Upper Bounds in Classical Discrepancy Theory

William Chen and Maxim Skriganov

Abstract We discuss some of the ideas behind the study of upper bound questions in classical discrepancy theory. The many ideas involved come from diverse areas of mathematics and include diophantine approximation, probability theory, number theory and various forms of Fourier analysis. We illustrate these ideas by largely restricting our discussion to two dimensions.

1 Introduction

Classical discrepancy theory, or irregularities of distribution, began as a branch of the theory of uniform distribution but has independent interest. It is often viewed as a quantitative and substantially more precise version of the theory of uniform distribution, in the sense that one seeks to obtain very accurate bounds on various quantities arising from the difference between the discrete and the continuous. Here the discrete concerns the actual point count in a given region, which clearly takes integer values, whereas the continuous refers to the expectation of the point count, which depends on the area or volume of the region concerned and therefore can take non-integer values.

We shall first state the problem in a rather general form. Let $k \geq 2$ be a fixed integer. Our domain U will be a set of unit Lebesgue measure in k-dimensional euclidean space $\mathbf{R}^k$.

Suppose that $\mathscr{A}$ is a set of measurable subsets of U, endowed with an integral geometric measure, normalized so that the total measure is equal to unity. Suppose

William Chen
Department of Mathematics, Macquarie University, Sydney, NSW 2109, Australia,
e-mail: william.chen@mq.edu.au

Maxim Skriganov
Steklov Mathematical Institute, Fontanka 27, St Petersburg 190011, Russia,
e-mail: skrig@pdmi.ras.ru

further that $\mathscr{P}$ is a set of N points in U. For every subset $A \in \mathscr{A}$ of U, let

$$Z[\mathscr{P};A] = \#(\mathscr{P} \cap A)$$

denote the number of points of $\mathscr{P}$ that fall into A. This is the actual point count of $\mathscr{P}$ in A, with corresponding expectation $N\mu(A)$. By the discrepancy of $\mathscr{P}$ in A, we mean the difference

$$D[\mathscr{P};A] = Z[\mathscr{P};A] - N\mu(A).$$

Often, we consider the extreme discrepancy of $\mathscr{P}$ in U, taken to be the L^∞-norm

$$\|D[\mathscr{P}]\|_\infty = \sup_{A \in \mathscr{A}} |D[\mathscr{P};A]|. \tag{1}$$

However, for upper bound considerations, it is far more interesting and challenging to consider the corresponding L^2-norm

$$\|D[\mathscr{P}]\|_2 = \left(\int_{\mathscr{A}} |D[\mathscr{P};A]|^2 \, \mathrm{d}A \right)^{1/2}, \tag{2}$$

as well as the corresponding L^q-norms where $2 < q < \infty$.

For any given choice of U and $\mathscr{A}$, we are interested in studying the growth of the functions (1) and (2) as a function of N, the number of points of $\mathscr{P}$. It is the cornerstone of discrepancy theory that these quantities become arbitrarily large in many interesting cases, following the early conjecture of van der Corput [17, 18] and the pioneering work of van Aardenne-Ehrenfest [1, 2] and Roth [33]. A lower bound result is thus of the form

$$\|D[\mathscr{P}]\|_\infty > f(N) \quad \text{for all sets } \mathscr{P} \text{ of } N \text{ points in } U,$$

or of the form

$$\|D[\mathscr{P}]\|_2 > f(N) \quad \text{for all sets } \mathscr{P} \text{ of } N \text{ points in } U.$$

For upper bounds, we first make a simple observation. Consider a set $\mathscr{P}$ of N points, where all the points coincide. Then clearly any subset $A \in \mathscr{A}$ of U either contains all points of $\mathscr{P}$ or contains no point of $\mathscr{P}$. In either case, we expect the discrepancy $D[\mathscr{P};A]$ to have rather large absolute value for many of these sets A. This is an example of an extremely badly distributed point set. Such examples must never be allowed to play a role in upper bound considerations. After all, if the lower bound asserts that all distributions are *bad*, then a complementary upper bound must say that some distributions are *close to as good as they possibly can be*. Hence an upper bound result must be of the form

$$\|D[\mathscr{P}]\|_\infty < g(N) \quad \text{for some sets } \mathscr{P} \text{ of } N \text{ points in } U,$$

or of the form

$$\|D[\mathscr{P}]\|_2 < g(N) \quad \text{for some sets } \mathscr{P} \text{ of } N \text{ points in } U.$$

Our task is therefore to construct such a point set $\mathscr{P}$, or to show that one exists.

Of course, the ultimate task is to establish lower and upper bounds where the two functions $f(N)$ and $g(N)$ have the same order of magnitude. This has been achieved in a few instances, and we shall discuss the upper bound aspects of some of these in some detail in this article.

There are well known choices of U and $\mathscr{A}$ where the quantities (1) and (2) exceed N^δ for some positive exponent δ. We refer to these as large discrepancy phenomena. On the other hand, there are also well known choices of U and $\mathscr{A}$ where, for suitably chosen point sets $\mathscr{P}$, the quantities (1) and (2) can be bounded above by $(\log N)^\delta$ for some positive exponent δ. We refer to these as small discrepancy phenomena. As a general rule, upper bound questions are somewhat harder for small discrepancy phenomena, as we shall attempt to illustrate in the course of this article.

Notation. For any complex-valued function f and any positive function g, we write $f = O(g)$ to denote that there exists a positive constant C such that $|f| \le Cg$, and write $f = O_\delta(g)$ if the positive constant C may depend on a parameter δ. We also use the Vinogradov notation, where $f \ll g$ if $f = O(g)$, and $f \ll_\delta g$ if $f = O_\delta(g)$. We also write $f \gg g$ and $f \gg_\delta g$ to denote respectively $g \ll f$ and $g \ll_\delta f$, but here both f and g must be positive functions. The letters $\mathbf{N}$, $\mathbf{Z}$ and $\mathbf{R}$ denote respectively the set of all natural numbers, *i.e.* positive integers, the set of all integers and the set of all real numbers. We also write $\mathbf{N}_0$ to denote the set of all non-negative integers. For any real number z, we write $e(z) = \mathrm{e}^{2\pi \mathrm{i} z}$, and write $[z]$ and $\{z\}$ to denote respectively the integer part and the fractional part of z, *i.e.*

$$[z] = \max\{n \in \mathbf{Z} : n \le z\} \quad \text{and} \quad \{z\} = z - [z].$$

For any finite set $\mathscr{S}$, we denote by $\#\mathscr{S}$ the cardinality of $\mathscr{S}$. For any probabilistic variable ξ, we denote by $\mathbf{E}\xi$ the expected value of ξ.

Acknowledgment. The research of the second author has been supported by RFFI Project No. 08-01-00182.

2 Large Discrepancy – Main Results

The work on large discrepancy problems can best be summarized by the following ground-breaking result of Beck [4]. Consider the k-dimensional euclidean space $\mathbf{R}^k$. We take as our domain U the unit cube $[0,1]^k$, treated as a torus for simplicity. Let $B \subseteq [0,1]^k$ be a compact and convex set that satisfies a technical condition

$$r(B) \ge N^{-1/k}, \tag{3}$$

where $r(B)$ denotes the radius of the largest inscribed ball in B, and N is the cardinality of the point sets $\mathscr{P}$ under consideration. While this technical condition does

not really affect the argument in a serious way, it is nevertheless necessary in order for us to avoid degenerate cases. Let $\mathscr{T}$ denote the group of all orthogonal transformations in $\mathbf{R}^k$, normalized so that the total measure is equal to unity. For any contraction $\lambda \in [0,1]$, orthogonal transformation $\tau \in \mathscr{T}$ and translation $\mathbf{x} \in [0,1]^k$, we consider the similar copy

$$B(\lambda, \tau, \mathbf{x}) = \lambda(\tau B) + \mathbf{x}$$

of B. We then consider the collection

$$\mathscr{A} = \{B(\lambda, \tau, \mathbf{x}) : \lambda \in [0,1],\, \tau \in \mathscr{T},\, \mathbf{x} \in [0,1]^k\}$$

of all similar copies of B, where the integral geometric measure is given by a natural combination of the standard Lebesgue measures of λ and $\mathbf{x}$ and the measure of $\mathscr{T}$. More precisely, for any set $\mathscr{P}$ of N points in $[0,1]^k$, we have the L^2-norm

$$\|D[\mathscr{P}]\|_2 = \left(\int_{[0,1]^k} \int_{\mathscr{T}} \int_0^1 |D[\mathscr{P}; B(\lambda, \tau, \mathbf{x})]|^2 \, \mathrm{d}\lambda \, \mathrm{d}\tau \, \mathrm{d}\mathbf{x} \right)^{1/2}. \tag{4}$$

We also have the simpler L^∞-norm

$$\|D[\mathscr{P}]\|_\infty = \sup_{\substack{\lambda \in [0,1] \\ \tau \in \mathscr{T} \\ \mathbf{x} \in [0,1]^k}} |D[\mathscr{P}; B(\lambda, \tau, \mathbf{x})]|. \tag{5}$$

The following result is due to Beck [4].

Theorem 1. *Suppose that $B \subseteq [0,1]^k$ is a compact and convex set that satisfies the condition (3). Then for every set $\mathscr{P}$ of N points in $[0,1]^k$, we have*

$$\|D[\mathscr{P}]\|_2 \gg_B N^{1/2 - 1/2k}. \tag{6}$$

This leads immediately to the corresponding statement for the L^∞-norm.

Theorem 2. *Suppose that $B \subseteq [0,1]^k$ is a compact and convex set that satisfies the condition (3). Then for every set $\mathscr{P}$ of N points in $[0,1]^k$, we have*

$$\|D[\mathscr{P}]\|_\infty \gg_B N^{1/2 - 1/2k}. \tag{7}$$

The lower bound (6) is essentially best possible, in view of the following result of Beck and Chen which can be established as a simple case of their more general result in [6].

Theorem 3. *Suppose that $B \subseteq [0,1]^k$ is a compact and convex set. Then for every natural number N, there exists a set $\mathscr{P}$ of N points in $[0,1]^k$ such that*

$$\|D[\mathscr{P}]\|_2 \ll_B N^{1/2 - 1/2k}. \tag{8}$$

The proof of Theorem 3 is an extension of the original ideas needed to establish the following result using ideas in Beck [3]; see also Beck and Chen [5, Section 8.1].

Theorem 4. *Suppose that $B \subseteq [0,1]^k$ is a compact and convex set. Then for every natural number $N \geq 2$, there exists a set $\mathscr{P}$ of N points in $[0,1]^k$ such that*

$$\|D[\mathscr{P}]\|_\infty \ll_B N^{1/2-1/2k}(\log N)^{1/2}. \tag{9}$$

We shall discuss Beck's ideas in Section 4 and sketch a proof of the special case $k = 2$ of Theorem 4. Most important of all, however, the argument gives us a very good understanding of the exponent in the estimates (6)–(9).

We shall then sketch a proof of the special case $k = 2$ of Theorem 3 in Section 5.

3 A Seemingly Trivial Argument

We start by making an inadequate attempt to establish the special case $k = 2$ of Theorem 4. Such simple and perhaps naive attempts often play an important role in the study of upper bounds. Remember that we need to find a *good* set of points, and we often start by toying with some specific set of points which we hope will be good. Often it is not, but sometimes it permits us to bring in some stronger techniques at a later stage of the argument.

For simplicity, let us assume that the number of points is a perfect square, so that $N = M^2$ for some natural number M. We may then choose to split the unit square $[0,1]^2$ in the natural way into a union of $N = M^2$ little squares of side length M^{-1}, and then place a point in the centre of each little square, as shown in Figure 1 below.

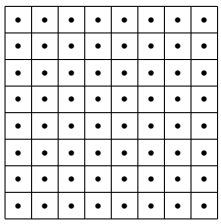

Fig. 1 A basic construction of $N = M^2$ points in the unit square

Suppose that $A = B(\lambda, \tau, \mathbf{x})$, where $\lambda \in [0,1]$, $\tau \in \mathscr{T}$ and $\mathbf{x} \in [0,1]^2$, is a similar copy of a given fixed compact and convex set B. We now attempt to estimate the discrepancy $D[\mathscr{P};A]$. Let $\mathscr{S}$ denote the collection of the $N = M^2$ little squares S of side length M^{-1}. The additive property of the discrepancy function then gives

$$D[\mathscr{P};A] = \sum_{S \in \mathscr{S}} D[\mathscr{P};S \cap A]. \tag{10}$$

Next, we make the simple observation that

$$D[\mathscr{P};S\cap A]=0 \quad \text{if } S\subseteq A \text{ or } S\cap A=\emptyset.$$

The identity (10) then becomes

$$D[\mathscr{P};A]=\sum_{\substack{S\in\mathscr{S}\\ S\cap\partial A\neq\emptyset}} D[\mathscr{P};S\cap A], \tag{11}$$

where ∂A denotes the boundary of A. Finally, observe that both $0\leq Z[\mathscr{P};S\cap A]\leq 1$ and $0\leq N\mu(S\cap A)\leq 1$, so that $|D[\mathscr{P};S\cap A]|\leq 1$, and it follows from (11) and the triangle inequality that

$$|D[\mathscr{P};A]|\leq \#\{S\in\mathscr{S}:S\cap\partial A\neq\emptyset\}\ll M=N^{1/2}. \tag{12}$$

This estimate is almost trivial, but very far from the upper bound $N^{1/4}(\log N)^{1/2}$ alluded to in Theorem 4.

We make an important observation here that the term $\#\{S\in\mathscr{S}:S\cap\partial A\neq\emptyset\}$ in (12) is intricately related to the length of the boundary curve ∂B of B; note that the set A is a similar copy of the given compact and convex set B. Indeed, in the general case of the problem in k-dimensional space, the corresponding term is intricately related to the $(k-1)$-dimensional volume of the boundary surface ∂B of B. It is worthwhile to record the important role played by boundary surface in large discrepancy problems.

4 A Large Deviation Technique

In this section, we continue our study of the special case $k=2$ of Theorem 4. Again, let us assume that the number of points is a perfect square, so that $N=M^2$ for some natural number M. Again, we choose to split the unit square $[0,1]^2$ in the natural way into a union of $N=M^2$ little squares of side length M^{-1}. As before, let $\mathscr{S}$ denote the collection of the $N=M^2$ little squares S of side length M^{-1}.

For every little square $S\in\mathscr{S}$, instead of placing a point in the centre of the square, we now associate a random point $\widetilde{\mathbf{p}}_S\in S$, uniformly distributed within the little square S and independent of all the other random points in the other little squares. We thus obtain a random point set

$$\widetilde{\mathscr{P}}=\{\widetilde{\mathbf{p}}_S:S\in\mathscr{S}\}. \tag{13}$$

Suppose that a fixed compact and convex set $B\subseteq[0,1]^2$ is given. Let

$$\mathscr{G}=\left\{B(\lambda,\tau,\mathbf{x}):\lambda\in\left[0,\frac{11}{10}\right],\tau\in\mathscr{T},\mathbf{x}\in[0,1]^2\right\}.$$

Note that the collection $\mathscr{G}$ contains the collection $\mathscr{A}$ and permits some similar copies of B which are a little bigger than B. Then one can find a subset $\mathscr{H}$ of $\mathscr{G}$ such that

$$\#\mathscr{H} \leq N^{C_1},$$

where C_1 is a positive constant depending at most on B, and such that for every $A \in \mathscr{A}$, there exist $A^-, A^+ \in \mathscr{H}$ such that

$$A^- \subseteq A \subseteq A^+ \quad \text{and} \quad \mu(A^+ \setminus A^-) \leq N^{-1}. \tag{14}$$

We comment that such a set $\mathscr{H}$ may not exist if we make the restriction $\mathscr{H} \subseteq \mathscr{A}$ instead of the more generous restriction $\mathscr{H} \subseteq \mathscr{G}$.

Suppose that $A \in \mathscr{H}$ is fixed. Then, analogous to the discrepancy function (10), we now consider the discrepancy function

$$D[\widetilde{\mathscr{P}}; A] = \sum_{S \in \mathscr{S}} D[\widetilde{\mathscr{P}}; S \cap A] = \sum_{\substack{S \in \mathscr{S} \\ S \cap \partial A \neq \emptyset}} D[\widetilde{\mathscr{P}}; S \cap A], \tag{15}$$

and note as before that

$$\#\{S \in \mathscr{S} : S \cap \partial A \neq \emptyset\} \ll M = N^{1/2}. \tag{16}$$

For every $S \in \mathscr{S}$, let

$$\phi_S = \begin{cases} 1, \text{ if } \widetilde{\mathbf{p}}_S \in A, \\ 0, \text{ otherwise.} \end{cases}$$

The observation

$$D[\widetilde{\mathscr{P}}; A] = \sum_{S \in \mathscr{S}} (\phi_S - \mathbf{E}\phi_S) = \sum_{\substack{S \in \mathscr{S} \\ S \cap \partial A \neq \emptyset}} (\phi_S - \mathbf{E}\phi_S) \tag{17}$$

sets us up to appeal to large deviation type inequalities in probability theory. For instance, we can use the following result attributed to Hoeffding; see, for instance, Pollard [31, Appendix B].

Lemma 1. *Suppose that $\phi_1, \ldots, \phi_m$ are independent random variables that satisfy $0 \leq \phi_i \leq 1$ for every $i = 1, \ldots, m$. Then for every real number $\gamma > 0$, we have*

$$\text{Prob}\left(\left| \sum_{i=1}^{m} (\phi_i - \mathbf{E}\phi_i) \right| \geq \gamma \right) \leq 2\mathrm{e}^{-2\gamma^2/m}.$$

In view of (17), we now apply Lemma 1 with

$$m = \#\{S \in \mathscr{S} : S \cap \partial A \neq \emptyset\} \leq C_2 N^{1/2},$$

where C_2 is a positive constant depending at most on the given set B, and with

$$\gamma = C_3 N^{1/4} (\log N)^{1/2},$$

where C_3 is a sufficiently large positive constant. Indeed,

$$\frac{\gamma^2}{m} \geq \frac{C_3^2}{C_2} \log N,$$

and it follows therefore that

$$2\mathrm{e}^{-2\gamma^2/m} \leq \frac{1}{2} N^{-C_1} \leq \frac{1}{2} (\#\mathscr{H})^{-1}$$

provided that C_3 is chosen sufficiently large in terms of C_1 and C_2. Then

$$\mathrm{Prob}\left(|D[\widetilde{\mathscr{P}};A]| \geq C_3 N^{1/4} (\log N)^{1/2} \right) \leq \frac{1}{2} (\#\mathscr{H})^{-1},$$

and so

$$\mathrm{Prob}\left(|D[\widetilde{\mathscr{P}};A]| \geq C_3 N^{1/4} (\log N)^{1/2} \text{ for some } A \in \mathscr{H} \right) \leq \frac{1}{2},$$

whence

$$\mathrm{Prob}\left(|D[\widetilde{\mathscr{P}};A]| \leq C_3 N^{1/4} (\log N)^{1/2} \text{ for all } A \in \mathscr{H} \right) \geq \frac{1}{2}.$$

In other words, there exists a set $\mathscr{P}^*$ of $N = M^2$ points in $[0,1]^2$ such that

$$|D[\mathscr{P}^*;A]| \leq C_3 N^{1/4} (\log N)^{1/2} \quad \text{for every } A \in \mathscr{H}.$$

Suppose now that $A \in \mathscr{A}$ is given. Then there exist $A^-, A^+ \in \mathscr{H}$ such that (14) is satisfied. It is not difficult to show that

$$\begin{aligned} |D[\mathscr{P}^*;A]| &\leq \max\left\{ |D[\mathscr{P}^*;A^-]|, |D[\mathscr{P}^*;A^+]| \right\} + N\mu(A^+ \setminus A^-) \\ &\leq C_3 N^{1/4} (\log N)^{1/2} + 1. \end{aligned}$$

Theorem 4 for $k = 2$ in the special case when $N = M^2$ is therefore established.

Finally, we can easily lift the restriction that N is a perfect square. By Lagrange's theorem, every positive integer N can be written as a sum

$$N = M_1^2 + M_2^2 + M_3^2 + M_4^2$$

of the squares of four non-negative integers. We can therefore superimpose up to four point distributions in $[0,1]^2$ where the number of points in each is a perfect square. This completes the proof of Theorem 4 for $k = 2$.

5 An Averaging Argument

In this section, we indicate how the argument in the previous section can be adapted to establish Theorem 3 in the case $k = 2$.

