Additive representation functions

Sándor Kiss

Doctoral thesis

Supervisor: **András Sárközy**

Professor

Member of the Hungarian Academy of Sciences

Doctoral school: Mathematics
Director: Professor Miklós Laczkovich

Doctoral program: Pure mathematics
Director: Professor András Szűcs

Eötvös Loránd University
Faculty of Natural Sciences
Department of Algebra and Number Theory

2009

Contents

Introduction

In this thesis we prove some results about the additive representation functions. Let $\mathbb{N}$ denote the set of positive integers, and let $k \geq 2$ be a fixed integer. Let $\mathcal{A} = \{a_1, a_2, \dots\}$ $(a_1 < a_2 < \dots)$ be an infinite sequence of positive integers. For $k \geq 2$ integer and $\mathcal{A} \subset \mathbb{N}$, and for $n = 0, 1, 2, \dots$ let $R_1(\mathcal{A}, n, k)$, $R_2(\mathcal{A}, n, k)$, $R_3(\mathcal{A}, n, k)$ denote the number of solutions of the equations

$$a_{i_1} + a_{i_2} + \dots + a_{i_k} = n, \quad a_{i_1} \in \mathcal{A}, \dots, a_{i_k} \in \mathcal{A},$$

$$a_{i_1} + a_{i_2} + \dots + a_{i_k} = n, \quad a_{i_1} < a_{i_2} < \dots < a_{i_k} \quad a_{i_1} \in \mathcal{A}, \dots, a_{i_k} \in \mathcal{A},$$

and

$$a_{i_1} + a_{i_2} + \dots + a_{i_k} = n, \quad a_{i_1} \leq a_{i_2} \leq \dots \leq a_{i_k}, \quad a_{i_1} \in \mathcal{A}, \dots, a_{i_k} \in \mathcal{A},$$

respectively. If $F(n) = O(G(n))$ then we write $F(n) << G(n)$. Put

$$A(n) = \sum_{\substack{a \in \mathcal{A} \\ a \leq n}} 1.$$

The research of the additive representation functions began in the 1950's with the famous Erdős - Fuchs theorem [12], which plays fundamental role in this topic, according to Erdős this theorem certainly will survive the authors by centuries [22]. The Erdős - Fuchs theorem states that if c is a positive constant, then

$$\sum_{n \leq N} R_1(\mathcal{A}, n, 2) = cN + o(N^{1/4}(\log N)^{-1/2})$$

cannot hold. This result have been generalized and extended by many people. As a corollary one can get an Ω - result for the error term in the circle problem. Starting from a problem of Sidon, P. Erdős proved that there exists a sequence $\mathcal{A} \subset \mathbb{N}$ so that there are two constans c_1 and c_2 for which for every n

$$c_1 \log n < R_1(\mathcal{A}, n, 2) < c_2 \log n.$$

On the other hand an old conjecture of Erdős states that for no sequence $\mathcal{A}$ can we have

$$\frac{R_1(\mathcal{A}, n, 2)}{\log n} \to c \quad (0 < c < +\infty).$$

There are some related questions in [3] and [12]. These problems led P. Erdős, A. Sárközy and V. T. Sós to study the regularity property and the monotonicity of the function $R_1(\mathcal{A}, n, 2)$ see in [6], [7], [8], [9]. In this thesis I study the regularity properties and the monotonicity of the representation function $R_1(\mathcal{A}, n, k)$ for $k > 2$ integer. I extend and generalize some result of P. Erdős, A. Sárközy and V. T. Sós by using the generator function method and the probabilistic method.

In chapter 1. I give a short survey about the probabilistic method we are working with. We use the definitions and notations of the Halberstam - Roth book [12]. This method plays a crucial role in this thesis. The next four chapters of the thesis consist my papers. In chapter 2. I study the monotonicity of $R_1(\mathcal{A}, n, k)$. For $k = 2$, P. Erdős, A. Sárközy and V. T. Sós studied the monotonicity of $R_1(\mathcal{A}, n, 2)$. I extend one of their results to any $k > 2$ by using the generator function method [18]. In chapter 3. I study the difference sequence of the additive representation functions. I extend and generalize some of the results of Erdős, Sárközy and V. T. Sós [16], [17], [20] by using the generator function method and the probabilistic method. In chapter 4. I study the regularity property of an additive representation function. I extend one of the result of Erdős and Sárközy by using probabilistic methods [19]. Finally in chapter 5. I study the connection between the asymptotic bases and Sidon sets. For $h \geq 2$ integer we say a set $\mathcal{A}$ of positive integers is an asymptotic basis of order h if every large enough positive integer can be represented as the sum of h terms from $\mathcal{A}$. A set of positive integers $\mathcal{A}$ is called Sidon set if all the sums $a + b$ with $a \in \mathcal{A}$, $b \in \mathcal{A}$, $a \leq b$ are distinct. In chapter 5. we prove the existence of Sidon sets which are asymptotic bases of order 5 by using probabilistic methods [21], especially the Janson inequality. In some chapter of this thesis the definitions sometimes repeated, which helps the reader to understand this thesis better.

Acknowledgement. I would like to thank my supervisor Professor András Sárközy, drawing my attention to the additive representation functions. I have learned a lot from our consultations, without his valuable advice, problems and questions I would never have been able to write my papers and this thesis. I would like to thank Professor Imre Ruzsa for the helpful and valuable discussions about Sidon sets and asymptotic bases.

Chapter 1

The Probabilistic Method

An important problem in additive number theory is to prove that a sequene with certain properties exists. One of the essential ways to obtain an affirmative answer for such a problem is to use the probabilistic method due to Erdős and Rényi. There is an excellent summary of this method in the Halberstam - Roth book [12]. In this thesis we use the notation and terminology of this book. To show that a sequence with a property $\mathcal{P}$ exists, it sufficies to show that a properly defined random sequence satisfies $\mathcal{P}$ with positive probability. Usually the property $\mathcal{P}$ requires that for all sufficiently large $n \in \mathbb{N}$, some relation $\mathcal{P}(n)$ holds. The general strategy to handle this situation is the following. For each n one first shows that $\mathcal{P}(n)$ fails with a small probability, say f_n. If f_n is sufficiently small so that $\sum_{n=1}^{+\infty} f_n$ converges, then by the Borel - Cantelli lemma, $\mathcal{P}(n)$ holds for all sufficiently large n with probability 1 (see also [26]).

Now we give a survey of the probabilistic tools and notations which we use in this thesis. Let Ω denote the set of strictly increasing sequences of positive integers. In this thesis we denote the probability of an event $\mathcal{E}$ by $P(\mathcal{E})$, and the expectation of a random variable ζ by $E(\zeta)$. The following Lemma plays an important role in our proofs.

1.1 Lemma *Let*

$$\alpha_1, \alpha_2, \alpha_3 \dots \tag{1.1}$$

be real numbers satisfying

$$0 \leq \alpha_n \leq 1 \quad (n = 1, 2, \dots). \tag{1.2}$$

Then there exists a probability space $(\Omega,\ X,\ P)$ *with the following two properties:*

(i) For every natural number n*, the event* $\mathcal{E}^{(n)} = \{\mathcal{A}\colon\ \mathcal{A} \in \Omega,\ n \in \mathcal{A}\}$ *is measurable, and* $P(\mathcal{E}^{(n)}) = \alpha_n$*.*

(ii) The events $\mathcal{E}^{(1)}$*,* $\mathcal{E}^{(2)}$*, ... are independent.*

See Theorem 13. in [12], p. 142. We denote the characteristic function of the event $\mathcal{E}^{(n)}$ by $\varrho(\mathcal{A}, n)$:

$$\varrho(\mathcal{A}, n) = \begin{cases} 1, \text{ if } n \in \mathcal{A} \\ 0, \text{ if } n \notin \mathcal{A}. \end{cases}$$

Furthermore, we denote the number of solutions of $a_{i_1} + a_{i_2} + \ldots + a_{i_k} = n$ by $r_k(\mathcal{A}, n)$, where $a_{i_1} \in \mathcal{A}$, $a_{i_2} \in \mathcal{A}$, ...,$a_{i_k} \in \mathcal{A}$, $1 \leq a_{i_1} < a_{i_2} \ldots < a_{i_k} < n$. Thus

$$r_k(\mathcal{A}, n) = r_k(n) = \sum_{\substack{(a_1, a_2, \dots, a_k) \in \mathbb{N}^k \\ 1 \leq a_1 < \ldots < a_k < n \\ a_1 + a_2 + \ldots + a_k = n}} \varrho(\mathcal{A}, a_1)\varrho(\mathcal{A}, a_2) \ldots \varrho(\mathcal{A}, a_k). \tag{1.3}$$

It is easy to see from (1.3) that $r_k(\mathcal{A}, n)$ is the sum of random variables. However for $k > 2$ these variables are not independent because the same $\varrho(\mathcal{A}, a_i)$ may appear in many terms. There are some probabilistic results which can help us to overcome this trouble. First we present a method of J. H. Kim and V. H. Vu. Interested reader can find more details in [15], [25], [26], [27]. Assume that $t_1, t_2, \dots, t_n$ are independent binary (i.e., all t_i's are in $\{0, 1\}$) random variables. Consider a polynomial Y in $t_1, t_2, \dots, t_n$ with degree k. We say a polynomial Y is positive if it can be written in the form $Y = \sum_i e_i \Gamma_i$, where the e_i's are positive and Γ_i is a product of some t_j's. Given a (multi-) set A, $\partial_A(Y)$ denotes the partial derivative of Y with respect to the variables with indices in A. For instance, if $Y = t_1 t_2^2$ and $A_1 = \{1,2\}$ and $A_2 = \{2,2\}$ then $\partial_{A_1}(Y) = 2t_2$ and $\partial_A Y = 2t_1$. If A is empty then $\partial_A(Y) = Y$. $E_A(Y)$ denotes the expectation of $\partial_A(Y)$. Furthermore, set $E_j(Y) = max_{|A|\geq j} E_A(Y)$, for all $j = 0, 1, \ldots, k$, thus $E_0(Y) = E(Y)$.

1.2 Theorem *(J. H. Kim - V. H. Vu) For every positive integer k there are positive constants d_k and b_k depending only on k such that the following holds. For any positive polynomial $Y = Y(t_1, t_2, \ldots, t_n)$ of degree k, where the t_i's are independent binary random variables,*

$$P\Big(|Y - E(Y)| \geq d_k \lambda^k \sqrt{E_0(Y)E_1(Y)}\Big) \leq b_k e^{-\lambda/4 + (k-1)\log n}.$$

See [15] for the proof. The following inequality due to S. Janson [10], [14], [25] which also plays important role in our proofs.

Consider a set $\{t_i\}_{i \in Q}$ of independent random indicator variables and for an index set Γ a family $\{Q(\gamma)\}_{\gamma \in \Gamma}$ of subsets of the index set Q, and define $I_\gamma = \prod_{i \in Q(\gamma)} t_i$ and $N = \sum_{\gamma \in \Gamma} I_\gamma$. (In other words N counts the number of the given sets $\{Q(\gamma)\}$ that are contained in the random set $\{i \in Q : t_i = 1\}$.) Let us write $\gamma \sim \delta$ if $Q(\gamma) \cap Q(\delta) \neq \emptyset$ but $\gamma \neq \delta$, and define

$$p_\gamma = E(I_\gamma),$$

$$\lambda = E(N) = \sum_\gamma p_\gamma,$$

$$\Delta = \frac{1}{\lambda} \sum_{\gamma \sim \delta} E(I_\gamma I_\delta).$$

1.3 Theorem *(Janson) With notations as above, if $0 \leq \varepsilon \leq 1$, then*

$$P(N \leq (1-\varepsilon)\lambda) \leq exp\Big(-\frac{1}{2(1+\Delta)}\varepsilon^2 \lambda\Big).$$

We will apply the following result due to Erdős and Tetali which is called disjointness lemma. We say events $G_1, \ldots, G_n$ are independent if for all subsets $I \subseteq \{1, \ldots, n\}$, $P(\cap_{i \in I} G_i) = \prod_{i \in I} P(G_i)$.

1.4 Lemma *Let $\{B_i\}$ be a sequence of events in a probability space. If $\sum_i P(B_i) \leq \mu$, then*

$$\sum_{\substack{(B_1,\ldots,B_l) \\ independent}} P(B_1 \cap \ldots \cap B_l) \leq \mu^l / l!.$$

See [10] for the proof.

We also need the Borel - Cantelli lemma (see in [12]):

1.5 Lemma *Let $\{B_i\}$ be a sequence of events in a probability space. If*

$$\sum_{j=1}^{+\infty} P(B_j) < \infty,$$

then with probability 1, at most a finite number of the events B_j can occur.

Finally we need the following combinatorial result due to Erdős and Rado, see [2]. Let r be a positive integer, $r \geq 3$. A collection of sets $A_1, A_2, \ldots A_r$ forms a Delta - system if the sets have pairwise the same intersection.

1.6 Lemma *If H is a collection of sets of size at most m and*

$$|H| > (r-1)^m m!$$

then H contains r sets forming a Delta - system.

Chapter 2

On the monotonicity of an additive representation function

2.1 Introduction

Let $k \geq 2$ be a fixed integer. For $i = 1, 2, 3$ we say $R_i(\mathcal{A}, n, k)$ is monotonous increasing in n from a certain point on, if there exists an integer n_0 with

$$R_i(\mathcal{A}, n+1, k) \geq R_i(\mathcal{A}, n, k) \quad for \quad n \geq n_0.$$

In a series of papers P. Erdős, A. Sárközy and V. T. Sós studied the monotonicity properties of the three representation functions $R_1(\mathcal{A}, n, 2)$, $R_2(\mathcal{A}, n, 2)$, $R_3(\mathcal{A}, n, 2)$. In [9] they proved the following theorems:

2.1 Theorem *(Erdős - Sárközy - T. Sós) The function $R_1(\mathcal{A}, n, 2)$ is monotonous increasing from a certain point on, if and only if the sequence $\mathcal{A}$ contains all the integers from a certain point on, i.e., there exists an integer n_1 with*

$$\mathcal{A} \cap \{n_1, n_1+1, n_1+2, \dots\} = \{n_1, n_1+1, n_1+2, \dots\}.$$

2.2 Theorem *(Erdős - Sárközy - T. Sós) If*

$$A(n) = o\Big(\frac{n}{\log n}\Big)$$

then the functions $R_2(\mathcal{A}, n, 2)$ *and* $R_3(\mathcal{A}, n, 2)$ *cannot be monotonous increasing from a certain point on, i.e., for* $i = 2$ *or* 3*, there does not exist an integer* n_0 *such that*

$$R_i(\mathcal{A}, n+1, 2) \geq R_i(\mathcal{A}, n, 2) \quad for \quad n \geq n_0.$$

A. Sárközy proposed the study of the monotonicity of the functions $R_i(\mathcal{A}, n, k)$ for $k > 2$ [24, Problem 5]. He conjectured [23, p. 337] that for any $k \geq 2$ integer, if $R_i(\mathcal{A}, n, k)$ $(i = 1, 2, 3)$ is monotonous increasing in n from a certain point on, then $A(n) = O(n^{2/k-\varepsilon})$ cannot hold. In this chapter I will prove the following slightly stronger result on $R_1(\mathcal{A}, n, k)$ by using similar methods as in [9]:

2.3 Theorem *If* $k \in \mathbb{N}$*,* $k \geq 2$*,* $\mathcal{A} \subset \mathbb{N}$ *and* $R_1(\mathcal{A}, n, k)$ *is monotonous increasing in* n *from a certain point on, then*

$$A(n) = o\left(\frac{n^{2/k}}{(\log n)^{2/k}}\right)$$

cannot hold.

