PROBABILISTIC AND CONSTRUCTIVE METHODS IN HARMONIC ANALYSIS AND ADDITIVE NUMBER THEORY

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MATHEMATICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Mihail N. Kolountzakis
May 1994

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Paul J. Cohen
(Principal Adviser)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Yitzhak Katznelson

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Rafe R. Mazzeo

Approved for the University Committee on Graduate Studies:

Abstract

We give several applications of the probabilistic method in harmonic analysis and additive number theory. We also give efficient constructions in place of previous probabilistic (existential) proofs.

1. Using the probabilistic method we prove that there exist integers $p_1, \ldots, p_N \geq 0$ for which
$$\left| \min_x \sum_{j=1}^{N} p_j \cos jx \right| = O(s^{1/3}),$$
as $s \to \infty$, where $s = \sum_{j=1}^{N} p_j$. This improves a result of Odlyzko who proved a similar inequality with the right hand side replaced by $O((s \log s)^{1/3})$.

2. Similarly we prove that there are frequencies $\lambda_1 < \cdots < \lambda_N \in \{1, \ldots, cN\}$, for $c = 2$, for which
$$\left| \min_x \sum_{j=1}^{N} \cos \lambda_j x \right| = O(N^{1/2})$$
and that this is impossible for smaller values of the positive constant c.

3. The previous result is used to prove easily a theorem of Erdős and Turán about the density of finite integer sequences with the property that any two elements have a different sum (B_2 sequences). We also generalize this to B_{2h} sequences (of which all sums of $2h$ elements are distinct). Some dense finite and infinite $B_2[2]$ sequences (only two pairs of elements are allowed to have the same sum) are also exhibited.

4. We prove that for any sequence of integers $n_1 \leq \ldots \leq n_N$ there is a subsequence $n_{m_1}, \ldots, n_{m_r}$ such that
$$\left| \min_x \sum_{j=1}^{r} \cos n_{m_j} x \right| \geq C \cdot N,$$

where $C > 0$ is an absolute constant. Uchiyama had previously proved this with the right hand side replaced by $C \cdot N^{1/2}$. Furthermore, our proof is constructive. We give a polynomial time algorithm for the selection of such a subsequence.

5. A set E of positive integers is called a *basis* if every positive integer can be written in at least one way as a sum of two elements of E. Using the probabilistic method, Erdős has proved the existence of such a basis E for which every positive integer x can be written as a sum of two elements of E, in at least $c_1 \log x$ and at most $c_2 \log x$ ways, where $c_1, c_2 > 0$ are absolute constants. We give an algorithm for the construction of such a basis which outputs the elements of E one by one, and which takes polynomial time to decide whether a certain integer is in E or not.

6. We employ the probabilistic method to improve on some recent results of Helm related to a conjecture of Erdős and Turán on the density of additive bases of the integers. We show that for a class of random sequences of positive integers A (which satisfy $|A \cap [1,x]| \geq C \cdot \sqrt{x}$), with probability 1, all integers in the interval $[1,N]$ can be written in at least $c_1 \log x$ and at most $c_2 \log x$ ways as a difference of elements of $A \cap [1,N^2]$. Furthermore, let m_k be a sequence of positive integers which satisfies the growth condition
$$\sum_{k=1}^{\infty} \frac{\log m_k}{\sqrt{m_k}} < \infty.$$
We show that, for the same class of random sequences and again almost surely, there is a subsequence $B \subseteq A$, $|B \cap [1,x]| \geq C \cdot \sqrt{x}$, such that, for k sufficiently large, each m_k can be written in exactly one way as a difference of two elements of B.

Acknowledgements

I want to express my deep gratitude to my advisor, Professor Paul Cohen, for all that I learned from him during the last four years of my graduate studies. The many discussions that we've had on various subjects, mathematical as well as not, have been invaluable to me. His generous encouragement during periods of frustration helped me greatly through this thesis.

I would like to thank Prof. Y. Katznelson for many useful discussions as well as for persistently running the Problem Seminar. I also want to thank Prof. R. Mazzeo for being on the reading committee of my dissertation. Profs. Joel Spencer of the Courant Institute and Andrew Odlyzko of Bell Laboratories have given me valuable advice and comments on my work, for which I am grateful.

The support of the Department of Mathematics has been great during my five years of studies. It is one of the friendliest places I have ever been.

My colleagues: Yannis Petridis, Andrew Stone, David Hurtubise, Sergei Makar-Limanov, Bryna Kra, Tatiana Toro, Maia Fraser, all have my thanks for making my life as a student a memorable experience.

One has to look back to find the people who deserve most of the blame that this thesis came to be. My parents Marilena and Nikos Kolountzakis clearly did not insist enough that I do not become a mathematician. It was too little, too late. After my father had already challenged me with problem solving for years, they tried, halfheartedly, to convince me to study a science that promised more financial security than Mathematics. Their love and support throughout my studies show that they were very pleased with their failure.

A failure which was partly due to the faculty of the Department of Mathematics at the University of Crete who managed to distract me from my Computer Science studies and turn me to one of their own. Among them: Profs. Souzana Papadopoulou, Vassilis Nestoridis, Nikos Tzanakis, and the late Stelios Pichorides, who left a clear mark on my mathematical tastes.

I owe much of what I am to Lydia Kavraki, my wife. Our parallel studies, for so many years now, would not be possible without her. Making it to this point together means to me that the end of our studies will be the greatest beginning for our life.

Αφιερωνεται στους γονεις μου και τη Λυδια.

Notations

The letter C will be used as an absolute constant. It need not represent the same constant in all its occurences, even in the same formula.
$a = O(b)$ is equivalent to $|a| \leq C|b|$, for some constant C.
$a \ll b$ means $a = O(b)$. $a \gg b$ means $b \ll a$.
$a = o(b)$ means that $a/b \to 0$ as a certain parameter, that will be implicit from the context, tends to a limit.
$\lfloor x \rfloor$ is the largest integer not larger than x and $\lceil x \rceil$ is the smallest integer not smaller than x.
x^- is defined as $\min\{0, x\}$ (the negative part of x).
(The L^p norms) $\|f\|_p = (\int |f|^p)^{1/p}$ for $1 \leq p < \infty$, and $\|f\|_\infty = \operatorname{esssup}|f|$, where the integration and the supremum are taken over the domain of definition of the function f, which will be implicit.
If E is a subset of a measure space then $|E|$ denotes its measure. If E is a finite set then $|E|$ denotes its cardinality.
We denote sequences of positive integers by capital letters and the individual elements of the sequence by lower case indexed letters:

$$A = \{a_1, a_2, \ldots\}$$

and we usually assume that $a_1 < a_2 < \cdots$. We also write

$$A(x) = |A \cap [1, x]|$$

for the counting function of the sequence A.
The Fourier coefficients are defined by

$$\widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx.$$

In particular $\widehat{f}(0) = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx$. ($f$ must be in $L^1[0, 2\pi]$.)

$\mathbf{N}$ denotes the set of natural numbers (positive integers), $\mathbf{Z}$ the integers, $\mathbf{R}$ the real numbers and $\mathbf{C}$ the complex numbers.

$\mathbf{E}[X]$ denotes the expectation of the random variable X.

$\Pr[A]$ denotes the probability of the event A, and $\Pr[A \mid B] = \Pr[A \cap B]/\Pr[B]$ denotes the conditional probability of A given B.

$\mathbf{1}(\ldots)$ is equal to 1 if the condition in the parentheses is true, otherwise it is equal to 0.

The acronym SIIRV stands for *Sum of Independent Indicator Random Variables.*

$i = \sqrt{-1}$ unless it is used as a running index.

Contents

Chapter 1

Introduction

1.1 The Prototype Average Value Argument

The *probabilistic* (*counting*) method in mathematics in its simplest form is the proof of the existence of a certain object by examining the average behavior of an appropriate collection of candidates. The prototype example of the probabilistic method in this form can be considered to be the following obvious statement.

Proposition 1.1 *If* $x_1, \ldots, x_n \in \mathbf{R}$ *and*

$$\frac{x_1 + \cdots + x_n}{n} \geq a \tag{1}$$

then for some j

$$x_j \geq a. \tag{2}$$

The usefulness of the method lies in the fact that the average (1) is often easier to compute than exhibiting a specific x_j for which (2) can be proved to hold.

This lack of the power to construct the solution to a specific problem is an inherent characteristic of the method. Very frequently the probabilistic proof of a theorem is extremely simple compared to a bare hands constructive proof, and that is to be expected since it furnishes less: the mere existence of a solution to a problem rather than the solution itself. Yet, one of the points that this thesis wants to make, is that very often a probabilistic proof can easily be turned into a construction, if one assumes the point of view, that an efficient algorithm is a construction.

Let us rephrase Proposition 1.1 in the following more useful form. The measure space Ω is equipped with a nonnegative measure dP of total mass 1, and a real random variable X on Ω is just a measurable function $X : \Omega \to \mathbf{R}$.

Proposition 1.2 *Let X be a real random variable on a probability space (Ω, dP) whose expected value*

$$\mathbf{E}[X] = \int_\Omega X(\omega) dP(\omega)$$

satisfies

$$\mathbf{E}[X] \geq a.$$

Then there is $\omega \in \Omega$ such that

$$X(\omega) \geq a.$$

We remark that because of the obvious linearity property of the expectation of a random variable

$$\mathbf{E}[\alpha X_1 + \beta X_2] = \alpha \mathbf{E}[X_1] + \beta \mathbf{E}[X_2]$$

(whenever the right hand side makes sense), the expected value of quantities of interest in the problems that follow are almost always very easy to compute or at least to estimate very well. Notice that no independence is required of the pair X_1, X_2. (For a definition of independence of a collection of random variables see for example [41].)

We proceed to give some examples.

1.1.1 An Edge-Colored K_n with Few Monochromatic K_m's

Let n, m, with $n \geq m \geq 3$, be two positive integers. We denote by K_n the complete graph on n vertices. We want to color the edges of K_n with two colors (say red and blue) so that it contains few monochromatic copies of K_m. Of course it is easy to have many monochromatic K_m's by coloring every edge with the same color. For each subset A of $[n] = \{1, \ldots, n\}$ with $|A| = m$ we define the function (of the coloring)

$$\chi_A = \begin{cases} 1 & \text{if } A \text{ is monochromatic,} \\ 0 & \text{otherwise.} \end{cases}$$

Then the number X of monochromatic K_m's is

$$X = \sum_{A \subseteq [n],\ |A|=m} \chi_A. \tag{3}$$

We color each edge of K_n red or blue with equal probability $1/2$ and independently of the other edges (we toss a fair coin for each edge). The expected value of χ_A is then $2(\frac{1}{2})^{\binom{m}{2}}$ and by the linearity of expectation and (3) we get

$$\mathbf{E}[X] = \binom{n}{m} 2^{1-\binom{m}{2}}.$$

We have proved:

Theorem 1.1 *There is a 2-coloring of the edges of the graph K_n which gives rise to no more than $\binom{n}{m}2^{1-\binom{m}{2}}$ monochromatic K_m's.*

We shall see later a way of turning this probabilistic proof into an efficient construction of such a coloring.

1.1.2 A Large Sum-Free Subset of a Given Set of Integers

A subset E of an additive group is called *sum-free* if

$$x + y \neq z, \quad \text{for all } x, y, z \in E. \tag{4}$$

The following theorem of Erdős [16], [2] has a beautiful probabilistic proof, which, to the best of my knowledge, is the only proof known to date. See also [32] for a similar, but computationally more efficient approach.

Theorem 1.2 *Let $A \subseteq \mathbf{N}$ be a set of N positive integers. Then there is a sum-free subset E of A with*

$$|E| > \frac{1}{3}N.$$

Proof: Let $A = \{n_1 < \cdots < n_N\}$ and choose any prime $p > n_N$ such that $p = 3k + 2$ for some $k \in \mathbf{N}$. View the set A as a subset of the multiplicative group of units of the field $\mathbf{Z}_p$ (the integers mod p). Write

$$S = \{k + 1, \ldots, 2k + 1\}$$

and notice that $|S| > (p - 1)/3$ and S is sum-free as a subset of $\mathbf{Z}_p$. Let t be uniformly distributed over $\mathbf{Z}_p^\times = \{1, \ldots, p - 1\}$ and write

$$X = |S \cap (t \cdot A)|,$$

where $t \cdot A = \{t \cdot n_1, \ldots, t \cdot n_N\}$ and the arithmetic is in $\mathbf{Z}_p$. Since

$$X = \sum_{j \in S} \mathbf{1}(t^{-1}j \in A)$$

and

$$\mathbf{E}[\mathbf{1}(t^{-1}j \in A)] = \frac{N}{p-1}, \quad \text{for all } j \in \mathbf{Z}_p^\times$$

($\mathbf{Z}_p^\times$ is a multiplicative group), we have

$$\mathbf{E}[X] = \frac{|S|N}{p-1} > \frac{N}{3}.$$

This implies that there is $t_0 \in \mathbf{Z}_p^\times$ for which $X > N/3$. Define then

$$E = A \cap (t_0^{-1}S).$$

It follows that E is sum-free as a set of integers (even more, it is sum-free mod p) and $|E| > N/3$, as we had to show. $\square$

1.2 The Prototype Large Deviation Argument

Often we associate several quantities $X_1, \ldots, X_n$ with a random object. Typically their averages $\mathbf{E}[X_1], \ldots, \mathbf{E}[X_n]$ will be easy to compute or estimate and their value will be in the desirable range. Our objective is to have the values of the random variables X_j themselves in that range, simultaneously for all j.

Having found a proper distribution of random objects, namely one for which the expected values $\mathbf{E}[X_j]$ are of the desirable magnitude, we still need to bound the probability that *some* X_j deviates too much from its expected value. That is, we want an upper bound on

$$\Pr\left[|X_j - \mathbf{E}[X_j]| > d_j, \text{ for some } j\right]. \tag{5}$$

The maximum allowed deviations d_j are problem dependent.

It is usually the case that the best upper bound for this probability that we know is

$$\sum_{j=1}^{n} \Pr\left[|X_j - \mathbf{E}[X_j]| > d_j\right].$$

So we aim for this sum (n can be infinite in some cases) to be less than 1. This implies that with positive probability none of the *bad events*

$$B_j = \{|X_j - \mathbf{E}[X_j]| > d_j\}, \ j = 1, \ldots, n,$$

holds. In particular there is an object for which the quantities X_j satisfy

$$\mathbf{E}[X_j] - d_j \;\leq\; X_j \;\leq\; \mathbf{E}[X_j] + d_j, \text{ for } j = 1, \ldots, n.$$

To achieve this we can use several well known Large Deviation Inequalities. The following two are straightforward to prove.

Proposition 1.3 (Markov's Inequality) *If X is any nonnegative random variable with finite expectation then for all $\alpha > 0$*

$$\Pr\left[X > \alpha \mathbf{E}[X]\right] \leq \frac{1}{\alpha}. \tag{6}$$

Proposition 1.4 (Chebyshev's Inequality) *If X is any real random variable with finite* variance $\sigma^2 = \mathbf{E}[(X - \mathbf{E}[X])^2]$ *then for all $\alpha > 0$*

$$\Pr\left[|X - \mathbf{E}[X]| > \alpha\sigma\right] \leq \frac{1}{\alpha^2}. \tag{7}$$

The inequalities of Markov and Chebyshev are rather weak in most cases, but they are applicable to virtually any random variable and this makes them very useful.

In the following theorems the random variable X is assumed to be of a special form: a sum of independent random variables.

Theorem 1.3 (Chernoff [9], [3, p. 239]) *If $X = X_1 + \cdots + X_k$, and the X_j are independent indicator random variables (that is $X_j \in \{0, 1\}$), then for all $\epsilon > 0$*

$$\Pr\left[|X - \mathbf{E}[X]| > \epsilon \mathbf{E}[X]\right] \leq 2e^{-c_\epsilon \mathbf{E}[X]},$$

where $c_\epsilon > 0$ is a function of ϵ alone

$$c_\epsilon = \min\{-\log\left(e^\epsilon (1+\epsilon)^{-(1+\epsilon)}\right), \epsilon^2/2\}.$$

We call a random variable X which, as above, is the *sum of indepenent indicator random variables* a SIIRV.

Remarks on Theorem 1.3:

1. Observe that if $X = X' + X''$, where X' and X'' are SIIRV then we have

$$\Pr\left[|X - \mathbf{E}[X]| > \epsilon \mathbf{E}[X]\right] \leq 4e^{-c_\epsilon \min\{\mathbf{E}[X'], \mathbf{E}[X'']\}}.$$

2. Since there is no dependence of the bound on k (the number of summands in X), it is easy to prove that the same bound holds for $X = \sum_{j=1}^{\infty} X_j$, provided that $\sum_{j=1}^{\infty} \mathbf{E}[X_k] < \infty$.

Theorem 1.4 [3, p. 236] *Let $p_1, \ldots, p_n \in [0,1]$ and the independent zero-mean random variables $X_1, \ldots, X_n$ have the distribution*

$$X_j = \begin{cases} 1 - p_j & \text{with probability } p_j, \\ -p_j & \text{with probability } 1 - p_j. \end{cases}$$

If $X = a_1 X_1 + \cdots + a_n X_n$ where $a_1, \ldots, a_n \in [-1,1]$ then we have for all $a > 0$

$$\Pr[|X| > a] \leq 2e^{-2a^2/n}.$$

Theorems 1.3 and 1.4 are extremely useful. In the next section we show a nice application of Theorem 1.3 to a problem in additive number theory.

1.2.1 An Asymptotic Additive Basis with Small Representation Function

A set E of positive integers is called an *asymptotic additive basis of order* 2 if the *representation* function

$$r(x) = r_E(x) = |\{(a,b) \ : \ a, b \in E \ \& \ a \leq b \ \& \ x = a + b\}|$$

is strictly positive for all sufficiently large integers x. In other words all sufficiently large x can be expressed as a sum of two elements of E. Examples of asymptotic additive bases are the set $\mathbf{N}$ itself and the set $\{1, 2, 4, 6, 8, \ldots\}$.