We construct the random point set $\mathscr{P}$, given by (13), as before. Suppose that a fixed compact and convex set $B \subseteq [0,1]^2$ is given. Let $A \in \mathscr{A}$ be fixed. Then (15), (16) and (17) are valid. If we write $\eta_S = \phi_S - \mathbf{E}\phi_S$, then

$$|D[\widetilde{\mathscr{P}};A]|^2 = \sum_{\substack{S_1,S_2\in\mathscr{S}\\ S_1\cap\partial A\neq\emptyset\\ S_2\cap\partial A\neq\emptyset}} \eta_{S_1}\eta_{S_2}.$$

Taking expectation over all the $N = M^2$ random points, we have

$$\mathbf{E}\left(|D[\widetilde{\mathscr{P}};A]|^2\right) = \sum_{\substack{S_1,S_2\in\mathscr{S}\\ S_1\cap\partial A\neq\emptyset\\ S_2\cap\partial A\neq\emptyset}} \mathbf{E}\left(\eta_{S_1}\eta_{S_2}\right). \tag{18}$$

The random variables η_S, where $S \in \mathscr{S}$, are independent since the distribution of the random points are independent of each other. If $S_1 \neq S_2$, then

$$\mathbf{E}\left(\eta_{S_1}\eta_{S_2}\right) = \mathbf{E}\left(\eta_{S_1}\right)\mathbf{E}\left(\eta_{S_2}\right) = 0.$$

It follows that the only non-zero contributions to the sum (18) come from those terms where $S_1 = S_2$, so that

$$\mathbf{E}\left(|D[\widetilde{\mathscr{P}};A]|^2\right) \leq \#\{S \in \mathscr{S} : S\cap\partial A \neq \emptyset\} \ll_B N^{1/2}.$$

Integrating now over all $A \in \mathscr{A}$ and changing the order of integration, we obtain

$$\mathbf{E}\left(\int_{\mathscr{A}} |D[\widetilde{\mathscr{P}};A]|^2\,\mathrm{d}A\right) \ll_B N^{1/2}.$$

It follows that there exists a set $\mathscr{P}^*$ of $N = M^2$ points in $[0,1]^2$ such that

$$\int_{\mathscr{A}} |D[\mathscr{P}^*;A]|^2\,\mathrm{d}A \ll_B N^{1/2},$$

establishing Theorem 3 for $k = 2$ in the special case when $N = M^2$.

The generalization to all positive integers N follow from Lagrange's theorem as before, and this completes the proof of Theorem 3 for $k = 2$.

We remark that the argument in Sections 3–5 can be extended in a reasonably straightforward manner to arbitrary dimensions $k \geq 2$. Also the argument in this section on Theorem 3 can be extended to L^q-norms for all even positive integers q, and hence all positive real numbers q, without too many complications.

6 A Comparison of Deterministic and Probabilistic Techniques

In this section, we make a digression and use Fourier transform techniques to try to understand and relate various approaches to upper bounds in large discrepancy problems.

Consider the unit cube $U = [0,1]^k$, treated as a torus, in euclidean space $\mathbf{R}^k$. Suppose that a natural number N is given, and that $N = M^k$ for some natural number M. We shall partition U into a union of $N = M^k$ cubes of sidelength M^{-1} in the natural way, and denote by $\mathscr{S}$ the collection of these small cubes. For every cube $S \in \mathscr{S}$, we denote by $\mathbf{p}_S$ the point in the centre of S. Then

$$\mathscr{P} = \{\mathbf{p}_S : S \in \mathscr{S}\} \tag{19}$$

is a collection of $N = M^k$ points in $U = [0,1]^k$.

Let ν be a probabilistic measure on U. For every cube $S \in \mathscr{S}$, let ν_S denote the translation of ν by $\mathbf{p}_S$, so that

$$\int_U f(\mathbf{u})\,\mathrm{d}\nu_S = \int_U f(\mathbf{u}-\mathbf{p}_S)\,\mathrm{d}\nu$$

for any integrable function f. Furthermore, let $\boldsymbol{\xi}_S$ denote the probabilistic variable associated with the probabilistic measure ν_S. Then

$$\widetilde{\mathscr{P}} = \{\boldsymbol{\xi}_S : S \in \mathscr{S}\}$$

is a random set of $N = M^k$ points in $U = [0,1]^k$.

Let A denote a compact and convex set in $U = [0,1]^k$. For every $\mathbf{x} \in [0,1]^k$, let $A(\mathbf{x}) = A + \mathbf{x}$ denote the translation of A by $\mathbf{x}$. Now consider the average

$$D_\nu^2(N;A) = \int_U \dots \int_U \left(\int_{[0,1]^k} |D[\widetilde{\mathscr{P}};A(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x} \right) \prod_{S\in\mathscr{S}} \mathrm{d}\nu_S. \tag{20}$$

In other words, for every realization of $\widetilde{\mathscr{P}}$, we consider the mean square average of the discrepancy function over all translations of A. We then average over all the different realizations of $\widetilde{\mathscr{P}}$, with the understanding that the probabilistic measures ν_S, where $S \in \mathscr{S}$, are independent.

We can describe $D_\nu^2(N;A)$ rather precisely in terms of the Fourier transforms of the measure ν and of the characteristic function χ_A of the set A.

Proposition 1. *For any natural number $N = M^k$, any compact and convex set A in $U = [0,1]^k$ and any probabilistic measure ν on U, we have*

$$D_\nu^2(N;A) = N \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2 (1 - |\widehat{\nu}(\mathbf{t})|^2) + N^2 \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(M\mathbf{t})|^2 |\widehat{\nu}(M\mathbf{t})|^2. \tag{21}$$

Before we proceed to establish this proposition, we shall first of all endeavour to understand the significance of the each of the two terms on the right hand side of (21).

Suppose first of all that ν is the Dirac measure δ_0 concentrated at the origin. Then the Fourier transform $\widehat{\nu}(\mathbf{t}) = 1$ identically, so the first term on the right hand side of the identity (21) vanishes, and we have

$$D^2_{\delta_0}(N;A) = N^2 \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(M\mathbf{t})|^2. \tag{22}$$

On the other hand, note that under this measure δ_0, the only realization of the random set $\widetilde{\mathscr{P}}$ is the set $\mathscr{P}$ given by (19). This represents a deterministic model.

Suppose next that ν is the uniform measure λ supported by the cube

$$\left[-\frac{1}{2M},\frac{1}{2M}\right]^k, \tag{23}$$

so that $\mathrm{d}\lambda = \lambda(\mathbf{u})\,\mathrm{d}\mathbf{u}$, where

$$\lambda(\mathbf{u}) = N\chi_{[-1/2M,1/2M]^k}(\mathbf{u})$$

denotes the characteristic function of the cube (23), suitably normalized. It is well known that for every $\mathbf{t} = (t_1,\ldots,t_k) \in \mathbf{Z}^k$, the Fourier transform

$$\widehat{\lambda}(\mathbf{t}) = N\prod_{i=1}^{k} \frac{\sin(\pi M^{-1}t_i)}{\pi t_i},$$

with suitable modification when $t_i = 0$ for some $i = 1,\ldots,k$. Since $\widehat{\lambda}(M\mathbf{t}) = 0$ for every non-zero $\mathbf{t} \in \mathbf{Z}^k$, the second term on the right hand side of the identity (21) vanishes, and we have

$$D^2_{\lambda}(N;A) = N \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2(1-|\widehat{\lambda}(\mathbf{t})|^2).$$

On the other hand, note that under this uniform measure λ, each of the probabilistic variables $\boldsymbol{\xi}_S$, where $S \in \mathscr{S}$, represents a random point uniformly distributed within the cube S. This represents a probabilistic model the special case $k = 2$ of which has been described earlier in Sections 4–5.

In summary, the two terms on the right hand side of the identity (21) may be interpreted as respectively the probabilistic and the deterministic part of the quantity $D^2_{\nu}(N;A)$.

Proof of Proposition 1. Applying Parseval's identity to the inner integral in (20), we obtain

$$D_\nu^2(N;A) = \int_U \dots \int_U \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2 \left| \sum_{X\in\mathscr{S}} e(\mathbf{t}\cdot\boldsymbol{\xi}_X)\right|^2 \prod_{S\in\mathscr{S}} \mathrm{d}\nu_S$$
$$= \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2 \int_U \dots \int_U \sum_{X,Y\in\mathscr{S}} e(\mathbf{t}\cdot\boldsymbol{\xi}_X) e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y) \prod_{S\in\mathscr{S}} \mathrm{d}\nu_S$$
$$= \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2 \sum_{X,Y\in\mathscr{S}} \int_U \dots \int_U e(\mathbf{t}\cdot\boldsymbol{\xi}_X) e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y) \prod_{S\in\mathscr{S}} \mathrm{d}\nu_S$$
$$= \sum_{\mathbf{0}\neq\mathbf{t}\in\mathbf{Z}^k} |\widehat{\chi_A}(\mathbf{t})|^2 \left(N + \sum_{\substack{X,Y\in\mathscr{S}\\ X\neq Y}} \int_U \int_U e(\mathbf{t}\cdot\boldsymbol{\xi}_X) e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y)\, \mathrm{d}\nu_X\, \mathrm{d}\nu_Y \right). \quad (24)$$

For $X \neq Y$, we clearly have

$$\int_U \int_U e(\mathbf{t}\cdot\boldsymbol{\xi}_X) e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y)\, \mathrm{d}\nu_X\, \mathrm{d}\nu_Y = \int_U \int_U e(\mathbf{t}\cdot(\boldsymbol{\xi}_X - \mathbf{p}_X)) e(-\mathbf{t}\cdot(\boldsymbol{\xi}_Y - \mathbf{p}_Y))\, \mathrm{d}\nu\, \mathrm{d}\nu$$
$$= e(-\mathbf{t}\cdot\mathbf{p}_X) e(\mathbf{t}\cdot\mathbf{p}_Y) \int_U e(\mathbf{t}\cdot\boldsymbol{\xi}_X)\, \mathrm{d}\nu \int_U e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y)\, \mathrm{d}\nu = |\widehat{\nu}(\mathbf{t})|^2 e(-\mathbf{t}\cdot\mathbf{p}_X) e(\mathbf{t}\cdot\mathbf{p}_Y),$$

and so

$$\sum_{\substack{X,Y\in\mathscr{S}\\ X\neq Y}} \int_U \int_U e(\mathbf{t}\cdot\boldsymbol{\xi}_X) e(-\mathbf{t}\cdot\boldsymbol{\xi}_Y)\, \mathrm{d}\nu_X\, \mathrm{d}\nu_Y = |\widehat{\nu}(\mathbf{t})|^2 \sum_{\substack{X,Y\in\mathscr{S}\\ X\neq Y}} e(-\mathbf{t}\cdot\mathbf{p}_X) e(\mathbf{t}\cdot\mathbf{p}_Y)$$
$$= |\widehat{\nu}(\mathbf{t})|^2 \left(\sum_{X,Y\in\mathscr{S}} e(-\mathbf{t}\cdot\mathbf{p}_X) e(\mathbf{t}\cdot\mathbf{p}_Y) - N \right)$$
$$= |\widehat{\nu}(\mathbf{t})|^2 \left(\left| \sum_{X\in\mathscr{S}} e(\mathbf{t}\cdot\mathbf{p}_X) \right|^2 - N \right). \quad (25)$$

The identity (21) follows easily on combining (24), (25) and the orthogonality relationship

$$\left| \sum_{X\in\mathscr{S}} e(\mathbf{t}\cdot\mathbf{p}_X) \right| = \begin{cases} N, \text{ if } \mathbf{t}\in M\mathbf{Z}^k, \\ 0, \text{ otherwise.} \end{cases}$$

This completes the proof. □

In the special cases when $N = M^k$ and when we have sufficient knowledge on the Fourier transform of the characteristic function of the given compact and convex set B, we expect to be able to establish the inequality (8) in Theorem 3 for the set $\mathscr{P}$ given by (19). This will give a deterministic proof of Theorem 3, an alternative to the probabilistic proof briefly described in Sections 4–5. However, there is virtually no documentation of results of this kind in the literature, apart from the special case when $N = M^2$ is odd and the set B is a cube, described in Chen [12, Section 3].

Nevertheless, the question arises as to whether a deterministic technique or a probabilistic technique gives a better upper bound. Much of the description in this section arises as a consequence of work done in this direction by Chen and Travaglini [16] for the case when B is a ball of fixed radius, so there is no contraction. Note also that since B is a ball, orthogonal transformation is redundant. Hence there is only translation.

Returning to the beginning of this section, we let A denote a ball in $U = [0,1]^k$, of fixed radius not exceeding $\frac{1}{2}$. We shall consider translations $A(\mathbf{x}) = A + \mathbf{x}$ of A, where $\mathbf{x} \in [0,1]^k$. We have the following surprising result.

Proposition 2. *Suppose that $k \not\equiv 1 \bmod 4$.*

a) If k is sufficiently large, then the inequality $D_\lambda(M^k;A) < D_{\delta_0}(M^k;A)$ holds for all sufficiently large natural numbers M.
b) For $k = 2$ and ball A of radius $\frac{1}{4}$, the inequality $D_{\delta_0}(M^k;A) < D_\lambda(M^k;A)$ holds for all sufficiently large natural numbers M.

Suppose that $k \equiv 1 \bmod 4$.

c) If k is sufficiently large, then the inequality $D_\lambda(M^k;A) < D_{\delta_0}(M^k;A)$ holds for infinitely many natural numbers M.
d) The inequality $D_{\delta_0}(M^k;A) < D_\lambda(M^k;A)$ holds for infinitely many natural numbers M.
e) For $k = 1$, the inequality $D_{\delta_0}(M^k;A) < D_\lambda(M^k;A)$ holds for every natural number M.

The case $k \not\equiv 1 \bmod 4$ is the standard case, whereas the case $k \equiv 1 \bmod 4$ is the exceptional case. This exceptional case is intimately related to the work of Konyagin, Skriganov and Sobolev [27] concerning the peculiar distribution of lattice points with respect to balls in these dimensions. We shall give a very brief description of the underlying ideas.

It is fairly straightforward to show that for every fixed dimension k, we have

$$D_\lambda^2(M^k;A) \leq \frac{\pi^{k/2} k^{3/2} r^{k-1} M^{k-1}}{2\Gamma(1+k/2)} \tag{26}$$

if M is sufficiently large, where r denotes the radius of the ball A.

To study the term $D_{\delta_0}^2(M^k;A)$, we make use of the identity (22). Suppose that A is a ball of radius r centred at the origin. Then the Fourier transform $\widehat{\chi_A}$ can be described in terms of Bessel functions. Roughly speaking, we can write

$$D_{\delta_0}^2(M^k;A) = M^k \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbf{Z}^k} r^k |\mathbf{t}|^{-k} J_{k/2}^2(2\pi r M |\mathbf{t}|), \tag{27}$$

where the Bessel function term $J_{k/2}^2(2\pi r M|\mathbf{t}|)$ is dominated by

$$\frac{1}{\pi^2 r M |\mathbf{t}|} \cos^2\left(2\pi r M |\mathbf{t}| - \frac{(k+1)\pi}{4}\right).$$

Suppose that $k \not\equiv 1 \bmod 4$. Then elementary calculation gives

$$\max\left\{\cos^2\left(2\pi rM - \frac{(k+1)\pi}{4}\right), \cos^2\left(4\pi rM - \frac{(k+1)\pi}{4}\right)\right\} > \frac{1}{100},$$

for instance, ensuring a significant contribution to the sum in (27) from those $\mathbf{t}$ satisfying $|\mathbf{t}| = 1$ or from those $\mathbf{t}$ satisfying $|\mathbf{t}| = 2$, sufficient for us to show that

$$D^2_{\delta_0}(M^k;A) \geq \frac{kr^{k-1}M^{k-1}}{1000\pi^2 2^k}. \tag{28}$$

For sufficiently large k, one has

$$\frac{\pi^{k/2}k^{3/2}}{2\Gamma(1+k/2)} < \frac{k}{1000\pi^2 2^k}.$$

Combining this with (26) and (28) gives part (a) of Proposition 2.

Suppose that $k \equiv 1 \bmod 4$. Then the Bessel function term $J^2_{k/2}(2\pi rM|\mathbf{t}|)$ in (27) is dominated by

$$\frac{1}{\pi^2 rM|\mathbf{t}|}\cos^2\left(2\pi rM|\mathbf{t}| \pm \frac{\pi}{2}\right) = \frac{1}{\pi^2 rM|\mathbf{t}|}\sin^2(2\pi rM|\mathbf{t}|).$$

For many values of M, the terms $\sin^2(2\pi rM|\mathbf{t}|)$ can be simultaneously small for all small $|\mathbf{t}|$, making $D_{\delta_0}(M^k;A)$ unusually small. This goes towards explaining parts (c) and (d) of Proposition 2.

7 Small Discrepancy – The Classical Problem

To illustrate the work on small discrepancy problems, we shall consider the pioneering work of Roth [33] on the classical problem in connection with aligned rectangular boxes in the unit cube. Consider the k-dimensional euclidean space $\mathbf{R}^k$. We take as our domain U the unit cube $[0,1]^k$. For every $\mathbf{x} = (x_1,\ldots,x_k) \in [0,1]^k$, we consider the aligned rectangular box

$$B(\mathbf{x}) = [0,x_1) \times \ldots \times [0,x_k),$$

anchored at the origin. Here the condition that the intervals do not include the right hand endpoints is unimportant but a very convenient technical device. On the other hand, the assumption that all such boxes are anchored at the origin is purely historical, but is necessary if one wants to have a deeper understanding of the problem. We then consider the collection

$$\mathscr{A} = \{B(\mathbf{x}) : \mathbf{x} \in [0,1]^k\}$$

of all such aligned rectangular boxes in U, where the integral geometric measure is given by the natural Lebesgue measure of $\mathbf{x}$. More precisely, for any set $\mathscr{P}$ of N points in $[0,1]^k$, we have the L^2-norm

$$\|D[\mathscr{P}]\|_2 = \left(\int_{[0,1]^k} |D[\mathscr{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \right)^{1/2}. \tag{29}$$

We also have the simpler L^∞-norm

$$\|D[\mathscr{P}]\|_\infty = \sup_{\mathbf{x} \in [0,1]^k} |D[\mathscr{P}; B(\mathbf{x})]|. \tag{30}$$

The following result is due to Roth [33].

Theorem 5. *For every set $\mathscr{P}$ of N points in $[0,1]^k$, we have*

$$\|D[\mathscr{P}]\|_2 \gg_k (\log N)^{(k-1)/2}. \tag{31}$$

This leads immediately to the corresponding statement for the l^∞-norm.

Theorem 6. *For every set $\mathscr{P}$ of N points in $[0,1]^k$, we have*

$$\|D[\mathscr{P}]\|_\infty \gg_k (\log N)^{(k-1)/2}. \tag{32}$$

It is well known that Theorem 6 is not sharp. In dimension $k = 2$, Schmidt [36] has shown that for every set $\mathscr{P}$ of N points in $[0,1]^2$, we have

$$\|D[\mathscr{P}]\|_\infty \gg \log N. \tag{33}$$

An alternative proof of this can be found in Halász [22]. On the other hand, the recent work of Bilyk and Lacey [8] and of Bilyk, Lacey and Vagharshakyan [9] has shown that for every dimension $k \geq 3$, there exists a positive constant $\delta(k)$ such that for every set $\mathscr{P}$ of N points in $[0,1]^k$, we have

$$\|D[\mathscr{P}]\|_\infty \gg (\log N)^{(k-1)/2+\delta(k)}. \tag{34}$$

See elsewhere in this volume for a detailed discussion of this question.

The lower bound (31) is essentially best possible, in view of the following result of Roth [35].

Theorem 7. *For every natural number $N \geq 2$, there exists a set $\mathscr{P}$ of N points in $[0,1]^k$ such that*

$$\|D[\mathscr{P}]\|_2 \ll_k (\log N)^{(k-1)/2}. \tag{35}$$

The special cases $k = 2$ and $k = 3$ have been established earlier by Davenport [19] and Roth [34] respectively.

The first proof of Theorem 7 is based on a probabilistic variant of the idea first used to establish the following result of Halton [23].

Theorem 8. *For every natural number $N \geq 2$, there exists a set $\mathscr{P}$ of N points in $[0,1]^k$ such that*

$$\|D[\mathscr{P}]\|_\infty \ll_k (\log N)^{k-1}. \tag{36}$$

The special case $k = 2$ has been known for over 100 years through the work of Lerch [28].

Note that in dimension $k = 2$, Theorem 8 shows that Schmidt's lower bound (33) is best possible. In dimensions $k \geq 3$, there remains a significant gap between the lower bound (34) and the upper bound (36). This is sometimes referred to as the *Great Open Problem*.

8 Diophantine Approximation and Davenport Reflection

We begin by making a fatally flawed attempt to establish[1] the special case $k = 2$ of Theorem 8.

Again, for simplicity, let us assume that the number of points is a perfect square, so that $N = M^2$ for some natural number M. We may then choose to split the unit square $[0,1]^2$ in the natural way into a union of $N = M^2$ little squares of sidelength M^{-1}, and then place a point in the centre of each little square. Let $\mathscr{P}$ be the collection of these $N = M^2$ points.

Let ξ be the second coordinate of one of the points of $\mathscr{P}$. Clearly, there are precisely M points in $\mathscr{P}$ sharing this second coordinate. Consider the discrepancy

$$D[\mathscr{P}; B(1,x_2)] \tag{37}$$

of the rectangle $B(1,x_2) = [0,1) \times [0,x_2)$. As x_2 increases from just less than ξ to just more than ξ, the value of (37) increases by M. It follows immediately that

$$\|D[\mathscr{P}]\|_\infty \geq \frac{1}{2} M \gg N^{1/2}.$$

Let us make a digression to the work of Hardy and Littlewood [24, 25] on the distribution of lattice points in a right angled triangle. Consider a large right angled triangle T with two sides parallel to the coordinate axes. We are interested in the number of points of the lattice $\mathbf{Z}^2$ that lie in T. For simplicity, the triangle T is placed so that the horizontal side is precisely halfway between two neighbouring rows of $\mathbf{Z}^2$ and the vertical side is precisely halfway between two neighbouring columns of $\mathbf{Z}^2$, as shown in Figure 2.

Note that the lattice $\mathbf{Z}^2$ has precisely one point per unit area, so we can think of the area of T as the expected number of lattice points in T. We therefore wish to understand the difference between the number of lattice points in T and the area of

[1] It was put to the first author by a rather preposterous engineering colleague many years ago that this could be achieved easily by a square lattice in the obvious way. Not quite the case, as an obvious way would be far from so to this colleague.

