomial (or an infinite series), and let g be a polynomial. Then*

$$H((f+g)^2 - f^2) \leqslant (2H(f) + H(g))L(g).$$

Proof: Using $H(fg) \leqslant H(f)L(g)$ and $H(g^2) \leqslant H(g)L(g)$ from the identity $(f+g)^2 - f^2 = 2fg + g^2$ we deduce

$$H((f+g)^2 - f^2) \leqslant 2H(fg) + H(g^2) \leqslant 2H(f)L(g) + H(g)L(g) = (2H(f) + H(g))L(g),$$

as claimed. ■

5 Proof of Theorem 1 and Corollary 3

Given $m \geqslant 3$, let E_m be the event which means that the random variable Z_m defined in (15) belongs to the union of intervals

$$[0, 0.1\varepsilon^2 \log m) \cup ((2e+\varepsilon)\log m, \infty).$$

Then $\cap_{i=M}^{\infty} E_i^c$ is the event consisting of each Z_m, $m = M, M+1, \ldots$, lying in the interval $[0.1\varepsilon^2 \log m, (2e+\varepsilon)\log m]$. Selecting

$$\lambda := \sqrt{1+0.18\varepsilon} \tag{28}$$

in (11) we will show that the probability of the event $\cap_{i=M}^{\infty} E_i^c$ is positive for some M.

By the definition of E_m,

$$\mathbb{P}(E_m) = \mathbb{P}(0 \leqslant Z_m < 0.1\varepsilon^2 \log m) + \mathbb{P}(Z_m > (2e+\varepsilon)\log m). \tag{29}$$

Let us first estimate the probability of the event $Z_m < 0.1\varepsilon^2 \log m$ from above. Observe that, by (15) and (16), we have $Z_m = 2V_m + Y_{m/2}$. So $V_m < 0.05\varepsilon^2 \log m - Y_{m/2}/2$ is the same event. Evidently, this event is contained in the event $V_m < 0.05\varepsilon^2 \log m$. By (11), (17) and Lemma 7, the latter event is contained in the event $V_m < 0.05\varepsilon^2 \mathbb{E}(V_m)$ for each sufficiently large m. Hence, by inequality (8) of Lemma 5 applied to the random variable $X = V_m$ which is the sum of independent Bernoulli trials with $\delta := 1 - \varepsilon^2/20$, in view of the inequality

$$\delta + (1-\delta)\log(1-\delta) = 1 - \varepsilon^2/20 + (\varepsilon^2/20)\log(\varepsilon^2/20) = 1 - (\varepsilon^2/20)\log(20e/\varepsilon^2) > 1 - 0.16\varepsilon$$

which holds for $0 < \varepsilon < 0.66$, we deduce that

$$\mathbb{P}(0 \leqslant Z_m < 0.1\varepsilon^2 \log m) \leqslant \mathbb{P}(V_m < 0.05\varepsilon^2 \mathbb{E}(V_m)) \leqslant e^{-(1-0.16\varepsilon)\mathbb{E}(V_m)}.$$

Note that, by (17), Lemma 7 and (28), for m large enough we must have

$$(1-0.16\varepsilon)\mathbb{E}(V_m) > (1-0.16\varepsilon)(1+0.1799\varepsilon)\log m > (1+\varepsilon/182)\log m,$$

since $0 < \varepsilon < 1/2$ implies $(1-0.16\varepsilon)(1+0.1799\varepsilon) > (1+\varepsilon/182)$. Consequently,

$$\mathbb{P}(0 \leqslant Z_m < \varepsilon \log m) \leqslant m^{-1-\varepsilon/182} < m^{-1-\varepsilon/200} \tag{30}$$

for each sufficiently large m.

We next estimate the probability of the event $Z_m > (2e+\varepsilon)\log m$ from above. Once again, by (15) and (16), we find that $V_m > (e+\varepsilon/2)\log m - Y_{m/2}/2$ is the same event. This is contained in the event $V_m > (e+0.49\varepsilon)\log m$ for m large enough. The latter event is contained in the event $V_m > e\mathbb{E}(V_m)$, since, by (28), we have

$$\lambda^2 = 1 + 0.18\varepsilon < 1 + 0.49e^{-1}\varepsilon = (e+0.49\varepsilon)/e,$$

