PROBLEMS IN COMBINATORIAL NUMBER THEORY

A Thesis
Presented to
The Academic Faculty

by

Gagik Amirkhanyan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
May 2014

PROBLEMS IN COMBINATORIAL NUMBER THEORY

Approved by:

Brett Wick, Committee Chair
School of Mathematics
Georgia Institute of Technology

Michael Lacey, Advisor
School of Mathematics
Georgia Institute of Technology

Dmitriy Bilyk
Department of Mathematics
University of Minnesota

Jeff Geronimo
School of Mathematics
Georgia Institute of Technology

Doron Lubinsky
School of Mathematics
Georgia Institute of Technology

Date Approved: 27 March 2014

To my parents.

ACKNOWLEDGEMENTS

I would like to thank my adviser Michael Lacey for his guidance and support throughout my study as well as for the numerous other things that he has done for me.

I would like to thank Ernie Croot for all the knowledge and research skills that I learnt from him.

I am also grateful my collaborators Dmitriy Bilyk, Albert Bush and Crish Pryby. I enjoyed and learnt a lot working with them.

Many thanks to Dmitriy Bilyk, Jeff Geronimo, Doron Lubinsky and Brett Wick for agreeing to be in my committee.

I am also thankful to the professors at Georgia Tech and my fellow graduate students for their help and friendship.

TABLE OF CONTENTS

SUMMARY

The thesis consists of two parts.

The first part is devoted to results in the discrepancy theory.

For integers $d \geq 2$, and $N \geq 1$, let $\mathcal{P}_N \subset [0,1]^d$ be a finite point set with cardinality $\sharp\mathcal{P}_N = N$. Define the associated discrepancy function by

$$D_N(X) = \sharp(\mathcal{P}_N \cap [0,X]) - N|[0,X]|, \tag{1}$$

where $X = (x_1, \dots, x_d)$ and $[0,X] = \prod_{j=1}^{d}[0,x_j]$ is a rectangle with antipodal corners at 0 and X, and $|\cdot|$ stands for the d-dimensional Lebesgue measure. The dependence upon the selection of points $\mathcal{P}_N$ will be suppressed, as we are interested in bounds that are only a function of $N = \sharp\mathcal{P}_N$. The discrepancy function measures equidistribution of the point set in the unit cube. A set of points is well-distributed if the discrepancy function is small in some appropriate function space.

In chapter 1 we consider geometric discrepancy in higher dimensions ($d > 2$) and with co-authors obtain estimates in Exponential Orlicz Spaces [1].

Theorem 0.0.1. *In all dimensions $n \geq 2$ for every integer $N \geq 1$ there exists a distribution $\mathcal{P} \subset [0,1]^n$ of N points such that*

$$\|\mathcal{D}_N\|_{\exp\left(L^{\frac{2}{n+1}}\right)} \lesssim (\log N)^{\frac{n-1}{2}} . \tag{2}$$

This has recently been proved by M. Skriganov, using random digit shifts of binary digital nets, building upon the remarkable examples of W. Chen and M. Skriganov. Our approach, developed independently, complements that of Skriganov. In dimension 2, analogs of this result are known. But in dimensions three and higher these are the only known results with close to optimal exponential Orlicz norms proved. The proof proceeds by a very delicate analysis. The example points sets are chosen to be optimal with respect to certain coding theory parameters. The Walsh-Paley expansion of the function $\mathcal{D}_N$ has a striking structure.

In chapter 2 we consider geometric discrepancy in L^1 space. It is a well-known conjecture in the theory of irregularities of distribution that the L^1 norm of the discrepancy function of an N-point set satisfies the same asymptotic lower bounds as its L^2 norm. In dimension $d = 2$ this fact has been established by Halász, while in higher dimensions the problem is wide open. With co-authors (see [2]), we establish a series of dichotomy-type results which state that if the L^1 norm of the discrepancy function is too small (smaller than the conjectural bound), then the discrepancy function has to be very large in some other function space. For instance, we show that

Theorem 0.0.2. *For all dimensions $d \geq 3$, there is an $\epsilon = \epsilon(d) > 0$ and $c = c(d) > 0$ such that for all integers $N \geq 1$, every $\mathcal{P}_N \subset [0,1]^d$ satisfies either*

$$\|D_N\|_1 \geq (\log N)^{(d-1)/2-\epsilon} \quad \text{or} \quad \|D_N\|_2 \geq \exp(c(\log N)^{\epsilon}) .$$

This result is one of only a very few that concern the L^1 endpoint. Curiously, the proofs proceed by analyzing carefully the incomplete information that is available about the L^∞ endpoint. Examples show that the dichotomy results are close to the limits of what can be proved in this direction.

The second part of the thesis, chapter 3, is devoted to results in the additive combinatorics. An order-preserving Freiman 2-isomorphism is a bijection $\phi : A \to B$, where A and B are finite subsets of $\mathbb{R}$, such that $\phi(a)+\phi(b) = \phi(c)+\phi(d)$ if and only if $a+b = c+d$ and $\phi(a) < \phi(b)$ if and only if $a < b$. We (see [3]) show that for any $A \subseteq \mathbb{Z}$, if $|A + A| \leq K|A|$, then there exists a subset $A' \subseteq A$ such that the following holds: (1) $|A'| \geq c_1|A|$, (2) there exists an order preserving Freiman 2-isomorphism $\phi : A' \to B'$, where $B' \subset [1, c_2|A|]$ and c_1, c_2 depend only on K. Informally, this states that a set with small doubling may, in a sense, be viewed as a dense subset of an interval. We also present several applications.

For a finite set of A of integers its *additive energy* is $E(A, A) := |\{(i, j, k, l) : a_i + a_j = a_k + a_l\}$. If we order the set we can define its *indexed energy.* If $A = \{a_1 < \ldots < a_n\}$ is a subset of the integers, the indexed energy of A is defined as $EI(A, A) := |\{(i, j, k, l) : a_i + a_j = a_k + a_l$ and $i + j = k + l\}|$. With co-authors [3], we prove the following theorem.

Theorem 0.0.3. *Let A be a finite set of integers with $|A+A| \leq c|A|$. Then, there exists c_1, c_2 depending only on c such that the following holds. There exists an $A' \subseteq A$ such that $EI(A', A') \geq c_1|A'|^3$ and $|A'| \geq c_2|A|$.*

We also give a construction of a set A with $E(A, A) = \Omega(n^3)$ and $EI(A, A) = O(n \log n)^2$.

The proof uses many techniques, first we handle the case when A is a dense subset of the interval $[0, n]$ using a greedy algorithm to choose A', then we use a celebrated theorem of Sanders which guarantees us a large generalized arithmetic progression P in $2A - 2A$ when A has small doubling. We use techniques from convex geometry to construct order preserving Freiman 2-isomorphism from a generalized arithmetic progression on the interval $[0, n]$.

CHAPTER I

ESTIMATES OF THE DISCREPANCY FUNCTION IN EXPONENTIAL ORLICZ SPACES

We prove that in all dimensions $n \geq 3$ for every integer $N \geq 1$ there exists a distribution of points $\mathcal{P} \subset [0,1]^n$ of cardinality N, for which the associated discrepancy function $\mathcal{D}_N$ satisfies the estimate

$$\|\mathcal{D}_N\|_{\exp\left(L^{\frac{2}{n+1}}\right)} \lesssim (\log N)^{\frac{n-1}{2}} .$$

This has recently been proved by M. Skriganov, using random digit shifts of binary digital nets, building upon the remarkable examples of W. Chen and M. Skriganov. Our approach, developed independently, complements that of Skriganov.

1.1 Introduction

For convenience, $\mathbb{N}$ denotes the set of positive integers, $\mathbb{N}_0$ denotes the set of non-negative integers and if S is a finite set, then $\sharp S$ denotes the number of elements of S.

Given a collection $\mathcal{P}$ of N points in the unit cube $[0,1]^n$ in dimension n, the discrepancy function associated to $\mathcal{P}$ is defined as

$$\mathcal{D}_N[\mathcal{P}, X] := \sharp(\mathcal{P} \cap [0,X]) - N|[0,X]| , \tag{3}$$

where $[0,X]$ is the rectangular box anchored at the origin and $X = (x_1, ..., x_n) \in [0,1]^n$. The optimal L^p estimates for the discrepancy function are well-known, aside from the endpoint cases of $p = 1, \infty$. In this article we continue the theme begun in [11] and extend it to higher dimensions, focusing on the exponential Orlicz space estimates for the discrepancy function in dimensions $n \geq 3$.

Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be an increasing convex function with $\psi(0) = 0$. The Orlicz space L^ψ associated to ψ is the class of functions for which the norm

$$\|f\|_{L^\psi} := \inf\left\{K > 0 : \int_{[0,1]^n} \psi\big(|f(x)|/K\big) dx \leq 1\right\} \tag{4}$$

is finite. The exponential Orlicz spaces $\exp(L^\alpha)$ are Orlicz spaces associated to the function ψ which equals $e^{x^\alpha} - 1$ for large x. Exponential Orlicz norms have different equivalent definitions. The one that is most important for this chapter is

$$\|f\|_{\exp(L^\alpha)} \simeq \sup_{q \geq 1} q^{\frac{-1}{\alpha}} \|f\|_q, \tag{5}$$

which allows one to estimate the exponential norm by estimating the L^q norms and carefully keeping track of the constants.

We prove the following theorem, which as this chapter was in final edits, we discovered had been proved by M. Skriganov [30].

Theorem 1.1.1. *In all dimensions $n \geq 2$ for every integer $N \geq 1$ there exists a distribution $\mathcal{P} \subset [0,1]^n$ of N points such that*

$$\|\mathcal{D}_N\|_{\exp\left(L^{\frac{2}{n+1}}\right)} \lesssim (\log N)^{\frac{n-1}{2}}. \tag{6}$$

It is well-known that the right-hand side of (6) is optimal (since it is the best bound for the L^q norms in dimension $n \geq 2$ [23, 24, 15]), however, the left-hand side does not seem to be. Skriganov indicates that this conjecture is indeed true:

Conjecture 1.1.2. *For dimensions $n \geq 2$, for all integers $N \geq 1$, there is a choice of $\mathcal{P}$ of cardinality N so that*

$$\|\mathcal{D}_N\|_{\exp\left(L^{\frac{2}{n-1}}\right)} \lesssim (\log N)^{\frac{n-1}{2}}.$$

In dimension $n = 2$ this statement has been proved in [11] using the digit shifts of the famous van der Corput set, and it has also been shown that it is best possible, i.e. all N-point distributions satisfy the reverse bound.

In dimensions $n \geq 3$, the first *explicit (non-random)* point distributions with $\|\mathcal{D}_N\|_p \lesssim (\log N)^{\frac{n-1}{2}}$ are the remarkable examples obtained by Chen and Skriganov [13] (in L^2) and Skriganov [29] (L^p, $1 < p < \infty$), see also [28]. In [14] Chen and Skriganov also considered *random* digit shifts of simpler constructions and showed that they too, on the average, have optimal L^2 norm of the discrepancy function.

The study of these constructions exhibits striking similarities to themes related to small ball problems and expansions of the Brownian sheet. The fine analysis of these objects is closely related to the (infamous) $p = \infty$ endpoint estimates of the discrepancy function [8], also see [7, 9] for more background on the subject and the techniques.

A heuristic informed by these connections suggests that the conjecture should be proved by estimating the L^q norm using Littlewood-Paley inequalities $n-1$ times, with each application giving one square root of q, see §1.6. This is just what we will do, but at a specific point in the proof we accumulate one more power of q. In the language of Skriganov, the Littlewood-Paley inequalities are the Khinchin inequalities; in that argument, he applies them n times.

The authors discovered the work of Skriganov at the final stages of the editing of this manuscript. The basic examples are of the same nature, but there are differences in the details of the proof. Certainly, the analysis of these examples is subtle, and it may take some time to tease out its different variants and details.

In an earlier breakthrough work [29] Skriganov showed that for each fixed $1 < q < \infty$ and integer N there is a *deterministic* distribution $\mathcal{P}$ with $\|\mathcal{D}_N\|_q \lesssim p^{2n} q^{\frac{n+1}{2}} (\log N)^{\frac{n-1}{2}}$, where p is a prime greater than qn^2, hence the real power of q is $\frac{5n+1}{2}$.[1] In [31], Skriganov studies the mean behaviour of the discrepancy function, in terms of the digit shifts. Remarkably, the L^q norms do not depend very much on the choice of the shift.

1.2 Linear Distributions

Our proofs will assume that N is a power of 2. A standard argument then implies the theorem as stated. If $2^{s-1} \le N < 2^s$, construct a distribution with 2^s points in $[0,1]^n$ with low discrepancy and take $a > 1/2$ such that the cube $[0,a]^n$ contains N points from the distribution. We get N points in $[0,1]^n$ with low discrepancy by scaling those points inside $[0,a]^n$ by the factor of $1/a < 2$ in each coordinate. See for instance the beginning of [13, §3].

Let $U = [0,1]$. We shall consider distributions $D \subset U^n$ which have the structure of a vector space over the finite field $\mathbb{F}_2$. (More general finite fields can be used, but with this simplest model, the more familiar Rademacher functions reveal themselves.) For $s \in \mathbb{N}_0$,

[1] In equation [13, (1.7)], the estimate is given in terms of a constant in a Littlewood-Paley inequality, which is no more than $C_n q^{\frac{n}{2}}$.

let $\mathbb{Q}(2^s) = \{m2^{-s} \ : \ 0 \le m < 2^s\} \subset U$. Each $x \in \mathbb{Q}(2^s)$ can be written in the form

$$x = \sum_{i=1}^{s} \xi_i(x) 2^{-s+i-1} = \sum_{i=1}^{s} \eta_i(x) 2^{-i} \tag{7}$$

with coefficients $\xi_i(x) = \eta_{s-i+1}(x) \in \mathbb{F}_2$ for $1 \le i \le s$. For $x, y \in \mathbb{Q}(2^s)$, and $\alpha, \beta \in \mathbb{F}_2$, define $\alpha x \oplus \beta y$ through

$$\eta_i(\alpha x \oplus \beta y) = \alpha \eta_i(x) + \beta \eta_i(y) \mod 2\,,$$

Then $\mathbb{Q}(2^s)$ is a vector space over $\mathbb{F}_2$ of dimension s.

In dimension $n \ge 2$ we consider $\mathbb{Q}^n(2^s)$ and extend the definition of $\oplus$ coordinate-wise, making $\mathbb{Q}^n(2^s)$ an ns-dimensional vector space over $\mathbb{F}_2$.

Definition 1.2.1. *We say that $D \subset \mathbb{Q}^n(2^s)$ is a linear distribution if D is a subspace of $\mathbb{Q}^n(2^s)$.*

The inner product on $\mathbb{Q}(2^s)$ is defined by

$$\langle x, y \rangle = \langle y, x \rangle = \sum_{i=1}^{s} \xi_i(x) \xi_{s-i+1}(y)\,.$$

This particular structure is dictated by the definition of Walsh functions, see §1.3. For $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$ in $\mathbb{Q}^n(2^s)$, we write

$$\langle X, Y \rangle = \langle Y, X \rangle = \sum_{j=1}^{n} \langle x_j, y_j \rangle\,.$$

We will frequently write vectors as capital letters and their coordinates as lower case letters, for example, $X = (x_1, \ldots, x_n)$, $K = (k_1, \ldots, k_n)$, $L = (\ell_1, \ldots, \ell_n)$, and this convention will be used without further explanation.

For any distribution D in $\mathbb{Q}^n(2^s)$, we define the *dual distribution* $D^\perp$ to be the set of $X \in \mathbb{Q}^n(2^s)$ with $\langle X, Y \rangle = 0$ for all $Y \in D$. It follows that $D^\perp$ is a subspace of $\mathbb{Q}^n(2^s)$, hence also a linear distribution. Furthermore, we have $(D^\perp)^\perp = D$, so that D and $D^\perp$ are mutually dual distributions.