We are interested in bases for which the representation function is small. Notice that in the previous two examples $r(x)$ can be as large as Cx.

We present Erdős' probabilistic proof that there is an asymptotic basis of order 2 such that

$$c_1 \log x \leq r(x) \leq c_2 \log x \tag{8}$$

for all sufficiently large x. The ratio of the two absolute constants c_1 and c_2 can be made arbitrarily close to 1.

Define the probabilities

$$p_x = K \cdot \left(\frac{\log x}{x}\right)^{1/2}$$

for the values of x for which the right hand side is in $[0,1]$, otherwise let $p_x = 0$. The constant K will be determined later in the proof. We define a random set E by letting

$$\Pr[x \in E] = p_x$$

independently for all x. We show that with high probability the random set E has the claimed property (8).

Define the indicator random variables

$$\chi_j = \mathbf{1}(j \in E)$$

with mean values $\mathbf{E}[\chi_j] = p_j$. We then have

$$r(x) = \sum_{j=1}^{\lfloor x/2 \rfloor} \chi_j \chi_{x-j}$$

from which and the independence of χ_j it follows that

$$\mathbf{E}[r(x)] = \sum_{j=1}^{\lfloor x/2 \rfloor} p_j p_{x-j}. \tag{9}$$

Notice also that, for each fixed x, $r(x)$ is a SIIRV. Easy calculations on the right hand side of (9) allow the asymptotic estimate

$$\mathbf{E}[r(x)] \sim IK^2 \log x,$$

where $I = \int_0^{1/2} (s(1-s))^{-1/2} ds$. We now define the bad events

$$A_x = \{|r(x) - \mathbf{E}[r(x)]| > \frac{1}{2}\mathbf{E}[r(x)]\},\ x = 1, 2, 3, \ldots.$$

Using Theorem 1.3 we can bound

$$\Pr[A_x] \leq 2 \exp(-\frac{1}{2} c_{1/2} IK^2 \log x) = 2x^{-\alpha}$$

where $\alpha = \frac{1}{2} c_{1/2} IK^2$. All we have to do now is to choose the constant K large enough to have $\alpha > 1$. We deduce that $\sum_x \Pr[A_x]$ is a convergent series and thus there is $n_0 \in \mathbf{N}$ for which

$$\sum_{x \geq n_0} \Pr[A_x] < 1,$$

which implies that with positive probability none of the events A_x, $x \geq n_0$, holds. This in turn implies the existence of a set $E \subseteq \mathbf{N}$ such that

$$\frac{1}{2} I K^2 \log x \leq r(x) \leq \frac{3}{2} I K^2 \log x$$

for all $x \geq n_0$ which concludes the proof.

We emphasize the structure of the proof. First we defined an appropriate class of random objects (random subsets of $\mathbf{N}$). We then showed that the quantities of interest (the numbers $r(x)$, $x \in \mathbf{N}$) have expected values of the desired size. The last step was to show that, with high probability, none of the quantities of interest deviates much from its expected value.

1.3 The Method of Conditional Probabilities for Derandomization

In Section 1.1.1 we saw that there is a 2-coloring of the edges of the complete graph K_n on n vertices such that the number of monochromatic copies of K_m is at most

$$\binom{n}{m} 2^{1-\binom{m}{2}}.$$

Assume that m is fixed and our task is to produce such a coloring of K_n. Trying every possible coloring clearly takes too much time since there are $2^{\binom{n}{2}}$ possible colorings. Let us describe a very general method of *derandomizing* the randomized construction that we gave in Section 1.1.1, to get an algorithm for finding such a coloring in time polynomial in n. We keep the same notation.

Let $A_1, \ldots, A_k$, where $k = \binom{n}{m}$, be all the copies of K_m in K_n (otherwise known as m-cliques), and enumerate all edges of K_n as $e_1, \ldots, e_{\binom{n}{2}}$. Let the color of edge e_j be the random variable c_j. We are going to define the colors $a_j \in \{RED, BLUE\}$ one by one, for $j = 1, \ldots, \binom{n}{2}$. Define the events

$$R_j = R_j(a_1, \ldots, a_j) = \{(c_1, \ldots, c_{\binom{n}{2}}) : c_1 = a_1, \ldots, c_j = a_j\}.$$

R_0 is the whole probability space. Intuitevely, R_j represents our choices of colors up to the j-th color.

As in Section 1.1.1 we define the 0-1-valued random variable χ_j to indicate whether A_j is monochromatic or not. We have $X = \sum_{j=1}^{k} \chi_j$ and we already computed

$$\mathbf{E}[X] = k 2^{1-\binom{m}{2}}.$$

We are going to choose the sequence of colors $a_1, \ldots, a_{\binom{n}{2}}$ so that the function of j

$$\mathbf{E}[X \mid R_j]$$

is non-increasing. This is possible for the following general reason (the nature of the variable X is immaterial here):

$$\begin{aligned}&\mathbf{E}[X \mid R_{j-1}(a_1, \ldots, a_{j-1})] = \\ &\quad \frac{1}{2}(\mathbf{E}[X \mid R_j(a_1, \ldots, a_{j-1}, RED)] + \mathbf{E}[X \mid R_j(a_1, \ldots, a_{j-1}, BLUE)]).\end{aligned}$$

This means that at least one of the choices $a_j = RED$ or $a_j = BLUE$ will yield $\mathbf{E}[X \mid R_j] \leq \mathbf{E}[X \mid R_{j-1}]$. Which of the two choices works can be decided (here the nature of X plays a role) since we can explicitly compute

$$\mathbf{E}[X \mid R_j(a_1, \ldots, a_j)]$$

for any colors $a_1, \ldots, a_j$. This computation clearly takes time polynomial in n. We proceed like this until all colors have been fixed. Then X is completely determined

$$X = \mathbf{E}[X \mid R_{\binom{n}{2}}] \leq \cdots \leq \mathbf{E}[X \mid R_0] = \mathbf{E}[X] \leq \binom{n}{m} 2^{1-\binom{m}{2}}$$

and our coloring $a_1, \ldots, a_{\binom{n}{2}}$ is thus a solution to our problem.

The very general applicability of the previous method, the so called *method of conditional probabilities*, should be obvious. The most general context is the following. We have n independent random variables $\epsilon_1, \ldots, \epsilon_n$ which, without loss of generality, can be assumed to have the distributions

$$\epsilon_j = \begin{cases} 1 & \text{with probability } p_j, \\ 0 & \text{with probability } 1 - p_j. \end{cases}$$

We also have a certain function

$$X = X(\epsilon_1, \ldots, \epsilon_n)$$

for which we can efficiently compute

$$\mathbf{E}[X \mid \epsilon_1 = v_1, \ldots, \epsilon_m = v_m]$$

for any m and $v_1, \ldots, v_m \in \{0, 1\}$. In particular we can compute $\mathbf{E}[X] = \mu$. We can then efficiently compute an assignment

$$\epsilon_1 = v_1, \ldots, \epsilon_n = v_n, \quad v_1, \ldots, v_n \in \{0, 1\},$$

for which $X \leq \mu$. See [3, p. 223] for a more detailed description of the method as well as other methods of derandomization.

1.4 Introduction to the Problems Studied and Other Related Problems

We proceed to describe some of the problems that are studied in this thesis.

1.4.1 Littlewood's Conjecture and the Cosine Problem

Both Littlewood's Conjecture (now a theorem) and the Cosine Problem concern norms of trigonometric polynomials whose coefficients are restricted.
Littlewood's Conjecture [26]*: For any set of N distinct integers $n_1 < \cdots < n_N$ we have*

$$\left\| e^{in_1x} + \cdots + e^{in_Nx} \right\|_1 \geq C \log N \tag{10}$$

for some absolute constant C.
Above we denote by $\|f\|_1$ the L^1 norm $\frac{1}{2\pi}\int_0^{2\pi} |f(x)|dx$. It is clear that $C \log N$ is the most we can expect in the right hand side of (10) since [30, 62]

$$\left\| e^{ix} + e^{i2x} + \cdots + e^{iNx} \right\|_1 \ll \log N.$$

The Cosine Problem*: For any set of N distinct positive integers $n_1 < \cdots < n_N$ we have*

$$\left| \min_x \left(\cos n_1 x + \cdots + \cos n_N x \right) \right| \geq C\sqrt{N} \tag{11}$$

for some absolute constant C.
This conjecture was stated by Chowla [11] in 1960.

It is easy to see that $C\sqrt{N}$ is the best lower bound that one can expect in (11). Indeed let

$$g(x) = \sum_{j=1}^{\sqrt{N}} e^{i2^jx}$$

and

$$f(x) = \sum_{1 \leq j < k \leq \sqrt{N}} \cos(2^k - 2^j)x.$$

Then f is a cosine sum since all the $\binom{\sqrt{N}}{2} \sim \frac{1}{2}N$ frequencies $2^k - 2^j, 1 \leq j < k \leq \sqrt{N}$, are distinct and

$$f(x) = \frac{1}{2}\left(|g(x)|^2 - \widehat{|g|^2}(0) \right).$$

This implies that $|\min_x f(x)| \le \widehat{|g|^2}(0)/2 = \sqrt{N}/2$. Thus the right hand side in (11) is best possible, up to a multiplicative constant.

Let $f(x) = \sum_{j=1}^N \cos n_j x$ and $M = |\min_x f(x)|$. Then $M + f(x) \ge 0$ so

$$M = \|M+f\|_1 \ge \|f\|_1 - M$$

which implies

$$M \ge \frac{1}{2}\|f\|_1. \tag{12}$$

So any lower bound for $\|f\|_1$ (Littlewood's Conjecture) implies a lower bound on $|\min f|$ (Cosine Problem). Also worth mentioning is the following result of Pichorides [50]:

$$M \log M + \left\| \sum_{j=1}^N e^{in_jx} \right\|_1 \gg \log N.$$

It was proved before the Littlewood Conjecture which obviously implies it.

The first result on Littlewood's Conjecture was by Cohen [12], who proved

$$\|f\|_1 \gg \left(\frac{\log N}{\log\log N}\right)^{1/8}.$$

The exponent was improved to $1/4$ by Davenport [13] and Pichorides [48] improved it further to $1/2$. The iterated logarithm was removed by Pichorides [51, 52] and independently by Fournier [22] who gave $\log^{1/2} N$ as a lower bound. In [53] Pichorides obtained the true power of the logarithm giving $\log N/(\log\log N)^2$ as a lower bound. The conjecture was finally proved in 1980 independently by Konjagin [36] and McGehee, Pigno and Smith [45]. It is still an open problem whether the actual minimum of $\|f\|_1$ is assumed by the function

$$D_N(x) = \sum_{k=1}^N e^{ikx}.$$

See also [49] for a history of both problems up to 1976.

According to our observation (12) above every new lower bound for the Littlewood Conjecture implied a new bound for the Cosine Problem. Roth [54] was the first to work directly on the C.P. and proved

$$\left|\min_x f(x)\right| \gg \left(\frac{\log N}{\log\log N}\right)^{1/2}$$

improving the current lower bound due to the bound on the L.C. The best bound for the C.P. known today is due to Bourgain [7]:

$$\left|\min_x f(x)\right| \gg 2^{\log^\epsilon N}$$

for some $\epsilon > 0$. This is super-logarithmic but it is not a power of N yet.

It should be pointed out that the two problems are not very similar. Littlewood's Conjecture is a translation invariant problem ($\|f\|_1$ stays the same if we add a fixed integer to all n_j) but the Cosine Problem is not. In fact it is easy to see that if we translate the frequencies n_j far away so that $(3-\epsilon)n_1 \ge n_N$ then $|\min f| \ge C(\epsilon)N$. Another case where Chowla's conjecture (11) is known to hold is when the set of frequencies is sum-free [54].

The following variant of the C.P. has been studied by Odlyzko [46]. Let

$$0 \le p(x) = p_0 + p_1 \cos x + \cdots + p_n \cos nx \tag{13}$$

with $p_1, \ldots, p_n$ nonnegative integers, and define

$$M(s) = \inf_{p(0) \ge s} p_0,$$

the infimum taken over all choices of $p_j \in \mathbf{N}$. How large must $M(s)$ be? In other words we want to have polynomials as in (13) with as small p_0 as possible. Odlyzko proved that $M(s) \ll (s \log s)^{1/3}$ and we shall improve this to $M(s) \ll s^{1/3}$ in Chapter 2.

1.4.2 The Salem-Zygmund Theorem

The following theorem is often used to estimate the size of a random trigonometric polynomial

Theorem 1.5 (Salem and Zygmund [56], [29, p. 69]) *Let $f_1(x), \ldots, f_n(x)$, be trigonometric polynomials of degree at most m, and $\xi_1, \ldots, \xi_n$ be independent zero-mean random variables*

$$\xi_j = \begin{cases} 1 - p_j & \text{with probability } p_j, \\ -p_j & \text{with probability } 1 - p_j, \end{cases} \tag{14}$$

for some $p_j \in [0,1]$. Write

$$f(x) = \sum_{j=1}^n \xi_j f_j(x).$$

Then, for some $C > 0$,

$$\Pr\left[\|f\|_\infty \le C\left(\sum_{j=1}^{n} \|f_j\|_\infty^2 \log m\right)^{1/2}\right] \to 1, \quad \textit{as } m \to \infty.$$

Theorem 1.5 was used in [46] to change the coefficients of a polynomial to integers without a big loss:

Corollary 1.1 *Let $p(x) = p_0 + \sum_{j=1}^{N} p_j \cos jx$ and define the random polynomial $r(x)$ so that $p(x) + r(x)$ has always integral coefficients (except perhaps the constant coefficient):*

$$r(x) = \sum_{j=1}^{N} \xi_j \cos jx,$$

with $\xi_j = 0$ if p_j is an integer, else

$$\xi_j = \begin{cases} \lfloor p_j \rfloor - p_j & \textit{with probability } \lceil p_j \rceil - p_j, \\ \lceil p_j \rceil - p_j & \textit{with probability } p_j - \lfloor p_j \rfloor. \end{cases}$$

Then $\Pr\left(\|r\|_\infty \ll (N \log N)^{1/2}\right) \to 1$, as $N \to \infty$.

For the proof of the Salem-Zygmund theorem we need the following.

Theorem 1.6 *Let a_{ij}, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, be a matrix of complex numbers, such that $|a_{ij}| \le 1$. Let also $p_1, \ldots, p_{n_2} \in [0, 1]$ and the random variables $\xi_1, \ldots, \xi_{n_2}$ be defined as in (14). Then with probability tending to 1 as $n_1 \to \infty$*

$$\left|\sum_{j=1}^{n_2} a_{ij}\xi_j\right| \le C\sqrt{n_2 \log n_1}, \quad \textit{for all } i = 1, \ldots, n_1,$$

where C is an absolute constant.

Proof: Define

$$L_i(\xi) = \sum_{j=1}^{n_2} a_{ij}\xi_j.$$

We can clearly work on the real and imaginary parts of the linear forms L_i separately, so we assume $a_{ij} \in \mathbf{R}$. Define the bad events

$$A_i = \left\{|L_i(\xi)| > C\sqrt{n_2 \log n_1}\right\}.$$

Using Theorem 1.4 we get

$$\Pr[A_i] \le 2e^{-2C^2 n_2 \log n_1/n_2} = 2n_1^{-2C^2}.$$

Now choose the constant $C = 1$ to get

$$\Pr\left[\bigcup_{i=1}^{n_1} A_i\right] \le \sum_{i=1}^{n_1} \Pr[A_i] \le \frac{2}{n_1},$$

which concludes the proof. □

To complete the proof of the Salem-Zygmund theorem we note that it is enough to ensure that $f(x_j)$ is small for a sufficiently dense set of points $x_j \in [0, 2\pi)$.

Since f is a trigonometric polynomial of degree at most m we can use Bernstein's inequality [30, p. 12]:

$$\|f'\|_\infty \le m\|f\|_\infty.$$

Define $x_i = i\frac{2\pi}{10m}$ for $i = 1, \ldots, 10m$ and the matrix

$$a_{ij} = f_j(x_i), \quad i = 1, \ldots, 10m, \; j = 1, \ldots, n.$$

Notice that for all $i = 1, \ldots, 10m$

$$f(x_i) = \sum_{j=1}^{n} \xi_j f_j(x_i) = \sum_{j=1}^{n} \xi_j a_{ij}.$$

From this and Theorem 1.6 follows that

$$\Pr\left[|f(x_i)| \le C(n \log m)^{1/2}, \text{ for all } i\right] \to 1 \tag{15}$$

as $m \to \infty$. But the event in (15) implies that $|f(x)| \le C(n \log m)^{1/2}$ for all $x \in [0, 2\pi)$ and for a larger constant C. For assume that $|f(x_0)| = \|f\|_\infty$ and that

$$|x_k - x_0| \le \frac{2\pi}{10m}.$$

Then, using Bernstein's inequality,

$$|f(x_0) - f(x_k)| \le \frac{2\pi}{10m}\|f'\|_\infty \le \frac{2\pi}{10}|f(x_0)|$$

and, since $2\pi/10 < 1$, we get

$$\|f\|_\infty = |f(x_0)| \le C|f(x_k)| \le C(n \log m)^{1/2}.$$

Remark: Theorem 1.6 has the following constructive equivalent [3, p. 225]. We give it only in the case $\epsilon_j = \pm 1$, $n_1 = n_2$ but it holds in general.

Theorem 1.7 *Let a_{ij} be a real $n \times n$ matrix, $|a_{ij}| \le 1$. We can then find in polynomial time signs $\epsilon_1, \ldots, \epsilon_n = \pm 1$ such that for every $i = 1, \ldots, n$ we have*

$$\left|\sum_{j=1}^{n} a_{ij}\epsilon_j\right| \le C(n \log n)^{1/2},$$

where C is an absolute constant.

