CHAPTER 1

Uniform Distribution

1.1. Introduction

Throughout, $I = [0,1)$ denotes the unit interval. For any set E, χ_E denotes its characteristic function.

Suppose that $(s_i)_{i\in\mathbb{N}}$ is a sequence of real numbers in the interval I. For every natural number n and any subset E of I, write

$$Z(E,n) = \sum_{i=1}^{n} \chi_E(s_i);$$

in other words, $Z(E,n)$ counts the number of terms among $s_1,\ldots,s_n$ that lie in the set E. The real sequence $(s_i)_{i\in\mathbb{N}}$ in I is said to be uniformly distributed in I if for every $\alpha,\beta\in\mathbb{R}$ satisfying $0\leqslant\alpha<\beta\leqslant 1$, we have

$$\lim_{n\to\infty} \frac{Z([\alpha,\beta),n)}{n} = \beta-\alpha. \tag{1.1}$$

Put simply, every subinterval of I should have its *fair share* of the terms of the sequence $(s_i)_{n\in\mathbb{N}}$.

Example. Let θ be a fixed real number. Consider the sequence $(s_i)_{i\in\mathbb{N}}$, where $s_i = \{i\theta\}$ for every $i\in\mathbb{N}$. Suppose first of all that θ is rational. Write $\theta = p/q$, where $p\in\mathbb{Z}$ and $q\in\mathbb{N}$. For every $i\in\mathbb{N}$, $\{i\theta\}$ is clearly an integer multiple of $1/q$, so that $Z([1/2q,1/q),n) = 0$ for every $n\in\mathbb{N}$. It follows that the sequence $(\{i\theta\})_{i\in\mathbb{N}}$ is not uniformly distributed in I if θ is rational. We shall show later that the sequence $(\{i\theta\})_{i\in\mathbb{N}}$ is uniformly distributed in I if θ is irrational.

Remark. Note that the real sequence $(s_i)_{i\in\mathbb{N}}$ in I is uniformly distributed in I if and only if for every $\alpha\in\mathbb{R}$ satisfying $0<\alpha\leqslant 1$, we have

$$\lim_{n\to\infty} \frac{Z([0,\alpha),n)}{n} = \alpha.$$

Suppose now that the real sequence $(s_i)_{i\in\mathbb{N}}$ in I is uniformly distributed in I. Let $f:[0,1]\to\mathbb{R}$ be a continuous function. Suppose that the natural number n is *large*. Then one may expect that the *discrete average*

$$\frac{1}{n}\sum_{i=1}^{n} f(s_i)$$

of the function f over the first n terms of the sequence $(s_i)_{i\in\mathbb{N}}$ may not differ substantially from the *continuous average*

$$\int_0^1 f(x)\,\mathrm{d}x$$

of the function f over the interval $[0,1]$. This simple observation leads us to the following characterization of sequences $(s_i)_{i\in\mathbb{N}}$ in I that are uniformly distributed in I.

THEOREM 1.1. *For any real sequence $(s_i)_{i\in\mathbb{N}}$ in I, the following statements are equivalent:*

(i) *The sequence $(s_i)_{i\in\mathbb{N}}$ is uniformly distributed in I.*
(ii) *For every step function $g:[0,1]\to\mathbb{R}$, we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} g(s_i)=\int_0^1 g(x)\,\mathrm{d}x.$$

(iii) *For every continuous function $f:[0,1]\to\mathbb{R}$ satisfying $f(0)=f(1)$, we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f(s_i)=\int_0^1 f(x)\,\mathrm{d}x.$$

REMARK. In (iii), it is not really necessary to impose the restriction $f(0)=f(1)$. However, we have made this specification here in order to facilitate our discussion of Weyl's criterion. We shall elaborate on this comment later.

We shall prove Theorem 1.1 by showing that (ii) follows from (i), (iii) follows from (ii), and (i) follows from (iii).

The first of these three steps is very simple. Note that (1.1) is equivalent to

$$\lim_{n\to\infty}\frac{Z([\alpha,\beta),n)}{n}=\int_0^1 \chi_{[\alpha,\beta)}(x)\,\mathrm{d}x. \tag{1.2}$$

Suppose that (i) holds. Let $g:[0,1]\to\mathbb{R}$ be a step function. By changing a finite number of values of g if necessary, we may assume that

$$g(x)=\sum_{j=1}^{k} c_j\chi_{[\alpha_{j-1},\alpha_j)}(x),$$

where $0=\alpha_0<\alpha_1<\ldots<\alpha_k=1$ is a dissection of $[0,1]$ and where $c_1,\ldots,c_k$ are real constants. Then as $n\to\infty$, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} g(s_i)&=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} c_j\chi_{[\alpha_{j-1},\alpha_j)}(s_i)=\sum_{j=1}^{k} c_j\left(\frac{1}{n}\sum_{i=1}^{n}\chi_{[\alpha_{j-1},\alpha_j)}(s_i)\right)\\&=\sum_{j=1}^{k} c_j\frac{Z([\alpha_{j-1},\alpha_j),n)}{n}\to\sum_{j=1}^{k} c_j\int_0^1\chi_{[\alpha_{j-1},\alpha_j)}(x)\,\mathrm{d}x\\&=\int_0^1\left(\sum_{j=1}^{k} c_j\chi_{[\alpha_{j-1},\alpha_j)}(x)\right)\mathrm{d}x=\int_0^1 g(x)\,\mathrm{d}x.\end{aligned}$$

This gives (ii).

However, the other two steps of the proof require ideas concerning approximation.

To show that (iii) follows from (ii), we use the idea of approximation by Riemann sums. Suppose that the function $f:[0,1]\to\mathbb{R}$ is continuous. Then it is Riemann integrable over $[0,1]$. It follows that given any $\epsilon > 0$, there exists a dissection $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_k = 1$ of $[0,1]$ such that

$$\sum_{j=1}^{k}(\alpha_j - \alpha_{j-1})\left(\sup_{x\in[\alpha_{j-1},\alpha_j]} f(x) - \inf_{x\in[\alpha_{j-1},\alpha_j]} f(x)\right) < \frac{\epsilon}{2}.$$

In other words, there exist two step functions $g_1 : [0,1] \to \mathbb{R}$ and $g_2 : [0,1] \to \mathbb{R}$ such that

$$g_1(x) \leqslant f(x) \leqslant g_2(x) \quad \text{for every } x \in [0,1], \tag{1.3}$$

and

$$\int_0^1 (g_2(x) - g_1(x))\,\mathrm{d}x < \frac{\epsilon}{2}. \tag{1.4}$$

For every $n \in \mathbb{N}$, the inequalities in (1.3) clearly lead to

$$\frac{1}{n}\sum_{i=1}^{n} g_1(s_i) \leqslant \frac{1}{n}\sum_{i=1}^{n} f(s_i) \leqslant \frac{1}{n}\sum_{i=1}^{n} g_2(s_i). \tag{1.5}$$

Suppose that (ii) holds. Then there exists $n_0 \in \mathbb{N}$ such that for every $n > n_0$, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n} g_1(s_i) - \int_0^1 g_1(x)\,\mathrm{d}x\right| < \frac{\epsilon}{2} \quad \text{and} \quad \left|\frac{1}{n}\sum_{i=1}^{n} g_2(s_i) - \int_0^1 g_2(x)\,\mathrm{d}x\right| < \frac{\epsilon}{2},$$

so that

$$\frac{1}{n}\sum_{i=1}^{n} g_1(s_i) > \int_0^1 g_1(x)\,\mathrm{d}x - \frac{\epsilon}{2} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} g_2(s_i) < \int_0^1 g_2(x)\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Combining these with (1.5), we conclude that for every $n > n_0$, we have

$$\int_0^1 g_1(x)\,\mathrm{d}x - \frac{\epsilon}{2} < \frac{1}{n}\sum_{i=1}^{n} f(s_i) < \int_0^1 g_2(x)\,\mathrm{d}x + \frac{\epsilon}{2}. \tag{1.6}$$

On the other hand, it clearly follows from (1.3) and (1.4) that

$$\int_0^1 g_1(x)\,\mathrm{d}x > \int_0^1 f(x)\,\mathrm{d}x - \frac{\epsilon}{2} \quad \text{and} \quad \int_0^1 g_2(x)\,\mathrm{d}x < \int_0^1 f(x)\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Combining these with (1.6), we conclude that for every $n > n_0$, we have

$$\int_0^1 f(x)\,\mathrm{d}x - \epsilon < \frac{1}{n}\sum_{i=1}^{n} f(s_i) < \int_0^1 f(x)\,\mathrm{d}x + \epsilon,$$

so that

$$\left|\frac{1}{n}\sum_{i=1}^{n} f(s_i) - \int_0^1 f(x)\,\mathrm{d}x\right| < \epsilon$$

as required. (iii) follows.

To show that (i) follows from (iii), we consider approximation of characteristic functions by piecewise linear continuous functions.

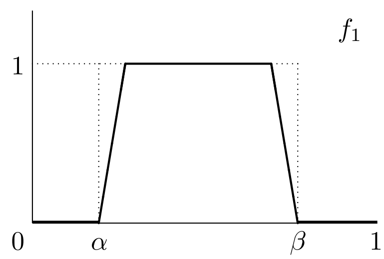
Suppose first of all that $0 < \alpha < \beta < 1$. Let $\epsilon > 0$ be given. We may clearly assume that $\epsilon < \max\{\alpha, \beta-\alpha, 1-\beta\}$. We define continuous functions $f_1 : [0,1] \to \mathbb{R}$ and $f_2 : [0,1] \to \mathbb{R}$ by writing

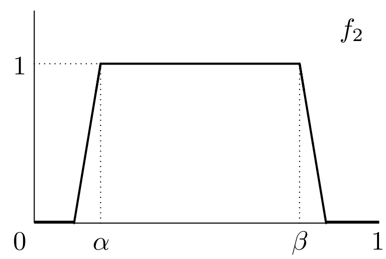
$$f_1(x) = \begin{cases} 0, & \text{if } 0 \leqslant x \leqslant \alpha, \\ 4(x-\alpha)/\epsilon, & \text{if } \alpha \leqslant x \leqslant \alpha + \epsilon/4, \\ 1, & \text{if } \alpha + \epsilon/4 \leqslant x \leqslant \beta - \epsilon/4, \\ -4(x-\beta)/\epsilon, & \text{if } \beta - \epsilon/4 \leqslant x \leqslant \beta, \\ 0, & \text{if } \beta \leqslant x \leqslant 1, \end{cases}$$

and

$$f_2(x) = \begin{cases} 0, & \text{if } 0 \leqslant x \leqslant \alpha - \epsilon/4, \\ 4(x-\alpha+\epsilon/4)/\epsilon, & \text{if } \alpha - \epsilon/4 \leqslant x \leqslant \alpha, \\ 1, & \text{if } \alpha \leqslant x \leqslant \beta, \\ -4(x-\beta-\epsilon/4)/\epsilon, & \text{if } \beta \leqslant x \leqslant \beta + \epsilon/4, \\ 0, & \text{if } \beta + \epsilon/4 \leqslant x \leqslant 1. \end{cases}$$

Note that both functions are piecewise linear.





Furthermore, it is easy to see that

$$f_1(x) \leqslant \chi_{[\alpha,\beta)} \leqslant f_2(x) \quad \text{for every } x \in [0,1], \tag{1.7}$$

and that

$$\int_0^1 (f_2(x) - f_1(x))\,\mathrm{d}x = \frac{\epsilon}{2}. \tag{1.8}$$

It now follows from (1.7) that for every $n \in \mathbb{N}$, we have

$$\frac{1}{n}\sum_{i=1}^n f_1(s_i) \leqslant \frac{1}{n}\sum_{i=1}^n \chi_{[\alpha,\beta)}(s_i) \leqslant \frac{1}{n}\sum_{i=1}^n f_2(s_i). \tag{1.9}$$

On the other hand, clearly $f_1(0) = f_1(1)$ and $f_2(0) = f_2(1)$, and so it follows from (iii) that there exists $n_0 \in \mathbb{N}$ such that for every $n > n_0$, we have

$$\left|\frac{1}{n}\sum_{i=1}^n f_1(s_i) - \int_0^1 f_1(x)\,\mathrm{d}x\right| < \frac{\epsilon}{2} \quad \text{and} \quad \left|\frac{1}{n}\sum_{i=1}^n f_2(s_i) - \int_0^1 f_2(x)\,\mathrm{d}x\right| < \frac{\epsilon}{2},$$

so that

$$\frac{1}{n}\sum_{i=1}^n f_1(s_i) > \int_0^1 f_1(x)\,\mathrm{d}x - \frac{\epsilon}{2} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^n f_2(s_i) < \int_0^1 f_2(x)\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Combining these with (1.9), we conclude that for every $n > n_0$, we have

$$\int_0^1 f_1(x)\,\mathrm{d}x - \frac{\epsilon}{2} < \frac{1}{n}\sum_{i=1}^n \chi_{[\alpha,\beta)}(s_i) < \int_0^1 f_2(x)\,\mathrm{d}x + \frac{\epsilon}{2}. \tag{1.10}$$

On the other hand, it clearly follows from (1.7) and (1.8) that

$$\int_0^1 f_1(x)\,\mathrm{d}x \geqslant \int_0^1 \chi_{[\alpha,\beta)}(x)\,\mathrm{d}x - \frac{\epsilon}{2} \quad \text{and} \quad \int_0^1 f_2(x)\,\mathrm{d}x \leqslant \int_0^1 \chi_{[\alpha,\beta)}(x)\,\mathrm{d}x + \frac{\epsilon}{2}.$$

Combining these with (1.10), we conclude that for every $n > n_0$, we have

$$\begin{aligned}(\beta-\alpha)-\epsilon = \int_0^1 \chi_{[\alpha,\beta)}(x)\,\mathrm{d}x - \epsilon &< \frac{1}{n}\sum_{i=1}^n \chi_{[\alpha,\beta)}(s_i) \\ &< \int_0^1 \chi_{[\alpha,\beta)}(x)\,\mathrm{d}x + \epsilon = (\beta-\alpha)+\epsilon,\end{aligned}$$

so that

$$\left|\frac{1}{n}\sum_{i=1}^n \chi_{[\alpha,\beta)}(s_i) - (\beta-\alpha)\right| < \epsilon$$

as required. (i) follows.

Suppose next that $\alpha = 0$ and $0 < \beta < 1$. As before, we may clearly assume that $\epsilon < \max\{\beta, 1-\beta\}$. In this case, it can be shown that the continuous functions $f_1 : [0,1] \to \mathbb{R}$ and $f_2 : [0,1] \to \mathbb{R}$, given by

$$f_1(x) = \begin{cases} 4x/\epsilon, & \text{if } 0 \leqslant x \leqslant \epsilon/4, \\ 1, & \text{if } \epsilon/4 \leqslant x \leqslant \beta - \epsilon/4, \\ -4(x-\beta)/\epsilon, & \text{if } \beta - \epsilon/4 \leqslant x \leqslant \beta, \\ 0, & \text{if } \beta \leqslant x \leqslant 1, \end{cases}$$

and

$$f_2(x) = \begin{cases} 1, & \text{if } 0 \leqslant x \leqslant \beta, \\ -4(x-\beta-\epsilon/4)/\epsilon, & \text{if } \beta \leqslant x \leqslant \beta + \epsilon/4, \\ 0, & \text{if } \beta + \epsilon/4 \leqslant x \leqslant 1 - \epsilon/4, \\ 4(x-1+\epsilon/4)/\epsilon, & \text{if } 1 - \epsilon/4 \leqslant x \leqslant 1, \end{cases}$$

will do.

The case $0 < \alpha < 1$ and $\beta = 1$ is left as a simple exercise.

1.2. Weyl's Criterion

Although Theorem 1.1 gives a nice characterization of uniform distribution in terms of continuous real valued functions on $[0,1]$, it is completely useless as it stands, for it is clearly impossible to estimate the discrete average

$$\frac{1}{n}\sum_{i=1}^n f(s_i)$$

for every continuous function $f : [0,1] \to \mathbb{R}$ satisfying $f(0) = f(1)$. It follows that we must seek to characterize uniform distribution in terms of special collections of functions f for which we are better able to handle the discrete averages that arise. The great achievement of Weyl is that he was able to look beyond the collection of real valued functions for such a special collection. Indeed, he chose a suitable collection of exponential functions, and Weyl's criterion is unquestionably the greatest result in the theory of uniform distribution.

THEOREM 1.2 (Weyl's criterion). *A real sequence $(s_i)_{i\in\mathbb{N}}$ in I is uniformly distributed in I if and only if for every non-zero integer h, we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}e(hs_i)=0.$$

We shall divide the proof of Weyl's criterion into two steps, summarized below as Theorems 1.3 and 1.4.

The first step of the proof is one of extension. We shall replace the collection of real valued functions in Theorem 1.1 by a collection of complex valued functions.

THEOREM 1.3. *A real sequence $(s_i)_{i\in\mathbb{N}}$ in I is uniformly distributed in I if and only if for every continuous function $f:[0,1]\to\mathbb{C}$ satisfying $f(0)=f(1)$, we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}f(s_i)=\int_0^1 f(x)\,\mathrm{d}x.$$

PROOF. We simply note that any continuous function $f:[0,1]\to\mathbb{C}$ satisfying $f(0)=f(1)$ can be written in the form $f=f_1+\mathrm{i}f_2$, where $f_1:[0,1]\to\mathbb{R}$ and $f_2:[0,1]\to\mathbb{R}$ are continuous and satisfy $f_1(0)=f_1(1)$ and $f_2(0)=f_2(1)$. The result now follows from Theorem 1.1. ◯

The major step in the proof of Weyl's criterion is one of reduction. Here we shall replace the collection of complex valued functions in Theorem 1.3 by Weyl's collection of exponential functions, clearly a subcollection of all continuous functions $f:[0,1]\to\mathbb{C}$ satisfying $f(0)=f(1)$. Note that for every non-zero integer h, we have $\int_0^1 e(hx)\,\mathrm{d}x=0$.