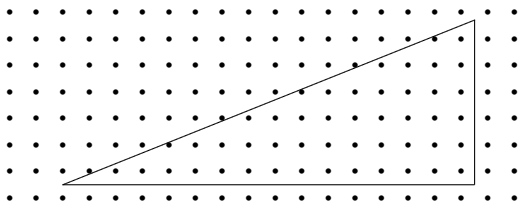

Fig. 2 Lattice points in a right angled triangle

T, and this is the discrepancy of $\mathbf{Z}^2$ in T. The careful placement of the horizontal and vertical sides of T means that the discrepancy comes solely from the third side of T. In the work of Hardy and Littlewood, it is shown that the size of the discrepancy when T is large is intimately related to the arithmetic properties of the slope of this third side of T. In particular, the discrepancy is essentially smallest when this slope is a badly approximable number[2].

Returning to our attempt to establish the special case $k = 2$ of Theorem 8, perhaps our approach is not quite fatally flawed as we have thought earlier, in view of our knowledge of the work of Hardy and Littlewood. Suppose that a positive integer $N \geq 2$ is given. The lattice

$$(N^{-1/2}\mathbf{Z})^2 \tag{38}$$

contains precisely N points per unit area. Inspired by Hardy and Littlewood, we now rotate (38) by an angle θ, chosen so that $\tan\theta$ is a badly approximable number. Let us denote the resulting lattice by Λ. Then $\Lambda \cap [0,1]^2$ has roughly N points. Deleting or adding a few points, we end up with a set $\mathscr{P}$ of precisely N points in $[0,1]^2$. It can then be shown that $\|D[\mathscr{P}]\|_\infty \ll \log N$, establishing Theorem 8 for $k = 2$. For the details, see the paper of Chen and Travaglini [15].

Indeed, this approach has been known for some time, as Beck and Chen [7] have already used this idea earlier in an alternative proof of Theorem 7 for $k = 2$. In fact, the first proof of this result by Davenport [19] makes essential use of diophantine approximation and badly approximable numbers, but in a slightly different and less obvious way. We now proceed to describe this.

Recall that $U = [0,1]^2$ in this case. For the sake of convenience, we shall assume that the intervals are closed on the left and open on the right. We are also going to rescale U. Suppose first of all that N is a given even positive integer, with $N = 2M$. We now rescale U in the vertical direction by a factor M to obtain

$$V = [0,1) \times [0,M).$$

Consider now the infinite lattice Λ_1 on $\mathbf{R}^2$ generated by the two vectors

$$(1,0) \quad \text{and} \quad (\theta,1),$$

[2] For those readers not familiar with the theory of diophantine approximation, just take any quadratic irrational like $\sqrt{2}$ or $\sqrt{3}$.

where the arithmetic properties of the non-zero number θ will be described later. It is not difficult to see that the set

$$\mathscr{Q}_1 = \Lambda_1 \cap V = \{(\{\theta n\}, n) : n = 0, 1, 2, \ldots, M-1\}$$

contains precisely M points. We now wish to study the discrepancy properties of the set $\mathscr{Q}_1$ in V. For every aligned rectangle

$$B(x_1, y) = [0, x_1) \times [0, y) \subseteq V,$$

we consider the discrepancy

$$E[\mathscr{Q}_1; B(x_1, y)] = \#(\mathscr{Q}_1 \cap B(x_1, y)) - x_1 y, \tag{39}$$

noting that the area of $B(x_1, y)$ is equal to $x_1 y$, and that there is an average of one point of $\mathscr{Q}_1$ per unit area in V. Suppose first of all that y is an integer satisfying $0 < y \leq M$. Then we can write

$$E[\mathscr{Q}_1; B(x_1, y)] = \sum_{0 \leq n < y} (\psi(\theta n - x_1) - \psi(\theta n)),$$

for all but finitely many x_1 satisfying $0 < x_1 \leq 1$, where $\psi(z) = z - [z] - \frac{1}{2}$ for every $z \in \mathbf{R}$. If we relax the condition that y is an integer, then for every real number y satisfying $0 < y \leq M$, we have the approximation

$$E[\mathscr{Q}_1; B(x_1, y)] = \sum_{0 \leq n < y} (\psi(\theta n - x_1) - \psi(\theta n)) + O(1)$$

for all but finitely many x_1 satisfying $0 < x_1 \leq 1$. For simplicity, let us write

$$E[\mathscr{Q}_1; B(x_1, y)] \approx \sum_{0 \leq n < y} (\psi(\theta n - x_1) - \psi(\theta n)).$$

The sawtooth function $\psi(z)$ is periodic, so it is natural to use its Fourier series, and we obtain the estimate

$$E[\mathscr{Q}_1; B(x_1, y)] \approx \sum_{0 \neq m \in \mathbf{Z}} \left(\frac{1 - e(-m x_1)}{2 \pi \mathrm{i} m} \right) \left(\sum_{0 \leq n < y} e(\theta n m) \right). \tag{40}$$

Ideally we would like to square the expression (40) and integrate with respect to the variable x_1 over $[0, 1]$. Unfortutnately, the term 1 in the numerator on the right hand side, arising from the assumption that the rectangles we consider are anchored at the origin, proves to be more than a nuisance.

To overcome this problem, Davenport's brilliant idea is to introduce a second lattice Λ_2 on $\mathbf{R}^2$ generated by the two vectors

$$(1, 0) \quad \text{and} \quad (-\theta, 1).$$

It is not difficult to see that the set

$$\mathscr{Q}_2 = \Lambda_2 \cap V = \{(\{-\theta n\}, n) : n = 0, 1, 2, \ldots, M-1\}$$

again contains precisely M points. Then the set

$$\mathscr{Q} = \mathscr{Q}_1 \cup \mathscr{Q}_2 = \{(\{\pm\theta n\}, n) : n = 0, 1, 2, \ldots, M-1\},$$

where the points are counted with multiplicity, contains precisely $2M$ points. Thus analogous to the discrepancy (39), we now consider the discrepancy

$$F[\mathscr{Q}; B(x_1, y)] = \#(\mathscr{Q} \cap B(x_1, y)) - 2x_1 y,$$

noting that there is now an average of two points of $\mathscr{Q}$ per unit area in V. The analogue of the estimate (40) is now

$$F[\mathscr{Q}; B(x_1, y)] \approx \sum_{0 \neq m \in \mathbf{Z}} \left(\frac{e(mx_1) - e(-mx_1)}{2\pi i m} \right) \left(\sum_{0 \leq n < y} e(\theta n m) \right).$$

Squaring this and integrating with respect to the variable x_1 over $[0,1]$, we have

$$\int_0^1 |F[\mathscr{Q}; B(x_1, y)]|^2 \, \mathrm{d}x_1 \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leq n < y} e(\theta n m) \right|^2. \tag{41}$$

To estimate the sum on the right hand side of (41), we need to make some assumptions on the arithmetic properties of the number θ. We shall assume that θ is a badly approximable number, so that there is a constant $c = c(\theta)$, depending only on θ, such that the inequality

$$m \|m\theta\| > c > 0 \tag{42}$$

holds for every natural number $m \in \mathbf{N}$, where $\|z\|$ denotes the distance of z to the nearest integer.

Lemma 2. *Suppose that the real number θ is badly approximable. Then*

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leq n < y} e(\theta n m) \right|^2 \ll_\theta \log(2y). \tag{43}$$

Proof. It is well known that

$$\left| \sum_{0 \leq n < y} e(\theta n m) \right| \ll \min\{y, \|m\theta\|^{-1}\},$$

so that

$$S = \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \le n < y} e(\theta nm) \right|^2 \ll \sum_{h=1}^{\infty} 2^{-2h} \sum_{2^{h-1} \le m < 2^h} \min\{y^2, \|m\theta\|^{-2}\}.$$

The condition (42) implies that if $2^{h-1} \le m < 2^h$, then the inequality

$$\|m\theta\| > c2^{-h}$$

holds. On the other hand, for any pair $h, p \in \mathbf{N}$, there are at most two values of m satisfying $2^{h-1} \le m < 2^h$ and

$$pc2^{-h} \le \|m\theta\| < (p+1)c2^{-h},$$

for otherwise the difference $(m_1 - m_2)$ of two of them would contradict (42). It follows that

$$\begin{aligned} S &\ll_\theta \sum_{h=1}^{\infty} \sum_{p=1}^{\infty} \min\{2^{-2h}y^2, p^{-2}\} \\ &= \sum_{2^h \le y} \sum_{p=1}^{\infty} \min\{2^{-2h}y^2, p^{-2}\} + \sum_{2^h > y} \sum_{p=1}^{\infty} \min\{2^{-2h}y^2, p^{-2}\} \\ &\ll \sum_{2^h \le y} \sum_{p=1}^{\infty} p^{-2} + \sum_{2^h > y} \left(2^{-2h}y^2 2^h y^{-1} + \sum_{p > 2^h y^{-1}} p^{-2} \right) \\ &\ll \sum_{2^h \le y} 1 + \sum_{2^h > y} 2^{-h} y \ll \log(2y). \end{aligned}$$

This completes the proof. □

Combining (41) and (43) and then integrating with respect to the variable y over $[0, M]$, we have

$$\int_0^M \int_0^1 |F[\mathscr{Q}; B(x_1, y)]|^2 \, \mathrm{d}x_1 \, \mathrm{d}y \ll_\theta M \log(2M).$$

Rescaling in the vertical direction by a factor M^{-1}, we see that the set

$$\mathscr{P} = \{(\{\pm\theta n\}, nM^{-1}) : n = 0, 1, 2, \ldots, M-1\}$$

of $N = 2M$ points in $[0,1)^2$ satisfies the conclusion of Theorem 7 for $k = 2$.

Finally, if N is a given odd number, then we can repeat the argument above with $2M = N + 1$ points. Removing one of the points causes an error of at most 1.

9 Roth's Probabilistic Technique – A Preview

In this section, we describe an ingenious variation of Davenport's argument by Roth [34]. This is a nice preview of his powerful probabilistic technique, which we shall describe in Section 11, and which has been generalized in many different ways and applied in many different situations by many other colleagues.

Let us return to the lattice Λ_1 on $\mathbf{R}^2$ generated by the two vectors $(1,0)$ and $(\theta,1)$. For any real number $t \in \mathbf{R}$, we consider the translated lattice

$$t(1,0)+\Lambda_1 = \{t(1,0)+\mathbf{v} : \mathbf{v} \in \Lambda_1\}.$$

In particular, we are interested in the set

$$\mathscr{Q}_1(t) = (t(1,0)+\Lambda_1) \cap V = \{(\{t+\theta n\}, n) : n = 0,1,2,\ldots,M-1\}$$

which clearly contains precisely M points. Thus analogous to the discrepancy (39), we now consider the discrepancy

$$E[\mathscr{Q}_1(t); B(x_1,y)] = \#(\mathscr{Q}_1(t) \cap B(x_1,y)) - x_1 y,$$

noting that there is now an average of one point of $\mathscr{Q}_1(t)$ per unit area in V. The analogue of the estimate (40) is now

$$E[\mathscr{Q}_1(t); B(x_1,y)] \approx \sum_{0 \neq m \in \mathbf{Z}} \left(\frac{1-e(-mx_1)}{2\pi \mathrm{i} m}\right) \left(\sum_{0 \le n < y} e(\theta nm)\right) e(tm).$$

Squaring this and integrating with respect to the new variable t over $[0,1]$, we have

$$\int_0^1 |E[\mathscr{Q}_1(t); B(x_1,y)]|^2 \, \mathrm{d}t \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left|\sum_{0 \le n < y} e(\theta nm)\right|^2. \tag{44}$$

Furthermore, if θ is a badly approximable number as in the last section, then integrating (44) trivially with respect to the variable x_1 over $[0,1]$ and with respect to the variable y over $[0,M]$, we have

$$\int_0^1 \int_0^M \int_0^1 |E[\mathscr{Q}_1(t); B(x_1,y)]|^2 \, \mathrm{d}x_1 \, \mathrm{d}y \, \mathrm{d}t \ll_\theta M \log(2M).$$

It follows that there exists $t^* \in [0,1]$ such that the set $\mathscr{Q}_1(t^*)$ satisfies

$$\int_0^M \int_0^1 |E[\mathscr{Q}_1(t^*); B(x_1,y)]|^2 \, \mathrm{d}x_1 \, \mathrm{d}y \ll_\theta M \log(2M).$$

Rescaling in the vertical direction by a factor M^{-1}, we see that the set

$$\mathscr{P}(t^*) = \{(\{t^*+\theta n\}, nM^{-1}) : n = 0,1,2,\ldots,M-1\}$$

of $N = M$ points in $[0,1)^2$ satisfies the requirements of Theorem 7 for $k = 2$.

10 Van der Corput Point Sets

In this section, we begin our discussion of those point sets which have been explored in great depth through our study of Theorems 7 and 8.

Our first step is to construct the simplest point sets which will allow us to establish Theorem 8 in the case $k = 2$.

The construction is based on the famous van der Corput sequence $c_0, c_1, c_2, \ldots$ defined as follows. For every non-negative integer $n \in \mathbf{N}_0$, we write

$$n = \sum_{j=1}^{\infty} a_j 2^{j-1} \tag{45}$$

as a dyadic expansion. Then we write

$$c_n = \sum_{j=1}^{\infty} a_j 2^{-j}. \tag{46}$$

Note that $c_n \in [0,1)$. Note also that only finitely many digits $a_1, a_2, a_3, \ldots$ are non-zero, so that the sums in (45) and (46) have only finitely many non-zero terms. For simplicity, we sometimes write

$$n = \ldots a_3 a_2 a_1 \quad \text{and} \quad c_n = 0.a_1 a_2 a_3 \ldots$$

in terms of the digits $a_1, a_2, a_3, \ldots$ of n. The infinite set

$$\mathscr{Q} = \{(c_n, n) : n = 0, 1, 2, \ldots\} \tag{47}$$

in $[0,1) \times [0,\infty)$ is known as the van der Corput point set.

The following is the most crucial property of the van der Corput point set.

Lemma 3. *For all non-negative integers s and ℓ such that $\ell < 2^s$ holds, the set*

$$\{n \in \mathbf{N}_0 : c_n \in [\ell 2^{-s}, (\ell+1)2^{-s})\}$$

contains precisely all the elements of a residue class modulo 2^s in $\mathbf{N}_0$.

Proof. There exist unique integers $b_1, b_2, b_3, \ldots$ such that $\ell 2^{-s} = 0.b_1 b_2 b_3 \ldots b_s$. Clearly $c_n = 0.a_1 a_2 a_3 \ldots \in [\ell 2^{-s}, (\ell+1)2^{-s})$ precisely when $0.a_1 a_2 a_3 \ldots a_s = \ell 2^{-s}$; in other words, precisely when $a_j = b_j$ for every $j = 1, \ldots, s$. The value of a_j for any $j > s$ is irrelevant. $\square$

We say that an interval of the form $[\ell 2^{-s}, (\ell+1)2^{-s}) \subseteq [0,1)$ for some integer ℓ is an elementary dyadic interval of length 2^{-s}. Hence Lemma 3 says that the van der

Corput sequence has very good distribution among such elementary dyadic intervals for all non-negative integer values of s.

Lemma 4. *For all non-negative integers s, ℓ and m such that $\ell < 2^s$ holds, the rectangle*

$$[\ell 2^{-s}, (\ell+1)2^{-s}) \times [m2^s, (m+1)2^s)$$

contains precisely one point of the van der Corput point set $\mathscr{Q}$.

It is clear that there is an average of one point of the van der Corput point set $\mathscr{Q}$ per unit area in $[0,1)\times[0,\infty)$. For any measurable set A in $[0,1)\times[0,\infty)$, let

$$E[\mathscr{Q};A] = \#(\mathscr{Q}\cap A) - \mu(A)$$

denote the discrepancy of $\mathscr{Q}$ in A.

Let $\psi(z) = z - [z] - \frac{1}{2}$ for every $z \in \mathbf{R}$.

Lemma 5. *For all non-negative integers s and ℓ such that $\ell < 2^s$ holds, there exist real numbers α_0, β_0, depending at most on s and ℓ, such that $|\alpha_0| \le \frac{1}{2}$ and*

$$E[\mathscr{Q};[\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,y)] = \alpha_0 - \psi(2^{-s}(y-\beta_0)) \tag{48}$$

at all points of continuity of the right hand side.

Proof. In view of Lemma 3, the second coordinates of the points of $\mathscr{Q}$ in the region $[\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,\infty)$ fall precisely into a residue class modulo 2^s. Let n_0 be the smallest of these second coordinates. Then $0 \le n_0 < 2^s$. We now study

$$E[\mathscr{Q};[\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,y)]$$

as a function of y. For simplicity, denote it by $f(y)$, say. Clearly $f(0) = E[\mathscr{Q};\emptyset] = 0$. On the other hand, note that

$$\mu([\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,y)) = 2^{-s}y$$

increases with y at the rate 2^{-s}, so that $f(y)$ decreases with y at the rate 2^{-s}, except when y coincides with the second coordinate of one of the points of the set $\mathscr{Q}$ in the region $[\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,\infty)$, in which case $f(y)$ jumps up by 1. The first instance of this jump occurs when $y = n_0$. See Figure 3.

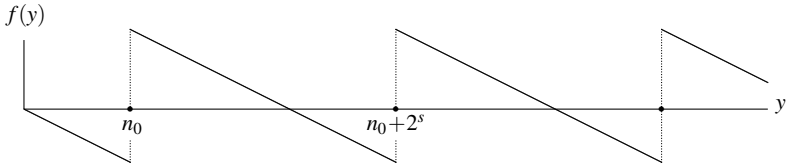

Fig. 3 The sawtooth function $E[\mathscr{Q};[\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,y)]$

With suitable α_0 and β_0, the right hand side of (48) fits all the requirements. □

We can now prove Theorem 8 for $k = 2$. Let $N \geq 2$ be a given integer. It follows immediately from the definition of $\mathscr{Q}$ that the set

$$\mathscr{Q}_0 = \mathscr{Q} \cap ([0,1) \times [0,N))$$

contains precisely N points. Let the integer h be determined uniquely by

$$2^{h-1} < N \leq 2^h. \tag{49}$$

Consider a rectangle of the form

$$B(x_1,y) = [0,x_1) \times [0,y) \subseteq [0,1) \times [0,N).$$

Let $x_1^{(0)} = 0$. For every $s = 1,\ldots,h$, let $x_1^{(s)} = 2^{-s}[2^s x_1]$ denote the greatest integer multiple of 2^{-s} not exceeding x_1. Then we can write $[0,x_1)$ as a union of disjoint intervals in the form

$$[0,x_1) = [x_1^{(h)},x_1) \cup \bigcup_{s=1}^{h} [x_1^{(s-1)},x_1^{(s)}).$$

It follows that

$$\begin{aligned} E[\mathscr{Q}_0;[0,x_1) \times [0,y)] &= E[\mathscr{Q};[0,x_1) \times [0,y)] \\ &= E[\mathscr{Q};[x_1^{(h)},x_1) \times [0,y)] + \sum_{s=1}^{h} E[\mathscr{Q};[x_1^{(s-1)},x_1^{(s)}) \times [0,y)]. \end{aligned} \tag{50}$$

Clearly $[x_1^{(h)},x_1) \times [0,y) \subseteq [x_1^{(h)},x_1^{(h)} + 2^{-h}) \times [0,2^h)$, and the latter rectangle has area 1 and is of the type under discussion in Lemma 4, hence contains precisely one point of $\mathscr{Q}$. It follows that

$$\#(\mathscr{Q} \cap ([x_1^{(h)},x_1) \times [0,y)) \leq 1 \quad \text{and} \quad \mu([x_1^{(h)},x_1) \times [0,y)) \leq 1,$$

and we have the bound

$$|E[\mathscr{Q};[x_1^{(h)},x_1) \times [0,y)]| \leq 1. \tag{51}$$

On the other hand, for every $s = 1,\ldots,k$, the rectangle

$$[x_1^{(s-1)},x_1^{(s)}) \times [0,y)$$

either is empty, in which case we have $E[\mathscr{Q};[x_1^{(s-1)},x_1^{(s)}) \times [0,y)] = 0$ trivially, or is of the type under discussion in Lemma 5, and we have the bound

$$|E[\mathscr{Q};[x_1^{(s-1)},x_1^{(s)}) \times [0,y)]| \leq 1. \tag{52}$$

Note that (52) still holds in the empty case. Combining (49)–(52), we arrive at an upper bound

$$|E[\mathscr{Q}_0;[0,x_1)\times[0,y)]| \le 1+h \ll \log N. \tag{53}$$

For comparison later in Section 14, let us summarize what we have done. We are approximating the interval $[0,x_1)$ by a subinterval $[0,x_1^{(h)})$, and consequently approximating the rectangle $B(x_1,y)$ by a smaller rectangle $B(x_1^{(h)},y)$. Then we show that the difference $B(x_1,y)\setminus B(x_1^{(h)},y)$ is contained in one of the rectangles under discussion in Lemma 4, and inequality (51) is the observation that

$$|E[\mathscr{Q};B(x_1,y)]-E[\mathscr{Q};B(x_1^{(h)},y]| \le 1.$$

To estimate $E[\mathscr{Q};B(x_1^{(h)},y)]$, we note that the interval $[0,x_1^{(h)})$ is a union of at most h disjoint elementary dyadic intervals. More precisely, if we write

$$x_1^{(h)} = \sum_{s=1}^{h} b_s 2^{-s}$$

as a dyadic expansion, then $[0,x_i^{(h)})$ can be written as a union of

$$\sum_{s=1}^{h} b_s \le h$$

elementary dyadic intervals, namely b_1 elementary dyadic intervals of length 2^{-1}, together with b_2 elementary dyadic intervals of length 2^{-2}, and so on. It follows that $B(x_1^{(h)},y)$ is a disjoint union of at most h rectangles discussed in Lemma 5, each of which satisfies inequality (52).