This of course means that the Salem-Zygmund theorem is equally effective, so that in most of the applications described below the trigonometric polynomials that are claimed to exist can actually be computed in time polynomial in the parameter of the problem.

1.4.3 The Salem-Zygmund Theorem – Applications

The Salem-Zygmund Theorem is used mostly in a form similar to Corollary 1.1 to prove the existence of trigonometric polynomials with small maximum norm and whose coefficients are restricted. Typically the situation is as follows.

We want to construct a polynomial with certain properties which satisfies certain restrictions on its coefficients. We are able to construct a polynomial which has the required properties except that its coefficients are not exactly what we want. We then add to that polynomial a random polynomial with appropriately chosen coefficients. Since our random polynomial will be small in size (by the Salem-Zygmund Theorem) it will not change the nice properties of our original polynomial by much, and at the same time will make its coefficients be what we want them to be.

We give several examples of this *random modification.*

Bourgain's Sum of Sines with Small L^∞ Norm

Our intention is to find sums of sines

$$f(x) = \sin n_1 x + \cdots + \sin n_N x$$

whose L^∞ norm is as small a function of N as possible. We present Bourgain's proof [6], [29, p. 79], that there is a set of integers $\{n_1, \ldots, n_N\}$ for which

$$\|f\|_\infty \ll N^{2/3}.$$

Notice that the best we can expect is $\|f\|_\infty \ll N^{1/2}$ since $\|f\|_\infty \ge \|f\|_2 = CN^{1/2}$.

We use the following variant of the Salem-Zygmund Theorem [29, p. 79] where the variances of the random variables ξ_j are taken into account.

Theorem 1.8 *If f_j, ξ_j are as in Theorem 1.5 and also $\|f_j\|_\infty \le 1$ and*

$$a_j^2 = \mathbf{E}[\xi_j^2] = p_j(1-p_j)$$

then

$$\Pr\left[\left\|\sum_{j=1}^n \xi_j f_j\right\|_\infty \le C(\sum_{j=1}^n a_j^2 \log m)^{1/2}\right] \to 1$$

as $m \to \infty$.

The polynomial that we are going to modify is

$$q(x) = a \sum_{j=M}^{M^2} \frac{\sin jx}{j},$$

where M is a large integer and the parameter $a > 0$ will be chosen later. It is well known [62, v. 1, p. 182] and easy to prove that there is an absolute constant A, independent of M, such that

$$\|q\|_\infty \le Aa.$$

So $q(x)$ strongly satisfies the requirement that it is a sine polynomial with small L^∞ norm. We now modify it to have coefficients 0 or 1 by adding to it a suitable random polynomial.

Define the independent, zero-mean random variables (a will be less than M)

$$\xi_j = \begin{cases} 1 - a/j & \text{with probability } a/j, \\ -a/j & \text{with probability } 1 - a/j, \end{cases}$$

for $j = M, \ldots, M^2$, with variance

$$\mathbf{E}[\xi_j^2] = \frac{a}{j}\left(1 - \frac{a}{j}\right) \le \frac{a}{j}.$$

Define also the random trigonometric polynomial

$$r(x) = \sum_{j=M}^{M^2} \xi_j \sin jx$$

and notice that the polynomial $r(x) + q(x)$ has coefficients 0 or 1:

$$r(x) + q(x) = \sin n_1 x + \cdots + \sin n_N x,$$

with $n_j \in \{M, \ldots, M^2\}$. The expected value of N is

$$\mathbf{E}[N] = \sum_{j=M}^{M^2} \Pr[\xi_j > 0] = \sum_{j=M}^{M^2} \frac{a}{j} \sim a \log M.$$

It is easy to see using either Chebyshev's inequality ($\mathbf{E}[N^2]$ can easily be estimated to be $\sim a^2 \log^2 M$) or Chernoff's inequality, that

$$\Pr\left[N \ge \frac{1}{2} a \log M\right] \to 1, \quad \text{as } N \to \infty.$$

Using Theorem 1.8 and the estimate on the variances of the ξ_j we get that, with high probability,

$$\|r\|_\infty \le C \left(\sum_{j=M}^{M^2} \frac{a}{j} \log\left(M^2\right) \right)^{1/2} \le C a^{1/2} \log M.$$

Thus with probability tending to 1 we have

$$N \ge C a \log M$$

and

$$\|r+q\|_\infty \le \|r\|_\infty + \|q\|_\infty \le C a^{1/2} \log M + C a.$$

The best choice for a is then $a = \log^2 M$ which implies $N \ge C \log^3 M$ while $\|r+q\|_\infty \le \log^2 M$, which concludes the proof.

To the best of my knowledge Bourgain's result has not been improved or proved best possible. I do not know of any other non-trivial (that is $o(N)$) upper bound for

$$\|\sin n_1 x + \cdots + \sin n_N x\|_\infty.$$

It is also not clear whether Bourgain's theorem can be made effective in polynomial time in N. This is so since we applied the Salem-Zygmund theorem on a polynomial of length exponential in N. This sparsity is unavoidable, as it can easily be seen that any sine sum with a small L^∞ norm has to be sparse (take inner product with a conjugate Dirichlet kernel).

Odlyzko's Nonnegative Cosine Polynomial

Here we give Odlyzko's proof [46] that there is a nonnegative cosine polynomial

$$0 \le p(x) = p_0 + p_1 \cos x + \cdots + p_n \cos nx$$

with $p_1, \ldots, p_n$ nonnegative integers, and such that

$$p_0 \ll (s \log s)^{1/3}$$

where

$$s = p_0 + \cdots + p_n = p(0).$$

Several similar results are presented in [46] with different restrictions on the coefficients p_j.

Consider the Fejér kernel

$$K_A(x) = \sum_{j=-A}^{A} \left(1 - \frac{|j|}{A+1}\right) e^{ijx} = 1 + 2\sum_{j=1}^{A} \left(1 - \frac{|j|}{A+1}\right) \cos jx$$

which is known to be nonnegative. Our initial polynomial will be $q(x) = \alpha K_A(x)$ where $\alpha > 1$ will be determined later. Write

$$q(x) = q_0 + q_1 \cos x + \cdots + q_A \cos Ax$$

(notice that $q_0 = \alpha$). We modify q so that it has integer coefficients by adding to it the random polynomial

$$r(x) = r_1 \cos x + \cdots + r_A \cos Ax,$$

where the independent zero-mean random variables r_j have the distribution

$$r_j = \begin{cases} \lceil q_j \rceil - q_j & \text{with probability } q_j - \lfloor q_j \rfloor, \\ \lfloor q_j \rfloor - q_j & \text{with probability } \lceil q_j \rceil - q_j. \end{cases}$$

Then the numbers $r_1 + q_1, \ldots, r_A + q_A$ are always nonnegative integers.

It follows by the Salem-Zygmund theorem that

$$\|r\|_\infty \le C(A \log A)^{1/2}$$

with high probability. Write

$$M = \left| \min_x \left((r_1 + q_1) \cos x + \cdots + (r_A + q_A) \cos Ax\right) \right|$$

and notice that

$$M \le q_0 + \|r\|_\infty \le \alpha + C(A \log A)^{1/2} \tag{16}$$

while if we write $s = q(0) + r(0)$ then (using the fact that $K_A(0) \sim A$)

$$s \ge \alpha A - C(A \log A)^{1/2} \sim \alpha A.$$

The best choice for α is the one which equates the two terms in the right hand side of (16), that is $\alpha = (A \log A)^{1/2}$. We can now check that

$$M \ll (s \log s)^{1/3},$$

as we had to show.

Körner's Flat Polynomials with Unimodular Coefficients

Littlewood [43] had asked whether there exist trigonometric polynomials

$$f(x) = \sum_{j=1}^{N} a_j e^{ijx}$$

with unimodular coefficients $|a_j| = 1, a_j \in \mathbf{C}$, such that

$$C_1 N^{1/2} \le |f(x)| \le C_2 N^{1/2}, \text{ for all } x \in [0, 2\pi), \tag{17}$$

where C_1 and C_2 are two absolute constants. If only the upper bound is required in (17) then these were known to exist even with $a_j = \pm 1$ (the Rudin-Shapiro polynomials that we shall see in Section 1.4.4).

Körner [37] proved that the answer to this question is indeed affrimative. His proof essentially consists of a random modification of a polynomial found by Byrnes which almost fits the requirements.

Theorem 1.9 (Byrnes [8], Körner [37, Lemma 4]) *If n is the square of an even integer, then there is $M \in \mathbf{N}$, absolute positive constants C_1, C_2 and complex numbers*

$$c_1, \ldots, c_{4n+2M+3}$$

such that

$$\begin{array}{ll} (i) & M = O(n^{3/4}), \\ (ii) & |c_1|, \ldots, |c_{2M+3}| \le 1, \quad |c_{4n-2M+4}|, \ldots, |c_{4n+2M+3}| \le 1, \\ (iii) & |c_{2M+4}|, \ldots, |c_{4n-2M+3}| = 1, \\ (iv) & C_1 n^{1/2} \le \left| \sum_{j=1}^{4n+2M+3} c_j e^{ijx} \right| \le C_2 n^{1/2}. \end{array}$$

The only problem with the polynomial

$$q(x) = \sum_{j=1}^{N} c_j e^{ijx}, \quad \text{with } N = 4n + 2M + 3,$$

is that $O(M) = O(N^{3/4})$ of its coefficients are of modulus at most 1, while the rest are unimodular as required. We solve this problem by adding to $q(x)$ two random polynomials

$$\begin{aligned} r(x) &= \sum_{j=1}^{2M+3} r_j e^{ijx}, \\ R(x) &= \sum_{j=4n-2M+4}^{4n+2M+3} r_j e^{ijx}, \end{aligned}$$

where the independent zero-mean random variables r_j are such that for each j in the set

$$\{1, \ldots, 2M+3\} \cup \{4n-2M+4, \ldots, 4n+2M+3\}$$

we always have

$$|r_j + c_j| = 1.$$

One way to accomplish this is by letting each r_j be equal to

$$\pm i\sqrt{1 - |c_j|^2}\, \frac{c_j}{|c_j|}$$

with equal probability. The Salem-Zygmund theorem then guarantees that

$$\|r\|_\infty + \|R\|_\infty \ll (M \log M)^{1/2} \ll N^{3/8} \log^{1/2} N,$$

with high probability. Consequently the polynomial with unimodular coefficients $r + q + R$, and of degree N, satisfies

$$(C_1 - \epsilon)N^{1/2} \le |(r + q + R)(x)| \le (C_2 + \epsilon)N^{1/2}$$

for any $\epsilon > 0$ and sufficiently large N. This solves the problem for this restricted set of N's and the general result is easy to obtain by concatenating such a polynomial and a much shorter random one (see [37]).

Using a much more elaborate argument, in which the Salem-Zygmund theorem again plays a central role, Kahane improved Körner's result.

Theorem 1.10 (Kahane [28], [29, p. 75]) *For n sufficiently large there is a trigonometric polynomial*

$$f(x) = \sum_{j=1}^{n} c_j e^{ijx}$$

with $|c_j| = 1$, which satisfies

$$|f(x)| = (1 + o(1))n^{1/2}$$

for all $x \in [0, 2\pi)$.

1.4.4 The Rudin-Shapiro Polynomials. Spencer's Theorem.

The Rudin-Shapiro sequence [55, 57], [30, p. 33], is a sequence of trigonometric polynomials of the form

$$P_m(x) = \sum_{j=0}^{2^m-1} \epsilon_j e^{ijx}, \quad \epsilon_j = \pm 1, \ m = 0, 1, 2, \ldots,$$

which have small maximum

$$\|P_m\|_\infty \le C\|P_m\|_2 = C2^{m/2},$$

for a constant C. We define the double sequence P_m, Q_m inductively as follows. For $m = 0$ we define $P_0(x) = Q_0(x) = 1$ and

$$\begin{aligned} P_{m+1}(x) &= P_m(x) + e^{i2^m x} Q_m(x), \\ Q_{m+1}(x) &= P_m(x) - e^{i2^m x} Q_m(x). \end{aligned}$$

It can be verfied that

$$|P_{m+1}(x)|^2 + |Q_{m+1}(x)|^2 = 2(|P_m(x)|^2 + |Q_m(x)|^2).$$

From this it follows immediately that for all m

$$|P_m(x)|^2 + |Q_m(x)|^2 = 2^{m+1}$$

and thus

$$\|P_m\|_\infty \le \sqrt{2} \cdot 2^{m/2}.$$

The Rudin-Shapiro polynomials have degrees which are powers of 2 but for every n we can get (by concatenating them) a polynomial of degree n, whose coefficients are ± 1 and whose maximum is bounded by a constant multiple of $\sqrt{n}$.

The Rudin-Shapiro polynomials are very explicitly constructed and their structure is perfectly clear from their simple definition. Yet, if one changes the problem just a little bit the inductive construction cannot be salvaged at all. Suppose that we want to construct polynomials of degree n, whose coefficients are 0 outside a given set $E \subseteq \{1, \ldots, n\}$ and on E their coefficients are restricted to be ± 1. Clearly the inductive definition cannot work here. Essentially the best we can do is take a random polynomial and hope for the best.

Let $r(x) = \sum_{j\in E} r_j e^{ijx}$ be a random polynomial with the independent zero-mean random variables r_j being ± 1 with equal probability. Then the Salem-Zygmund theorem gives that, with high probability,

$$\|r\|_\infty \le C(|E| \log n)^{1/2}.$$

In the particular case where $|E| \gg n$ the result is clearly suboptimal, at least in the case where E is an arithmetic progression and thus we can have Rudin-Shapiro polynomials on it. It turns out that for any E we can find signs r_j for which $\|r\|_\infty \le Cn^{1/2}$.

For this we need the following important theorem of J. Spencer which in certain cases is an improvement over plain randomization.

Theorem 1.11 (Spencer [59], Gluskin [23]) *Let a_{ij}, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$ be such that $|a_{ij}| \le 1$. Then there are signs $\epsilon_1, \ldots, \epsilon_{n_2} \in \{-1, 1\}$ such that for all i*

$$\left|\sum_{j=1}^{n_2} \epsilon_j a_{ij}\right| \le C n_1^{1/2}. \tag{18}$$

Notice that there is no dependence of the bound on n_2. Compare with Theorem 1.6.

Corollary 1.2 *Let $f_1(x), \ldots, f_n(x)$, $\|f_j\|_\infty \le C$, be trigonometric polynomials of degree at most m. Then there is a choice of signs $\epsilon_1, \ldots, \epsilon_n \in \{-1, 1\}$ such that*

$$\left\|\sum_{j=1}^{n} \epsilon_j f_j\right\|_\infty \le C m^{1/2}.$$

Proof of Corollary 1.2: For $i = 1, \ldots, 10m$, $j = 1, \ldots, n$ define $a_{ij} = f_j(x_i)$, where $x_i = i\frac{2\pi}{10m}$. Let $\epsilon_1, \ldots, \epsilon_n$ be the sequence of signs given by Theorem 1.11 for the matrix a_{ij} and write $f = \sum_{j=1}^n \epsilon_j f_j$. There is $x_0 \in [0, 2\pi]$ such that $|f(x_0)| = \|f\|_\infty$. For some k we have $|x_k - x_0| \le \frac{2\pi}{10m}$. By Bernstein's inequality, $\|f'\|_\infty \le m\|f\|_\infty$, we get

$$|f(x_0) - f(x_k)| \le \frac{2\pi}{10m}\|f'\|_\infty \le \frac{2\pi}{10}|f(x_0)|$$

which, since $2\pi/10 < 1$, implies

$$\|f\|_\infty = |f(x_0)| \le C|f(x_k)| = C\left|\sum_{j=1}^{n} \epsilon_j a_{kj}\right| \le Cm^{1/2}$$

and the proof is complete. □

Using Corollary 1.2 on the functions $f_j(x) = e^{ijx}$ for all $j \in E$ we conclude that there exist signs ϵ_j such that

$$\left\|\sum_{j\in E} \epsilon_j e^{ijx}\right\|_\infty \le C\sqrt{n}.$$

Thus on any subset E of $\{1, \ldots, n\}$ there is a polynomial with coefficients ± 1 with maximum less than $C\sqrt{n}$.

Remarks

1. The proof of Spencer's theorem uses in an essential way the fact that the numbers ϵ_j can take on the values ± 1. The proof does not work if ϵ_j can take the values $1 - p_j$ or p_j for some $p_j \in [0, 1]$, while randomization (Theorem 1.6) applies to this case too. This situation can be remedied a little by the following generalization of Spencer's theorem [59, 44].

Theorem 1.12 *Let a_{ij}, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, be such that $|a_{ij}| \le 1$. Let also $p_1, \ldots, p_{n_2} \in [0, 1]$. Then there is a choice of $\epsilon_j \in \{-p_j, 1 - p_j\}$, $j = 1, \ldots, n_2$, such that, for all i,*

$$\left|\sum_{j=1}^{n_2} \epsilon_j a_{ij}\right| \le Cn_1^{1/2}. \qquad (19)$$

We essentially prove this theorem in Chapter 2, Section 2.2, in the form of a result on trigonometric polynomials.

Despite Theorem 1.12 we do not have a generalization of the form of Theorem 1.8. That is, while Theorem 1.12 is a generalization in the nonsymmetric case, the bound that we get does not involve the "variances" of the ϵ_j's. In other words we cannot take advantage of the fact that the p_j's might be small.