THEOREM 1.4. *Suppose that $(s_i)_{i\in\mathbb{N}}$ is a real sequence in I. Suppose further that for every non-zero integer h, we have*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}e(hs_i)=0.$$

Then for every continuous function $f:[0,1]\to\mathbb{C}$ satisfying $f(0)=f(1)$, we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}f(s_i)=\int_0^1 f(x)\,\mathrm{d}x.$$

PROOF. The technical result that underpins our proof is the Weierstrass approximation theorem. Suppose that $f:[0,1]\to\mathbb{C}$ is continuous and satisfies $f(0)=f(1)$. Then given any $\epsilon>0$, there exists a trigonometric polynomial $p:[0,1]\to\mathbb{C}$, *i.e.* a linear combination of functions of the type $e(hx)$ where $h\in\mathbb{Z}$, such that

$$\sup_{x\in[0,1]}|f(x)-p(x)|<\frac{\epsilon}{3}. \tag{1.11}$$

Note first of all that

$$\left|\frac{1}{n}\sum_{i=1}^{n}f(s_i)-\int_0^1 f(x)\,\mathrm{d}x\right|\leqslant\left|\frac{1}{n}\sum_{i=1}^{n}f(s_i)-\frac{1}{n}\sum_{i=1}^{n}p(s_i)\right| \tag{1.12}$$
$$+\left|\frac{1}{n}\sum_{i=1}^{n}p(s_i)-\int_0^1 p(x)\,\mathrm{d}x\right|+\left|\int_0^1 p(x)\,\mathrm{d}x-\int_0^1 f(x)\,\mathrm{d}x\right|.$$

In view of (1.11), clearly the first and last terms on the right hand side of (1.12) are each less than $\epsilon/3$. It therefore remains to show that for all sufficiently large n, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n} p(s_i) - \int_0^1 p(x)\,\mathrm{d}x\right| < \frac{\epsilon}{3}.$$

Suppose that

$$p(x) = c_0 + \sum_{j=1}^{k} c_j e(h_j x), \tag{1.13}$$

where $c_0 \in \mathbb{C}$, $c_1, \ldots, c_k \in \mathbb{C} \setminus \{0\}$ and $h_1, \ldots, h_k \in \mathbb{Z} \setminus \{0\}$. Then clearly

$$\frac{1}{n}\sum_{i=1}^{n} p(s_i) - \int_0^1 p(x)\,\mathrm{d}x = \sum_{j=1}^{k} c_j \left(\frac{1}{n}\sum_{i=1}^{n} e(h_j s_i) - \int_0^1 e(h_j x)\,\mathrm{d}x\right).$$

Furthermore, we have $\int_0^1 e(h_j x)\,\mathrm{d}x = 0$ for every $j = 1, \ldots, k$, so that

$$\left|\frac{1}{n}\sum_{i=1}^{n} p(s_i) - \int_0^1 p(x)\,\mathrm{d}x\right| \leqslant \sum_{j=1}^{k} |c_j| \left|\frac{1}{n}\sum_{i=1}^{n} e(h_j s_i)\right|. \tag{1.14}$$

For every $j = 1, \ldots, k$, it follows from the hypotheses that there exists $n_j \in \mathbb{N}$ such that for every $n > n_j$, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n} e(h_j s_i)\right| < \frac{\epsilon}{3k|c_j|}. \tag{1.15}$$

Let $n_0 = \max\{n_1, \ldots, n_k\}$. Then for every $n > n_0$, in view of (1.14) and (1.15), we have

$$\left|\frac{1}{n}\sum_{i=1}^{n} p(s_i) - \int_0^1 p(x)\,\mathrm{d}x\right| < \frac{\epsilon}{3}$$

as required. ○

REMARK. In Theorem 1.1, we impose the restriction $f(0) = f(1)$ in (iii). This is essential in the proof of Theorem 1.4, as the Weierstrass approximation theorem would not otherwise be applicable. This is clear, since any function of the type (1.13) satisfies $p(0) = p(1)$.

EXAMPLE. Let θ be a fixed real number. The sequence $(\{i\theta\})_{i\in\mathbb{N}}$ is uniformly distributed in I if θ is irrational. To prove this, we shall use Weyl's criterion. For every non-zero integer h, we have

$$\begin{aligned}\left|\frac{1}{n}\sum_{i=1}^{n} e(h\{i\theta\})\right| &= \left|\frac{1}{n}\sum_{i=1}^{n} e(hi\theta)\right| = \left|\frac{1}{n}\frac{e(h(n+1)\theta) - e(h\theta)}{1 - e(h\theta)}\right| \\ &\leqslant \frac{2}{n|1 - e(h\theta)|} = \frac{1}{n|\sin \pi h\theta|} \to 0\end{aligned}$$

as $n \to \infty$, since $\sin \pi h\theta \neq 0$.

Remark. Note that in Weyl's criterion, we have used the exponential functions, and our proof is underpinned by the Weierstrass approximation theorem. Other variants of Weyl's criterion using collections other than the exponential functions can also be established, so long as the proof is supported by a suitable analogue of the Weierstrass approximation theorem for those collections.

1.3. Further Comments

There are a number of criteria for uniform distribution. We may, for instance, confront the problem directly and attempt to estimate the exponential sums that arise from Weyl's criterion. An early example goes back to van der Corput. The estimation of exponential sums, however, is an extremely vast subject, and it is not our purpose here to discuss such problems. Alternatively, we may devise some indirect applications of Weyl's criterion; in other words, we may devise tests for uniform distribution which ultimately depend on Weyl's criterion. An early example of such an approach is due to Fejér. Another example concerns difference theorems due to van der Corput, for instance.

We may also consider metric results. Here we consider sequences of the type $(u_i(x))_{i\in\mathbb{N}}$, where the parameter x lies in some given interval J which may be bounded or unbounded. The sequence may be uniformly distributed modulo 1 for some values $x \in J$ and not so for other values $x \in J$. We are interested in cases where the exceptional set is of Lebesgue measure 0; in other words, we are interested in results where the sequence is uniformly distributed modulo 1 for almost all values $x \in J$. Some early results of this type are due to Weyl and to Koksma.

CHAPTER 2

The Classical Discrepancy Problem

2.1. Introduction

We shall first of all show that the notion of uniform distribution is a relatively weak one. Recall that a real sequence $(s_i)_{i\in\mathbb{N}}$ in $I=[0,1)$ is uniformly distributed in I if for every $\alpha\in\mathbb{R}$ satisfying $0<\alpha\leqslant 1$, we have

$$\lim_{n\to\infty}\frac{Z([0,\alpha),n)}{n}=\alpha,$$

where $Z(E,n)$ counts the number of terms among $s_1,\ldots,s_n$ that lie in a set E, so that the counting function $Z([0,\alpha),n)$, as a function of n, is asymptotic to its expectation $n\alpha$ in the limit.

Instead of studying the ratio of these two functions, let us consider instead their difference, and study the discrepancy

$$D([0,\alpha),n)=Z([0,\alpha),n)-n\alpha$$

of the truncated sequence $s_1,\ldots,s_n$. Then it is easy to see that a real sequence $(s_i)_{i\in\mathbb{N}}$ in I is uniformly distributed in I if and only if, for every $\alpha\in\mathbb{R}$ satisfying $0<\alpha\leqslant 1$, we have

$$D([0,\alpha),n)=o(n)\quad\text{as } n\to\infty. \tag{2.1}$$

We shall show that there are sequences $(s_i)_{i\in\mathbb{N}}$ in I that satisfy bounds that are substantially sharper than (2.1). Equally importantly, we shall also exhibit limitations to such sharpness by showing that in some sense, every sequence $(s_i)_{i\in\mathbb{N}}$ in I has some minimal irregularity. This study of irregularities of distribution can be viewed as a quantitative version of uniform distribution.

2.2. History of the Problem and Roth's Reformulation

We have shown earlier that the sequence $(\{i\theta\})_{i\in\mathbb{N}}$ of fractional parts is uniformly distributed in I for every fixed irrational number $\theta\in\mathbb{R}$. In fact, it can be shown that in the case when $\theta=\sqrt{2}$, the sequence $(\{i\sqrt{2}\})_{i\in\mathbb{N}}$ satisfies the bound

$$D([0,\alpha),n)=O(\log n)\quad\text{as } n\to\infty, \tag{2.2}$$

for every fixed $\alpha\in\mathbb{R}$ satisfying $0<\alpha\leqslant 1$.

It can also be shown that the famous van der Corput sequence, which we shall define later, satisfies a similar bound.

These examples raise the question of whether the bound (2.2) can be improved. The following conjecture of van der Corput can be found in his work on distribution functions.

Conjecture (van der Corput 1935). *Suppose that* $(s_i)_{i\in\mathbb{N}}$ *is a real sequence in* $I = [0,1)$. *Corresponding to any arbitrarily large real number* κ, *there exist a positive integer* n *and two subintervals* I_1 *and* I_2, *of equal length, of* I *such that*

$$|Z(I_1,n) - Z(I_2,n)| > \kappa.$$

In short, this conjecture expresses the fact that no sequence can, in a certain sense, be too evenly distributed.

This conjecture is true, as shown by van Aardenne-Ehrenfest in 1945. Indeed, we have the following refinement.

Theorem 2.1 (van Aardenne-Ehrenfest 1949). *Suppose that* $(s_i)_{i\in\mathbb{N}}$ *is a real sequence in* $I = [0,1)$. *Suppose further that* $N \in \mathbb{N}$ *is sufficiently large. Then*

$$\sup_{\substack{1\leqslant n\leqslant N\\ 0<\alpha\leqslant 1}} |D([0,\alpha),n)| \gg \frac{\log\log N}{\log\log\log N}. \tag{2.3}$$

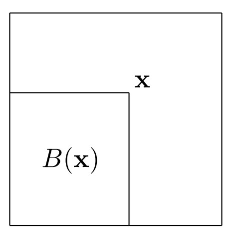
This result immediately raises the question of which functions $f(N)$ satisfy the following assertion.

Assertion A. *For every real sequence* $(s_i)_{i\in\mathbb{N}}$ *in* $I = [0,1)$ *and every* $N \in \mathbb{N}$, *we have*

$$\sup_{\substack{1\leqslant n\leqslant N\\ 0<\alpha\leqslant 1}} |D([0,\alpha),n)| \gg f(N). \tag{2.4}$$

Next, we consider Roth's formulation of the problem in 1954.

Suppose that $\mathcal{P}$ is a distribution of N points in the unit square $[0,1]^2$. For every aligned rectangle $B(\mathbf{x}) = [0,x_1) \times [0,x_2)$, where $\mathbf{x} = (x_1,x_2) \in [0,1]^2$,



let $Z[\mathcal{P};B(\mathbf{x})]$ denote the number of points of $\mathcal{P}$ that fall into $B(\mathbf{x})$, and consider the discrepancy

$$D[\mathcal{P};B(\mathbf{x})] = Z[\mathcal{P};B(\mathbf{x})] - Nx_1x_2, \tag{2.5}$$

noting that Nx_1x_2 represents the expected number of points of $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x})$.

We now consider the corresponding question of which functions $g(N)$ satisfy the following assertion.

Assertion B. *For every distribution* $\mathcal{P}$ *of* N *points in the unit square* $[0,1]^2$, *we have*

$$\sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \gg g(N). \tag{2.6}$$

In fact, the two assertions are equivalent, as shown by Roth in 1954. More precisely, if $f(N)$ satisfies the inequality (2.4), then the inequality (2.6) holds with $g(N) = f(N)$. Similarly, if $g(N)$ satisfies the inequality (2.6), then the inequality (2.4) holds with $f(N) = g(N)$.

The following astonishing result represents arguably Roth's greatest work.

THEOREM 2.2 (Roth 1954). *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \gg \log N. \tag{2.7}$$

The following two results are immediate consequences.

THEOREM 2.3. *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \gg (\log N)^{\frac{1}{2}}. \tag{2.8}$$

THEOREM 2.4. *For every real sequence $(s_i)_{i\in\mathbb{N}}$ in $I = [0,1)$ and every $N \in \mathbb{N}$, we have*

$$\sup_{\substack{1\leqslant n\leqslant N\\ 0<\alpha\leqslant 1}} |D([0,\alpha),n)| \gg (\log N)^{\frac{1}{2}}.$$

Theorem 2.4, having been obtained as a consequence of Theorem 2.2, is not best possible. We have a substantially stronger result.

THEOREM 2.5 (Schmidt 1972). *For every real sequence $(s_i)_{i\in\mathbb{N}}$ in $I = [0,1)$ and every $N \in \mathbb{N}$, we have*

$$\sup_{\substack{1\leqslant n\leqslant N\\ 0<\alpha\leqslant 1}} |D([0,\alpha),n)| \gg \log N.$$

This leads immediately to the following corresponding result.

THEOREM 2.6. *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \gg \log N. \tag{2.9}$$

An alternative proof of Theorem 2.6 is due to Halász in 1981, as is the following new result.

THEOREM 2.7 (Halász 1981). *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}\mathbf{x} \gg (\log N)^{\frac{1}{2}}. \tag{2.10}$$

This is an improvement of the following result of Schmidt which itself is an improvement of Theorem 2.2.

THEOREM 2.8 (Schmidt 1977). *Suppose that $q \in \mathbb{R}$ is fixed, where $1 < q < \infty$. For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]|^q \,\mathrm{d}\mathbf{x} \gg_q (\log N)^{\frac{1}{2}q}. \tag{2.11}$$

Theorems 2.5 and 2.6 are essentially best possible, in view of the following upper bound, due to Lerch, and its analogue.

THEOREM 2.9 (Lerch 1904). *There exists a real sequence $(s_i)_{i\in\mathbb{N}}$ in $I = [0,1)$ such that for every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, we have*

$$\sup_{\substack{1\leqslant n\leqslant N\\ 0<\alpha\leqslant 1}} |D([0,\alpha),n)| \ll \log N.$$

THEOREM 2.10. *For every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \ll \log N. \tag{2.12}$$

In fact, Theorems 2.2, 2.7 and 2.8 are also sharp.

THEOREM 2.11 (Davenport 1956). *For every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \ll \log N. \tag{2.13}$$

THEOREM 2.12 (Chen 1980). *Suppose that $q \in \mathbb{R}$ is fixed, where $0 < q < \infty$. For every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]|^q \,\mathrm{d}\mathbf{x} \ll_q (\log N)^{\frac{1}{2}q}. \tag{2.14}$$

We complete this section by demonstrating the equivalence of Assertions A and B.

Suppose first of all that Assertion B holds for a function $g(N)$, and we may clearly assume that $g(N) \to \infty$ as $N \to \infty$. Let $(s_i)_{i\in\mathbb{N}}$ be a given sequence in $I = [0,1)$. We now apply Assertion B with the set

$$\mathcal{P} = \{(s_i, N^{-1}(i-1)) : i = 1,\ldots,N\}.$$

Then there exists $\mathbf{x} = (x_1,x_2) \in [0,1]^2$ such that

$$|D[\mathcal{P};B(\mathbf{x})]| > c_1 g(N),$$

where c_1 is a positive absolute constant. It then follows that

$$|D([0,\alpha),n)| > c_2 g(N),$$

where $\alpha = x_1$ and $n = -[-Nx_2]$, and where $c_2 < c_1$ is a positive constant depending at most on the function $g(N)$ but independent of N. Hence Assertion A is satisfied with $f(N) = g(N)$.

Suppose next that Assertion A holds for a function $f(N)$, and again we may clearly assume that $f(N) \to \infty$ as $N \to \infty$. Suppose that

$$\mathcal{P} = \{(y_1(i), y_2(i)) : i = 1,\ldots,N\},$$

with

$$y_2(1) \leqslant y_2(2) \leqslant \ldots \leqslant y_2(N). \tag{2.15}$$

We now consider the modified point set

$$\mathcal{P}^* = \{(y_1(i), N^{-1}(i-1)) : i = 1,\ldots,N\},$$

and write

$$M = \sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \quad \text{and} \quad M^* = \sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P}^*; B(\mathbf{x})]|. \tag{2.16}$$

Clearly, it follows immediately from (2.16) that

$$|Z[\mathcal{P}; B(1, y_2(i))] - Ny_2(i)| \leqslant M.$$

On the other hand, (2.15) can have at most $2M - 1$ consecutive equalities, and so

$$|Z[\mathcal{P}; B(1, y_2(i))] - (i-1)| \leqslant 2M.$$

It follows that

$$|(i-1) - Ny_2(i)| \leqslant 3M. \tag{2.17}$$

Observe next that the difference $|Z[\mathcal{P}; B(\mathbf{x})] - Z[\mathcal{P}^*; B(\mathbf{x})]|$ does not exceed the number of integers i for which exactly one of the two numbers $y_2(i)$ and $N^{-1}(i-1)$ is strictly less than x_2. By (2.17), there are at most $6M$ such integers i. Hence

$$|D[\mathcal{P}^*; B(\mathbf{x})]| \leqslant |Z[\mathcal{P}; B(\mathbf{x})] - Z[\mathcal{P}^*; B(\mathbf{x})]| + |D[\mathcal{P}; B(\mathbf{x})]| \leqslant 7M,$$

and so $M^* \leqslant 7M$. Clearly $M^* \gg f(N)$, so it follows immediately that Assertion B holds with $g(N) = f(N)$.

2.3. Higher Dimensions

The two formulations that we have discussed earlier can be extended to higher dimensions. Then the corresponding Assertion A in dimension k will concern the distribution of sequences $(\mathbf{s}_i)_{i\in\mathbb{N}}$ in $I^k = [0,1)^k$ with respect to aligned rectangular boxes of the type $[\mathbf{0}, \boldsymbol{\alpha}) = [0, \alpha_1) \times \ldots \times [0, \alpha_k)$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k) \in (0,1]^k$, while the corresponding Assertion B in dimension K will concern distributions $\mathcal{P}$ of points in the unit cube $[0,1]^K$ with respect to aligned rectangular boxes of the type $B(\mathbf{x}) = [0, x_1) \times \ldots \times [0, x_K)$, where $\mathbf{x} = (x_1, \ldots, x_K) \in [0,1]^K$. One can also show without too much difficulty that Assertion A in dimension k is equivalent to Assertion B in dimension $K = k+1$.

Furthermore, Theorems 2.2 and 2.8 can be extended in a straightforward way.

Theorem 2.13 (Schmidt 1977). *Suppose that $q \in \mathbb{R}$ is fixed, where $1 < q < \infty$. For every distribution $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\int_{[0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]|^q \, \mathrm{d}\mathbf{x} \gg_{K,q} (\log N)^{\frac{1}{2}(K-1)q}.$$

The case $q = 2$ is part of Roth's original work in 1954.