Consider the Rosenbloom–Tsfasman weight defined by

$$\rho(x) = \begin{cases} 0, & \text{if } x = 0, \\ \max\{i \ : \ \xi_i(x) \ne 0\}, & \text{if } x \ne 0, \end{cases} \tag{8}$$

i.e. the index of the first non-zero binary digit in the expansion of x. It is easy to see that these satisfy the triangle inequality on $\mathbb{Q}(2^s)$. They are extended to $\mathbb{Q}^n(2^s)$ by the formula $\rho(X) = \sum_{i=1}^n \rho(x_i)$ for $X = (x_1, \ldots, x_n) \in \mathbb{Q}^n(2^s)$. Obviously, $\rho(X) = 0$ iff $X = 0$.

If D is a linear distribution, we define its Rosenbloom–Tsfasman weight $\rho(D)$ to be the minimum of $\rho(X)$ over $X \in D \setminus \{0\}$.

Remark 1.2.1. *In the works of Chen–Skriganov [13] and Skriganov [29], the more familiar Hamming metric is also used in order to gain (super) orthogonality relations for integrals of Walsh functions. In our work, as in [30, 14], orthogonality is achieved by averaging over random digit shifts instead.*

1.3 Walsh Functions

We write $\mathbb{Q}(2^\infty) = \bigcup_{s\in\mathbb{N}_0} \mathbb{Q}(2^s)$. The notion of $\oplus$ addition can be defined on this set, making $\mathbb{Q}(2^\infty)$ an infinite dimensional vector space over $\mathbb{F}_2$. Each $\lambda \in \mathbb{N}_0$ can be written as $\sum_{i=1}^\infty \lambda_i(\ell)2^{i-1}$, where the coefficients $\lambda_i(\ell) \in \mathbb{F}_2$ for every $i \in \mathbb{N}$ and only finitely many are non-zero. With this notation, we can extend the notion of $\oplus$ to $\mathbb{N}_0$: $\ell \oplus k$ is the integer j such that for all $i \in \mathbb{N}$,

$$\lambda_i(j) = \lambda_i(\ell) + \lambda_i(k) \mod 2\,.$$

We define the Walsh functions on U by

$$w_\ell(x) = \exp\left(\pi i \sum_{i=1}^\infty \lambda_i(\ell)\eta_i(x)\right) = (-1)^{\sum_{i=1}^\infty \lambda_i(\ell)\eta_i(x)} \tag{9}$$

where $\eta_i(x)$ are as in (7). A detailed study of these functions can be found in [26]. The set of functions $\{w_\ell \,:\, \ell \in \mathbb{N}_0\}$ forms an orthonormal basis for $L^2(U)$: for every $f \in L^2(U)$

$$f \simeq \sum_{\ell\in\mathbb{N}_0} \langle f, w_\ell\rangle w_\ell$$

with $\simeq$ indicating that the sum on the right converges to f in the L^2 metric. It is also relevant for us that there is an explicit formula connecting Walsh expansions and conditional expectations.

$$\sum_{\ell=0}^{2^s-1} \langle f, w_\ell\rangle w_\ell = 2^s \sum_{t=1}^{2^s} \int_{(t-1)2^{-s}}^{t2^{-s}} f(y)\,dy \cdot \mathbf{1}_{[(t-1)2^{-s}, t2^{-s})}\,. \tag{10}$$

It is also the case that w_ℓ are the characters of the group $\mathbb{Q}(2^\infty)$. In particular, $w_\ell(x \oplus y) = w_\ell(x) w_\ell(y)$, and $w_{\ell \oplus k}(x) = w_\ell(x) \cdot w_k(x)$ for all $\ell, k \in \mathbb{N}_0$ and $x, y \in U$.

In dimension n, the notion of $\oplus$ can be extended coordinate-wise to $\mathbb{N}_0^n$ and likewise to $\mathbb{Q}^n(2^\infty)$. For $L = (\ell_1, \ldots, \ell_n) \in \mathbb{N}_0^n$ and $X \in \mathbb{Q}^n(2^\infty)$, we define

$$W_L(X) = \prod_{j=1}^{n} w_{\ell_j}(x_j)\,.$$

The properties mentioned above continue to hold for these Walsh functions. The collection $\{W_L \ : \ L \in \mathbb{N}_0\}$ forms an orthonormal basis for $L^2(U^n)$, and the W_L are group characters with respect to $\oplus$. In particular, for all $L, K \in \mathbb{N}_0^n$

$$\langle W_L, W_K \rangle = \int_{U^n} W_L W_K \; dx = \int_{U^n} W_{L \ominus K} \; dx = \begin{cases} 1, & L = K, \\ 0, & L \neq K. \end{cases}$$

There are some useful consequences of W_L being the group characters, which we collect here. Consider the vector space over $\mathbb{F}_2$ given by

$$\mathbb{N}_0^n(2^s) := \{L = (\ell_1, \ldots, \ell_n) \in \mathbb{N}_0^n \ : \ 0 \leq \ell_i < 2^s, \ 1 \leq i \leq n\}\,.$$

Obviously, the map

$$\theta \ : \ \mathbb{Q}^n(2^s) \to \mathbb{N}_0^n(2^s) \ : \ (x_1, \ldots, x_n) \mapsto (2^s x_1, \ldots, 2^s x_n)$$

is a vector space isomorphism. The following variant of the Poisson summation formula holds.

Lemma 1.3.1. *For every linear distribution $D \subset \mathbb{Q}^n(2^s)$ and every $L \in \mathbb{N}_0^n(2^s)$, it holds that*

$$\sum_{X \in D} W_L(X) = \begin{cases} \sharp D, & L \in \theta(D^\perp), \\ 0, & L \notin \theta(D^\perp). \end{cases}$$

And for every $X \in \mathbb{Q}^n(2^s)$

$$\sum_{L \in \theta(D)} W_L(X) = \begin{cases} \sharp D, & X \in D^\perp, \\ 0, & X \notin D^\perp. \end{cases}$$

Using the isomorphism θ, we can define $\rho(\ell)$ and $\rho(L)$. In particular, for $\ell \in \mathbb{N}_0(2^s)$, we write $\ell = \sum_{i=1}^{s} \lambda_i(\ell) p^{i-1}$, and then $\rho(\ell)$ is the largest i with $\lambda_i(\ell) \neq 0$. We furthermore set

$$\lambda(\ell) := \lambda_{\rho(\ell)}(\ell)\,, \qquad \tau(\ell) := \ell - \lambda(\ell) p^{\rho(\ell)-1}\,. \tag{11}$$

So, $\lambda(\ell)$ is the most significant digit of ℓ, and $\tau(\ell)$ is ℓ less its most significant term in the dyadic expansion of ℓ (we shall say that $\tau(\ell)$ is the *truncation* of ℓ). For $L \in \mathbb{N}_0^n(2^s)$ we set

$$\rho(L) = \sum_{i=1}^{n} \rho(\ell_i), \qquad \overline{\rho}(L) = \big(\rho(\ell_1), ..., \rho(\ell_n)\big), \qquad \tau(L) = \big(\tau(\ell_1), ..., \tau(\ell_n)\big). \tag{12}$$

1.4 Approximation of the Discrepancy Function

Let $\chi(y, \cdot)$ be the indicator of the interval $[0, y) \subset U$, i.e.

$$\chi(y, x) := \begin{cases} 1 & 0 \leq x < y \\ & \\ 0 & y \leq x < 1 \end{cases}.$$

This function has the following Walsh expansion

$$\chi(y, x) \simeq \sum_{\ell \in \mathbb{N}_0} \widetilde{\chi}_\ell(y) w_\ell(x),$$

where $\widetilde{\chi}(y) = \langle \chi(y, \cdot), w_\ell \rangle = \displaystyle\int_0^y w_\ell(x)dx$, and in particular, $\widetilde{\chi}_0(y) = y$. For $s \in \mathbb{N}_0$, we truncate the Walsh expansion above to

$$\chi_s(y, x) = \sum_{\ell \in \mathbb{N}_0(2^s)} \widetilde{\chi}_\ell(y) w_\ell(x)\,.$$

This is extended to n dimensions in an obvious way. For $X, Y \in U^n$, we write

$$\chi(Y, X) := \prod_{j=1}^{n} \chi(y_j, x_j)\,,$$

$$\chi_s(Y, X) := \prod_{j=1}^{n} \chi_s(y_j, x_j)\,,$$

$$\mathcal{M}[D; Y] := \sum_{X \in D} \chi_s(Y, X) - 2^s \prod_{j=1}^{n} y_j\,.$$

The first is the indicator of the box in U^n, anchored at the origin and Y; the second is a truncation of the Walsh expansion of the first; and the third is an approximation of the discrepancy function $\mathcal{D}_N[D,Y]$, since according to (3)

$$\mathcal{D}_N[D;Y] := \sum_{X\in D} \chi(Y,X) - 2^s \prod_{j=1}^{n} y_j \,.$$

For $T \in \mathbb{Q}^n(2^s)$ the digit shift $D \oplus T$ is defined as $D \oplus T = \{X \oplus T : \ X \in D\}$. The following important observation of Chen and Skriganov [13, Lemma 6A] shows that $\mathcal{M}[D \oplus T; Y]$ is indeed a good approximation to the discrepancy function.

Lemma 1.4.1. *Suppose that $D \subset \mathbb{Q}^n(2^s)$ is a linear distribution of $N = 2^s$ points with dual linear distribution $D^\perp$ satisfying the bound $\rho(D^\perp) \geq s - \delta + 1$, and let $T \in \mathbb{Q}^n(2^s)$. Then*

$$\|\mathcal{D}_N[D\oplus T, Y] - \mathcal{M}[D \oplus T, Y]\|_{L^\infty(X)} \leq n2^\delta \lesssim 1\,.$$

Below, constants that only depend upon the dimension n will not be systematically tracked. The usefulness of this approximation is that $\mathcal{M}[D \oplus T; Y]$ can be expressed by a remarkably succinct formula. Using Poisson summation, Lemma 1.3.1, we obtain

$$\begin{aligned}
\mathcal{M}[D\oplus T;Y] &= \sum_{X\in D} \sum_{L\in\mathbb{N}_0^n(2^s)} \widetilde{\chi}_L(Y) W_L(X\oplus T) - 2^s \prod_{j=1}^{n} y_j \\
&= \sum_{L\in\mathbb{N}_0^n(2^s)} \Bigg\{ \sum_{X\in D} W_L(X\oplus T) \Bigg\} \widetilde{\chi}_L(Y) - 2^s \prod_{j=1}^{n} y_j \\
&= 2^s \sum_{L\in\theta(D^\perp)\setminus\{0\}} W_L(T)\widetilde{\chi}_L(Y).
\end{aligned} \tag{13}$$

since $W_L(X\oplus T) = W_L(X)W_L(T)$ and $\chi_{\overline{0}}(Y) = \prod_{j=1}^{n} y_j$.

Recall that $\widetilde{\chi}_L(Y) = \prod_{j=1}^{n} \widetilde{\chi}_{\ell_j}(y_j)$. Formulas of Fine [18] (later extended by Price [22] to p-adic Walsh functions and known as Fine-Price formulas) give a precise expansion of the functions $\tilde{\chi}_\ell$. For every $\ell \in \mathbb{N}_0$, we have

$$\widetilde{\chi}_\ell(y) = 2^{-\rho(\ell)} u_\ell(y), \quad \text{where} \quad u_\ell(y) = \tfrac{1}{2}\Big(w_{\tau(\ell)}(y) - \sum_{i=1}^{\infty} 2^{-i} w_{\ell+2^{\rho(\ell)+i-1}}(y)\Big)\,. \tag{14}$$

The equality above holds for $\ell = 0$ as well, with the understanding that $\tau(0) = \rho(0) = 0$.

Recall that for $x \in U$, we write $x = \sum_{i=1}^{\infty} \eta_i 2^{-i}$, where $\eta_i(x) \in \{0,1\}$. The Rademacher functions are defined as

$$r_i(x) = (-1)^{\eta_i(x)}. \tag{15}$$

In particular, $w_{2^k} = r_{k+1}$. We then have the following representation.

Lemma 1.4.2. *For any $\ell \in N_0$ we have*

$$\widetilde{\chi}_\ell(y) = 2^{-\rho(\ell)-1} w_{\tau(\ell)}(y)\omega_{\rho(\ell)}(y), \ \ \textit{where}$$
$$\omega_{\rho(\ell)}(y) = 1 - \sum_{i=1}^{\infty} 2^{-i} r_{\rho(\ell)}(y) r_{\rho(\ell)+i}(y),$$

and $r_k(y)$ are the Rademacher functions.

The function $\omega_{\rho(\ell)}(y)$ is continuous and piecewise linear with a period of $2^{-\rho(\ell)+1}$.

Proof. As $\ell + 2^{\rho(\ell)+i-1} = \tau(\ell) \oplus 2^{\rho(\ell)-1} \oplus 2^{\rho(\ell)+i-1}$ then

$$w_{\ell+2^{\rho(\ell)+i-1}}(y) = w_{\tau(\ell)}(y) w_{2^{\rho(\ell)-1}}(y) w_{2^{\rho(\ell)+i-1}}(y) = w_{\tau(\ell)}(y) r_{\rho(\ell)}(y) r_{\rho(\ell)+i}(y).$$

Which along with (14) proves Lemma 1.4.2. □

Remark 1.4.1. *Lemma 1.4.2 may be explained and proved without appealing to the Fine-Price formula* (14)*. Indeed, the integral of a Rademacher function $\int_0^y r_k(x)dx$ is the 2^{-k+1}-periodic "saw-tooth" function. Hence, the integral of the Walsh function $w_\ell = r_{\rho(\ell)} \cdot w_{\tau(\ell)}$ also has this structure, but with sign changes on dyadic intervals of length 2^{-k+1} dictated by the sign of $w_{\tau(\ell)}$. One can easily check that on $[0,1]$ $x = \frac{1}{2}\big(1 - \sum_{i=1}^{\infty} 2^{-i} r_i(x)\big)$ and therefore the* 1*-periodic "saw-tooth" function $|||x|||$, i.e. the distance from x to the nearest integer, satisfies*

$$|||x||| = \frac{1}{2}\big(1 - \sum_{i=1}^{\infty} 2^{-i} r_1(x) r_i(x)\big) = \frac{1}{2^2}\big(1 - \sum_{i=1}^{\infty} 2^{-i} r_1(x) r_{1+i}(x)\big).$$

The rest follows by rescaling.

1.5 The Rademacher Functions and Shifts

We say that a distribution D with $N = 2^s$ points is a dyadic net with deficiency δ if each dyadic box of volume $2^{-s+\delta}$ in U^n contains precisely 2^δ points of D. It is well known that this is equivalent to the fact that $\rho(D^\perp) \geq s - \delta + 1$ (see e.g. Lemma 2C in [13]). While dyadic nets with deficiency zero do not exist in dimensions $n > 3$, one can construct dyadic nets with deficiency δ of the order $n \log n$ in any dimension. See the book [16] for a detailed treatment of digital nets.