An example might clarify the situation more. Suppose that we want to find a polynomial

$$f(x) = \sum_{j=1}^{N} \epsilon_j e^{ijx}$$

with small $\|f\|_\infty$ and such that $\epsilon_j \in \{-p_j, 1-p_j\}$ for all $j = 1, \ldots, N$. Then the Salem-Zygmund theorem (in the form of Theorem 1.8) gives

$$\|f\|_\infty \ll \left(\sum_{j=1}^{N} p_j(1-p_j)\log N\right)^{1/2}$$

while Spencer's theorem (in the form of Theorem 1.12) gives

$$\|f\|_\infty \ll N^{1/2}.$$

If the numbers p_j are bounded away from 0 and 1, then clearly Spencer's theorem does better, but not necessarily if

$$\sum_{j=1}^{N} p_j(1-p_j) = o(\frac{N}{\log N}).$$

It would be very interesting to know what is the best possible result when

$$\sum_{j=1}^{N} p_j(1-p_j) = o(N).$$

2. Unlike plain randomization (Theorem 1.6) it is not known whether Spencer's theorem can be efficiently derandomized. The reason for this is that the proof is not purely probabilistic. At some point in the proof the existence of two ± 1 assignments of the ϵ_j's is claimed such that all sums (for all i) $\sum_{j=1}^{n_2} a_{ij}\epsilon_j$ take approximately the same value for both assignments. No way is known to make this existential claim constructive.

1.4.5 Uchiyama's Theorem

The following theorem was proved by Uchiyama [61]. It is related to both the Littlewood Conjecture and the Cosine Problem. In Chapter 4 we shall give some constructive analogs and improvements.

Theorem 1.13 *Let $A = \{n_1 < \cdots < n_N\}$ be a set of N positive integers. Then there is a subset $E \subseteq A$ such that*

$$\left\|\sum_{j\in E} e^{ijx}\right\|_1 \geq C\sqrt{N}, \tag{20}$$

where C is a positive constant.

Proof: Let $g(x) = \sum_{j\in A} e^{ijx}$ and

$$f(x) = \sum_{j\in A} \epsilon_j e^{ijx},$$

where $\epsilon_j = \pm 1$ with equal probability and independently. By the triangle inequality it suffices to show that there is an assignment to ϵ_j that makes $\|f\|_1 \gg \sqrt{N}$. To this end we use Hölder's inequality in the form

$$\|f\|_2^2 \le \|f\|_1^{2/3} \|f\|_4^{4/3}.$$

We always have $\|f\|_2^2 = N$ and

$$|f|^4 = \left| \sum_{j,k} \epsilon_j \epsilon_k e^{i(j-k)x} \right|^2.$$

Thus writing (the indices j, k always run through A)

$$r(x) = \sum_{x=j-k} \epsilon_j \epsilon_k$$

we get

$$\|f\|_4^4 = \frac{1}{2\pi} \int_0^{2\pi} |f|^4 = \sum_{x\in\mathbf{N}} r^2(x)$$

and

$$\mathbf{E}[\|f\|_4^4] = \sum_{x\in\mathbf{N}} \mathbf{E}[r^2(x)] = \sum_{j-k=j'-k'} \mathbf{E}[\epsilon_j \epsilon_k \epsilon'_j \epsilon'_k]$$

and the only terms that will survive are those with $j = j', k = k'$, thus

$$\mathbf{E}[\|f\|_4^4] = N^2.$$

This implies the existence of an assignment of the ϵ_j such that $\|f\|_4^4 \le N^2$. Using Hölder's inequality for this assignment we get $\|f\|_1 \ge N^{1/2}$ which concludes the proof. □

A very similar proof but applied on cosine sums, instead of sums of exponentials, shows that there is always a subset $E \subseteq A$ such that

$$\left\| \sum_{j\in E} \cos jx \right\|_1 \gg \sqrt{N}.$$

Since the absolute value of the minimum of a cosine sum dominates its L^1 norm we also get

$$\left| \min_x \sum_{j\in E} \cos jx \right| \gg \sqrt{N}.$$

1.4.6 $B_h[g]$ Sets of Integers

We introduce some concepts of Additive Number Theory. Let $E = \{n_1 < n_2 < \cdots\}$ be a finite or infinite set of nonnegative integers. We define a corresponding representation function on $\mathbf{N}$

$$r(x) = r_E(x;2) = |\{(a,b) \ : \ a,b \in E \ \& \ a \le b \ \& \ x = a+b\}|. \tag{21}$$

A set E is in the class B_2 if $r(x) \le 1$ for all $x \in \mathbf{N}$. In other words all sums of the form

$$a+b, \ a,b \in E, \tag{22}$$

are distinct except for permutation of a and b. It is not hard to see that this condition is equivalent to requiring that all differences

$$a-b, \ a,b \in E, \ a \ne b, \tag{23}$$

are distinct. The terminology "Sidon set" is sometimes used to describe B_2 sets but we will avoid it since it has a rather different meaning in classical harmonic analysis.

An immediate generalization of (21) is the following

$$r_h(x) = r_E(x;h) = |\{(a_1,\ldots,a_h) \ : \ a_j \in E \ \& \ a_1 \le \cdots \le a_h \ \& \ x = a_1 + \cdots + a_h\}|. \tag{24}$$

We call a set E a B_h set if $r_h(x) \le 1$ for all $x \in \mathbf{N}$. We call it a $B_h[g]$ set if $r_h(x) \le g$ for all $x \in \mathbf{N}$. Thus a B_h set is a set of which all sums of the form

$$a_1 + \cdots + a_h, \ a_j \in E, \ a_1 \le \cdots \le a_h,$$

are distinct.

The Density of Finite $B_h[g]$ Sets

The main question that we are interested in is "How big can a $B_h[g]$ subset of $\{1,\ldots,n\}$ be?" Considerably more is known when $g = 1$ (B_h sets) and among B_h sets much more is known about B_2 sets. The reason is roughly that when $g > 1$ sets of the type $B_h[g]$ do not admit any characterisation in terms of distinct differences like (23). Even for $B_2[2]$ sets very little is known.

Define $F_h(n)$ to be the maximum size of a B_h subset of $\{1,\ldots,n\}$. Also let $F_{h,g}(n)$ be the maximum size of a $B_h[g]$ subset of the same set. It is easy to see that

$$F_{h,g}(n) \le C_{h,g} n^{1/h}. \tag{25}$$

Indeed by just counting distinct sums we can get

$$\binom{F_{g,h}(n)}{h} \leq ghn$$

which implies

$$F_{h,g}(n) \leq (ghh!)^{1/h} n^{1/h} + o(n^{1/h}).$$

The quantity of interest is the constant $C_{h,g}$ in (25).

For $h = 2, g = 1$ we can do better by counting the distinct positive differences instead of the sums. We have

$$\binom{F_2(n)}{2} \leq n$$

which gives

$$F_2(n) \leq \sqrt{2} n^{1/2} + o(n^{1/2}).$$

This is far from best possible though. Erdős and Turán [20] have proved

$$F_2(n) \leq n^{1/2} + O(n^{1/4}). \tag{26}$$

Not only is the coefficient of the major term better but the error term is non-trivial too. On the other hand it has been shown by Chowla [10] and Erdős [20, Addendum] that

$$F_2(n) \geq n^{1/2} - o(n^{1/2}). \tag{27}$$

Their method uses a result of Singer [58].

For B_4 sets Lindström [42] has proved

$$F_4(n) \leq (8n)^{1/4} + O(n^{1/8}). \tag{28}$$

In Chapter 3 we will prove that for all m

$$F_{2m}(n) \leq (m(m!))^{1/2m}\, n^{1/2m} + O(n^{1/4m}). \tag{29}$$

The coefficient of the major term that can be obtained by just counting distinct sums or differences is $(2m(m!))^{1/2m}$. Notice also that we get a non-trivial error term, and that (26) and (28) are subsumed by our result. This result has also been proved independently and by a completely different method by Jia [39]. Graham [24] has proved an analogous result for $F_{2m-1}(n)$:

$$F_{2m-1}(n) \leq (m!)^{2/(2m-1)} n^{1/(2m-1)} + O(n^{1/(4m-2)}). \tag{30}$$

The Density of Infinite $B_h[g]$ Sets

While it is possible to have a B_2 subset of $\{1,\ldots,n\}$ with about $\sqrt{n}$ elements, the following theorem of Erdős shows that the situation is quite different if we look at infinite B_2 sequences of high lower density.

Theorem 1.14 *If the sequence $\{n_1 < n_2 < \cdots\} \subseteq \mathbf{N}$ is B_2 then we have*

$$\limsup_j \frac{n_j}{j^2 \log j} \geq 0. \tag{31}$$

Thus we cannot have a (finite or infinite – the infinite sequence can be obtained from finite sequences by a diagonal argument) B_2 sequence which satisfies for all j

$$n_j \ll j^2.$$

On the other hand it is easy to see that we can have an infinite B_2 sequence with large upper density.

Theorem 1.15 *There is a B_2 sequence $\{n_1 < n_2 < \cdots\} \subseteq \mathbf{N}$ for which*

$$\liminf_j \frac{n_j}{j^2} \leq A,$$

where A is an absolute positive constant.

Erdős [60], [25, p. 88] proved Theorem 1.15 with $A = 4$ and Krückeberg [38], [25, p. 89] gave $A = 2$. This is still the best known constant, and it is not known whether it can be lowered to $A = 1$ to match the Erdős-Turán bound (26).

For a long time the B_2 sequence with the highest lower density known was the one produced by the so called greedy method. Let $n_1 = 1$ and having found $n_1, \ldots, n_k$ choose n_{k+1} to be the smallest positive integer x that is not in the set

$$\{a + b - c \ : \ a, b, c \in \{n_1, \ldots, n_k\}\}.$$

It then follows easily that the sequence n_j is B_2 and that $n_j \leq j^3$. The gap between this sequence and Theorem 1.14 still stands except for the following result of Ajtai, Komlós and Szemerédi [1].

Theorem 1.16 *There is a B_2 sequence $\{n_1 < n_2 < \cdots\} \subseteq \mathbf{N}$ such that*

$$n_j \ll \left(\frac{j^3}{\log j}\right).$$

The following theorem of Erdős and Rényi deals with dense infinite $B_2[g]$ sequences. The proof is once again probabilistic.

Theorem 1.17 (Erdős and Rényi [18]) *For every $\delta > 0$ there is an integer g and a $B_2[g]$ sequence $A = \{a_1 < a_2 < \cdots\}$ such that*

$$a_j \ll j^{2+\delta}, \tag{32}$$

for all $j > 0$.

Proof: Let $\delta \in (0,1)$ be given. Let A be a random set with

$$\Pr[x \in A] = p_x,$$

independently for all $x \in \mathbf{N}$, where

$$p_x = x^{-1/2-\delta/2}.$$

Then with high probability $A(x) \gg x^{1/2-\delta/2}$ for all x, which implies (32). Write, as usual, $\chi_j = \mathbf{1}(j \in A)$. Then we have

$$r(x) = \sum_{j=1}^{\lfloor x/2 \rfloor} \chi_j \chi_{x-j}$$

and we can estimate

$$\mathbf{E}[r(x)] \le C x^{-\delta},$$

where $C = \int_0^{1/2} (s(1-s))^{-1/2-\delta/2} ds$. Define the bad events

$$A_x = \{r(x) > g\} = \{r(x) > (\frac{g}{\mathbf{E}[r(x)]})\mathbf{E}[r(x)]\}.$$

We now use Theorem 1.3 with

$$\epsilon = \frac{g}{\mathbf{E}[r(x)]} - 1 \ge C g x^{\delta},$$

observing that

$$c_\epsilon \sim \epsilon \log \epsilon, \text{ as } \epsilon \to \infty.$$

We get

$$\Pr[A_x] \le 2e^{-2c_\epsilon \mathbf{E}[r(x)]} \le 2e^{-3\log\epsilon(\epsilon \mathbf{E}[r(x)])} = 2e^{-3g\log\epsilon},$$

and using the estimate on $\mathbf{E}[r(x)]$ we get

$$\Pr[A_x] \le C e^{-Cg\delta \log x} = C x^{-Cg\delta}.$$

Choose now $g = C/\delta$, for large enough C, to get $\Pr[A_x] \ll x^{-2}$ and thus $\sum_x \Pr[A_x] < \infty$. So there is $n_0 \in \mathbf{N}$ for which

$$\sum_{x \ge n_0} \Pr[A_x] < 1,$$

so that with positive probability none of the bad events A_x, $x \ge n_0$, holds. Now discard all elements of the random set A up to n_0 to get a $B_2[g]$ set with the desired growth. $\Box$

See also the paper [17] by Erdős for further results and conjectures on additive number theory.

Chapter 2

On Nonnegative Cosine Polynomials with Nonnegative Integral Coefficients

2.1 Introduction

We consider nonnegative cosine polynomials of the form

$$0 \le p(x) = p_0 + p_1 \cos x + p_2 \cos 2x + \cdots + p_N \cos Nx \ , \quad x \in [0, 2\pi],$$

where $p_j \ge 0$. We also write $\widehat{p}(0) = p_0$. Notice that $p(0) = \sum_{j=0}^{N} p_j$ is the maximum of $p(x)$. We are interested in estimating the size of

$$M(s) = \inf_{p(0) \ge s} \widehat{p}(0)$$

for $s \to \infty$. That is, we want to find polynomials of this form for which $p_0 = \frac{1}{2\pi} \int_0^{2\pi} p(x) dx$ is small compared to the maximum of $p(x)$.

If no more restrictions are imposed on the cosine polynomial $p(x)$ then $M(s) = 0$ for all s. This is because the Fejér kernel

$$K_A(x) = \sum_{j=-A}^{A} \left(1 - \frac{|j|}{A+1}\right) e^{ijx} = 1 + \sum_{j=1}^{A} 2 \left(1 - \frac{j}{A+1}\right) \cos jx$$

has constant coefficient 1, has $K_A(0) \gg A$ and is nonnegative.

If we restrict the coefficients $p_1, \ldots, p_N$ to be either 0 or 1 we have the classical *cosine problem* (see Section 1.4.1), about which we know that for some $\epsilon > 0$

$$2^{\log^\epsilon s} \ll M(s) \ll s^{1/2}. \tag{33}$$

The upper bound in (33) is easily proved by considering the polynomial

$$\begin{aligned} f(x) &= (\sum_1^A \cos 3^j x)^2 && (34)\\ &= A + \frac{1}{2}\sum_{j=1}^A \cos(2\cdot 3^j x) + \sum_{\substack{k,l=1\\k>l}}^A \left(\cos(3^k+3^l)x + \cos(3^k-3^l)x\right). && (35)\end{aligned}$$

All cosines in (35) have distinct frequencies. Define

$$f_1(x) = f(x) + \frac{1}{2}A - \frac{1}{2}\sum_{j=1}^A \cos(2\cdot 3^j x).$$

Then $f_1(x) \ge 0$, $f_1(0) \gg A^2$, $\widehat{f_1}(0) \ll A$ and f_1 has non-constant coefficients which are either 0 or 1. The lower bound in (33) is much harder to prove and is due to Bourgain [7]. Earlier, Roth [54] had obtained $M(s) \gg (\log s/\log\log s)^{1/2}$.

From this point on, we will study the case of $p_1, \ldots, p_N$ being arbitrary nonnegative integers. This case was studied by Odlyzko [46] who showed that

$$M(s) \ll (s\log s)^{1/3}. \tag{36}$$

See Chapter 1, p. 17, for the proof.

Odlyzko studied this problem in connection with a problem posed by Erdős and Szekeres [19]. The problem is to estimate

$$E(n) = \inf \max_{|z|=1} \left| \prod_{k=1}^n (1 - z^{a_k}) \right|,$$

where $a_1, \ldots, a_n$ may be any positive integers. The following inequality holds (see [46])

$$\log E(n) \ll M(n)\log(n) \tag{37}$$

so that Odlyzko's result implies $\log E(n) \ll n^{1/3}\log^{4/3} n$.

In this chapter we replace the random modification in Odlyzko's argument with a more careful modification, based, again, partly on randomization. We use Spencer's theorem

(Corollary 1.2) which in some cases does better than the Salem-Zygmund theorem. We show in Section (2.2) that, when $p_1, \ldots, p_N$ are restricted to be nonnegative integers, we have

$$M(s) \ll s^{1/3}.$$

By (37) this implies $\log E(n) \ll n^{1/3} \log n$. The method employed has appeared in [59, 44] and is also similar to that used by Beck [4] on a different problem, posed by Littlewood.

In Section 2.3 we give a deterministic procedure which, given a polynomial $p(x)$ with nonnegative, integral Fourier coefficients (in other words p_j is a nonnegative even integer, for $j \geq 1$) and with $\widehat{p}(0) \leq (p(0))^{\alpha}$, for some $\alpha > 0$, produces a sequence of polynomials $p = p^{(0)}, p^{(1)}, p^{(2)}, \ldots$, such that $\deg p^{(n)} \to \infty$, $p^{(n)}(0) \to \infty$ and

$$(p^{(n)})\widehat{\ }(0) \leq (p^{(n)}(0))^{\alpha}.$$

This shows $M(s) \leq C s^{1/\alpha}$, with C, α dependent on the initial p only.

This has appeared in [31].

2.2 Proof of the Inequality $M(s) \ll s^{1/3}$

Since Corollary 1.2 only allows us to choose random signs, we cannot use it directly (as we used the Salem-Zygmund theorem) to modify the coefficients of a polynomial to integers, while controlling the size of the change. In this section we show how to modify the coefficients little by little, to achieve the same result.