Unfortunately I have not been able to prove the conjecture for $R_2(\mathcal{A}, n, k)$ and $R_3(\mathcal{A}, n, k)$, thus the conjecture remains open in these cases.

2.2 Proof of Theorem 2.3

We write $R_1(\mathcal{A}, n, k) = R_k(n)$. We prove the result by contradiction. Assume that $R_k(n)$ is monotonous increasing from a certain point on and $A(n) = o\left(\frac{n^{2/k}}{(\log n)^{2/k}}\right)$. First we show that there exist infinitely many integers N satisfying

$$A(N+j) < A(N)\left(\frac{N+j}{N}\right)^2 \quad for \;\; j = 1, 2, \ldots. \tag{2.1}$$

If (2.1) holds only for finitely many N, then there exists an integer N_0 such that

$$A(N_0) > 1$$

and for $N \geq N_0$, there exists an integer $N' = N'(N)$ satisfying $N' > N$ and

$$A(N') \geq A(N)\left(\frac{N'}{N}\right)^2.$$

Then we get by induction that there exist integers $N_1 < N_2 < \ldots < N_j < \ldots$ such that

$$A(N_{j+1}) \geq A(N_j)\left(\frac{N_{j+1}}{N_j}\right)^2 \quad for \;\; j = 0, 1, 2, \ldots,$$

hence

$$A(N_{l+1}) = A(N_0)\prod_{j=0}^{l}\frac{A(N_{j+1})}{A(N_j)} \geq A(N_0)\prod_{j=0}^{l}\left(\frac{N_{j+1}}{N_j}\right)^2 \tag{2.2}$$

$$= A(N_0)\left(\frac{N_{l+1}}{N_0}\right)^2 > \left(\frac{N_{l+1}}{N_0}\right)^2 > N_{l+1}^{3/2}$$

for large enough l. On the other hand, clearly we have

$$A(N_{l+1}) = \sum_{\substack{a\in\mathcal{A}\\ a\leq N_{l+1}}} 1 \leq \sum_{a\leq N_{l+1}} 1 = N_{l+1} \tag{2.3}$$

(2.2) and (2.3) cannot hold simultaneously and this contradiction proves the existence of infinitely many integers N satisfying (2.1).

Throughout the remaining part of the proof of Theorem 2.3 we use the following notations: N denotes a large integer satisfying (2.1). We write $e^{2i\pi\alpha} = e(\alpha)$ and we put $r = e^{-1/N}$, $z = re(\alpha)$ where α is a real variable (so that a function of form $p(z)$ is a function of the real variable $\alpha : p(z) = p(re(\alpha)) = P(\alpha)$). We write

$$f(z) = \sum_{a\in\mathcal{A}} z^a.$$

(Since $r < 1$, this infinite series and all the other infinite series in the remaining part of the proof are absolutely convergent.) Then we have

$$f^k(z) = \sum_{n=1}^{+\infty} R_k(n) z^n.$$

Let I denote

$$I = \int_0^1 |f(z)|^k d\alpha.$$

We will give lower and upper bound for I. The lower bound will be greater then the upper bound, and this contradiction will prove that our indirect assumption cannot hold which will complete the proof of Theorem 2.3.

First we will give lower bound for I. Using Hölder's inequality and Parseval's formula we have

$$\begin{aligned} I^{2/k} &= \left(\int_0^1 |f(z)|^k d\alpha \right)^{2/k} \left(\int_0^1 1 d\alpha \right)^{1-2/k} \geq \int_0^1 |f(z)|^2 d\alpha \\ &= \sum_{a \in \mathcal{A}} r^{2a} \geq \sum_{\substack{a \in \mathcal{A} \\ a \leq N}} r^{2N} = e^{-2} \sum_{\substack{a \in \mathcal{A} \\ a \leq N}} 1 = e^{-2} A(N) \end{aligned}$$

hence

$$I \geq e^{-k} (A(N))^{k/2}. \tag{2.4}$$

Now we will give upper bound for I. First we will estimate $R_k(n)$ in terms of $A(2n)$. Since $R_k(n)$ is monotonous increasing from a certain point on, i.e., there exists an integer n_0 such that $R_k(n+1) \geq R_k(n)$ for $n \geq n_0$, we have

$$\begin{aligned} (A(2n))^k &= \Big(\sum_{\substack{a \in \mathcal{A} \\ a \leq 2n}} 1 \Big)^k = \sum_{\substack{a_1 \in \mathcal{A}, a_2 \in \mathcal{A}, \ldots, a_k \in \mathcal{A} \\ a_1 \leq 2n, a_2 \leq 2n, \ldots, a_k \leq 2n}} 1 \geq \sum_{\substack{a_1 + a_2 + \ldots + a_k \leq 2n \\ a_1 \in \mathcal{A}, \ldots, a_k \in \mathcal{A}}} 1 \\ &\geq \sum_{i=1}^{2n} R_k(i) \geq \sum_{i=n+1}^{2n} R_k(i) \geq \sum_{i=n+1}^{2n} R_k(n) = n R_k(n) \end{aligned}$$

hence

$$\frac{(A(2n))^k}{n} \geq R_k(n) \tag{2.5}$$

for $n \geq n_0$. In view of the monotonicity of $R_k(n)$, and since $\mathcal{A}$ is infinite, we have $R_k(n) \geq 1$ for n large enough. Thus we obtain from (2.5) that

$$(A(2n))^k \geq n \tag{2.6}$$

for n large enough. We have

$$I = \int_0^1 |f(z)|^k d\alpha = \int_0^1 |f^k(z)| d\alpha = \int_0^1 \left| \sum_{n=1}^{+\infty} R_k(n) z^n \right| d\alpha \tag{2.7}$$

$$= \int_0^1 |(1-z) \sum_{n=1}^{+\infty} R_k(n) z^n| |1-z|^{-1} d\alpha.$$

By the monotonicity, and if N and n_0 are large enough we have

$$\left|(1-z) \sum_{n=1}^{+\infty} R_k(n) z^n\right| = \left| \sum_{n=1}^{+\infty} (R_k(n) - R_k(n-1)) z^n \right|$$

$$\leq \sum_{n=1}^{n_0} |R_k(n) - R_k(n-1)| r^n + \sum_{n=n_0+1}^{+\infty} |R_k(n) - R_k(n-1)| r^n$$

$$< \sum_{n=1}^{n_0} |R_k(n) - R_k(n-1)| + \sum_{n=n_0+1}^{+\infty} |R_k(n) - R_k(n-1)| r^n$$

$$= \sum_{n=1}^{n_0} |R_k(n) - R_k(n-1)| + \sum_{n=n_0+1}^{+\infty} (R_k(n) - R_k(n-1)) r^n$$

$$< 2 \sum_{n=1}^{n_0} |R_k(n) - R_k(n-1)| + \sum_{n=1}^{+\infty} (R_k(n) - R_k(n-1)) r^n$$

$$= c_1 + \sum_{n=1}^{+\infty} R_k(n)(r^n - r^{n+1}) = c_1 + (1-r) \sum_{n=1}^{+\infty} R_k(n) r^n$$

$$< c_1 + \sum_{n=1}^{n_0-1} R_k(n) + (1-r) \sum_{n=n_0}^{+\infty} R_k(n) r^n$$

$$< c_2 + (1 - e^{-1/N}) \left(\sum_{n=n_0}^{N} R_k(N) + \sum_{n=N+1}^{+\infty} R_k(n) r^n \right).$$

Thus by (2.1), (2.5) and (2.6) we have

$$\left|(1-z)\sum_{n=1}^{+\infty} R_k(n)z^n\right| < c_2 + N^{-1}\left(N\frac{(A(2N))^k}{N} + \sum_{n=N+1}^{+\infty}\frac{(A(2n))^k}{n}r^n\right)$$

$$< c_2 + N^{-1}\left((A(N))^k\left(\frac{2N}{N}\right)^{2k} + \sum_{n=N+1}^{+\infty}\left(A(N)\left(\frac{2n}{N}\right)^2\right)^k\frac{1}{n}r^n\right)$$

$$< c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}\sum_{n=1}^{+\infty} n^{2k-1}r^n\right)$$

$$< c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}\sum_{n=1}^{+\infty}(n+1)(n+2)\dots(n+2k-1)r^n\right)$$

$$= c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}\sum_{m=2k}^{+\infty} m(m-1)\dots(m-2k+2)r^{m-2k+1}\right)$$

$$< c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}(\sum_{m=0}^{+\infty} r^m)^{(2k-1)}\right)$$

$$= c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}\left(\frac{1}{1-r}\right)^{(2k-1)}\right)$$

$$= c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}}{N^{2k+1}}(2k-1)!(1-r)^{-2k}\right)$$

$$= c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}(2k-1)!}{N^{2k+1}}(1-e^{-1/N})^{-2k}\right).$$

Since

$$1-e^{-x} = x - \frac{x^2}{2!} + \frac{x^3}{3!} - \dots > x - \frac{x^2}{2!} = x(1-\frac{x}{2}) > \frac{x}{2}$$

for $0 < x < 1$, it follows by (2.5) that

$$\left|(1-z)\sum_{n=1}^{+\infty} R_k(n)z^n\right| < c_2 + (A(N))^k\left(2^{2k}N^{-1} + \frac{2^{2k}(2k-1)!}{N^{2k+1}}(2N)^{2k}\right)$$

$$= c_2 + (A(N))^k N^{-1}(2^{2k} + 2^{4k}(2k-1)!) < c_3(A(N))^k N^{-1}. \qquad (2.8)$$

Furthermore we have

$$|1-z| = ((1-z)(1-\bar{z}))^{1/2} = (1+|z|^2 - 2Rez)^{1/2} = \tag{2.9}$$

$$(1+r^2-2r\cos 2\pi\alpha)^{1/2} = ((1-r)^2 + 2r(1-\cos 2\pi\alpha))^{1/2} >$$

$$(2r(1-\cos 2\pi\alpha))^{1/2} = (2e^{-1/N} 2\sin^2 \pi\alpha)^{1/2}$$

$$\geq (2(2\alpha)^2)^{1/2} \geq 2\alpha$$

for $0 \leq \alpha \leq \frac{1}{2}$ and for large N, and

$$|1-z| = ((1-r)^2 + 2r(1-\cos 2\pi\alpha))^{1/2} \geq ((1-r)^2)^{1/2} \tag{2.10}$$

$$= 1 - r = 1 - e^{-1/N} > 1/2N$$

for all α. It follows from (2.7), (2.8), (2.9) and (2.10) that

$$I \leq \int_0^1 c_3 (A(N))^k N^{-1} |1-z|^{-1} d\alpha \tag{2.11}$$

$$= 2c_3 (A(N))^k N^{-1} \int_0^{1/2} |1-z|^{-1} d\alpha$$

$$= c_4 (A(N))^k N^{-1} \left(\int_0^{1/N} |1-z|^{-1} d\alpha + \int_{1/N}^{1/2} |1-z|^{-1} d\alpha \right)$$

$$< c_4 (A(N))^k N^{-1} \left(\int_0^{1/N} 2N d\alpha + \int_{1/N}^{1/2} (2\alpha)^{-1} d\alpha \right)$$

$$< c_4 (A(N))^k N^{-1} (2 + \frac{1}{2}\log N) < c_5 (A(N))^k N^{-1} \log N.$$

In view of (2.4), (2.11) and our indirect assumption we have

$$e^{-k} (A(N))^{k/2} \leq I < c_5 (A(N))^k N^{-1} \log N,$$

$$N < c_6 (A(N))^{k/2} \log N = o\left(\left(\frac{N^{2/k}}{(\log N)^{2/k}} \right)^{k/2} \log N \right) = o(N).$$

This contradiction completes the proof of Theorem 2.3.

Chapter 3

On the difference sequence of an additive representation function

3.1 Introduction

In this chapter we write $R_1(\mathcal{A}, n, k) = R_k(n)$. Let $k \geq 2$, $t \geq 1$ be fixed integers. If $s_0, s_1, s_2 \dots$ is a given sequence of real numbers, then let $\Delta_t s_n$ denote the t-th difference of the sequence $s_0, s_1, s_2 \dots$ defined by $\Delta_1 s_n = s_{n+1} - s_n$ and $\Delta_t s_n = \Delta_1(\Delta_{t-1} s_n)$. It is well-known and it is easy to see by induction that

$$\Delta_t s_n = \sum_{i=0}^{t} (-1)^{t-i} \binom{t}{i} s_{n+i}. \tag{3.1}$$

Let $B(\mathcal{A}, N)$ denote the number of blocks formed by consecutive integers in $\mathcal{A}$ up to N, i.e.,

$$B(\mathcal{A}, N) = \sum_{\substack{a \leq N \\ a-1 \notin \mathcal{A}, a \in \mathcal{A}}} 1.$$

We will consider the following problem : what condition is needed to guarantee that $|\Delta_t R_k(n)|$ cannot be bounded. P. Erdős, A. Sárközy and V. T. Sós proved in [8] that if $k = 2$, $t = 1$ then

3.1 Theorem *(Erdős, Sárközy, T.Sós): If* $\lim_{N\to\infty} \frac{B(\mathcal{A},N)}{\sqrt{N}} = \infty$, *then* $|\Delta_1(R_2(n))| = |R_2(n+1) - R_2(n)|$ *cannot be bounded.*

They also proved in [8] that the above result is nearly best possible:

3.2 Theorem *(Erdős, Sárközy, T.Sós): For all* $\varepsilon > 0$, *there exists an infinite sequence* $\mathcal{A}$ *such that*

(i) $B(\mathcal{A}, N) \gg N^{1/2-\varepsilon}$,

(ii) $R_2(n)$ *is bounded so that also* $\Delta_1 R_2(n)$ *is bounded.*

In [16] I extended Theorem 3.1 to any $k > 2$:

3.3 Theorem *If* $k \geq 2$ *is an integer and* $\lim_{N\to\infty} \frac{B(\mathcal{A},N)}{\sqrt[k]{N}} = \infty$, *and* $t \leq k$, *then* $|\Delta_t R_k(n)|$ *cannot be bounded.*

I also proved in [20] that the above result is nearly best possible:

3.4 Theorem *For all* $\varepsilon > 0$, *there exists an infinite sequence* $\mathcal{A}$ *such that*

(i) $B(\mathcal{A}, N) \gg N^{1/k-\varepsilon}$,

(ii) $R_k(n)$ *is bounded so that also* $\Delta_t R_k(n)$ *is bounded if* $t \leq k$.