Assume that D is a dyadic net with deficiency δ and return to formula (13):

$$\mathcal{M}[D \oplus T; Y] = 2^s \sum_{L \in \theta(D^\perp)\setminus\{0\}} W_L(T)\,\widetilde{\chi}_L(Y). \tag{16}$$

Switch to the vector notation, setting $Y = (y_1, ..., y_n)$, $L = (\ell_1, ..., \ell_n)$, $\overline{\rho}(L) = (\rho(\ell_1), ..., \rho(\ell_n))$, $\omega_{\overline{\rho}(L)}(Y) = \prod_{i=1}^{n} \omega_{\rho(\ell_i)}(y_i)$, and $r_{\overline{\rho}(L)}(Y) = \prod_{i=1}^{n} r_{\rho(\ell_i)}(y_i)$. Applying Lemma 1.4.2 to the summands above, we obtain

$$\begin{aligned} W_L(T)\,\widetilde{\chi}_L(Y) &= 2^{-n-\rho(L)} W_L(T) W_{\tau(L)}(Y) \omega_{\overline{\rho}(L)}(Y), \\ &= 2^{-n-\rho(L)} r_{\overline{\rho}(L)}(Y) W_L(T) W_L(Y) \omega_{\overline{\rho}(L)}(Y), \qquad (\text{since } W_{\tau(L)} = r_{\overline{\rho}(L)} W_L) \\ &= 2^{-n-\rho(L)} r_{\overline{\rho}(L)}(Y) \omega_{\overline{\rho}(L)}(Y) W_L(Y \oplus T). \end{aligned}$$

Whence we have

$$\mathcal{M}[D \oplus T; Y] = \sum_{L \in \theta(D^\perp)\setminus\{0\}} 2^{s-n-\rho(L)} r_{\overline{\rho}(L)}(Y) \omega_{\overline{\rho}(L)}(Y) W_L(Y \oplus T). \tag{17}$$

This leads to the following consequence for the L^q norms of this sum.

Lemma 1.5.1. *Let the distribution D with $N = 2^s$ points be a dyadic net with deficiency δ. For any $1 \le q < \infty$ we have*

$$\|\mathcal{M}[D \oplus T; Y]\|_{L^q[Y \times T]} \le C q^{\frac{n+1}{2}} s^{\frac{n-1}{2}},$$

where the implicit constant depends only on the dimension n and deficiency δ.

In view of Lemma 1.4.1, it clearly completes the proof of our main theorem, Theorem 1.5.4. Indeed, this inequality implies that $\mathcal{D}_N[D \oplus T, Y]$ satisfies the $\exp\left(L^{\frac{2}{n+1}}\right)$ bound as a function of two variables, Y and T. Therefore, for some T it has to satisfy this bound in Y.

Proof. It is convenient to prove the lemma for q replaced by $2q$, with $q \in \mathbb{N}$. The following elementary fact will be used: for an integrable function $f : U^n \to \mathbb{R}$ and fixed $Z \in U^n$ we have

$$\int_{U^n} f(Y) dY = \int_{U^n} f(Y \oplus Z) dY. \tag{18}$$

According to this, it suffices to estimate the $L^{2q}[Y \times T]$ norm of $\mathcal{M}[D \oplus T; Y \oplus T]$.

The latter has a more symmetric expansion. From (17) we get

$$\mathcal{M}[D \oplus T; Y \oplus T] = \sum_{L \in \theta(D^\perp)\setminus\{0\}} 2^{s-n-\rho(L)} r_{\overline{\rho}(L)}(Y \oplus T)\omega_{\overline{\rho}(L)}(Y \oplus T) W_L(Y), \tag{19}$$

and by grouping the summands in (19) which have the same $\overline{\rho}(L)$ we obtain

$$= \sum_{\overline{\rho} \in \mathbb{N}(s)^n,\ |\overline{\rho}| > s-\delta} 2^{s-n-|\overline{\rho}|} r_{\overline{\rho}}(Y \oplus T)\omega_{\overline{\rho}}(Y \oplus T) \sum_{L \in \Lambda(\overline{\rho})} W_L(Y), \tag{20}$$

where $|\overline{\rho}| = \rho_1 + ... + \rho_n$ is the ℓ^1 norm of $\overline{\rho}$ and $\Lambda(\overline{\rho}) = \{L \in \theta(D^\perp) :\ \overline{\rho}(L) = \overline{\rho}\}$. The latter is an affine copy of the subspace

$$\Lambda_0(\overline{\rho}) = \{L \in \theta(D^\perp) :\ \overline{\rho}(L) < \overline{\rho}\}\,. \tag{21}$$

Lemma 1.5.2. *If $\rho(D^\perp) \geq s - \delta + 1$, then the cardinality of $\Lambda_0(\overline{\rho})$ satisfies*

$$2^{|\overline{\rho}|-s} \leq \sharp\Lambda_0(\overline{\rho}) \leq 2^{|\overline{\rho}|-s+\delta}. \tag{22}$$

Proof. Observe that $\Lambda_0(\overline{\rho})$ is $\theta(D^\perp)$ restricted to a dyadic box of area $2^{|\overline{\rho}|}$. Divide the box into $2^{|\overline{\rho}|-s+\delta}$ congruent boxes of volume $2^{s-\delta}$. Since $\rho(D^\perp) \geq s - \delta + 1$, each such box contains no more than one point of $\theta(D^\perp)$, for otherwise the difference of the two points would yield a non-zero point of $L = L_1 \ominus L_2 \in \theta(D^\perp)$ with $\rho(L) \leq s - \delta$. This proves the right inequality.

On the other hand, divide the cube $\theta(U^n)$ into $2^{ns-|\overline{\rho}|}$ disjoint dyadic boxes of dimensions $2^{\rho_1} \times \cdots \times 2^{\rho_n}$, which have volume $2^{|\overline{\rho}|}$. The intersection of each such box with $\theta(D^\perp)$, if non-empty, is an affine copy (digit shift) of $\Lambda_0(\overline{\rho})$. Therefore, $2^{(n-1)s} = \sharp D^\perp \leq \sharp\Lambda_0(\overline{\rho}) \cdot 2^{ns-|\overline{\rho}|}$, which proves the left inequality. $\square$

The sum $\sum_{L \in \Lambda(\overline{\rho})} W_L(Y)$ can be rewritten using the Poisson summation formula (Lemma 1.3.1).

$$\sum_{L \in \Lambda(\overline{\rho})} W_L(Y) = W_{L_{\overline{\rho}}}(Y) \sum_{L \in \Lambda_0(\overline{\rho})} W_L(Y) = W_{L_{\overline{\rho}}}(Y) \cdot \sharp\Lambda_0(\overline{\rho}) \cdot \delta(\overline{\rho}, Y), \tag{23}$$

where $L_{\overline{\rho}}$ is any point in $\Lambda(\overline{\rho})$ and

$$\delta(\overline{\rho}, Y) = \begin{cases} 1, & Y \perp \{X \in D^\perp :\ \overline{\rho}(X) < \overline{\rho}\},\ \text{i.e. } Y \in \theta^{-1}(\Lambda_0(\overline{\rho}))^\perp, \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

The orthogonality condition in (24) is understood by truncating the extra binary digits (above s^{th}) of Y in each coordinate. We can easily see from (22) that

$$\int_{[0,1]^n} \delta(\overline{\rho}, Y)dY = \sharp\big[\Lambda_0(\overline{\rho})\big]^{\perp} \cdot 2^{-ns} \leq 2^{s-|\overline{\rho}|}. \tag{25}$$

Combining (20) and (23) we obtain

$$\begin{aligned}\mathcal{M}[D \oplus T; Y \oplus T] &= \sum_{\overline{\rho}\in\mathbb{N}^n(s),\ |\overline{\rho}|>s} 2^{s-n-|\overline{\rho}|} r_{\overline{\rho}}(Y\oplus T)\omega_{\overline{\rho}}(Y\oplus T)W_{L_{\overline{\rho}}}(Y)\cdot\sharp\Lambda_0(\overline{\rho})\delta(\overline{\rho}, Y) \\ &= 2^{-n}\sum_{k=s-\delta+1}^{ns} M_k(T, Y)\,, \end{aligned}\tag{26}$$

$$\text{where}\quad M_k(T, Y) := \sum_{\overline{\rho}\in\mathbb{N}^n(s),\ |\overline{\rho}|=k} 2^{s-k}\, r_{\overline{\rho}}(Y\oplus T)\omega_{\overline{\rho}}(Y\oplus T)W_{L_{\overline{\rho}}}(Y)\cdot\sharp\Lambda_0(\overline{\rho})\delta(\overline{\rho}, Y).$$

The variables Y and T can be decoupled. By (18), the $L^{2q}[Y \times T]$ norm of $M_k[D \oplus T; Y \oplus T]$ equals the $L^{2q}[Y \times T]$ norm of

$$M_k'(T, Y) = \sum_{\overline{\rho}\in\mathbb{N}^n(s),\ |\overline{\rho}|=k} 2^{s-k}\, r_{\overline{\rho}}(T)\omega_{\overline{\rho}}(T)W_{L_{\overline{\rho}}}(Y)\cdot\sharp\Lambda_0(\overline{\rho})\cdot\delta(\overline{\rho}, Y)\,.$$

Let us rewrite the function $r_{\overline{\rho}}(T)\omega_{\overline{\rho}}(T)$. Using Lemma 1.4.2, since $r_m^2 = 1$, we have

$$r_m(t)\omega_m(t) = r_m(t) - \sum_{i=1}^{\infty} 2^{-i} r_{m+i}(t).$$

This implies that

$$r_{\overline{\rho}}(T)\omega_{\overline{\rho}}(T) = \sum_{\overline{\imath}\in\mathbb{N}_0^n} \epsilon_{\overline{\imath}} 2^{-|\overline{\imath}|} r_{\overline{\rho}+\overline{\imath}}(T).$$

Here, $\epsilon_{\overline{\imath}}$ is -1 raised to the number of non-zero entries of $\overline{\imath}$. Therefore

$$\begin{aligned} M_k'(T, Y) &= \sum_{\overline{\imath}\in\mathbb{N}_0^n} \epsilon_{\overline{\imath}} 2^{-|\overline{\imath}|} \sum_{\overline{\rho}\in\mathbb{N}^n(s),\ |\overline{\rho}|=k} 2^{s-k}\, r_{\overline{\rho}+\overline{\imath}}(T)W_{L_{\overline{\rho}}}(Y)\cdot\sharp\Lambda_0(\overline{\rho})\cdot\delta(\overline{\rho}, Y) \\ &=: \sum_{\overline{\imath}\in\mathbb{N}_0^n} \epsilon_{\overline{\imath}} 2^{-|\overline{\imath}|} M_k^{\overline{\imath}}(T, Y). \end{aligned}\tag{27}$$

We estimate the $L^{2q}(Y \times T)$ norm of $M_k^{\overline{\imath}}(T, Y)$, which the Littlewood–Paley inequalities are ideally suited for. Applying Lemma 1.6.2 in T and using the fact that $q \in \mathbb{N}$, we obtain

$$\begin{aligned}
\|M_k^{\bar{i}}(T,Y)\|_{L^{2q}[T\times Y]}^{2q} &\leq (Cq)^{q(n-1)}\, 2^{(s-k)2q} \int \Bigg(\sum_{\overline{\rho}\in\mathbb{N}^n(s),\ |\overline{\rho}|=k} [\sharp\Lambda_0(\overline{\rho})]^2 \cdot \delta^2(\overline{\rho},Y) \Bigg)^q \, dY \\
&= (Cq)^{q(n-1)}\, 2^{(s-k)2q} \times \\
&\quad \times \sum_{|\overline{\rho}_1|,\ldots,|\overline{\rho}_q|=k} \prod_{j=1}^{q} \left[\sharp\Lambda_0(\overline{\rho}_j)\right]^2 \int \delta^2(\overline{\rho}_1,Y)\cdots\delta^2(\overline{\rho}_q,Y)\, dY \\
&\leq (Cq)^{q(n-1)}\, 2^{(s-k)2q} 2^{(k-s+\delta)2q} \cdot \sum_{|\overline{\rho}_1|,\ldots,|\overline{\rho}_q|=k} \int \delta(\overline{\rho}_1,Y)\cdots\delta(\overline{\rho}_q,Y)\, dY \\
&\leq (Cq)^{q(n-1)} \sum_{|\overline{\rho}_1|,\ldots,|\overline{\rho}_q|=k} \int \delta(\overline{\rho}_1,Y)\cdots\delta(\overline{\rho}_q,Y)\, dY \\
&\leq (Cq)^{q(n-1)}\, s^{q(n-1)} \int \delta(\overline{\rho}_1,Y)\, dY \\
&\leq (Cq)^{q(n-1)}\, s^{q(n-1)}\, 2^{s-k}. \qquad (28)
\end{aligned}$$

The constant C changes from line to line above. The first line is the Littlewood–Paley inequality (Lemma 1.6.2); in the following we use the facts that (i) the number of $\overline{\rho}$ with $|\overline{\rho}| = k$ is at most $Ck^{n-1} < C's^{n-1}$; (ii) that $\sharp\Lambda_0(\overline{\rho}) \leq 2^{|\overline{\rho}|-s+\delta}$, cf. (22); (iii) estimate (25) together with $\delta^2(\overline{\rho},Y) = \delta(\overline{\rho},Y)$.

Because of the geometric decay in (27), the above computation yields

$$\|M_k'(T,Y)\|_{L^{2q}[T\times Y]} \leq Cq^{\frac{n-1}{2}} s^{\frac{n-1}{2}} 2^{\frac{s-k}{2q}}.$$

Note in particular the exponent of 2 above, which will lead to one additional power of q in our estimate. Recall that $\|M_k(T,Y)\|_{L^{2q}[T\times Y]} = \|M_k'(T,Y)\|_{L^{2q}[T\times Y]}$, thus from (26) we obtain

$$\begin{aligned}
\|\mathcal{M}[D\oplus T; Y\oplus T]\|_{L^{2q}[T\times Y]} &\leq Cq^{\frac{n-1}{2}} s^{\frac{n-1}{2}} \sum_{k=s+1}^{ns} 2^{\frac{s-k}{2q}} \\
&\leq Cq^{\frac{n+1}{2}} s^{\frac{n-1}{2}} \leq Cq^{\frac{n+1}{2}} s^{\frac{n-1}{2}}, \qquad (29)
\end{aligned}$$

since the sum in the first line is $\mathcal{O}(q)$. This completes the proof. □

The reader interested in further improvements in arguments of this type will quickly focus on the fact that this method of proof uses the Rademacher structure, but exploits very little information (essentially just (25)) about the coefficients of the Rademacher functions. The first point where one would like to do much better is estimate (28) above: here the integral of the q-fold product of $\delta(\overline{\rho}, Y)$ is estimated by the integral of a single $\delta(\overline{\rho}, Y)$. However we have only found incremental improvements on this point and we leave the topic to the future.

Before we proceed, it might be convenient to point out why Conjecture 1.1.2 represents a natural goal, and why the possible extensions are far from clear. For integers k, one has

$$\Big\| \sum_{\overline{\rho}\,:\,|\overline{\rho}|=k} r_{\overline{\rho}} \Big\|_k \gtrsim k^{(n-1)} 2^{-n},$$

since on the cube $[0, 2^{-k}]^n$, the summands are all of the same sign. On the other hand, the Littlewood–Paley immediately show that $\Big\| \sum_{\overline{\rho}\,:\,|\overline{\rho}|=k} r_{\overline{\rho}} \Big\|_{\exp(L^{\frac{2}{n-1}})} \lesssim k^{\frac{n-1}{2}}$, i.e. according to (5) $\Big\| \sum_{\overline{\rho}\,:\,|\overline{\rho}|=k} r_{\overline{\rho}} \Big\|_q \lesssim q^{\frac{n-1}{2}} k^{\frac{n-1}{2}}$, which by the above is not improvable.

Lemma 1.5.3. *Let If $\rho(D^{\perp}) \geq s - \delta + 1$, then for any $\epsilon > 0$ there is a constant $C = C(\epsilon, n)$ so that for all integers s, $s < k \leq ns$, and $1 \leq q \leq s^{1/2-\epsilon}$, the following inequality holds (see 28)*

$$\int \sum_{|\overline{\rho}_1|,\dots,|\overline{\rho}_q|=k} \delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y)\, dY \leq s^{q(n-1)} a_k, \tag{30}$$

and $\sum_{k=s+1}^{ns} a_k^{1/q} < C$.