On the other hand, Theorems 2.11 and 2.12 can also be extended, albeit with substantial new ideas. The case $q = 2$ of the following result is due to Roth in 1980.

Theorem 2.14 (Chen 1980). *Suppose that $q \in \mathbb{R}$ is fixed, where $0 < q < \infty$. For every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\int_{[0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]|^q \, \mathrm{d}\mathbf{x} \ll_{K,q} (\log N)^{\frac{1}{2}(K-1)q}.$$

The *Great open problem* in discrepancy theory concerns the study of the function

$$\sup_{\mathbf{x}\in[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]|$$

in dimensions $K > 2$. Here the best results known so far are summarized below.

THEOREM 2.15 (Halton 1960). *For every $N \in \mathbb{N}$ satisfying $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\sup_{\mathbf{x}\in[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \ll_K (\log N)^{K-1}. \tag{2.18}$$

THEOREM 2.16 (Bilyk, Lacey and Vagharshakyan 2008). *There exists $\delta_K \in (0, \frac{1}{2})$ such that for every distribution $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\sup_{\mathbf{x}\in[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \gg_K (\log N)^{\frac{1}{2}(K-1)+\delta_K}. \tag{2.19}$$

There clearly remains a substantial gap between the lower bound (2.19) and the upper bound (2.18). There have been different conjectures concerning the correct answer to this question.

Another difficult question concerns the study of the function

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}\mathbf{x}$$

in dimensions $K > 2$. Here we have the following result.

THEOREM 2.17 (Halász 1981). *For every distribution $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}\mathbf{x} \gg_K (\log N)^{\frac{1}{2}}.$$

Here the conjecture is that the exponent $\frac{1}{2}$ should be $\frac{1}{2}(K-1)$, consistent with the conclusion of Theorem 2.13.

CHAPTER 3

Generalization of the Problem

3.1. The Work of Schmidt and Beck

Roth's reformulation of the original discrepancy problem brings geometry into consideration, and enables us to pose the following far more general question.

We initially restrict our discussion to the unit square $U = [0,1]^2$ or the torus $U = \mathbb{T}^2$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Let $\mathcal{A}$ denote an infinite collection of subsets of U, endowed with an integral geometric measure $\mathrm{d}A$, suitably normalized so that the total measure is equal to 1. Suppose that $\mathcal{P}$ is a distribution of N points in U. For every set $A \in \mathcal{A}$, let $Z[\mathcal{P}; A]$ denote the number of points of $\mathcal{P}$ that fall into A, and consider the discrepancy

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

where $\mu(A)$ denotes the usual area of A, so that $N\mu(A)$ represents the expected number of points of $\mathcal{P}$ that fall into the set A.

Example. For the classical discrepancy problem, the sets in $\mathcal{A}$ are of the form $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, where $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, and the measure is the usual Lebesgue measure $\mathrm{d}\mathbf{x}$.

We may then consider an average of the form

$$\int_A |D[\mathcal{P}; A]|^q \,\mathrm{d}A$$

for fixed $q \in \mathbb{R}$ satisfying $1 \leqslant q < \infty$. We may also consider the extreme discrepancy

$$\sup_{A \in \mathcal{A}} |D[\mathcal{P}; A]|.$$

The study of the problem in this more general setting originates from the work of Schmidt in the 1960s and 1970s. We state here a few of his remarkable results.

Theorem 3.1 (Schmidt 1969). *Let $\mathcal{A}$ denote the collection of discs of diameter not exceeding 1 in the unit torus $\mathbb{T}^2$. For every distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ and every positive real number ϵ, we have*

$$\sup_{A \in \mathcal{A}} |D[\mathcal{P}; A]| \gg_\epsilon N^{\frac{1}{4} - \epsilon}. \tag{3.1}$$

Theorem 3.2 (Schmidt 1977). *Let $\mathcal{A}$ denote the collection of rotated rectangles of diameter not exceeding 1 in the unit torus $\mathbb{T}^2$. For every distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ and every positive real number ϵ, we have*

$$\sup_{A \in \mathcal{A}} |D[\mathcal{P}; A]| \gg_\epsilon N^{\frac{1}{4} - \epsilon}. \tag{3.2}$$

THEOREM 3.3 (Schmidt 1973). *Let $\mathcal{A}$ denote the collection of convex sets of diameter not exceeding* 1 *in the unit square* $[0,1]^2$. *For every distribution* $\mathcal{P}$ *of* N *points in* $[0,1]^2$, *we have*

$$\sup_{A\in\mathcal{A}} |D[\mathcal{P};A]| \gg_\epsilon N^{\frac{1}{3}}.$$

These estimates are close to best possible. Indeed, the following results can be established using large deviation techniques in probability theory.

THEOREM 3.4 (Beck 1981). *Let $\mathcal{A}$ denote the collection of discs of diameter not exceeding* 1 *in the unit torus* $\mathbb{T}^2$ *or the collection of rotated rectangles of diameter not exceeding* 1 *in the unit torus* $\mathbb{T}^2$. *For every natural number* $N \geqslant 2$, *there exists a distribution* $\mathcal{P}$ *of* N *points in* $\mathbb{T}^2$ *such that*

$$\sup_{A\in\mathcal{A}} |D[\mathcal{P};A]| \ll N^{\frac{1}{4}}(\log N)^{\frac{1}{2}}.$$

THEOREM 3.5 (Beck 1988). *Let $\mathcal{A}$ denote the collection of convex sets of diameter not exceeding* 1 *in the unit square* $[0,1]^2$. *For every natural number* $N \geqslant 2$, *there exists a distribution* $\mathcal{P}$ *of* N *points in* $[0,1]^2$ *such that*

$$\sup_{A\in\mathcal{A}} |D[\mathcal{P};A]| \ll N^{\frac{1}{3}}(\log N)^4.$$

The following result, arguably the greatest single contribution to discrepancy theory, is due to Beck and is established via Fourier transform techniques used by Roth in the 1960s to study discrepancy problems for integer sequences.

Let B denote a compact and convex set in the unit torus $\mathbb{T}^2$. For every real number $\lambda \in [0,1]$, every rotation $\theta \in [0,2\pi]$ and every translation $\mathbf{x} \in \mathbb{T}^2$, let

$$B(\lambda,\theta,\mathbf{x}) = \{\theta(\lambda\mathbf{y}) + \mathbf{x} : \mathbf{y} \in B\} \tag{3.3}$$

denote the similar copy of B obtained from B by a contraction by factor λ about the origin, followed by an anticlockwise rotation by angle θ about the origin and then by a translation by vector $\mathbf{x}$. We denote by $\mathcal{A}(B)$ the collection of all similar copies of B obtained this way.

THEOREM 3.6 (Beck 1987). *Suppose that* $\mathcal{P}$ *is an arbitrary distribution of* N *points in the unit torus* $\mathbb{T}^2$. *Let* B *denote a compact and convex set in* $\mathbb{T}^2$ *such that* $r(B) \geqslant N^{-\frac{1}{2}}$, *where* $r(B)$ *denotes the radius of the largest inscribed ball in* B. *Then*

$$\int_{\mathbb{T}^2}\int_0^{2\pi}\int_0^1 |D[\mathcal{P};B(\lambda,\theta,\mathbf{x})]|^2 \,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\mathbf{x} \gg_B N^{\frac{1}{2}}. \tag{3.4}$$

REMARK. The estimate (3.4) gives, as an immediate consequence, the estimate

$$\sup_{A\in\mathcal{A}(B)} |D[\mathcal{P};A]| \gg_B N^{\frac{1}{4}},$$

leading to an improvement of the estimates (3.1) and (3.2). In fact, the technique of Theorem 3.4 applies in this more general setting, so that for every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that

$$\sup_{A\in\mathcal{A}(B)} |D[\mathcal{P};A]| \ll_B N^{\frac{1}{4}}(\log N)^{\frac{1}{2}}.$$

In fact, Theorem 3.6 is best possible, in view of the following result.

THEOREM 3.7 (Beck and Chen 1990). *Let B denote a compact and convex set in the unit torus $\mathbb{T}^2$. For every natural number N, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2}\int_0^{2\pi}\int_0^1 |D[\mathcal{P};B(\lambda,\theta,\mathbf{x})]|^2\,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\mathbf{x} \ll_B N^{\frac{1}{2}}. \tag{3.5}$$

3.2. The Disc Segment Problem of Roth

Next, we turn our attention to an interesting problem raised by Roth.

Let U_0 denote the closed disc of unit area in $\mathbb{R}^2$, centred at the origin. For every non-negative real number $r \in \mathbb{R}$ and every angle $\theta \in [0,2\pi]$, let $H(r,\theta)$ denote the closed half plane

$$H(r,\theta) = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}\cdot\mathbf{e}(\theta) \geqslant r\},$$

where $\mathbf{e}(\theta) = (\cos\theta, \sin\theta)$ and $\mathbf{x}\cdot\mathbf{y}$ denotes the scalar product of $\mathbf{x}$ and $\mathbf{y}$, and write

$$S(r,\theta) = H(r,\theta) \cap U_0.$$

It is not difficult to see that $S(r,\theta)$ is a disc segment in U_0, with its straight boundary a perpendicular distance r away from the origin, and where the perpendicular from the origin to this boundary is at an angle θ from the positive horizontal direction.

Suppose that $\mathcal{P}$ is a distribution of N points in U_0. For every real number $r \in [0,\pi^{-\frac{1}{2}}]$ and every angle $\theta \in [0,2\pi]$, let $Z[\mathcal{P};S(r,\theta)]$ denote the number of points of $\mathcal{P}$ that fall into $S(r,\theta)$, and consider the discrepancy

$$D[\mathcal{P};S(r,\theta)] = Z[\mathcal{P};S(r,\theta)] - N\mu(S(r,\theta)),$$

where μ is the usual area measure in $\mathbb{R}^2$. The question of Roth concerns whether there exists a positive function $g(N)$, satisfying $g(N)\to\infty$ as $N\to\infty$, such that for every distribution $\mathcal{P}$ of N points in U_0, we have

$$\sup_{\substack{0\leqslant r\leqslant \pi^{-\frac{1}{2}}\\ 0\leqslant\theta\leqslant 2\pi}} |D[\mathcal{P};S(r,\theta)]| \gg g(N).$$

Roth's question has been answered in the affirmative by Beck and improved by Alexander.

THEOREM 3.8 (Beck 1983). *For every distribution $\mathcal{P}$ of N points in U_0, we have*

$$\sup_{\substack{0\leqslant r\leqslant \pi^{-\frac{1}{2}}\\ 0\leqslant\theta\leqslant 2\pi}} |D[\mathcal{P};S(r,\theta)]| \gg N^{\frac{1}{4}}(\log N)^{-\frac{7}{2}}.$$

THEOREM 3.9 (Alexander 1990). *For every distribution $\mathcal{P}$ of N points in U_0, we have*

$$\int_0^{2\pi}\int_0^{\pi^{-\frac{1}{2}}} |D[\mathcal{P};S(r,\theta)]|^2\,\mathrm{d}r\,\mathrm{d}\theta \gg N^{\frac{1}{2}}. \tag{3.6}$$

In particular, we have

$$\sup_{\substack{0\leqslant r\leqslant \pi^{-\frac{1}{2}}\\ 0\leqslant\theta\leqslant 2\pi}} |D[\mathcal{P};S(r,\theta)]| \gg N^{\frac{1}{4}}. \tag{3.7}$$

The estimate (3.7) is sharp, in view of the following amazing result.

THEOREM 3.10 (Matoušek 1995). *For every natural number N, there exists a distribution $\mathcal{P}$ of N points in U_0 such that*

$$\sup_{\substack{0\leqslant r\leqslant \pi^{-\frac{1}{2}}\\ 0\leqslant\theta\leqslant 2\pi}} |D[\mathcal{P};S(r,\theta)]| \ll N^{\frac{1}{4}}.$$

On the other hand, we can establish the following surprising result.

THEOREM 3.11 (Beck and Chen 1993). *For every natural number $N\geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in U_0 such that*

$$\int_0^{2\pi}\int_0^{\pi^{-\frac{1}{2}}} |D[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\,\mathrm{d}\theta \ll (\log N)^2. \tag{3.8}$$

Note that any distribution $\mathcal{P}$ of N points that satisfies the inequality (3.8) must also at the same time satisfy the inequalities (3.6) and (3.7). This shows that while the discrepancy can be very large, it does not happen very often.

3.3. Convex Polygons

Let us return to the unit torus $\mathbb{T}^2$.

Suppose now that B is a closed convex polygon in $\mathbb{T}^2$. For every real number $\lambda\in[0,1]$, every rotation $\theta\in[0,2\pi]$ and every translation $\mathbf{x}\in\mathbb{T}^2$, we can consider similar copies of B given by (3.3). For every distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$, the inequality (3.4) holds. On the other hand, for every natural number N, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that the inequality (3.5) holds.

Note next that a convex polygon is the intersection of a finite number of half planes. But then the disc segment problem is really a problem about half planes. It is therefore perhaps not too surprising that we have the following analogue of Theorem 3.11.

THEOREM 3.12 (Beck and Chen 1993). *Let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every natural number $N\geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2}\int_0^{2\pi}\int_0^1 |D[\mathcal{P};B(\lambda,\theta,\mathbf{x})]|\,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\mathbf{x} \ll_B (\log N)^2.$$

The point set $\mathcal{P}$ used in the proof of Theorem 3.12 is a suitably scaled copy of a square lattice. Here one notes that the discrepancy $D[\mathcal{P};B(\lambda,\theta,\mathbf{x})]$ is large for some rotations θ and small for other rotations θ. This leads us to investigate what happens if we do not permit rotations and restrict our investigation to homothetic copies of a given closed convex polygon B.

More precisely, let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every real number $\lambda\in[0,1]$ and every translation $\mathbf{x}\in\mathbb{T}^2$, let

$$B(\lambda,\mathbf{x})=\{\lambda\mathbf{y}+\mathbf{x}:\mathbf{y}\in B\}$$

denote the homothetic copy of B obtained from B by a contraction by factor λ about the origin, followed by a translation by vector $\mathbf{x}$.

By rotating a suitably scaled copy of a square lattice by a carefully chosen angle, we can establish the following generalization of Davenport's theorem concerning aligned rectangles.

THEOREM 3.13 (Beck and Chen 1997). *Let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2}\int_0^1 |D[\mathcal{P};B(\lambda,\mathbf{x})]|^2 \,\mathrm{d}\lambda\,\mathrm{d}\mathbf{x} \ll_B \log N.$$

3.4. Higher Dimensions

The questions discussed in Section 3.1 can be generalized to higher dimensions without extra difficulties. Let B denote a compact and convex set in the unit torus $\mathbb{T}^K$. For every real number $\lambda \in [0,1]$, every orthogonal transformation $\tau \in \mathcal{T}$ and every translation $\mathbf{x} \in \mathbb{T}^K$, let

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

denote the similar copy of B obtained from B by a contraction by factor λ about the origin, followed by an orthogonal transformation τ about the origin and then by a translation by vector $\mathbf{x}$. Here $\mathcal{T}$ denotes the group of all orthogonal transformations in $\mathbb{T}^K$, with measure $\mathrm{d}\tau$ normalized so that the total measure is equal to 1.

We summarize the main results as follows.

THEOREM 3.14 (Beck 1987). *Suppose that $\mathcal{P}$ is an arbitrary distribution of N points in the unit torus $\mathbb{T}^K$. Let B denote a compact and convex set in $\mathbb{T}^K$ such that $r(B) \geqslant N^{-\frac{1}{K}}$, where $r(B)$ denotes the radius of the largest inscribed ball in B. Then*

$$\int_{\mathbb{T}^K}\int_{\mathcal{T}}\int_0^1 |D[\mathcal{P};B(\lambda,\tau,\mathbf{x})]|^2 \,\mathrm{d}\lambda\,\mathrm{d}\tau\,\mathrm{d}\mathbf{x} \gg_B N^{1-\frac{1}{K}}.$$

THEOREM 3.15 (Beck and Chen 1990). *Let B denote a compact and convex set in the unit torus $\mathbb{T}^K$. For every natural number N, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^K$ such that*

$$\int_{\mathbb{T}^K}\int_{\mathcal{T}}\int_0^1 |D[\mathcal{P};B(\lambda,\tau,\mathbf{x})]|^2 \,\mathrm{d}\lambda\,\mathrm{d}\tau\,\mathrm{d}\mathbf{x} \ll_B N^{1-\frac{1}{K}}.$$

THEOREM 3.16 (Beck 1981). *Let B denote a compact and convex set in the unit torus $\mathbb{T}^K$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^K$ such that*

$$\sup_{\substack{\lambda\in[0,1]\\ \tau\in\mathcal{T}\\ \mathbf{x}\in\mathbb{T}^K}} |D[\mathcal{P};B(\lambda,\tau,\mathbf{x})]| \ll_B N^{\frac{1}{2}-\frac{1}{2K}}(\log N)^{\frac{1}{2}}.$$

The disc segment problem discussed in Section 3.2 can be generalized to higher dimensions as follows. Let U_0 denote the closed sphere of unit volume in $\mathbb{R}^K$, centred at the origin. The radius R of U_0 satisfies the formula

$$R^K = \Gamma\left(\frac{K}{2}+1\right)\pi^{-\frac{K}{2}}.$$

For every non-negative real number $r \in \mathbb{R}$ and every unit vector $\mathbf{e} \in \mathbb{R}^n$, let $H(r,\mathbf{e})$ denote the half space in $\mathbb{R}^n$, not containing the origin and with its boundary hyperspace a perpendicular distance r from the origin in the direction $\mathbf{e}$, and let $S(r,\mathbf{e}) = H(r,\mathbf{e}) \cap U_0$. Let Σ denote the set of unit vectors in $\mathbb{R}^K$ and $\mathrm{d}\mathbf{e}$ its measure normalized so that the total measure is equal to 1.

THEOREM 3.17 (Alexander 1991). *For every distribution $\mathcal{P}$ of N points in the closed sphere U_0 in $\mathbb{R}^K$, we have*

$$\int_\Sigma \int_0^R |D[\mathcal{P}; S(r, \mathbf{e})]|^2 \,\mathrm{d}r \,\mathrm{d}\mathbf{e} \gg N^{1-\frac{1}{K}}.$$

Using a variant of the technique for the proof of Theorem 3.7, one can show that the above result is essentially best possible.