In the case $t > k$ I have only a partial result ([17]):

3.5 Theorem *If* $t \geq 2$ *is an integer and* $\lim_{N\to\infty} \frac{B(\mathcal{A},N)}{\sqrt{N}} = \infty$, *then* $|\Delta_t(R_2(n))|$ *cannot be bounded.*

In the next part of this chapter I prove Theorem 3.3 and Theorem 3.4. I omit the proof of Theorem 3.5 because it is similar to the proof Theorem 3.3. Interested reader can find it in [17].

3.2 Proof of Theorem 3.3

Clearly it sufficies to prove the assertion of the theorem in the special case $t = k$. We prove by contradiction. Assume that contrary to the conclusion of the theorem there is a positive constant $C > 0$ such that $|\Delta_k R_k(n)| < C$ for every n. Throughout the remaining part of the proof of the theorem we use the following notations: N denotes a large integer. We write $e^{2i\pi\alpha} = e(\alpha)$ and we put $r = e^{-1/N}$, $z = re(\alpha)$ where α is a real variable (so that a function of form $p(z)$ is a function of the real variable $\alpha : p(z) = p(re(\alpha)) = P(\alpha)$). We write $f(z) = \sum\limits_{a\in\mathcal{A}} z^a$. (By $r < 1$, this infinite series and all the other infinite series in the remaining part of the proof are absolutely convergent).

We start out from the integral $I = \int\limits_0^1 |f(z)(1-z)|^k d\alpha$. We will give lower and upper bound for I. The comparison of these bounds will show that $\frac{B(\mathcal{A},N)}{\sqrt[k]{N}}$ is bounded which contradicts the assumption of the theorem. This contradiction will prove that our indirect assumption on $|\Delta_k R_k(n)|$ cannot hold which will complete the proof of the theorem.

First we will give a lower bound for I. We write $f(z)(1-z) = \sum\limits_{n=1}^{\infty} \beta_n z^n$. Then for $n-1 \notin \mathcal{A}, n \in \mathcal{A}$ we have $\beta_n = 1$, thus by the Hölder inequality and the Parseval formula, we have

$$I^{2/k} = \Big(\int_0^1 |f(z)(1-z)|^k d\alpha\Big)^{2/k} \Big(\int_0^1 1 d\alpha\Big)^{1-2/k} \geq \int_0^1 |f(z)(1-z)|^2 d\alpha$$

$$= \int_0^1 \Big|\sum_{n=1}^{\infty} \beta_n z^n\Big|^2 d\alpha = \sum_{n=1}^{\infty} \beta_n^2 r^{2n} \geq r^{2N} \sum_{\substack{n\leq N\\ n-1\notin\mathcal{A}, n\in\mathcal{A}}} \beta_n^2 =$$

$$= e^{-2} \sum_{\substack{n\leq N\\ n-1\notin\mathcal{A}, n\in\mathcal{A}}} 1 = e^{-2} B(\mathcal{A}, N).$$

whence

$$I \geq e^{-k}(B(\mathcal{A}, N))^{k/2}.$$

Now we will give an upper bound for I. By (3.1), our indirect assumption, the Cauchy inequality and the Parseval formula we have

$$I = \int_0^1 |f(z)(1-z)|^k d\alpha = \int_0^1 |f^k(z)(1-z)^k| d\alpha = \int_0^1 \Big|(\sum_{a\in\mathcal{A}} z^a)^k(1-z)^k\Big| d\alpha$$

$$= \int_0^1 \Big|(\sum_{n=1}^{\infty} R_k(n)z^n)(1-z)^k\Big| d\alpha = \int_0^1 \Big|(\sum_{n=1}^{\infty} R_k(n)z^n)(\sum_{i=0}^{k}(-1)^i\binom{k}{i}z^i)\Big| d\alpha$$

$$= \int_0^1 \Big|\sum_{m=1}^{\infty}\sum_{i=0}^{k}(-1)^i\binom{k}{i}R_k(m-i)z^m\Big| d\alpha = \int_0^1 \Big|\sum_{m=1}^{\infty}\Delta_k R_k(m-k)z^m\Big| d\alpha$$

$$\leq \left(\int_0^1 \Big|\sum_{m=1}^{\infty}\Delta_k R_k(m-k)z^m\Big|^2 d\alpha\right)^{1/2} = \left(\sum_{m=1}^{\infty}\big|\Delta_k R_k(m-k)\big|^2 r^{2m}\right)^{1/2}$$

$$\leq C\left(\sum_{m=1}^{\infty} r^{2m}\right)^{1/2}$$

$$= C\left(\frac{1}{1-r^2}\right)^{1/2} \leq C\left(\frac{1}{1-r}\right)^{1/2}$$

$$= C\left(\frac{1}{1-e^{-\frac{1}{N}}}\right)^{1/2} < C\sqrt{2N}$$

since we have

$$1-e^{-x} = x - \frac{x^2}{2!} + \frac{x^3}{3!} - \cdots > x - \frac{x^2}{2!} = x(1-\frac{x}{2}) > \frac{x}{2}$$

for $0 < x < 1$.

Now we will complete the proof of the theorem. We have

$$e^{-k}\Big(B(\mathcal{A}, N)\Big)^{k/2} \leq I < C\sqrt{2N}$$

hence

$$\frac{B(\mathcal{A},N)}{\sqrt[k]{N}} < e^2\sqrt[k]{2C^2}.$$

This contradicts our assumption on $B(\mathcal{A},N)$ which completes the proof of Theorem 3.3.

3.3 Proof of Theorem 3.4

The proof of Theorem 3.4 is based on the probabilistic method due to Erdős and Rényi we introduced in chapter 2.

First we proof part (i) of Theorem 3.4. The proof is similar as in [8]. To do this, we need the following important lemma:

3.6 Lemma *If the sequence (1.1) satisfies (1.2) and*

$$\alpha_j = \delta j^{-c} \quad for \quad j \geq j_0,$$

where δ, c *are constants such that* $0 < \delta$, $0 < c < 1$, *then with probability 1, we have*

$$A(n) \sim \frac{\delta}{1-c} n^{1-c}.$$

This lemma is a consequence of Lemmas 10 and 11 in [12], pp. 144 - 145. For $\mathcal{A} \in \Omega$, we write

$$T(\mathcal{A},n) = \sum_{\substack{a \leq n \\ a-1 \in \mathcal{A}, a \in \mathcal{A}}} 1$$

so that

$$B(\mathcal{A},n) + T(\mathcal{A},n) = \sum_{\substack{a \leq n \\ a-1 \notin \mathcal{A}, a \in \mathcal{A}}} 1 + \sum_{\substack{a \leq n \\ a-1 \in \mathcal{A}, a \in \mathcal{A}}} 1$$

$$= \sum_{\substack{a \in \mathcal{A} \\ a \leq n}} 1 = A(n).$$

The following lemma will play a crucial role in the proof.

3.7 Lemma *If the sequence (1.1) satisfies (1.2) and*

$$\sum_{j=1}^{+\infty} \alpha_j \alpha_{j+1} < +\infty, \tag{3.2}$$

then, with probability 1,

$$T(\mathcal{A}, n) < 4 \log n \quad for \quad n > n_2(\mathcal{A})$$

(where n_2 may depend on both the sequence (1.1) and $\mathcal{A}$).

See this lemma and the proof in [8]. Define the sequence (1.1) by

$$\alpha_j = \frac{1}{k} j^{1/k-1-\varepsilon}. \tag{3.3}$$

Thus by Lemma 3.6 with probability 1, we have

$$A(n) \sim \frac{1}{k}\Big(\frac{1}{k} - \varepsilon\Big)^{-1} n^{\frac{1}{k}-\varepsilon}$$

so that, with probability 1,

$$A(n) > \frac{1}{k} k n^{\frac{1}{k}-\varepsilon} = n^{\frac{1}{k}-\varepsilon}, \tag{3.4}$$

for n large enough. By Lemma 3.7 (clearly, the sequence (3.3) satisfies (3.2)), with probability 1,

$$B(\mathcal{A}, n) = A(n) - T(\mathcal{A}, n) > n^{\frac{1}{k}-\varepsilon} - 4\log n > \frac{1}{k} n^{\frac{1}{k}-\varepsilon}$$

for $n > n_3(\varepsilon, \mathcal{A})$. In the next section we will prove part (ii) of Theorem 3.4. Remember that

$$r_k(n) = \sum_{\substack{(a_1, a_2, \dots, a_k) \in \mathbb{N}^k \\ 1 \le a_1 < \dots < a_k < n \\ a_1 + a_2 + \dots + a_k = n}} \varrho(\mathcal{A}, a_1)\varrho(\mathcal{A}, a_2) \dots \varrho(\mathcal{A}, a_k).$$

(see in (1.3) in Chapter 2). Let $r_k^*(\mathcal{A}, n)$ denote the number of those representations of n as the sum of k terms from $\mathcal{A}$ in which there are at least two equal terms. Thus we have

$$R_1(\mathcal{A}, n, k) = k! r_k(\mathcal{A}, n) + r_k^*(\mathcal{A}, n). \tag{3.5}$$

Write $r_k^*(\mathcal{A}, n) = r_k^*(n)$. It follows that we have to show that with probability 1, both $r_k(n)$ and $r_k^*(n)$ are bounded by a constant. First we prove $r_k(n)$ is bounded by using similar methods as Erdős and Tetali in [10]. Let $S_1 = \{a_1, a_2, \dots a_k\}$ and $S_2 = \{b_1, b_2, \dots b_k\}$, be two different representations of n as the sum of k terms from $\mathcal{A}$, that is, $S_1 \neq S_2$ and $S_1, S_2 \subset \mathcal{A}$ and

$$a_1 + a_2 + \ldots + a_k = b_1 + b_2 + \ldots + b_k = n.$$

We say S_1 and S_2 are disjoint if they share no element in common. Let $h(n)$ denote the size of a maximal collection of pairwise disjoint representations of n as the sum of k distinct numbers from $\mathcal{A}$. We can see in (1.3) that $r_k(n)$ is the sum of random variables. However, for $k > 2$ these variables are not independent because any $\varrho(\mathcal{A}, a_i)$ may appear in many terms. To overcome this problem we will prove that with probability 1, $h(n)$ and $r_{k-1}(n)$ are bounded by a constant, i. e., almost always there exist constans g and c_1 such that $f(n) < g$ and $r_{k-1}(n) < c_1$. The following argument shows that this implies $r_k(n)$ almost always bounded by a constant. Let S be any maximal collection of pairwise disjoint representations of n as a sum of k distinct numbers. Clearly $|S| = h(n)$. It is clear that if $h(n) < g$ then there are at most $k \times g$ numbers in our collection S. As S is maximal, any representation of n must use at least one number from the collection. However, the number of representations of n which use x is precisely $r_{k-1}(n - x)$. If $r_{k-1}(n) < c_1$ then the total number of representations of n is at most $c_1 \times k \times g$. Now we give an upper estimation for $h(n)$. Let $E(r_k(n))$ denote the expectation of $r_k(n)$. We need an upper estimation for $E(r_k(n))$. Clearly $a_k > n/k$, thus we have

$$E(r_k(n)) = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ 1 \le a_1 < a_2 < \ldots < a_k < n}} P(a_1 \in \mathcal{A}) P(a_2 \in \mathcal{A}) \ldots P(a_k \in \mathcal{A}) \qquad (3.6)$$

$$= \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ 1\le a_1<a_2<\ldots<a_k<n}} \frac{1}{(a_1\ldots a_k)^{1+\varepsilon-1/k}}$$

$$< \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k}\Big(\sum_{\substack{a_1+a_2+\ldots+a_k=n \\ 1\le a_1<a_2<\ldots<a_k<n}} \frac{1}{(a_1\ldots a_{k-1})^{1+\varepsilon-1/k}}\Big)$$

$$\le \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k} \sum_{\substack{1\le a_i\le n \\ i=1\ldots k-1}} \frac{1}{(a_1\ldots a_{k-1})^{1+\varepsilon-1/k}}$$

$$= \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k}\Big(\sum_{1\le a_1\le n}\frac{1}{a_1^{1+\varepsilon-1/k}}\Big)^{k-1} = \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k}\Big(\int_1^n \frac{1}{a_1^{1+\varepsilon-1/k}}da_1+O(1)\Big)^{k-1}$$

$$= \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k}\Big(\Big[\frac{a_1^{1/k-\varepsilon}}{\frac{1}{k}-\varepsilon}\Big]_1^n + O(1)\Big)^{k-1}$$

$$= \Big(\frac{k}{n}\Big)^{1+\varepsilon-1/k}\Bigg(\Big(\frac{k}{1-k\varepsilon}\Big)^{k-1} n^{\frac{(1-k\varepsilon)(k-1)}{k}} + o(n^{\frac{(1-k\varepsilon)(k-1)}{k}})\Bigg)$$

$$= \frac{k^{1+\varepsilon-1/k}}{n^{k\varepsilon}}\Bigg(\Big(\frac{k}{1-k\varepsilon}\Big)^{k-1} + o(1)\Bigg) < C(k,\varepsilon)n^{-k\varepsilon} = c_2 n^{-k\varepsilon},$$