Let $\alpha > 0$ and define

$$a(x) = \alpha K_A(x) = \sum_{j=0}^{A} a_j \cos jx.$$

Suppose $\epsilon > 0$ and the nonnegative integer k_0 is such that for some nonnegative integers b_j

$$|a_j - b_j 2^{-k_0}| \leq \epsilon \ \text{ for all } j = 1, \ldots, A.$$

We shall define a finite sequence of polynomials

$$a^{(0)}(x) = a_0 + \sum_{j=1}^{A} b_j 2^{-k_0} \cos jx,\ a^{(1)}(x),\ \ldots\ ,\ a^{(k_0)}(x)$$

inductively, so that if

$$a^{(k)}(x) = a_0 + \sum_{j=1}^{A} a_j^{(k)} \cos jx,$$

then for each $j = 1 \ldots A$,

$$a_j^{(k)} = b_j^{(k)} 2^{k-k_0} \tag{38}$$

for some nonnegative integers $b_j^{(k)}$. We define inductively the coefficients of $a^{(k+1)}$ as follows. If $b_j^{(k)}$, $j > 0$, is even then $a_j^{(k+1)} = a_j^{(k)}$. Else define

$$a_j^{(k+1)} = a_j^{(k)} + \epsilon_j^{(k)} 2^{k-k_0}, \tag{39}$$

where $\epsilon_j^{(k)} \in \{-1, 1\}$ are such that

$$\left\| \sum_{b_j^{(k)} \text{odd}} \epsilon_j^{(k)} \cos jx \right\|_\infty \le C A^{1/2}. \tag{40}$$

The existence of the signs $\epsilon_j^{(k)}$ is guaranteed by Corollary 1.2. Notice that (39) implies the preservation of (38) by the inductive definition. We deduce from (40) that

$$\left\| a^{(k+1)} - a^{(k)} \right\|_\infty \le C 2^{k-k_0} A^{1/2}. \tag{41}$$

The polynomial $a^{(k_0)}$ has integral coefficients (except perhaps for the constant coefficient). Summing (41) we get

$$\left\| a - a^{(k_0)} \right\|_\infty \le \left\| a - a^{(0)} \right\|_\infty + \left\| a^{(0)} - a^{(k_0)} \right\|_\infty \le A\epsilon + C A^{1/2}.$$

Choose $\epsilon = 1/A$ to get

$$\left\| a - a^{(k_0)} \right\|_\infty \le C A^{1/2}.$$

On the other hand, the coefficients of a and $a^{(k_0)}$ differ by at most 1 and this implies that for the nonnegative polynomial

$$p(x) = a^{(k_0)}(x) + \left\| a - a^{(k_0)} \right\|_\infty$$

we have

$$p(0) \ \ge \ a(0) - A \ge C\alpha A - A, \tag{42}$$

$$\widehat{p}(0) \ = \ \alpha + \left\| a - a^{(k_0)} \right\|_\infty \le \alpha + C A^{1/2}. \tag{43}$$

Select $\alpha = A^{1/2}$ to get $\widehat{p}(0) \ll A^{1/2}$ and $p(0) \gg A^{3/2}$. Since p has integral coefficients, we have exhibited a polynomial that achieves $M(s) \ll s^{1/3}$, and the proof is complete.

Remark on Cosine Sums

Applying the method of the preceding proof on the coefficients of the Fejér kernel $K_A(x)$, one ends up with a nonnegative polynomial of degree at most A, which is of the form

$$p(x) = p_0 + 2\sum_{j=1}^{k} \cos \lambda_j x,$$

where $\lambda_j \in \{1, \ldots, A\}$ are distinct. We have $\|K_A - p\|_\infty \ll A^{1/2}$ which, since $p_0 = \frac{1}{2\pi}\int_0^{2\pi} p(x)dx$, implies

$$p_0 \ll A^{1/2} \text{ and } p(0) \gg A.$$

Thus p is a new example of a cosine sum that achieves the upper bound in (33). It is not as simple as the one mentioned in Section 2.1 but the spectrum of it is much denser: $\frac{1}{2}A + O(A^{1/2})$ cosines with frequencies from 1 to A.

Since the Dirichlet kernel

$$D_A(x) = \sum_{j=-A}^{A} e^{ijx} \tag{44}$$

$$= 1 + 2\sum_{j=1}^{A} \cos jx \tag{45}$$

$$= \frac{\sin{(A+\frac{1}{2})x}}{\sin \frac{x}{2}} \tag{46}$$

has a minimum asymptotically equal to $-\frac{4}{3\pi}A$, it is conceivable that one may be able to raise this number of cosines from $\frac{1}{2}A + O(A^{1/2})$ to

$$(1 - \frac{2}{3\pi})A + o(A).$$

In other words, since

$$\min_x \sum_{j=1}^{A} \cos jx = -\frac{2}{3\pi}A + o(A),$$

one must remove at least $\frac{2}{3\pi}A$ cosines from the sum, in order to make its minimum be $o(A)$, in absolute value. However we show in Section 3.2, that $\frac{1}{2}A + O(A^{1/2})$ is best possible.

2.3 The construction

Suppose we are given a polynomial $p(x) \geq 0$ of degree d, whose non-constant coefficients are even nonnegative integers, which satisfies

$$\widehat{p}(0) \leq (p(0))^{\alpha}$$

for some $\alpha > 0$. Define the sequence of nonnegative polynomials $p = p^{(1)}, p^{(2)}, p^{(3)}, \ldots$, with the recursive formula

$$p^{(k+1)}(x) = p^{(k)}((2d+3)x) \cdot p(x). \tag{47}$$

Since p has even non-constant coefficients, the Fourier coefficients of all $p^{(k)}$ are nonnegative integers. The spectrum of the first factor in (47) is supported by the multiples of $2d+3$, and that of the second factor is supported by the interval $[-d, d]$. This implies that

$$(p^{(k+1)})\widehat{\ }(0) = (p^{(k)})\widehat{\ }(0)\widehat{p}(0).$$

We obviously have $p^{(k+1)}(0) = p^{(k)}(0)p(0)$. We conclude that for all $k \geq 0$

$$(p^{(k)})\widehat{\ }(0) = (\widehat{p}(0))^k \text{ and } p^{(k)}(0) = (p(0))^k$$

and consequently

$$(p^{(k)})\widehat{\ }(0) \leq (p^{(k)}(0))^\alpha.$$

So, if s is a power of $p(0)$, we have $M(s) \leq s^\alpha$, and for any s we have $M(s) \leq Cs^\alpha$, where $C = (p(0))^\alpha$.

As an example we give

$$p(x) = 4 + 4\cos x + 2\sum_{j=2}^{10} \cos jx$$

which can be checked numerically to be positive and has constant coefficient $\widehat{p}(0) = 4$ and $p(0) = 26$. This gives $\alpha = \log 4/\log 26 = .42549\cdots$.

In view of the construction above, finding a single polynomial p with $\widehat{p}(0) \leq (p(0))^\alpha$, with $\alpha < 1/3$, will prove that the result in this Chapter is not the best possible. This example was actually found by a computer but if no more insight is gained into how these good "seed" polynomials look like, the computing time grows dramatically as we increase the degree of the polynomial.

Chapter 3

The Density of $B_h[g]$ Sequences and the Minimum of Dense Cosine Sums

3.1 Introduction

Let E be a set of integers. For any integer x we denote by $r_E(x;h)$ the number of ways x can be written as a sum of h (not necessarily distinct) elements of E. Two sums $a_1+\cdots+a_h$ and $b_1+\cdots+b_h$ are considered the same if the a_j's are a permutation of the b_j's. A set E of integers is called a $B_h[g]$ set if $r_E(x;h)\le g$ for all x. A $B_h[1]$ set is also called a B_h set.

We are interested here in the density of $B_h[g]$ sets. Considerably more is known about B_2 sets than general $B_h[g]$ sets ([25, Ch. 2] is the principal reference). The main reason for this is that a set E is B_2 if and only if all differences $x-y$, for $x,y\in E$, $x\ne y$, are distinct. Nothing similar is true for $B_2[g]$ sets, for example.

Let $F_h(n)$ be the maximum size of a B_h set contained in $\{1,\ldots,n\}$.

It is obvious that $F_h(n)\le Cn^{1/h}$ and it is the size of this constant C that we care about in this chapter. For $h=2$ Erdős and Turán [20], using a counting method, have proved

$$F_2(n)\le\sqrt{n}+O(n^{1/4}). \tag{48}$$

(The constant one obtains by just counting differences is $\sqrt{2}$.) For $h=4$ Lindström [42], using van der Corput's lemma, has proved

$$F_4(n)\le(8n)^{1/4}+O(n^{1/8}). \tag{49}$$

In the other direction it has been shown by Chowla [10] and by Erdős [20, Addendum] using a theorem of Singer [58] that

$$F_2(n) \geq \sqrt{n} - o(\sqrt{n}) \tag{50}$$

and more generally it has been proved by Bose and Chowla [5] that

$$F_h(n) \geq n^{1/h} - o(n^{1/h}). \tag{51}$$

The Cosine Problem: Chowla [11] has conjectured that for any distinct positive integers $n_1 < \cdots < n_N$

$$-\min_x \sum_{j=1}^{N} \cos n_j x \geq C\sqrt{N}. \tag{52}$$

This conjecture has remained unproved and the best result known to date is due to Bourgain [7]:

$$-\min_x \sum_{j=1}^{N} \cos n_j x \geq 2^{\log^\epsilon N}$$

for some $\epsilon > 0$.

It is easy to see that there are sequences $\{n_j\}$ for which the left hand side of (52) is bounded above by $C\sqrt{N}$. We proved in Chapter 2, p. 35, that there are very dense such sequences: we can have $n_N \leq 2N$. (This can also be proved using (50).) In Section 3.2 we prove that the density above is best possible.

In Section 3.3 we use this result to prove the following theorem.

Theorem 3.1 *Let $h = 2m \geq 2$ be an even integer. Then*

$$F_h(n) \leq \left(m(m!)^2\right)^{1/h} n^{1/h} + O(n^{1/2h}). \tag{53}$$

Theorem 3.1 contains the results of Erdős and Turán (48) and Lindström (49) as special cases. It was also proved recently by Jia [39] who used an elementary combinatorial argument.

In Sections 3.4 and 3.5 we show that allowing $g > 1$ indeed helps. We exhibit a $B_2[2]$ subset of $\{1, \ldots, n\}$ with $\sqrt{2n} + o(\sqrt{n})$ elements and an infinite $B_2[2]$ sequence $1 \leq n_1 < \cdots < n_j < \cdots$ for which

$$\liminf_j \frac{n_j}{j^2} = 1.$$

3.2 Dense Cosine Sums

It was proved in Section 2.2 that for every positive integer N there are positive integers $1 \le \lambda_1 < \cdots < \lambda_N \le 2N$ such that

$$- \min_x \sum_1^N \cos \lambda_j x \le C\sqrt{N}. \tag{54}$$

We now prove that we cannot have more dense cosine sums whose minimum is small in absolute value.

Theorem 3.2 *Let $0 \le f(x) = M + \sum_1^N \cos \lambda_j x$, with $1 \le \lambda_1 < \cdots < \lambda_N \le (2-\epsilon)N$, for some $\epsilon > 3/N$. Then*

$$M > C\epsilon^2 N. \tag{55}$$

Proof: We use the following theorem of Fejér [21]:

Theorem 3.3 *Let $p(x)$ be a nonnegative trigonometric polynomial of degree n and constant term $\widehat{p}(0) = 1$. Then $p(0) \le n+1$.*

The obvious inequality above is $p(0) \le 2n+1$. We note that Theorem 3.3 is a corollary of the well known theorem of Fejér and Riesz which states that every nonnegative trigonometric polynomial can be written as the square of the modulus of a polynomial of the same degree.

To use Theorem 3.3 we first need to "smooth" $\widehat{f}$. Define $p(x) = f(x)K_a(x) \ge 0$, where

$$K_a(x) = \sum_{j=-a}^{a} \left(1 - \frac{|j|}{(a+1)}\right) e^{ijx} \ge 0$$

is the Fejér kernel of degree a (the parameter a will be determined later). Then

$$\begin{aligned} \deg p &\le (2-\epsilon)N + a, \\ \widehat{p}(0) &= M + \frac{1}{a} \sum_{\lambda_j \le a} (a - \lambda_j), \\ p(0) &= (M+N)(a+1). \end{aligned}$$

Observe that $\widehat{p}(0) \le M + a/2$ and apply Theorem 3.3 to get

$$p(0) \le (1 + \deg\, p)\widehat{p}(0)$$

or

$$\begin{aligned} M + \frac{a}{2} &\geq \frac{(M+N)a}{(2-\epsilon)N + a + 1} \\ &\geq \frac{1}{2 - \epsilon + (a+1)/N} a. \end{aligned}$$

Let $a = \epsilon N/2$ to get

$$\begin{aligned} M &\geq \left(\frac{1}{4 - \epsilon + 2/N} - \frac{1}{4} \right) \epsilon N \\ &\geq \frac{(\epsilon - 2/N)\epsilon}{16} N \\ &\geq \frac{\epsilon^2}{3 \cdot 16} N, \end{aligned}$$

since $\epsilon > 3/N$. □

3.3 An Upper Bound for $F_h(n)$, h Even

Let $h = 2m \geq 2$ be a fixed even integer. We shall give an upper bound for the size of B_h sets contained in $\{1, \ldots, n\}$. Theorem 3.2 is the main tool. In this section C denotes an arbitrary positive constant which may depend on h only.

Let $E = \{n_1, \ldots, n_k\}$, $1 \leq n_1 < \cdots < n_k \leq n$, be a B_h set. This means that all sums $a_1 + \cdots + a_h$ with $a_j \leq a_{j+1}$ and $a_j \in E$ are distinct. Consequently the sums of the form

$$a_1 + \cdots + a_m - b_1 - \cdots - b_m$$

with

$$a_j, b_j \in E,\ a_j < a_{j+1},\ b_j < b_{j+1} \text{ and } a_i \neq b_j \tag{56}$$

are all different. Indeed, if $\sum_1^m a_j - \sum_1^m b_j = \sum_1^m a'_j - \sum_1^m b'_j$ we have

$$\sum_1^m a_j + \sum_1^m b'_j = \sum_1^m a'_j + \sum_1^m b_j$$

and, since $\{n_j\}$ is a B_h sequence, the collection of terms in the left hand side is the same as that in the right hand side. But the a_j's have been assumed different from the b_j's, so we must have $a_j = a'_j$ and similarly $b_j = b'_j$, for all j.

Define the nonnegative polynomial

$$\begin{aligned} f(x) &= \left|\sum_{j=1}^{k} e^{in_j x}\right|^h \\ &= \left(\sum_{j=1}^{k} e^{in_j x}\right)^m \left(\sum_{j=1}^{k} e^{-in_j x}\right)^m \\ &= r(x) + 2(m!)^2 \left(\sum_{a_j,\, b_j \text{ satisfy (56)}} \cos\left(\sum a_j - \sum b_j\right)x\right). \end{aligned}$$

The polynomial $r(x)$ consists of $O(k^{h-1})$ terms with coefficient 1, and thus, for some $C > 0$,

$$Ck^{h-1} + \sum_{a_j,\, b_j \text{ satisfy (56)}} \cos\left(\sum_{j=1}^{m} a_j - \sum_{j=1}^{m} b_j\right)x \geq 0.$$

Write $\lambda_1 < \cdots < \lambda_N$, $N = k^h/(2(m!)^2) - O(k^{h-1})$, for the positive sums of the form $\sum_1^m a_j - \sum_1^m b_j$ (they are all different). Using Theorem 3.2 we conclude that

$$mn \geq \lambda_N \geq \left(2 - Ck^{-1/2}\right)\left(\frac{1}{2(m!)^2}k^h - O(k^{h-1})\right). \tag{57}$$

This implies

$$k \leq (m(m!)^2)^{1/h} n^{1/h} + o(n^{1/h}).$$

The error term can also be bounded as follows. Write $k = C_1 n^{1/h} + R$, where $R \geq 0$ and $C_1 = (m(m!)^2)^{1/h}$. Then $n = ((k-R)/C_1)^h$, and substituting this in (57) and matching the second largest terms we get

$$R = O(k^{1/2}) = O(n^{1/2h}),$$

which concludes the proof of Theorem 3.1.

Theorem 3.1 improves the estimate one gets by just counting the λ_j's. Indeed, there are $N = k^h/(2(m!)^2) - O(k^{h-1})$ different λ_j's in $\{1, \ldots, mn\}$ so we only get

$$k \leq \left(2m(m!)^2\right)^{1/h} n^{1/h} + o(n^{1/h}). \tag{58}$$

3.4 Dense Finite $B_2[2]$ Sequences

As mentioned in Section 3.1, for each n there is a B_2 sequence $1 \leq n_1 < \cdots < n_k \leq n$ with $k = \sqrt{n} + o(\sqrt{n})$. In this Section we show that if one allows up to 2 sums to coincide we can have denser sequences. We do this by interleaving two dense B_2 sequences.

Theorem 3.4 *For each n there is a $B_2[2]$ set $B \subseteq \{1, \ldots, n\}$ with*

$$|B| = \sqrt{2n} + o(\sqrt{n}). \tag{59}$$

Proof: By (50) there is a B_2 set $A \subseteq \{1, \ldots, \lfloor n/2 \rfloor - 1\}$, with $|A| = \sqrt{n/2} + o(\sqrt{n})$. We shall show that the subset

$$B = 2A \cup (2A+1)$$

of $\{1, \ldots, n\}$ is $B_2[2]$ which proves the theorem.

The proof is by contradiction. Assume that we have the non-trivial relations

$$x_1 + y_1 = x_2 + y_2 = x_3 + y_3, \tag{60}$$

with $x_j, y_j \in B$ and let $z = x_1 + y_1$. Look at $x_j + y_j \bmod 2$. There are three possible patterns: $0+0$, $1+1$ and $0+1$.

If z is even then only $0+0$ and $1+1$ may appear in (60) and we have either a relation of the pattern $0+0 = 0+0$ or a relation of the pattern $1+1 = 1+1$. Both cases contradict the fact that A is B_2, the first after just dividing by 2, the second after canceling the remainders and then dividing by 2.