Finally, no generalization of any result in Section 3.3 to higher dimension is known. Any attempt is likely to bring the investigator towards the famous but yet unresolved Littlewood conjecture on diophantine approximation.

CHAPTER 4

Introduction to Lower Bounds

4.1. Trivial Errors

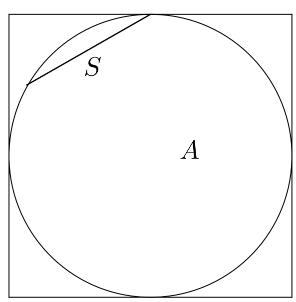
One of the key ideas in the study of discrepancy theory is the very simple notion that there is no integer between 0 and 1. Indeed, we exploit these gaps between integers, as they give us the *trivial discrepancies* or *trivial errors*. The main part of any discrepancy argument is then to find ways to *blow these up*. More precisely, we try to find ways of accumulating these trivial errors and ensure that they do not cancel among themselves.

The best illustration of the above description is an elegant yet simple proof of the following surprisingly sharp result.

THEOREM 3.3 (Schmidt 1973). *Let $\mathcal{A}$ denote the collection of convex sets of diameter not exceeding 1 in the unit square $[0,1]^2$. For every distribution $\mathcal{P}$ of N points in $[0,1]^2$, we have*

$$\sup_{A\in\mathcal{A}} |D[\mathcal{P};A]| \gg_{\epsilon} N^{\frac{1}{3}}.$$

PROOF. Let A denote the closed disc of diameter 1 and centred at the centre of the unit square $[0,1]^2$. We now consider disc segments S of area $(2N)^{-1}$ as shown in the picture below.



Any such disc segment S may contain no points of $\mathcal{P}$ or contain at least one point of $\mathcal{P}$. But then $N\mu(S) = \frac{1}{2}$, so that we expect each of these disc segments to contain precisely half a point, which is clearly impossible. Consequently, corresponding to these two cases, we have respectively the trivial discrepancies

$$D[\mathcal{P};S] = -\frac{1}{2} \quad \text{and} \quad D[\mathcal{P};S] \geqslant \frac{1}{2}.$$

Simple geometric consideration will show that there are $\gg N^{\frac{1}{3}}$ mutually disjoint disc segments of this type. Suppose that among these, $S_1,\ldots,S_k$ contain no point of $\mathcal{P}$, while $T_1,\ldots,T_m$ each contains at least one point of $\mathcal{P}$. We now consider the two convex sets

$$A \setminus (S_1 \cup \ldots \cup S_k) \quad \text{and} \quad A \setminus (T_1 \cup \ldots \cup T_m).$$

Then

$$D[\mathcal{P}; A \setminus (S_1 \cup \ldots \cup S_k)] = D[\mathcal{P}; A] - \sum_{i=1}^{k} D[\mathcal{P}; S_i]$$

and

$$D[\mathcal{P}; A \setminus (T_1 \cup \ldots \cup T_m)] = D[\mathcal{P}; A] - \sum_{j=1}^{m} D[\mathcal{P}; T_j],$$

so that

$$\begin{aligned} & D[\mathcal{P}; A \setminus (S_1 \cup \ldots \cup S_k)] - D[\mathcal{P}; A \setminus (T_1 \cup \ldots \cup T_m)] \\ & = \sum_{j=1}^{m} D[\mathcal{P}; T_j] - \sum_{i=1}^{k} D[\mathcal{P}; S_i] \geqslant \frac{m+k}{2} \gg N^{\frac{1}{3}}. \end{aligned}$$

It follows that

$$\max\{|D[\mathcal{P}; A \setminus (S_1 \cup \ldots \cup S_k)]|, |D[\mathcal{P}; A \setminus (T_1 \cup \ldots \cup T_m)]|\} \gg N^{\frac{1}{3}}.$$

The result follows. ○

Remark. This result is sometimes affectionately known as *Schmidt's chocolate theorem*. The reader may find the email below amusing:

Dear William,

Recently I came upon some old writing of yours about me and chocolate. Actually my son had found it someplace on the internet and forwarded it to me. It is the note which contains two lemmas.

Lemma 1. Wolfgang Schmidt loves chocolate.

Lemma 2. Pat Schmidt makes lovely chocolate cake.

I am very touched by your kind comments. Am I forgetful or what, but I don't remember hearing you talk about this at a conference or reading it before. My son talked about your writing to my grandson (8 years) who then wrote about it in a school project, saying he liked me because I like chocolate and I am funny. Unfortunately I now have to eat less chocolate. I had kidney stones and nutritionists (they are bad people) say I should avoid chocolate and some other food to prevent kidney stones from recurring ...

Best wishes, Wolfgang.

Proof of Schmidt's chocolate theorem. The proof of Lemma 1 is obvious. The proof of Lemma 2 is obvious to any reader who has been to the Schmidt residence in Boulder, Colorado. For others, try to get an invitation to visit the great man.

On Wolfgang's N-th birthday, Pat had made a beautiful round chocolate cake of diameter 1 and placed it on a square plate of area 1. She then decorated this with N chocolates, some of these on top of the cake and others on the plate.

When Wolfgang entered the kitchen while Pat was out, and when he saw the cake, he remembered Lemma 2. It follows from Lemma 1 that he decided to cut a small piece. By instinct, he chose to cut a small segment of area $(2N)^{-1}$, realizing that the remainder would remain convex and that he could repeat this operation $\gg N^{\frac{1}{3}}$ times without destroying the convexity of (what remained of) the cake.

Naturally, Lemma 1 dictates that those segments that Wolfgang preferred to cut each contained at least one chocolate. After a while, he realized that the remainder of the cake was rather deficient of chocolates. In any case, when Pat returned and discovered that some chocolates were missing, she decided to make another cake, rather similar to the first one. After all, this was Wolfgang's birthday. However, she did put the chocolates closer to the centre of the cake.

Later that day, when Wolfgang saw the second cake, he realized that if he chose again to cut a small segment of area $(2N)^{-1}$ and repeat this operation a reasonable number of times, these small pieces would now not contain any chocolates, with the result that (what remained of) the cake was still convex but now rather abundant of chocolates.

One way or other, the number of chocolates would differ from the expected number by $\gg N^{\frac{1}{3}}$. ○

4.2. Roth's Orthogonal Function Method

In this section, we shall study the ideas underpinning the pioneering result of Roth.

THEOREM 2.2 (Roth 1954). *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \gg \log N. \tag{4.1}$$

Recall that the rectangle $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$ for every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$.

Since the point set $\mathcal{P}$ is arbitrary, we have no precise information on these points, and so it is hard to extract discrepancy near these points. On the other hand, note that parts of $[0,1]^2$ are short of points of $\mathcal{P}$, giving rise to *trivial discrepancies*, so we try to exploit these.

Suppose that $\mathcal{P}$ has N points. If we partition the unit square $[0,1]^2$ into more than $2N$ subsets, then at least half of these subsets are devoid of points of $\mathcal{P}$. More precisely, choose n to satisfy

$$2N \leqslant 2^n < 4N, \tag{4.2}$$

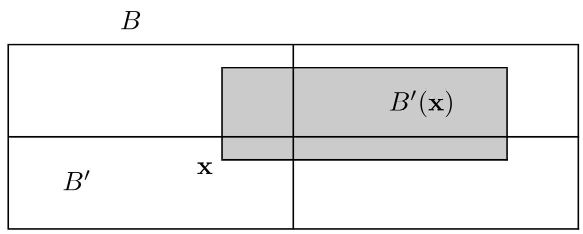
and partition $[0,1]^2$ into similar rectangles of area 2^{-n}. Then at least half of these rectangles contain no points of $\mathcal{P}$. We shall extract discrepancy from such "empty" rectangles.

A typical rectangle of area $2^{-n} = 2^{-r_1} \times 2^{-r_2}$ is of the form

$$B = \prod_{j=1}^{2} [m_j 2^{-r_j}, (m_j + 1) 2^{-r_j}), \tag{4.3}$$

where $m_1, m_2 \in \mathbb{Z}$ satisfy $0 \leqslant m_j < 2^{r_j}$ for $j = 1, 2$. For convenience, we shall call this an $\mathbf{r}$-rectangle, where $\mathbf{r} = (r_1, r_2)$.

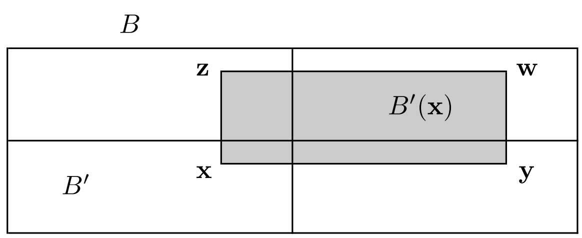
Let B' denote the bottom left quarter of the $\mathbf{r}$-rectangle B. For any $\mathbf{x} \in B'$, consider the rectangle $B'(\mathbf{x})$ of area 2^{-n-2} and with bottom left vertex $\mathbf{x}$ as shown.



Suppose that B contains no point of $\mathcal{P}$. Then $B'(\mathbf{x})$ contains no point of $\mathcal{P}$, and so has trivial discrepancy $-N2^{-n-2}$. A device to pick up this trivial discrepancy is provided by the Rademacher function $R_{r_1,r_2}(\mathbf{x})$ defined locally on B as follows.

-1	$+1$
$+1$	-1

For any $\mathbf{x} \in B'$, let $\mathbf{y}, \mathbf{z}, \mathbf{w}$ denote the other vertices of $B'(\mathbf{x})$ as shown.



Writing $D(\mathbf{x})$ as an abbreviation for $D[\mathcal{P}; B(\mathbf{x})]$, we have

$$\begin{aligned}\int_B D[\mathcal{P}; B(\mathbf{x})] R_{r_1,r_2}(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \int_B D(\mathbf{x}) R_{r_1,r_2}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ &= \int_{B'} (D(\mathbf{x}) - D(\mathbf{y}) - D(\mathbf{z}) + D(\mathbf{w}))\,\mathrm{d}\mathbf{x} \\ &= \int_{B'} D[\mathcal{P}; B'(\mathbf{x})]\,\mathrm{d}\mathbf{x} = -N2^{-2n-4}.\end{aligned}$$

To avoid those $\mathbf{r}$-rectangles B with points of $\mathcal{P}$ undoing what we have achieved, we choose to kill off the effects of those $\mathbf{r}$-rectangles. Accordingly, we define an auxiliary function on $[0,1]^2$ as follows. Let B be any $\mathbf{r}$-rectangle of area $2^{-n} = 2^{-r_1} \times 2^{-r_2}$ and of the form (4.3), where $\mathbf{r} = (r_1, r_2)$. Write

$$f_{r_1,r_2}(\mathbf{x}) = \begin{cases} -R_{r_1,r_2}(\mathbf{x}), & \text{if } B \cap \mathcal{P} = \emptyset, \\ 0, & \text{if } B \cap \mathcal{P} \neq \emptyset. \end{cases} \tag{4.4}$$

Then

$$\int_B D(\mathbf{x}) f_{r_1,r_2}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \begin{cases} N2^{-2n-4}, & \text{if } B \cap \mathcal{P} = \emptyset, \\ 0, & \text{if } B \cap \mathcal{P} \neq \emptyset. \end{cases} \tag{4.5}$$

Summing over all similar $\mathbf{r}$-rectangles, and noting that at least 2^{n-1} $\mathbf{r}$-rectangles B satisfy $B \cap \mathcal{P} = \emptyset$, we conclude that

$$\int_{[0,1]^2} D(\mathbf{x}) f_{r_1,r_2}(\mathbf{x})\,\mathrm{d}\mathbf{x} = N 2^{-2n-4} \#\{B : B \cap \mathcal{P} = \emptyset\} \gg 1. \tag{4.6}$$

There are $n+1$ choices of integers $r_1, r_2 \geqslant 0$ with $r_1 + r_2 = n$. Accordingly, we consider the auxiliary function

$$F(\mathbf{x}) = \sum_{\substack{r_1,r_2 \geqslant 0 \\ r_1+r_2=n}} f_{r_1,r_2}(\mathbf{x}). \tag{4.7}$$

Then it follows from (4.6) that

$$\int_{[0,1]^2} D(\mathbf{x}) F(\mathbf{x})\,\mathrm{d}\mathbf{x} \gg n+1. \tag{4.8}$$

The Cauchy–Schwarz inequality gives

$$\left| \int_{[0,1]^2} D(\mathbf{x}) F(\mathbf{x})\,\mathrm{d}\mathbf{x} \right| \leqslant \left(\int_{[0,1]^2} |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \right)^{\frac{1}{2}} \left(\int_{[0,1]^2} |F(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \right)^{\frac{1}{2}}. \tag{4.9}$$

We therefore need

$$\int_{[0,1]^2} |F(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \ll n+1, \tag{4.10}$$

but this follows easily from the orthogonality condition

$$\int_{[0,1]^2} f_{r_1',r_2'}(\mathbf{x}) f_{r_1'',r_2''}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0 \quad \text{if } (r_1', r_2') \neq (r_1'', r_2''), \tag{4.11}$$

easily seen by drawing a suitable picture of the two functions in the integrand. Thus

$$\int_{[0,1]^2} |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \gg \log N,$$

and this is an abbreviated form of the desired inequality (4.1).

We shall complete this section by making some remarks concerning the proof of Theorem 2.8. We shall not give all the details, since this result is superseded by Theorem 2.7.

We shall use the same auxiliary function (4.7), so that (4.8) still holds. Now, instead of the Cauchy–Schwarz inequality (4.9), we use the Hölder inequality

$$\left| \int_{[0,1]^2} D(\mathbf{x}) F(\mathbf{x})\,\mathrm{d}\mathbf{x} \right| \leqslant \left(\int_{[0,1]^2} |D(\mathbf{x})|^q\,\mathrm{d}\mathbf{x} \right)^{\frac{1}{q}} \left(\int_{[0,1]^2} |F(\mathbf{x})|^t\,\mathrm{d}\mathbf{x} \right)^{\frac{1}{t}},$$

valid for any positive $q, t \in \mathbb{R}$ satisfying $q^{-1} + t^{-1} = 1$. The inequality (4.10) now needs to be replaced by the stronger inequality

$$\int_{[0,1]^2} |F(\mathbf{x})|^{2m}\,\mathrm{d}\mathbf{x} \ll_m (n+1)^m,$$

and this can be established for every $m \in \mathbb{N}$, based on some super-orthogonality condition which represents a generalization of (4.11) to the case where the integrand is a product of $2m$ terms rather than just 2 terms.

REMARK. We comment at this point that while Theorem 2.8 extends in a natural way to higher dimensions, this is not the case for Theorem 2.7. Indeed, the higher dimensional analogue of Theorem 2.7 is one of the greatest open problems in the subject.

4.3. Halász's Variation of Roth's Method

Suppose that the integer n is defined by (4.2) as before. A shrewd observation by Halász in 1981 is that the auxiliary function

$$H(\mathbf{x}) = \prod_{\substack{r_1,r_2\geqslant 0\\ r_1+r_2=n}} (1+\alpha f_{r_1,r_2}(\mathbf{x})) - 1, \tag{4.12}$$

where α is a suitable fixed positive constant satisfying $0 < \alpha < \frac{1}{2}$, to be chosen later, can be used to establish the following result.

THEOREM 2.6 (Schmidt 1972). *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\sup_{\mathbf{x}\in[0,1]^2} |D[\mathcal{P};B(\mathbf{x})]| \gg \log N.$$

The Cauchy–Schwarz inequality (4.9) is now replaced by the simpler inequality

$$\left|\int_{[0,1]^2} D(\mathbf{x})H(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \leqslant \left(\sup_{\mathbf{x}\in[0,1]^2} |D(\mathbf{x})|\right) \int_{[0,1]^2} |H(\mathbf{x})|\,\mathrm{d}\mathbf{x}. \tag{4.13}$$

To handle the integral on the right hand side, we need a simple generalization of the orthogonality condition (4.11). The following result follows almost immediately from the definition (4.4).

LEMMA 4.1. *Suppose that for each $i = 1,\ldots,j$, the pair $\mathbf{r}_i = (r_{1i}, r_{2i})$ of non-negative integers satisfies $r_{1i} + r_{2i} = n$. Suppose further that the pairs $\mathbf{r}_1,\ldots,\mathbf{r}_j$ are distinct. Then if $\mathbf{s} = (s_1,s_2)$, where*

$$s_1 = \max_{1\leqslant i\leqslant j} r_{1i} \quad \textit{and} \quad s_2 = \max_{1\leqslant i\leqslant j} r_{2i},$$

then for any $\mathbf{s}$-rectangle B, precisely one of the following three conditions holds:

(i) $f_{\mathbf{r}_1}(\mathbf{x})\ldots f_{\mathbf{r}_j}(\mathbf{x}) = R_{\mathbf{s}}(\mathbf{x})$; *or*
(ii) $f_{\mathbf{r}_1}(\mathbf{x})\ldots f_{\mathbf{r}_j}(\mathbf{x}) = -R_{\mathbf{s}}(\mathbf{x})$; *or*
(iii) $f_{\mathbf{r}_1}(\mathbf{x})\ldots f_{\mathbf{r}_j}(\mathbf{x}) = 0$.

Furthermore, we have

$$\int_{[0,1]^2} f_{\mathbf{r}_1}(\mathbf{x})\ldots f_{\mathbf{r}_j}(\mathbf{x}) = 0. \tag{4.14}$$

It now follows that

$$\int_{[0,1]^2} |H(\mathbf{x})|\,\mathrm{d}\mathbf{x} \leqslant \int_{[0,1]^2} \prod_{\substack{r_1,r_2\geqslant 0\\ r_1+r_2=n}} (1+\alpha f_{r_1,r_2}(\mathbf{x}))\,\mathrm{d}\mathbf{x} + 1 = 2. \tag{4.15}$$

In view of (4.13), it remains to prove that

$$\int_{[0,1]^2} D(\mathbf{x})H(\mathbf{x})\,\mathrm{d}\mathbf{x} \gg n+1, \tag{4.16}$$

the analogue of (4.8).