Proof. Based the estimate in (28) we can assume that $q > 2/\epsilon$ and $a_k \leq 1$.

The sum above is over the set

$$\sharp\{(\overline{\rho}_1, \dots, \overline{\rho}_q) \;:\; |\overline{\rho}_u| = k\,, 1 \leq u \leq q\} \leq k^{q(n-1)} \lesssim s^{q(n-1)}\,,$$

since $s < k \leq ns$. Converting the estimate of the lemma to expectations over this set, it's equivalent to show that

$$\mathbb{E} \int_{U^n} \delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y)\, dY \leq a_k,$$

where the expectation is taken over independent random vectors $\overline{\rho}_1, \overline{\rho}_2, \dots, \overline{\rho}_q$ with $|\overline{\rho}_1|, \dots, |\overline{\rho}_q| = k$.

Let $E_{\overline{\rho}} := \{X \in D^{\perp} : \ \overline{\rho}(X) < \overline{\rho}\}$. Then,

$$\delta(\overline{\rho}, Y) = \begin{cases} 1, & Y \perp E_{\overline{\rho}}, \\ 0, & \text{otherwise.} \end{cases}$$

It follows from (25) that $\int_{U^n} \delta(\overline{\rho}, Y) \, dY \leq 2^{s-k}$.

And,

$$\delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y) = \begin{cases} 1, & Y \perp E_{\overline{\rho}_1} \cup \cdots \cup E_{\overline{\rho}_q}, \\ 0, & \text{otherwise.} \end{cases}$$

Of course Y is perpendicular to $V = E_{\overline{\rho}_1} \cup \cdots \cup E_{\overline{\rho}_q}$ if and only if it is perpendicular to the linear span of the V. Our strategy is to show that the dimension of $\operatorname{span}(V)$ is large with high probability.

For $\overline{\rho} = (\rho_1, ..., \rho_n)$ with $|\overline{\rho}_q| = k$, $E_{\overline{\rho}}$ has dimension at least $k - s = \Delta - \delta$, for $\Delta = k - s + \delta$. Hence one can choose linearly independent elements of $E_{\overline{\rho}}$ of size $\Delta - \delta$, let it be $X_{\overline{\rho}}$. Notice that for $Y = (y_1, ..., y_n) \in X_{\overline{\rho}}$, $\rho(y_1) \in [\rho_1, \rho_1 - \Delta) =: I(\overline{\rho})$ because $\rho(y_1) \leq \rho_1 - \Delta$ will imply that $\rho(Y) \leq s - \delta$.

Let's assume that $k > s + 1 + \delta$ and construct linearly independent elements belonging to V.

Let A be the event that there are at least m ($m < q$) disjoint intervals among $I(\overline{\rho}_i)$, $i = 1, 2..., q$. If we take the corresponding sets $X_{\overline{\rho}}$ for disjoint intervals $I(\overline{\rho}_i)$, $i = 1, 2..., q$ we will get linearly independent elements of V. If A holds the dimension of $\operatorname{span}(V)$ is at least $m(\Delta - \delta) > \frac{m}{2}\Delta$ and the dimension of $V^{\perp}$ is at most $ns - \frac{m}{2}\Delta$. Provided that A holds we get

$$\int_{U^n} \delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y) \, dY = \sharp(V^{\perp}) \cdot 2^{-ns} \leq 2^{-\frac{m}{2}\Delta}.$$

Now let's consider the case when A doesn't hold, we want to show that it has small probability. Let $\overline{\rho} = (\rho_1, ..., \rho_n)$ be uniformly random vector with $|\overline{\rho}| = k$. We want to get the distribution of the first coordinate ρ_1 inside $[0, s]$. Let $T_n(k)$ be the number of all such vectors, i.e.

$$T_n(k) = \#\{(\rho_1, \rho_2, ..., \rho_n) \in [0, s]^d : \ \rho_1 + \cdots + \rho_n = k\}.$$

The values of $\rho_1, ..., \rho_{n-1}$ will determine the value of ρ_n, so one can easily get the following bounds for $T_n(k)$:

$$\left(\frac{k}{n}\right)^{n-1} < T_n(k) < k^{n-1}.$$

From the definition of $T_n(k)$ we get

$$Pr[\rho_1 = i] = \frac{T_{n-1}(k-i)}{T_n(k)}.$$

And using the bounds above we get

$$Pr[\rho_1 = i] < \frac{(k-i)^{n-2} n^{n-1}}{k^{n-1}}.$$

For $s < k \le ns$ we get

$$Pr[\rho_1 = i] < \frac{C_n}{s}. \tag{31}$$

We can choose m disjoint intervals out of $I(\overline{\rho}_i) = [\rho_{i1}, \rho_{i1} - \Delta),\ i = 1, ..., q$ by choosing m distinct integers out of $\left\lfloor \frac{\rho_{i1}}{2\Delta} \right\rfloor,\ i = 1, ..., q$ which are in the range $[0, \frac{s}{2\Delta}]$ and then considering the corresponding intervals which are easy to verify to be disjoint.

Hence using (31) we get

$$Pr[\neg A] < \binom{s/2\Delta}{m} \left(\frac{C_n 2\Delta m}{s}\right)^q.$$

Using an upper bound for the binomial coefficient we get

$$Pr[\neg A] < \left(\frac{es}{2\Delta m}\right)^m \left(\frac{C_n 2\Delta m}{s}\right)^q < (C_n')^q \left(\frac{\Delta m}{s}\right)^{q-m}.$$

Combining both estimates we get

$$\mathbb{E} \int_{U^n} \delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y)\, dY = Pr[A] \mathbb{E}_A \int_{U^n} ...dY + Pr[\neg A] \mathbb{E}_{\neg A} \int_{U^n} ...dY,$$

$$\mathbb{E} \int_{U^n} \delta(\overline{\rho}_1, Y) \cdots \delta(\overline{\rho}_q, Y)\, dY \le 2^{-\frac{m}{2}\Delta} + (C_n')^q \left(\frac{\Delta m}{s}\right)^{q-m} 2^{-\Delta} =: b_k + c_k =: a_k.$$

Let's take $m = \left\lceil \frac{4q \ln \Delta}{\Delta} \right\rceil$. Then $b_k \le 2^{-2q \ln \Delta}$ and hence $\sum_{k=s+1}^{ns} b_k^{1/q} < C$.

Now let's estimate c_k, we have

$$c_k = (C'_n)^q \left(\frac{\Delta m}{s}\right)^{q-m} 2^{-\Delta} \le (C'_n)^q \left(\frac{4q\ln\Delta + \Delta}{s}\right)^{q-\frac{4q\ln\Delta}{\Delta}+1} 2^{-\Delta}.$$

Hence

$$c_k^{1/q} \le C'_n \left(\frac{4q\ln\Delta + \Delta}{s}\right)^{1-\frac{4\ln\Delta}{\Delta}-1/q} 2^{-\Delta/q}.$$

As $1/q < \epsilon/2$ we can choose Δ_0 depending on ϵ such that if $\Delta > \Delta_0$ then $1-\frac{4\ln\Delta}{\Delta}-1/q > 1-\epsilon$. Hence for $\Delta > \Delta_0$ we have

$$c_k^{1/q} \le C'_n \left(\frac{4q\ln\Delta + \Delta}{s}\right)^{1-\epsilon} 2^{-\Delta/q}.$$

Using this inequality we get

$$\sum_{k=s+1}^{ns} c_k^{1/q} \le \Delta_0 + C'_n s^{-1+\epsilon} \sum_{r=1}^{ns/q} 2^{-r+1} \sum_{\Delta=(r-1)q}^{rq-1} (4q\ln\Delta + \Delta)^{1-\epsilon}. \tag{32}$$

We have

$$\begin{aligned}\sum_{\Delta=(r-1)q}^{rq-1} (4q\ln\Delta + \Delta)^{1-\epsilon} \le &\sum_{\Delta=(r-1)q}^{rq-1} (4q\ln\Delta)^{1-\epsilon} + \sum_{\Delta=(r-1)q}^{rq-1} (\Delta)^{1-\epsilon} \\ \le &4rq^2 \ln rq + rq^2 \lesssim rq^2(\ln r + \ln q).\end{aligned}$$

By plugging in back into (32) we get

$$\begin{aligned}\sum_{k=s+1}^{ns} c_k^{1/q} \lesssim &1 + s^{-1+\epsilon} \sum_{r=1}^{ns/q} 2^{-r+1} rq^2(\ln r + \ln q) \\ \lesssim &1 + s^{-1+\epsilon} q^2 \ln q \lesssim 1 + s^{-\epsilon}\ln s \lesssim 1.\end{aligned}$$

Combining the last estimate with the estimate $\sum\limits_{k=s+1}^{ns} b_k^{1/q} < C$ we get $\sum\limits_{k=s+1}^{ns} a_k^{1/q} < C$.

$\square$

From Lemma 1.5.3 and following the steps after the estimate (28) we get the following theorem.

Theorem 1.5.4. *In all dimensions $n \geq 2$ for every integer $N \geq 1$ there exists a distribution $\mathcal{P} \subset [0,1]^n$ of N points such that for any $0 < \epsilon < 1/2$ there is a constant $C = C(\epsilon, n)$ for which the following holds*

$$\|\mathcal{D}_N\|_p \leq C q^{\frac{n-1}{2}} (\log N)^{\frac{n-1}{2}}\,, \tag{33}$$

for all $1 \leq q < (\ln N)^{1/2-\epsilon}$.

1.6 The Littlewood–Paley Inequalities

We start with the following version of the Littlewood–Paley inequalities (which is just the Hilbert space-valued Khinchin inequality):

Lemma 1.6.1. *For coefficients c_i in a Hilbert space $\mathcal{H}$ and for any $q \geq 2$, there holds*

$$\Big\|\sum_i c_i r_i\Big\|_{L_q(U)} \leq C\sqrt{q}\Big[\sum_i |c_i|^2\Big]^{\frac{1}{2}}\,,$$

where r_i are the Rademacher functions as defined in (15).

There is a hyperbolic extension of this inequality that we will need. For $X = (x_1, \ldots, x_n) \in U^n$ and $I = (i_1, \ldots, i_n) \in \mathbb{N}^n$, set

$$r_I(X) = \prod_{t=1}^{n} r_{i_t}(x_t)\,.$$

Lemma 1.6.2. *For coefficients $c_I \in \mathbb{R}$, for any $k \in \mathbb{N}$, and $K \in \mathbb{N}^n$*

$$\Big\|\sum_{I \in \mathbb{N}^n \,:\, |I|=k} c_I r_{I+K}(X)\Big\|_{L_q(U^n)} \leq [C\sqrt{q}]^{n-1}\Bigg[\sum_{I \,:\, |I|=k} |c_I|^2\Bigg]^{\frac{1}{2}}\,.$$

Proof. The point of the estimate is that we need only apply the Littlewood–Paley $n-1$ times. That we can do so recursively, follows from the Hilbert space structure associated with square functions.

Indeed, apply the Littlewood–Paley inequality in the first coordinate only. We have

$$\int_U \Big|\sum_{I \,:\, |I|=k} c_I r_{I+K}(X)\Big|^q \, dx_1 \leq [C\sqrt{q}]^q \Bigg[\sum_{t \in \mathbb{N}} \Big|\sum_{I' \,:\, |I'|=k-t} c_{(t,I')} r_{I'+K'}(X')\Big|^2\Bigg]^{\frac{q}{2}}\,.$$

On the right, we set $K' = (k_2, \ldots, k_n)$, and similarly for I' and X'. Note that the length of I' is prescribed to be $k-t$, and as well, that the sum on the right is a Hilbert space

(ℓ^2) norm of a Hilbert space-valued Rademacher series in $n-1$ variables. In particular, the Littlewood-Paley inequalities apply to the sum on the right etc. Moreover, since the length of the vectors I is fixed, in $n-1$ applications the process terminates.

Such arguments are common in *product harmonic analysis* and have been used in the context of discrepancy in e.g. [8], see also [7]. $\square$

CHAPTER II

L^1 DICHOTOMY FOR THE DISCREPANCY FUNCTION

It is a well-known conjecture in the theory of irregularities of distribution that the L^1 norm of the discrepancy function of an N-point set satisfies the same asymptotic lower bounds as its L^2 norm. In dimension $d = 2$ this fact has been established by Halász, while in higher dimensions the problem is wide open. In this chapter, we establish a series of dichotomy-type results which state that if the L^1 norm of the discrepancy function is too small (smaller than the conjectural bound), then the discrepancy function has to be large in some other function space.

2.1 Introduction

2.1.1 Preliminaries

For integers $d \geq 2$, and $N \geq 1$, let $\mathcal{P}_N \subset [0,1]^d$ be a finite point set with cardinality $\sharp\mathcal{P}_N = N$. Define the associated discrepancy function by

$$D_N(X) = \sharp(\mathcal{P}_N \cap [0,X]) - N|[0,X]|,$$

where $X = (x_1, \ldots, x_d)$ and $[0,X] = \prod_{j=1}^{d}[0,x_j]$ is a rectangle with antipodal corners at 0 and X, and $|\cdot|$ stands for the d-dimensional Lebesgue measure. The dependence upon the selection of points $\mathcal{P}_N$ will be suppressed, as we are interested in bounds that are only a function of $N = \sharp\mathcal{P}_N$. The discrepancy function D_N measures equidistribution of $\mathcal{P}_N$: a set of points is *well-distributed* if D_N is small in some appropriate function space.

It is a basic fact of the theory of irregularities of distribution that relevant norms of this function in dimensions 2 and higher must tend to infinity as N grows. The classic results are due to Roth [23] in the case of the L^2 norm and Schmidt [27] for L^p, $1 < p < 2$:

Theorem 2.1.1. *For $1 < p < \infty$ and any collection of points $\mathcal{P}_N \subset [0,1]^d$, we have*

$$\|D_N\|_p \gtrsim (\log N)^{(d-1)/2}\,. \tag{34}$$

Moreover, we have the endpoint estimate

$$\|D_N\|_{L(\log L)^{(d-2)/2}} \gtrsim (\log N)^{(d-1)/2}. \tag{35}$$

In dimension $d = 2$ the L^1 endpoint estimate above was established by Halász [20], while its Orlicz space generalization for dimensions $d \geq 3$ is due to the last author Lacey [21] (notice that, when $d = 2$, we have $L(\log L)^{(d-2)/2} = L^1$).

The symbol "$\gtrsim$" in this paper stands for "greater than a constant multiple of", and the implied constant may depend on the dimension, the function space, but *not* on the configuration $\mathcal{P}_N$ or the number of points N. $A \simeq B$ means $A \lesssim B \lesssim A$.

Estimate (34) is sharp, i.e. there exist sets $\mathcal{P}_N$ that meet the L^p bounds (34) in all dimensions. This remarkable fact is established by beautiful and quite non-trivial constructions of point distributions $\mathcal{P}_N$. We refer the reader to one of the very good references [6, 16, 17] for more information about low-discrepancy sets, which is an important complement to the theme of this note.

The subject of our paper is the L^1 endpoint. Halász's original argument yields the following very weak extension to higher dimensions.

Theorem 2.1.2. *In all dimensions $d \geq 3$, we have*

$$\|D_N\|_1 \gtrsim \sqrt{\log N}. \tag{36}$$

No improvements of (36) have been obtained thus far – embarrassingly, it is not even known whether the L^1 norm of D_N should grow as the dimension increases. It is widely believed that the correct bound for the L^1 norm matches Roth's L^2 estimates (34).