If z is odd then only the pattern $0+1$ appears in (60) which can be rewritten as

$$2a_1 + (2a_1' + 1) = 2a_2 + (2a_2' + 1) = 2a_3 + (2a_3' + 1) \tag{61}$$

with $a_j, a_j' \in A$. By canceling 1 and dividing by 2 we have

$$a_1 + a_1' = a_2 + a_2' = a_3 + a_3'.$$

But A is B_2 so for at least one of these relations, say the first one, we have $a_1 = a_2$ and $a_1' = a_2'$ which contradicts the fact that the first relation in (61) is non-trivial. □

Jia [40] recently improved (59). He constructed a $B_2[2]$ set $B \subseteq \{1, \ldots, n\}$ with

$$|B| = \sqrt{3n} + o(\sqrt{n}). \tag{62}$$

3.5 Infinite $B_2[2]$ Sequences with Large Upper Density

The situation is rather different for infinite $B_2[g]$ sequences. Erdős [60] has proved that there is no infinite B_2 sequence $\{n_j\}$ with $n_j = O(j^2)$. Infinity is not the problem here but the fact that we require $n_j \le Cj^2$ for all j and not just for the last one, as we did with finite sequences.

For the upper density of the sequence $\{n_j\}$ Erdős [60] proved that it is possible to have

$$\liminf_j \frac{n_j}{j^2} \le 4$$

and Krückeberg [38] later improved this to

$$\liminf_j \frac{n_j}{j^2} \le 2. \tag{63}$$

It is still unknown whether the number 2 in the right hand side of (63) can be reduced (it cannot be less than 1 by (48)).

We now show that for a $B_2[2]$ infinite sequence this is possible.

Theorem 3.5 *There is a $B_2[2]$ sequence $\{n_j\}$ with*

$$\liminf_j \frac{n_j}{j^2} = 1. \tag{64}$$

Proof: The theorem will be proved if we show that any $B_2[2]$ sequence $1 \le n_1 < \cdots < n_k$ can be extended to a sequence $1 \le n_1 < \cdots < n_k < n_{k+1} < \cdots < n_l$, such that $n_l = l^2 + o(l^2)$.

Write $A = \{n_1, \ldots, n_k\}$ and $x = n_k$. Take $B \subseteq \{2x+1, \ldots, x^4\}$ to be a B_2 set with $|B| = x^2 + o(x^2)$ (this is possible by (50)). In what follows $a_j \in A$, $b_j \in B$ and $d_j \in D$ (to be defined below).

Consider the relations of the form

$$a_1 + b_1 = a_2 + b_2. \tag{65}$$

Such a relation may be written as $a_1 - a_2 = b_2 - b_1$. But B is a B_2 set, so all differences $b_2 - b_1$ are distinct, which implies that a pair $a_1, a_2 \in A$ may appear in (65) only once. Thus there are $O(k^2) = O(x)$ of these relations which may involve $O(x)$ elements of B. Let then

$$D = \{b \in B : b \text{ does not appear in any relation of the form (65)}\} \tag{66}$$

and $E = A \cup D$. Obviously $|E| = x^2 + o(x^2)$. We show that E is a $B_2[2]$ set.

First note that the relations of the form

$$\begin{aligned} a_1 + a_2 &= a_3 + d_1 \\ a_1 + a_2 &= d_1 + d_2 \end{aligned}$$

are not possible (the left hand side is too small) and A is itself $B_2[2]$. This proves $r_E(a_1 + a_2; 2) \leq 2$ for all $a_1, a_2 \in A$.

It remains to be checked that $r_E(a_1+d_1;2) \leq 2$ and $r_E(d_1+d_2;2) \leq 2$. By passing from B to D we eliminated all relations of the form (65) and so the only remaining non-trivial relations that we have to check are of the form

$$a_1 + d_1 = d_2 + d_3. \tag{67}$$

These are indeed possible. Assume $y = a_1 + d_1 = d_2 + d_3$. We have to show that these are the only ways that y can be written as a sum of two elements of E. But this is obvious since $y = d_2' + d_3'$ is impossible (this would mean $d_2 + d_3 = d_2' + d_3'$ which contradicts D in B_2), $y = a_1' + a_2'$ is impossible because of size and $y = a_1' + d_1'$ would mean that $a_1' + d_1' = a_1 + d_1$ which we took care to eliminate in (66). $\Box$

Remark: Because of the result of Section 3.4 the previous theorem is not necessarily best possible.

Chapter 4

A Construction Related to the Cosine Problem

In a result related to the Cosine Problem (see Chapter 1, p. 10) Uchiyama [61] proved that for any sequence of N distinct positive integers $n_1 < \cdots < n_N$ there is always a subsequence $m_1, \ldots, m_r$ for which

$$-\min_x \sum_{j=1}^{r} \cos m_j x \geq C\sqrt{N}. \tag{68}$$

He actually proved the stronger statement

$$\frac{1}{2\pi}\int_0^{2\pi} \left| \sum_{j=1}^{r} \cos m_j x \right| dx \geq C\sqrt{N}. \tag{69}$$

In this chapter we improve (68). (To appear in [33].)

Theorem 4.1 *For any sequence $n_1, \ldots, n_N$ of positive integers there is a subsequence $m_1, \ldots, m_r$ such that*

$$-\min_x \sum_{j=1}^{r} \cos m_j x \geq CN. \tag{70}$$

Theorem 4.1 is an obvious corollary of the more general theorem that follows.

Theorem 4.2 *Let $w_k \geq 0$ and $w = \sum_1^\infty w_k < \infty$. Then there is a set E of positive integers for which*

$$-\min_x \sum_{k\in E} w_k \cos kx \geq Cw. \tag{71}$$

The essential content of this chapter is that the proof of Theorem 4.2 (and consequently of Theorem 4.1) we give is constructive. Indeed there is a simple non-constructive proof of our theorem.

Proof of Theorem 4.2 – Non-constructive: (Odlyzko [46]) Define

$$f(x) = \sum_1^\infty w_k (\cos kx)^-.$$

Then

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)dx = \sum_{k=1}^\infty w_k \frac{1}{2\pi}\int_0^{2\pi}(\cos kx)^- dx \tag{72}$$

$$= -\frac{1}{\pi}w. \tag{73}$$

Thus there is $x_0 \in [0, 2\pi)$ with $f(x_0) \le -\frac{1}{\pi}w$. Let $E = \{k \in \mathbf{N} : \cos kx_0 \le 0\}$. Then obviously

$$\sum_{k\in E} \cos kx_0 \le -\frac{1}{\pi}w.$$

$\square$

We now give a constructive proof of Theorem 4.2 with a worse constant. (See Remark 1 after the proof for the exact meaning of the word "constructive".)

We shall need two lemmas.

Lemma 4.1 *Let $I_k = (a_k, b_k) \subseteq (0,1)$, $k = 1, 2, \ldots$, be intervals of length at least $\theta > 0$ and w_k be nonnegative weights associated with them. Let also $w = \sum_1^\infty w_k < \infty$. Then there is an interval $J \subseteq (0,1)$, with $|J| = \theta/2$, for which*

$$\sum_{J\subseteq I_k} w_k \ge \frac{1}{2}\theta w. \tag{74}$$

Proof of Lemma 4.1: Let $m = \lfloor 2/\theta \rfloor$ and $J_\nu = [\nu\theta/2, (\nu+1)\theta/2)$, for $\nu = 0, 1, \ldots, m-1$. Write also $s_\nu = \sum_{a_k \in J_\nu} w_k$. Since $w = \sum_0^{m-2} s_\nu$ there is some $\nu_0 \le m-2$ for which

$$s_{\nu_0} \ge \frac{w}{m-1} \ge \frac{1}{2}\theta w.$$

Let $J = J_{\nu_0+1}$. Then J satisfies (74) since $a_k \in J_{\nu_0}$ implies $J_{\nu_0+1} \subseteq I_k$. $\square$

The following lemma is a useful special case of Theorem 4.2.

Lemma 4.2 *Let $a > 0$, $\sigma > 1$, $\rho \ge 24\sigma$,*

$$E'_j = [\rho^j a, \sigma\rho^j a) \cap \mathbf{N},$$

for $j = 0, 1, 2, \ldots$, and

$$E' = \bigcup_{j=0}^{\infty} E'_j.$$

Assume also $w_k \ge 0$, $w = \sum_1^\infty w_k < \infty$ and $w_k = 0$ outside E'. Then there is a set $E \subseteq E'$ for which

$$-\min_x \sum_{k\in E} w_k \cos kx \ge \frac{1}{48\sigma} w. \tag{75}$$

Proof of Lemma 4.2: First observe that in any interval of length at least $2\pi/k$ there is a subinterval of length $2\pi/(12k)$ in which $\cos kx \le -1/2$. According to this observation, for all $k \in E'_0$ there is an interval I_k contained in $(0, 2\pi/a)$, of length at least $2\pi/(12\sigma a)$, in which $\cos kx \le -1/2$. By Lemma 4.1 ($\theta = 1/(12\sigma)$) there is an interval $J_0 \subseteq (0, 2\pi/a)$ of length $2\pi/(24\sigma a)$ for which

$$\sum_{J_0 \subseteq I_k} w_k \ge \frac{1}{24\sigma} \sum_{k\in E'_0} w_k. \tag{76}$$

Let $E_0 = \{k \in E'_0 : J_0 \subseteq I_k\}$. Then

$$\sum_{k\in E_0} w_k \cos kx \le -\frac{1}{48\sigma} \sum_{k\in E'_0} w_k, \text{ for all } x \in J_0. \tag{77}$$

Similarly we can find an interval $J_1 \subseteq J_0$, with $|J_1| = 2\pi/(12\sigma\rho a)$, and $E_1 \subseteq E'_1$, such that

$$\sum_{k\in E_1} w_k \cos kx \le -\frac{1}{48\sigma} \sum_{k\in E'_1} w_k, \text{ for all } x \in J_1.$$

This is possible since $\rho \ge 24\sigma$ and therefore J_0 is big enough to accomodate all frequencies in E'_1. In the same fashion we define $J_2 \supseteq J_3 \supseteq \ldots$, and $E_2, E_3, \ldots$. Finally we set $E = \bigcup_0^\infty E_j$. It follows that (75) is true. □

We can now complete the proof of the theorem.

Proof of Theorem 4.2 – Constructive: Let $\sigma = 2$, $\rho = 64$ and write for $\nu = 0, \ldots, 5$

$$A_\nu = \bigcup_{j=0}^{\infty} [\rho^j \sigma^\nu, \rho^j \sigma^{\nu+1}) \cap \mathbf{N}.$$

Since $\mathbf{N} = \bigcup_0^5 A_\nu$ there is some ν_0 for which

$$\sum_{k\in A_{\nu_0}} w_k \ge \frac{1}{6} w. \tag{78}$$

An application of Lemma 4.2 with $\sigma = 2$, $\rho = 64$, $a = 1$ and the collection of w_k for $k \in A_{\nu_0}$ furnishes a set $E \subseteq A_{\nu_0}$ for which

$$-\min_x \sum_{k \in E} w_k \cos kx \geq \frac{1}{6 \cdot 48 \cdot 2} w.$$

$\square$

Remarks

1. The simple proof of Theorem 4.1 mentioned can of course be made constructive by looking for an x that satisfies

$$\sum_1^N (\cos n_k x)^- \leq -\frac{1}{2\pi} N$$

among the points $x_j = jh$, for $j = 0, \ldots, \lfloor 1/h \rfloor$. But h has to be smaller than Cn_N^{-1} and this leads to an algorithm which in the worst case takes time exponential in the size of the input (which is considered to be the number of binary digits required to write down all $n_1, \ldots, n_N$). For example if $n_N = 2^N$ then the algorithm needs time at least $C2^N$ but the size of the input is at most N^2.

In contrast, our construction takes time which is polynomial in the size of the input (in other words, polynomial in $N \log n_N$). Assume that we are given N positive integers $n_1 \leq \cdots \leq n_N$ and let $L = \lceil \log_2 n_N \rceil$. Define $w_j = |\{k \in \mathbf{N} : j = n_k\}|$. The algorithm we described consists of the following steps. The notation of Lemma 4.2 is used throughout.

1. Find for which $\nu_0 \in \{0, \ldots, 5\}$ inequality (78) is true.

2. Construct the sequence of intervals $J_0 \supseteq J_1 \supseteq \cdots$ and the sequence of sets $E_0, E_1, \ldots$. This proceeds inductively. Having constructed the interval J_{m-1} and the set E_{m-1} we

   a. construct the intervals I_{n_k} for all $n_k \in E'_m$,

   b. find (as described in Lemma 4.1) a subinterval J_m of J_{m-1} which is big and is contained in many of the I_{n_k}'s. The set E_m consists of those $n_k \in E'_m$ for which $J_m \subseteq I_{n_k}$.

   Notice that the sequences J_m and E_m have length $O(L)$.

After observing that we never need to perform arithmetic with more than $O(L)$ binary digits, it is easy to see that all the above can be carried out in time $O(N \cdot L^2)$, since an algebraic operation on two numbers, with $O(L)$ binary digits each, takes $O(L^2)$ time.

2. Uchiyama's proof of (68) is probabilistic (see Section 1.4.5). We give an even simpler constructive proof of (68). (Of course Uchiyama proved the stronger statement (69) about the L^1 norm of a subseries.) Assume $n_1 \le n_2 \le \cdots \le n_N$ and let ρ be any fixed number between 2 and 3, say $\rho = 5/2$. Observe that if $n_N \le \rho n_1$ then

$$-\min_x \sum_{j=1}^{N} \cos n_j x \ge CN,$$

as can be seen by evaluating the function $\sum_{j=1}^{N} \cos n_j x$ for $x = (\pi/2+\epsilon)/n_1$, where $\epsilon = \epsilon(\rho)$ is a small positive constant.

Let $\lambda_1 = n_1$ and define $\lambda_k \in \{n_1, \ldots, n_N\}$ recursively by

$$\lambda_k = \min \{n_j : n_j > \rho\lambda_{k-1}\} \cup \{n_N\}.$$

Let L be the length of the sequence λ_k. That is let λ_L be the first λ equal to n_N. Then either $L \ge \sqrt{N}$ or there is some k for which the set

$$A = \{n_j : \lambda_k \le n_j < \lambda_{k+1}\}$$

has more than $\sqrt{N}$ elements. In the first case we have

$$-\min_x \sum_{j=1}^{L-1} \cos \lambda_j x \ge CL \ge C\sqrt{N},$$

since the λ_j's form a lacunary sequence with ratio $\rho > 2$. Otherwise, according to the observation above, we have

$$-\min_x \sum_{n_j \in A} \cos n_j x \ge C|A| \ge C\sqrt{N},$$

which completes the proof.

3. It is easy to see that Theorem 4.2 holds also for complex w_k, with $w = \sum |w_k| < \infty$ and writing e^{ikx} in place of $\cos kx$. Also the minimum in (71) has to be interpreted as the minimum (or maximum) of the real part.

Chapter 5

An Effective Additive Basis for the Integers

5.1 Introduction

A set E of nonnegative integers is called a basis if every nonnegative integer can be written as a sum of two elements of E. We write $r(x) = r_E(x)$ for the number of representations of x as $a + b$, with $a, b \in E$ and $a \le b$.

Erdős [14, 15] has proved that there is a basis E such that

$$C_1 \log x \le r(x) \le C_2 \log x \tag{79}$$

for all positive integers x and for some absolute constants C_1, C_2. For the proof see Chapter 1, p. 6 (see also [2, p. 106] and [25, Ch. 3]). The most widely known proof (in Chapter 1 and [2, 15, 25]) is probabilistic. It is proved that if we let $x \in E$ with a certain probability p_x, independently for all x, then the random set E is such an asymptotic basis (that is (79) is true for sufficiently large x) with probability 1. Since the probability space used is infinite, the question of whether such a basis exists which is also computable is not addressed by this proof.

The original [14] proof though, which has been stated using counting arguments and not probability, uses an existential argument on a finite interval at a time and can thus be readily turned into a construction by examining all possible intersections of E with the interval. But the algorithm which we get this way takes time exponential in n to decide whether n is in E or not.

In this Chapter, we give an algorithm which produces the elements of E one by one and in increasing order, and which takes time polynomial in n in order to produce all the elements of E not greater than n. We use the method of conditional probabilities (Chapter 1, p. 8 and [2, p. 223]) in order to "derandomize" a modified proof. The method is not directly applicable to Erdős's probabilistic proof. We will only show that (79) holds for x large enough, since, then, with the addition of a finite number of elements to the set E we can have it hold true for all positive x.

In Section 5.2 we give a probabilistic proof of the existence of a basis with certain properties. In Section 5.3 we apply the method of conditional probabilities to derandomize the proof and arrive to our algorithm.

This will appear in [34].

5.2 Probabilistic Proof of Existence

We define the modified representation function $r'(x) = r'_E(x)$ as the number of representations of the nonegative integer x as a sum $a + b$, with $a, b \in E$, $g(x) \le a \le b$, where $g(x) = (x \log x)^{1/2}$. (This is our main difference from Erdős's proof. By doing this modification we have achieved that the presence or absence of a certain number n in our set E affects $r'(x)$ for only a finite number of nonnegative integers x.)

Theorem 5.1 *There are positive constants c_1, c_2, c_3, with $c_2 < c_3$, and a set E of positive integers such that*

$$c_2 \log x \le r'(x) \le c_3 \log x$$

and

$$|E \cap [x - g(x), x]| \le c_1 \log x$$

for all large enough $x \in \mathbf{N}$.

Proof: In what follows x is assumed to be sufficiently large. We define the random set E by letting

$$\Pr[x \in E] = p_x = K \cdot \left(\frac{\log x}{x}\right)^{1/2}$$

independently for all $x \in \mathbf{N}$, where K is a positive constant that will be specified later. We are going to show that with positive probability (in fact almost surely but we do not need this here) the random set E satisfies Theorem 5.1. Let

$$\mu = \mathbf{E}[r'(x)] = \sum_{t=g(x)}^{x/2} p_t p_{x-t}.$$

Define also

$$s(x) = |E \cap [x-g(x), x]|$$

and

$$\nu = \mathbf{E}[s(x)] = \sum_{t=x-g(x)}^{x} p_t.$$

First we estimate μ and ν for large x. We have

$$\begin{aligned}
\mu &\geq \sum_{t=x/\log x}^{x/2} p_t p_{x-t} \\
&\geq K^2 \log \frac{x}{\log x} \sum_{t=x/\log x}^{x/2} (t(x-t))^{-1/2} \\
&= (1+o(1)) I K^2 \log x,
\end{aligned}$$

where $I = \int_0^{1/2} (s(1-s))^{-1/2} ds$, and

$$\begin{aligned}
\mu &\leq \sum_1^{x/2} p_t p_{x-t} \\
&\leq K^2 \log \frac{x}{2} \sum_1^{x/2} (t(x-t))^{-1/2} \\
&= (1+o(1)) I K^2 \log x,
\end{aligned}$$

which proves $\mu = (1+o(1)) I K^2 \log x$.