For simplicity, we write

$$\mathcal{I} = \{\mathbf{r} = (r_1, r_2) \in \mathbb{Z}^2 : r_1, r_2 \geqslant 0 \text{ and } r_1 + r_2 = n\},$$

so that $\mathcal{I}$ has precisely $n+1$ elements.

It is easy to see that

$$H(\mathbf{x}) = \alpha F(\mathbf{x}) + \sum_{j=2}^{n+1} \alpha^j F_j(\mathbf{x}), \tag{4.17}$$

where, for every $j = 2, \ldots, n+1$,

$$F_j(\mathbf{x}) = \sum_{\substack{\mathbf{r}_1,\ldots,\mathbf{r}_j \in \mathcal{I} \\ \mathbf{r}_1,\ldots,\mathbf{r}_j \text{ distinct}}} f_{\mathbf{r}_1}(\mathbf{x}) \ldots f_{\mathbf{r}_j}(\mathbf{x}).$$

Here, the first term on the right hand side of (4.17) is the main term, and we think of the rest as remainder terms.

We shall use a crude argument to prove that the contribution of the remainder terms on the right hand side of (4.17) to the integral (4.16) is $\ll \alpha^2 n$. However, the contribution of the first term on the right hand side of (4.17) to the integral (4.16) is $\gg \alpha(n+1)$, in view of Roth's estimate (4.8). A sufficiently small fixed positive value for α will therefore give (4.16).

Suppose first of all that $j = 2, \ldots, n+1$ is fixed and the pairs $\mathbf{r}_1, \ldots, \mathbf{r}_j \in \mathcal{I}$ are distinct. Then in the notation of Lemma 4.1, for every $\mathbf{s}$ rectangle B, we have

$$\left| \int_B D(\mathbf{x}) f_{\mathbf{r}_1}(\mathbf{x}) \ldots f_{\mathbf{r}_j}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right| \leqslant N 2^{-2(s_1+s_2)-4}.$$

Note that this is a crude upper bound, since we do not care whether the $\mathbf{s}$-rectangle B contains any point of $\mathcal{P}$ or not. Summing over the $2^{s_1+s_2}$ $\mathbf{s}$-rectangles B, we deduce that

$$\left| \int_{[0,1]^2} D(\mathbf{x}) f_{\mathbf{r}_1}(\mathbf{x}) \ldots f_{\mathbf{r}_j}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right| \leqslant N 2^{-s_1-s_2-4}. \tag{4.18}$$

We may further assume without loss of generality that $r_{11} < \ldots < r_{1j}$, and write $h = r_{1j} - r_{11}$. Then (4.18) becomes

$$\left| \int_{[0,1]^2} D(\mathbf{x}) f_{\mathbf{r}_1}(\mathbf{x}) \ldots f_{\mathbf{r}_j}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right| \leqslant N 2^{-n-h-4}.$$

If we fix $r_{11} = r$ and $r_{1j} = r + h$, then there are precisely $\binom{h-1}{j-2}$ choices for $r_{12}, \ldots, r_{1(j-1)}$. It then follows that

$$\left| \int_{[0,1]^2} D(\mathbf{x}) F_j(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right| \leqslant \sum_{r=0}^{n-j+1} \sum_{h=1}^{n-r} N 2^{-n-h-4} \binom{h-1}{j-2}.$$

To complete the proof of Theorem 2.6, we now simply sum over all $j = 2, \ldots, n+1$ and obtain

$$\left|\sum_{j=2}^{n+1} \alpha^j \int_{[0,1]^2} D(\mathbf{x}) F_j(\mathbf{x}) \, \mathrm{d}\mathbf{x}\right| \leqslant \sum_{j=2}^{n+1} \sum_{r=0}^{n-j+1} \sum_{h=1}^{n-r} \alpha^j N 2^{-n-h-4} \binom{h-1}{j-2}$$

$$= \sum_{r=0}^{n-1} \sum_{h=1}^{n-r} \sum_{j=2}^{h+1} \alpha^2 N 2^{-n-h-4} \binom{h-1}{j-2} \alpha^{j-2} \leqslant N \sum_{r=0}^{n-1} \sum_{h=1}^{n-r} \alpha^2 2^{-n-h-4} (1+\alpha)^h$$

$$\leqslant N n 2^{-n-4} \alpha^2 \sum_{h=0}^{\infty} \left(\frac{1+\alpha}{2}\right)^h \leqslant N n 2^{-n-2} \alpha^2 \ll \alpha^2 n.$$

Suppose again that the integer n is defined by (4.2) as before. Halász can also show that the auxiliary function

$$K(\mathbf{x}) = \prod_{\substack{r_1, r_2 \geqslant 0 \\ r_1 + r_2 = n}} (1 + \mathrm{i} n^{-\frac{1}{2}} f_{r_1, r_2}(\mathbf{x})) - 1 \tag{4.19}$$

can be used to establish the following result.

THEOREM 2.7 (Halász 1981). *For every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \, \mathrm{d}\mathbf{x} \gg (\log N)^{\frac{1}{2}}.$$

Here the analogue of the inequalities (4.9) and (4.13) is the simple inequality

$$\left|\int_{[0,1]^2} D(\mathbf{x}) K(\mathbf{x}) \, \mathrm{d}\mathbf{x}\right| \leqslant \left(\sup_{\mathbf{x} \in [0,1]^2} |K(\mathbf{x})|\right) \int_{[0,1]^2} |D(\mathbf{x})| \, \mathrm{d}\mathbf{x}. \tag{4.20}$$

It is easy to see that

$$|K(\mathbf{x})| \leqslant \left(1 + \frac{1}{n}\right)^{\frac{1}{2}(n+1)} + 1$$

is bounded by an absolute positive constant. One now handles the left hand side of (4.20) as in the earlier part of this section, although the details are a little different.

4.4. A Haar Wavelet Approach

Let $\varphi(x)$ denote the characteristic function of the interval $[0,1)$, so that

$$\varphi(x) = \begin{cases} 1, & \text{if } 0 \leqslant x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\vartheta(x) = \varphi(2x) - \varphi(2x-1)$ for every $x \in \mathbb{R}$, so that

$$\vartheta(x) = \begin{cases} 1, & \text{if } 0 \leqslant x < \frac{1}{2}, \\ -1, & \text{if } \frac{1}{2} \leqslant x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

For every $n, k \in \mathbb{Z}$ and $x \in \mathbb{R}$, write

$$\varphi_{n,k}(x) = 2^{\frac{1}{2}n} \varphi(2^n x - k) \quad \text{and} \quad \vartheta_{n,k}(x) = 2^{\frac{1}{2}n} \vartheta(2^n x - k).$$

Note that for every $n \in \mathbb{N}_0$ and $k = 0, 1, 2, \ldots, 2^n - 1$, the function $\varphi(2^n x - k)$ denotes the characteristic function of the interval $[2^{-n}k, 2^{-n}(k+1)) \subseteq [0,1)$. It is well known that an orthonormal basis for the space $L^2([0,1])$ is given by the collection of functions

$$\vartheta_{n,k}(x), \quad n \in \mathbb{N}_0 \text{ and } k = 0,1,2,\ldots,2^n-1,$$

together with the function $\varphi(x)$. This is known as the wavelet basis for $L^2([0,1])$.

Let us now extend this to two dimensions. For every $\mathbf{n} = (n_1, n_2)$ and $\mathbf{k} = (k_1, k_2)$ in $\mathbb{Z}^2$ and every $\mathbf{x} = (x_1, x_2)$ in $\mathbb{R}^2$, write

$$\Theta_{\mathbf{n},\mathbf{k}}(\mathbf{x}) = \vartheta_{n_1,k_1}(x_1)\vartheta_{n_2,k_2}(x_2).$$

Then an orthonormal basis for $L^2([0,1]^2)$ is given by the collection of functions

$$\Theta_{\mathbf{n},\mathbf{k}}(\mathbf{x}), \quad \mathbf{n} \in \mathbb{N}_0^2,\ k_1 = 0,1,2,\ldots,2^{n_1}-1 \text{ and } k_2 = 0,1,2,\ldots,2^{n_2}-1,$$

together with the two collections of functions

$$\begin{cases} \varphi(x_1)\vartheta_{n_2,k_2}(x_2), & n_2 \in \mathbb{N}_0 \text{ and } k_2 = 0,1,2,\ldots,2^{n_2}-1, \\ \vartheta_{n_1,k_1}(x_1)\varphi(x_2), & n_1 \in \mathbb{N}_0 \text{ and } k_1 = 0,1,2,\ldots,2^{n_1}-1, \end{cases}$$

and the function $\varphi(x_1)\varphi(x_2)$. This is usually known as the rectangular wavelet basis for $L^2([0,1]^2)$.

We now give an alternative proof of Theorem 2.2 due to Pollington. First of all, note that the discrepancy function $D(\mathbf{x}) = D[\mathcal{P}; B(\mathbf{x})]$ can be written in the form $D(\mathbf{x}) = Z(\mathbf{x}) - Nx_1x_2$, where

$$Z(\mathbf{x}) = \sum_{\mathbf{p}\in\mathcal{P}} \chi_{[0,x_1)}(p_1)\chi_{[0,x_2)}(p_2) = \sum_{\mathbf{p}\in\mathcal{P}} \chi_{(p_1,1)}(x_1)\chi_{(p_2,1)}(x_2).$$

We now make use of the rectangular wavelet basis for $L^2([0,1]^2)$.

For every $\mathbf{n} = (n_1,n_2) \in \mathbb{N}_0^2$ and every $\mathbf{k} = (k_1,k_2)$, where $k_1 = 0,1,2,\ldots,2^{n_1}-1$ and $k_2 = 0,1,2,\ldots,2^{n_2}-1$, consider the wavelet coefficients

$$a_{\mathbf{n},\mathbf{k}} = \int_{[0,1]^2} Nx_1x_2\Theta_{\mathbf{n},\mathbf{k}}(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \text{and} \quad b_{\mathbf{n},\mathbf{k}} = \int_{[0,1]^2} Z(\mathbf{x})\Theta_{\mathbf{n},\mathbf{k}}(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

It is easy to see that

$$a_{\mathbf{n},\mathbf{k}} = N\left(\int_0^1 x_1\vartheta_{n_1,k_1}(x_1)\,\mathrm{d}x_1\right)\left(\int_0^1 x_2\vartheta_{n_2,k_2}(x_2)\,\mathrm{d}x_2\right).$$

Simple calculation gives

$$\int_0^1 x\vartheta_{n,k}(x)\,\mathrm{d}x = 2^{\frac{1}{2}n}\int_{2^{-n}k}^{2^{-n}(k+1)} x\vartheta(2^nx-k)\,\mathrm{d}x = \frac{1}{2^n2^{\frac{1}{2}n}(-4)}.$$

It follows that writing $|\mathbf{n}| = n_1 + n_2$, we have

$$a_{\mathbf{n},\mathbf{k}} = \frac{N}{2^{|\mathbf{n}|+4}2^{\frac{1}{2}|\mathbf{n}|}}.$$

On the other hand, we have

$$\begin{aligned} b_{\mathbf{n},\mathbf{k}} &= \sum_{\mathbf{p}\in\mathcal{P}}\left(\int_0^1 \chi_{(p_1,1)}(x_1)\vartheta_{n_1,k_1}(x_1)\,\mathrm{d}x_1\right)\left(\int_0^1 \chi_{(p_2,1)}(x_2)\vartheta_{n_2,k_2}(x_2)\,\mathrm{d}x_2\right) \\ &= \sum_{\mathbf{p}\in\mathcal{P}}\left(\int_{p_1}^1 \vartheta_{n_1,k_1}(x_1)\,\mathrm{d}x_1\right)\left(\int_{p_2}^1 \vartheta_{n_2,k_2}(x_2)\,\mathrm{d}x_2\right). \end{aligned}$$

Note that the only non-zero contributions to $b_{\mathbf{n},\mathbf{k}}$ come from those $\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}$, the support of $\Theta_{\mathbf{n},\mathbf{k}}$. If $\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}$, then $2^{-n_i}k_i \leqslant p_i < 2^{-n_i}(k_i+1)$, or $[2^{n_i}p_i] = k_i$, for both $i = 1, 2$. Simple calculation now shows that if $2^{-n}k \leqslant p < 2^{-n}(k+1)$, so that $2^n p - k = \{2^n p\}$, then

$$\int_p^1 \vartheta_{n,k}(x)\,\mathrm{d}x = 2^{\frac{1}{2}n} \int_p^{2^{-n}(k+1)} \vartheta(2^n x - k)\,\mathrm{d}x$$
$$= \frac{2^{\frac{1}{2}n}}{2^n} \int_{\{2^n p\}}^1 \vartheta(y)\,\mathrm{d}y = -\frac{2^{\frac{1}{2}n}}{2^n} \|2^n p\|,$$

where for every $\beta \in \mathbb{R}$, $\{\beta\}$ and $\|\beta\|$ denote respectively the fractional part of β and the distance of β to the nearest integer. It follows that

$$b_{\mathbf{n},\mathbf{k}} = \frac{1}{2^{\frac{1}{2}|\mathbf{n}|}} \sum_{\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}} \|2^{n_1} p_1\| \|2^{n_2} p_2\|.$$

Combining the above, we then have the wavelet coefficients

$$c_{\mathbf{n},\mathbf{k}} = \int_{[0,1]^2} D(\mathbf{x}) \Theta_{\mathbf{n},\mathbf{k}}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{1}{2^{\frac{1}{2}|\mathbf{n}|}} \left(\sum_{\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}} \|2^{n_1} p_1\| \|2^{n_2} p_2\| - \frac{N}{2^{|\mathbf{n}|+4}} \right).$$

Note in particular that the functions $\Theta_{\mathbf{n},\mathbf{k}}$ form a subcollection of the rectangular basis for $L^2([0,1]^2)$. It follows from Parseval's identity that

$$\int_{[0,1]^2} |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}$$
$$\geqslant \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \frac{1}{2^{|\mathbf{n}|}} \sum_{k_1=0}^{2^{n_1}-1} \sum_{k_2=0}^{2^{n_2}-1} \left(\sum_{\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}} \|2^{n_1} p_1\| \|2^{n_2} p_2\| - \frac{N}{2^{|\mathbf{n}|+4}} \right)^2.$$

To complete the proof, we now choose n so that $2N \leqslant 2^n < 4N$. Then for every fixed $\mathbf{n}$ satisfying $|\mathbf{n}| = n$, at least 2^{n-1} of the rectangles $B_{\mathbf{n},\mathbf{k}}$ do not contain any point of $\mathcal{P}$, so that

$$\sum_{\mathbf{p} \in B_{\mathbf{n},\mathbf{k}}} \|2^{n_1} p_1\| \|2^{n_2} p_2\| = 0.$$

It follows that

$$\int_{[0,1]^2} |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} \geqslant \sum_{n_1=0}^{\infty} \sum_{\substack{n_2=0 \\ |\mathbf{n}|=n}}^{\infty} \frac{1}{2^n} 2^{n-1} \left(\frac{N}{2^{n+4}} \right)^2 \gg n + 1 \gg \log N.$$

CHAPTER 5

Introduction to Upper Bounds

5.1. A Seemingly Trivial Argument

Let B denote a compact and convex set in the unit torus $\mathbb{T}^2$. For every real number $\lambda \in [0,1]$, every rotation $\theta \in [0,2\pi]$ and every translation $\mathbf{x} \in \mathbb{T}^2$, let

$$B(\lambda,\theta,\mathbf{x}) = \{\theta(\lambda\mathbf{y}) + \mathbf{x} : \mathbf{y} \in B\}$$

denote the similar copy of B obtained from B by a contraction by factor λ about the origin, followed by an anticlockwise rotation by angle θ about the origin and then by a translation by vector $\mathbf{x}$. We denote by $\mathcal{A}(B)$ the collection of all similar copies of B obtained this way.

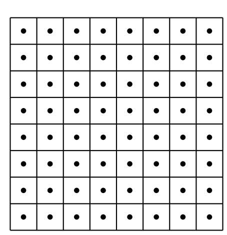
We begin our discussion here by making an inadequate attempt to establish the following variant of Theorem 3.4.

THEOREM 5.1. *Let B denote a compact and convex set in $\mathbb{T}^2$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\sup_{A\in\mathcal{A}(B)} |D[\mathcal{P};A]| \ll_B N^{\frac{1}{4}}(\log N)^{\frac{1}{2}}.$$

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 torus $\mathbb{T}^2$ 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.



Suppose that $A \in \mathcal{A}(B)$ is a similar copy of a given fixed compact and convex set B. We now attempt to estimate the discrepancy $D[\mathcal{P};A]$. Let $\mathcal{S}$ 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

$$D[\mathcal{P};A] = \sum_{S\in\mathcal{S}} D[\mathcal{P};S\cap A]. \tag{5.1}$$

Next, we make the simple observation that

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

The identity (5.1) then becomes

$$D[\mathcal{P};A] = \sum_{\substack{S\in\mathcal{S}\\ S\cap\partial A\neq\emptyset}} D[\mathcal{P};S\cap A], \tag{5.2}$$

where ∂A denotes the boundary of A. Finally, observe that both $0\leqslant Z[\mathcal{P};S\cap A]\leqslant 1$ and $0\leqslant N\mu(S\cap A)\leqslant 1$, so that $|D[\mathcal{P};S\cap A]|\leqslant 1$, and it follows from (5.2) and the triangle inequality that

$$|D[\mathcal{P};A]|\leqslant \#\{S\in\mathcal{S} : S\cap\partial A\neq\emptyset\} \ll M = N^{\frac{1}{2}}. \tag{5.3}$$

This estimate is almost trivial, but very far from the upper bound $N^{\frac{1}{4}}(\log N)^{\frac{1}{2}}$ alluded to in Theorem 5.1.

We make an important observation here that the term $\#\{S\in\mathcal{S} : S\cap\partial A\neq\emptyset\}$ in (5.3) 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.

5.2. A Large Deviation Technique

Let us continue our study of Theorem 5.1.