Finally, rescaling the second coordinate of the points of $\mathscr{Q}_0$ by a factor N^{-1}, we obtain a set

$$\mathscr{P} = \{(c_n,N^{-1}n) : n = 0,1,2,\ldots,N-1\} \tag{54}$$

of precisely N points in $[0,1)^2$. For every $\mathbf{x}=(x_1,x_2)\in[0,1]^2$, we have

$$D[\mathscr{P};B(\mathbf{x})] = E[\mathscr{Q}_0;[0,x_1)\times[0,Nx_2)] \ll \log N,$$

in view of (53) and noting that $0 \le Nx_2 \le N$. This now completes the proof of Theorem 8 for $k=2$.

11 Roth's Probabilistic Technique

We now attempt to extend the ideas in the last section to obtain a proof of Theorem 7 for $k=2$.

Let us first of all consider the special case when $N = 2^h$. Then the set (54) used to establish Theorem 8 for $k = 2$ becomes

$$\begin{aligned}\mathscr{P}(2^h) &= \{(c_n, 2^{-h}n) : n = 0,1,2,\dots,2^h-1\} \\ &= \{(0.a_1a_2a_3\dots a_h, 0.a_h\dots a_3a_2a_1) : a_1,\dots,a_h \in \{0,1\}\}, \end{aligned} \tag{55}$$

in terms of binary digits. We have the following unhelpful result[3] of Halton and Zaremba [26].

Proposition 3. *For every positive integer h, we have*

$$\int_{[0,1]^2} |D[\mathscr{P}(2^h); B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = 2^{-6}h^2 + O(h). \tag{56}$$

Clearly the order of magnitude is $(\log N)^2$, and not $\log N$ as we would have liked. Hence any unmodified van der Corput point set is not sufficient to establish our desired result. To understand the problem, we return to our discussion in the last section. Assume that $N = 2^h$. Consider a rectangle of the form

$$B(x_1, y) = [0, x_1) \times [0, y) \subseteq [0,1) \times [0, 2^h).$$

For simplicity, let us assume that x_1 is an integer multiple of 2^{-h}, so that $x_1 = x_1^{(h)}$ and (50) simplifies to

$$D[\mathscr{P}; B(x_1, 2^{-h}y)] = E[\mathscr{Q}_0; [0,x_1) \times [0,y)] = \sum_{s=1}^{h\,*} E[\mathscr{Q}; [x_1^{(s-1)}, x_1^{(s)}) \times [0,y)],$$

where the $*$ in the summation sign denotes that the sum includes only those terms where $x_1^{(s-1)} \neq x_1^{(s)}$. Note that when $x_1^{(s-1)} \neq x_1^{(s)}$, we have

$$[x_1^{(s-1)}, x_1^{(s)}) = [\ell 2^{-s}, (\ell+1)2^{-s})$$

for some integer ℓ, so it follows from Lemma 5 that

$$D[\mathscr{P}; B(x_1, 2^{-h}y)] = \sum_{s=1}^{h\,*} (\alpha_s - \psi(2^{-s}(y - \beta_s))), \tag{57}$$

where, for each $s = 1,\dots,h$, the real numbers α_s and β_s satisfy $|\alpha_s| \leq \frac{1}{2}$. If we square the expression (57), then the right hand side becomes

$$\sum_{s'=1}^{h\,*} \sum_{s''=1}^{h\,*} (\alpha_{s'} - \psi(2^{-s'}(y - \beta_{s'})))(\alpha_{s''} - \psi(2^{-s''}(y - \beta_{s''}))).$$

Expanding the summand, this gives rise eventually to a constant term

[3] In their paper, Halton and Zaremba have an exact expression for the integral under study.

$$\sum_{s'=1}^{h}{}^{*}\sum_{s''=1}^{h}{}^{*}\alpha_{s'}\alpha_{s''}$$

which ultimately leads to the term $2^{-6}h^2$ in (56).

Note that this constant term arises from our assumption that all the aligned rectangles under consideration are anchored at the origin, and recall that Roth's attempt to overcome this handicap, discussed in Section 9, involves the introduction of a translation variable t. So let us attempt to describe Roth's incorporation of this idea of a translation variable into the argument here.

To pave the way for a smooth introduction of a probabilistic variable, we shall modify the van der Corput point set somewhat. Let $N \geq 2$ be a given integer, and let the integer h be determined uniquely by

$$2^{h-1} < N \leq 2^h. \tag{58}$$

For every $n = 0, 1, 2, \ldots, 2^h - 1$, we define c_n as before by (45) and (46). We then extend the definition of c_n to all other integers using periodicity by writing

$$c_{n+2^h} = c_n \quad \text{for every } n \in \mathbf{Z},$$

and consider the extended van der Corput point set

$$\mathscr{Q}_h = \{(c_n, n) : n \in \mathbf{Z}\}.$$

Furthermore, for every real number $t \in \mathbf{R}$, we consider the translated van der Corput point set

$$\mathscr{Q}_h(t) = \{(c_n, n+t) : n \in \mathbf{Z}\}.$$

It is clear that there is an average of one point of the translated van der Corput point set $\mathscr{Q}_h(t)$ per unit area in $[0,1) \times (-\infty, \infty)$. For any measurable set A in $[0,1) \times (-\infty, \infty)$, we now let

$$E[\mathscr{Q}_h(t); A] = \#(\mathscr{Q}_h(t) \cap A) - \mu(A)$$

denote the discrepancy of $\mathscr{Q}_h(t)$ in A.

Consider a rectangle of the form

$$B(x_1, y) = [0, x_1) \times [0, y) \subseteq [0,1) \times [0, N).$$

As before, let $x_1^{(0)} = 0$. For every $s = 1, \ldots, h$, let $x_1^{(s)} = 2^{-s}[2^s x_1]$ denote the greatest integer multiple of 2^{-s} not exceeding x_1. Then, analogous to (51), we have the trivial bound

$$|E[\mathscr{Q}_h(t); [x_1^{(h)}, x_1) \times [0, y)]| \leq 1, \tag{59}$$

so we shall henceforth assume that $x_1 = x_1^{(h)}$, so that

$$E[\mathscr{Q}_h(t);B(x_1,y)] = \sum_{s=1}^{h}{}^{*} E[\mathscr{Q}_h(t);[x_1^{(s-1)},x_1^{(s)})\times[0,y)]. \tag{60}$$

Corresponding to Lemma 5, we can establish the following result without too much difficulty.

Lemma 6. *For all positive real numbers y and all non-negative integers s and ℓ such that $s \le h$ and $\ell < 2^s$ hold, there exist real numbers β_0 and γ_0, depending at most on s, ℓ and y, such that*

$$E[\mathscr{Q}_h(t);[\ell 2^{-s},(\ell+1)2^{-s})\times[0,y)] = \psi(2^{-s}(t-\beta_0)) - \psi(2^{-s}(t-\gamma_0))$$

at all points of continuity of the right hand side.

Combining (60) and Lemma 6, we have

$$E[\mathscr{Q}_h(t);B(x_1,y)] = \sum_{s=1}^{h}{}^{*} (\psi(2^{-s}(t-\beta_s)) - \psi(2^{-s}(t-\gamma_s))) \tag{61}$$

for some real numbers β_s and γ_s depending at most on x_1 and y. We shall square this expression and integrate with respect to the translation variable t over the interval $[0,2^h)$, an interval of length equal to the period of the set $\mathscr{Q}_h(t)$. We therefore need to study integrals of the form

$$\int_0^{2^h} \psi(2^{-s'}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t,$$

or when either or both of $\beta_{s'}$ and $\beta_{s''}$ are replaced by $\gamma_{s'}$ and $\gamma_{s''}$ respectively.

Lemma 7. *Suppose that the integers s' and s'' satisfy $0 \le s', s'' \le h$, and that the real numbers $\beta_{s'}$ and $\beta_{s''}$ are fixed. Then*

$$\int_0^{2^h} \psi(2^{-s'}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t = O(2^{h-|s'-s''|}).$$

Proof. The result is obvious if $s' = s''$. Without loss of generality, let us assume that $s' > s''$. For every $a = 0,1,2,\ldots,2^{s'-s''}-1$, in view of periodicity, we have

$$\begin{aligned}
&\int_0^{2^h} \psi(2^{-s'}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t \\
&\quad = \int_0^{2^h} \psi(2^{-s'}(t+a2^{s''}-\beta_{s'}))\psi(2^{-s''}(t+a2^{s''}-\beta_{s''}))\,\mathrm{d}t \\
&\quad = \int_0^{2^h} \psi(2^{-s'}(t+a2^{s''}-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t,
\end{aligned}$$

with the last equality arising from the observation that

$$\psi(2^{-s''}(t+a2^{s''}-\beta_{s''}))=\psi(a+2^{-s''}(t-\beta_{s''}))=\psi(2^{-s''}(t-\beta_{s''})).$$

It follows that

$$\begin{aligned}
&2^{s'-s''}\int_0^{2^h}\psi(2^{-s'}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t\\
&\quad=\sum_{a=0}^{2^{s'-s''}-1}\int_0^{2^h}\psi(2^{-s'}(t+a2^{s''}-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t\\
&\quad=\int_0^{2^h}\left(\sum_{a=0}^{2^{s'-s''}-1}\psi(2^{-s'}(t+a2^{s''}-\beta_{s'}))\right)\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t.
\end{aligned}$$

It is not difficult to see that

$$\sum_{a=0}^{2^{s'-s''}-1}\psi(2^{-s'}(t+a2^{s''}-\beta_{s'}))=\psi(2^{-s''}(t-\beta_{s'})) \tag{62}$$

at all points of continuity, as shown in Figure 4.

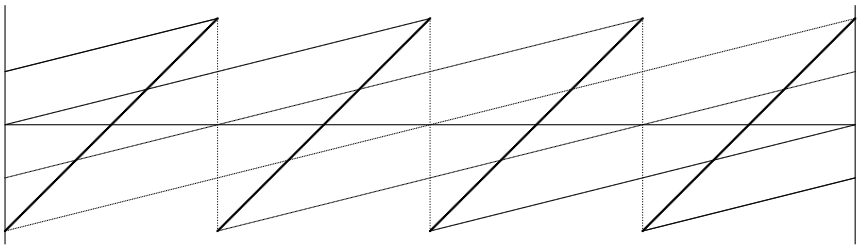

Fig. 4 An illustration of the summation (62)

We therefore conclude that

$$\begin{aligned}
&2^{s'-s''}\int_0^{2^h}\psi(2^{-s'}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t\\
&\quad=\int_0^{2^h}\psi(2^{-s''}(t-\beta_{s'}))\psi(2^{-s''}(t-\beta_{s''}))\,\mathrm{d}t=O(2^h),
\end{aligned}$$

and the desired result follows immediately. □

It now follows from (61) and Lemma 7 that

$$\int_0^{2^h}|E[\mathscr{Q}_h(t);B(x_1,y)]|^2\,\mathrm{d}t\ll\sum_{s'=1}^{h}{}^{*}\sum_{s''=1}^{h}{}^{*}2^{h-|s'-s''|}\ll 2^h h, \tag{63}$$

noting that the diagonal terms contribute $O(2^h h)$, and the contribution from the off-diagonal terms decays geometrically.

Note that the estimate (63) is independent of the choice of x_1 and y. We also recall the trivial estimate (59). It follows that integrating (63) trivially with respect to x_1 over the interval $[0,1)$ and with respect to y over the interval $[0,N)$, we conclude that

$$\begin{aligned}&\int_0^N \int_0^1 \int_0^{2^h} |E[\mathscr{Q}_h(t);B(x_1,y)]|^2 \,\mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}y\\ &\quad = \int_0^{2^h} \left(\int_0^N \int_0^1 |E[\mathscr{Q}_h(t);B(x_1,y)]|^2 \,\mathrm{d}x_1\,\mathrm{d}y \right) \mathrm{d}t \ll 2^h hN.\end{aligned}$$

Hence there exists $t^* \in [0,2^h)$ such that

$$\int_0^N \int_0^1 |E[\mathscr{Q}_h(t^*);B(x_1,y)]|^2 \,\mathrm{d}x_1\,\mathrm{d}y \ll hN. \tag{64}$$

Finally, we note that the set $\mathscr{Q}_h(t^*) \cap ([0,1) \times [0,N))$ contains precisely N points. Rescaling in the vertical direction by a factor N^{-1}, we observe that the set

$$\mathscr{P}^* = \{(z_1, N^{-1} z_2) : (z_1,z_2) \in \mathscr{Q}_h(t^*)\}$$

contains precisely N points in $[0,1)^2$, and the estimate (64) now translates to

$$\int_{[0,1]^2} |D[\mathscr{P}^*;B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \ll h \ll \log N,$$

in view of (58). This completes the proof of Theorem 7 for $k=2$.

We conclude this section by trying to obtain a different interpretation of the effect of the translation variable t. Consider a typical term

$$E[\mathscr{Q}_h(t); [x_1^{(s-1)}, x_1^{(s)}) \times [0,y)]$$

in the sum (60). If $x_1^{(s-1)} \neq x_1^{(s)}$, then $x_1^{(s)}$ cannot be an integer multiple of $2^{-(s-1)}$ and therefore must be an odd integer multiple of 2^{-s}, and so

$$[x_1^{(s-1)}, x_1^{(s)}) = [\ell 2^{-s}, (\ell+1)2^{-s}) \subset \left[\frac{\ell}{2} 2^{-(s-1)}, (\frac{\ell}{2}+1) 2^{-(s-1)} \right)$$

for some even integer ℓ. One can then show that

$$E[\mathscr{Q}_h(2^{s-1}); [\ell 2^{-s}, (\ell+1)2^{-s}) \times [0,y)] = E[\mathscr{Q}_h; [(\ell+1)2^{-s}, (\ell+2)2^{-s}) \times [0,y)].$$

This means that instead of translating vertically, as on the left hand side above, one may shift horizontally, as on the right hand side above. Another way to see this is to note from Lemma 3 that the interval $[\ell 2^{-s}, (\ell+1)2^{-s})$ is associated with a residue class R_s modulo 2^s, whereas the interval $[\ell 2^{-s}, (\ell+2)2^{-s})$ is associated

with a residue class R_{s-1} modulo 2^{s-1}, so the interval $[(\ell+1)2^{-s},(\ell+2)2^{-s})$ must be associated with the residue class $R_{s-1}\setminus R_s$ modulo 2^s. But then $R_{s-1}\setminus R_s$ is clearly R_s translated by 2^{s-1}.

12 Digit Shifts

In this section, we shall attempt to replace the vertical translation studied in the last section by horizontal shifts, as pioneered by Chen [11].

Let $N \geq 2$ be a given integer, and let the integer h be determined uniquely by

$$2^{h-1} < N \leq 2^h. \tag{65}$$

For every $n = 0,1,2,\ldots,2^h-1$, we define c_n as before by (45) and (46). As we are not translating vertically, there is no need[4] to extend the definition of c_n to other integers as in the last section, and we consider the set[5]

$$\begin{aligned}\mathscr{Q}_h &= \{(c_n,n) : n = 0,1,2,\ldots,2^h-1\}\\ &= \{(0.a_1a_2a_3\ldots a_h, a_h\ldots a_3a_2a_1) : a_1,\ldots,a_h \in \{0,1\}\},\end{aligned}$$

in terms of binary digits. Furthermore, for every $\mathbf{t} = (t_1,\ldots,t_h) \in \mathbf{Z}_2^h$, where $\mathbf{Z}_2 = \{0,1\}$, write

$$c_n^{(\mathbf{t})} = 0.(a_1\oplus t_1)(a_2\oplus t_2)(a_3\oplus t_3)\ldots(a_h\oplus t_h) \quad \text{if} \quad c_n = 0.a_1a_2a_3\ldots a_h$$

in binary notation, where $\oplus$ denotes addition modulo 2, and consider the shifted van der Corput point set

$$\mathscr{Q}_h^{(\mathbf{t})} = \{(c_n^{(\mathbf{t})},n) : n = 0,1,2,\ldots,2^h-1\},$$

obtained from $\mathscr{Q}_h$ by a digit shift $\mathbf{t}$.

It is clear that there is an average of one point of the shifted van der Corput point set $\mathscr{Q}_h^{(\mathbf{t})}$ per unit area in $[0,1)\times[0,2^h)$. For any measurable set A in $[0,1)\times[0,2^h)$, we study the discrepancy function

$$E[\mathscr{Q}_h^{(\mathbf{t})};A] = \#(\mathscr{Q}_h^{(\mathbf{t})}\cap A) - \mu(A).$$

Consider a rectangle of the form

$$B(x_1,y) = [0,x_1)\times[0,y) \subseteq [0,1)\times[0,N).$$

[4] This is not the case if we wish to study Theorem 7 for $k > 2$.

[5] Note that the set $\mathscr{Q}_h$ here is different from that in the last section. However, since we are working with rectangles inside $[0,1)\times[0,2^h)$, our statements here concerning $\mathscr{Q}_h$ remain valid for the set $\mathscr{Q}_h$ defined in the last section.

Analogous to (59), we have the trivial bound

$$|E[\mathscr{Q}_h^{(\mathbf{t})};[x_1^{(h)},x_1)\times[0,y)]|\le 1, \tag{66}$$

so we shall henceforth assume that $x_1 = x_1^{(h)}$, so that

$$E[\mathscr{Q}_h^{(\mathbf{t})};B(x_1,y)] = \sum_{s=1}^{h}{}^{*} E[\mathscr{Q}_h^{(\mathbf{t})};[x_1^{(s-1)},x_1^{(s)})\times[0,y)]. \tag{67}$$

We now square this expression and sum it over all digit shifts $\mathbf{t}\in\mathbf{Z}_2^h$. For simplicity and convenience, let us omit reference to $\mathscr{Q}_h$ and y, and write

$$E[\mathscr{Q}_h^{(\mathbf{t})};[x_1^{(s-1)},x_1^{(s)})\times[0,y)] = E_s[t_1,\dots,t_h].$$

Then we need to study sums of the form

$$\sum_{\mathbf{t}\in\mathbf{Z}_2^h} E_{s'}[t_1,\dots,t_h]E_{s''}[t_1,\dots,t_h].$$

Analogous to Lemma 7, we have the following estimate.

Lemma 8. *Suppose that the real number $y\in[0,N)$ is fixed, and that the integers s' and s'' satisfy $0\le s',s''\le h$. Then*

$$\sum_{\mathbf{t}\in\mathbf{Z}_2^h} E_{s'}[t_1,\dots,t_h]E_{s''}[t_1,\dots,t_h] = O(2^{h-|s'-s''|}). \tag{68}$$

Proof. First of all, for fixed $t_1,\dots,t_s$, the value of $E_s[t_1,\dots,t_h]$ remains the same for every choice of $t_{s+1},\dots,t_h$, as these latter variables only shift the digits of c_n after the s-th digit, and so

$$c_n^{(t_1,\dots,t_s,t_{s+1},\dots,t_h)}\in[x_1^{(s-1)},x_1^{(s)}) \quad\text{if and only if}\quad c_n^{(t_1,\dots,t_s,0,\dots,0)}\in[x_1^{(s-1)},x_1^{(s)}).$$

Next, the case when $x_1^{(s'-1)} = x_1^{(s')}$ or $x_1^{(s''-1)} = x_1^{(s'')}$ is also trivial, as the summand is clearly equal to zero, so we shall assume that $x_1^{(s'-1)}\ne x_1^{(s')}$ and $x_1^{(s''-1)}\ne x_1^{(s'')}$. Now the case when $s' = s''$ is easy, since we have $E[t_1,\dots,t_h;x_1^{(s-1)},x_1^{(s)}] = O(1)$ trivially. Without loss of generality, let us assume that $s' > s''$. For fixed $t_1,\dots,t_{s''}$, in view of the comment at the beginning of the proof, we have

$$\begin{aligned}&\sum_{t_{s''+1},\dots,t_h\in\mathbf{Z}_2} E_{s'}[t_1,\dots,t_h]E_{s''}[t_1,\dots,t_h]\\ &\quad= 2^{h-s'}\left(\sum_{t_{s''+1},\dots,t_{s'}\in\mathbf{Z}_2} E_{s'}[t_1,\dots,t_{s'},0,\dots,0]\right)E_{s''}[t_1,\dots,t_{s''},0,\dots,0].\end{aligned}$$

We shall show that

$$\begin{aligned}
&\sum_{t_{s''+1},\ldots,t_{s'}\in\mathbf{Z}_2} E_{s'}[t_1,\ldots,t_{s'},0,\ldots,0]\\
&\quad=\sum_{t_{s''+1},\ldots,t_{s'}\in\mathbf{Z}_2} E[\mathscr{Q}_h^{(t_1,\ldots,t_{s''},t_{s''+1},\ldots,t_{s'},0,\ldots,0)};[x_1^{(s'-1)},x_1^{(s')})\times[0,y)]\\
&\quad=E[\mathscr{Q}_h^{(t_1,\ldots,t_{s''},0,\ldots,0)};[\ell 2^{-s''},(\ell+1)2^{-s''})\times[0,y)],
\end{aligned}\tag{69}$$

where ℓ is an integer and $[x_1^{(s'-1)},x_1^{(s')})\subset[\ell 2^{-s''},(\ell+1)2^{-s''})$. Then

$$\sum_{t_{s''+1},\ldots,t_h\in\mathbf{Z}_2} E_{s'}[t_1,\ldots,t_h]E_{s''}[t_1,\ldots,t_h]=O(2^{h-s'}),$$

from which it follows that

$$\sum_{t_1,\ldots,t_h\in\mathbf{Z}_2} E_{s'}[t_1,\ldots,t_h]E_{s''}[t_1,\ldots,t_h]=O(2^{h-s'+s''}),$$

giving the desired result. To establish (69), simply note that for fixed $t_1,\ldots,t_{s''}$, if a point

$$c_n^{(t_1,\ldots,t_{s''},0,\ldots,0)}\in[x_1^{(s'-1)},x_1^{(s')}),$$

then each distinct choice of $t_{s''+1},\ldots,t_{s'}$ will shift this point into one of the $2^{s'-s''}$ distinct intervals of length $2^{-s'}$ that make up the interval $[\ell 2^{-s''},(\ell+1)2^{-s''})$. $\square$

It now follows from (67) and Lemma 8 that

$$\sum_{\mathbf{t}\in\mathbf{Z}_2^h}|E[\mathscr{Q}_h^{(\mathbf{t})};B(x_1,y)]|^2\ll\sum_{s'=1}^{h\,*}\sum_{s''=1}^{h\,*}2^{h-|s'-s''|}\ll 2^h h,\tag{70}$$

noting that the diagonal terms contribute $O(2^h h)$, and the contribution from the off-diagonal terms decays geometrically.