For ν we have

$$K g(x) \left(\frac{\log(x-g(x))}{x} \right)^{1/2} \leq \nu \leq K g(x) \left(\frac{\log x}{x-g(x)} \right)^{1/2},$$

which implies

$$\nu = (1+o(1)) K \log x.$$

We define the "bad" events

$$\begin{aligned}
A_x &= \{|r'(x) - \mu| > \epsilon \mu\} \\
B_x &= \{s(x) - \nu > \epsilon \nu\}
\end{aligned}$$

for a positive constant ϵ. To bound their probabilities we need the following Theorem 1.3. Since both $r'(x)$ and $s(x)$ are sums of independent indicator random variables we can use the theorem to get

$$\Pr[A_x] \le 2e^{-c_\epsilon \mu} \le 2e^{-\frac{1}{2}c_\epsilon I K^2 \log x} = 2x^{-\alpha}$$

and

$$\Pr[B_x] \le 2e^{-c_\epsilon \nu} \le 2e^{-\frac{1}{2}c_\epsilon K \log x} = 2x^{-\beta}$$

where $\alpha = \frac{1}{2}c_\epsilon I K^2$ and $\beta = \frac{1}{2}c_\epsilon K$. We now let $\epsilon = 1/2$ and choose K large enough to make both α and β greater than 1.

Then

$$\sum_{x=1}^{\infty} \Pr[A_x] + \Pr[B_x] < \infty$$

which implies the existence of $n_0 \in \mathbf{N}$ such that, with positive probability, none of the events A_x and B_x, $x \ge n_0$, holds. In particular there exists a set E for which

$$\mu/2 \le r'(x) \le 3\mu/2$$

and

$$s(x) \le 3\nu/2,$$

for all $x \ge n_0$. This implies the conclusion of Theorem 5.1 with $c_1 = \frac{1}{2}K$, $c_2 = \frac{1}{2}IK^2$ and $c_3 = \frac{3}{2}IK^2$. $\square$

Observe that $r'(x) \le r(x) \le r'(x) + s(x)$. We deduce that for the set E of Theorem 5.1 we have

$$c_2 \log x \le r(x) \le (c_1 + c_3) \log x$$

so that (79) is true for E.

5.3 Derandomization of the Proof

We keep the notation of the previous section. We showed that for some $n_0 \in \mathbf{N}$ the complement of the bad event

$$B = \bigcup_{x \ge n_0} (A_x \cup B_x)$$

has positive probability, by establishing the inequality

$$\sum_{x \ge n_0} \Pr[A_x] + \Pr[B_x] < 1.$$

This implies the existence of a point E in our probability space $\{0,1\}^{\mathbf{N}}$ which is not in B (there is a natural identification between points in the probability space and subsets of $\mathbf{N}$). In this section we are going to show how to construct efficiently such a point E. We give an algorithm which at the n-th step outputs 0 or 1 to denote the absence or presence of n in our set E.

Denote by $\chi \in \{0,1\}^{\mathbf{N}}$ a generic element in our space and by $R(a_1,\ldots,a_k)$ the event $\chi_1 = a_1,\ldots,\chi_k = a_k$, where $a_1,\ldots,a_k \in \{0,1\}$. It is obvious that for any event $D \subseteq \{0,1\}^{\mathbf{N}}$

$$\begin{aligned} &\Pr\left[D \mid R(a_1,\ldots,a_{n-1})\right] = \\ &\qquad p_n \Pr\left[D \mid R(a_1,\ldots,a_{n-1},1)\right] + (1-p_n)\Pr\left[D \mid R(a_1,\ldots,a_{n-1},0)\right]. \end{aligned} \tag{80}$$

We are going to define the sequence $a_n \in \{0,1\}$ so that the function

$$b_n = b_n(a_1,\ldots,a_n) = \sum_{x \ge n_0} \Pr\left[A_x \mid R(a_1,\ldots,a_n)\right] + \Pr\left[B_x \mid R(a_1,\ldots,a_n)\right]$$

is non-increasing in n. (Notice that the function $\Pr\left[A_x \mid R(a_1,\ldots,a_n)\right]$ is constant in n when $n > x$, and is equal to either 0 or 1. The same is true for the events B_x.) Since $b_0 = \sum_{x \ge n_0} \Pr\left[A_x\right] + \Pr\left[B_x\right] < 1$, the monotonicity of b_n implies that

$$\sum_{x \ge n_0} \Pr\left[A_x \mid R(a_1,\ldots,a_n,\ldots)\right] + \Pr\left[B_x \mid R(a_1,\ldots,a_n,\ldots)\right] < 1.$$

The probabilities above are either 0 or 1, so they are all 0, and the point $E = (a_1,\ldots,a_n,\ldots)$ is not in B.

So all that remains to be done is to ensure that b_n does not increase. Adding up (80) we get

$$b_{n-1}(a_1,\ldots,a_{n-1}) = p_n b_n(a_1,\ldots,a_{n-1},1) + (1-p_n) b_n(a_1,\ldots,a_{n-1},0),$$

which implies that at least one of $b_n(a_1,\ldots,a_{n-1},1)$, $b_n(a_1,\ldots,a_{n-1},0)$ is not greater than $b_{n-1}(a_1,\ldots,a_{n-1})$. We let $a_n = 1$ if the first number is smaller than the latter, otherwise we let $a_n = 0$.

Notice that

$$\begin{aligned} \Delta &= b_n(a_1,\ldots,a_{n-1},1) - b_n(a_1,\ldots,a_{n-1},0) \\ &= \sum_{x=n}^{G(n)} \Pr\left[A_x \mid R(a_1,\ldots,a_{n-1},1)\right] - \Pr\left[A_x \mid R(a_1,\ldots,a_{n-1},0)\right] + \\ &\qquad + \Pr\left[B_x \mid R(a_1,\ldots,a_{n-1},1)\right] - \Pr\left[B_x \mid R(a_1,\ldots,a_{n-1},0)\right], \end{aligned}$$

where $G(n) = (1 + o(1))n^2/\log n$ is the greatest integer k such that $g(k) \le n$. This is so because the events A_x and B_x, with $x > G(n)$ are independent of $\chi_1, \ldots, \chi_n$ and their probabilities cancel out in the difference above. We have to decide in time polynomial in n whether $\Delta \ge 0$. This is indeed possible since the expression for Δ has $(4 + o(1))n^2/\log n$ terms, each of which can be computed in polynomial time as the following Lemma claims.

Lemma 5.1 *Let $X_k = \xi_1 + \cdots + \xi_k$ be a sum of k independent indicator random variables with $\Pr[\xi_j = 1] = p_j$, $j = 1, \ldots, k$. Then the distribution of X_k can be computed in time polynomial in k.*

Proof: The distribution of X_k is a vector of length $k + 1$, where the j-th coordinate in the vector, $j = 0, \ldots, k$, is equal to $\Pr[X_k = j]$. To compute the distribution of X_k from that of X_{k-1} we use the obvious formulas

$$\Pr[X_k = j] = p_k \Pr[X_{k-1} = j - 1] + (1 - p_k)\Pr[X_{k-1} = j], \quad \text{for } j = 1, \ldots, k - 1,$$

$\Pr[X_k = 0] = (1 - p_k)\Pr[X_{k-1} = 0]$ and $\Pr[X_k = k] = p_k \Pr[X_{k-1} = k - 1]$. It is obvious now that the computation of the distribution of X_k can be carried out in time polynomial in k. (Here we are really assuming that arithmetic operations on the numbers p_j can be done in time polynomial in k. See the Remarks at the end of the section for a justification of this assumption.) $\Box$

Thus all probabilities of the form $\Pr[\alpha < X_k < \beta]$ can be efficiently computed. Observe that having fixed $\chi_1 = a_1, \ldots, \chi_n = a_n$ we have

$$\begin{aligned} r'(x) &= \sum_{t=g(x)}^{x/2} \chi_t \chi_{x-t} \\ &= \sum_{g(x)}^{n} a_t \chi_{x-t} + \sum_{n+1}^{x/2} \chi_t \chi_{x-t} \end{aligned}$$

for $x - g(x) > n$, otherwise $r'(x)$ has already been completely determined by the assigned values of $\chi_1, \ldots, \chi_n$. This means that $r'(x)$ is a sum of independent indicator random variables and so is $s(x)$. Thus the probabilities of A_x and B_x conditioned on $R(a_1, \ldots, a_{n-1}, 1)$ and $R(a_1, \ldots, a_{n-1}, 0)$ can be efficiently computed and $\Delta \ge 0$ can be decided in polynomial time, as we had to show.

Remarks:

1. Our definition of the probabilities $p_x = K(\log x/x)^{1/2}$ has to be modified so that the numbers p_x can be represented with a number of digits polynomial in x and can also be computed in polynomial time, given x. One such modification is to use the probabilities $q_x = K2^{-\lfloor L/2 \rfloor}S$, where $L = \lfloor \log_2 x \rfloor$ and $S = \lfloor \sqrt{L} \rfloor$. The number S can for example be computed in time polynomial in $\log L$ (and in particular in x) using a simple binary search of the interval $[0, L]$. Since $p_x < Cq_x < Cp_x$ one can easily prove asymptotic estimates of the form $CIK^2 < \mu < CIK^2$ and $CK \log x < \nu < CK \log x$, which is all our existential proof needs.

2. Ignoring polylogarithmic factors, the time our algorithm needs to decide whether $n \in E$, having already found the set E up to $n - 1$, is $O(n^6)$. This is so since the distribution of X_k in Lemma 5.1 can be computed in time $O(k^2)$. So the computation of any probability of the form $\Pr[\alpha < X_k < \beta]$ can be computed in time $O(k^2)$. For the computation of Δ we need to evaluate $O(n^2)$ such probabilities with $k = O(n^2)$, thus the total time is $O(n^6)$.

Chapter 6

On a Problem of Erdős and Turán and Some Related Results

6.1 Introduction

All sequences we consider are sequences of distinct positive integers. We denote by the lower case indexed letter the members of the sequence and by the capital letter the sequence as a set as well as its counting function. For example $A = \{a_1, a_2, \ldots\}$ denotes a sequence of distinct positive integers and $A(x) = |A \cap [1, x]|$ denotes its counting function. We define

$$\delta_A(x) = |\{(a, b) : a, b \in A,\ x = a - b\}|,$$

$$h_{A,N}(x) = \left|\{(a, b) : a, b \in A \cap [1, N^2],\ x = a - b\}\right|,$$

$$H_A(N) = \sum_{x=1}^{N} h_{A,N}(x),$$

and

$$r_A(x) = |\{(a, b) : a, b \in A,\ a \leq b,\ x = a + b\}|.$$

A conjecture of Erdős and Turán [20] asserts that for any asymptotic basis (of order 2) of the positive integers, that is for any set $E \subseteq \mathbf{N} = \{1, 2, 3, \ldots\}$ for which $r_E(x) > 0$ for all sufficiently large x, we must have

$$\limsup_{x \to \infty} r_E(x) = \infty.$$

Erdős [60] has proved that it cannot eventually be true that $r_E(x) = 1$, by showing that for any sequence E, with $E(x) \gg \sqrt{x}$, we have

$$H_E(N) \gg N \log N \tag{81}$$

for all N. Indeed, any asymptotic basis E satisfies $E(x) \gg \sqrt{x}$ and, if $r_E(x) = 1$, all sums we can form with two elements of E (with the exception of a finite number of elements of E) are distinct. This in turn implies that so are all the differences, that is $\delta_E(x) \leq 1$ for all x, which makes (81) impossible.

Recently Helm [27] proved that (81) is best possible by explicitly constructing a sequence A, with $A(x) \gg \sqrt{x}$, for which

$$H_A(N) \ll N \log N \tag{82}$$

for all sufficiently large N. Helm's proof does not provide any upper or lower bound on the individual $h_{A,N}(x)$ for $x \in [1, N]$, but only describes the average behaviour.

In addition to this result Helm [27] constructed two sequences B and M, with $B(x) \gg \sqrt{x}$ and $M(x) \gg \log x$, for which $\delta_B(m_k) = 1$, for all k sufficiently large.

In this Chapter we shall use the probabilistic method to improve both results of Helm. Throughout this Chapter a random sequence A is defined by letting $x \in A$ with probability

$$p_x = \begin{cases} \frac{K}{\sqrt{x}} & \text{if } x \geq K^2, \\ 0 & \text{otherwise,} \end{cases}$$

for appropriately chosen K, independently for all x.

We prove

Theorem 6.1 *Let A be a random sequence as defined above. Then, with probability 1, there is an integer N_0 and positive constants c_1, c_2, c_3, c_4 such that*

$$c_1\sqrt{x} \leq A(x) \leq c_2\sqrt{x} \tag{83}$$

and

$$c_3 \log N \leq h_{A,N}(m) \leq c_4 \log N \tag{84}$$

for all $x, N \geq N_0$ and $1 \leq m \leq N$.

This implies the first result of Helm and with upper and lower estimates on the individual $h_{A,N}(m)$. We also prove

Theorem 6.2 *Let M be any sequence which satisfies the growth condition*

$$\sum_{1}^{\infty} \frac{\log m_k}{\sqrt{m_k}} < \infty. \tag{85}$$

Let also A be a random sequence. Then, with probability 1, there is a subsequence B of A, an integer N_0 and positive constants c_3, c_4 such that

$$c_3\sqrt{x} \le B(x) \le c_4\sqrt{x} \tag{86}$$

for all $x \ge N_0$ and

$$\delta_B(m_k) = 1 \tag{87}$$

for all $k \ge N_0$.

Helm's second result follows from Theorem 6.2, but Theorem 6.2 is much stronger, since we are free to choose the sequence M, subject only to the growth condition (85).

This work will appear in [35].

6.2 Proofs

Proof of Theorem 6.1: Write $\chi_j = 1$ if $j \in A$, $\chi_j = 0$ otherwise, so that $\mathbf{E}[\chi_j] = p_j$. Notice that

$$\begin{aligned} A(x) &= \sum_{j=1}^{x} \chi_j, \\ h_{A,N}(m) &= \sum_{j=1}^{N^2-m} \chi_j \chi_{j+m}, \end{aligned}$$

so that $A(x)$ is a SIIRV and $h_{A,N}(m)$ is the sum of two SIIRV:

$$h_{A,N}(m) = h^e_{A,N}(m) + h^o_{A,N}(m),$$

where

$$h^e_{A,N}(m) = \sum_{j=1}^{m} \sum_{k \text{ even}} \chi_{j+km}\chi_{j+(k+1)m}$$

and

$$h^o_{A,N}(m) = \sum_{j=1}^{m} \sum_{k \text{ odd}} \chi_{j+km}\chi_{j+(k+1)m}.$$

(We broke up $h_{A,N}(m)$ so that each χ_j appears at most once in each $h^e_{A,N}(m)$ and $h^o_{A,N}(m)$.) Then, as $x \to \infty$,

$$\begin{aligned}
\mathbf{E}[A(x)] &= \sum_{j=1}^{x} p_j \sim \sum_{j=1}^{x} \frac{K}{\sqrt{j}} \\
&= K\sqrt{x} \sum_{j=1}^{x} \frac{1}{x} \frac{1}{\sqrt{j/x}} \\
&\sim 2K\sqrt{x},
\end{aligned}$$

since $2 = \int_0^1 ds/\sqrt{s}$. We also have, for $m \le N$ and $N \to \infty$,

$$\begin{aligned}
\mathbf{E}[h_{A,N}(m)] &= \sum_{j=1}^{N^2-m} p_j p_{j+m} \\
&\sim K^2 \sum_{j=1}^{N^2-m} \frac{1}{\sqrt{j(j+m)}} \\
&\le K^2 \sum_{j=1}^{N^2-m} \frac{1}{j} \sim 2K^2 \log N,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}[h_{A,N}(m)] &\ge (1+o(1))K^2 \sum_{j=1}^{N^2-m} \frac{1}{j+m} \\
&\ge (1+o(1))K^2 \log N.
\end{aligned}$$

So we have

$$\mathbf{E}[A(x)] \sim 2K\sqrt{x} \tag{88}$$

as $x \to \infty$ and

$$(1+o(1))K^2 \log N \le \mathbf{E}[h_{A,N}(m)] \le (1+o(1))2K^2 \log N \tag{89}$$

as $N \to \infty$, and for all $m \le N$. Notice that $\mathbf{E}[h^e_{A,N}(m)] \sim \frac{1}{2}\mathbf{E}[h_{A,N}(m)]$ and $\mathbf{E}[h^o_{A,N}(m)] \sim \frac{1}{2}\mathbf{E}[h_{A,N}(m)]$.