where c_2 is a constant depending on k and ε. Let

$$\mathcal{B} = \{(a_1,\ldots,a_k) : a_1+\ldots+a_k = n, a_1\in\mathcal{A},\ldots,a_k\in\mathcal{A}, 1\le a_1<\ldots<a_k<n\},$$

and let $H(\mathcal{B}) = \{\mathcal{T}\subset\mathcal{B}$: all the $S\in\mathcal{T}$ are pairwise disjoint$\}$. It is clear that the pairwise disjointness of the sets implies the independence of the associated events, i. e., if S_1 and S_2 are pairwise disjoint representations, the events $S_1\subset\mathcal{A}$, $S_2\subset\mathcal{A}$ are independent. Thus by (3.6) and Lemma 1.4 for $g = \left[\frac{1}{\varepsilon}\right]$ we have

$$P(h(n)>g) \le P\Big(\bigcup_{\substack{\mathcal{T}\subset H(\mathcal{B}) \\ |\mathcal{T}|=g+1}}\bigcap_{S\in\mathcal{T}} S\Big) \le \sum_{\substack{\mathcal{T}\subset H(\mathcal{B}) \\ |\mathcal{T}|=g+1}} P\Big(\bigcap_{S\in\mathcal{T}} S\Big) \tag{3.7}$$

$$= \sum_{\substack{(S_1,\ldots,S_{g+1}) \\ Pairwise \\ disjoint}} P(S_1\cap\ldots\cap S_{g+1}) \le \frac{1}{(g+1)!}(E(h(n)))^{g+1}$$

$$\leq \frac{1}{(g+1)!}(E(r_k(n)))^{g+1} \leq \frac{1}{(g+1)!}c_2^{g+1}n^{-k(g+1)\varepsilon} \leq c_3(\varepsilon)n^{-k},$$

where $c_3(\varepsilon)$ is a constant depending on ε. Using the Borel - Cantelli lemma, it follows that with probability 1, there exists an n_0 such that

$$h(n) \leq g \ \ for \ \ n > n_0. \tag{3.8}$$

In the next step we will give an upper bound for $r_{k-1}(n)$. Before doing this, we introduce some new notations. Let $r_l(n)$ denote the number of representations of n as the sum of l distinct numbers from $\mathcal{A}$ and let $h_l(n)$ denote the size of a maximal collection of pairwise disjoint such representations. We will give an upper estimation for $h_l(n)$ similarly as in (3.7). First we give an upper estimation for $E(r_l(n))$ similarly to (3.6): Let $2 \leq l \leq k-1$ be fixed. Then using the definition, we have $n/l < a_l$, thus

$$E(r_l(n)) = \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\leq a_1<a_2<\ldots<a_l<n}} P(a_1 \in \mathcal{A})P(a_2 \in \mathcal{A})\ldots P(a_l \in \mathcal{A}) \tag{3.9}$$

$$= \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\leq a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_l)^{1+\varepsilon-1/k}}$$

$$< n^{-1-\varepsilon+1/k+o(1)} \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\leq a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_{l-1})^{1+\varepsilon-1/k}}$$

$$< n^{-1-\varepsilon+1/k+o(1)} \sum_{\substack{1\leq a_i\leq n \\ i=1\ldots l-1}} \frac{1}{(a_1\ldots a_{l-1})^{1+\varepsilon-1/k}}$$

$$= n^{-1-\varepsilon+1/k+o(1)}\Big(\sum_{1\leq a_1\leq n} \frac{1}{a_1^{1+\varepsilon-1/k}}\Big)^{l-1}$$

$$= n^{-1-\varepsilon+1/k+o(1)}(n^{1/k-\varepsilon+o(1)})^{l-1} = n^{-1+l/k-(l\varepsilon)+o(1)}.$$

Let $S^{[l]}$ denote a representation of n as a sum of l distinct numbers. When $S_i^{[l]}$ and $S_j^{[l]}$ are disjoint $S_i^{[l]} \subset \mathcal{A}$ and $S_j^{[l]} \subset \mathcal{A}$ are independent events. For

$2 \leq l \leq (k-1)$, applying Lemma 1.4, using an argument similar to (3.7), and in view of (3.9) we have

$$P(h_l(n) > 2k) < \sum_{\substack{(S_1^{[l]},\dots,S_{2k+1}^{[l]})\\ Pairwise\\ disjoint}} P(S_1^{[l]} \cap \ldots \cap S_{2k+1}^{[l]})$$

$$< \frac{(E(r_l(n))^{2k}}{(2k+1)!} < \frac{1}{(2k+1)!}(n^{-1+l/k-l\varepsilon+o(1)})^{2k} = n^{-2k+2l(1-k\varepsilon)+o(1)}.$$

By $l \leq (k-1)$ it follows that

$$P(h_l(n) > 2k) < n^{-2+o(1)}.$$

Thus by the Borel - Cantelli lemma with probability 1, the above assertion implies that almost always for $2 \leq l \leq (k-1)$ there exists n_l such that if $n > n_l$ then $h_l(n) \leq 2k$. But for any finite n_l, there are at most a finite number of representations as a sum of l numbers. Therefore, almost always for $2 \leq l \leq (k-1)$ there exists a C_l such that for every n, $h_l(n) < C_l$. Set $c_{max} = max_l\{C_l\}$. Now we show similarly as in [10] that almost always there exists $c_4 = c_4(\mathcal{A})$ such that for every n,

$$r_{k-1}(n) < c_4. \tag{3.10}$$

The proof of (3.10) is purely combinatorial. We show that (whenever every C_l exists), for every n

$$r_{k-1}(n) \leq (c_{max})^{k-1}(k-1)!. \tag{3.11}$$

We prove by contradiction. Suppose (3.11) is false for some $n = n'$, i. e.,

$$r_{k-1}(n') > (c_{max})^{k-1}(k-1)!. \tag{3.12}$$

We want to apply Lemma 1.6. Let H be the set of representations of n' as the sum of $k-1$ distinct numbers from $\mathcal{A}$. Clearly $|H| = r_{k-1}(n')$, thus by (3.12) and applying Lemma 1.6 we get that H contains $c_{max}+1$ representations of $n^{'}$ as the sum of $k-1$ distinct numbers which form a Delta - system $\{S_1^{k-1},\dots,S_{c_{max}+1}^{k-1}\}$. If the common intersection of these sets is empty then this $c_{max}+1$ set form a family of disjoint $k-1$ representations of $n^{'}$, which contradicts the definition of c_{max}. Otherwise let the common intersection of the system be $\{x_1,\dots,x_v\}$, where $0\le v\le k-2$. If $\sum_i x_i=m$, then removing the common intersection each set will yield $h_{k-1-v}(n^{'}-m)\ge c_{max}+1$. This is impossible in view of $h_l(n)<C_l$ and the definition of c_{max}. This proves (3.11), and in fact, also shows that $c_4\le c_{max}^{k-1}(k-1)!$.

In the last section we will give an upper estimation for $r_k^*(n)$. It can be prove similarly to the estimate of $r_k(n)$ that is $r_k^*(n)$ is also bounded by a constant. For the sake of completeness I sketch the proof leaving the details to the reader. If we collect the equal terms, we have

$$u_1a_1+u_2a_2+\ldots+u_ha_h=n, \tag{3.13}$$

where the u_i's are natural numbers, and

$$u_1+u_2+\ldots+u_h=k. \tag{3.14}$$

Thus $r_k^*(n)$ denotes the number of representations of n in the form (3.13), where the a_i's are different. Similarly to the estimate of $r_k(n)$, we show that with probability 1, $r_k^*(n)$ is also bounded by a constant. Let $2\le h\le k-1$ be fixed. For a fixed $u_1,\dots,u_h$ denote $w_h(n)$ the number of representations of n in the form (3.13). We show that with probability 1, $w_h(n)$ is bounded by a constant. (Note that in the previous section we proved the case when all u_i's equal to one, and $h=k$). First we will give an upper estimation for $E(w_h(n))$, with a calculation similar to (3.9). Using the definition, and

$n/k < a_h$, we have

$$E(w_h(n)) = \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n\\ 1\le a_1<a_2<\ldots<a_h<n}} P(a_1\in\mathcal{A})P(a_2\in\mathcal{A})\ldots P(a_h\in\mathcal{A}) \tag{3.15}$$

$$= \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n\\ 1\le a_1<a_2<\ldots<a_h<n}} \frac{1}{(a_1\ldots a_h)^{1+\varepsilon-1/k}}$$

$$\le n^{-1-\varepsilon+1/k+o(1)} \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n\\ 1\le a_1<a_2<\ldots<a_h<n}} \frac{1}{(a_1\ldots a_{h-1})^{1+\varepsilon-1/k}}$$

$$< n^{-1+h/k-(h\varepsilon)+o(1)}.$$

Let $w_h^*(n)$ denote the size of a maximal collection of pairwise disjoint representations in the form (3.13). The same argument as in (3.7) and (3.8) shows that almost always there exists a d_h constant such that for every large enough n, $w_h^*(n)\le d_h$. In view of (3.15), and applying Lemma 1.4 we have

$$P(w_h^*(n) > d_h) < n^{-2+o(1)},$$

thus by the Borel - Cantelli lemma we get that with probability 1, $w_h^*(n) < d_h$ if n is large enough. We say that a m - tuple $(a_1,\ldots,a_m)$ $(m\le h)$ is an m - representation of n in the form (3.13) if there is a permutation π of the numbers $\{1,2,\ldots,h\}$ such that $\sum_{i=1}^m u_{\pi(i)}a_i = n$. For all $m<h$, let $w_m^*(n)$ denote the size of a maximal collection of pairwise disjoint such representations of n. The same argument as above shows that almost always there exists p_m constant such that for all large enough n, $w_m^*(n) < p_m$. In the last step we apply Lemma 1.6 to prove that $w_h(n)$ is bounded by a constant. Let $D = \Big(max(p_mh!, d_h)\Big)^h h!$. Let H in Lemma 1.6 is the collection of representations of n in the form (3.13). Clearly $|H| = w_h(n)$. If $w_h(n) > D$, and n is sufficiently large then by Lemma 1.6, H contains a Delta - system with $max(p_mh!, d_h)+1$ sets. If the intersection of these sets is empty, then they form a family of disjoint h - representations in the form (3.13). Otherwise let the common intersection of the sets be $\{y_1, \dots, y_s\}$, where $1 \le s \le h-1$. By the pigeon hole principle, there exists a permutation π of the numbers $\{1, 2, \dots, h\}$ such that we can find p_m+1 $(k-s)$ representations of $n'' = n - \sum_{i=1}^{s} u_{\pi(i)} y_i$. These $p_m + 1$ sets are disjoint, thus in both cases we obtain a contradiction. Since there are only finite number of partitions of k in the form (3.13), we get that $r_k^*(n)$ is bounded by a constant. From (3.5) we get that $R_k(n)$ is also bounded by a constant. Thus with probability 1, both (i) and (ii) in Theorem 3.4 hold, so that there exists infinitely many sequences satisfying both (i) and (ii), which proves Theorem 3.4.

Chapter 4

On the regularity property of an additive representation function

4.1 Introduction

Let $k \geq 2$ be a fixed integer. In this chapter we write $R_1(\mathcal{A}, n, k) = R_k(n)$. For $k = 2$, P. Erdős and A. Sárközy studied how regular can be the behaviour of the function $R_2(n)$. In [6] they proved the following theorem:

4.1 Theorem *(Erdős-Sárközy) If $F(n)$ is an arithmetic function such that*

$$F(n) \to +\infty,$$

$$F(n+1) \geq F(n) \quad for \quad n \geq n_0,$$

$$F(n) = o\Big(\frac{n}{(\log n)^2}\Big),$$

and we write

$$\Gamma(N) = \sum_{n=1}^{N} (R_2(n) - F(n))^2,$$

then

$$\Gamma(N) = o(NF(N))$$

cannot hold.

In [7] they showed that the above result is nearly best possible:

4.2 Theorem *(Erdős-Sárközy) If $F(n)$ is an arithmetic function satisfying*

$$F(n) > 36 \log n \quad for \quad n > n_0,$$

and there exist a real function $g(x)$, defined for $0 < x < +\infty$, and real numbers x_0, n_1 such that

(i) $g^{'}(x)$ exists and it is continuous for $0 < x < +\infty$,

(ii) $g^{'}(x) \le 0$ for $x \ge x_0$,

(iii) $0 < g(x) < 1$ for $x \ge x_0$,

(iv) $|F(n) - 2\int_0^{n/2} g(x)g(n-x)dx| < (F(n)\log n)^{1/2}$ for $n > n_1$,

then there exists a sequence $\mathcal{A}$ such that

$$|R_2(n) - F(n)| < 8(F(n)\log n)^{1/2} \quad for \quad n > n_2.$$

In [13] G. Horváth extended Theorem 4.1 to any $k > 2$:

4.3 Theorem *(G. Horváth) If $F(n)$ is an arithmetic function such that*

$$F(n) \to +\infty,$$

$$F(n+1) \ge F(n) \quad for \quad n \ge n_0,$$
$$F(n) = o\Big(\frac{n}{(\log n)^2}\Big),$$

and we write

$$\Gamma(N) = \sum_{n=1}^{N} (R_k(n) - F(n))^2,$$

then

$$\Gamma(N) = o(NF(N))$$

cannot hold.