Note that the estimate (70) is independent of the choice of x_1 and y. We also recall the trivial estimate (66). It follows that integrating (70) trivially with respect to x_1 over the interval $[0,1)$ and with respect to y over the interval $[0,N)$, we conclude that

$$\begin{aligned}
&\int_0^N\int_0^1\sum_{\mathbf{t}\in\mathbf{Z}_2^h}|E[\mathscr{Q}_h^{(\mathbf{t})};B(x_1,y)]|^2\,\mathrm{d}x_1\,\mathrm{d}y\\
&\quad=\sum_{\mathbf{t}\in\mathbf{Z}_2^h}\int_0^N\int_0^1|E[\mathscr{Q}_h^{(\mathbf{t})};B(x_1,y)]|^2\,\mathrm{d}x_1\,\mathrm{d}y\ll 2^h hN.
\end{aligned}$$

Hence there exists $\mathbf{t}^*\in\mathbf{Z}_2^h$ such that

$$\int_0^N\int_0^1|E[\mathscr{Q}_h^{(\mathbf{t}^*)};B(x_1,y)]|^2\,\mathrm{d}x_1\,\mathrm{d}y\ll hN.\tag{71}$$

Finally, we note that the set $\mathscr{Q}_h^{(\mathbf{t}^*)} \cap ([0,1) \times [0,N))$ contains precisely N points. Rescaling in the vertical direction by a factor N^{-1}, we observe that the set

$$\mathscr{P}^* = \{(z_1, N^{-1} z_2) : (z_1, z_2) \in \mathscr{Q}_h^{(\mathbf{t}^*)}\}$$

contains precisely N points in $[0,1)^2$, and the estimate (71) now translates to

$$\int_{[0,1]^2} |D[\mathscr{P}^*; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \ll h \ll \log N,$$

in view of (65). This completes the proof of Theorem 7 for $k = 2$.

13 A Fourier–Walsh Approach to van der Corput Sets

In this section, we sketch yet another proof of Theorem 7 for $k = 2$ by highlighting the interesting group structure of the van der Corput point set

$$\mathscr{P}(2^h) = \{(0.a_1 a_2 a_3 \dots a_h, 0.a_h \dots a_3 a_2 a_1) : a_1, \dots, a_h \in \{0,1\}\}.$$

This is a finite abelian group isomorphic to the group $\mathbf{Z}_2^h$. We shall make use of the characters of these groups. These are the Walsh functions.

To define the Walsh functions, we first consider binary representation of any integer $\ell \in \mathbf{N}_0$, written uniquely in the form

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell) 2^{i-1}, \tag{72}$$

where the coefficient $\lambda_i(\ell) \in \{0,1\}$ for every $i \in \mathbf{N}$. On the other hand, every real number $y \in [0,1)$ can be represented in the form

$$y = \sum_{i=1}^{\infty} \eta_i(y) 2^{-i}, \tag{73}$$

where the coefficient $\eta_i(y) \in \{0,1\}$ for every $i \in \mathbf{N}$. This representation is unique if we agree that the series in (73) is finite for every $y = m2^{-s}$ where $s \in \mathbf{N}_0$ and $m \in \{0,1,\dots,2^s - 1\}$.

For every $\ell \in \mathbf{N}_0$ of the form (72), we define the Walsh function $w_\ell : [0,1) \to \mathbf{R}$ by writing

$$w_\ell(y) = (-1)^{\sum\limits_{i=1}^{\infty} \lambda_i(\ell)\eta_i(y)}. \tag{74}$$

Since (72) is essentially a finite sum, the Walsh function is well defined, and takes the values ± 1. It is easy to see that $w_0(y) = 1$ for every $y \in [0,1)$. It is well known that under the inner product

$$\langle w_k, w_\ell \rangle = \int_0^1 w_k(y) w_\ell(y)\,\mathrm{d}y,$$

the collection of Walsh functions form an orthonormal basis of $L^2[0,1]$.

For every $\ell, k \in \mathbf{N}_0$, we can define $\ell \oplus k$ by setting

$$\lambda_i(\ell \oplus k) = \lambda_i(\ell) + \lambda_i(k) \bmod 2$$

for every $i \in \mathbf{N}$. Then it is easy to see that for every $y \in [0,1)$, we have

$$w_{\ell \oplus k}(y) = w_\ell(y) w_k(y). \tag{75}$$

For every $x, y \in [0,1)$, we can define $x \oplus y$ be setting

$$\eta_i(x \oplus y) = \eta_i(x) + \eta_i(y) \bmod 2$$

for every $i \in \mathbf{N}$. Then it is easy to see that for every $\ell \in \mathbf{N}_0$, we have

$$w_\ell(x \oplus y) = w_\ell(x) w_\ell(y). \tag{76}$$

We shall be concerned with the characteristic function

$$\chi_{B(\mathbf{x})}(\mathbf{y}) = \begin{cases} 1, \text{ if } \mathbf{y} \in B(\mathbf{x}), \\ 0, \text{ otherwise,} \end{cases}$$

of the aligned rectangle $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, where $\mathbf{x} = (x_1, x_2)$. Then we have the discrepancy function

$$D[\mathscr{P}(2^h); B(\mathbf{x})] = \sum_{\mathbf{p} \in \mathscr{P}(2^h)} \chi_{B(\mathbf{x})}(\mathbf{p}) - 2^h x_1 x_2. \tag{77}$$

Clearly the characteristic function in question can be written as a product of one-dimensional characteristic functions in the form

$$\chi_{B(\mathbf{x})}(\mathbf{y}) = \chi_{[0,x_1)}(y_1) \chi_{[0,x_2)}(y_2),$$

where $\mathbf{y} = (y_1, y_2)$. Since the Walsh functions form an orthonormal basis for the space $L^2[0,1]$, we shall use Fourier–Walsh analysis[6] to study a characteristic function of the form $\chi_{[0,x)}(y)$. We have the Fourier–Walsh series

$$\chi_{[0,x)}(y) \sim \sum_{\ell=0}^{\infty} \widetilde{\chi}_\ell(x) w_\ell(y),$$

where, for every $\ell \in \mathbf{N}_0$, the Fourier–Walsh coefficients are given by

[6] Simply imagine that we use Fourier analysis but with the Walsh functions replacing the exponential functions.

$$\widetilde{\chi}_\ell(x) = \int_0^x w_\ell(y)\,\mathrm{d}y.$$

In particular, we have $\widetilde{\chi}_0(x) = x$ for every $x \in [0,1)$.

Instead of using the full Fourier–Walsh series, we shall truncate it and use the approximation

$$\chi^{(h)}_{[0,x)}(y) = \sum_{\ell=0}^{2^h-1} \widetilde{\chi}_\ell(x) w_\ell(y). \tag{78}$$

Note that there exists a unique $m \in \mathbf{N}_0$ such that $m2^{-h} \le x < (m+1)2^{-h}$. Then

$$\chi^{(h)}_{[0,x)}(y) = \begin{cases} 1, & \text{if } 0 \le y < m2^{-h}, \\ 2^h x - m, & \text{if } m2^{-h} \le y < (m+1)2^{-h}, \\ 0, & \text{if } (m+1)2^{-h} \le y < 1, \end{cases}$$

where the quantity

$$2^h x - m = 2^h \int_{m2^{-h}}^{(m+1)2^{-h}} \chi_{[0,x)}(y)\,\mathrm{d}y$$

represents the average value of $\chi_{[0,x)}(y)$ in the interval $[m2^{-h},(m+1)2^{-h})$.

The approximation (78) in turn leads to the approximation

$$\chi^{(h)}_{B(\mathbf{x})}(\mathbf{y}) = \chi^{(h)}_{[0,x_1)}(y_1)\chi^{(h)}_{[0,x_2)}(y_2) = \sum_{\ell_1=0}^{2^h-1}\sum_{\ell_2=0}^{2^h-1} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) W_{\mathbf{l}}(\mathbf{y})$$

of the characteristic function $\chi_{B(\mathbf{x})}(\mathbf{y})$. Here $\mathbf{l} = (\ell_1,\ell_2)$,

$$\widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) = \widetilde{\chi_{\ell_1}}(x_1)\widetilde{\chi_{\ell_2}}(x_2) \quad \text{and} \quad W_{\mathbf{l}}(\mathbf{y}) = w_{\ell_1}(y_1)w_{\ell_2}(y_2). \tag{79}$$

Corresponding to this, we approximate the discrepancy function (77) by

$$\begin{aligned} D^{(h)}[\mathscr{P}(2^h);B(\mathbf{x})] &= \sum_{\mathbf{p}\in\mathscr{P}(2^h)} \chi^{(h)}_{B(\mathbf{x})}(\mathbf{p}) - 2^h x_1 x_2 \\ &= \sum_{\mathbf{p}\in\mathscr{P}(2^h)} \sum_{\ell_1=0}^{2^h-1}\sum_{\ell_2=0}^{2^h-1} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) W_{\mathbf{l}}(\mathbf{p}) - 2^h \widetilde{\chi}_{\mathbf{0}}(\mathbf{x}) \\ &= \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{2^h-1}\sum_{\ell_2=0}^{2^h-1} \left(\sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) \right) \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}), \end{aligned}$$

noting that

$$\sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{0}}(\mathbf{p}) = \#\mathscr{P}(2^h) = 2^h. \tag{80}$$

It is well known in the theory of abelian groups that the sum

$$\sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) \in \{0, 2^h\}; \tag{81}$$

see, for instance, [29, Chapters 5 and 9] or [30, Chapter 5]. We therefore need to have some understanding on the set

$$L(h) = \left\{ \mathbf{l} \in [0,2^h) \times [0,2^h) : \mathbf{l} \neq \mathbf{0} \text{ and } \sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) = 2^h \right\}.$$

Then

$$D^{(h)}[\mathscr{P}(2^h); B(\mathbf{x})] = 2^h \sum_{\mathbf{l}\in L(h)} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}). \tag{82}$$

Recall the discussion at the beginning of Section 11. The estimate (56) shows that the set $\mathscr{P}(2^h)$ is insufficient for us to establish Theorem 7 in the case $k = 2$. To overcome this problem, we use digit shifts in Section 12. Here, for every $\mathbf{t} \in \mathbf{Z}_2^{2h}$, we consider the set

$$\mathscr{P}(2^h) \oplus \mathbf{t} = \{\mathbf{p} \oplus \mathbf{t} : \mathbf{p} \in \mathscr{P}(2^h)\}$$

where, for every

$$\mathbf{p} = (0.a_1 \dots a_h, 0.a_h \dots a_1) \in \mathscr{P}(2^h) \quad \text{and} \quad \mathbf{t} = (t_1', \dots, t_h', t_h'', \dots, t_1'') \in \mathbf{Z}_2^{2h},$$

we have the shifted point[7]

$$\mathbf{p} \oplus \mathbf{t} = (0.b_1' \dots b_h', 0.b_h'' \dots b_1''),$$

with the digits $b_1', \dots, b_h', b_1'', \dots, b_h'' \in \{0,1\}$ satisfying

$$b_s' \equiv a_s + t_s' \bmod 2 \quad \text{and} \quad b_s'' \equiv a_s + t_s'' \bmod 2$$

for every $s = 1, \dots, h$. Then

$$\begin{aligned} D^{(h)}[\mathscr{P}(2^h) \oplus \mathbf{t}; B(\mathbf{x})] &= \sum_{\mathbf{p}\in\mathscr{P}(2^h)} \chi_{B(\mathbf{x})}^{(h)}(\mathbf{p} \oplus \mathbf{t}) - 2^h x_1 x_2 \\ &= \sum_{\substack{\ell_1=0}}^{2^h-1} \sum_{\substack{\ell_2=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{2^h-1} \left(\sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p} \oplus \mathbf{t}) \right) \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) \\ &= \sum_{\substack{\ell_1=0}}^{2^h-1} \sum_{\substack{\ell_2=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{2^h-1} W_{\mathbf{l}}(\mathbf{t}) \left(\sum_{\mathbf{p}\in\mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) \right) \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}), \end{aligned}$$

in view of (76) and the second identity in (79). It follows that

[7] Here we somewhat abuse notation, as $\mathbf{t}$ clearly has more coordinates than $\mathbf{p}$. In the sequel, $W_{\mathbf{l}}(\mathbf{t})$ is really $W_{\mathbf{l}}(\mathbf{0} \oplus \mathbf{t})$, notation abused again.

$$D^{(h)}[\mathscr{P}(2^h);B(\mathbf{x})] = 2^h \sum_{\mathbf{l}\in L(h)} W_{\mathbf{l}}(\mathbf{t})\widetilde{\chi}_{\mathbf{l}}(\mathbf{x}).$$

Squaring this expression and summing over all $\mathbf{t} \in \mathbf{Z}_2^{2h}$, we obtain

$$\begin{aligned}\sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} |D^{(h)}[\mathscr{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2 &= 4^h \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} \left(\sum_{\mathbf{l}\in L(h)} W_{\mathbf{l}}(\mathbf{t})\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})\right)^2\\ &= 4^h \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} \sum_{\mathbf{l}',\mathbf{l}''\in L(h)} W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t})\widetilde{\chi_{\mathbf{l}'}}(\mathbf{x})\widetilde{\chi_{\mathbf{l}''}}(\mathbf{x})\\ &= 4^h \sum_{\mathbf{l}',\mathbf{l}''\in L(h)} \left(\sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t})\right)\widetilde{\chi_{\mathbf{l}'}}(\mathbf{x})\widetilde{\chi_{\mathbf{l}''}}(\mathbf{x}). \end{aligned} \tag{83}$$

Lemma 9. *For every $\mathbf{l}',\mathbf{l}'' \in \mathbf{N}_0^2$, we have*

$$\sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t}) = \begin{cases} 4^h, & \text{if } \mathbf{l}' = \mathbf{l}'',\\ 0, & \text{otherwise.}\end{cases}$$

Proof. Note first of all that in view of (75) and the second identity in (79), with $\mathbf{l}' \oplus \mathbf{l}'' = (\ell_1',\ell_2') \oplus (\ell_1'',\ell_2'') = (\ell_1' \oplus \ell_1'', \ell_2' \oplus \ell_2'')$, we have $W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t}) = W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t})$. For simplicity, write

$$S = \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t}) = \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}).$$

If $\mathbf{l}' = \mathbf{l}''$, so that $\mathbf{l}' \oplus \mathbf{l}'' = \mathbf{0}$, then $W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}) = W_{\mathbf{0}}(\mathbf{t}) = 1$ for every $\mathbf{t} \in \mathbf{Z}_2^{2h}$, and so clearly $S = \#\mathbf{Z}_2^{2h} = 4^h$. If $\mathbf{l}' \neq \mathbf{l}''$, so that $\mathbf{l}' \oplus \mathbf{l}'' \neq \mathbf{0}$, then there exists $\mathbf{t}_0 \in \mathbf{Z}_2^{2h}$ such that $W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}_0) \neq 1$. As $\mathbf{t}$ runs through the group $\mathbf{Z}_2^{2h}$, so does $\mathbf{t} \oplus \mathbf{t}_0$, so that

$$S = \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}\oplus\mathbf{t}_0) = \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t})W_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}_0) = SW_{\mathbf{l}'\oplus\mathbf{l}''}(\mathbf{t}_0),$$

in view of (76) and the second identity in (79). Clearly $S = 0$ in this case. □

Combining (83) and Lemma 9, we deduce that

$$\frac{1}{4^h} \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} |D^{(h)}[\mathscr{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2 = 4^h \sum_{\mathbf{l}\in L(h)} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2, \tag{84}$$

so that on integrating trivially with respect to $\mathbf{x} \in [0,1]^2$, we have

$$\frac{1}{4^h} \sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}} \int_{[0,1]^2} |D^{(h)}[\mathscr{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x} = 4^h \sum_{\mathbf{l}\in L(h)} \int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}. \tag{85}$$

To estimate the right hand side of (85), we need to use a formula of Fine [21] on the Fourier–Walsh coefficients of the characteristic function $\chi_{[0,x)}(y)$.

Let $\rho(0) = 0$. For any integer $\ell \in \mathbf{N}$ with representation (72), let

$$\rho(\ell) = \max\{i \in \mathbf{N} : \lambda_i(\ell) \neq 0\}, \quad \text{so that} \quad 2^{\rho(\ell)-1} \leq \ell < 2^{\rho(\ell)}. \tag{86}$$

Then the formula of Fine gives

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2 \, \mathrm{d}x = \frac{4^{-\rho(\ell)}}{3}.$$

If we write $\rho(\mathbf{l}) = \rho(\ell_1) + \rho(\ell_2)$ for $\mathbf{l} = (\ell_1, \ell_2)$, then in view of the first identity in (79), we have

$$\int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} = \frac{4^{-\rho(\mathbf{l})}}{9},$$

and the identity (85) becomes

$$\frac{1}{4^h} \sum_{\mathbf{t} \in \mathbf{Z}_2^{2h}} \int_{[0,1]^2} |D^{(h)}[\mathscr{P}(2^h) \oplus \mathbf{t}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = \frac{4^h}{9} \sum_{\mathbf{l} \in L(h)} 4^{-\rho(\mathbf{l})}. \tag{87}$$

To estimate the sum on the right hand side of (87), we need some reasonably precise information on the set $L(h)$. The following result is rather useful.

Lemma 10. *For every $y \in [0,1)$ and every $s \in \mathbf{N}_0$, we have*

$$\sum_{\ell=0}^{2^s-1} w_\ell(y) = 2^s \chi_{[0,2^{-s})}(y).$$

Proof. If $y \in [0, 2^{-s})$, then it follows from (73) that $\eta_i(y) = 0$ whenever $1 \leq i \leq s$. On the other hand, for every $\ell = 0, 1, 2, \ldots, 2^s - 1$, it follows from (72) that $\lambda_i(\ell) = 0$ for every $i > s$. It follows that for every $\ell = 0, 1, 2, \ldots, 2^s - 1$, we have

$$\sum_{i=1}^{\infty} \lambda_i(\ell) \eta_i(y) = 0,$$

and so $w_\ell(y) = 1$. On the other hand, if $y \in [2^{-s}, 1)$, then it follows from (73) that there exists some $j \in \{1, \ldots, s\}$ such that $\eta_j(y) = 1$. We now choose $k \in \{1, 2, \ldots, 2^s - 1\}$ such that $\lambda_j(k) = 1$ and $\lambda_i(k) = 0$ for every $i \neq j$. Then $w_k(y) \neq 1$. It is easy to see that as ℓ runs through the set $0, 1, 2, \ldots, 2^s - 1$, then so does $\ell \oplus k$, so that

$$\sum_{\ell=0}^{2^s-1} w_\ell(y) = \sum_{\ell=0}^{2^s-1} w_{\ell \oplus k}(y) = w_k(y) \sum_{\ell=0}^{2^s-1} w_\ell(y),$$

in view of (75). The result follows immediately. □

Lemma 11. *For every $s_1, s_2 \in \{0, 1, \ldots, h\}$, let*

$$\Xi(s_1,s_2)=\sum_{\ell_1=0}^{2^{s_1}-1}\sum_{\ell_2=0}^{2^{s_2}-1}\sum_{\mathbf{p}\in\mathscr{P}(2^h)}W_{\mathbf{l}}(\mathbf{p}).$$

Then

$$\Xi(s_1,s_2)=\begin{cases}2^{s_1+s_2}, & \text{if } s_1+s_2\geq h,\\ 2^h, & \text{if } s_1+s_2\leq h.\end{cases}$$

Proof. Writing $\mathbf{p}=(p_1,p_2)$ and $\mathbf{l}=(\ell_1,\ell_2)$ and noting the second identity in (79) and Lemma 10, we have

$$\begin{aligned}\sum_{\ell_1=0}^{2^{s_1}-1}\sum_{\ell_2=0}^{2^{s_2}-1}\sum_{\mathbf{p}\in\mathscr{P}(2^h)}W_{\mathbf{l}}(\mathbf{p})&=\sum_{\mathbf{p}\in\mathscr{P}(2^h)}\left(\sum_{\ell_1=0}^{2^{s_1}-1}w_{\ell_1}(p_1)\right)\left(\sum_{\ell_2=0}^{2^{s_2}-1}w_{\ell_2}(p_2)\right)\\&=2^{s_1+s_2}\sum_{\mathbf{p}\in\mathscr{P}(2^h)}\chi_{[0,2^{-s_1})}(p_1)\chi_{[0,2^{-s_2})}(p_2)\\&=2^{s_1+s_2}\sum_{\mathbf{p}\in\mathscr{P}(2^h)}\chi_{[0,2^{-s_1})\times[0,2^{-s_2})}(\mathbf{p}).\end{aligned}$$

It is not difficult to deduce from Lemma 4 that every rectangle of the form

$$[m_1 2^{-s},(m+1)2^{-s})\times[m_2 2^{s-h},(m_2+1)2^{s-h})\subseteq[0,1)^2$$

where $m_1,m_2\in\mathbf{N}_0$, and area 2^{-h}, contains precisely one point of $\mathscr{P}(2^h)$. Let us say that such a rectangle is an elementary rectangle. Suppose first of all that $s_1+s_2\geq h$. Then the rectangle $[0,2^{-s_1})\times[0,2^{-s_2})$ is contained in one elementary rectangle anchored at the origin, and so contains at most one point of $\mathscr{P}(2^h)$. Clearly it contains the point $\mathbf{0}\in\mathscr{P}(2^h)$, and so

$$\sum_{\mathbf{p}\in\mathscr{P}(2^h)}\chi_{[0,2^{-s_1})\times[0,2^{-s_2})}(\mathbf{p})=1.$$

Suppose then that $s_1+s_2\leq h$. Then the rectangle $[0,2^{-s_1})\times[0,2^{-s_2})$ is a union of precisely $2^{h-s_1-s_2}$ elementary rectangles, and so contains precisely $2^{h-s_1-s_2}$ points of $\mathscr{P}(2^h)$, whence

$$\sum_{\mathbf{p}\in\mathscr{P}(2^h)}\chi_{[0,2^{-s_1})\times[0,2^{-s_2})}(\mathbf{p})=2^{h-s_1-s_2}.$$

This completes the proof. □

Note that with $s_1=s_2=h$, Lemma 11 gives

$$\sum_{\ell_1=0}^{2^h-1}\sum_{\ell_2=0}^{2^h-1}\sum_{\mathbf{p}\in\mathscr{P}(2^h)}W_{\mathbf{l}}(\mathbf{p})=4^h.$$

In view of (80) and (81), we conclude that $\#L(h)=2^h-1$. We now study the set $L(h)$ in greater detail.