and so, by (17) and Lemma 7,

$$(e+0.49\varepsilon)\log m > e\mathbb{E}(V_m)$$

for each sufficiently large m. Therefore, selecting $\delta := e-1$ in inequality (7) of Lemma 5 and using $(1+\delta)\log(1+\delta)-\delta = 1$, from Lemma 7 and (28) we obtain

$$\mathbb{P}(Z_m > (2e+\varepsilon)\log m) \leqslant \mathbb{P}(V_m > e\mathbb{E}(V_m)) \leqslant e^{-\mathbb{E}(V_m)} < e^{-\lambda \log m} = m^{-\lambda} < m^{-1-\varepsilon/12}$$

for each sufficiently large m. (Here, we used the inequality $\lambda = \sqrt{1+0.18\varepsilon} > 1+\varepsilon/12$ for $0 < \varepsilon < 1$.)

Combining this upper bound with (29) and (30) we deduce the inequality

$$\mathbb{P}(E_m) \leqslant m^{-1-\varepsilon/200} + m^{-1-\varepsilon/12} \leqslant 2m^{-1-\varepsilon/200}$$

for each $m \geqslant m_0$, and so the series $\sum_{m=m_0}^{\infty} \mathbb{P}(E_m)$ are convergent. In particular, Lemma 6 implies that for some $M = M(\varepsilon)$ the event $\cap_{i=M}^{\infty} E_i^c$ occurs with positive probability. By the definition of E_m, this means that there exist a series $f(x) := \sum_{n=0}^{\infty} a_n x^n$, where $a_n \in \{0,1\}$, $a_0 = a_1 = a_2 = 1$, such that the coefficients b_n of its square $f(x)^2 = \sum_{n=0}^{\infty} b_n x^n$ satisfy

$$0.1\varepsilon^2 \log n \leqslant b_n \leqslant (2e+\varepsilon)\log n$$

for every $n \geqslant M \geqslant 3$.

To complete the proof of the theorem we replace $f(x)$ by the series

$$f_1(x) := \sum_{n=0}^{M-1} x^n + \sum_{n=M}^{\infty} a_n x^n.$$

Note that the coefficients c_n of its square $f_1(x)^2 = \sum_{n=0}^{\infty} c_n x^n$ are all integers, so they are all at least 1, because $c_n \geqslant b_n > 0$ for $n \geqslant M$ and $c_n = n+1 > 0$ for $0 \leqslant n \leqslant M-1$. As $\log n < n+1$ for $2 \leqslant n \leqslant M$, we clearly have $c_n \geqslant \log n > 0.1\varepsilon^2 \log n$ for each $n \geqslant 2$. Since the difference $g(x) = f_1(x) - f(x)$ is a Newman polynomial of length at most $M-3$, by Lemma 9, we obtain

$$0 \leqslant c_n - b_n \leqslant (2H(f)+H(g))L(g) \leqslant 3L(g) \leqslant 3(M-3)$$

for every $n \geqslant 0$. Thus, setting $b := \max\{b_0, b_1, \ldots, b_{M-1}\}$, we find that

$$c_n \leqslant b_n + 3(M-3) \leqslant (2e+\varepsilon)\log n + b + 3(M-3)$$

for each $n \geqslant 0$. This proves (3) and completes the proof of Theorem 1.

To prove Corollary 3 we assume without restriction of generality that $\varepsilon < 1$ and select

$$K := \lceil 2/\varepsilon \rceil, \quad q_0 := \lceil 4/\varepsilon \rceil.$$

To prove the corollary it suffices to show it that holds with some positive constant C for each $n \geqslant Kq_0$. Write $n = Kq + r$ with integers $q \geqslant q_0$, $r \in \{0, 1, \ldots, K-1\}$ and consider

the Newman polynomial $p(x) := a_0 + a_1 x + \cdots + a_{q-1}x^{q-1}$ which is the beginning of the series $f_1(x)$ that are just found in the proof of Theorem 1. It is clear that the polynomial $p(x)^2$ has the coefficients for x^j at least 1 for $j = 0, 1, \ldots, q-1$. Also,

$$H(p^2) \leqslant c \log(2 \deg p) < c \log(2q)$$

for some positive number c and each $q \geqslant 2$.