A. Sárközy proposed to prove the analogue of Theorem 4.2 for $k > 2$ [23, Problem 3]. In this chapter my goal is to extend Theorem 4.2 to any $k > 2$, i. e., to show that Theorem 4.3 is nearly best possible. In fact I will prove the following theorem:

4.4 Theorem *If $k > 2$ is a positive integer, c_8 is a constant large enough in terms of k, $F(n)$ is an arithmetic function satisfying*

$$F(n) > c_8 \log n \quad for \quad n > n_0,$$

and there exists a real function $g(x)$, defined for $0 < x < +\infty$, and real numbers x_0, n_1 and constants c_7, c_9 such that

(i) $0 < g(x) \le \frac{(\log x)^{\frac{1}{k}}}{x^{1-\frac{k+1}{k^2}}} < 1$ *for* $x \ge x_0$,

(ii) $\left| F(n) - k! \sum_{\substack{x_1+x_2+\ldots+x_k=n \\ 1\le x_1<x_2<\ldots<x_k<n}} g(x_1)g(x_2)\ldots g(x_k) \right| < c_7(F(n)\log n)^{1/2}$ *for* $n > n_1$,

then there exists a sequence $\mathcal{A}$ such that

$$|R_k(n) - F(n)| < c_9(F(n)\log n)^{1/2} \quad for \quad n > n_2.$$

It is easy to see that the following fuctions satisfy the conditions of Theorem 4.4: $g(x) = c_{10}\left(\frac{(\log x)^\beta}{x^\alpha}\right)$, where c_{10} is a positive constant, $\alpha > 1 - \frac{k+1}{k^2}$, or $\alpha = 1 - \frac{k+1}{k^2}$ and $\beta \le 1/k$. It follows that for $F(n) = n^\delta(\log n)^\gamma$ with $0 < \delta \le 1/k$, or $0 \le \gamma < 1$ there is a sequence $\mathcal{A}$ for which $R_k(n)$ satisfies the coclusion of the theorem. For $k = 2$ in [6] P. Erdős and A. Sárközy used probabilistic method to construct a sequence $\mathcal{A}$. In the case $k = 2$, in their paper certain events were mutually independent. For $k > 2$ the independency fails, thus in order to prove Theorem 4.4 we need deeper probabilistic tools.

4.2 Proof of Theorem 4.4

Fix a number n and write

$$S_n = \{(a_1, a_2, \ldots a_k) \in \mathbb{N}^k : 0 < a_1 < a_2 \ldots < a_k < n, a_1+a_2+\ldots+a_k = n\}.$$

Define the sequence (1.1) of real numbers by

$$\alpha_n = \begin{cases} g(n) \text{ if } n \ge x_0, \\ 0 \text{ otherwise }, \end{cases}$$

and let (Ω, X, P) be the probability space as described in Lemma 1.1. Clearly the sequence α_n satisfies (1.2). Thus

$$r_k(n, \mathcal{A}) = r_k(n) = \sum_{(a_1, a_2, \ldots, a_k) \in S_n} t_{a_1} t_{a_2} \ldots t_{a_k},$$

where

$$t_{a_i} = \begin{cases} 1, \text{ if } a_i \in \mathcal{A} \\ 0, \text{ if } a_i \notin \mathcal{A} \end{cases}.$$

Then we have

$$\lambda_n = E(r_k(n)) = \sum_{(a_1, a_2, \ldots, a_k) \in S_n} P(a_1 \in \mathcal{A}) P(a_2 \in \mathcal{A}) \ldots P(a_k \in \mathcal{A}),$$

where $E(\zeta)$ denotes the expectation of the random variable ζ. To prove Theorem 4.4 we will give an upper estimation for $|R_k(n) - k!\lambda_n|$. As Vu in [26] we split $r_k(n)$ into two parts, as follows. Let a be a small positive constant say $a < \frac{1}{2(k+1)}$ and let $S_n^{[1]}$ be the subset of S_n consisting of all k-tuples whose smallest element is at least n^a, i. e., $S_n^{[1]} = \{(a_1, a_2, \ldots a_k) \in \mathbb{N}^k : n^a \leq a_1 < a_2 \ldots < a_k < n, a_1 + a_2 + \ldots + a_k = n\}$ and $S_n^{[2]} = S_n \setminus S_n^{[1]}$. We split $r_k(n)$ into the sum of two terms corresponding to $S_n^{[1]}$ and $S_n^{[2]}$, respectively:

$$r_k(n) = r_k^{[1]}(n) + r_k^{[2]}(n),$$

where

$$r_k^{[j]}(n) = \sum_{(a_1, a_2, \ldots, a_k) \in S_n^{[j]}} t_{a_1} t_{a_2} \ldots t_{a_k}, \tag{4.1}$$

and set

$$\lambda_n^{[j]} = E(r_k^{[j]}(n)).$$

Let $r_k^*(\mathcal{A}, n)$ denote the number of those representations of n as the sum of k terms from $\mathcal{A}$ in which there are at least two equal terms. Thus we have

$$R_k(n) = k! r_k(\mathcal{A}, n) + r_k^*(\mathcal{A}, n). \tag{4.2}$$

Clearly

$$|R_k(n) - k!\lambda_n| \leq |R_k(n) - k!r_k(n)| + k!|r_k(n) - \lambda_n| \tag{4.3}$$
$$= r_k^*(n) + k!|r_k^{[1]}(n) + r_k^{[2]}(n) - \lambda_n^{[1]} - \lambda_n^{[2]}|$$
$$\leq r_k^*(n) + k!|r_k^{[1]}(n) - \lambda_n^{[1]}| + k!|r_k^{[2]}(n) - \lambda_n^{[2]}|$$
$$= r_k^*(n) + I_1 + I_2.$$

The rest of the proof of Theorem 4.4 has four parts. In the first part we give an upper estimation for I_1, in the second part we give an upper estimation for I_2, in the third part we give an upper estimation for $r_k^*(n)$, and in the last part we complete the proof of Theorem 4.4.

To estimate I_1 we will apply Theorem 1.2 so we need an upper bound for $E_1(r_k^{[1]}(n))$. To do this, it is clear from the definition of E_1 that we need the following lemma, which guarantees that every partial derivative of $r_k^{[1]}(n)$ has small expectation.

4.5 Lemma *For all non-empty multi-sets A of size at most $k-1$,*

$$E(\partial_A(r_k^{[1]}(n))) = O(n^{-a/2k^2}).$$

Proof. This can be proved similarly to Lemma 5.3 in [26]. For the sake of completeness I will present the proof. Consider a multi-set A of $k-l$ elements and $\sum_{x \in A} x = n - m$. There exists a constant $c(k)$ such that

$$\partial_A(r_k^{[1]}(n)) \leq c(k) \sum_{\substack{n^a < a_1 < a_2 < \ldots, < a_l \\ a_1 + \ldots + a_l = m}} t_{a_1} t_{a_2} \ldots t_{a_l}.$$

As $a_l \geq m/l$, and using the fact that $\sum_{x=1}^m x^{1/k-1} \approx \int_1^m z^{1/k-1} dz \approx m^{1/k}$, and (i) of Theorem 4, we have

$$E(\partial_A(r_k^{[1]}(n))) = O\left(\sum_{\substack{n^a < a_1 < a_2 < \ldots, < a_l \\ a_1 + \ldots + a_l = m}} P(a_1 \in \mathcal{A}) \ldots P(a_l \in \mathcal{A}) \right)$$

$$= O\Bigg(\sum_{\substack{n^a<a_1<a_2<\ldots,<a_l\\ a_1+\ldots+a_l=m}} g(a_1)g(a_2)\ldots g(a_l)\Bigg)$$

$$= O(\log n)\sum_{\substack{n^a<a_1<a_2<\ldots,<a_l\\ a_1+\ldots+a_l=m}} a_1^{\frac{k+1}{k^2}-1} a_2^{\frac{k+1}{k^2}-1}\ldots a_l^{\frac{k+1}{k^2}-1}$$

$$= O(\log n)O\Bigg(\Bigg(\sum_{x=1}^{m} x^{\frac{k+1}{k^2}-1}\Bigg)^{l-1}(m/l)^{\frac{k+1}{k^2}-1}\Bigg)$$

$$= O(\log n)O(m^{\frac{(l-1)(k+1)}{k^2}}(m/l)^{\frac{k+1}{k^2}-1}) = O(\log n)O(m^{\frac{l(k+1)-k^2}{k^2}}) = O(n^{-a/2k^2}),$$

since $k-1 \geq l$ and $m \geq n^a$. The proof of Lemma 4.5 is completed.

By the definition of $E_1(r_k^{[1]}(n))$, and from Lemma 4.5 it is clear that $E_1(r_k^{[1]}(n)) = max_{|A|\geq 1}E_A(r_k^{[1]}(n)) \leq cn^{-a/2k^2}$, where c is a constant. It is clear from (4.1) that $r_k^{[1]}(n)$ is a positive polynomial of degree k. Now we apply Theorem 1.2 with $\lambda = \left(\frac{\log n}{E_1(r_k^{[1]}(n))}\right)^{\frac{1}{2k}}$. If n is large enough we have

$$P\Bigg(|r_k^{[1]}(n) - \lambda_n^{[1]}| \geq d_k\sqrt{\frac{\log n}{E_1(r_k^{[1]}(n))}}\sqrt{\lambda_n^{[1]}E_1(r_k^{[1]}(n))}\Bigg) \leq$$

$$\leq b_k exp\Bigg(-\frac{1}{4}\sqrt[2k]{\frac{\log n}{E_1(r_k^{[1]}(n))}}+(k-1)\log n\Bigg) \leq b_k exp\Bigg(-\frac{1}{4}\sqrt[2k]{\frac{\log n}{n^{-a/2k^2}}}+(k-1)\log n\Bigg)$$

$$< exp(-2\log n) = \frac{1}{n^2}.$$

Applying the above result we obtain

$$\sum_{n=1}^{+\infty} P\Big(|r_k^{[1]}(n) - \lambda_n^{[1]}| \geq d_k\sqrt{\lambda_n^{[1]}\log n}\Big) < \sum_{n=1}^{+\infty}\frac{1}{n^2} < +\infty.$$

By the Borel - Cantelli lemma with probability 1, there exists a number n_0 such that

$$|r_k^{[1]}(n) - \lambda_n^{[1]}| < d_k\sqrt{\lambda_n^{[1]}\log n} \quad for \quad n > n_0. \tag{4.4}$$

In the next section we will give an upper estimation for I_2. We will prove similarly to the proof in [26] that for almost every sequence $\mathcal{A}$, there is a finite number $c_{11}(\mathcal{A})$ such that $r_k^{[2]}(n) \le c_{11}(\mathcal{A})$ for all sufficiently large n. Let $r_l(n)$ denote the number of representations of n as the sum of l distinct numbers from $\mathcal{A}$. First we give an upper estimation for $E(r_l(n))$ similarly to the estimate in [10]. Fix $2 \le l \le (k-1)$. As $n/l < a_l$, and (i) of Theorem 4.4, we have

$$E(r_l(n)) \le \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} P(a_1 \in \mathcal{A}) \ldots P(a_l \in \mathcal{A}) \tag{4.5}$$

$$< \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} g(a_1)g(a_2)\ldots g(a_l)$$

$$\le \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} \frac{(\log a_1)^{\frac{1}{k}}}{a_1^{1-\frac{k+1}{k^2}}} \ldots \frac{(\log a_l)^{\frac{1}{k}}}{a_l^{1-\frac{k+1}{k^2}}} = n^{o(1)} \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_l)^{1-\frac{k+1}{k^2}}}$$

$$\le n^{o(1)} \left(n^{\frac{k+1}{k^2}-1+o(1)} \sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_{l-1})^{1-\frac{k+1}{k^2}}} \right)$$

$$\le n^{\frac{k+1}{k^2}-1+o(1)} \sum_{\substack{1\le a_i\le n \\ i=1\ldots l-1}} \frac{1}{(a_1\ldots a_{l-1})^{1-\frac{k+1}{k^2}}}$$

$$= n^{\frac{k+1}{k^2}-1+o(1)} \Big(\sum_{1\le a_1\le n} \frac{1}{a_1^{1-\frac{k+1}{k^2}}} \Big)^{l-1}$$

$$= n^{\frac{k+1}{k^2}-1+o(1)} (n^{\frac{k+1}{k^2}+o(1)})^{l-1} = n^{-1+l\frac{k+1}{k^2}+o(1)}.$$

Let $T_1 = \{a_1, a_2, \ldots a_k\}$, $T_2 = \{b_1, b_2, \ldots b_k\}$ be two different representations of n as the sum of k terms from $\mathcal{A}$, that is, $T_1 \ne T_2$, $T_1, T_2 \subset \mathcal{A}$ and

$$a_1 + a_2 + \ldots + a_k = b_1 + b_2 + \ldots + b_k = n.$$

We say these representations are disjoint if they share no element in common. Let $f_l(n)$ denote the maximum number of pairwise disjoint representations of n as the sum of l distinct numbers from $\mathcal{A}$. We show that with probability 1, $f_l(n)$ is bounded. Let

$$\mathcal{B} = \{(a_1, \dots, a_l) : a_1 + \dots + a_l = n, a_1 \in \mathcal{A}, \dots, a_l \in \mathcal{A}, 1 \le a_1 < \dots < a_l < n\}.$$