Lemma 12. *For every $s_1, s_2 \in \{1, \dots, h\}$, let*

$$L(s_1, s_2) = \left\{ \mathbf{l} \in [2^{s_1-1}, 2^{s_1}) \times [2^{s_2-1}, 2^{s_2}) : \sum_{\mathbf{p} \in \mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) = 2^h \right\}.$$

Then

a) for every $\mathbf{l} \in L(s_1, s_2)$, we have $\rho(\mathbf{l}) = s_1 + s_2$;
b) we have

$$\#L(s_1, s_2) = \begin{cases} 2^{s_1+s_2-h-2}, & \text{if } s_1 + s_2 \geq h+2, \\ 1, & \text{if } s_1 + s_2 = h+1, \\ 0, & \text{otherwise}. \end{cases}$$

Furthermore, every $\mathbf{l} \in L(h)$ belongs to $L(s_1, s_2)$ for some $s_1, s_2 \in \{1, \dots, h\}$ that satisfy $s_1 + s_2 \geq h+1$.

Proof. Note that if $\mathbf{l} \in L(s_1, s_2)$, then $\rho(\mathbf{l}) = \rho(\ell_1) + \rho(\ell_2) = s_1 + s_2$, in view of (86). This establishes part (a). To prove part (b), note that in view of (81), we have, in the notation of Lemma 11,

$$\begin{aligned} \#L(s_1, s_2) &= 2^{-h} \sum_{\ell_1 = 2^{s_1-1}}^{2^{s_1}-1} \sum_{\ell_2 = 2^{s_2-1}}^{2^{s_2}-1} \sum_{\mathbf{p} \in \mathscr{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) \\ &= 2^{-h} (\Xi(s_1, s_2) - \Xi(s_1 - 1, s_2) - \Xi(s_1, s_2 - 1) + \Xi(s_1 - 1, s_2 - 1)). \end{aligned}$$

Part (b) now follows easily from Lemma 11. Finally, it is easily checked that

$$\sum_{\substack{s_1=1 \\ s_1+s_2=h+1}}^{h} \sum_{s_2=1}^{h} 1 + \sum_{\substack{s_1=1 \\ s_1+s_2 \geq h+2}}^{h} \sum_{s_2=1}^{h} 2^{s_1+s_2-h-2} = 2^h - 1 = \#L(h).$$

The last assertion follows immediately. □

Using Lemma 12, we deduce that

$$\begin{aligned} \sum_{\mathbf{l} \in L(h)} 4^{-\rho(\mathbf{l})} &= \sum_{\substack{s_1=1 \\ s_1+s_2=h+1}}^{h} \sum_{s_2=1}^{h} 4^{-h-1} + \sum_{\substack{s_1=1 \\ s_1+s_2 \geq h+2}}^{h} \sum_{s_2=1}^{h} 2^{s_1+s_2-h-2} 4^{-s_1-s_2} \\ &= \sum_{\substack{s_1=1 \\ s_1+s_2=h+1}}^{h} \sum_{s_2=1}^{h} 4^{-h-1} + \sum_{\substack{s_1=1 \\ s_1+s_2 \geq h+2}}^{h} \sum_{s_2=1}^{h} 2^{-s_1-s_2-h-2} \\ &= \sum_{\substack{s_1=1 \\ s_1+s_2=h+1}}^{h} \sum_{s_2=1}^{h} 4^{-h-1} + \sum_{k=2}^{h} \sum_{\substack{s_1=1 \\ s_1+s_2=h+k}}^{h} \sum_{s_2=1}^{h} 2^{-h-k-h-2} \end{aligned}$$

$$= 4^{-h-1}h + 4^{-h-1}\sum_{k=2}^{h}\sum_{\substack{s_1=1\\ s_1+s_2=h+k}}^{h}\sum_{s_2=1}^{h} 2^{-k}$$

$$< 4^{-h-1}h + 4^{-h-1}h\sum_{k=2}^{h} 2^{-k} < 4^{-h}h.$$

Combining this with (87), we obtain

$$\frac{1}{4^h}\sum_{\mathbf{t}\in\mathbf{Z}_2^{2h}}\int_{[0,1]^2}|D^{(h)}[\mathscr{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x} < \frac{h}{9} \ll \log N,$$

noting that $N = 2^h$ in this case. Hence there is a digit shift $\mathbf{t}^* \in \mathbf{Z}_2^{2h}$ such that

$$\int_{[0,1]^2}|D^{(h)}[\mathscr{P}(2^h)\oplus\mathbf{t}^*;B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x} \ll \log N,$$

essentially establishing Theorem 7 in the case $k = 2$, apart from our not having properly analyzed the effect of the approximation of the certain characteristic functions by their truncated Fourier–Walsh series.

We complete this section by making an important comment for later use. Let us return to (82) and make the hypothetical assumption that the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l} \in L(h)$, are orthogonal. Then

$$\int_{[0,1]^2}|D^{(h)}[\mathscr{P}(2^h);B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x} = 4^h\sum_{\mathbf{l}\in L(h)}\int_{[0,1]^2}|\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}.$$

Note that the right hand side is exactly the same as the right hand side of (85), so that we can analyze this as before.

Unfortunately, the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l} \in L(h)$, are not orthogonal in this instance, so we cannot proceed in this way. Our technique in overcoming this handicap is to make use of the digit shifts $\mathbf{t} \in \mathbf{Z}_2^{2h}$, and bring into the argument, one may say through the back door, some orthogonality in the form of Lemma 9. We shall return to this in Sections 15 and 16.

14 Generalizations of van der Corput Point Sets

In our discussion of the van der Corput sequence and van der Corput point sets in Sections 10 and 11, we have restricted our discussion to dimension $k = 2$. Indeed, historically, the van der Corput sequence is constructed dyadically, and offers no generalization to the multi-dimensional case without going beyond dyadic constructions, except for one instance which we shall describe later in this section.

To study the general case in Theorems 7 and 8, one way is to generalize the van der Corput sequence. Here we know two ways of doing so, one by Halton [23]

and the other by Faure [20]. The Halton construction enables Halton to establish Theorem 8 in its generality and forms the basis for the proof of Theorem 7 in its generality by Roth [35]. The Faure construction enables Faure to give an alternative proof of Theorem 8 in its generality, enables Chen [11] soon afterwards to give an alternative proof of Theorem 7 in its generality and, more recently, forms the basis for the explicit construction proof of Theorem 7 by Chen and Skriganov [13, 14].

The generalizations by Halton and by Faure both require the very natural p-adic generalization of the van der Corput construction. The difference is that while Halton uses many different primes p, Faure uses only one such prime p but chosen to be sufficiently large.

14.1 Halton Point Sets

We first discuss Halton's contribution. Recall the dyadic construction (45) and (46) of the classical van der Corput sequence. Suppose now that we wish to study Theorem 7 or 8 in arbitrary dimension $k \geq 2$. Let p_i, where $i = 1, \ldots, k-1$, denote the first $k-1$ primes, with $p_1 < \ldots < p_{k-1}$. For every non-negative integer $n \in \mathbf{N}_0$ and every $i = 1, \ldots, k-1$, we write

$$n = \sum_{j=1}^{\infty} a_j^{(i)} p_i^{j-1} \tag{88}$$

as a p_i-adic expansion. Then we write

$$c_n^{(i)} = \sum_{j=1}^{\infty} a_j^{(i)} p_i^{-j}. \tag{89}$$

Finally we write

$$\mathbf{c}_n = (c_n^{(1)}, \ldots, c_n^{(k-1)}).$$

Note that $\mathbf{c}_n \in [0,1)^{k-1}$. The infinite sequence $\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \ldots$ is usually called a Halton sequence, and the infinite set

$$\mathscr{H} = \{(\mathbf{c}_n, n) : n = 0, 1, 2, \ldots\} \tag{90}$$

in $[0,1)^{k-1} \times [0,\infty)$ is usually called a Halton point set.

Corresponding to Lemma 3, we have the following multi-dimensional version.

Lemma 13. *For all non-negative integers $s_1, \ldots, s_{k-1}$ and $\ell_1, \ldots, \ell_{k-1}$ satisfying $\ell_i < p_i^{s_i}$ for every $i = 1, \ldots, k-1$, the set*

$$\left\{ n \in \mathbf{N}_0 : \mathbf{c}_n \in \prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \right\}$$

contains precisely all the elements of a residue class modulo $p_1^{s_1} \dots p_{k-1}^{s_{k-1}}$ *in* $\mathbf{N}_0$.

Proof. For fixed $i = 1, \dots, k-1$, the p_i-adic version of Lemma 3 says that the set

$$\{n \in \mathbf{N}_0 : c_n^{(i)} \in [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i})\}$$

contains precisely all the elements of a residue class modulo $p_i^{s_i}$ in $\mathbf{N}_0$. The result now follows from the Chinese remainder theorem. $\square$

We say that a rectangular box of the form

$$\prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \subseteq [0,1)^{k-1}$$

for some integers $\ell_1, \dots, \ell_{k-1}$ is an elementary $(p_1, \dots, p_{k-1})$-adic box of volume $p_1^{-s_1} \dots p_{k-1}^{-s_{k-1}}$. Hence Lemma 13 says that the given Halton sequence has very good distribution among such elementary $(p_1, \dots, p_{k-1})$-adic boxes for all non-negative integer values of $s_1, \dots, s_{k-1}$.

Lemma 14. *For all non-negative integers* $s_1, \dots, s_{k-1}$, $\ell_1, \dots, \ell_{k-1}$ *and* m *satisfying* $\ell_i < p_i^{s_i}$ *for every* $i = 1, \dots, k-1$, *the rectangular box*

$$\prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \times \left[m \prod_{i=1}^{k-1} p_i^{s_i}, (m+1) \prod_{i=1}^{k-1} p_i^{s_i} \right)$$

contains precisely one point of the Halton point set $\mathscr{H}$.

Clearly there is an average of one point of the Halton point set $\mathscr{H}$ per unit volume in $[0,1)^{k-1} \times [0, \infty)$. For any measurable set A in $[0,1)^{k-1} \times [0, \infty)$, let

$$E[\mathscr{H}; A] = \#(\mathscr{H} \cap A) - \mu(A)$$

denote the discrepancy of $\mathscr{H}$ in A.

We have the following generalization of Lemma 5.

Lemma 15. *For all non-negative integers* $s_1, \dots, s_{k-1}$ *and* $\ell_1, \dots, \ell_{k-1}$ *satisfying* $\ell_i < p_i^{s_i}$ *for every* $i = 1, \dots, k-1$, *there exist real numbers* α_0, β_0, *depending at most on* $s_1, \dots, s_{k-1}$ *and* $\ell_1, \dots, \ell_{k-1}$, *such that* $|\alpha_0| \le \frac{1}{2}$ *and*

$$E\left[\mathscr{H}; \prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \times [0, y)\right] = \alpha_0 - \psi(p_1^{-s_1} \dots p_{k-1}^{-s_{k-1}} (y - \beta_0)) \quad (91)$$

at all points of continuity of the right hand side.

We can now prove Theorem 8. Let $N \ge 2$ be a given integer. It follows at once from the definition of $\mathscr{H}$ that the set

$$\mathscr{H}_0 = \mathscr{H} \cap ([0,1)^{k-1} \times [0, N))$$

contains precisely N points. Let the integer h be determined uniquely by

$$p_1^{h-1} < N \leq p_1^h. \tag{92}$$

Consider a rectangular box of the form

$$B(x_1,\ldots,x_{k-1},y) = [0,x_1) \times \ldots \times [0,x_{k-1}) \times [0,y) \subseteq [0,1)^{k-1} \times [0,N).$$

Similar to our technique in Section 10, we shall approximate each interval $[0,x_i)$, where $i = 1,\ldots,k-1$, by the subinterval $[0,x_i^{(h)})$, where $x_i^{(h)} = p_i^{-h}[p_i^h x_i]$ is the greatest integer multiple of p_i^{-h} not exceeding x_i, and then consider the smaller rectangular box

$$B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y) = [0,x_1^{(h)}) \times \ldots \times [0,x_{k-1}^{(h)}) \times [0,y)$$

as an approximation of $B(x_1,\ldots,x_{k-1},y)$. A slight elaboration of the corresponding argument in Section 10 will show that the difference

$$B(x_1,\ldots,x_{k-1},y) \setminus B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)$$

is contained in a union of at most $k-1$ sets of the type discussed in Lemma 14, and so

$$|E[\mathscr{H};B(x_1,\ldots,x_{k-1},y)] - E[\mathscr{H};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]| \leq k-1; \tag{93}$$

note that since $y \leq N$, it makes no difference whether we write $\mathscr{H}$ or $\mathscr{H}_0$ in our argument.

It remains to estimate $E[\mathscr{H};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]$. To do so, we need to write each interval $[0,x_i^{(h)})$, where $i = 1,\ldots,h-1$, as a union of elementary p_i-adic intervals, each of length p_i^{-s} for some integer s satisfying $0 \leq s \leq h$.

If $x_i^{(h)} = 1$, then $[0,x_i^{(h)})$ is a union of precisely one elementary p_i-adic interval of unit length, so we now assume that $0 \leq x_i^{(h)} < 1$.

Lemma 16. *Suppose that* $0 \leq x_i^{(h)} < 1$, *with*

$$x_i^{(h)} = \sum_{s=1}^{h} b_s p_i^{-s}$$

as a p_i*-adic expansion. Then* $[0,x_i^{(h)})$ *can be written as a union of*

$$\sum_{s=1}^{h} b_s < hp_i$$

elementary p_i*-adic intervals, namely* b_1 *elementary* p_i*-adic intervals of length* p_i^{-1}, *together with* b_2 *elementary* p_i*-adic intervals of length* p_i^{-2}, *and so on.*

Hence the set $B(x_1^{(h)},\dots,x_{k-1}^{(h)},y)$ is a disjoint union of fewer than $h^{k-1}p_1\dots p_{k-1}$ sets of the type discussed in Lemma 15. Hence

$$|E[\mathscr{H};B(x_1^{(h)},\dots,x_{k-1}^{(h)},y)]| < h^{k-1}p_1\dots p_{k-1} \ll_k (\log N)^{k-1}. \tag{94}$$

Combining (93) and (94), we conclude that

$$|E[\mathscr{H};B(x_1,\dots,x_{k-1},y)]| \ll_k (\log N)^{k-1}. \tag{95}$$

Finally, rescaling the second coordinate of the points of $\mathscr{H}_0$ by a factor N^{-1}, we obtain a set

$$\mathscr{P} = \{(\mathbf{c}_n, N^{-1}n) : n = 0,1,2,\dots,N-1\}$$

of precisely N points in $[0,1)^k$. For every $\mathbf{x} = (x_1,\dots,x_k) \in [0,1]^k$, we have

$$D[\mathscr{P};B(\mathbf{x})] = E[\mathscr{H}_0;[0,x_1)\times\dots\times[0,x_{k-1})\times[0,Nx_k)] \ll_k (\log N)^{k-1},$$

in view of (95) and noting that $0 \le Nx_k \le N$. This now completes the proof of Theorem 8.

Next we discuss Roth's ideas in shaping this Halton construction to give a proof of Theorem 7. As in the special case $k = 2$, one needs to introduce a probabilistic variable. To pave the way for this, we shall modify the Halton point set somewhat. Let $N \ge 2$ be a given integer, and let the integer h be determined uniquely by

$$p_1^{h-1} < N \le p_1^h, \tag{96}$$

as before. For every $i = 1,\dots,k-1$ and every $n = 0,1,2,\dots,p_i^h - 1$, we define $c_n^{(i)}$ as before by (88) and (89). We then extend the definition of $c_n^{(i)}$ to all other integers using periodicity by writing

$$c_{n+p_i^h} = c_n \quad \text{for every } n \in \mathbf{Z},$$

write $\mathbf{c}_n = (c_n^{(1)},\dots,c_n^{(k-1)})$, and consider the extended Halton point set

$$\mathscr{H}_h = \{(c_n,n) : n \in \mathbf{Z}\}.$$

Remark. In Roth [35], as well as Chen [10], the construction of the set $\mathscr{H}_h$ is slightly different, but the difference does not affect the argument in any way. Let $M = p_1\dots p_{k-1}$. One then defines $c_n^{(i)}$ for $n = 0,1,2,\dots,M^h-1$ by (88) and (89), write $\mathbf{c}_n = (c_n^{(1)},\dots,c_n^{(k-1)})$ for these values of n, and define $\mathbf{c}_n$ for all other integer values of n by the periodicity relationship $\mathbf{c}_{n+M^h} = \mathbf{c}_n$ for every $n \in \mathbf{Z}$.

Furthermore, for every real number $t \in \mathbf{R}$, we consider the translated Halton point set

$$\mathscr{H}_h(t) = \{(\mathbf{c}_n, n+t) : n \in \mathbf{Z}\}.$$

It is clear that there is an average of one point of the translated Halton point set $\mathscr{H}_h(t)$ per unit volume in $[0,1)^{k-1} \times (-\infty,\infty)$. For any measurable set A in $[0,1)^{k-1} \times (-\infty,\infty)$, we now let

$$E[\mathscr{H}_h(t);A] = \#(\mathscr{H}_h(t) \cap A) - \mu(A)$$

denote the discrepancy of $\mathscr{H}_h(t)$ in A.

Consider a rectangular box of the form

$$B(x_1,\ldots,x_{k-1},y) = [0,x_1) \times \ldots \times [0,x_{k-1}) \times [0,y) \subseteq [0,1)^{k-1} \times [0,N).$$

As in the earlier proof of Theorem 8, we shall consider the smaller rectangular box $B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)$ and, corresponding to (93), we have

$$|E[\mathscr{H}_h(t);B(x_1,\ldots,x_{k-1},y)] - E[\mathscr{H};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]| \le k-1. \tag{97}$$

Next, we study $E[\mathscr{H}_h(t);B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]$ in detail, and require an analogue of the expansion (60). It is not difficult to see that

$$E[\mathscr{H};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)] = \sum_{I_1 \in \mathscr{I}_1} \ldots \sum_{I_{k-1} \in \mathscr{I}_{k-1}} E[\mathscr{H}_h(t);\mathbf{I} \times [0,y)],$$

where $\mathbf{I} = I_1 \times \ldots \times I_{k-1}$ and where, for every $i = 1,\ldots,k-1$, $\mathscr{I}_i$ denotes the collection of elementary p_i-adic intervals in the union that makes up the interval $[0,x_i^{(h)})$ in Lemma 16.