Now fix $\epsilon = 1/2$ and define the bad events

$$\begin{aligned}
P_x &= \{|A(x) - \mathbf{E}[A(x)]| > \epsilon \mathbf{E}[A(x)]\}, \\
Q_{N,m} &= \{|h_{A,N}(m) - \mathbf{E}[h_{A,N}(m)]| > \epsilon \mathbf{E}[h_{A,N}(m)]\},
\end{aligned}$$

for all x, N and $m \le N$. Using Theorem 1.3 and the remarks following it we have

$$\Pr[P_x] \le 2\exp(-c_\epsilon \mathbf{E}[A(x)]) \le 2\exp(-\frac{1}{2}c_\epsilon 2K\sqrt{x})$$

and

$$\Pr[Q_{N,m}] \le 4\exp(-\frac{1}{3}c_\epsilon \mathbf{E}[h_{A,N}(m)]) \le 4\exp(-\frac{1}{6}c_\epsilon K^2 \log N) = 4N^{-\frac{1}{6}c_\epsilon K^2}$$

for x and N sufficiently large. Thus

$$\sum_{x=1}^{\infty} \Pr[P_x] + \sum_{N=1}^{\infty}\sum_{m=1}^{N} \Pr[Q_{N,m}] \ll \sum_{x=1}^{\infty} \exp(-c_\epsilon K\sqrt{x}) + \sum_{N=1}^{\infty} N^{1-\frac{1}{6}c_\epsilon K^2}.$$

The first term in the right hand side is finite, and we choose K large enough to make the second term also finite, that is large enough to make $1 - \frac{1}{6}c_\epsilon K^2 < -1$. Let now $\epsilon' \in (0,1)$ be arbitrary. Since the right hand side above is finite, we can find N_0 so that

$$\sum_{x \ge N_0} \Pr[P_x] + \sum_{N \ge N_0}\sum_{m=1}^{N} \Pr[Q_{N,m}] < \epsilon' \tag{90}$$

which means that, with probability at least $1-\epsilon'$, none of the events which appear in (90) holds. We conclude that, with probability at least $1-\epsilon'$,

$$(1+o(1))K\sqrt{x} \le A(x) \le (1+o(1))3K\sqrt{x},$$

and

$$(1+o(1))\frac{1}{2}K^2 \log N \le h_{A,N}(m) \le (1+o(1))3K^2 \log N,$$

for all $x, N \ge N_0$ and $1 \le m \le N$. Since ϵ' was arbitrary this concludes the proof.

Proof of Theorem 6.2: Let $\epsilon > 0$. By the proof of Theorem 6.1, with probability at least $1-\epsilon$, the random sequence A satisfies

$$c_1(\epsilon)\sqrt{x} \le A(x) \le c_2(\epsilon)\sqrt{x}.$$

We have

$$\delta_A(m) = \sum_{j=1}^{\infty} \chi_j \chi_{j+m} = \delta_A^e(m) + \delta_A^o(m),$$

where

$$\delta_A^e(m) = \sum_{r=1}^{n} \sum_{k \text{ even}} \chi_{r+km}\chi_{r+(k+1)m}$$

and

$$\delta_A^o(m) = \sum_{r=1}^{n} \sum_{k \text{ odd}} \chi_{r+km}\chi_{r+(k+1)m}.$$

Notice that $\delta_A^e(m)$ and $\delta_A^o(m)$ are both SIIRV and that $\mathbf{E}[\delta_A^e(m)] = \mathbf{E}[\delta_A^o(m)] = \infty$. An application of the Borel-Cantelli Lemma shows that they are both ∞, for all m, almost surely (a.s.), and so is $\delta_A(m)$.

We shall say that *a is used by m* if $a, a+m \in A$. Let l_k be a sequence of integers which tends to ∞, and whose rate of growth will be determined later. The sequence l_k will depend on the sequence m_k only. We define $f(m_k)$ to be the least integer $a \geq l_k$ which is used by m_k (it exists a.s. by the previous argument).

For each m_k we will remove every $a \in A$ which is used by m_k except if $a = f(m_k)$. We then want to ensure that $f(m_k)$ and $f(m_k) + m_k$ will not be removed by any m_j. Define

$$w(y) = \sum_{j=1}^{\infty} \frac{1}{\sqrt{y + m_j}}.$$

We first prove that the growth condition on the sequence m_k implies that $\sum_{y=1}^{\infty} w(y)/y < \infty$:

$$\begin{aligned}
\sum_{y=1}^{\infty} \frac{w(y)}{y} &= \sum_{j=1}^{\infty}\sum_{y=1}^{\infty} \frac{1}{y\sqrt{y+m_j}} \\
&= \sum_{j=1}^{\infty} \left(\sum_{y=1}^{m_j} \frac{1}{y\sqrt{y+m_j}} + \sum_{y=m_j+1}^{\infty} \frac{1}{y\sqrt{y+m_j}} \right) \\
&\leq \sum_{j=1}^{\infty} \left(\sum_{y=1}^{m_j} \frac{1}{y\sqrt{m_j}} + \sum_{y=m_j+1}^{\infty} \frac{1}{y^{3/2}} \right) \\
&\ll \sum_{j=1}^{\infty} \left(\frac{\log m_j}{\sqrt{m_j}} + \frac{1}{\sqrt{m_j}} \right) \\
&< \infty.
\end{aligned}$$

Then define the bad event

$$E_1 = \bigcup_{k=1}^{\infty} \bigcup_{j \neq k} \bigcup_{y \geq l_k} \{y \in A \ \&\ y + m_k \in A \ \&\ y + m_j \in A\}.$$

Clearly no $f(m_k)$ will be removed by any m_j, $j \neq k$, if E_1 does not hold. We have to bound the probability of E_1 by ϵ:

$$\Pr[E_1] \leq K^3 \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \sum_{y \geq l_k} \frac{1}{\sqrt{y(y+m_k)(y+m_j)}}$$

$$
\begin{aligned}
&\le \quad K^3 \sum_{k=1}^{\infty} \sum_{y \ge l_k} \sum_{j=1}^{\infty} \frac{1}{y\sqrt{y+m_j}} \\
&= \quad K^3 \sum_{k=1}^{\infty} \sum_{y \ge l_k} \frac{w(y)}{y}.
\end{aligned}
$$

But since $\sum_{y=1}^{\infty} w(y)/y < \infty$ we can choose the numbers l_k large enough to have

$$\sum_{y \ge l_k} \frac{w(y)}{y} \le K^{-3} 10^{-k} \epsilon$$

and then $\Pr[E_1] \le \epsilon$ follows.

Similarly we can bound by ϵ the probability of the event

$$E_2 = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcup_{y \ge l_k} \{y \in A \ \&\ y - m_k \in A \ \&\ y + m_j \in A\}.$$

Avoiding E_2 implies that no integer of the form $f(m_k) + m_k$ can be removed by any m_j, $j = 1, 2, \ldots$.

We can now form the sequences

$$A' = \{a \in A \ : \ a \text{ is used by some } m_k,\ k \ge N_0 \text{ and } a \ne f(m_k)\}$$

and

$$B = A \backslash A'.$$

If E_1, E_2 are false then $\delta_B(m_k) = 1$ for all $k \ge N_0$ and all that remains is to ensure that $B(x) \ge c_3(\epsilon)\sqrt{x}$.

We shall prove that $A'(x) \le \frac{1}{2} c_1(\epsilon)\sqrt{x}$. Note that $A'(x)$ is not a SIIRV and we cannot use Theorem 1.3 to bound the probability of large deviations from its mean value (which is $\ll \sqrt{x}$). Instead, we shall use Markov's inequality: $\Pr[X \ge \lambda \mathbf{E}[X]] \le 1/\lambda$, for $X \ge 0$. It is not really essential in the first part of the proof, that is in bounding the probabilities of E_1, E_2.)

For $n \ge 0$ write

$$s_n = \left| A' \cap [2^n, 2^{n+1}) \right|.$$

It is enough to show $s_n \le c(\epsilon) 2^{n/2}$, with $c(\epsilon)$ sufficiently small. We have

$$s_n \le \sum_{j=2^n}^{2^{n+1}-1} \sum_{k=N_0}^{\infty} \chi_j \chi_{j+m_k}$$

which implies, for n sufficiently large,

$$\begin{aligned}
\mathbf{E}[s_n] &= K^2 \sum_{j=2^n}^{2^{n+1}-1} \sum_{k=N_0}^{\infty} \frac{1}{\sqrt{j(j+m_k)}} \\
&= K^2 \sum_{j=2^n}^{2^{n+1}-1} \frac{1}{\sqrt{j}} \sum_{k=N_0}^{\infty} \frac{1}{\sqrt{j+m_k}} \\
&\leq K^2 2^{n/2} \sum_{k=N_0}^{\infty} \frac{1}{\sqrt{2^n+m_k}}.
\end{aligned}$$

We want to bound by, say, ϵ the probability of the event

$$E_3 = \bigcup_{n=1}^{\infty} \{s_n \geq c(\epsilon) 2^{n/2}\}.$$

By Markov's inequality, it suffices to show

$$c^{-1}(\epsilon) K^2 \sum_{n=1}^{\infty} \sum_{k=N_0}^{\infty} \frac{1}{\sqrt{2^n+m_k}} \leq \epsilon.$$

But the sum on the left can be written as

$$S_1 + S_2 = c^{-1}(\epsilon) K^2 \sum_{n=1}^{\infty} \sum_{\substack{k \geq N_0 \\ m_k \leq 2^n}} \frac{1}{\sqrt{2^n+m_k}} + c^{-1}(\epsilon) K^2 \sum_{n=1}^{\infty} \sum_{m_k > \max\{2^n, N_0\}} \frac{1}{\sqrt{2^n+m_k}}.$$

If one writes $C(K,\epsilon)$ for an arbitrary constant that depends at most on K and ϵ, then

$$\begin{aligned}
S_1 &\leq C(K,\epsilon) \sum_{k=N_0}^{\infty} \sum_{n \geq \log_2 m_k} 2^{-n/2} \\
&\leq C(K,\epsilon) \sum_{k=N_0}^{\infty} \frac{2^{-\lceil \log_2 m_k \rceil/2}}{1-2^{-1/2}} \\
&\leq C(K,\epsilon) \sum_{k=N_0}^{\infty} m_k^{-1/2},
\end{aligned}$$

and

$$\begin{aligned}
S_2 &\leq C(K,\epsilon) \sum_{k=N_0}^{\infty} \sum_{2^n < m_k} \frac{1}{\sqrt{m_k}} \\
&\leq C(K,\epsilon) \sum_{k=N_0}^{\infty} \frac{\log_2 m_k}{\sqrt{m_k}}.
\end{aligned}$$

We now choose N_0 large enough to make both S_1 and S_2 smaller than $\epsilon/2$. Since the probabilites of E_1, E_2 and E_3 have been shown smaller than the arbitrary ϵ, the proof is complete.

Bibliography

[1] M. Ajtai, J. Komlós and E. Szemerédi, *A dense infinite Sidon sequence*, Europ. J. Comb. 2 (1981), 1-11.

[2] N. Alon and D. J. Kleitman, *Sum-free subsets*, in A. Baker, B. Bollobás, and A. Hajnál, eds., A Tribute to Paul Erdős, Cambridge University Press (1990), 13-26.

[3] N. Alon and J. Spencer, *The probabilistic method*, Wiley Interscience Series in Discrete Mathematics and Optimization, 1992.

[4] J. Beck, *Flat polynomials on the unit circle – note on a problem of Littlewood*, Bull. London Math. Soc. 23 (1991), 269-277.

[5] R. C. Bose and S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv. 37 (1962-63), 141-147.

[6] J. Bourgain, *Sur les sommes de sinus*, Sém. Anal. Harm., Publ. Math. d'Orsay 84-01 (1984), exp. no 3.

[7] J. Bourgain, *Sur le minimum d'une somme de cosinus*, Acta Arith. 45 (1986), 381-389.

[8] J. S. Byrnes, *On polynomials with coefficients of modulus one*, Bull. London Math. Soc. 9 (1977), 171-176.

[9] H. Chernoff, *A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Stat. 23 (1952), 493-509.

[10] S. Chowla, *Solution of a problem of Erdős and Turán in additive number theory*, Proc. Nat. Acad. Sci. India 14 (1944) 1-2.

[11] S. Chowla, *Some applications of a method of A. Selberg*, J. Reine Angew. Math. 217 (1965), 128-132.

[12] P. J. Cohen, *On a conjecture of Littlewood and idempotent measures*, Amer. J. Math. 82 (1960), 191-212.

[13] H. Davenport, *On a theorem of P. J. Cohen*, Mathematica 7 (1960), 93-97.

[14] P. Erdős, *On a problem of Sidon in additive number theory*, Acta Sci. Math. (Szeged), 15 (1953-54), 255-259.

[15] P. Erdős, *Problems and results in additive number theory*, Colloque sur la Théorie des Nombres (CBRM, Bruxelles), 127-137.

[16] P. Erdős, *Extremal problems in number theory*, Proceedings of the Symp. Pure Math. VIII, AMS (1965), 181-189.

[17] P. Erdős, *Some applications of probability theory to Number Theory. Successes and limitations*, in *Sequences*, Capocelli (editor), Springer-Verlag.

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

[19] P. Erdős and G. Szekeres, *On the product* $\prod_{k=1}^{n}(1-z^{a_k})$, Acad. Serbe Sci., Publ. Inst. Math. 13 (1959), 29-34.

[20] 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; *Addendum* (by P. Erdős), ibid. 19 (1944), 208.

[21] L. Fejér, *Über trigonometrische Polynome*, J. Reine Angew. Math. 146 (1915), 53-82.

[22] J.-F. Fournier, *On a theorem of Paley and the Littlewood conjecture*, Arkiv för Mat. 17 (1979), 199-216.

[23] E. D. Gluskin, *Extremal properties of orthogonal parellelepipeds and their applications to the geometry of Banach spaces*, Math. USSR Sbornik 64 (1989), 1, p. 85-96.

[24] S. W. Graham, *Upper bounds for Sidon sequences*, preprint.

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

[26] G. H. Hardy and J. E. Littlewood, *A new proof of a theorem on rearrangements*, J. London Math. Soc. 23 (1948), 163-168.

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

[28] J.-P. Kahane, *Sur les polynômes à coefficients unimodulaires*, Bull. London Math. Soc. 12 (1980), 321-342.

[29] J.-P. Kahane, *Some random series of functions*, Cambridge Studies in Advanced Mathematics 5, 1985, Second Edition.

[30] Y. Katznelson, *An introduction to Harmonic Analysis*, Wiley, 1968, and Dover, 1976, New York.

[31] M. N. Kolountzakis, *On nonnegative cosine polynomials with nonnegative, integral coefficients*, Proc. Amer. Math. Soc. 120 (1994), 157-163.

[32] M. N. Kolountzakis, *Selection of a large sum-free subset in polynomial time*, Inf. Proc. Letters 49 (1994), 255-256.

[33] M. N. Kolountzakis, *A construction related to the cosine problem*, Proc. Amer. Math. Soc. , to appear.

[34] M. N. Kolountzakis, *An effective additive basis for the integers*, Discr. Math., to appear.

[35] M. N. Kolountzakis *On a problem of Erdős and Turán and some related results*, J. Number Th., to appear.

[36] S. Konjagin, *O probleme Littlewood'a*, Izv. A.N. SSSR, ser. mat. 45, 2 (1981), 243-265.

[37] T. W. Körner, *On a polynomial of J. S. Byrnes*, Bull. London Math. Soc. 12 (1980), 219-224.

[38] F. Krückeberg, B_2 *Folgen und verwandte Zahlenfolgen*, J. Reine Angew. Math. 106 (1961), 53-60.

[39] X.-D. Jia, *On finite Sidon sequences*, J. Number Th. 44 (1993), 84-92.

[40] X.-D. Jia, *$B_h[g]$-sequences with large upper density*, preprint.

[41] J. Lamperti, *Probability*, Benjamin, New York, 1966.

[42] B. Lindström, *A remark on B_4 sequences*, J. Comb. Theory, 7 (1969), 276-277.

[43] J. E. Littlewood, *On polynomials $\sum^n \pm z^m, \sum^n e^{\alpha_m i} z^m, z = e^{\theta i}$*, J. London Math. Soc. 41 (1966), 367-376.

[44] L. Lovász, J. Spencer and K. Vesztergombi, *Discrepancy of set systems and matrices*, European J. Combin. 7 (1986), 151-160.

[45] O. C. McGehee, L. Pigno, B. Smith, *Hardy's inequality and L^1-norms of exponential sums*, Ann. Math. 113 (1981), 613-618.

[46] A. M. Odlyzko, *Minima of cosine sums and maxima of polynomials on the unit circle*, J. London Math. Soc. (2) 26 (1982), 412-420.

[47] A. M. Odlyzko, personal communication.

[48] S. K. Pichorides, *A lower bound for the L^1 norm of exponential sums*, Mathematika 21 (1974), 155-159.

[49] S. K. Pichorides, *L^p norms of exponential sums*, Sém. Anal. Harm., Publ. Math. d'Orsay 77-73 (1976), 1-65.

[50] S. K. Pichorides, *A remark on exponential sums*, Bull. Amer. Math. Soc. 83 (1977), 283-285.

[51] S. K. Pichorides, *On a conjecture of Littlewood concerning exponential sums I*, Bull. Greek Math. Soc. 18 (1977), 8-16.

[52] S. K. Pichorides, *On a conjecture of Littlewood concerning exponential sums II*, Bull. Greek Math. Soc. 19 (1978), 274-277.

[53] S. K. Pichorides, *On the L^1 norms of exponential sums*, Ann. Inst. Fourier 30 (1980), 79-89.

[54] K. F. Roth, *On cosine polynomials corresponding to sets of integers*, Acta Arith. 24 (1973), 347-355.

[55] W. Rudin, *Some theorems on Fourier coefficients*, Proc. Amer. Math. Soc. 10 (1959), 855-859.

[56] R. Salem and A. Zygmund, *Some properties of trigonometric series whose terms have random signs*, Acta Math. 91 (1954), 245-301

[57] H. S. Shapiro, *Extremal problems for polynomials and power series*, Thesis for S.M. degree, MIT (1957).

[58] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. 43 (1938), 377-385.

[59] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. 289, 2 (1985), 679-706.

[60] A. Stöhr, *Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe II*, J. Reine Angew. Math. 194 (1955), 111-140.

[61] S. Uchiyama, *On the mean modulus of trigonometric polynomials whose coefficients have random signs*, Proc. Amer. Math. Soc. 16 (1965), 1185-1190.

[62] A. Zygmund, *Trigonometric series*, Cambridge Univ. Press, 1959, vol. 1 and 2.