For every little square $S\in\mathcal{S}$, 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

$$\widetilde{\mathcal{P}} = \{\widetilde{\mathbf{p}_S} : S\in\mathcal{S}\}. \tag{5.4}$$

Suppose that a fixed compact and convex set $B\subseteq\mathbb{T}^2$ is given. Let

$$\mathcal{G}(B) = \{B(\lambda,\theta,\mathbf{x}) : \lambda\in[0,\frac{11}{10}],\ \theta\in[0,2\pi],\ \mathbf{x}\in\mathbb{T}^2\}.$$

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

$$\#\mathcal{H}(B)\leqslant N^{C_1},$$

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

$$A^-\subseteq A\subseteq A^+ \quad\text{and}\quad \mu(A^+\setminus A^-)\leqslant N^{-1}. \tag{5.5}$$

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

Suppose that $A \in \mathcal{H}(B)$ is fixed. Then, analogous to the discrepancy function (5.1), we now consider the discrepancy function

$$D[\widetilde{\mathcal{P}};A] = \sum_{S\in\mathcal{S}} D[\widetilde{\mathcal{P}};S\cap A] = \sum_{\substack{S\in\mathcal{S}\\ S\cap\partial A\neq\emptyset}} D[\widetilde{\mathcal{P}};S\cap A], \tag{5.6}$$

and note as before that

$$\#\{S\in\mathcal{S} : S\cap\partial A\neq\emptyset\} \ll M = N^{\frac{1}{2}}. \tag{5.7}$$

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

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

The observation

$$D[\widetilde{\mathcal{P}};A] = \sum_{S\in\mathcal{S}} (\phi_S - \mathbb{E}\phi_S) = \sum_{\substack{S\in\mathcal{S}\\ S\cap\partial A\neq\emptyset}} (\phi_S - \mathbb{E}\phi_S) \tag{5.8}$$

sets us up to appeal to large deviation type inequalities in probability theory. For instance, we can use the following result attributed to Hoeffding.

LEMMA 5.2. *Suppose that $\phi_1,\ldots,\phi_m$ are independent random variables such that $0\leqslant\phi_i\leqslant 1$ for every $i=1,\ldots,m$. Then for every real number $\gamma>0$, we have*

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

In view of (5.8), we now apply Lemma 5.2 with

$$m = \#\{S\in\mathcal{S} : S\cap\partial A\neq\emptyset\} \leqslant C_2 N^{\frac{1}{2}},$$

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

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

where C_3 is a sufficiently large positive constant. Indeed,

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

and it follows therefore that

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

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

$$\text{Prob}\left(|D[\widetilde{\mathcal{P}};A]| \geqslant C_3 N^{\frac{1}{4}}(\log N)^{\frac{1}{2}}\right) \leqslant \frac{1}{2}(\#\mathcal{H}(B))^{-1},$$

and so

$$\text{Prob}\left(|D[\widetilde{\mathcal{P}};A]| \geqslant C_3 N^{\frac{1}{4}}(\log N)^{\frac{1}{2}} \text{ for some } A\in\mathcal{H}(B)\right) \leqslant \frac{1}{2},$$

whence

$$\text{Prob}\left(|D[\widetilde{\mathcal{P}};A]| \leqslant C_3 N^{\frac{1}{4}}(\log N)^{\frac{1}{2}} \text{ for all } A\in\mathcal{H}(B)\right) \geqslant \frac{1}{2}.$$

In other words, there exists a set $\mathcal{P}^*$ of $N=M^2$ points in $\mathbb{T}^2$ such that

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

Suppose now that $A \in \mathcal{A}(B)$ is given. Then there exist $A^-, A^+ \in \mathcal{H}(B)$ such that (5.5) is satisfied. It is not difficult to show that

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

Theorem 5.1 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 of the squares of four non-negative integers. We can therefore superimpose up to four point distributions in $\mathbb{T}^2$ where the number of points in each is a perfect square. This completes the proof of Theorem 5.1.

5.3. An Averaging Argument

We now indicate how the argument in the previous section can be adapted to establish Theorem 3.7.

We construct the random point set $\mathcal{P}$, given by (5.4), as before. Suppose that a fixed compact and convex set $B \subseteq \mathbb{T}^2$ is given. Let $A \in \mathcal{A}(B)$ be fixed. Then (5.6)–(5.8) are valid. If we write $\eta_S = \phi_S - \mathbb{E}\phi_S$, then

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

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

$$\mathbb{E}\left(|D[\widetilde{\mathcal{P}}; A]|^2\right) = \sum_{\substack{S_1, S_2 \in \mathcal{S}\\ S_1 \cap \partial A \neq \emptyset\\ S_2 \cap \partial A \neq \emptyset}} \mathbb{E}\left(\eta_{S_1}\eta_{S_2}\right). \tag{5.9}$$

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

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

It follows that the only non-zero contributions to the sum (5.9) come from those terms where $S_1 = S_2$, so that

$$\mathbb{E}\left(|D[\widetilde{\mathcal{P}}; A]|^2\right) \leqslant \#\{S \in \mathcal{S} : S \cap \partial A \neq \emptyset\} \ll_B N^{\frac{1}{2}}.$$

Integrating over all $A \in \mathcal{A}(B)$ and changing the order of integration, we obtain

$$\mathbb{E}\left(\int_{\mathcal{A}(B)} |D[\widetilde{\mathcal{P}}; A]|^2 \,\mathrm{d}A\right) \ll_B N^{\frac{1}{2}}.$$

It follows that there exists a set $\mathcal{P}^*$ of $N = M^2$ points in $\mathbb{T}^2$ such that

$$\int_{\mathcal{A}(B)} |D[\mathcal{P}^*; A]|^2 \,\mathrm{d}A \ll_B N^{\frac{1}{2}},$$

establishing Theorem 3.7 in the special case when $N = M^2$.

We remark that the argument in this chapter can be extended in a reasonably straightforward manner to arbitrary dimensions $k \geqslant 2$. Also the argument in this section on Theorem 3.7 can be extended to L^q-norms for all even positive integers q, and hence all positive real numbers q, without too many complications.

CHAPTER 6

Fourier Transform Techniques

6.1. Introduction

Suppose that $\mathcal{P}$ is a distribution of N points in the unit torus $\mathbb{T}^2$. Let Z denote the counting measure in $\mathbb{T}^2$ of $\mathcal{P}$, and let μ denote the usual Lebesgue area mesaure in $\mathbb{T}^2$.

For any set $A \subseteq \mathbb{T}^2$, let χ_A denote the characteristic function of A, and consider the convolution

$$F = \chi_A * (\mathrm{d}Z - N\mathrm{d}\mu).$$

More precisely, for every $\mathbf{x} \in \mathbb{T}^2$, we have

$$F(\mathbf{x}) = \int_{\mathbb{T}^2} \chi_A(\mathbf{x}-\mathbf{y})(\mathrm{d}Z - N\mathrm{d}\mu)(\mathbf{y}). \tag{6.1}$$

For simplicity, let us assume that A is symmetric across the origin. Then

$$\mathbf{x}-\mathbf{y} \in A \quad \text{if and only if} \quad \mathbf{y}-\mathbf{x} \in A \quad \text{if and only if} \quad \mathbf{y} \in A+\mathbf{x},$$

where $A+\mathbf{x}$ denotes the translation of A by $\mathbf{x}$. Then

$$F(\mathbf{x}) = \int_{\mathbb{T}^2} \chi_{A+\mathbf{x}}(\mathbf{y})(\mathrm{d}Z - N\mathrm{d}\mu)(\mathbf{y}) = D[\mathcal{P}; A+\mathbf{x}], \tag{6.2}$$

the discrepancy of $\mathcal{P}$ in the translated copy $A+\mathbf{x}$ of A.

Note that the characteristic function χ_A gives the geometric input, and does not depend on the point set $\mathcal{P}$. On the other hand, the function $Z - N\mu$ gives the measure-theoretic input, and does not depend on the set A. Thus we can say that discrepancy under translation is a *convolution of geometry and measure*.

For lower bound problems, we do not have information on the point set $\mathcal{P}$ apart from the number of points that it contains. To make any progress, we need to concentrate on the geometric part, and therefore need to separate this from the measure part.

Let $L^1(\mathbb{T}^2)$ denote the set of all measurable complex valued functions that are absolutely integrable over $\mathbb{T}^2$, with Fourier transform $\widehat{f}$ defined for every $\mathbf{t} \in \mathbb{Z}^2$ by

$$\widehat{f}(\mathbf{t}) = \int_{\mathbb{T}^2} f(\mathbf{x}) e(-\mathbf{t}\cdot\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

It is well known that for any two functions $f, g \in L^1(\mathbb{T}^2)$, we have $f * g \in L^1(\mathbb{T}^2)$ and the Fourier transforms $\widehat{f}$ and $\widehat{g}$ satisfy

$$\widehat{f*g} = \widehat{f}\,\widehat{g}.$$

Let $L^2(\mathbb{T}^2)$ denote the set of all measurable complex valued functions f that are square integrable over $\mathbb{T}^2$. Then the Parseval theorem states that for every function

$f \in L^1(\mathbb{T}^2) \cap L^2(\mathbb{T}^2)$, the Fourier transform $\widehat{f} \in \ell^2(\mathbb{Z}^2)$ and satisfies

$$\int_{\mathbb{T}^2} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{t} \in \mathbb{Z}^2} |\widehat{f}(\mathbf{t})|^2.$$

Write $\phi = \widehat{Z - N\mu}$, so that for every $\mathbf{t} \in \mathbb{Z}^2$, we have

$$\phi(\mathbf{t}) = \int_{\mathbb{T}^2} e(-\mathbf{t} \cdot \mathbf{x})(\mathrm{d}Z - N\mathrm{d}\mu)(\mathbf{x}).$$

Then

$$\phi(\mathbf{0}) = \int_{\mathbb{T}^2} (\mathrm{d}Z - N\mathrm{d}\mu)(\mathbf{x}) = 0,$$

and for every non-zero $\mathbf{t} \in \mathbb{Z}^2$, it is easy to see that

$$\phi(\mathbf{t}) = \int_{\mathbb{T}^2} e(-\mathbf{t} \cdot \mathbf{x}) \,\mathrm{d}Z(\mathbf{x}), \tag{6.3}$$

and this is the Fourier transform of the point count measure Z. Furthermore, we have

$$\int_{\mathbb{T}^2} |D[\mathcal{P}; A + \mathbf{x}]|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} |\widehat{\chi}_A(\mathbf{t})|^2 |\phi(\mathbf{t})|^2. \tag{6.4}$$

6.2. A Lower Bound Argument

We shall illustrate our Fourier transform technique by establishing Theorem 3.6 in the special case when B is a square in $\mathbb{T}^2$ of side length $\frac{1}{2}$.

For every $r \in [0, \frac{1}{4}]$ and $\theta \in [0, 2\pi]$, let $A(r, \theta)$ denote the square $[-r, r]^2$ rotated anticlockwise about the origin by an angle θ. For every $\mathbf{x} \in \mathbb{T}^2$, let

$$A(r, \theta, \mathbf{x}) = A(r, \theta) + \mathbf{x}$$

denote the translation of $A(r, \theta)$ by $\mathbf{x}$.

Theorem 6.1. *Suppose that $\mathcal{P}$ is an arbitrary distribution of N points in the unit torus $\mathbb{T}^2$. Then*

$$\int_{\mathbb{T}^2} \int_0^{2\pi} \int_0^{\frac{1}{4}} |D[\mathcal{P}; A(r, \theta, \mathbf{x})]|^2 \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\mathbf{x} \gg N^{\frac{1}{2}}. \tag{6.5}$$

For every $r \in [0, \frac{1}{4}]$ and $\theta \in [0, 2\pi]$, let $\chi_{r,\theta}$ denote the characteristic function of $A(r, \theta)$. Then corresponding to (6.1) and (6.2), we have

$$D[\mathcal{P}; A(r, \theta, \mathbf{x})] = \int_{\mathbb{T}^2} \chi_{r,\theta}(\mathbf{x} - \mathbf{y})(\mathrm{d}Z - N\mathrm{d}\mu)(\mathbf{y}).$$

Corresponding to (6.4), we therefore have

$$\int_{\mathbb{T}^2} |D[\mathcal{P}; A(r, \theta, \mathbf{x}]|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} |\widehat{\chi}_{r,\theta}(\mathbf{t})|^2 |\phi(\mathbf{t})|^2.$$

We shall first of all establish some trivial discrepancies by exploiting the gaps between integers.

Lemma 6.2. *Suppose that a set $S \subseteq \mathbb{T}^2$ satisfies $0 < \delta \leqslant N\mu(S) \leqslant 1 - \delta$ for some $\delta > 0$. Then*

$$\int_{\mathbb{T}^2} |D[\mathcal{P}; S + \mathbf{x}]|^2 \,\mathrm{d}\mathbf{x} \geqslant \delta^3.$$

Note that $N\mu(S)$ is the expected number of points of $\mathcal{P}$ in S. The hypothesis of Lemma 6.2 thus says that this expectation is bounded away from an integer.

PROOF OF LEMMA 6.2. Either $Z[\mathcal{P}; S+\mathbf{x}] = 0$ or $Z[\mathcal{P}; S+\mathbf{x}] \geqslant 1$. Suppose first of all that $Z[\mathcal{P}; S+\mathbf{x}] \geqslant 1$. Then

$$D[\mathcal{P}; S+\mathbf{x}] \geqslant Z[\mathcal{P}; S+\mathbf{x}] + \delta - 1 \geqslant \delta Z[\mathcal{P}; S+\mathbf{x}],$$

and so

$$|D[\mathcal{P}; S+\mathbf{x}]| \geqslant \delta Z[\mathcal{P}; S+\mathbf{x}].$$

Note that this last inequality remains valid if $Z[\mathcal{P}; S+\mathbf{x}] = 0$, and therefore holds always. It then follows that

$$\begin{aligned}\int_{\mathbb{T}^2} |D[\mathcal{P}; S+\mathbf{x}]|^2 \,\mathrm{d}\mathbf{x} &\geqslant \delta^2 \int_{\mathbb{T}^2} Z^2[\mathcal{P}; S+\mathbf{x}] \,\mathrm{d}\mathbf{x} \geqslant \delta^2 \int_{\mathbb{T}^2} Z[\mathcal{P}; S+\mathbf{x}] \,\mathrm{d}\mathbf{x} \\ &= \delta^2 \int_{\mathbb{T}^2} \sum_{\mathbf{p}\in\mathcal{P}} \chi_{S+\mathbf{x}}(\mathbf{p}) \,\mathrm{d}\mathbf{x} = \delta^2 \sum_{\mathbf{p}\in\mathcal{P}} \int_{\mathbb{T}^2} \chi_{\mathbf{p}-S}(\mathbf{x}) \,\mathrm{d}\mathbf{x} \\ &= \delta^2 \sum_{\mathbf{p}\in\mathcal{P}} \mu(\mathbf{p}-S) = \delta^2 N\mu(S) \geqslant \delta^3\end{aligned}$$

as required. ○

We next study the function $\widehat{\chi}_{r,\theta}$. Ideally, one would like an inequality of the type

$$\frac{|\widehat{\chi}_{r,\theta}(\mathbf{t})|^2}{|\widehat{\chi}_{s,\theta}(\mathbf{t})|^2} \gg \frac{r}{s},$$

but this makes use only of one rotated square with no extra rotation or contraction, and is therefore too good to be true. Instead, we shall consider averages of the form

$$\omega_q(\mathbf{t}) = \frac{1}{q} \int_{\frac{1}{2}q}^{q} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} |\widehat{\chi}_{r,\theta}(\mathbf{t})|^2 \,\mathrm{d}\theta \,\mathrm{d}r.$$

We have the following amplification result.

LEMMA 6.3. *Suppose that $0 < p < q$. Uniformly for all $\mathbf{t} \in \mathbb{Z}^2$, we have*

$$\frac{\omega_q(\mathbf{t})}{\omega_p(\mathbf{t})} \gg \frac{q}{p}.$$

For every $r \in [0, \frac{1}{4}]$, let χ_r denote the characteristic function of $[-r, r]^2 = A(r, 0)$. Then it is not difficult to check that for every $\mathbf{t} = (t_1, t_2) \in \mathbb{Z}^2$, we have

$$\widehat{\chi}_{r,\theta}(\mathbf{t}) = \widehat{\chi}_r(t_1\cos\theta + t_2\sin\theta, -t_1\sin\theta + t_2\cos\theta). \tag{6.6}$$

Furthermore, for every $\mathbf{u} = (u_1, u_2) \in \mathbb{Z}^2$, we have

$$\widehat{\chi}_r(\mathbf{u}) = \int_{\mathbb{T}^2} \chi_r(\mathbf{x}) e(-\mathbf{u}\cdot\mathbf{x}) \,\mathrm{d}\mathbf{x} = \frac{\sin(2\pi r u_1)\sin(2\pi r u_2)}{\pi^2 u_1 u_2}. \tag{6.7}$$

Lemma 6.3 follows easily from the following result.

LEMMA 6.4. *Uniformly for all non-zero $\mathbf{t} \in \mathbb{Z}^2$, we have*

$$\omega_q(\mathbf{t}) \asymp \min\left\{q^4, \frac{q}{|\mathbf{t}|^3}\right\}.$$

PROOF. We shall show that (6.6) holds uniformly for all non-zero $\mathbf{t} \in \mathbb{R}^2$. In view of the integration over θ in the definition of $\omega_q(\mathbf{t})$, it suffices to show that uniformly for all $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$ satisfying $t_1 > 0$ and $t_2 = 0$, we have

$$\omega_q(t_1, 0) \asymp \min \left\{ q^4, \frac{q}{t_1^3} \right\}.$$

Using (6.6) and (6.7), we have

$$\omega_q(t_1, 0) \asymp \frac{1}{q} \int_{\frac{1}{2}q}^{q} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \frac{\sin^2(2\pi r t_1 \cos\theta) \sin^2(2\pi r t_1 \sin\theta)}{t_1^4 \cos^2\theta \sin^2\theta} \, \mathrm{d}\theta \, \mathrm{d}r.$$

Since $-\frac{1}{4}\pi \leqslant \theta \leqslant \frac{1}{4}\pi$, we have $\sin\theta \asymp \theta$ and $\cos\theta \asymp 1$, and so

$$\omega_q(t_1, 0) \asymp \frac{1}{q} \int_{\frac{1}{2}q}^{q} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \frac{\sin^2(2\pi r t_1 \cos\theta) \sin^2(2\pi r t_1 \sin\theta)}{t_1^4 \theta^2} \, \mathrm{d}\theta \, \mathrm{d}r. \tag{6.8}$$

We consider two cases. Let C be a fixed constant satisfying $2/\pi^2 < C < \frac{1}{4}$.