Corresponding to Lemma 6, one can show that each summand

$$E[\mathscr{H}_h(t);\mathbf{I} \times [0,y)]$$

can be written in the form

$$\psi(p_1^{-s_1} \ldots p_{k-1}^{-s_{k-1}}(t-\beta_{\mathbf{I}})) - \psi(p_1^{-s_1} \ldots p_{k-1}^{-s_{k-1}}(t-\gamma_{\mathbf{I}})),$$

where the real numbers $\beta_{\mathbf{I}}$ and $\gamma_{\mathbf{I}}$ depend at most on $\mathbf{I}$ and y, and where, for every $i = 1,\ldots,k-1$, the elementary p_i-adic interval I_i has length $p_i^{-s_i}$. Making use of this, one can then proceed to show, corresponding to Lemma 7, that

$$\int_0^{M^h} E[\mathscr{H}_h(t);\mathbf{I}' \times [0,y)] E[\mathscr{H}_h(t);\mathbf{I}'' \times [0,y)] \, \mathrm{d}t = O\left(M^h \prod_{i=1}^{k-1} p_i^{-|s_i'-s_i''|} \right)$$

for any $\mathbf{I}' = I_1' \times \ldots \times I_{k-1}'$ and $\mathbf{I}'' = I_1'' \times \ldots \times I_{k-1}''$ where, for every $i = 1,\ldots,k-1$, the elementary p_i-adic intervals $I_i', I_i'' \in \mathscr{I}_i$ have lengths $p_i^{-s_i'}$ and $p_i^{-s_i''}$ respectively. One then goes on to show that

$$\int_0^{M^h} |E[\mathscr{H};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]|^2 \, \mathrm{d}t$$

$$\ll \sum_{I_1' \in \mathscr{I}_1} \dots \sum_{I_{k-1}' \in \mathscr{I}_{k-1}} \sum_{I_1'' \in \mathscr{I}_1} \dots \sum_{I_{k-1}'' \in \mathscr{I}_{k-1}} M^h \prod_{i=1}^{k-1} p_i^{-|s_i'-s_i''|}$$
$$\ll_k M^h h^{k-1}.$$

Taking the bound (97) into account and then integrating trivially with respect to $x_1,\dots,x_{k-1}$, each over the interval $[0,1)$, and with respect to y over the interval $[0,N)$, we conclude that

$$\int_0^N \int_0^1 \dots \int_0^1 \int_0^{M^h} |E[\mathscr{H}_h(t);B(x_1,\dots,x_{k-1},y)]|^2 \,\mathrm{d}t\,\mathrm{d}x_1\dots\mathrm{d}x_{k-1}\,\mathrm{d}y$$
$$= \int_0^{M^h} \left(\int_0^N \int_0^1 \dots \int_0^1 |E[\mathscr{H}_h(t);B(x_1,\dots,x_{k-1},y)]|^2 \,\mathrm{d}x_1\dots\mathrm{d}x_{k-1}\,\mathrm{d}y \right) \mathrm{d}t$$
$$\ll_k M^h h^{k-1} N.$$

Hence there exists $t^* \in [0,M^h)$ such that

$$\int_0^N \int_0^1 \dots \int_0^1 |E[\mathscr{H}_h(t^*);B(x_1,\dots,x_{k-1},y)]|^2 \,\mathrm{d}x_1\dots\mathrm{d}x_{k-1}\,\mathrm{d}y$$
$$\ll_k h^{k-1} N. \tag{98}$$

Finally, we note that the set $\mathscr{H}_h(t^*) \cap ([0,1)^{k-1} \times [0,N))$ contains precisely N points. Rescaling in the vertical direction by a factor N^{-1}, we observe that the set

$$\mathscr{P}^* = \{(z_1,\dots,z_{k-1},N^{-1}z_k) : (z_1,\dots,z_k) \in \mathscr{H}_h(t^*)\}$$

contains precisely N points in $[0,1)^k$, and the estimate (98) now translates to

$$\int_{[0,1]^k} |D[\mathscr{P}^*;B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \ll_k h^{k-1} \ll_k (\log N)^{k-1},$$

in view of (96). This completes our brief sketch of the proof of Theorem 7.

14.2 Faure Point Sets

We now discuss Faure's contribution. Suppose again that we wish to study Theorem 7 or 8 in arbitrary dimension $k \geq 2$. Let p denote a prime such that[8] $p \geq k-1$. For every non-negative integer $n \in \mathbf{N}_0$, we write

$$n = \sum_{j=1}^{\infty} a_j^{(1)} p^{j-1} \tag{99}$$

[8] The assumption that $p \geq k-1$ cannot be relaxed, as noted by Chen [11].

as a p-adic expansion. Then we write

$$c_n^{(1)} = \sum_{j=1}^{\infty} a_j^{(1)} p^{-j}. \tag{100}$$

For $i = 2, \ldots, k-1$, we shall write

$$c_n^{(i)} = \sum_{j=1}^{\infty} a_j^{(i)} p^{-j}, \tag{101}$$

where the coefficients $a_j^{(i)}$ are defined inductively using the infinite upper triangular matrix

$$\mathscr{B} = \begin{bmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \cdots \\ & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \cdots \\ & & \binom{2}{2} & \binom{3}{2} & \cdots \\ & & & \binom{3}{3} & \cdots \\ & & & & \ddots \end{bmatrix} \tag{102}$$

made up of binomial coefficients.

It is convenient to use matrix multiplication modulo p to define the coefficients $a_j^{(i)}$ when $i > 1$. For every $i = 1, \ldots, k-1$, consider the infinite column matrix

$$\mathbf{a}^{(i)} = \begin{bmatrix} a_1^{(i)} \\ a_2^{(i)} \\ a_3^{(i)} \\ a_4^{(i)} \\ \vdots \end{bmatrix}.$$

Then for every $i = 2, \ldots, k-1$, we write

$$\mathbf{a}^{(i)} \equiv \mathscr{B}\mathbf{a}^{(i-1)} \bmod p;$$

in other words, we write

$$\begin{bmatrix} a_1^{(i)} \\ a_2^{(i)} \\ a_3^{(i)} \\ a_4^{(i)} \\ \vdots \end{bmatrix} \equiv \begin{bmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \cdots \\ & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \cdots \\ & & \binom{2}{2} & \binom{3}{2} & \cdots \\ & & & \binom{3}{3} & \cdots \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} a_1^{(i-1)} \\ a_2^{(i-1)} \\ a_3^{(i-1)} \\ a_4^{(i-1)} \\ \vdots \end{bmatrix} \bmod p.$$

For every $n \in \mathbf{N}_0$, write

$$\mathbf{c}_n = (c_n^{(1)}, \ldots, c_n^{(k-1)}).$$

The set

$$\mathscr{F} = \{(\mathbf{c}_n, n) : n = 0, 1, 2, \ldots\}$$

in $[0,1)^{k-1} \times [0,\infty)$ is usually called a Faure point set.

Analogous to Lemma 14, we have the following result.

Lemma 17. *For all non-negative integers $s_1, \ldots, s_{k-1}$, $\ell_1, \ldots, \ell_{k-1}$ and m such that $\ell_i < p^{s_i}$ holds for every $i = 1, \ldots, k-1$, the rectangular box*

$$\prod_{i=1}^{k-1} [\ell_i p^{-s_i}, (\ell_i + 1) p^{-s_i}) \times [m p^{s_1 + \ldots + s_{k-1}}, (m+1) p^{s_1 + \ldots + s_{k-1}}) \tag{103}$$

contains precisely one point of the Faure point set $\mathscr{F}$.

To prove Lemma 17, we need a simple result concerning the matrix $\mathscr{B}$.

Lemma 18. *For the matrix $\mathscr{B}$ given by (102), we have, for every $i = 1, \ldots, k-1$,*

$$\mathscr{B}^{i-1} = \begin{bmatrix} \binom{0}{0} & \binom{1}{0}(i-1) & \binom{2}{0}(i-1)^2 & \binom{3}{0}(i-1)^3 & \cdots \\ & \binom{1}{1} & \binom{2}{1}(i-1) & \binom{3}{1}(i-1)^2 & \cdots \\ & & \binom{2}{2} & \binom{3}{2}(i-1) & \cdots \\ & & & \binom{3}{3} & \cdots \\ & & & & \ddots \end{bmatrix}.$$

Proof (Proof of Lemma 17). Suppose that suitable integers $s_1, \ldots, s_{k-1}$, $\ell_1, \ldots, \ell_{k-1}$ and m are chosen and fixed. For a point $(\mathbf{c}_n, n)$ to lie in the rectangle (103), we must have

$$c_n^{(i)} \in [\ell_i p^{-s_i}, (\ell_i + 1) p^{-s_i}) \tag{104}$$

for every $i = 1, \ldots, k-1$, as well as

$$n \in [m p^{s_1 + \ldots + s_{k-1}}, (m+1) p^{s_1 + \ldots + s_{k-1}}). \tag{105}$$

Comparing (99) and (105), it is clear that the value of the coefficient $a_j^{(1)}$ for every $j > s_1 + \ldots + s_{k-1}$ is uniquely determined. It therefore remains to show that there is one choice of the vector

$$(a_1^{(1)}, \ldots, a_{s_1 + \ldots + s_{k-1}}^{(1)})$$

that satisfies the requirement (104) for every $i = 1, \ldots, k-1$.

Note next that for every $i = 1, \ldots, k-1$, we have

$$\begin{bmatrix} a_1^{(i)} \\ a_2^{(i)} \\ a_3^{(i)} \\ a_4^{(i)} \\ \vdots \end{bmatrix} \equiv \begin{bmatrix} \binom{0}{0} & \binom{1}{0}(i-1) & \binom{2}{0}(i-1)^2 & \binom{3}{0}(i-1)^3 & \cdots \\ & \binom{1}{1} & \binom{2}{1}(i-1) & \binom{3}{1}(i-1)^2 & \cdots \\ & & \binom{2}{2} & \binom{3}{2}(i-1) & \cdots \\ & & & \binom{3}{3} & \cdots \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \\ a_3^{(1)} \\ a_4^{(1)} \\ \vdots \end{bmatrix} \bmod p.$$

Let us consider the p-adic expansion

$$\ell_i\, p^{-s_i} = \beta_1^{(i)} p^{-1} + \ldots + \beta_{s_i}^{(i)} p^{-s_i}.$$

If (104) holds, then in view of (100) or (101), we must have $a_j^{(i)} = \beta_j^{(i)}$ for every $j = 1, \ldots, s_i$. This can be summarized by writing

$$\mathscr{W}_i \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \\ a_3^{(1)} \\ a_4^{(1)} \\ \vdots \end{bmatrix} \equiv \begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \\ \beta_3^{(i)} \\ \vdots \\ \beta_{s_i}^{(i)} \end{bmatrix} \bmod p, \tag{106}$$

where the matrix $\mathscr{W}_i$ contains precisely the first s_i rows of the matrix $\mathscr{B}^{i-1}$. Now recall that $a_j^{(1)}$ are already uniquely determined for every $j > S = s_1 + \ldots + s_{k-1}$ by (105), and clearly there are at most finitely many non-zero terms among these. The system (106) can therefore be simplified to one of the form

$$\mathscr{V}_i \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \\ a_3^{(1)} \\ \vdots \\ a_S^{(1)} \end{bmatrix} \equiv \begin{bmatrix} \gamma_1^{(i)} \\ \gamma_2^{(i)} \\ \gamma_3^{(i)} \\ \vdots \\ \gamma_{s_i}^{(i)} \end{bmatrix} \bmod p, \tag{107}$$

where the matrix $\mathscr{V}_i$ contains precisely the first S columns of the matrix $\mathscr{W}_i$. On combining (107) for every $i = 1, \ldots, k-1$, we arrive at a system of S linear congruences in the S variables $a_1^{(1)}, \ldots, a_S^{(1)}$, with the matrix given by

$$\mathscr{V} = \begin{bmatrix} \mathscr{V}_1 \\ \vdots \\ \mathscr{V}_{k-1} \end{bmatrix}.$$

It is not difficult to see that for every $i = 1, \ldots, k-1$, we have

$$\mathscr{V}_i = \begin{bmatrix} \binom{0}{0} & \binom{1}{0}(i-1) & \binom{2}{0}(i-1)^2 & \cdots & \binom{S-1}{0}(i-1)^{S-1} \\ & \binom{1}{1} & \binom{2}{1}(i-1) & \cdots & \binom{S-1}{1}(i-1)^{S-2} \\ & & \ddots & & \vdots \\ & & & \binom{s_i-1}{s_i-1} \cdots & \binom{S-1}{s_i-1}(i-1)^{S-s_i} \end{bmatrix},$$

a matrix with s_i rows and S columns. It follows that the matrix $\mathscr{V}$ is of generalized Vandermonde type, with determinant

$$\prod_{1 \le i' < i'' \le k-1} (i''-i')^{s_{i'} s_{i''}} \not\equiv 0 \bmod p,$$

in view of the assumption that $p \ge k-1$. Hence the system of S linear congruences in the S variables $a_1^{(1)}, \ldots, a_S^{(1)}$ has unique solution. Recall once again that the coefficients $a_j^{(1)}$ are already uniquely determined for every $j > S$, we conclude that there is precisely one value of n that satisfies all the requirements. $\square$

The following analogue of Lemma 15 is a simple consequence of Lemma 18.

Lemma 19. *For all non-negative integers $s_1, \ldots, s_{k-1}$ and $\ell_1, \ldots, \ell_{k-1}$ satisfying $\ell_i < p^{s_i}$ for every $i = 1, \ldots, k-1$, and for every real number $y > 0$, we have*

$$\left| E\left[\mathscr{F}; \prod_{i=1}^{k-1} [\ell_i p^{-s_i}, (\ell_i+1)p^{-s_i}) \times [0,y) \right] \right| \le 1.$$

To study Theorem 8, let $N \ge 2$ be a given integer. It follows at once from the definition of $\mathscr{F}$ that the set

$$\mathscr{F}_0 = \mathscr{F} \cap ([0,1)^{k-1} \times [0,N))$$

contains precisely N points. Let the integer h be determined uniquely by

$$p^{h-1} < N \le p^h.$$

We can now deduce Theorem 8 from Lemma 17 and Lemma 19 in a way similar to our deduction of the same result from Lemma 14 and Lemma 15 in Section 14.1, noting that Lemma 16 remains valid with p_i replaced by p. Indeed, rescaling the second coordinate of the points of $\mathscr{F}_0$ by a factor N^{-1}, we obtain a set

$$\mathscr{P} = \{(\mathbf{c}_n, N^{-1}n) : n = 0, 1, 2, \ldots, N-1\},$$

of precisely N points in $[0,1)^k$ and which satisfies the conclusion of Theorem 8.

14.3 A General Point Set and a Digit Shift Argument

In this section, we briefly describe a rather general digit shift argument developed by Chen [11] which enables us to establish Theorem 7 using Halton point sets discussed in Section 14.1 or Faure point sets discussed in Section 14.2. Recall that these point sets satisfy Lemma 14 and Lemma 17 respectively.

Let $p_1 \leq \ldots \leq p_{k-1}$ be primes, not necessarily distinct, and let h be a non-negative integer. We shall say that a set of the form

$$\mathscr{Z} = \{(\mathbf{c}_n, n) : n = 0, 1, 2, \ldots\} \tag{108}$$

in $[0,1)^{k-1} \times [0,\infty)$ is a 1-set of order h with respect to the primes $p_1, \ldots, p_{k-1}$ if the following condition is satisfied. For all non-negative integers $s_1, \ldots, s_{k-1}$, $\ell_1, \ldots, \ell_{k-1}$ and m satisfying $s_i \leq h$ and $\ell_i < p_i^{s_i}$ for every $i = 1, \ldots, k-1$, the rectangular box

$$\prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \times \left[m \prod_{i=1}^{k-1} p_i^{s_i}, (m+1) \prod_{i=1}^{k-1} p_i^{s_i} \right)$$

contains precisely one point of $\mathscr{Z}$.

If the primes $p_1, \ldots, p_{k-1}$ are distinct, then the Halton set $\mathscr{H}$ is a 1-set of every non-negative order with respect to $p_1, \ldots, p_{k-1}$. If the primes $p_1, \ldots, p_{k-1}$ are all identical and equal to p, then the Faure set $\mathscr{F}$ is 1-set of every non-negative order with respect to $p, \ldots, p$, provided that $p \geq k-1$.

The property below follows almost immediately from the definition.

Lemma 20. *Suppose that h be a non-negative integer, and that $\mathscr{Z}$ is a 1-set of order h with respect to the primes $p_1, \ldots, p_{k-1}$. Then for all non-negative integers $s_1, \ldots, s_{k-1}$ and $\ell_1, \ldots, \ell_{k-1}$ satisfying $s_i \leq h$ and $\ell_i < p_i^{s_i}$ for every $i = 1, \ldots, k-1$, and for every real number $y > 0$, we have*

$$\left| E\left[\mathscr{Z}; \prod_{i=1}^{k-1} [\ell_i p_i^{-s_i}, (\ell_i + 1) p_i^{-s_i}) \times [0, y) \right] \right| \leq 1.$$

Let $N \geq 2$ be a given integer, and let the integer h be determined uniquely by

$$p_1^{h-1} < N \leq p_1^h. \tag{109}$$

For any 1-set (108) of order h with respect to the primes $p_1, \ldots, p_{k-1}$, the set

$$\mathscr{Z}_0 = \mathscr{Z} \cap ([0,1)^{k-1} \times [0, N))$$

contains precisely N points. Then it can be shown easily that the set

$$\mathscr{P} = \{(\mathbf{c}_n, N^{-1} n) : n = 0, 1, 2, \ldots, N-1\},$$

of precisely N points in $[0,1)^k$ and which satisfies the conclusion of Theorem 8.

To study Theorem 7, we again choose the integer h to satisfy (109). However, we need to modify the 1-set $\mathscr{Z}$.

Let $\mathscr{M}$ denote the collection of all $(k-1)\times h$ matrices $\mathbf{T}=(t_{i,j})$ where, for every $i=1,\ldots,k-1$ and $j=1,\ldots,h$, the entry $t_{i,j}\in\{0,1,2,\ldots,p_i-1\}$. Clearly the collection $\mathscr{M}$ has $(p_1\ldots p_{k-1})^h$ elements.

For every $n=0,1,2,\ldots$, let us write

$$\mathbf{c}_n=(c_1(n),\ldots,c_{k-1}(n)).$$

For every $i=1,\ldots,k-1$, we consider the base p_i expansion

$$c_i(n)=0.a_{i,1}a_{i,2}\ldots a_{i,h}a_{i,h+1}\ldots.$$

For every $\mathbf{T}\in\mathscr{M}$ and every $n=0,1,2,\ldots,$, we shall write

$$\mathbf{c}_n^{\mathbf{T}}=(c_1^{\mathbf{T}}(n),\ldots,c_{k-1}^{\mathbf{T}}(n)),$$

where, for every $i=1,\ldots,k-1$, we have

$$c_i^{\mathbf{T}}(n)=0.(a_{i,1}\oplus t_{i,1})(a_{i,2}\oplus t_{i,2})\ldots(a_{i,h}\oplus t_{i,h})a_{i,h+1}\ldots,$$

where $\oplus$ denotes addition modulo p_i. It is not difficult to show that the shifted set

$$\mathscr{Z}^{\mathbf{T}}=\{(\mathbf{c}_n^{\mathbf{T}},n):n=0,1,2,\ldots\}$$

in $[0,1)^{k-1}\times[0,\infty)$ is also a 1-set of order h with respect to the primes $p_1,\ldots,p_{k-1}$.

Consider a rectangular box of the form

$$B(x_1,\ldots,x_{k-1},y)=[0,x_1)\times\ldots\times[0,x_{k-1})\times[0,y)\subseteq[0,1)^{k-1}\times[0,N).$$

As in the earlier proof of Theorem 7, we shall again consider the smaller rectangular box $B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)$, where, for every $i=1,\ldots,k-1$, we replace the point x_i by $x_i^{(h)}=p_i^{-h}[p_i^hx_i]$, the greatest integer multiple of p_i^{-h} not exceeding x_i. Then for every $\mathbf{T}\in\mathscr{M}$, we have

$$|E[\mathscr{Z}^{\mathbf{T}};B(x_1,\ldots,x_{k-1},y)]-E[\mathscr{Z}^{\mathbf{T}};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]|\le k-1,$$

so it remains to study $E[\mathscr{Z}^{\mathbf{T}};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]$ in detail. It can be shown that

$$\sum_{\mathbf{T}\in\mathscr{M}}|E[\mathscr{Z}^{\mathbf{T}};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]|^2\ll_k(p_1\ldots p_{k-1})^hh^{k-1},$$

from which it follows that

$$\int_0^N\int_0^1\ldots\int_0^1\left(\sum_{\mathbf{T}\in\mathscr{M}}|E[\mathscr{Z}^{\mathbf{T}};B(x_1^{(h)},\ldots,x_{k-1}^{(h)},y)]|^2\right)\mathrm{d}x_1\ldots\mathrm{d}x_{k-1}\,\mathrm{d}y$$

$$= \sum_{\mathbf{T}\in\mathscr{M}} \left(\int_0^N \int_0^1 \dots \int_0^1 |E[\mathscr{Z}^{\mathbf{T}}; B(x_1^{(h)}, \dots, x_{k-1}^{(h)}, y)]|^2 \, \mathrm{d}x_1 \dots \mathrm{d}x_{k-1} \, \mathrm{d}y \right)$$
$$\ll_k (p_1 \dots p_{k-1})^h h^{k-1} N.$$

Hence there exists $\mathbf{T}^* \in \mathscr{M}$ such that

$$\int_0^N \int_0^1 \dots \int_0^1 |E[\mathscr{Z}^{\mathbf{T}^*}; B(x_1, \dots, x_{k-1}, y)]|^2 \, \mathrm{d}x_1 \dots \mathrm{d}x_{k-1} \, \mathrm{d}y \ll_k h^{k-1} N.$$

Finally, we note that the set $\mathscr{Z}^{\mathbf{T}^*} \cap ([0,1)^{k-1} \times [0,N))$ contains precisely N points. Rescaling in the vertical direction by a factor N^{-1}, we observe that the set

$$\mathscr{P}^* = \{(z_1, \dots, z_{k-1}, N^{-1} z_k) : (z_1, \dots, z_k) \in \mathscr{Z}^{\mathbf{T}^*}\}$$

contains precisely N points in $[0,1)^k$, and satisfies the conclusion of Theorem 7.