Conjecture 2.1.3. *In all dimensions $d \geq 3$, the following estimate holds*

$$\|D_N\|_{L^1([0,1]^d)} \gtrsim (\log N)^{(d-1)/2}. \tag{37}$$

Observe that (35) supports this conjecture.

2.1.2 Main results

While the conjectural bound (37) does not seem accessible at this point, we shall prove several dichotomy-type results for the L^1 norm, which essentially say that either the L^1 norm is large, or some larger norm has to be very large.

We start with a very simple result, valid in all dimensions, which states that if a point distribution has optimally small (according to (34)) L^p norm of the discrepancy, then it has to satisfy the conjectured L^1 estimate (37). In other words, if there exist sets with L^1-discrepancy so small as to violate Conjecture 2.1.3, they cannot simultaneously have low L^p-discrepancy.

Theorem 2.1.4. *Let $p \in (1, \infty)$. For every constant $C_1 > 0$, there exists $C_2 > 0$ such that whenever $\mathcal{P}_N \subset [0,1]^d$ satisfies $\|D_N\|_p \leq C_1 (\log N)^{(d-1)/2}$, it implies that*

$$\|D_N\|_1 \geq C_2 (\log N)^{(d-1)/2}. \tag{38}$$

The next theorem, also true for general dimensions, amplifies this effect. It states that if the L^1-discrepancy fails Conjecture 2.1.3 by a small exponent, then the L^2-discrepancy is not just suboptimal, but huge.

Theorem 2.1.5. *For all dimensions $d \geq 3$, there is an $\epsilon = \epsilon(d) > 0$ and $c = c(d) > 0$ such that for all integers $N \geq 1$, every $\mathcal{P}_N \subset [0,1]^d$ satisfies either*

$$\|D_N\|_1 \geq (\log N)^{(d-1)/2 - \epsilon} \quad \text{or} \quad \|D_N\|_2 \geq \exp(c (\log N)^{\epsilon}) .$$

Thus a putative example of a distribution $\mathcal{P}_N$ with D_N very small in the L^1 norm must be very far from extremal in the L^2-norm. The proof will show that one can take $\epsilon(d)$ as large as a fixed multiple of $1/d$. Specializing to the case of dimension $d = 3$, we can replace the L^2 norm above by a much smaller norm.

Theorem 2.1.6. *In dimension $d = 3$, there holds*

$$\|D_N\|_1 \cdot \|D_N\|_{L(\log L)} \gtrsim (\log N)^2 .$$

Unfortunately, this estimate is consistent with a putative distribution $\mathcal{P}_N$, for which $\|D_N\|_1 \lesssim (\log N)^{1/2}$. The last theorem of this series addresses possible examples, where D_N is less that $(\log N)^{1/2}$ in the L^1 norm.

Theorem 2.1.7. *For all dimensions $d \geq 3$ and all $C_1 > 0$, there is a $C_2 > 0$ so that if $\|D_N\|_1 \leq C_1\sqrt{\log N}$, then $\|D_N\|_2 \gtrsim N^{C_2}$.*

Finally, the dichotomies above are of an essentially optimal nature in light of the examples in this next result.

Theorem 2.1.8. *For all dimensions $d \geq 2$, there is a distribution such that*

$$\|D_N\|_1 \lesssim (\log N)^{(d-1)/2} \quad \text{and} \quad \|D_N\|_2 \gtrsim N^{1/4}.$$

The proofs are based upon the detailed information used to obtain non-trivial improvement in the L^∞ endpoint estimates in [8, 10]. We recall the required estimates in the next section and then turn to the proofs of Theorems 2.1.4–2.1.8 in §2.3.

2.2 The Orthogonal Function Method

All progress on these universal lower bounds has been based upon the orthogonal function method, initiated by Roth [23], with the modifications of Schmidt [27], as presented here. Denote the family of all dyadic intervals $I \subset [0,1]$ by $\mathcal{D}$. Each dyadic interval I is the union of two dyadic intervals I_- and I_+, each of exactly half the length of I, representing the left and right halves of I respectively. Define the Haar function associated to I by $h_I = -\chi_{I_-} + \chi_{I_+}$. Here and throughout we will use the L^∞ (rather than L^2) normalization of the Haar functions.

In dimension d, the d-fold product $\mathcal{D}^d$ is the collection of dyadic intervals in $[0,1]^d$. Given $R = R_1 \times \cdots \times R_d \in \mathcal{D}^d$, the Haar function associated with R is the tensor product

$$h_R(x_1, \ldots, x_d) = \prod_{j=1}^{d} h_{R_j}(x_j)\,.$$

These functions are pairwise orthogonal as $R \in \mathcal{D}^d$ varies.

For a d-dimensional vector $r = (r_1, \ldots, r_d)$ with non-negative integer coordinates let $\mathcal{D}_r$ be the set of those $R \in \mathcal{D}^d$ that for each coordinate $1 \leq j \leq d$, we have $|R_j| = 2^{-r_j}$. These rectangles partition $[0,1]^d$. We call f_r an r-function (a generalized Rademacher function) if for some choice of signs $\{\varepsilon_R \,:\, R \in \mathcal{D}_r\}$, we have

$$f_r(x) = \sum_{R \in \mathcal{D}_r} \varepsilon_R h_R(x)\,.$$

The following is the crucial lemma of the method, see [23, 27, 7]. Given an integer N, we set $n = \lceil 1 + \log_2 N \rceil$, where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

Lemma 2.2.1. *In all dimensions $d \geq 2$ there is a constant $c_d > 0$ such that for each r with $|r| := \sum_{j=1}^{d} r_j = n$, there is an r-function f_r with $\langle D_N, f_r \rangle \geq c_d$. Moreover, for all r-functions there holds $|\langle D_N, f_r \rangle| \lesssim N 2^{-|r|}$.*

Heuristically, this lemma quantifies the fact that most of the information about the discrepancy function is encoded by the Haar coefficients corresponding to boxes $R \in \mathcal{D}^d$ with volume $|R| \approx 1/N$. The proofs of most known lower bounds for the discrepancy function have been guided by this idea. We briefly outline the argument leading to (34).

For integer vectors $\vec{r} \in \mathbb{N}^d$, let $f_{\vec{r}}$ be an $\vec{r}$-function as in the previous lemma. Set

$$Z := \frac{1}{n^{(d-1)/2}} \sum_{\vec{r}\,:\,|\vec{r}|=n} f_{\vec{r}}\,.$$

It is easy to see that, due to orthogonality and the fact that the number of vectors $\vec{r} \in \mathbb{N}^d$ with $|\vec{r}| = n$ is of the order n^{d-1}, we have $\|Z\|_2 \simeq 1$. Moreover, it also satisfies $\|Z\|_p \lesssim 1$ for all $1 < p < \infty$. This extension can be derived using Littlewood–Paley theory or, as originally done in [27], using combinatorial arguments if p is an even integer. This is enough to establish (34): Hölder inequality and Lemma 2.2.1 yield

$$n^{\frac{d-1}{2}} \lesssim \langle D_N, Z \rangle \lesssim \|D_N\|_p \cdot \|Z\|_{p'} \lesssim \|D_N\|_p. \tag{39}$$

The following is a deep exponential-squared distributional estimate for Z – indeed, it is a key estimate behind the main theorems of [10] on the L^∞ norm of the discrepancy function.

Theorem 2.2.2. *[10, Theorem 6.1] There is an absolute constant $0 < c < 1$, such that in all dimensions $d \geq 3$, for $\epsilon = c/d$ we have*

$$|\{x \,:\, |Z(x)| > t\}| \lesssim \exp(-ct^2)\,, \qquad 0 < t < cn^{\frac{1-2\epsilon}{4d-2}}$$

2.3 Proofs

We now proceed to the proofs of the main theorems.

Proof of Theorem 2.1.4. Assume that for a given $1 < p < \infty$ we have $\|D_N\|_p \le C_1 \big(\log N\big)^{\frac{d-1}{2}}$. The Roth–Schmidt bound (34) states that $\|D_N\|_{2p/(p+1)} \ge c_{2p/(p+1)} \big(\log N\big)^{\frac{d-1}{2}}$. Interpolating between 1 and p using Hölder's inequality we find that $\|D_N\|_{2p/(p+1)} \le \|D_N\|_1^{1/2} \, \|D_N\|_p^{1/2}$. Therefore

$$\|D_N\|_1 \ge \frac{\|D_N\|^2_{2p/(p+1)}}{\|D_N\|_p} \ge \frac{c^2_{2p/(p+1)} \big(\log N\big)^{d-1}}{C_1 \big(\log N\big)^{\frac{d-1}{2}}} = C_2 \big(\log N\big)^{\frac{d-1}{2}}, \tag{40}$$

which proves (38) with $C_2 = \frac{c^2_{2p/(p+1)}}{C_1}$. ☐

Proof of Theorem 2.1.5. Set $q = n^{\varepsilon}$, where $\varepsilon \simeq 1/d$, and define

$$Y := \frac{1}{n^{(d-1)/2} q} \sum_{\vec{r} \,:\, |\vec{r}| = n} f_{\vec{r}} \,.$$

Then $\|Y\|_p \lesssim q^{-1}$ for $1 < p < \infty$. Besides, one has $\langle D_N, Y \rangle \ge c \frac{n^{(d-1)/2}}{q}$. But unfortunately Y is not bounded, preventing an immediate conclusion about the L^1 norm of D_N.

On the other hand, from Theorem 2.2.2 we get

$$|\{|Y| > 1\}| \lesssim \exp(-cq^2) \,.$$

Using a trilinear Hölder's inequality, we obtain

$$\begin{aligned} \int_{\{|Y|>1\}} |D_N \cdot Y| \, dx &\le |\{|Y| > 1\}|^{1/4} \|Y\|_4 \|D_N\|_2 \\ &\lesssim \exp(-c' q^2) \cdot q^{-1} \|D_N\|_2 \,. \end{aligned}$$

This last quantity will be at most $\frac{1}{2} \langle D_N, Y \rangle$, if $\|D_N\|_2 \lesssim \exp(c'' q^2)$. Then

$$\begin{aligned} \|D_N\|_1 &\ge \big| \langle D_N, Y \cdot \mathbf{1}_{\{|Y| \le 1\}} \rangle \big| \\ &\ge \langle D_N, Y \rangle - \int_{\{|Y|>1\}} |D_N \cdot Y| \, dx \ge \frac{1}{2} \langle D_N, Y \rangle \gtrsim n^{\frac{d-1}{2} - \varepsilon} \end{aligned}$$

and this proves Theorem 2.1.5 ☐

Proof of Theorem 2.1.6. Define

$$Y = \frac{1}{\sqrt{n}} \sum_{j=1}^{n/2} \sin\Bigl(cn^{-1/2} \sum_{\vec r\,:\, r_1 = j} f_{\vec r} \Bigr)$$

where $0 < c < 1$ is a sufficiently small constant.

Lemma 2.3.1. *The following two estimates hold. First, $\langle D_N, Y\rangle \gtrsim n$, and second,*

$$\mathbb{P}(|Y| > \alpha) \lesssim \exp(-c\alpha^2) \qquad \alpha > 1\,.$$

Proof. Modify, in a straight forward way, [21]*§3 to see that for c sufficiently small,

$$\Bigl\langle D_N, \sin\Bigl(cn^{-1/2} \sum_{\vec r\,:\, r_1 = j} f_{\vec r} \Bigr) \Bigr\rangle \gtrsim \sqrt{n}\,, \qquad 1 \le j \le n/2\,.$$

Sum this over j to prove the first claim of the Lemma.

The second claim, the distributional estimate, is equivalent to the bound $\|Y\|_p \lesssim C\sqrt{p}$ for $2 \le p < \infty$. This is estimate (4.1) in [21]. □

Set $E = \{|Y| > \alpha\}$, where $\alpha > 1$ is to be chosen. We consider the inner product

$$\begin{aligned} cn \le \langle D_N, Y\rangle &\le \langle D_N, Y\mathbf{1}_{E^c}\rangle + \langle D_N, Y\mathbf{1}_E\rangle \\ &\le \alpha\|D_N\|_1 + \|D_N\|_{L(\log L)}\|Y\mathbf{1}_E\|_{\exp(L)} \\ &\le \alpha\|D_N\|_1 + \alpha^{-1}\|D_N\|_{L(\log L)}\,, \end{aligned}$$

where we have used the duality of the spaces $L(\log L)$ and $\exp(L)$. The last estimate depends upon the calculation

$$\|Y\mathbf{1}_E\|_{\exp(L)} \simeq \sup_{t\ge 1} t\cdot \bigl|\log|\{|Y| > \max\{t,\alpha\}\}|\bigr|^{-1} \lesssim \sup_{t\ge1} \min\left\{\frac{1}{t}, \frac{t}{\alpha^2}\right\} \simeq \alpha^{-1}\,.$$

Choose $\alpha^2 \simeq \|D_N\|_{L\log L}/\|D_N\|_1 \ge 1$. We then have

$$n \lesssim \|D_N\|_{L\log L}^{1/2}\|D_N\|_1^{1/2}\,,$$

and this proves Theorem 2.1.6. □

Proof of Theorem 2.1.7. Assume that $\|D_N\|_1 \le C_1 \log N$. We shall utilize the main result of [21], namely (35). Consider the probability measure $\mathbb{P}_N$ which is the normalized $|D_N|\, dx$, i.e. $d\mathbb{P}_N(x) = \frac{|D_N(x)|}{\|D_N\|_1} dx$. We see that

$$\int (\log_+ |D_N|)^{\frac{d-2}{2}}\, d\mathbb{P}_N(x) \ge \frac{\|D_N\|_{L(\log L)^{\frac{d-2}{2}}}}{\|D_N\|_1} \ge Cn^{\frac{d-2}{2}}.$$

It is obvious that $|D_N(x)| \le N$, therefore $\log |D_N| \le n$. It follows from a Paley–Zygmund-type inequality that for some $c > 0$

$$\mathbb{P}_N\{\log_+ |D_N| > cn\} \gtrsim 1. \tag{41}$$

Indeed, denoting $f = \log_+ |D_N|$ and $\alpha = (d-2)/2$, using Cauchy–Schwarz inequality we get

$$\begin{aligned} Cn^\alpha \le \mathbb{E}|f|^\alpha &\le \mathbb{E}|f|^\alpha \mathbf{1}_{\{|f|>cn\}} + \mathbb{E}|f|^\alpha \mathbf{1}_{\{|f|\le cn\}} \\ &\le \left(\mathbb{E}|f|^{2\alpha}\right)^{1/2} \cdot \mathbb{P}_N^{1/2}\{|f| > cn\} + c^\alpha n^\alpha \\ &\le n^\alpha \cdot \left(\mathbb{P}_N^{1/2}\{|f| > cn\} + c^\alpha\right), \end{aligned}$$

which yields (41) if c is small enough. From this, using the fact that $\|D_N\|_1 \gtrsim \sqrt{n}$ (Theorem 2.1.2), we deduce that

$$\|D_N\|_2^2 \gtrsim \int_{\{\log D_N > cn\}} D_N^2(x)\, dx \gtrsim \sqrt{n} \cdot \int_{\{\log D_N > cn\}} |D_N(x)|\, d\mathbb{P}_N(x) \gtrsim N^{C'},$$

which is the conclusion of Theorem 2.1.7. □

For the last proof we need an additional definition.

Definition 2.3.1. *A distribution $\mathcal{P}_N$ of $N = p^s$ points is called a p-adic net, if any p-adic rectangle*

$$\Delta = \prod_{j=1}^{d} [m_j p^{-a_j}, (m_j+1)p^{-a_j}), \quad 0 \le m_j < a_j$$

of volume $\frac{1}{N}$ contains exactly one point of $\mathcal{P}_N$.