Consider the Newman polynomial

$$f(x) := p(x)(1 + x^q + x^{2q} + \cdots + x^{(K-1)q})$$

of degree

$$(K-1)q + \deg p \leqslant (K-1)q + q - 1 < Kq \leqslant n.$$

The corresponding set A consists of the indices j, where the coefficients of f are equal to 1. Its square

$$f(x)^2 = p(x)^2(1 + 2x^q + \cdots + Kx^{(K-1)q} + \cdots + 2x^{(2K-3)q} + x^{(2K-2)q})$$

has the coefficients at least 1 for x^j for $j = 0, 1, \ldots, (2K-1)q - 1$. Using $n < K(q+1)$, we obtain

$$(2K-1)q - 1 \geqslant (2-\varepsilon)K(q+1) > (2-\varepsilon)n \geqslant \lfloor (2-\varepsilon)n \rfloor,$$

because, by the choice of K and q_0,

$$(2K-1)q - 1 - (2-\varepsilon)K(q+1) = \varepsilon Kq - 2K - q + \varepsilon K - 1$$

$$= K(\varepsilon q/2 - 2) + q(\varepsilon K/2 - 1) + \varepsilon K - 1 \geqslant 0 + 0 + \varepsilon(2/\varepsilon) - 1 = 1.$$

It follows that $A + A$ contains the set $\{0, 1, \ldots, \lfloor (2-\varepsilon)n \rfloor\}$. Finally, as $K \geqslant 2$, we obtain

$$H(f^2) \leqslant 2KH(p^2) < 2Kc \log(2q) \leqslant 2Kc \log(2n/K) < C \log n$$

with the constant $C := 2Kc$. This proves Corollary 3.

6 Proof of Theorem 2

Fix a positive constant $\varepsilon < 1/80$. We shall split the set $\{2, 3, \ldots, n\}$ into two sets $\mathcal{S}_1$ and $\mathcal{S}_2$ depending on whether for $m \in \{2, 3, \ldots, n\}$ we have $T_m \leqslant \varepsilon^{-3} \log m$ (set $\mathcal{S}_1$) or $T_m > \varepsilon^{-3} \log m$ (set $\mathcal{S}_2$). Both sets are nonempty, because, by Lemmas 7 and 8, for n large enough $\lfloor n/2 \rfloor \in \mathcal{S}_1$ and $n \in \mathcal{S}_2$.

This time, we select

$$\lambda := \sqrt{2}(1 + \varepsilon/3), \tag{31}$$

so that (11) is satisfied. Let E_m, $2 \leqslant m \leqslant n$, be the event that the random variable V_m which is defined in (18), (19), (20) and (22) belongs to the union of intervals

$$[0, (\varepsilon/4)T_m) \cup (T_m, \infty) \quad \text{in case} \quad m \in \mathcal{S}_1,$$

and to the union of intervals

$$[0,(1-\varepsilon/5)T_m)\cup((1+\varepsilon/5)T_m,\infty)\quad\text{in case}\quad m\in\mathcal{S}_2.$$

Suppose first that $m\in\mathcal{S}_1$ and that m is large enough. Then, applying inequality (8) to the sum of Bernoulli trials V_m with $\delta:=1-\varepsilon/4$ and using

$$T_m\geqslant S_m>2(1+2\varepsilon/3)\log m$$

(see Lemma 7, (25), (26), (31)), we obtain

$$\mathbb{P}(V_m<(\varepsilon/4)T_m)\leqslant e^{-(1-\varepsilon/4)^2T_m/2}<e^{-(1-\varepsilon/4)^2(1+2\varepsilon/3)\log m}<m^{-1-\varepsilon/7},$$

because $1+\varepsilon/7<(1-\varepsilon/4)^2(1+2\varepsilon/3)$ for $0<\varepsilon<0.08$. Similarly, applying (7) with $\delta:=2$, we derive that

$$\mathbb{P}(V_m>3T_m)\leqslant e^{-(3\log 3-2)T_m}<e^{-2\log m}=m^{-2},$$

since $2/(3\log 3-2)<2<T_m/\log m$. Hence

$$\mathbb{P}(E_m)=\mathbb{P}(V_m\notin[(\varepsilon/4)T_m,3T_m])\leqslant m^{-1-\varepsilon/7}+m^{-2}\leqslant 2m^{-1-\varepsilon/7}\tag{32}$$

for each sufficiently large $m\in\mathcal{S}_1$.