Let $H(\mathcal{B}) = \{\mathcal{T} \subset \mathcal{B}$: all the $K \in \mathcal{T}$ are pairwise disjoint$\}$ and c_1 be a constant. It is clear that the pairwise disjointness of the sets implies the independence of the associated events, i. e., if K_1 and K_2 are pairwise disjoint representations, the events $K_1 \subset \mathcal{A}$, $K_2 \subset \mathcal{A}$ are independent. Thus by (4.5) and Lemma 1.4 we have

$$P(f_l(n) > c_1) \le P\Big(\bigcup_{\substack{\mathcal{T} \subset H(\mathcal{B}) \\ |\mathcal{T}| = c_1 + 1}} \bigcap_{K \in \mathcal{T}} K \Big) \le \sum_{\substack{\mathcal{T} \subset H(\mathcal{B}) \\ |\mathcal{T}| = c_1 + 1}} P\Big(\bigcap_{K \in \mathcal{T}} K \Big) \tag{4.6}$$

$$= \sum_{\substack{(K_1, \dots, K_{c_1+1}) \\ Pairwise \\ disjoint}} P(K_1 \cap \dots \cap K_{c_1+1}) \le \frac{1}{(c_1+1)!} (E(f_l(n)))^{c_1+1}$$

$$\le \frac{1}{(c_1+1)!} (E(r_l(n)))^{c_1+1} \le \frac{1}{(c_1+1)!} n^{-2+o(1)},$$

if c_1 large enough. By the Borel - Cantelli lemma, with probability 1 for almost every random sequence $\mathcal{A}$ there is a finite number $c_1(\mathcal{A})$ such that for any $l < k$ and all n, the maximal number of disjoint l - representations of n from $\mathcal{A}$ is at most $c_1(\mathcal{A})$. In the next step we give an upper estimation for $E(r_k^{[2]}(n))$ similarly as in Lemma 4.5. Using also the fact that $\sum_{x=1}^{m} x^{1/k-1} \approx \int_1^m z^{1/k-1} dz \approx m^{1/k}$, and $a_k \ge n/k$, $a < \frac{1}{2(k+1)}$, and (i) of Theorem 4.4, we have

$$E(r_k^{[2]}(n)) = E\left(\sum_{(a_1, a_2, \dots, a_k) \in S_n^{[2]}} t_{a_1} \dots t_{a_k} \right)$$

$$= O\left(\sum_{(a_1, a_2, \dots, a_k) \in S_n^{[2]}} P(a_1 \in \mathcal{A}) \dots P(a_k \in \mathcal{A}) \right)$$

$$= O(\log n) \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_1 \le n^a}} a_1^{\frac{k+1}{k^2}-1} a_2^{\frac{k+1}{k^2}-1} \ldots a_k^{\frac{k+1}{k^2}-1}$$

$$= O(\log n) O\left(\sum_{x=1}^{n^a} x^{\frac{k+1}{k^2}-1} \left(\sum_{x=1}^{n} x^{\frac{k+1}{k^2}-1} \right)^{k-2} (n/k)^{\frac{k+1}{k^2}-1} \right)$$

$$= O(n^{\frac{a(k+1)-1}{k^2}} \log n) = O(n^{-1/2k^2}).$$

Thus by Lemma 1.4 and the Borel - Cantelli lemma, with probability 1, there is a constant c_2 such that almost surely the maximum number of disjoint representations of n in $r_k^{[2]}(n)$ is at most c_2 for all large n. The proof is similar to (4.6). To finish the proof it suffices to show that $r_k^{[2]}(n)$ is bounded by a constant. The proof is purely combinatorial. Set $C(\mathcal{A}) = \Big(max(c_1(\mathcal{A}), c_2)\Big)^k k!$ and assume that n is sufficiently large. To each representation of n counted in $r_k^{[2]}(n)$ we assign the set formed by the k terms occuring in this representation. We will apply Lemma 1.6 with the collection of these sets in place of H. It is clear that if $r_k^{[2]}(n) > C(\mathcal{A})$, then by Lemma 1.6, $r_k^{[2]}(n)$ contains a Delta - system with $c_3 = max(c_1(\mathcal{A}), c_2) + 1$ sets. If the intersection of these sets is empty, then they form a family of c_3 disjoint k-representations of n, which contradicts the definition of c_3. Otherwise, assume that the intersection of these sets is $\{y_1, y_2 \ldots y_j\}$, where $1 \le j \le k-1$, and $\sum_{i=1}^{j} y_i = m$. Removing the common intersection of these sets we can find $c_1(\mathcal{A}) + 1$ $(k-j)$ representations of $n - m = n - \sum_{i=1}^{j} y_i$. These $c_1(\mathcal{A}) + 1$ sets are disjoint due to the definition of the Delta - system. Therefore in both cases we obtain a contradiction.

In the next section we will give an upper estimation for $r_k^*(n)$. If we collect the equal terms, we have

$$u_1 a_1 + u_2 a_2 + \ldots + u_h a_h = n, \tag{4.7}$$

where the u_i's are positive integers, and

$$u_1 + u_2 + \ldots + u_h = k. \tag{4.8}$$

Thus $r_k^*(n)$ denotes the number of representations of n in the form (4.7), where the a_i's are different. It can be proved similarly to the estimate of $r_k^{[2]}(n)$, that $r_k^*(n)$ is also bounded by a constant. For the sake of completeness we sketch the proof and we leave the details to the reader. Let $2 \le h \le k-1$ be fixed. For a fixed $u_1, \ldots, u_h$ let $s_h(n)$ denote the number of representations of n in the form (4.7). We show that $s_h(n)$ is bounded by a constant. (Note that in the previous section we proved this in the case when all u_i's are equal to one, and $h = k$). First we will give an upper estimation for $E(s_h(n))$, with a calculation similar to (4.5). Using the definition of $s_h(n)$, and $n/k < a_h$, we have

$$\begin{aligned} E(s_h(n)) &\le \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n \\ 1\le a_1<a_2<\ldots<a_h<n}} P(a_1 \in \mathcal{A})P(a_2 \in \mathcal{A}) \ldots P(a_h \in \mathcal{A}) \\ &= \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n \\ 1\le a_1<a_2<\ldots<a_h<n}} g(a_1)g(a_2)\ldots g(a_h) \\ &\le \sum_{\substack{u_1a_1+u_2a_2+\ldots+u_ha_h=n \\ 1\le a_1<a_2<\ldots<a_h<n}} \frac{(\log a_1)^{\frac{1}{k}}}{a_1^{1-\frac{k+1}{k^2}}} \cdots \frac{(\log a_h)^{\frac{1}{k}}}{a_h^{1-\frac{k+1}{k^2}}} \\ &= n^{-1+h\frac{k+1}{k^2}+o(1)}. \end{aligned} \tag{4.9}$$

Let $s_h^*(n)$ denote the size of a maximal collection of pairwise disjoint representations in the form (4.7). The same argument as in (4.6) shows that there exists a constant v_h such that for n large enough $s_h^*(n) < v_h$. In view of (4.9), and applying Lemma 1.4 we have

$$P(s_h^*(n) > v_h) < n^{-2+o(1)},$$

if v_h is large enough. Thus by the Borel - Cantelli lemma, with probability 1, $s_h^*(n) < v_h$ for every n large enough. We say that an m - tuple $(a_1, \ldots, a_m)$ $(m \leq h)$ is an m - representation of n in the form (4.7) if there is a permutation π of the numbers $\{1, 2, \ldots, h\}$ such that $\sum_{i=1}^{m} u_{\pi(i)} a_i = n$. For all $m < h$, let $s_m^*(n)$ denote the size of a maximal collection of pairwise disjoint such representations of n. The same argument as above shows that almost always there exists a constant p_m such that for every n, $s_m^*(n) < p_m$. In the last step we apply Lemma 1.6 to prove that $s_h(n)$ is bounded by a constant. Let $C = \Big(max(p_m h!, v_h)\Big)^h h!$. Let H in Lemma 1.6 is the collection of representations of n in the form (4.7). Clearly $|H| = s_h(n)$. If $s_h(n) > C$, and n is sufficiently large then by Lemma 1.6, H contains a Delta - system with $max(p_m h!, v_h) + 1$ sets. If the intersection of these sets is empty, then they form a family of disjoint h - representations in the form (4.7). Otherwise let the common intersection of the sets be $\{y_1, \ldots, y_s\}$, where $1 \leq s \leq h - 1$. By the pigeon hole principle there exists a permutation π of the numbers $\{1, 2, \ldots, h\}$ such that we can find $p_m + 1$ $(k - s)$ representations of $n'' = n - \sum_{i=1}^{s} u_{\pi(i)} y_s$. These $p_m + 1$ sets are disjoint, thus in both cases we obtain a contradiction. Since there are only finite number of partitions of k in the form (4.8), we get that $r_k^*(n)$ is bounded by a constant, i.e., there exists a constant C_3 such that $r_k^*(n) < C_3$. Let c_4, c_5, c_6 be constants. Thus by (4.3) and (4.4) we have

$$|R_k(n) - k!\lambda_n| \leq |R_k(n) - k! r_k(n)| + k!|r_k(n) - \lambda_n| < C_3 + k!|r_n^{[1]} + r_n^{[2]} - \lambda_n^{[1]} - \lambda_n^{[2]}|$$

$$\leq C_3 + k!|r_n^{[1]} - \lambda_n^{[1]}| + k!|r_n^{[2]} - \lambda_n^{[2]}| \leq C_3 + d_k k! \sqrt{\lambda_n^{[1]} \log n} + 2k! c_4$$

$$\leq c_5 + d_k k! \sqrt{\lambda_n \log n}.$$

In the last section we complete the proof of Theorem 4.4, similarly as in [7].

In view of the estimate above and (ii) in Theorem 4.4, for large n we have

$$|R_k(n) - F(n)| \le |R_k(n) - k!\lambda_n| + |k!\lambda_n - F(n)|$$

$$< c_5 + d_k k!(\lambda_n \log n)^{1/2} + |k!\lambda_n - F(n)|$$

$$\le c_5 + c_6 \left(\left(\frac{1}{k!} F(n) + \frac{1}{k!} |k!\lambda_n - F(n)| \right) \log n \right)^{1/2} + |k!\lambda_n - F(n)|$$

$$< c_5 + c_6 \left(\left(\frac{1}{k!} F(n) + \frac{c_7}{k!} (F(n) \log n)^{1/2} \right) \log n \right)^{1/2} + c_7 (F(n) \log n)^{1/2}$$

$$< c_5 + c_6 \left(\left(\frac{1}{k!} F(n) + \frac{c_7}{k!} \left(F(n) \frac{F(n)}{c_8} \right)^{1/2} \right) \log n \right)^{1/2} + c_7 (F(n) \log n)^{1/2}$$

$$= c_5 + c_6 \left(\left(\frac{1}{k!} + \frac{c_7}{\sqrt{c_8 k!}} \right) F(n) \log n \right)^{1/2} + c_7 (F(n) \log n)^{1/2} < c_9 (F(n) \log n)^{1/2}.$$

The proof of Theorem 4.4 is completed.

Chapter 5

On Sidon sets which are asymptotic bases

5.1 Introduction

A (finite or infinite) set $\mathcal{A}$ of positive integers is said to be a Sidon set if all the sums $a+b$ with $a \in \mathcal{A}$, $b \in \mathcal{A}$, $a \leq b$ are distinct. In other words $\mathcal{A}$ is a Sidon set if for every n positive integer $R_3(\mathcal{A}, n, 2) \leq 1$. We say a set $\mathcal{A} \subset \mathbb{N}$ is an asymptotic basis of order h, if every large enough positive integer n can be represented as a sum of h terms from $\mathcal{A}$, i.e., if there exists a positive integer n_0 such that $R_3(\mathcal{A}, n, h) > 0$ for $n > n_0$. In [4] and [5] P. Erdős, A. Sárközy and V. T. Sós asked if there exists a Sidon set which is an asymptotic basis of order 3. The problem was also appears in [24] (with a typo in it: order 2 is written instead of order 3). In [11] G. Grekos, L. Haddad, C. Helou and J. Pihko proved that a Sidon set cannot be an asymptotic basis of order 2. Recently J. M. Deshouillers and A. Plagne in [1] constructed a Sidon set which is an asymptotic basis of order at most 7. In this chapter I will prove that there exists an asymptotic basis of order 5 which is a Sidon set by using probabilistic methods. In fact I will prove the following theorem:

5.1 Theorem *There exists an asymptotic basis of order 5 which is a Sidon set.*

5.2 Proof of Theorem 5.1

Let $\frac{1}{5} < \alpha < \frac{3}{14}$ be real number. Define the sequence α_n in Lemma 1.1 by

$$\alpha_n = \frac{1}{n^{1-\alpha}},$$

so that $P(\{\mathcal{A}\colon \mathcal{A} \in \Omega,\, n \in \mathcal{A}\}) = \frac{1}{n^{1-\alpha}}$. The proof of Theorem 5.1 has three parts. In the first part we prove similarly as in [10] that with probability 1, $\mathcal{A}$ is asymptotic basis of order 5, i.e., with probability 1, $R_3(\mathcal{A}, n, 5) > 0$ if n is large enough. In the second part we show that deleting finitely many elements from $\mathcal{A}$ we obtain a Sidon set. Finally in the third part we prove that the above deletion does not destroy the asymptotic basis property, therefore we obtain the desired set.

Let $T_1 = \{a_1, a_2, \dots a_5\}$, $T_2 = \{b_1, b_2, \dots b_5\}$ be two different representations of n, that is $T_1 \neq T_2$, $T_1, T_2 \subset \mathcal{A}$ and

$$a_1 + a_2 + \ldots + a_5 = b_1 + b_2 + \ldots + b_5 = n.$$

We say T_1 and T_2 are disjoint if they share no element in common. To prove that $\mathcal{A}$ is asymptotic basis of order 5 we apply Theorem 1.3. We use the theorem with $Q = \mathbb{N}$. In our case t_i in Theorem 1.3 is $\varrho(\mathcal{A}, i)$. For a fixed n the sets $\{Q(\gamma)\}_{\gamma\in\Gamma}$ denote all the representations of n as the sum of 5 distinct positive integers, i.e.,

$$\{Q(\gamma)\}_{\gamma\in\Gamma} = \{(a_1, \dots, a_5) : a_1 + \ldots + a_5 = n, 1 \leq a_1 < \ldots < a_5 < n\}.$$

Thus $I_\gamma = \prod_{a_i \in Q(\gamma)} \varrho(a_i, \mathcal{A})$. In other words I_γ is the indicator variable that $Q(\gamma)$ i.e., a representation of n as the sum of 5 terms is in $\mathcal{A}$. Then it is clear that

$$N=\sum_{\gamma\in\Gamma} I_\gamma=\sum_{\gamma\in\Gamma}\prod_{a_i\in Q(\gamma)}\varrho(\mathcal{A},a_i)$$

$$=\sum_{\substack{(a_1,a_2,\dots,a_5)\in\mathbb{N}^5\\ 1\le a_1<\dots<a_5<n\\ a_1+a_2+\dots+a_5=n}}\varrho(\mathcal{A},a_1)\varrho(\mathcal{A},a_2)\dots\varrho(\mathcal{A},a_5)=r_5(n).$$

If $Q(\gamma)$, $Q(\delta)$ are two different representations of n as the sum of 5 terms and $\gamma\neq\delta$, then $\gamma\sim\delta$ implies that they have at least 1 but at most 3 common terms. It is clear that $E(I_\gamma I_\delta)=P(\{Q(\gamma)\in\mathcal{A}\}\cap\{Q(\delta)\in\mathcal{A}\})$. To apply Theorem 1.3 we have to estimate $E(r_5(n))$ and calculate Δ.