For any dimension $d \ge 2$ and a prime $p \ge d-1$, there exist nets with p^s points for all values of $s \ge 2$. One can show that if $\mathcal{P}_N$ is a p-adic net of $N = p^s$ points, then for any rectangle $R \subset [0,1]^d$

$$\left|\sharp(\mathcal{P}_N \cap R) - |R|N\right| \le s^{d-1}.$$

A similar inequality can be obtained for arbitrary N.

Proof of Theorem 2.1.8. Let us take a net $\mathcal{P}_N$ with small L_2 discrepancy, i.e.

$$\|D_N\|_2 \lesssim (\log N)^{(d-1)/2}.$$

The existence of such nets is well-known [13, 16]. Then clearly we also have $\|D_N\|_1 \lesssim (\log N)^{(d-1)/2}$. For $\delta > 0$ we define the cube $Q = [1 - N^{-\delta}, 1]^d$, which lies at the top right corner of $[0,1]^d$. As $|Q| = N^{-\delta d}$ and the distribution $\mathcal{P}_N$ is a net, it follows that Q contains about $N^{1-\delta d}$ points of $\mathcal{P}_N$.

Let $\mathcal{P}'_N$ be a new distribution obtained from $\mathcal{P}_N$ by replacing the points inside Q with $(1, 1, \ldots, 1)$ and keeping the points outside Q unchanged. Let D'_N be the associated discrepancy function. Then $D_N(x) = D'_N(x)$ for $x \notin Q$, and D'_N has no contribution from the distribution of points inside Q. Hence for $x \in Q$

$$|D_N(x) - D'_N(x)| \lesssim N^{1-\delta d}.$$

Because $\mathcal{P}_N$ is a net, in a positive proportion of Q we will also have

$$|D_N(x) - D'_N(x)| \gtrsim N^{1-\delta d}.$$

Therefore we have

$$\|D_N - D'_N\|_1 \simeq N^{1-2\delta d} \quad \text{and} \quad \|D_N - D'_N\|_2^2 \simeq N^{2-3\delta d}.$$

If we take $\delta = \frac{1}{2d}$, we obtain

$$\|D_N - D'_N\|_1 \simeq 1 \quad \text{and} \quad \|D_N - D'_N\|_2 \simeq N^{1/4},$$

which implies that

$$\|D'_N\|_1 \lesssim (\log N)^{(d-1)/2} \quad \text{and} \quad \|D'_N\|_2 \gtrsim N^{1/4}.$$

$\square$

CHAPTER III

ORDER-PRESERVING FREIMAN ISOMORPHISMS

An order-preserving Freiman 2-isomorphism is a bijection $\phi : A \to B$, where A and B are finite subsets of $\mathbb{Z}$, such that $\phi(a) + \phi(b) = \phi(c) + \phi(d)$ if and only if $a + b = c + d$ and $\phi(a) < \phi(b)$ if and only if $a < b$. We show that for any $A \subseteq \mathbb{Z}$, if $|A + A| \leq K|A|$, then there exists a subset $A' \subseteq A$ such that the following holds: (1) $|A'| \geq c_1|A|$, (2) there exists an order preserving Freiman 2-isomorphism $\phi : A' \to B'$, where $B' \subset [1, c_2|A|]$ and c_1, c_2 depend only on K. Informally, this states that a set with small doubling may, in a sense, be viewed as a dense subset of an interval. We also present several applications.

3.1 Introduction

Let G, H be additive groups, and let $A \subseteq G$ and $B \subseteq H$. A Freiman k-homomorphism is a map $\phi : A \to B$ such that

$$\phi(x_1) + \ldots + \phi(x_k) = \phi(y_1) + \ldots + \phi(y_k)$$

if

$$x_1 + \ldots + x_k = y_1 + \ldots + y_k.$$

Such a map ϕ is called a Freiman k-isomorphism if the converse holds as well. If G and H have an ordering, then ϕ is order preserving if

$$\phi(a) < \phi(b) \iff a < b.$$

Frequently, an additive problem is easier in certain ambient groups. In such situations, it is often desirable to find a mapping from one group to another that preserves the additive structure even when no group homomorphism is suitable. Freiman isomorphisms serve such a purpose. We refer the interested reader to [33] for a more detailed exposition.

In certain applications, the additive structure may not be the only property one wants to preserve – mainly, one would also like that the elements stay in the same order under the mapping. The main tool we introduce in this paper, what we call a 'Condensing Lemma,' allows one, under certain restrictions, to find an order preserving Freiman 2-isomorphism from a set of n integers to $[0, cn]$. This lemma allows one to, in a sense, treat sets with small doubling as if they were a dense subset of an interval by finding an order-preserving Freiman 2-Isomorphism from a large subset to a subset of a reasonably sized interval.

Lemma 3.1.1. *[Condensing Lemma] Let $A \subseteq \mathbb{Z}$ be such that $|A + A| = c|A|$. Then, there exists constants c_1, c_2 depending only on c such that the following holds: there exists $A' \subseteq A$ with $|A'| \geq c_1|A|$, and there exists an order-preserving Freiman 2-isomorphism $\phi : A' \to B'$ where $B' \subset [1, c_2|A|]$.*

In order to prove the Condensing Lemma, we need a celebrated theorem of Sanders which guarantees us a large generalized arithmetic progression P in $2A - 2A$ when A has small doubling.

Theorem 3.1.2 (Sanders [25])**.** *Suppose $A \subseteq \mathbb{Z}$ satisfies $|A+A| \leq K|A|$. Then, there exists absolute constants c_1, c_2, c_3 such that $2A - 2A$ contains a proper generalized arithmetic progression P of dimension at most $c_1(\log 2K)^6$ and size at least $\exp(-c_2(\log K)^6 \log 2 \log 2K)|A|$. Moreover, for each $x \in P$, there are at least $\exp(-c_3(\log K)^6 \log 2 \log 2K)|A|^3$ quadruples $(a, b, c, d) \in A^4$ with $x = a + b - c - d$.*

The proof of the Condensing Lemma consists of first applying Sanders' theorem so that we may approximate A by a generalized arithmetic progression $G = \{\sum_{i=1}^{k} x_i d_i : |x_i| \leq L_i\}$. Then, after passing to certain subsets, we use some basic theory from convex geometry to show that there is a generalized arithmetic progression $G' = \{\sum_{i=1}^{k} x_i d_i' : |x_i| \leq cL_i\}$ that shares the additive properties of G, but is contained in an interval of length $O(|G|)$.

After we prove the Condensing Lemma, we provide some applications. Let $A = \{a_1 < a_2 < \ldots < a_n\}$ be a finite subset of the integers. We prove a result on

$$EI(A, A) := \{(i, j, k, l) : a_i + a_j = a_k + a_l \text{ and } i + j = k + l\}.$$

We call this quantity the *indexed energy* of A and we denote it as $EI(A, A)$. This is related to the well-known additive energy of a set which is related to the sumset via an application of Cauchy-Schwarz:

$$E(A,A) = |\{(i,j,k,l) : a_i + a_j = a_k + a_l\}| \geq \frac{|A|^4}{|A+A|}.$$

Although the indexed energy of a set has not been directly studied, the additive properties of a set and how they interact with the related indices has appeared in various forms. Solymosi [32] studied the situation when $a_i + a_j \neq a_k + a_l$ for $i - j = k - l = c$ for a fixed constant c, and in particular when a set A has the property that $a_{i+1} + a_i \neq a_{j+1} + a_j$ for all pairs i, j. Brown et al [12] asked if one finitely colors the integers $\{1, \ldots, n\}$, must one be forced to find a monochromatic 'double' 3-term arithmetic progression $a_i + a_j = 2a_k$ where $i + j = 2k$? In this paper, we determine the relationship between the indexed energy of a set and the additive energy.

Layout and Notation. In section 2, we prove the Condensing Lemma. In section 3, we study the indexed energy of a set, providing both an extremal construcion of a set with large additive energy and small indexed energy as well as proving a Balog-Szemerédi-Gowers type theorem to find a subset with large indexed energy. Section 4 contains further applications and conjectures related to both the Condensing Lemma and the indexed energy. All sets are assumed to be subsets of $\mathbb{Z}$. In particular, we write $[a,b]$ for $[a,b] \cap \mathbb{Z}$, and similarly for $[a,b), (a,b)$, and $(a,b]$. For two functions f, g, we write $f \gtrsim g$ if $f = \Omega(g)$, and if $f = O(g)$, then we write $f \lesssim g$. We write $[n]$ for the set of integers $\{0, \ldots, n-1\}$. The doubling of a set A is $\frac{|A+A|}{|A|}$. A set has *small doubling* if its doubling is $O(1)$. A *generalized arithmetic progression* G is a set $\{a + x_1 d_1 + \ldots x_k d_k : -L_i \leq x_i \leq L_i\}$; we call k the dimension of G; $|G|$ is the volume of G. Moreover, G is proper if the volume of G is maximal – $(2L_i + 1)^k$.

3.2 Condensing Lemma

The following lemma in conjunction with Theorem 3.1.2 will allow us to prove Lemma. 3.1.1

Lemma 3.2.1. *Let G be a proper generalized arithmetic progression of the form $G := \{\sum_{i=1}^k a_i d_i : -4L_i \leq a_i \leq 4L_i\}$. Let $G' := \{\sum_{i=1}^k a_i d_i : -L_i \leq a_i \leq L_i\}$. There exists a constant $c = c(k)$ such that there exists a map $\phi : G' \to [-c|G'|, c|G'|]$ with the following properties:*

- *$\phi(\sum_{i=1}^{k} a_i d_i) = \sum_{i=1}^{k} a_i d_i'$ for some $d_1', \ldots, d_k'$.*

- *ϕ is an order preserving Freiman 2-isomorphism.*

In order to prove this lemma we need some definitions and results from convex geometry, from which we use [5] as a reference.

A set $K \subset \mathbb{R}^n$ is said to be a convex cone if for all $\alpha \geq 0$, $\beta \geq 0$ and $x, y \in K$ we have $\alpha x + \beta y \in K$. For points $x_1, ..., x_m \in \mathbb{R}^n$ and non-negative real numbers $\alpha_1, ..., \alpha_m$, the point

$$x = \sum_{i=1}^{m} \alpha_i x_i$$

is called a conic combination of the points $x_1, ..., x_m$. The set $co(D)$ is defined as all conic combinations of points in $D \subset \mathbb{R}^n$ and is called the convex hull of the set D. For a non-zero $x \in \mathbb{R}^n$ the convex hull of x is called a ray spanned by x. A ray R of the cone K is called an extreme ray if whenever $\alpha x + \beta y \in R$ for $\alpha > 0$, $\beta > 0$ and $x, y \in K$ then $x, y \in R$. An extreme ray is a 1-dimensional face of the cone. A set $B \subset K$ is called a base of K if $0 \notin B$ and for every point $x \in K$, $x \neq 0$, there is a unique representation $x = \lambda y$ with $y \in B$ and $\lambda > 0$.

Theorem 3.2.2 ([5]). *If K is a convex cone with a compact base. Then every point $x \in K$ can be written as a conic combination*

$$x = \sum_{i=1}^{m} \lambda_i x_i, \quad \lambda_i \geq 0, \quad i = 1, ..., m,$$

where points x_i span extreme rays of K.

The broad idea of the proof of Lemma 3.2.1 is as follows. We are given a generalized arithmetic progression $G := \{\sum_{i=1}^{k} a_i d_i : -4L_i \leq y_i \leq 4L_i\}$. In a sense, this can be indentified with the point $(d_1, \ldots, d_k)$. What we would like to find is another generalized arithmetic progression, $H := \{\sum_{i=1}^{k} b_i d_i' : -L_i' \leq b_i \leq L_i'\}$ which maintains the same additive structure as G, but is much more compact. Viewed another way, we want to find a point $(d_1', \ldots, d_k')$ much closer to the origin than $(d_1, \ldots, d_k)$ that also satisfies certain inequalities (these are what maintain the additive structure and order). Hence, we reduce our problem to finding an integer solution, relatively close to the origin, to a set of linear inequalities.

Proof. Given $G := \{\sum_{i=1}^k a_i d_i : -4L_i \leq a_i \leq 4L_i\}$, consider the following set of inequalities:

$$\{\sum_{i=1}^{k} a_i x_i > 0 : a_1 d_1 + \ldots + a_k d_k > 0; -4L_i \leq a_i \leq 4L_i\}. \tag{42}$$

If $d'_1, \ldots, d'_k$ is an integer solution to the above system of inequalities, then the map ϕ induced by $d'_1, \ldots, d'_k$ is an order preserving Freiman 2-isomorphism. We prove this below.

If

$$\sum_{i=1}^{k} a_i d_i < \sum_{i=1}^{k} b_i d_i$$

for two points in G', then

$$\sum_{i=1}^{k} (b_i - a_i) d_i > 0$$

is one of the inequalities that $d'_1, \ldots, d'_k$ must satisfy; so

$$\sum_{i=1}^{k} a_i d'_i < \sum_{i=1}^{k} b_i d'_i.$$

The converse is also clear.

If we have points in G' such that

$$\sum_{i=1}^{k} a_i d_i + \sum_{i=1}^{k} b_i d_i = \sum_{i=1}^{k} s_i d_i + \sum_{i=1}^{k} t_i d_i$$

then

$$\sum_{i=1}^{k} (a_i + b_i) d_i = \sum_{i=1}^{k} (s_i + t_i) d_i. \tag{43}$$

These correspond to points in G, and by the fact that G is proper, we must have that $a_i + b_i = s_i + t_i$ for $i = 1, \ldots, k$. Hence,

$$\sum_{i=1}^{k} a_i d'_i + \sum_{i=1}^{k} b_i d'_i = \sum_{i=1}^{k} s_i d'_i + \sum_{i=1}^{k} t_i d'_i. \tag{44}$$

For the converse, if (44) holds and (43) does not, then without loss of generality

$$\sum_{i=1}^{k} (a_i + b_i - s_i - t_i) d_i > 0.$$

However, $a_i + b_i - s_i - t_i \in [-4L_i, 4L_i]$, and so the above inequality is satisfied by $d_1', \ldots, d_k'$ which is a contradiction.

Now, we bound the image of ϕ. Consider the system of inequalities defined in (42). First, note that the solution space to this system of linear inequalities is the interior of a closed convex cone defined by $k-1$-dimensional hyperplanes (that all pass through the origin). Moreover, this interior is nonempty since there is a solution – $d_1, \ldots, d_k$. For visual reference, note that $x_i > 0$ is one of our inequalities for all $i = 1, \ldots, k$ so we are in the positive quadrant of $\mathbb{R}^k$. Let K be the closure of the cone defined by inequalities (42). Since K is in the positive quadrant of $\mathbb{R}^k$, it has a compact base, e.g. we can take as a base B the intersection of cone K with the hyperplane $x_1 + \cdots + x_k = 1$. Hence cone K can be represented as conic combinations of the points on its extreme rays. Because all extreme rays have dimension 1, they are intersections of $k-1$ linearly independent hyperplanes corresponding to the system (42). For each extreme ray, we show how to find an integer point on it; then, taking a conic combination of these integer points will allow us to find an integer point in the interior of the cone.