We next give an upper bound for the event $V_m\notin[(1-\varepsilon/5)T_m,(1+\varepsilon/5)T_m]$ when $m\in\mathcal{S}_2$ is large enough. By (9), (23) and $T_m>\varepsilon^{-3}\log m$, we find that

$$\mathbb{P}(E_m)=\mathbb{P}(|V_m-T_m|>(\varepsilon/5)T_m))=\mathbb{P}(|V_m-\mathbb{E}(V_m)|>(\varepsilon/5)\mathbb{E}(V_m))$$

$$\leqslant 2e^{-\varepsilon^2\mathbb{E}(V_m)/75}=2e^{-\varepsilon^2T_m/75}<2m^{-1/75\varepsilon}.$$

Therefore,

$$\mathbb{P}(E_m)=\mathbb{P}(V_m\notin[(1-\varepsilon/5)T_m,(1+\varepsilon/5)T_m])<2m^{-1/75\varepsilon}<m^{-1-\varepsilon/7}\tag{33}$$

for each sufficiently large $m\in\mathcal{S}_2$, because $1/75\varepsilon>1+\varepsilon/7$ for $0<\varepsilon<1/80$.

Observe that, by (20) and (22), $Z_m-2V_m\in\{Y_{m/2},Y^*_{n-m/2}\}$. So the event E_m, i.e., $V_m\notin[(\varepsilon/4)T_m,3T_m]$ (for $m\in\mathcal{S}_1$) contains the event $Z_m\notin[(\varepsilon/3)T_m,7T_m]$ for each sufficiently large m. Also, as $m\in\mathcal{S}_1$, we have $2\log m<T_m\leqslant\varepsilon^{-3}\log m$, so the latter event contains the event $Z_m\notin[(\varepsilon/2)\log m,\varepsilon^{-4}\log m]$. Using (32) we obtain

$$\mathbb{P}(Z_m\notin[(\varepsilon/2)\log m,\varepsilon^{-4}\log m])\leqslant 2m^{1-\varepsilon/7}\tag{34}$$

for each sufficiently large $m\in\mathcal{S}_1$, $m\leqslant n$.

Similarly, for $m\in\mathcal{S}_2$ the event E_m, i.e., $V_m\notin[(1-\varepsilon/5)T_m,(1+\varepsilon/5)T_m]$ contains the event $Z_m\notin[2(1-\varepsilon/4)T_m,2(1+\varepsilon/4)T_m]$ for each sufficiently large m. Note that

$$2\log m<T_m<(1+7\varepsilon/10)(4/\pi)(\log n)^2,$$

by (26), (27), (31), because $\lambda^2 = 2(1 + 2\varepsilon/3 + \varepsilon^2/9) < 2(1 + 7\varepsilon/10)$ for $0 < \varepsilon < 1/4$. Therefore, from (33) and $(1+7\varepsilon/10)(1+\varepsilon/4) < 1+39\varepsilon/40 < 1+0.98\varepsilon$ (for $0 < \varepsilon < 1/7$) we find that

$$\mathbb{P}(Z_m \notin [3\log m, (1+0.98\varepsilon)(4/\pi)(\log n)^2]) < m^{-1-\varepsilon/7} \tag{35}$$

for each sufficiently large $m \in \mathcal{S}_2$, $m \leqslant n$. By (21), inequalities (34) and (35) also hold for Z_m replaced with Z_{2n-m}. From (34) and (35) we derive that there is a positive integer M_1 such that

$$\mathbb{P}(Z_m \notin [(\varepsilon/2)\log m, (1+0.98\varepsilon)(4/\pi)(\log n)^2]) \leqslant 2m^{-1-\varepsilon/7}$$

for each $m = M_1 + 1, \dots, n$ and the same holds for Z_m replaced with Z_{2n-m}.

Since

$$\sum_{m=M_1+1}^{2n-M_1-1} 2m^{-1-\varepsilon/7} < 2\sum_{m=M_1+1}^{\infty} m^{-1-\varepsilon/7} < 1/2$$

for M_1 large enough, by (24), we conclude that there is an integer $M = M(\varepsilon)$ such that the probability that all the events

$$\varepsilon \log\min(m, 2n-m) \leqslant Z_m \leqslant (1+0.98\varepsilon)(4/\pi)(\log n)^2$$

from $m = M$ to $m = n - M$ hold is positive. Thus there exists a polynomial $f(x) := \sum_{j=0}^{n} a_j x^j$, where $a_j \in \{0,1\}$, $a_0 = a_1 = a_2 = a_{n-2} = a_{n-1} = a_n = 1$, such that the coefficients b_m of its square $f(x)^2 = \sum_{m=0}^{2n} b_m x^m$ satisfy

$$\varepsilon \log m \leqslant b_m, b_{2n-m} \leqslant (1+0.98\varepsilon)(4/\pi)(\log n)^2$$

for $m \geqslant M \geqslant 3$ and $m \leqslant n$. Here, M depends on ε, but it does not depend on n.