15 Group Structure and p-adic Fourier–Walsh Analysis

In Section 13, we exploit the group structure of the van der Corput set $\mathscr{P}(2^h)$ to sketch a proof of Theorem 7 for $k = 2$. The central argument there is to use Fourier–Walsh analysis to show that an approximation $D^{(h)}[\mathscr{P}(2^h); B(\mathbf{x})]$ of the discrepancy function $D[\mathscr{P}(2^h); B(\mathbf{x})]$ satisfies the identity (85) which involves digit shifts. Under certain hypothetical orthogonality assumptions, we can further deduce the simpler identity

$$\int_{[0,1]^2} |D^{(h)}[\mathscr{P}(2^h); B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = 4^h \sum_{\mathbf{l}\in L(h)} \int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x}.$$

Unfortunately, these hypothetical orthogonality assumptions do not hold.

To have a better understanding of the underlying ideas, it is necessary to study p-adic versions of the analysis carried out earlier.

For simplicity, let us again restrict our attention to Theorem 7 for $k = 2$. Let p be a prime, and consider the base p van der Corput point set

$$\mathscr{P}(p^h) = \{(0.a_1 a_2 a_3 \dots a_h, 0.a_h \dots a_3 a_2 a_1) : a_1, \dots, a_h \in \{0, 1, \dots, p-1\}\}.$$

This is a finite abelian group isomorphic to the group $\mathbf{Z}_p^h$. We shall make use of the characters of these groups. These are the base p Walsh functions, usually known as the Chrestenson or Chrestenson–Levy functions. For simplicity, we refer to them all as Walsh functions here.

To define these Walsh functions, we first consider p-ary representation of any integer $\ell \in \mathbf{N}_0$, written uniquely in the form

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell) p^{i-1}, \tag{110}$$

where the coefficient $\lambda_i(\ell) \in \{0,1,\dots,p-1\}$ for every $i \in \mathbf{N}$. On the other hand, every real number $y \in [0,1)$ can be represented in the form

$$y = \sum_{i=1}^{\infty} \eta_i(y) p^{-i}, \tag{111}$$

where the coefficient $\eta_i(y) \in \{0,1,\dots,p-1\}$ for every $i \in \mathbf{N}$. This representation is unique if we agree that the series in (111) is finite for every $y = mp^{-s}$ where $s \in \mathbf{N}_0$ and $m \in \{0,1,\dots,p^s-1\}$.

For every $\ell \in \mathbf{N}_0$ of the form (110), we define the Walsh function $w_\ell : [0,1) \to \mathbf{R}$ by writing

$$w_\ell(y) = e_p\left(\sum_{i=1}^{\infty} \lambda_i(\ell)\eta_i(y)\right), \tag{112}$$

where $e_p(z) = \mathrm{e}^{2\pi \mathrm{i} z/p}$ for every real number z. Since (110) is essentially a finite sum, the Walsh function is well defined, and takes the p-th roots of unity as its values. It is easy to see that $w_0(y) = 1$ for every $y \in [0,1)$. It is well known that under the inner product

$$\langle w_k, w_\ell \rangle = \int_0^1 w_k(y)\overline{w_\ell(y)}\,\mathrm{d}y,$$

the collection of Walsh functions form an orthonormal basis of $L^2[0,1]$.

The operation $\oplus$ defined modulo 2 previously can easily be suitably modified to an operation modulo p. Then (75) and (76) remain valid in this new setting.

As before, we shall use Fourier–Walsh analysis to study characteristic functions of the form $\chi_{[0,x)}(y)$. We have the Fourier–Walsh series

$$\chi_{[0,x)}(y) \sim \sum_{\ell=0}^{\infty} \widetilde{\chi}_\ell(x)\overline{w_\ell(y)},$$

where, for every $\ell \in \mathbf{N}_0$, the Fourier–Walsh coefficients are given by

$$\widetilde{\chi}_\ell(x) = \int_0^x w_\ell(y)\,\mathrm{d}y.$$

In particular, we have $\widetilde{\chi}_0(x) = x$ for every $x \in [0,1)$. Again, as before, instead of using the full Fourier–Walsh series, we shall truncate it and use the approximation

$$\chi^{(h)}_{[0,x)}(y) = \sum_{\ell=0}^{p^h-1} \widetilde{\chi}_\ell(x)\overline{w_\ell(y)}.$$

This approximation in turn leads to the approximation

$$\chi^{(h)}_{B(\mathbf{x})}(\mathbf{y}) = \chi^{(h)}_{[0,x_1)}(y_1)\chi^{(h)}_{[0,x_2)}(y_2) = \sum_{\ell_1=0}^{p^h-1}\sum_{\ell_2=0}^{p^h-1} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x})\overline{W_{\mathbf{l}}(\mathbf{y})}$$

of the characteristic function $\chi_{B(\mathbf{x})}(\mathbf{y})$. Here $\mathbf{l} = (\ell_1, \ell_2)$,

$$\widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) = \widetilde{\chi_{\ell_1}}(x_1)\widetilde{\chi_{\ell_2}}(x_2) \quad \text{and} \quad W_{\mathbf{l}}(\mathbf{y}) = w_{\ell_1}(y_1)w_{\ell_2}(y_2).$$

Consequently, we approximate the discrepancy function

$$D[\mathscr{P}(p^h); B(\mathbf{x})] = \sum_{\mathbf{p}\in\mathscr{P}(p^h)} \chi_{B(\mathbf{x})}(\mathbf{p}) - p^h x_1 x_2$$

by

$$\begin{aligned} D^{(h)}[\mathscr{P}(p^h); B(\mathbf{x})] &= \sum_{\mathbf{p}\in\mathscr{P}(p^h)} \chi^{(h)}_{B(\mathbf{x})}(\mathbf{p}) - p^h x_1 x_2 \\ &= \sum_{\mathbf{p}\in\mathscr{P}(p^h)} \sum_{\ell_1=0}^{p^h-1} \sum_{\ell_2=0}^{p^h-1} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) W_{\mathbf{l}}(\mathbf{p}) - p^h \widetilde{\chi}_{\mathbf{0}}(\mathbf{x}) \\ &= \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{p^h-1} \sum_{\ell_2=0}^{p^h-1} \left(\sum_{\mathbf{p}\in\mathscr{P}(p^h)} W_{\mathbf{l}}(\mathbf{p}) \right) \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}), \end{aligned}$$

noting that

$$\sum_{\mathbf{p}\in\mathscr{P}(p^h)} W_{\mathbf{0}}(\mathbf{p}) = \#\mathscr{P}(p^h) = p^h.$$

It is well known in the theory of abelian groups that the sum

$$\sum_{\mathbf{p}\in\mathscr{P}(p^h)} W_{\mathbf{l}}(\mathbf{p}) \in \{0, p^h\}.$$

We therefore need to have some understanding on the set

$$L(h) = \left\{ \mathbf{l} \in [0,p^h) \times [0,p^h) : \mathbf{l} \neq \mathbf{0} \text{ and } \sum_{\mathbf{p}\in\mathscr{P}(p^h)} W_{\mathbf{l}}(\mathbf{p}) = p^h \right\}.$$

Then

$$D^{(h)}[\mathscr{P}(p^h); B(\mathbf{x})] = p^h \sum_{\mathbf{l}\in L(h)} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}).$$

We have the following special case of a general result of Skriganov [38].

Lemma 21. *Suppose that the prime p satisfies $p \geq 8$. Then the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l} \in L(h)$, are orthogonal, so that*

$$\int_{[0,1]^2} |D^{(h)}[\mathscr{P}(p^h); B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = p^{2h} \sum_{\mathbf{l}\in L(h)} \int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x}. \tag{113}$$

To progress further, we need to estimate each of the integrals

$$\int_{[0,1]^2} |\widetilde{\chi_{\mathbf{l}}}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} = \left(\int_0^1 |\widetilde{\chi_{\ell_1}}(x_1)|^2 \, \mathrm{d}x_1 \right) \left(\int_0^1 |\widetilde{\chi_{\ell_2}}(x_2)|^2 \, \mathrm{d}x_2 \right) \tag{114}$$

on the right hand side of (113).

Lemma 22. *We have*

$$\int_0^1 |\widetilde{\chi}_0(x)|^2 \, \mathrm{d}x = \frac{1}{4} + \frac{1}{4(p^2-1)} \sum_{j=1}^{p-1} \csc^2 \frac{\pi j}{p}. \tag{115}$$

Furthermore, for every $\ell \in \mathbf{N}$, we have

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2 \, \mathrm{d}x = p^{-2\rho(\ell)} \left(\frac{1}{2} \csc^2 \frac{\pi \lambda(\ell)}{p} - \frac{1}{4} + \frac{1}{4(p^2-1)} \sum_{j=1}^{p-1} \csc^2 \frac{\pi j}{p} \right), \tag{116}$$

where

$$\rho(\ell) = \begin{cases} 0, & \text{if } \ell = 0, \\ \max\{i \in \mathbf{N} : \lambda_i(\ell) \neq 0\}, & \text{if } \ell \in \mathbf{N}, \end{cases}$$

denotes the position of the leading coefficient of ℓ given by (110) and $\lambda(\ell) = \lambda_{\rho(\ell)}(\ell)$ denotes its value.

Proof. We have the Fine–Price formula, that for every $\ell \in \mathbf{N}_0$,

$$\widetilde{\chi}_\ell(x) = p^{-\rho(\ell)} u_\ell(x), \tag{117}$$

where

$$u_0(x) = \frac{1}{2} w_0(x) + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} w_{jp^{i-1}}(x), \tag{118}$$

and where for every $\ell \in \mathbf{N}$,

$$\begin{aligned} u_\ell(x) = {} & (1-\zeta^{\lambda(\ell)})^{-1} w_{\tau(\ell)}(x) + \left(\frac{1}{2} - (1-\zeta^{\lambda(\ell)})^{-1} \right) w_\ell(x) \\ & + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} w_{\ell + jp^{\rho(\ell)+i-1}}(x). \end{aligned} \tag{119}$$

Here $\tau(\ell) = \ell - \lambda(\ell) p^{\rho(\ell)-1}$, and $\zeta = \mathrm{e}^{2\pi \mathrm{i}/p}$ is a primitive p-th root of unity. For details, see Fine [21] and Price [32]. The right hand side of (119) is a linear combination of distinct Walsh functions. It follows that for every $\ell \in \mathbf{N}$, we have

$$\begin{aligned} \int_0^1 |u_\ell(x)|^2 \, \mathrm{d}x = {} & \frac{1}{(1-\zeta^{\lambda(\ell)})(1-\zeta^{-\lambda(\ell)})} + \left(\frac{1}{2} - \frac{1}{1-\zeta^{\lambda(\ell)}} \right) \left(\frac{1}{2} - \frac{1}{1-\zeta^{-\lambda(\ell)}} \right) \\ & + \sum_{i=1}^{\infty} p^{-2i} \sum_{j=1}^{p-1} |1-\zeta^j|^{-2} \end{aligned}$$

$$= 2|1-\zeta^{\lambda(\ell)}|^{-2} - \frac{1}{4} + \frac{1}{p^2-1}\sum_{j=1}^{p-1}|1-\zeta^j|^{-2}. \tag{120}$$

The identity (116) follows on combining (117) and (120) with the observation

$$|1-\zeta^j|^2 = \left(1-\cos\frac{2\pi j}{p}\right)^2 + \sin^2\frac{2\pi j}{p} = 4\sin^2\frac{\pi j}{p}. \tag{121}$$

Similarly, we have

$$\int_0^1 |u_0(x)|^2\,\mathrm{d}x = \frac{1}{4} + \sum_{i=1}^{\infty} p^{-2i}\sum_{j=1}^{p-1}|1-\zeta^j|^{-2} = \frac{1}{4} + \frac{1}{p^2-1}\sum_{j=1}^{p-1}|1-\zeta^j|^{-2}. \tag{122}$$

The identity (115) follows on combining (117), (121) and (122). □

Lemma 23. *For every $\ell \in \mathbf{N}_0$, we have*

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2\,\mathrm{d}x \le \frac{p^{2-2\rho(\ell)}}{4}.$$

Proof. Suppose first of all that $\ell \neq 0$. Then using the inequality that

$$\csc^2\frac{\pi j}{p} \le \frac{p^2}{4}$$

for every $j = 1,\ldots,p-1$, we see from (116) that

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2\,\mathrm{d}x \le p^{-2\rho(\ell)}\left(\frac{p^2}{8} + \frac{1}{4} + \frac{p^2(p-1)}{16(p^2-1)}\right) \le \frac{p^{2-2\rho(\ell)}}{4}.$$

On the other hand, it follows similarly from (115) that

$$\int_0^1 |\widetilde{\chi}_0(x)|^2\,\mathrm{d}x \le \frac{1}{4} + \frac{p^2(p-1)}{16(p^2-1)} \le \frac{p^2}{4} = \frac{p^{2-2\rho(0)}}{4}$$

as required. □

Combining (114) and Lemma 23, we conclude that

$$\int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \le \frac{p^{4-2\rho(\mathbf{l})}}{16},$$

where $\rho(\mathbf{l}) = \rho(\ell_1) + \rho(\ell_2)$. Thus we need to estimate the sum

$$\sum_{\mathbf{l}\in L(h)} p^{-2\rho(\mathbf{l})}. \tag{123}$$

Here $\rho(\mathbf{l})$ is a non-Hamming weight that arises from the Rosenblum–Tsfasman weight in coding theory. The idea here is that if the distribution dual to $\mathscr{P}(p^h)$ has sufficiently large Rosenblum–Tsfasman weight, then we can obtain a good estimate for the sum (123).

For a brief discussion on how we may complete our proof, the reader is referred to the paper of Chen and Skriganov [14].

16 Explicit Constructions and Orthogonality

The first proof of Theorem 7 for arbitrary $k \geq 2$ by Roth [35] is probabilistic in nature, as are the subsequent proofs by Chen [11] and Skriganov [37]. The disadvantage of such probabilistic arguments is that while we can show that a good point set exists, we cannot describe it explicitly.

On the other hand, the proof by Davenport [19] of Theorem 7 in dimension $k = 2$ is not probabilistic in nature, and one can describe the point set explicitly. However, finding explicit constructions in dimensions $k \geq 3$ turns out to be rather hard. Its eventual solution by Chen and Skriganov [13] is based on the observation that provided that the prime p is sufficiently large, then the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l} \in L(h)$, are *quasi-orthogonal*, so that some weaker version of Lemma 21 in arbitrary dimensions holds.

However, if we are not able to establish any orthogonality or quasi-orthogonality, then our techniques thus far fail to give any explicit constructions in dimensions $k \geq 3$. To establish an appropriate upper bound, we may resort to digit shifts, and our argument is underpinned by the general result below for arbitrary dimensions $k \geq 2$ for some suitably defined Walsh function $W_{\mathbf{l}}(\mathbf{t})$.

Lemma 24. *For every* $\mathbf{l}', \mathbf{l}'' \in \mathbf{N}_0^k$, *we have*

$$\sum_{\mathbf{t} \in \mathbf{Z}_p^{kh}} W_{\mathbf{l}'}(\mathbf{t}) W_{\mathbf{l}''}(\mathbf{t}) = \begin{cases} p^{kh}, & \text{if } \mathbf{l}' = \mathbf{l}'', \\ 0, & \text{otherwise}. \end{cases}$$

This result can be viewed as an orthogonality result. We may therefore conclude that orthogonality or quasi-orthogonality in some form is central to our upper bound arguments here, whether we consider explicit constructions or otherwise.

References

1. van Aardenne-Ehrenfest, T.: Proof of the impossibility of a just distribution of an infinite sequence of points over an interval. Proc. Kon. Ned. Akad. v. Wetensch. **48**, 266–271 (1945)
2. van Aardenne-Ehrenfest, T.: On the impossibility of a just distribution. Proc. Kon. Ned. Akad. v. Wetensch. **52**, 734–739 (1949)
3. Beck, J.: Roth's estimate of discrepancy of integer sequences is nearly sharp. Combinatorica **1**, 319–325 (1981)

4. Beck, J.: Irregularities of distribution I. Acta Math. **159**, 1–49 (1987)
5. Beck, J., Chen, W.W.L.: Irregularities of Distribution. Cambridge Tracts in Mathematics **89**, Cambridge University Press, Cambridge (1987)
6. Beck, J., Chen, W.W.L.: Note on irregularities of distribution II. Proc. London Math. Soc. **61**, 251–272 (1990)
7. Beck, J., Chen, W.W.L.: Irregularities of point distribution relative to convex polygons III. J. London Math. Soc. **56**, 222–230 (1997)
8. Bilyk, D., Lacey, M.T.: On the small ball inequality in three dimensions. Duke Math. J. **143**, 81–115 (2008)
9. Bilyk, D., Lacey, M.T., Vagharshakyan, A.: On the small ball inequality in all dimensions. J. Funct. Anal. **254**, 2470–2502 (2008)
10. Chen, W.W.L.: On irregularities of distribution. Mathematika **27**, 153–170 (1980)
11. Chen, W.W.L.: On irregularities of distribution II. Q. J. Math. **34**, 257–279 (1983)
12. Chen, W.W.L.: Fourier techniques in the theory of irregularities of point distribution. In: Brandolini, L., Colzani, L., Iosevich, A., Travaglini, G. (eds.) Fourier Analysis and Convexity, pp. 59–82. Applied and Numerical Harmonic Analysis, Birkhäuser, Boston (2004)
13. Chen, W.W.L., Skriganov, M.M.: Explicit constructions in the classical mean squares problem in irregularities of point distribution. J. reine angew. Math. **545**, 67–95 (2002)
14. Chen, W.W.L., Skriganov, M.M.: Orthogonality and digit shifts in the classical mean squares problem in irregularities of point distribution. In: Schlickewei, H.P., Schmidt, K., Tichy, R.F. (eds.) Diophantine Approximation: Festschrift for Wolfgang Schmidt, pp. 141–159. Developments in Mathematics **16**, Springer, Wien (2008)
15. Chen, W.W.L., Travaglini, G.: Discrepancy with respect to convex polygons. J. Complexity **23**, 662–672 (2007)
16. Chen, W.W.L., Travaglini, G.: Deterministic and probabilistic discrepancies. Ark. Mat. **47**, 273-293 (2009)
17. van der Corput, J.G.: Verteilungsfunktionen I. Proc. Kon. Ned. Akad. v. Wetensch. **38**, 813–821 (1935)
18. van der Corput, J.G.: Verteilungsfunktionen II. Proc. Kon. Ned. Akad. v. Wetensch. **38**, 1058–1066 (1935)
19. Davenport, H.: Note on irregularities of distribution. Mathematika **3**, 131–135 (1956)
20. Faure, H.: Discrépance de suites associées à un système de numération (en dimension *s*). Acta Arith. **41**, 337–351 (1982)
21. Fine, N.J.: On the Walsh functions. Trans. Amer. Math. Soc. **65**, 373–414 (1949)
22. Halász, G.: On Roth's method in the theory of irregularities of point distributions. In: Halberstam, H., Hooley, C. (eds.) Recent Progress in Analytic Number Theory, volume 2, pp. 79–94. Academic Press, London (1981)
23. Halton, J.H.: On the efficiency of certain quasirandom sequences of points in evaluating multidimensional integrals. Numer. Math. **2**, 84–90 (1960)
24. Hardy, G.H., Littlewood, J.E.: The lattice points of a right angled triangle I. Proc. London Math. Soc. **3**, 15–36 (1920)
25. Hardy, G.H., Littlewood, J.E.: The lattice points of a right angled triangle II. Abh. Math. Sem. Hamburg **1**, 212–249 (1922)
26. Halton, J.H., Zaremba, S.K.: The extreme and L^2 discrepancies of some plane sets. Monatsh. Math. **73**, 316–328 (1969)
27. Konyagin, S.V., Skriganov, M.M., Sobolev, A.V.: On a lattice point problem arising in the spectral analysis of periodic operators. Mathematika **50**, 87–98 (2003)
28. Lerch, M.: Question 1547. L'Intermediaire Math. **11**, 144–145 (1904)
29. Lidl, R., Niederreiter, H.: Finite Fields. Addison-Wesley, Reading (1983)
30. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error Correcting Codes. North-Holland, Amsterdam (1977)
31. Pollard, D.: Convergence of Stochastic Processes. Springer, New York (1984)
32. Price, J.J.: Certain groups of orthonormal step functions. Canadian J. Math. **9**, 413–425 (1957)
33. Roth, K.F.: On irregularities of distribution. Mathematika **1**, 73–79 (1954)

34. Roth, K.F.: On irregularities of distribution III. Acta Arith. **35**, 373–384 (1979)
35. Roth, K.F.: On irregularities of distribution IV. Acta Arith. **37**, 67–75 (1980)
36. Schmidt, W.M.: Irregularities of distribution VII. Acta Arith. **21**, 45–50 (1972)
37. Skriganov, M.M.: Constructions of uniform distributions in terms of geometry of numbers. Algebra i Analiz **6**, 200–230 (1994); St. Petersburg Math. J. **6**, 635–664 (1995)
38. Skriganov, M.M.: Harmonic analysis on totally disconnected groups and irregularities of point distributions. J. reine angew. Math. **600**, 25–49 (2006)