Case 1. Suppose that $t_1 \leqslant C/q$. Then for all r and θ satisfying $\frac{1}{2}q \leqslant r \leqslant q$ and $-\frac{1}{4}\pi \leqslant \theta \leqslant \frac{1}{4}\pi$, we have $2\pi r t_1 \leqslant 2\pi C < \frac{1}{2}\pi$, so $\sin(2\pi r t_1 \cos\theta) \asymp q t_1$ and $\sin(2\pi r t_1 \sin\theta) \asymp q t_1 \theta$. Hence

$$\omega_q(t_1, 0) \asymp \frac{1}{q} \int_{\frac{1}{2}q}^{q} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \frac{(q t_1)^2 (q t_1 \theta)^2}{t_1^4 \theta^2} \, \mathrm{d}\theta \, \mathrm{d}r \asymp q^4 \asymp \min \left\{ q^4, \frac{q}{t_1^3} \right\}.$$

Case 2. Suppose that $t_1 > C/q$. Then we write

$$\left[-\frac{1}{4}\pi, \frac{1}{4}\pi \right] = \left[-\frac{1}{4}\pi, -\frac{1}{2\pi q t_1} \right] \cup \left[-\frac{1}{2\pi q t_1}, \frac{1}{2\pi q t_1} \right] \cup \left[\frac{1}{2\pi q t_1}, \frac{1}{4}\pi \right]$$

as an essentially disjoint union. For the two intervals on the side, we have

$$\begin{aligned} &\int_{1/2\pi q t_1 \leqslant |\theta| \leqslant \frac{1}{4}\pi} \frac{\sin^2(2\pi r t_1 \cos\theta) \sin^2(2\pi r t_1 \sin\theta)}{t_1^4 \theta^2} \, \mathrm{d}\theta \, \mathrm{d}r \\ &\qquad \leqslant \int_{1/2\pi q t_1 \leqslant |\theta| \leqslant \frac{1}{4}\pi} \frac{\mathrm{d}\theta}{t_1^4 \theta^2} = \frac{2}{t_1^4} \left(2\pi q t_1 - \frac{4}{\pi} \right) \asymp \frac{q}{t_1^3}. \end{aligned} \tag{6.9}$$

For the last inequalities in (6.9), the upper bound is obvious. For the lower bound, note that

$$2\pi q t_1 - \frac{4}{\pi} = \left(2\pi - \frac{4}{C\pi} \right) q t_1 + \frac{4}{C\pi} q t_1 - \frac{4}{\pi} > \left(2\pi - \frac{4}{C\pi} \right) q t_1.$$

Suppose next that $-1/2\pi q t_1 \leqslant \theta \leqslant 1/2\pi q t_1$. Then $\sin(2\pi r t_1 \sin\theta) \asymp q t_1 \theta$ and

$$\frac{1}{q} \int_{\frac{1}{2}q}^{q} \sin^2(2\pi r t_1 \cos\theta) \, \mathrm{d}r \asymp 1. \tag{6.10}$$

For the inequalities (6.10), the upper bound is obvious. For the lower bound, note that as r runs through the interval $[\frac{1}{2}q, q]$, the quantity $2\pi r t_1 \cos\theta$ runs through an interval of length $\pi q t_1 \cos\theta > 2\pi^{-1} \cos \frac{1}{4}\pi$. Combining (6.8)–(6.10), we conclude that

$$\omega_q(t_1, 0) \asymp \frac{q}{t_1^3} + \int_{-1/2\pi q t_1}^{1/2\pi q t_1} \frac{q^2}{t_1^2} \, \mathrm{d}\theta \asymp \frac{q}{t_1^3} \asymp \min \left\{ q^4, \frac{q}{t_1^3} \right\}.$$

This completes the proof. ○

Proof of Theorem 6.1. We now choose $p = \frac{1}{3}N^{-\frac{1}{2}}$ and $q = \frac{1}{4}$. Then for all r and θ satisfying $\frac{1}{2}p \leqslant r \leqslant p$ and $-\frac{1}{4}\pi \leqslant \theta \leqslant \frac{1}{4}\pi$, we have $\mu(A(r,\theta)) = 4r^2$, and so $\frac{1}{9} \leqslant N\mu(A(r,\theta)) \leqslant \frac{4}{9}$. It then follows from Lemma 6.2 with $S = A(r,\theta)$ and $\delta = \frac{1}{9}$ that

$$\int_{\mathbb{T}^2} |D[\mathcal{P}; A(r,\theta,\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \gg 1,$$

and so

$$\sum_{\mathbf{0}\neq\mathbf{t}\in\mathbb{Z}^2} \omega_p(\mathbf{t})|\phi(\mathbf{t})|^2 \gg 1.$$

We now amplify this discrepancy by using Lemma 6.3 and conclude that

$$\sum_{\mathbf{0}\neq\mathbf{t}\in\mathbb{Z}^2} \omega_q(\mathbf{t})|\phi(\mathbf{t})|^2 \gg N^{\frac{1}{2}}.$$

This in turn implies that

$$\int_{\mathbb{T}^2} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \int_{\frac{1}{8}}^{\frac{1}{4}} |D[\mathcal{P}; A(r,\theta,\mathbf{x})]|^2 \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\mathbf{x} \gg N^{\frac{1}{2}},$$

from which the desired inequality (6.5) follows easily. ○

6.3. An Upper Bound Argument

We next revisit our Fourier transform technique by studying Theorem 3.7 in the special case when B is a square in $\mathbb{T}^2$ of side length $\frac{1}{2}$. More precisely, we shall establish a special case of the following result. Throughout, we retain our notation in the last section.

Theorem 6.5. *For every natural number N, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2} \int_{0}^{2\pi} \int_{0}^{\frac{1}{4}} |D[\mathcal{P}; A(r,\theta,\mathbf{x})]|^2 \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\mathbf{x} \ll N^{\frac{1}{2}}. \tag{6.11}$$

In the special case when $N = (2M+1)^2$ for some integer M, it is very tempting to see whether the set

$$\mathcal{P} = \left\{ \left(\frac{m_1}{2M+1}, \frac{m_2}{2M+1} \right) : m_1, m_2 \in \{-M, \ldots, 0, \ldots, M\} \right\}$$

of N points satisfies the inequality (6.11). Since we have precise information of the set $\mathcal{P}$, this allows us to calculate $\phi(\mathbf{t})$.

We know that $\phi(\mathbf{0}) = 0$. For any non-zero $\mathbf{t} \in \mathbb{Z}^2$, it follows from (6.3) that

$$\phi(\mathbf{t}) = \sum_{\mathbf{p}\in\mathcal{P}} e(-\mathbf{t}\cdot\mathbf{p}) = \sum_{m_1=-M}^{M} e\left(-\frac{t_1 m_1}{2M+1}\right) \sum_{m_2=-M}^{M} e\left(-\frac{t_2 m_2}{2M+1}\right).$$

It is easy to check that for every $t \in \mathbb{Z}$, we have

$$\sum_{m=-M}^{M} e\left(-\frac{tm}{2M+1}\right) = \begin{cases} 2M+1, & \text{if } t \in (2M+1)\mathbb{Z}, \\ 0, & \text{if } t \notin (2M+1)\mathbb{Z}. \end{cases}$$

Hence for every non-zero $\mathbf{t} \in \mathbb{Z}^2$, we have

$$\phi(\mathbf{t}) = \begin{cases} (2M+1)^2, & \text{if } t_1, t_2 \in (2M+1)\mathbb{Z}, \\ 0, & \text{otherwise.} \end{cases} \tag{6.12}$$

On the other hand, we know that for every q satisfying $0 < q \leqslant \frac{1}{4}$, we have

$$\int_{\mathbb{T}^2} \int_{-\frac{1}{4}\pi}^{\frac{1}{4}\pi} \int_q^{\frac{1}{2}q} |D[\mathcal{P}; A(r,\theta,\mathbf{x})]|^2 \,\mathrm{d}r \,\mathrm{d}\theta \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2.$$

In view of symmetry, we see that to establish the inequality (6.11), it suffices to show that

$$\sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2 \ll q N^{\frac{1}{2}}.$$

Recall Lemma 6.4 and (6.12). We have

$$\begin{aligned} \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2 &\ll q \sum_{\mathbf{0} \neq \mathbf{t} \in \mathbb{Z}^2} \frac{|\phi(\mathbf{t})|^2}{|\mathbf{t}|^3} = q \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} \frac{(2M+1)^4}{(2M+1)^3 |\mathbf{k}|^3} \\ &= q(2M+1) \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} \frac{1}{|\mathbf{k}|^3} \ll q N^{\frac{1}{2}} \end{aligned}$$

as required.

CHAPTER 7

Upper Bounds in the Classical Problem

7.1. Diophantine Approximation and Davenport Reflection

We begin by making a fatally flawed attempt to establish[1] Theorem 2.10.

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 $\mathcal{P}$ be the collection of these $N = M^2$ points.

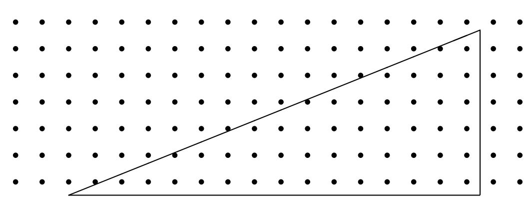
Let ξ be the second coordinate of one of the points of $\mathcal{P}$. Clearly, there are precisely M points in $\mathcal{P}$ sharing this second coordinate. Consider the discrepancy

$$D[\mathcal{P}; B(1, x_2)] \tag{7.1}$$

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 (7.1) increases by M. It follows immediately that

$$\|D[\mathcal{P}]\|_\infty \geqslant \frac{1}{2}M \gg N^{\frac{1}{2}}.$$

Let us make a digression to the work of Hardy and Littlewood 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 $\mathbb{Z}^2$ that lie in T. For simplicity, the triangle T is placed so that the horizontal side is precisely halfway between two neighbouring rows of $\mathbb{Z}^2$ and the vertical side is precisely halfway between two neighbouring columns of $\mathbb{Z}^2$.



Note that the lattice $\mathbb{Z}^2$ 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 T, and this is the discrepancy of $\mathbb{Z}^2$ 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

[1] It was put to the 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.

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 Theorem 2.10, 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 $N \geqslant 2$ is given. The lattice

$$(N^{-\frac{1}{2}}\mathbb{Z})^2 \tag{7.2}$$

contains precisely N points per unit area. Inspired by Hardy and Littlewood, we now rotate (7.2) 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 $\mathcal{P}$ of precisely N points in $[0,1]^2$. It can then be shown that the inequality (2.12) is valid, establishing Theorem 2.10.

Indeed, this approach has been known for some time, as Beck and Chen have already used this idea earlier in an alternative proof of Theorem 2.11. In fact, the original proof of Theorem 2.11 by Davenport 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.

Write $U = [0,1]^2$. 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 $\mathbb{R}^2$ generated by the two vectors

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

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

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

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

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

we consider the discrepancy

$$E[\mathcal{Q}_1; B(x_1,y)] = \#(\mathcal{Q}_1 \cap B(x_1,y)) - x_1 y, \tag{7.3}$$

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 $\mathcal{Q}_1$ per unit area in V. Suppose first of all that y is an integer satisfying $0 < y \leqslant M$. Then we can write

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

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

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

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

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

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

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

$$E[\mathcal{Q}_1; B(x_1, y)] \approx \sum_{0 \neq m \in \mathbb{Z}} \left(\frac{1 - e(-m x_1)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leqslant n < y} e(\theta n m) \right). \tag{7.4}$$

Ideally we would like to square the expression (7.4) 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 $\mathbb{R}^2$ generated by the two vectors

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

It is not difficult to see that the set

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

again contains precisely M points. Then the set

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

where the points are counted with multiplicity, contains precisely $2M$ points. Thus analogous to the discrepancy (7.3), we now consider the discrepancy

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

noting that there is now an average of two points of $\mathcal{Q}$ per unit area in V. The analogue of the estimate (7.4) is now

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

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

$$\int_0^1 |F[\mathcal{Q}; B(x_1, y)]|^2 \, \mathrm{d}x_1 \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leqslant n < y} e(\theta n m) \right|^2. \tag{7.5}$$

To estimate the sum on the right hand side of (7.5), 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

$$m \| m\theta \| > c > 0 \tag{7.6}$$

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

LEMMA 7.1. *Suppose that the real number θ is badly approximable. Then*

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leqslant n < y} e(\theta n m) \right|^2 \ll_\theta \log(2y). \tag{7.7}$$

PROOF. It is well known that

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

so that

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

The condition (7.6) implies that if $2^{h-1} \leqslant m < 2^h$, then the inequality

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

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

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

for otherwise the difference $(m_1 - m_2)$ of two of them would contradict (7.6). It follows that

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

as required. ◯

Combining (7.5) and (7.7) and then integrating with respect to the variable y over $[0, M]$, we have

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

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

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

of $N = 2M$ points in $[0,1)^2$ satisfies the conclusion of Theorem 2.11.

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.

7.2. Roth's Probabilistic Technique – A Preview

In this section, we describe an ingenious variation of Davenport's argument by Roth. This is a nice preview of his powerful probabilistic technique, which we shall describe in Section 7.4, 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 $\mathbb{R}^2$ generated by the two vectors $(1,0)$ and $(\theta,1)$. For any real number $t \in \mathbb{R}$, 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

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

which clearly contains precisely M points. Thus analogous to the discrepancy (7.3), we now consider the discrepancy

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

noting that there is now an average of one point of $\mathcal{Q}_1(t)$ per unit area in V. The analogue of the estimate (7.4) is now

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

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

$$\int_0^1 |E[\mathcal{Q}_1(t); B(x_1, y)]|^2 \, \mathrm{d}t \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leqslant n < y} e(\theta n m) \right|^2. \tag{7.8}$$

Furthermore, if θ is a badly approximable number as in the last section, then integrating (7.8) trivially with respect to the variable x_1 over $[0,1]$ and with respect to the variable y over $[0, M]$, we have

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

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

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

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

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

of $N = M$ points in $[0,1)^2$ satisfies the requirements of Theorem 2.11.

7.3. 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 2.10 and 2.11.

Our first step is to construct the simplest point sets which will allow us to establish Theorem 2.10.

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

$$n = \sum_{j=1}^{\infty} a_j 2^{j-1} \tag{7.9}$$

as a dyadic expansion. Then we write

$$c_n = \sum_{j=1}^{\infty} a_j 2^{-j}. \tag{7.10}$$

Note that $c_n \in [0,1)$. Note also that only finitely many digits $a_1, a_2, a_3, \dots$ are non-zero, so that the sums in (7.9) and (7.10) have only finitely many non-zero terms. For simplicity, we sometimes write

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

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

$$\mathcal{Q} = \{(c_n, n) : n = 0, 1, 2, \dots\} \tag{7.11}$$

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

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

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

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

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 7.2 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 7.3. *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 $\mathcal{Q}$.

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

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

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

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

LEMMA 7.4. *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 $|\alpha_0| \leqslant \frac{1}{2}$ and*

$$E[\mathcal{Q}; [\ell\, 2^{-s}, (\ell+1)2^{-s}) \times [0,y)] = \alpha_0 - \psi(2^{-s}(y - \beta_0)) \tag{7.12}$$

at all points of continuity of the right hand side.

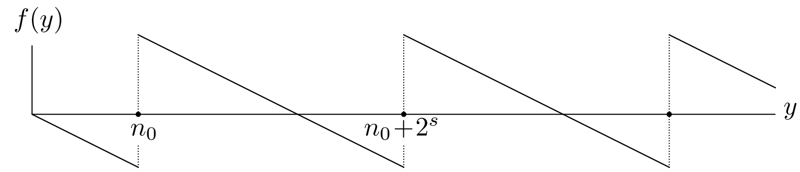
PROOF. In view of Lemma 7.2, the second coordinates of the points of $\mathcal{Q}$ in the region $[\ell\, 2^{-s}, (\ell+1)2^{-s}) \times [0,\infty)$ fall precisely into a residue class modulo 2^s. Let n_0 be the smallest of these second coordinates. Then $0 \leqslant n_0 < 2^s$. We now study

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

as a function of y. For simplicity, denote it by $f(y)$, say. Clearly $f(0) = E[\mathcal{Q};\emptyset] = 0$. 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 $\mathcal{Q}$ 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 $y = n_0$.



With suitable α_0 and β_0, the right hand side of (7.12) fits all the requirements. ○

We can now prove Theorem 2.10. Let $N \geqslant 2$ be a given integer. It follows immediately from the definition of $\mathcal{Q}$ that the set

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

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

$$2^{h-1} < N \leqslant 2^h. \tag{7.13}$$

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

$$\begin{aligned} E[\mathcal{Q}_0;[0,x_1)\times[0,y)] &= E[\mathcal{Q};[0,x_1)\times[0,y)] \\ &= E[\mathcal{Q};[x_1^{(h)},x_1)\times[0,y)] + \sum_{s=1}^{h} E[\mathcal{Q};[x_1^{(s-1)},x_1^{(s)})\times[0,y)]. \end{aligned} \tag{7.14}$$

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 7.3, hence contains precisely one point of $\mathcal{Q}$. It follows that

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

and we have the bound

$$|E[\mathcal{Q};[x_1^{(h)},x_1)\times[0,y)]| \leqslant 1. \tag{7.15}$$

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 $E[\mathcal{Q}; [x_1^{(s-1)}, x_1^{(s)}) \times [0, y)] = 0$ trivially, or is of the type under discussion in Lemma 7.4, and we have the bound

$$|E[\mathcal{Q}; [x_1^{(s-1)}, x_1^{(s)}) \times [0, y)]| \leqslant 1. \tag{7.16}$$

Note that (7.16) still holds in the empty case. Combining (7.13)–(7.16), we arrive at an upper bound

$$|E[\mathcal{Q}_0; [0, x_1) \times [0, y)]| \leqslant 1 + h \ll \log N. \tag{7.17}$$

For comparison later in Section 7.6, 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 7.3, and inequality (7.15) is the observation that

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

To estimate $E[\mathcal{Q}; B(x_1^{(h)}, y)]$, 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

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

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 7.4, each of which satisfies inequality (7.16).