First we give lower estimation to $E(r_5(n))$. Let a be a small positive constant. By $a_5<n$, we have

$$\begin{aligned}
E(r_5(n)) &= \sum_{\substack{a_1+a_2+a_3+a_4+a_5=n\\ 1\le a_1<a_2<a_3<a_4<a_5<n}} P(a_1\in\mathcal{A})\dots P(a_5\in\mathcal{A}) \qquad (5.1)\\
&= \sum_{\substack{a_1+a_2+a_3+a_4+a_5=n\\ 1\le a_1<a_2<a_3<a_4<a_5<n}} \frac{1}{(a_1a_2a_3a_4a_5)^{1-\alpha}}\\
&\ge \sum_{\substack{a_1+a_2+a_3+a_4+a_5=n\\ n^a\le a_1<a_2<a_3<a_4<a_5<n}} \frac{1}{(a_1a_2a_3a_4a_5)^{1-\alpha}}\\
&> \frac{1}{n^{1-\alpha}}\sum_{n^a<a_1<\frac{n}{20}}\frac{1}{a_1^{1-\alpha}}\sum_{\frac{n}{20}<a_2<\frac{2n}{20}}\frac{1}{a_2^{1-\alpha}}\sum_{\frac{2n}{20}<a_3<\frac{3n}{20}}\frac{1}{a_3^{1-\alpha}}\sum_{\frac{3n}{20}<a_4<\frac{4n}{20}}\frac{1}{a_4^{1-\alpha}}\\
&= \frac{1}{n^{1-\alpha}}\Big(\int_{n^a}^{\frac{n}{20}}\frac{1}{a_1^{1-\alpha}}+O(1)\Big)\Big(\int_{\frac{n}{20}}^{\frac{2n}{20}}\frac{1}{a_2^{1-\alpha}}+O(1)\Big)\\
&\quad\times\Big(\int_{\frac{2n}{20}}^{\frac{3n}{20}}\frac{1}{a_3^{1-\alpha}}+O(1)\Big)\Big(\int_{\frac{3n}{20}}^{\frac{4n}{20}}\frac{1}{a_4^{1-\alpha}}+O(1)\Big)\\
&= \frac{1}{n^{1-\alpha}}\Big(\frac{n^\alpha}{20^\alpha\alpha}-\frac{n^{a\alpha}}{\alpha}+O(1)\Big)\Big(\frac{n^\alpha(2^\alpha-1)}{20^\alpha\alpha}+O(1)\Big)\\
&\quad\times\Big(\frac{n^\alpha(3^\alpha-2^\alpha)}{20^\alpha\alpha}+O(1)\Big)\Big(\frac{n^\alpha(4^\alpha-3^\alpha)}{20^\alpha\alpha}+O(1)\Big)
\end{aligned}$$

$$= \frac{1}{n^{1-\alpha}} n^{4\alpha}(1+o(1))c_1(1-n^{\alpha(a-1)}) > c_2 n^{5\alpha-1},$$

if n large enough, and c_1, c_2 are constants depending on α.

For $1 \le l \le 4$, denote by $r_l(n)$ the number of representations of n as the sum of l distinct numbers from $\mathcal{A}$. Let $E(r_l(n)) = \lambda_l(n)$. In the next step we give upper estimation for $E(r_l(n))$. By $n/l < a_l$, we have

$$\begin{aligned} \lambda_l(n) = E(r_l(n)) = &\sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} P(a_1 \in \mathcal{A})P(a_2 \in \mathcal{A})\ldots P(a_l \in \mathcal{A}) \\ = &\sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_l)^{1-\alpha}} \\ \le n^{-1+\alpha+o(1)} &\sum_{\substack{a_1+a_2+\ldots+a_l=n \\ 1\le a_1<a_2<\ldots<a_l<n}} \frac{1}{(a_1\ldots a_{l-1})^{1-\alpha}} \\ \le n^{-1+\alpha+o(1)} &\sum_{\substack{1\le a_i\le n \\ i=1\ldots l-1}} \frac{1}{(a_1\ldots a_{l-1})^{1-\alpha}} \\ \le n^{-1+\alpha+o(1)} &\Big(\sum_{1\le a_1\le n} \frac{1}{a_1^{1-\alpha}}\Big)^{l-1} \\ = n^{-1+\alpha+o(1)} &(n^{\alpha+o(1)})^{l-1} = n^{-1+l\alpha+o(1)}. \end{aligned} \tag{5.2}$$

Let $Q(i)$ and $Q(j)$ be two different representations of n as the sum of 5 terms. Let F_i denote the event that $Q(i) \subset \mathcal{A}$. The following lemma shows that the above events have low correlation in the following sense:

5.2 Lemma

$$\sum_{i\sim j} P(F_i \cap F_j) = o(1).$$

Proof. The proof of this lemma is similar to Lemma 11 in [10]. Note that $i \sim j$ implies that $Q(i)$ and $Q(j)$ share at least 1 number and at most 3 numbers.

$$\sum_{i\sim j} P(F_i \cap F_j) = \sum_{l=1}^{3} \sum_{|Q(i)\cap Q(j)|=l} P(F_i \cap F_j).$$

Consider $Q(i), Q(j)$ such that $|Q(i) \cap Q(j)| = l$. Say,

$$Q(i) = (z_1, \ldots, z_l, x_1, x_2, \ldots, x_{5-l})$$

and

$$Q(j) = (z_1, \ldots, z_l, y_1, y_2, \ldots, y_{5-l}).$$

Let $\sum_i z_i = m$. Then $\sum_i x_i = \sum_i y_i = n - m$. Write $P(x_i \in \mathcal{A}) = P(x_i)$. So

$$\sum_{|Q(i) \cap Q(j)| = l} P(F_i \cap F_j) =$$

$$= \sum_m \sum_{\substack{z_1 + \ldots + z_l = m \\ x_1 + \ldots + x_{5-l} = n-m \\ y_1 + \ldots + y_{5-l} = n-m}} (P(z_1) \ldots P(z_l))(P(x_1) \ldots P(x_{5-l}))(P(y_1) \ldots P(y_{5-l}))$$

$$= \sum_m \Big(\sum_{z_1 + \ldots + z_l = m} P(z_1) \ldots P(z_l) \Big) \Big(\sum_{x_1 + \ldots + x_{5-l} = n-m} P(x_1) \ldots P(x_{5-l}) \Big)^2$$

$$= \sum_m \lambda_l(m) [\lambda_{5-l}(n-m)]^2.$$

We already made the estimates in (5.2) that $\lambda_l(n) < n^{-1 + l\alpha + o(1)}$, for $1 \leq l \leq 4$. Fix $\varepsilon < 1/28$. Then there exists an m_0 such that

$$\lambda_l(m) < m^{-1 + l\alpha + \varepsilon},$$

for $m > m_0$. Since m_0 is a constant, $\lambda_l(m) < C$, where C is a constant, for $m \leq m_0$. We split the above summation in four parts:

$$\sum_m \lambda_l(m)[\lambda_{5-l}(n-m)]^2 = \sum_{m \leq m_0} \lambda_l(m)[\lambda_{5-l}(n-m)]^2 + \sum_{m_0 < m \leq n/2} \lambda_l(m)[\lambda_{5-l}(n-m)]^2 +$$

$$+ \sum_{n/2 < m \leq n - m_0} \lambda_l(m)[\lambda_{5-l}(n-m)]^2 + \sum_{n - m_0 < m} \lambda_l(m)[\lambda_{5-l}(n-m)]^2$$

$$= \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4.$$

First we estimate Δ_1:

$$\Delta_1 = \sum_{m \leq m_0} \lambda_l(m)[\lambda_{5-l}(n-m)]^2$$

$$< (n^{-1+(5-l)\alpha+o(1)})^2 \sum_{m \le m_0} C = n^{-2-2l\alpha+10\alpha+o(1)} = o(1).$$

In the next step we estimate Δ_2:

$$\Delta_2 = \sum_{m_0 < m \le n/2} \lambda_l(m)[\lambda_{5-l}(n-m)]^2$$

$$< (n^{-1+(5-l)\alpha+o(1)})^2 \sum_{m_0 < m \le n/2} m^{-1+l\alpha+\varepsilon}$$

$$= n^{-2-2l\alpha+10\alpha+o(1)} \sum_{m_0 < m \le n/2} m^{-1+l\alpha+\varepsilon}.$$

Now we estimate by integrals over the full range:

$$\Delta_2 < n^{-2-2l\alpha+10\alpha+o(1)} \Big(\int_0^n m^{-1+l\alpha+\varepsilon} dm + O(1) \Big)$$

$$= n^{-2-l\alpha+10\alpha+o(1)+\varepsilon} = o(1).$$

In the next step we estimate Δ_3:

$$\Delta_3 = \sum_{n/2 < m \le n-m_0} \lambda_l(m)[\lambda_{5-l}(n-m)]^2$$

$$< (n^{-1+l\alpha+o(1)}) \sum_{n/2 < m \le n-m_0} [(n-m)^{-1+(5-l)\alpha+\varepsilon}]^2$$

$$= (n^{-1+l\alpha+o(1)}) \sum_{n/2 < m \le n-m_0} (n-m)^{-2-2l\alpha+10\alpha+2\varepsilon}.$$

Once again estimating by integral over the full range

$$\Delta_3 < n^{-1+l\alpha+o(1)} \Big(\int_0^n (n-m)^{-2-2l\alpha+10\alpha+2\varepsilon} dm + O(1) \Big)$$

$$= n^{-2-l\alpha+10\alpha+2\varepsilon+o(1)} = o(1).$$

In the last step we estimate Δ_4:

$$\Delta_4 = \sum_{n-m_0 < m} \lambda_l(m)[\lambda_{5-l}(n-m)]^2$$

$$< (n^{-1+l\alpha+o(1)}) \sum_{n-m_0<m} C^2 = n^{-1-l\alpha+o(1)} = o(1).$$

Thus we have

$$\sum_{i\sim j} P(F_i \cap F_j) = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 = o(1).$$

The proof of Lemma 5.2 is completed.

Then it follows from Theorem 1.3 that for $0 \leq c_3 \leq 1$ constant, we have

$$P(r_5(n) \leq c_3\lambda) \leq e^{-1/2(1+\Delta)(1-c_3)^2\lambda}.$$

It follows from Lemma 5.2 that $\Delta = o(1)$. Thus in view of (5.1) it follows from Theorem 1.3 that

$$P(r_5(n) \leq c_3 E(r_5(n))) \leq e^{-1/2(1+o(1))(1-c_3)^2 c_2 n^{5\alpha-1}} < e^{-c_4 \log n},$$

where c_4 is a constant. Note that c_3 can be chosen arbitrarily small, thus if c_4 is large enough we have

$$P(r_5(n) \leq c_3 E(r_5(n))) \leq n^{-2+o(1)}.$$

Thus by (5.1) and the Borel - Cantelli lemma we get that with probability 1, there exists an $n_0 = n_0(\mathcal{A})$ such that

$$r_5(n) > c_3 n^{5\alpha-1} \quad for \quad n > n_0. \tag{5.3}$$

Let $r_k^*(n)$ denote the number of those representations of n as the sum of k terms from $\mathcal{A}$ in which there are at least two equal terms. Thus we have

$$R_3(\mathcal{A}, n, k) = k! r_k(n) + r_k^*(n).$$

It is clear that with probability 1, $R_3(\mathcal{A}, n, 5) > c_3 n^{5\alpha-1}$ for $n > n_0$ because $r_5^*(n) \geq 0$, thus $\mathcal{A}$ is asymptotic basis of order 5.

In the next section we prove similarly as in [5] that with probability 1,

$\mathcal{A}$ is almost Sidon set in the sense that it is enough to discard finitely many elements from $\mathcal{A}$ in order to get a Sidon set. It is clear from the definition that we have to prove that with probability 1, $R_3(\mathcal{A}, n, 2) \leq 1$ if n is large enough.

Let G_n denote the event

$$G_n = \{\mathcal{A} : \mathcal{A} \in \Omega, R_3(\mathcal{A}, n, 2) > 1\},$$

and write

$$\mathcal{F} = \Omega \setminus \bigcap_{j=1}^{+\infty} \Big(\bigcup_{n=j}^{+\infty} G_n \Big) \tag{5.4}$$

so that $\mathcal{A} \in \mathcal{F}$ if and only if there exists a number $n_1 = n_1(\mathcal{A})$ such that we have

$$R_3(\mathcal{A}, n, 2) \leq 1 \quad for \quad n \geq n_1. \tag{5.5}$$

We will prove that

$$P(\mathcal{F}) = 1. \tag{5.6}$$

For $1 \leq i < j \leq n/2$, let $U_n(i,j)$ denote the event

$$U_n(i,j) = \{\mathcal{A} : \mathcal{A} \in \Omega, i \in \mathcal{A}, n-i \in \mathcal{A}, j \in \mathcal{A}, n-j \in \mathcal{A}\}.$$

Then clearly,

$$G_n \subset \bigcup_{1 \leq i < j \leq n/2} U_n(i,j)$$

whence

$$P(G_n) \leq \sum_{1 \leq i < j \leq n/2} P(U_n(i,j)). \tag{5.7}$$

By (i) and (ii) in Lemma 1.1 we have

$$P(U_n(i,j)) = \begin{cases} \alpha_i \alpha_{n-i} \alpha_j \alpha_{n-j}, \text{ for } 1 \leq i < j < n/2 \\ \alpha_i \alpha_{n-i} \alpha_{n/2}, \text{ for } 1 \leq i < j = n/2. \end{cases}$$

Let $\delta_n = 1$, if n is even and $\delta_n = 0$ if n is odd. Thus we have

$$\sum_{1\le i<j\le n/2} P(U_n(i,j)) = \sum_{1\le i<j<n/2} \alpha_i\alpha_{n-i}\alpha_j\alpha_{n-j} + \delta_n\alpha_{n/2}\sum_{1\le i<n/2}\alpha_i\alpha_{n-i}$$
$$\le \Big(\sum_{1\le i<n/2}\alpha_i\alpha_{n-i} + \delta_n\alpha_{n/2}\Big)^2$$
$$\le \Big(\Big(\frac{2}{n}\Big)^{1-\alpha}\Big(\sum_{1\le i<n}\frac{1}{i^{1-\alpha}}\Big) + \delta_n\alpha_{n/2}\Big)^2$$
$$= \Big(\Big(\frac{2}{n}\Big)^{1-\alpha}\Big(\int_1^n \frac{1}{i^{1-\alpha}}di + O(1)\Big) + \delta_n\alpha_{n/2}\Big)^2$$
$$\Big(\Big(\frac{2}{n}\Big)^{1-\alpha}\Big(\Big[\frac{i^\alpha}{\alpha}\Big]_1^n + O(1)\Big) + \delta_n\alpha_{n/2}\Big)^2$$
$$= \Big(\Big(\frac{2}{n}\Big)^{1-\alpha}\Bigg(n^\alpha + O(1)\Bigg) + \delta_n\alpha_{n/2}\Big)^2$$
$$= \Big(\frac{2^{1-\alpha}}{n^{1-\alpha}}n^\alpha(1+o(1)) + \delta_n\alpha_{n/2}\Big)^2$$
$$< \Big(c_5 n^{2\alpha-1} + \delta_n 2^{1-\alpha}n^{\alpha-1}\Big)^2 < c_{15}n^{4\alpha-2}, \tag{5.8}$$

where c_{15} depends on α. By the definition of α, (5.6) and (5.7) we have

$$\sum_{n=1}^{+\infty} P(G_n) < +\infty.$$

Thus by the Borel - Cantelli lemma, with probability 1 at most a finite number of the events G_n can occur which, by (5.4), proves (5.6). By $\mathcal{A} \in \mathcal{F}$, there exists a number $n_1 = n_1(\mathcal{A})$ such that (5.5) holds. Let

$$\mathcal{C} = \mathcal{A} \cap [n_1, +\infty).$$

It follows from (5.5) that $\mathcal{C}$ is a Sidon set.
In the following lemma we estimate $r_4(n)$.