Let the following hyperplanes define one of our extreme rays:

$$\{a_{i,1}x_1 + \ldots + a_{i,k}x_k = 0 : i = 1, \ldots, k-1\}.$$

This system of equations will have all the points along our extreme ray as a solution. Hence, we may treat one of the variables x_i as a free variable while the other variables depend on it. Without loss of generality, assume that x_k is the free variable, and let us solve the system for the case when $x_k = 1$. Let

$$\Delta := \begin{vmatrix} a_{1,1} & \cdots & a_{1,k-1} \\ a_{2,1} & \cdots & a_{2,k-1} \\ \vdots & & \\ a_{k-1,1} & \cdots & a_{k-1,k-1} \end{vmatrix}$$

and

$$\Delta_i := \begin{vmatrix} a_{1,1} & \cdots & a_{1,i-1} & -a_{1,k} & a_{1,i+1} & \cdots & a_{1,k-1} \\ \vdots & \ddots & & & & & \\ a_{k-1,1} & \cdots & a_{k-1,i-1} & -a_{k-1,k} & a_{k-1,i+1} & \cdots & a_{k-1,k-1} \end{vmatrix}.$$

By Cramer's rule, the solution to the system is given by $x_i = \frac{\Delta_i}{\Delta}$ for $i = 1, \ldots, k-1$. Any multiple of this is also a solution. Hence, $(|\Delta_1|, \ldots, |\Delta_{k-1}|, |\Delta|)$ is an integer solution to our system that lies along our edge. For convenience, let $\Delta_k := \Delta$.

Now, we may get such an integer solution for each of our extreme rays. Because cone K has interior points, then not all extreme rays belong to the same face, in particular, we may take a set of $k+1$ of such rays that do not all lie along the same face and get $k+1$ integer solutions as we did above. Call these solutions $P_1, \ldots, P_{k+1}$. We can bound the entries of P_r quite easily. To see that, note that for $i = 1, \ldots, k$, we have that since each entry $|a_{i,j}| \leq 4L_j$, the determinant is bounded as follows:

$$|\Delta_i| \leq 4^k k! \prod_{j \neq i} L_j$$

Moreover, their sum, $P := P_1 + \ldots + P_{k+1} = (d'_1, \ldots, d'_k)$ does not belong to any of the faces of K; so, it belongs to the interior of the cone. Lastly, this implies our bounds on the image of ϕ since any element $y_1 d'_1 + \ldots + y_k d_k$ must satisfy the following:

$$|y_1 d'_1 + \ldots + y_k d'_k| \leq (k+1) 4^k k! \prod_{j=1}^{k} L_j.$$

So if $g \in G'$, $\phi(g) \in [-4^k (k+1)! |G|, 4^k (k+1)! |G|]$. $\square$

The proof of Lemma 3.1.1 follows easily from applying the theorem of Sanders to a set with small doubling.

Proof. Let $A \subseteq \mathbb{Z}$ be such that $|A + A| \leq c|A|$. All constants c_i in the following depend only on c. We may apply Theorem 3.1.2 to A to get a generalized arithmetic progression $G \subseteq 2A - 2A$ with $|G| \geq c_1 |A|$, dimension at most c_2, and for each $x \in G$, there are at least $c_3 |A|^3$ quadruples $(a, b, c, d) \in A^4$ with $x = a + b - (c + d)$. Hence,

$$\sum_{a,b,c,d \in A} \mathbb{1}_G(a + b - c - d) \geq c_3 |A|^3 |G|.$$

So, we can find a triple (b, c, d) such that

$$\sum_{a \in A} \mathbb{1}_G(a + b - c - d) \geq c_3 |G|.$$

Let $A' := \{a \in A : a + b - c - d \in G\}$. Let $G' = G + c + d - b$. So, $A' \subseteq G'$, $|A'| \geq c_3|G'|$, and G' is a proper generalized arithmetic progression of the same size and dimension as G.

Denote G' as

$$G' = \{u + \sum_{i=1}^{k} x_i d_i : -L_i \leq x_i \leq L_i\}.$$

First, we may assume $u = 0$, else simply shift everything in A' and G' by $-u$, and work with those sets instead. Let $G'' := \{\sum_{i=1}^{k} x_i d_i : \lceil L_i/4 \rceil \leq x_i \leq \lfloor L_i/4 \rfloor\}$. Apply Lemma 3.2.1 to G' to get a map $\phi : G'' \to [-c_4|G''|, c_4|G''|]$. We may have that A' does not intersect G'', but it will certainly intersect one of the 4^k different translates of it. In otherwords, since $A' \subseteq G'$, there exists an integer v such that $|A' \cap (G'' + v)| \geq \frac{|A'|}{4^k}$. Hence, we can assume that $|A' \cap G''| \geq \frac{|A'|}{4^k}$ since, otherwise, we could simply replace G' with $G' + v$. Letting $A'' := A' \cap G''$ defines our subset such that $\phi(A'')$ is an order-preserving Freiman 2-isomorphism into $[0, c_6|A|]$, proving the lemma. $\square$

3.3 Indexed Energy

One always has the following relationship between the additive energy and indexed energy:

$$|A|^2 \leq EI(A, A) \leq E(A, A) \leq |A|^3.$$

If A is an arithmetic progression the relationship is strengthened to $EI(A, A) = E(A, A)$. Moreover, for an arithmetic progression A, $E(A, A)$ is maximized. Thus, it is natural to wonder if one loosens the restriction to $E(A, A) \gtrsim |A|^3$ then is $EI(A, A) \gtrsim |A|^3$? We provide a counterexample to show that this is not true.

Theorem 3.3.1. *There exists an integer N such that for every $n \geq N$, there exists $A \subset [n]$ such that, $E(A, A) \geq \frac{1}{18}|A|^3$ and $EI(A, A) \leq 2000|A|^2(\log |A|)^2$.*

Thus, one can indeed have the additive energy nearly maximal, $\Omega(|A|^3)$, while the indexed energy is nearly minimal, $O((|A| \log |A|)^2)$. However, our main theorem states that when the additive energy is large, one can still pass to a large subset $A' \subseteq A$, $|A'| = \Omega(|A|)$, which has indexed energy $\Omega(|A'|^3)$. We note that when passing to a subset, the subset *does not* inherit the same indices as the superset, but rather it is reindexed in the natural way.

Theorem 3.3.2. *Let A be a finite set of integers with $|A+A| \leq c|A|$. Then, there exists c_1, c_2 depending only on c such that the following holds. There exists an $A' \subseteq A$ such that $EI(A', A') \geq c_1|A'|^3$ and $|A'| \geq c_2|A|$.*

The condition that $|A+A| \lesssim |A|$ may be loosened to $E(A, A) \lesssim |A|^3$ by applying the following well-known result of Balog-Szemerédi [4] and Gowers [19] to pass to a subset with small doubling.

Theorem 3.3.3 (Balog-Szemerédi[4], Gowers[19])**.** *Suppose $A \subseteq \mathbb{Z}$ is such that $E(A, A) \geq c|A|^3$. Then, there exists c_1, c_2 dependent on c such that there exists $A' \subseteq A$ with $|A'| \geq c_1|A|$ and $|A' + A'| \leq c_2|A'|$.*

3.3.1 Finding a subset with large indexed energy

It turns out that if A is a dense subset of an interval, then there is a simple algorithm that can find a subset $A' \subseteq A$ with $|A'| \gtrsim |A|$ and $EI(A', A') \gtrsim |A'|^3$. Thus, the general case may then be quickly deduced by applying the Condensing Lemma. We first begin with a lemma that states, loosely speaking, that if A is a dense subset of $[n]$, then one can choose a large subset $A' \subseteq A$ that is equidistributed over the interval.

Lemma 3.3.4. *If $A \subseteq [n]$ with n sufficiently large and $|A| = \delta n$, then there exists c_1, c_2, c_3 dependent only on δ such that the following holds. There exists an $A' \subseteq A$, $|A'| \geq c_1|A|$ and for $c_3|A|^2$ pairs of integers $0 \leq i, j < n/c_2$, we have that*

$$|A' \cap [ic_2, jc_2)| = j - i. \tag{45}$$

It is easy to establish that a set with property (45) has large indexed energy.

Lemma 3.3.5. *If $A \subseteq [n]$ with n sufficiently large and $|A| = \delta n$, then there exists a c_0, c_1 dependent on δ such that A has a subset $A' \subseteq A$ with $|A'| \geq c_1|A|$ and $EI(A', A') \geq c_0|A|^3$.*

Proof of Lemma 3.3.4. It suffices to prove that there exists an $A' \subseteq A$ and constants c_1, c_2, c_3 such that the following holds: $|A'| \geq c_1|A|$, for $c_2|A|$ integers $0 \leq i < n/c_2$,

$$|A' \cap [0, ic_2)| = i.$$

Once this statement is established, then for any pair of integers i, j satisfying the above, we have $|A' \cap [ic_2, jc_2)| = j - i$. This would prove the statement of the lemma.

Denote $A = \{a_1 < a_2 < \ldots < a_{\delta n}\}$. Let $d = \lfloor \frac{2}{\delta} \rfloor$. We may assume $d|n$, if not, delete the largest elements from A and work with this slightly modified set. Let $I_j = [(j-1)d, jd)$ for all $j = 1, \ldots, \frac{n}{d}$. Let $A_j = A \cap I_j$. We pick our subset A' as follows:

- *Step 1:* If $A_1 \neq \emptyset$ then let $X_1 = \{a_1\}$. Else, $X_1 = \emptyset$.
- *Step k:* If $|A_k \cup X_{k-1}| \leq k$, then $X_k = A_k \cup X_{k-1}$. Else, arbitrarily choose $B \subseteq A_k$ so that $|B \cup X_{k-1}| = k$.

It is clear this algorithm ends after $\frac{n}{d}$ steps. Let $A' = X_{\frac{n}{d}}$.

To prove that A' satisfies the conclusion of the lemma, we analyze the algorithm as follows. First, note that $X_1 \subseteq X_2 \subseteq \ldots \subseteq X_{\frac{n}{d}} = A'$ and $|X_i| \leq i$ for all i. Now, the sets X_i for which $|X_i| = i$ we will call good, and the others we will call bad. Note that if X_i is good, then $|A' \cap [0, id)| = i$; hence, showing that lots of X_i are good will prove the lemma. Let $J = \{j_1, j_2, \ldots, j_k\}$ be the set of indices such that $|X_{j_i}| = j_i$. Observe that for indices between j_i and j_{i+1}, we must not have enough elements to make any of those corresponding sets good. More precisely,

$$|A_{j_i+k}| \leq k - 1 - \sum_{s=1}^{k-1} |A_{j_i+s}|.$$

This implies that

$$\left| \bigcup_{k=1}^{j_{i+1}-j_i-1} A_{j_i+k} \right| \leq j_{i+1} - j_i - 2.$$

So, we must have that

$$\left| \bigcup_{i=1}^{k-1} \bigcup_{s=1}^{j_{i+1}-j_i-1} A_{j_i+s} \right| \leq \sum_{i=1}^{k-1} j_{i+1} - j_i - 2 = j_k - j_1 - 2(k-1) \leq j_k \leq \frac{n}{d}.$$

Thus, we have that $\delta n - \frac{n}{d} \geq \delta n/2$ elements of A are distributed over good intervals. Since each interval is of length d, then we must have that k, the number of good intervals, is at least

$$\frac{\delta n}{2d} \geq \frac{n\delta^2}{4}.$$

This in turn gives us a lower bound on $|A'| = j_k \geq k \geq \frac{n\delta^2}{4} = \frac{\delta}{4}|A|$. $\square$

Proof of Lemma 3.3.5. Apply Lemma 3.3.4 to A to get A', c_1, c_2, c_3 as in the lemma. Let $A' = \{b_1 < b_2 < \ldots < b_m\}$. Let $J = \{j : |A' \cap [0, c_2 j)| = j\}$. We know that $|J| \geq c_2|A|$. Now, let $A'' = \{b_j : j \in J\}$. Since $EI(A', A') \geq |\{(i,j,k,l) \in J^4 : b_i + b_j = b_k + b_l \text{ and } i+j = k+l\}|$, we will simply work with these quadruples from A''. However, our final set will still be A' since we need to keep the indices of elements the same as they were in A'.

For all of the following, b_j will be assumed to be from A''. Let $t \in \{2, \ldots, 2m\}$. For $t \leq m$, there are $t-1$ pairs $(i,j) \in [m] \times [m]$ such that $i + j = t$. For $t > m$, there are $2m - (t-1)$ pairs $(i,j) \in [m] \times [m]$ such that $i + j = t$. Let α_t be defined so that for $t \in \{2, \ldots, 2m\}$ there are $\alpha_t(t-1)$ pairs $(i,j) \in J \times J$ with $i + j = t$ and there are $\alpha_t(2m-(t-1))$ such pairs for $t \in \{m+1, \ldots, 2m\}$. Observe that for such pairs $(i,j) \in J \times J$, we have $b_i + b_j \in [(t-2)d, td)$. Thus, there are only $2d$ values that $b_i + b_j$ can take. For every $i \in [0, 2d-1]$, let t_i denote the number of pairs $(i,j) \in J \times J$ with $b_i + b_j = (t-2)d + i$. We can bound the indexed energy of A' as follows:

$$EI(A', A') \geq \sum_t \sum_{i=0}^{2d-1} t_i^2 = \sum_{t=2}^{m} \sum_{i=0}^{2d-1} t_i^2 + \sum_{t=m+1}^{2m} \sum_i t_i^2.$$

Using Cauchy-Schwarz, one has

$$\geq \frac{1}{2d} \left(\sum_{t=2}^{m} (\alpha_t(t-1))^2 + \sum_{t=m+1}^{2m} (\alpha_t(2m-t+1))^2 \right).$$

Using Cauchy-Shwarz again,

$$\geq \frac{1}{2d} \frac{1}{m} \left(\left(\sum_{t=2}^{m} \alpha_t(t-1) \right)^2 + \left(\sum_{t=m+1}^{2m} \alpha_t(2m-t+1) \right)^2 \right).$$

Since

$$\sum_{t=2}^{m} \alpha_t(t-1) + \sum_{t=m+1}^{2m} \alpha_t(2m-t+1) = |J|^2$$

one of the sums must be at least $|J|^2/2$. Hence, we have that

$$EI(A', A') \geq \frac{|J|^4}{2md} = c_0|A|^3$$

for some constant c_0 depending only on δ. $\square$

Now, we are ready to prove Theorem 3.3.2.

Proof of Theorem 3.3.2. Let A be a finite subset of integers with $|A+A| \leq c|A|$. All constants c_i in the following depend only on c. Apply Lemma 3.1.1 to A to get a set $A' \subseteq A$ with $|A'| \geq c_1|A|$ and an order preserving Freiman $\phi : A' \to [0, c_2|A'|]$. Apply Lemma 3.3.5 to $\phi(A')$ to conclude that $EI(\phi(A'), \phi(A')) \geq c_3|\phi(A')|^3 = c_3|A'|^3$. It is easy to see that $EI(\phi(A'), \phi(A')) = EI(A', A')$ since ϕ is an order preserving Freiman 2-isomorphism, so the result follows. $\square$

3.3.2 An Extremal Construction

The proof of Theorem 3.3.1 follows from the following lemma.