As above, to complete the proof of the theorem let us replace the Newman polynomial $f(x)$ by the Newman polynomial

$$f_1(x) := \sum_{m=0}^{M-1}(x^m + x^{2n-m}) + \sum_{m=M}^{n-M} a_m x^m.$$

The coefficients of its square $f_1(x)^2$ are all integers and positive numbers, so they are all at least 1. Since $L(f_1 - f) \leqslant 2(M-3)$, Lemma 9 implies that the largest coefficient of $f_1(x)^2$ does not exceed

$$H(f^2) + 3(2M-6) < (1+0.98\varepsilon)(4/\pi)(\log n)^2 + 6M < (1+\varepsilon)(4/\pi)(\log n)^2$$

for each sufficiently large integer n. This completes the proof of the theorem.

Note that, by (10), (18), (31), it is easy to see that the length of the Newman polynomial f_1 whose existence is just established will be close to

$$2\sum_{k=3}^{\lfloor n/2\rfloor} \lambda\sqrt{\frac{2\log k}{\pi k}} \sim 4(1+\varepsilon/3)\sqrt{\frac{2n\log n}{\pi}}$$

as $n \to \infty$.

References

[1] N. Alon and M.N. Kolountzakis, *On a problem of Erdős and Turán and some related results,* J. Number Theory, **55** (1995), 82–93.

[2] P. Borwein, S. Choi and F. Chu, *An old conjecture of Erdős-Turán on additive basis,* Math. Comp., **75** (2006), 475–484.

[3] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations,* Ann. Math. Stat., **23** (1952), 493–507.

[4] A. Dubickas, *Additive bases of positive integers and related problems,* Uniform Distribution Theory, **3** (2) (2008), 81–90.

[5] A. Dubickas and G. Šemetulskis, *On polynomials with flat squares,* Acta Arith., **146** (2011), 247–255.

[6] P. Erdős, *On a problem of Sidon in additive number theory,* Acta Sci. Math., **15** (1954), 255–259.

[7] P. Erdős, *Some old and new problems on additive and combinatorial number theory,* Combinatorial Mathematics: Proc. of the Third Intern. Conf. (New York, 1985), New York Acad. Sci., New York, 1989, pp. 181–186.

[8] P. Erdős and P. Tetali, *Representations of integers as the sum of k terms,* Random Struct. Algorithms, **1** (1990), 245–261.

[9] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory and some related problems,* J. London Math. Soc., **16** (1941), 212–215.

[10] W. Feller, *An introduction to probability theory and its applications. I,* J. Wiley and Sons, 3rd. ed., New York, London, 1968.

[11] G. Grekos, L. Haddad, C. Helou and J. Pihko, *On the Erdős-Turán conjecture,* J. Number Theory, **102** (2003), 339–352.

[12] M. Helm, *Some remarks on the Erdős-Turán conjecture,* Acta Arith., **63** (1993), 373–378.

[13] M. Helm, *A generalization of a theorem of Erdős on asymptotic basis of order 2,* J. Théor. Nombres Bordeaux, **6** (1994), 9–19.

[14] R. Motwani and P. Raghavan, *Randomized algorithms,* Cambridge Univ. Press, New York, NY, 1995.

[15] D. J. Newman, *An L^1 extremal problem for polynomials,* Proc. Amer. Math. Soc., **16** (1965), 1287–1290.

[16] I. Ruzsa, *A just basis,* Monatsh. Math., **109** (1990), 145–151.

[17] C. Sándor, *A note on a conjecture of Erdős-Turán,* Integers, **8** (2008), #A30, 4 p.

[18] M. Tang, *A note on a result of Ruzsa,* Bull. Austral. Math. Soc., **77** (2008), 91–98.

[19] M. Tang, *A note on a result of Ruzsa, II,* Bull. Austral. Math. Soc., **82** (2010), 340–347.

[20] M. Tang and Y. G Chen, *A basis of $\mathbb{Z}_m$,* Colloq. Math., **104** (2006), 99–103.