5.3 Lemma *Almost always there exists a constant $c_6 = c_6(\mathcal{A})$ such that for every n positive integer*

$$r_4(n) < c_6. \tag{5.9}$$

Proof. The proof of this lemma is similar to the proof of Lemma 10 in [10]. Let $S^{[l]}$ denote a representation of n as a sum of l distinct numbers. When $S_i^{[l]}$ and $S_j^{[l]}$ are disjoint $S_i^{[l]} \subset \mathcal{A}$ and $S_j^{[l]} \subset \mathcal{A}$ are independent events. For $2 \leq l \leq 4$, let $f_l(n)$ denote the size of a maximal collection of pairwise disjoint such representations. Let

$$\mathcal{G} = \{(a_1, \ldots, a_l) : a_1 + \ldots + a_l = n, a_1 \in \mathcal{A}, \ldots, a_l \in \mathcal{A}, 1 \leq a_1 < \ldots < a_l < n\}.$$

In view of Lemma 1.4 and (5.2) we have

$$\begin{aligned} P(f_l(n) > 10) &\leq P\Big(\bigcup_{\substack{\mathcal{T} \subset \mathcal{G} \\ |\mathcal{T}|=11}} \bigcap_{K \in \mathcal{T}} K\Big) \leq \sum_{\substack{\mathcal{T} \subset \mathcal{G} \\ |\mathcal{T}|=11}} P\Big(\bigcap_{K \in \mathcal{T}} K\Big) \\ &= \sum_{\substack{(S_1^{[l]}, \ldots, S_{11}^{[l]}) \\ Pairwise \\ disjoint}} P(S_1^{[l]} \cap \ldots \cap S_{11}^{[l]}) \\ &\leq \frac{(E(r_l(n)))^{11}}{11!} < \frac{1}{11!}(n^{-1+l\alpha+o(1)})^{11} = n^{-11+11l\alpha+o(1)}. \end{aligned}$$

By $l \leq 4$ it follows that

$$P(f_l(n) > 10) < n^{-1.1+o(1)}.$$

Thus by the Borel - Cantelli lemma the above assertion implies that almost always for $2 \leq l \leq 4$ there exists n_l such that if $n > n_l$ then $f_l(n) \leq 10$. But for any finite n_l, there are at most a finite number of representations as the sum of l numbers. Therefore, almost always for $2 \leq l \leq 4$ there exists a C_l such that for every n, $f_l(n) < C_l$. Set $C_{max} = max_l\{C_l\}$. We show that (whenever every C_l is exist), for every n

$$r_4(n) \leq (C_{max})^4 4!. \tag{5.10}$$

We prove by contradiction. Suppose (5.10) is false for some $n = n'$, i.e.,

$$r_4(n') > (C_{max})^4 4!. \tag{5.11}$$

We want to apply Lemma 1.6. Let H be the set of representations of $n^{'}$ as the sum of 4 distinct numbers from $\mathcal{A}$. Clearly $|H| = r_4(n^{'})$, thus by (5.11) and applying Lemma 1.6 we get that H contains $C_{max}+1$ representations of $n^{'}$ as the sum of 4 distinct numbers which form a Delta - system $\{S_1^4, \ldots, S_{C_{max}+1}^4\}$. If the common intersection of these sets is empty then this $C_{max}+1$ set form a family of disjoint 4 representations of $n^{'}$, which contradicts the definition of C_{max}. Otherwise let the common intersection of the system be $\{v_1, \ldots, v_r\}$, where $0 \le r \le 2$. If $\sum_i v_i = s$, then removing the common intersection each set will yield $f_{4-1-r}(n^{'} - s) \ge C_{max} + 1$. This is impossible in view of $f_l(n) < C_l$ and the definition of C_{max}. This proves (5.10), and in fact, also shows that $c_6 \le C_{max}^4 4!$. The proof of Lemma 5.3 is completed.

Now we complete the proof of Theorem 5.1. Let $\mathcal{J}$ denote the event

$$\mathcal{J} = \{\mathcal{A} : \mathcal{A} \in \Omega, \exists n_0 = n_0(\mathcal{A}), such\ that\ r_5(n) > c_3 n^{5\alpha-1}\ for\ n > n_0\}.$$

By (5.3), (5.6) we have

$$P(\mathcal{J} \cap \mathcal{F}) = 1,$$

so that $\mathcal{J} \cap \mathcal{F}$ is non - empty. Consider a set $\mathcal{A} \in \mathcal{J} \cap \mathcal{F}$. By $\mathcal{A} \in \mathcal{F}$, there exists a number $n_1 = n_1(\mathcal{A})$ such that (5.5) holds. Let

$$\mathcal{C} = \mathcal{C}(\mathcal{A}) = \mathcal{A} \cap [n_1, +\infty),$$

and $\mathcal{D} = \{u_1, \ldots, u_t\} = \mathcal{A} \setminus \mathcal{C}$. It follows from (5.5) that $\mathcal{C}$ is a Sidon set. We prove that with probability 1, $\mathcal{C}$ is an asymptotic basis of order 5, i.e., the deletion of the "small" elements of $\mathcal{A}$ does not destroy its asymptotic basis property. We prove by contradiction. Assume that with positive probability there exist infinitely many positive integers which cannot be represented as the sum of 5 numbers from $\mathcal{C}$. Choose such an M large enough. By $\mathcal{A} \in \mathcal{J}$, we have $r_5(M) > c_3(\mathcal{A}) M^{5\alpha-1}$. It follows from our assumption that every representations of M as the sum of 5 numbers from $\mathcal{A}$ contains at least one element from $\mathcal{D}$. By the pigeon hole principle there exists an $y \in \mathcal{D}$ which is in at least $\frac{r_5(M)}{t}$ representations of M. Then it follows from Lemma 5.3 that with probability 1,

$$\frac{c_2 M^{5\alpha-1}}{t} < \frac{r_5(M)}{t} \leq r_4(M-y) < c_6,$$

which is a contradiction if M is large enough.

Összefoglalás

Ebben a disszertációban additív reprezentációfüggvényekkel és Sidon - sorozatokkal foglalkozunk. Megvizsgáljuk, hogyan lehet a kéttagú összegre vonatkozó eredményeket kiterjeszteni többtagú összegekre. A dolgozat bevezető részében röviden ismertetem a szükséges definíciókat, fogalmakat, jelöléseket valamint a kezdeti eredményeket. Az első fejezetben rövid áttekintést adok az Erdős és Rényi által bevezetett valószínűségszámítási módszerről. Ez a módszer fontos szerepet játszik a disszertációban. A módszer alapjainak ismertetése után adom meg a felhasznált tételeket. A második fejezetben egy az additív reprezentációfüggvény monotonitására vonatkozó eredményt tárgyalok, amely Sárközy András egy korábbi sejtése volt. Az eredmény Erdős, Sárközy és T. Sós egy korábbi tételének kiterjesztése kéttagú összegről többtagúra. A bizonyításban a generátorfüggvény módszert használom. A harmadik fejezetben foglalkozom az additív reprezentációfüggvény differenciájának korlátosságával, itt Erdős, Sárközy és T.Sós Vera eredményeit élesítem, és terjesztem ki kéttagú összegről többtagúra. Valószínűségszámítási módszerrel bebizonyítom, hogy létezik olyan sorozat amely mutatja, hogy az ebben a fejezetben szereplő egyik eredményem lényegében a legjobb. A negyedik fejezetben Erdős és Sárközy egy tételét általánosítom többtagú összegekre, ehhez V. H. Vu tételét használom. Az ötödik fejezetben Sidon - sorozatokkal foglalkozom. Nemrégiben Deshouillers és Plagne konstruáltak olyan Sidon - sorozatot, amely hetedrendű aszimptotikus bázis. Én javítottam ezt az eredményt, és valószínúségszámítási módszerekkel, mégpedig a Janson - egyenlőtlenséget felhasználva bebizonyítom, hogy létezik olyan Sidon - sorozat, amely ötödrendű aszimptotikus bázis.

Summary

In this thesis we devoted to the additive representation functions and Sidon sequences. We extend and generalize some results of Erdős, Sárközy and V. T. Sós. In the Introduction we give a short survey about the definitions and notations. In chapter 1. we give a short survey about the probabilistic method due to Erdős and Rényi. This method plays an important role in this thesis. First I introduce the probability space we are working with, and then I give some important theorems. In chapter 2. I study the monotonicity of an additive representation function. I extend one of the results of Erdős, Sárközy and V. T. Sós, by using the generating function method. In chapter 3. I generalized and sharpen the results of Erdős, Sárközy and V. T. Sós about the boundary of the difference sequence of an additive representation function. In this chapter I also prove, that one of my result is nearly best possible by using probabilistic methods. In chapter 4. I prove that one of the results of Erdős and Sárközy about the behaviour of an additive representation function is nearly best possible by using probabilistic methods, especially the theorem of V. H. Vu. We say a set $\mathcal{A} \subset \mathbb{N}$ is an asymptotic basis of order k if every large enough positive integer can be represented as the sum of k terms from $\mathcal{A}$. We say a set $\mathcal{A} \subset \mathbb{N}$ is a Sidon set if every sum of two terms from the set $\mathcal{A}$ are different. In chapter 5. I prove the existence of Sidon sets, which are asimptotic bases of order 5. Recently Deshouillers and Plagne constructed a Sidon set which is asymptotic basis of order 7. My proof is based on the probabilistic methods especially the Janson's inequality.

Bibliography

[1] J. M. Deshouillers, A. Plagne. *A Sidon basis*, Acta Mathematica Hungarica, **123** (2009), 233-238.

[2] P. Erdős, R. Rado. *Intersection theorems for system of sets*, Journal of London Mathematical Society, **35** (1960), 85-90.

[3] P. Erdős, A. Rényi. *Additive properties of random sequences of positive integers*, Acta Arithmetica, **6** (1960), 83-110.

[4] P. Erdős, A. Sárközy, V. T. Sós. *On additive properties of general sequences*, Discrete Mathematics, **136** (1994), 75-99.

[5] P. Erdős, A. Sárközy, V. T. Sós. *On sum sets of Sidon sets I.*, Journal of Number Theory, **47** (1994), 329-347.

[6] P. Erdős, A. Sárközy. *Problems and results on additive properties of general sequences I.*, Pacific Journal, **118** (1985), 347-357.

[7] P. Erdős, A. Sárközy. *Problems and results on additive properties of general sequences II.*, Acta Mathematica Hungarica, **48** (1986), 201-211.

[8] P. Erdős, A. Sárközy, V. T. Sós. *Problems and results on additive properties of general sequences III.*, Studia Scientiarum Mathematicarum Hungarica, **22** (1987), 53-63.

[9] P. Erdős, A. Sárközy, V. T. Sós. *Problems and results on additive properties of general sequences IV.*, in: Number Theory, Proceedings, Ootacamund, India, 1984, Lecture Notes in Mathematics **1122**, Springer-Verlag, (1985), 85-104.

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

[11] G. Grekos, L. Haddad, C. Helou, J. Pihko. *Representation functions, Sidon sets and bases*, Acta Arithmetica, **130** (2007), 149-156.

[12] H. Halberstam, K. F. Roth. *Sequences*, Springer - Verlag, New York, 1983.

[13] G. Horváth. *On an additive property of sequences of nonnegative integers*, Periodica Mathematica Hungarica, **45** (2002), 73-80.

[14] S. Janson. *Poisson approximation for large deviations*, Random Structures and Algorithms, **1** (1990), 221-229.

[15] J. H. Kim, V. H. Vu. *Concentration of multivariate polynomials and its applications*, Combinatorica, **20** (2000), 417-434.

[16] S. Kiss. *Generalization of a theorem on additive representation functions*, Annales Universitas Scientiarium Budapestinesis Eötvös, **48** (2005), 15-18.

[17] S. Kiss. *On a regularity property of additive representation functions*, Periodica Mathematica Hungarica, **51** (2005), 31-35.

[18] S. Z. Kiss. *On the monotonicity of an additive representation function*, Publicationes Mathematicae Debrecen, **73** (2008), 489-495.

[19] S. Z. Kiss. *On the number of representations of integers as the sum of k terms*, Acta Arithmetica, **139** (2009), 395-406.

[20] S. Z. Kiss. *On the k - th difference of an additive representation function*, Studia Scientiarum Mathematicarum Hungarica, to appear.

[21] S. Z. Kiss. *On Sidon sets which are asymptotic bases*, Acta Mathematica Hungarica, to appear.

[22] A. Sárközy. *Paul Erdős's (1913-1996)*, Acta Arithmetica, **81** (1997), 299-317.

[23] A. Sárközy. *On the number of additive representations of integers*, in: More Sets, Graphs and Numbers, A Salute to Vera T. Sós and András Hajnal, eds. E. Győry et al., Bolyai Soc. Math. Studies, **15**, Conference on Finite and Infinite Sets, 2006, J. Bolyai Math. Soc. and Springer; 329-339.

[24] A. Sárközy. *Unsolved problems in number theory*, Periodica Mathematica Hungarica, **42** (2001), 17-35.

[25] T. Tao, V. H. Vu. *Additive Combinatorics*, Cambridge University Press, 2006.

[26] V. H. Vu. *Chernoff type bounds for sum of dependent random variables and applications in additive number theory*, Number theory for the millennium, III (Urbana, IL, 2000), 341-356, A K Peters, Natick, MA, 2002.

[27] V. H. Vu. *On the concentration of multi - variatepolynomials with small expectation*, Random Structures and Algorithms, **16** (2000), 344-363.