Lemma 3.3.6. *Let $n \in \mathbb{N}$, and let $p \in (1,2)$ and denote $p = 1+\epsilon$. Let $A = \{\lfloor a^p \rfloor : 1 \leq a \leq \lfloor n^{1/p} \rfloor\}$. Then, $EI(A,A) \leq 16\epsilon^{-1}n^2 \log n$.*

Proof of Lemma 3.3.6. Let $x, y \in [1, \lfloor n^{1/p} \rfloor]$ with $x+1 < y$. The main part of the argument is to establish the following bound:

$$x^p + y^p - (x+1)^p - (y-1)^p > \frac{\epsilon(y-x)}{2y} \tag{46}$$

For now, assume (46) holds. If $x + y = z + w$, then by convexity, $x^p + y^p \neq z^p + w^p$ unless $z = x$ and $y = w$ or vice versa. However, it may happen that $x + y = z + w$ and $\lfloor x^p \rfloor + \lfloor y^p \rfloor = \lfloor z^p \rfloor + \lfloor w^p \rfloor$. Since $\lfloor a^p \rfloor = a^p - [a^p]$, where $[a^p]$ is the noninteger part of a^p, we must have that if $x + y = z + w$ and

$$\lfloor x^p \rfloor + \lfloor y^p \rfloor = \lfloor z^p \rfloor + \lfloor w^p \rfloor$$

then

$$|x^p + y^p - z^p - w^p| < 2.$$

So, fixing an x and a y, we can bound how many other pairs z and w can have $z+w = x+y$ and $\lfloor z^p \rfloor + \lfloor w^p \rfloor = \lfloor x^p \rfloor + \lfloor y^p \rfloor$. More specifically, we find the largest t such that

$$x^p + y^p - (x+t)^p - (y-t)^p < 2.$$

Using (46), the triangle inequality, and letting $k = y - x$ we get that

$$x^p + y^p - (x+t)^p - (y-t)^p \geq \frac{\epsilon k}{2y} + \frac{\epsilon(k+2)}{2(y-1)} + \ldots + \frac{\epsilon(k+2(t-1))}{2(y-(t-1))}.$$

Each term in the sum is greater than or equal to $\frac{\epsilon k}{2y}$, so we get a lower bound of $\frac{t\epsilon k}{2y}$. So, if $t \geq \frac{4y}{\epsilon(y-x)}$, then we cannot have

$$\lfloor x^p \rfloor + \lfloor y^p \rfloor = \lfloor (x+t)^p \rfloor + \lfloor (y-t)^p \rfloor.$$

This allows us to conclude that any quadruple (x, y, z, w) with $x + y = z + w$, with $x < z < w < y$, $z < x < y < w$, $w < y < x < z$, or $y < w < z < x$ we must have that $|z - x| < \frac{4y}{\epsilon(y-x)}$. Accounting for an extra factor of 2 for when $x < w < z < y$ and so on, we can bound the indexed energy of A

$$EI(A, A) \leq 2 \sum_{y} \sum_{x<y} \frac{4y}{\epsilon(y-x)}.$$

Estimating this summation by using the harmonic series gets us that

$$EI(A, A) \leq \frac{16}{\epsilon} n^2 \log n$$

concluding the proof.

Now, we work to establish (46). First, since $f(x) = x^p$ is convex for $p > 1$, it is easy to establish the following bound for any $k > 1$:

$$p(x+k)^{p-1} > (x+1)^p - x^p > px^{p-1}.$$

Assuming $p = 1 + \epsilon < 2$, we have that x^{p-1} is concave. Doing a similar analysis for $g(x) = x^{p-1}$, we get that

$$(p-1)x^{p-2} > (x+k)^{p-1} - x^{p-1} > (p-1)(x+k)^{p-2}.$$

Let $k = y - x$, and we have

$$x^p + y^p - (x+1)^p - (y-1)^p =$$

$$= y^p - (y-1)^p - ((x+1)^p - x^p) > p(y-1)^{p-1} - p(x+1)^{p-1}.$$

Since $x = y - k$, we have

$$p[(y-1)^{p-1} - (y-k+1)^{p-1} > p[(k-2)(p-1)(y-1)^{p-2}] > \frac{\epsilon k}{2y}$$

where we remind the reader $p = 1 + \epsilon$, $\epsilon \in (0, 1)$. □

Theorem 3.3.1 follows by letting $\epsilon = \frac{1}{\log n}$.

Proof of Theorem 3.3.1. Let A be as in the above lemma, let $\epsilon = \frac{1}{\log n}$. Then, for n sufficiently large

$$|A| = \lfloor n^{\frac{1}{1+\epsilon}} \rfloor = \left\lfloor n^{\frac{1}{1+\frac{1}{\log n}}} \right\rfloor = \left\lfloor \frac{n}{e} \cdot n^{\frac{1}{1+\log n}} \right\rfloor \geq \left\lfloor \frac{n}{e^2} \right\rfloor \geq \frac{n}{9}.$$

So, $A \subseteq [n]$, $|A| = \frac{n}{9}$, and $A + A \subseteq [2n]$. Thus, $|A + A| \leq 2n \leq 18|A|$. Hence,

$$E(A, A) \geq \frac{|A|^4}{|A + A|} \geq \frac{|A|^3}{18}.$$

By the lemma above, for A sufficiently large,

$$EI(A, A) \leq 16n^2(\log n)^2 \leq 16 \cdot (9|A|)^2(\log 9|A|)^2 \leq 1296|A|^2(\log 9|A|)^2 \leq 2000|A|^2(\log |A|)^2.$$

□

3.4 Further Applications and Conjectures

Since $|(A \times B) + (A \times B)| = |A + A||B + B|$, it is obvious that if $|A + A| \leq K|A|$ and $|B + B| \leq K|A|$, then for any $C \subseteq A \times B$ of size $\delta|A||B|$, one has $|C + C| \lesssim |C|$. However, if $|C| = O(\sqrt{|A||B|})$, one has no control of $|C + C|$. Does there exist a $C \subseteq A \times B$ with $|C| = c\sqrt{|A||B|}$, and $|C + C| \lesssim |C|$? Clearly one could simply take $C = \{(a, b) : a \in A\}$. If we additionally require that for any distinct $(x, y), (z, w) \in C$ we have $(x - z)(y - w) > 0$, the answer is not as obvious.

For a set $C \subseteq A_1 \times \ldots \times A_k$, call C a diagonal set if for any distinct pairs of elements $(x_1, \ldots, x_k), (y_1, \ldots, y_k) \in C$, one has $x_i - y_i > 0$ for all i or $x_i - y_i < 0$ for all i.

Theorem 3.4.1. *Let $A_1, \ldots, A_k \subseteq \mathbb{Z}$ be sets of size n such that $|A_i + A_i| \leq K|A_i|$ for all $i = 1, \ldots, k$. Then, there exists a diagonal set $C \subset A_1 \times \ldots \times A_k$ and constants c_1, c_2 such that $|C + C| \leq c_1|C|$ and $|C| \geq c_2(|A_1| \ldots |A_k|)^{1/k}$.*

Proof. We may apply the Condensing Lemma to each A_i to find constants c_1, c_2 depending on K and subsets $A_i' \subseteq A_i$ such that A_i' is Freiman 2-isomorphic to a set $B_i \subseteq [0, c_1 n]$, and $|A_i'| \geq c_2 n$. Next, we claim that there exists $t_1, \ldots, t_k \in \mathbb{Z}$ such that

$$| \cap_{i=1}^k B_i + t_i| \geq \frac{c_2^k}{2^{k-1}} n.$$

We prove this by induction on k. For $k = 1$, it is trivial. For the induction step, let $X, Y \subset [n]$ be of size $\delta_1 n$ and $\delta_2 n$ respectively. Then,

$$\sum_{t=-(n-1)}^{n-1} |X + t \cap Y| = |X||Y| = \delta_1 \delta_2 n^2.$$

Hence, there exists a t such that

$$|(X + t) \cap Y| \geq \frac{\delta_1 \delta_2}{2} n.$$

Now, let $C' = \cap_{i=1}^{k} B_i + t_i$ for such a set of t_i, $i = 1, \ldots, k$. Denote $C' := \{x_1 < \ldots < x_m\}$. We let C be the following set:

$$C := \{(x_i - t_1, x_i - t_2, \ldots, x_i - t_m) : i = 1, \ldots, m\}.$$

Since $x_i - t_j \in B_j$, we have that $C \subseteq B_1 \times \ldots \times B_k$. Since $x_i - t_j > x_\ell - t_j$ for $i > \ell$, C must be diagonal. Also, $|C| = |C'| \in [\frac{c_2^k}{2^{k-1}} n, n]$. Lastly, it is easy to see that

$$|C + C| = |C' + C'| \leq 2n = \frac{2^k}{c_2^k} |C'|.$$

□

Although the above application is similar in spirit to the indexed energy problem – letting $A \times B := A \times [1, |A|]$ – there are several subtle differences. Mainly, in the indexed energy problem, when we pass to a subset, we are forced to reindex the set in a very specific way. The following conjecture however would be general enough to imply Theorem 3.3.2.

Conjecture 3.4.2. *Let $A, B \subseteq \mathbb{Z}$ be sets of size N such that $|A+A|, |B+B| \leq KN$. Then, there exists an $A' \subseteq A$ and constants c_1, c_2 depending only on K such that the following holds.*

Denote $A' := \{a'_1 < \ldots < a'_k\}$ and $B := \{b_1 < \ldots < b_n\}$. If $C := \{(a'_i, b_i) : i = 1, \ldots, k\}$, then $|C + C| \leq c_2 |C|$, and $|A'| \geq c_1 |A|$.

Conjecture 3.4.2 is true in the case where $B = [1, |A|]$ (or any arithmetic progression of size $|A|$) since this then becomes the indexed energy result. It would be interesting to know whether the conjecture is even true in the case where B is a generalized arithmetic progression of dimension 2.

Another problem closely related to the indexed energy is as follows. Let $A \subseteq \mathbb{Z}$ and let $f : A \to \mathbb{Z}$ be such that $|f(A) + f(A)| \leq c|A|$, and $|A + A| \leq c|A|$. Let

$$E_f(A) := \{(a, b, c, d) : a + b = c + d, f(a) + f(b) = f(c) + f(d)\}.$$

Does there exist an $A' \subseteq A$ with $|A'| \gtrsim |A|$, and $E_f(A') \gtrsim |A|^3$? The answer is no since $E_f(A') \leq E_f(A)$ when $A' \subseteq A$. Thus, if we let f be the indexing function for A, and let A be as in Lemma 3.3.6, then $E_f(A') = EI(A', A') \leq EI(A, A) \lesssim |A|^2 \log |A|$. Moreover, $|\{(a+a', f(a)+f(a')) : a, a' \in A\}| \gtrsim |A|^2/\log |A|$. Are there any reasonable conditions that we can impose on f or A to arrive at a different conclusion?

Lastly, we remark that the content of Lemma 3.3.4 is making a statement about equidistribution of a set in an interval. This has been a well-studied topic in discrepancy theory; however, we are not aware of it appearing in this specific, combinatorial form – where one is allowed to pass to a subset of the original set, and one only requires that for lots of interval, the subset is well-distributed. We conjecture a generalization of Lemma 3.3.4 to higher dimensions.

Conjecture 3.4.3. *Let $A \subseteq [n] \times [n]$ be of size $|A| = \delta n^2$. There exists constants c_1, c_2, c_3 depending only on δ such that the following holds. There exists an $A' \subseteq A$ such that $|A'| \geq c_1|A|$ and for $c_2 n^2$ pairs $0 \leq i, j \leq n/c_3$, $|A' \cap [0, ic_3) \times [0, jc_3)| = ij$.*

REFERENCES

[1] Amirkhanyan, G., Bilyk, D., and Lacey, M., "Estimates of the discrepancy function in exponential orlicz spaces," *http://arxiv.org/abs/1306.1766.*

[2] Amirkhanyan, G., Bilyk, D., and Lacey, M., "Dichotomy results for the l_1 norm of the discrepancy function," *Journal of Mathematical Analysis and Applications*, vol. 410, pp. 1–6, 2014.

[3] Amirkhanyan, G., Bush, A., and Croot, E., "Order-preserving freiman isomorphisms," *In preparation.*

[4] Balog, A. and Szemerédi, E., "A statistical theorem of set addition.," *Combinatorica*, vol. 14, no. 3, pp. 263–268, 1994.

[5] Barvinok, A., *A Course in Convexity.* 2002.

[6] Beck, J. and Chen, W., "Irregularities of distribution," vol. 89, pp. xiv+294, 1987.

[7] Bilyk, D., "On roth's orthogonal function method in discrepancy theory," *Unif. Distrib. Theory*, vol. 6, no. 1, pp. 143–184, 2011.

[8] Bilyk, D. and Lacey, M., "On the small ball inequality in three dimensions," *Duke Math. J.*, vol. 143, no. 1, pp. 81–115, 2008.

[9] Bilyk, D. and Lacey, M., "The supremum norm of the discrepancy function: Recent results and connections," *Proceedings of MCQMC*, 2012.

[10] Bilyk, D., Lacey, M., and Vagharshakyan, A., "On the small ball inequality in all dimensions," *J. Funct. Anal.*, vol. 254, no. 9, pp. 2470–2502, 2008.

[11] Bilyk, D., Lacey, M. T., Parissis, I., and Vagharshakyan, A., "Exponential squared integrability of the discrepancy function in two dimensions," *Mathematika*, vol. 55, no. 1-2, pp. 1–27, 2009.

[12] Brown, T., Jungiċ, V., and Poelstra, A., "On double 3-term progressions," *arXiv:1304.1829.*

[13] Chen, W. and M., S., "Explicit constructions in the classical mean squares problem in irregularities of point distribution," *J. Reine Angew. Math.*, vol. 545, pp. 67–95, 2002.

[14] Chen, W. and M., S., "Orthogonality and digit shifts in the classical mean squares problem in irregularities of point distribution," pp. 141–159, 2008.

[15] Chen, W. W. L., "On irregularities of distribution," *Mathematika*, vol. 27, no. 2, pp. 153–170 (1981), 1980.

[16] Dick, J. and Pillichshammer, W., "Digital nets and sequences," pp. xviii+600, 2010. Discrepancy theory and quasi-Monte Carlo integration.

[17] Drmota, M. and Tichy, R., "Sequences, discrepancies and applications," vol. 1651, pp. xiv+503, 1997.

[18] Fine, N. J., "On the walsh functions," *Trans. Amer. Math. Soc.*, vol. 65, pp. 372–414, 1949.

[19] Gowers, W., "A new proof of szemerédi's theorem for arithmetic progressions of length four.," *Geom. Funct. Anal.*, vol. 8, no. 3, pp. 529–551, 1998.

[20] Halász, G., "On roth's method in the theory of irregularities of point distributions," pp. 79–94, 1981.

[21] Lacey, M. T., "On the discrepancy function in arbitrary dimension, close to l^1," *Anal. Math.*, vol. 34, no. 2, pp. 119–136, 2008.

[22] Price, J. J., "Certain groups of orthonormal step functions," *Canad. J. Math.*, vol. 9, pp. 413–425, 1957.

[23] Roth, K. F., "On irregularities of distribution," *Mathematika*, vol. 1, pp. 73–79, 1954.

[24] Roth, K. F., "On irregularities of distribution. iii," *Acta Arith.*, vol. 35, no. 4, pp. 373–384, 1979.

[25] Sanders, T., "On the bogolyubov-ruzsa lemma.," *Anal. PDE*, vol. 3, no. 3, pp. 627–655, 2012.

[26] Schipp, F., Wade, W. R., and Simon, P., *Walsh series.* Adam Hilger Ltd., 1990. An introduction to dyadic harmonic analysis; With the collaboration of J. Pál.

[27] Schmidt, W. M., "Irregularities of distribution. x," pp. 311–329, 1977.

[28] Skriganov, M. M., "Ergodic theory on sl(n), diophantine approximations and anomalies in the lattice point problem," *Invent. Math.*, vol. 132, no. 1, pp. 1–72, 1998.

[29] Skriganov, M. M., "Harmonic analysis on totally disconnected groups and irregularities of point distributions," *J. Reine Angew. Math.*, vol. 600, pp. 25–49, 2006.

[30] Skriganov, M. M., "The khinchin inequality and chen's theorem," *Algebra i Analiz*, vol. 23, no. 4, pp. 179–204, 2011.

[31] Skriganov, M. M., "On the mean values of l_q-discrepancies of point distributions," *St. Petersburg Math. J., to appear*, 2011.

[32] Solymosi, J., "On distinct consecutive differences.," *arXiv:math/0503069.*

[33] Tao, T. and Vu, V., *Additive Combinatorics*, vol. 105 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2006.