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

$$\mathcal{P} = \{(c_n, N^{-1}n) : n = 0, 1, 2, \ldots, N-1\} \tag{7.18}$$

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

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

in view of (7.17) and noting that $0 \leqslant Nx_2 \leqslant N$. This now completes the proof of Theorem 2.10.

7.4. Roth's Probabilistic Technique

We now attempt to extend the ideas in the last section to obtain a proof of Theorem 2.11.

Let us first of all consider the special case when $N = 2^h$. Then the set (7.18) used to establish Theorem 2.10 becomes

$$\begin{aligned}\mathcal{P}(2^h) &= \{(c_n, 2^{-h}n) : n = 0, 1, 2, \ldots, 2^h - 1\} \\ &= \{(0.a_1a_2a_3\ldots a_h, 0.a_h\ldots a_3a_2a_1) : a_1, \ldots, a_h \in \{0,1\}\},\end{aligned} \tag{7.19}$$

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

THEOREM 7.5. *For every positive integer h, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}(2^h); B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = 2^{-6}h^2 + O(h). \tag{7.20}$$

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 (7.14) simplifies to

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

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 7.4 that

$$D[\mathcal{P}; B(x_1, 2^{-h}y)] = \sum_{s=1}^{h}{}^{*} (\alpha_s - \psi(2^{-s}(y - \beta_s))), \tag{7.21}$$

where, for each $s = 1, \ldots, h$, the real numbers α_s and β_s satisfy $|\alpha_s| \leqslant \frac{1}{2}$. If we square the expression (7.21), 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

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

which ultimately leads to the term $2^{-6}h^2$ in (7.20).

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 7.2, 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.

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

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

$$2^{h-1} < N \leqslant 2^h. \tag{7.22}$$

For every $n = 0, 1, 2, \ldots, 2^h - 1$, we define c_n as before by (7.9) and (7.10). We then extend the definition of c_n to all other integers using periodicity by writing

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

and consider the extended van der Corput point set

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

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

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

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

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

denote the discrepancy of $\mathcal{Q}_h(t)$ 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 (7.15), we have the trivial bound

$$|E[\mathcal{Q}_h(t); [x_1^{(h)}, x_1) \times [0, y)]| \leqslant 1, \tag{7.23}$$

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

$$E[\mathcal{Q}_h(t); B(x_1, y)] = \sum_{s=1}^{h}{}^{*} E[\mathcal{Q}_h(t); [x_1^{(s-1)}, x_1^{(s)}) \times [0, y)]. \tag{7.24}$$

Corresponding to Lemma 7.4, we can establish the following result without too much difficulty.

Lemma 7.6. *For all positive real numbers y and all non-negative integers s and ℓ such that $s \leqslant h$ and $\ell < 2^s$ hold, there exist real numbers β_0 and γ_0, depending at most on s, ℓ and y, such that*

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

at all points of continuity of the right hand side.

Combining (7.24) and Lemma 7.6, we have

$$E[\mathcal{Q}_h(t); B(x_1, y)] = \sum_{s=1}^{h}{}^{*} (\psi(2^{-s}(t - \beta_s)) - \psi(2^{-s}(t - \gamma_s))) \tag{7.25}$$

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 $\mathcal{Q}_h(t)$. 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 7.7. *Suppose that the integers s' and s'' satisfy $0 \leqslant s', s'' \leqslant h$, 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

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

with the last equality arising from the observation that

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

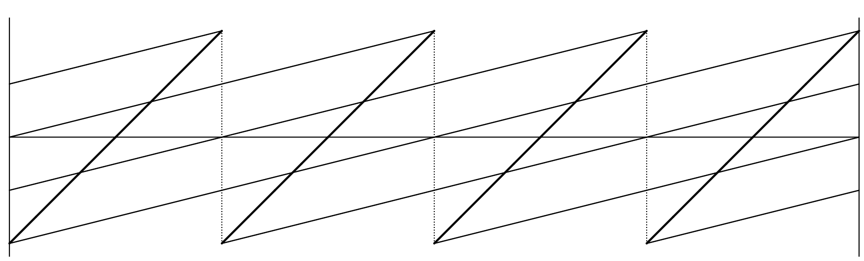
It follows that

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

It is not difficult to see that

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

at all points of continuity.



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 (7.25) and Lemma 7.7 that

$$\int_0^{2^h}|E[\mathcal{Q}_h(t);B(x_1,y)]|^2\,\mathrm{d}t\ll\sum_{s'=1}^{h}{}^{*}\sum_{s''=1}^{h}{}^{*}2^{h-|s'-s''|}\ll 2^h h, \tag{7.26}$$

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

Note that the estimate (7.26) is independent of the choice of x_1 and y. We also recall the trivial estimate (7.23). It follows that integrating (7.26) trivially with respect to x_1 over the interval $[0,1)$ and with respect to y over the interval $[0,N)$, we conclude that

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

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

$$\int_0^N\int_0^1|E[\mathcal{Q}_h(t^*);B(x_1,y)]|^2\,\mathrm{d}x_1\,\mathrm{d}y\ll hN. \tag{7.27}$$

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

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

contains precisely N points in $[0,1)^2$, and the estimate (7.27) now translates to

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

in view of (7.22). This completes the proof of Theorem 2.11.

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

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

in the sum (7.24). 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

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

for some even integer ℓ. One can then show that

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

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 7.2 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}.

7.5. 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.

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

$$2^{h-1} < N \leqslant 2^h. \tag{7.28}$$

For every $n = 0, 1, 2, \ldots, 2^h - 1$, we define c_n as before by (7.9) and (7.10). 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]

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

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

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

obtained from $\mathcal{Q}_h$ 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 $\mathcal{Q}_h^{(\mathbf{t})}$ 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

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

Consider a rectangle of the form

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

Analogous to (7.23), we have the trivial bound

$$|E[\mathcal{Q}_h^{(\mathbf{t})}; [x_1^{(h)}, x_1) \times [0, y)]| \leqslant 1, \tag{7.29}$$

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

$$E[\mathcal{Q}_h^{(\mathbf{t})}; B(x_1, y)] = \sum_{s=1}^{h}{}^{*}\, E[\mathcal{Q}_h^{(\mathbf{t})}; [x_1^{(s-1)}, x_1^{(s)}) \times [0, y)]. \tag{7.30}$$

---

[4]This is not the case if we wish to study higher dimensional analogues of Theorem 2.11.

[5]Note that the set $\mathcal{Q}_h$ 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 $\mathcal{Q}_h$ remain valid for the set $\mathcal{Q}_h$ defined in the last section.

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

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

Then we need to study sums of the form

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

Analogous to Lemma 7.7, we have the following estimate.

LEMMA 7.8. *Suppose that the real number $y \in [0, N)$ is fixed, and that the integers s' and s'' satisfy $0 \leqslant s', s'' \leqslant h$. Then*

$$\sum_{\mathbf{t} \in \mathbb{Z}_2^h} E_{s'}[t_1, \ldots, t_h] E_{s''}[t_1, \ldots, t_h] = O(2^{h - |s' - s''|}). \tag{7.31}$$

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

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

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 $x_1^{(s'-1)} \neq x_1^{(s')}$ and $x_1^{(s''-1)} \neq x_1^{(s'')}$. Now the case when $s' = s''$ is easy, since we have $E[t_1, \ldots, t_h; x_1^{(s-1)}, x_1^{(s)}] = O(1)$ trivially. Without loss of generality, let us assume that $s' > s''$. For fixed $t_1, \ldots, t_{s''}$, in view of the comment at the beginning of the proof, we have

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

We shall show that

$$\begin{aligned} &\sum_{t_{s''+1}, \ldots, t_{s'} \in \mathbb{Z}_2} E_{s'}[t_1, \ldots, t_{s'}, 0, \ldots, 0] \\ &\qquad = \sum_{t_{s''+1}, \ldots, t_{s'} \in \mathbb{Z}_2} E[\mathcal{Q}_h^{(t_1, \ldots, t_{s''}, t_{s''+1}, \ldots, t_{s'}, 0, \ldots, 0)}; [x_1^{(s'-1)}, x_1^{(s')}) \times [0, y)] \\ &\qquad = E[\mathcal{Q}_h^{(t_1, \ldots, t_{s''}, 0, \ldots, 0)}; [\ell\, 2^{-s''}, (\ell+1) 2^{-s''}) \times [0, y)], \end{aligned} \tag{7.32}$$

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

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

from which it follows that

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

giving the desired result. To establish (7.32), 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 $[\ell\, 2^{-s''}, (\ell+1)2^{-s''})$. ○

It now follows from (7.30) and Lemma 7.8 that

$$\sum_{\mathbf{t}\in\mathbb{Z}_2^h} |E[\mathcal{Q}_h^{(\mathbf{t})}; B(x_1, y)]|^2 \ll \sum_{s'=1}^{h}{}^{*} \sum_{s''=1}^{h}{}^{*} 2^{h-|s'-s''|} \ll 2^h h, \tag{7.33}$$

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

Note that the estimate (7.33) is independent of the choice of x_1 and y. We also recall the trivial estimate (7.29). It follows that integrating (7.33) trivially with respect to x_1 over the interval $[0,1)$ and with respect to y over the interval $[0,N)$, we conclude that

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

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

$$\int_0^N \int_0^1 |E[\mathcal{Q}_h^{(\mathbf{t}^*)}; B(x_1, y)]|^2 \,\mathrm{d}x_1\,\mathrm{d}y \ll hN. \tag{7.34}$$

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

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

contains precisely N points in $[0,1)^2$, and the estimate (7.34) now translates to

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

in view of (7.28). This completes the proof of Theorem 2.11.

7.6. Generalizations of van der Corput Point Sets

In our discussion of the van der Corput sequence and van der Corput point sets in Sections 7.3 and 7.4, we have restricted our discussion to dimension 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 generalizations of Theorems 2.10 and 2.11 to higher dimensions, one way is to generalize the van der Corput sequence. Here we know two ways of doing so, one by Halton and the other by Faure. The Halton construction enables Halton to establish Theorem 2.15 and forms the basis for the proof of Theorem 2.14 in the case $q = 2$ by Roth and in general by Chen. The Faure construction enables

Faure to give an alternative proof of Theorem 2.15, enables Chen soon afterwards to give an alternative proof of Theorem 2.14 and, more recently, forms the basis for the explicit construction proof of Theorem 2.14 in the case $q = 2$ by Chen and Skriganov and in general by Skriganov.

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.

7.6.1. Halton Point Sets. We first discuss Halton's contribution. Recall the dyadic construction (7.9) and (7.10) of the classical van der Corput sequence. Suppose now that we wish to study the higher dimensional analogues of Theorem 2.10 or 2.11. 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 $n \in \mathbb{N}_0$ and every $i = 1, \ldots, K-1$, we write

$$n = \sum_{j=1}^{\infty} a_j^{(i)} p_i^{j-1} \tag{7.35}$$

as a p_i-adic expansion. Then we write

$$c_n^{(i)} = \sum_{j=1}^{\infty} a_j^{(i)} p_i^{-j}. \tag{7.36}$$

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

$$\mathcal{H} = \{(\mathbf{c}_n, n) : n = 0, 1, 2, \ldots\} \tag{7.37}$$

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

Corresponding to Lemma 7.2, we have the following multi-dimensional version.

LEMMA 7.9. *For all non-negative integers $s_1, \ldots, s_{K-1}$ and $\ell_1, \ldots, \ell_{K-1}$ such that $\ell_i < p_i^{s_i}$ holds for every $i = 1, \ldots, K-1$, the set*

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

contains precisely all the elements of a residue class modulo $p_1^{s_1} \ldots p_{K-1}^{s_{K-1}}$ in $\mathbb{N}_0$.

PROOF. For fixed $i = 1, \ldots, K-1$, the p_i-adic version of Lemma 7.2 says that the set

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

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

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 $\ell_1,\ldots,\ell_{K-1}$ is an elementary $(p_1,\ldots,p_{K-1})$-adic box of volume $p_1^{-s_1}\ldots p_{K-1}^{-s_{K-1}}$. Hence Lemma 7.9 says that the given Halton sequence has very good distribution among such elementary $(p_1,\ldots,p_{K-1})$-adic boxes for all non-negative integer values of $s_1,\ldots,s_{K-1}$.

LEMMA 7.10. *For all non-negative integers $s_1,\ldots,s_{K-1}$, $\ell_1,\ldots,\ell_{K-1}$ and m such that $\ell_i<p_i^{s_i}$ holds 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 the Halton point set $\mathcal{H}$.

Clearly there is an average of one point of the Halton point set $\mathcal{H}$ 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

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

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

We have the following generalization of Lemma 7.4.

LEMMA 7.11. *For all non-negative integers $s_1,\ldots,s_{K-1}$ and $\ell_1,\ldots,\ell_{K-1}$ such that $\ell_i<p_i^{s_i}$ holds for every $i=1,\ldots,K-1$, there exist real numbers α_0,β_0, depending at most on $s_1,\ldots,s_{K-1}$ and $\ell_1,\ldots,\ell_{K-1}$, such that $|\alpha_0|\leqslant\frac{1}{2}$ and*

$$E\left[\mathcal{H};\prod_{i=1}^{K-1}[\ell_i\,p_i^{-s_i},(\ell_i+1)p_i^{-s_i})\times[0,y)\right]=\alpha_0-\psi(p_1^{-s_1}\ldots p_{K-1}^{-s_{K-1}}(y-\beta_0))\tag{7.38}$$

at all points of continuity of the right hand side.

We can now prove Theorem 2.15. Let $N\geqslant 2$ be a given integer. It follows at once from the definition of $\mathcal{H}$ that the set

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

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

$$p_1^{h-1}<N\leqslant p_1^h.\tag{7.39}$$

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 7.3, we shall approximate each interval $[0,x_i)$, where $i=1,\ldots,K-1$, by the subinterval $[0,x_i^{(h)})$, where $x_i^{(h)}=p_i^{-h}[p_i^hx_i]$ 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 7.3 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 7.10, and so

$$|E[\mathcal{H};B(x_1,\ldots,x_{K-1},y)]-E[\mathcal{H};B(x_1^{(h)},\ldots,x_{K-1}^{(h)},y)]|\leqslant K-1;\tag{7.40}$$

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

It remains to estimate $E[\mathcal{H}; B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)]$. 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 $0 \leqslant s \leqslant h$.

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 $0 \leqslant x_i^{(h)} < 1$.

LEMMA 7.12. *Suppose that* $0 \leqslant x_i^{(h)} < 1$, *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*

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

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.*

It follows that the set $B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)$ is a disjoint union of fewer than $h^{K-1} p_1 \ldots p_{K-1}$ sets of the type discussed in Lemma 7.11. Hence

$$|E[\mathcal{H}; B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)]| < h^{K-1} p_1 \ldots p_{K-1} \ll_K (\log N)^{K-1}. \tag{7.41}$$

Combining (7.40) and (7.41), we conclude that

$$|E[\mathcal{H}; B(x_1, \ldots, x_{K-1}, y)]| \ll_K (\log N)^{K-1}. \tag{7.42}$$

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

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

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

$$D[\mathcal{P}; B(\mathbf{x})] = E[\mathcal{H}_0; [0, x_1) \times \ldots \times [0, x_{K-1}) \times [0, N x_K)] \ll_K (\log N)^{K-1},$$

in view of (7.42) and noting that $0 \leqslant N x_K \leqslant N$. This now completes the proof of Theorem 2.15.

Next we discuss Roth's ideas in shaping this Halton construction to give a proof of Theorem 2.14 in the case $q = 2$. As in the proof of Theorem 2.11, one needs to introduce a probabilistic variable. To pave the way for this, we shall modify the Halton point set somewhat. Let $N \geqslant 2$ be a given integer, and let the integer h be determined uniquely by

$$p_1^{h-1} < N \leqslant p_1^h, \tag{7.43}$$

as before. For every $i = 1, \ldots, K-1$ and every $n = 0, 1, 2, \ldots, p_i^h - 1$, we define $c_n^{(i)}$ as before by (7.35) and (7.36). We then extend the definition of $c_n^{(i)}$ to all other integers using periodicity by writing

$$c_{n+p_i^h}^{(i)} = c_n^{(i)} \quad \text{for every } n \in \mathbb{Z},$$

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

$$\mathcal{H}_h = \{(\mathbf{c}_n, n) : n \in \mathbb{Z}\}.$$

REMARK. In the original proofs of Roth and Chen, the construction of the set $\mathcal{H}_h$ is slightly different, but the difference does not affect the argument in any way. Let $M = p_1 \ldots p_{K-1}$. One then defines $c_n^{(i)}$ for $n = 0, 1, 2, \ldots, M^h - 1$ by (7.35) and (7.36), write $\mathbf{c}_n = (c_n^{(1)}, \ldots, c_n^{(K-1)})$ 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 $n \in \mathbb{Z}$.

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

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

It is clear that there is an average of one point of the translated Halton point set $\mathcal{H}_h(t)$ 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

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

denote the discrepancy of $\mathcal{H}_h(t)$ 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 2.15, we shall consider the smaller rectangular box $B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)$ and, corresponding to (7.40), we have

$$|E[\mathcal{H}_h(t); B(x_1, \ldots, x_{K-1}, y)] - E[\mathcal{H}; B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)]| \leqslant K - 1. \tag{7.44}$$

Next, we study $E[\mathcal{H}_h(t); B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)]$ in detail, and require an analogue of the expansion (7.24). It is not difficult to see that

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

where $\mathbf{I} = I_1 \times \ldots \times I_{K-1}$ and where, for every $i = 1, \ldots, K-1$, $\mathcal{I}_i$ denotes the collection of elementary p_i-adic intervals in the union that makes up the interval $[0, x_i^{(h)})$ in Lemma 7.12.

Corresponding to Lemma 7.6, one can show that each summand

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

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 7.7, that

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

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 $I_i', I_i'' \in \mathcal{I}_i$ have lengths $p_i^{-s_i'}$ and $p_i^{-s_i''}$ respectively. One then goes on to show that

$$\begin{aligned}
&\int_0^{M^h} |E[\mathcal{H}; B(x_1^{(h)}, \ldots, x_{K-1}^{(h)}, y)]|^2 \,\mathrm{d}t \\
&\quad \ll \sum_{I_1' \in \mathcal{I}_1} \cdots \sum_{I_{K-1}' \in \mathcal{I}_{K-1}} \sum_{I_1'' \in \mathcal{I}_1} \cdots \sum_{I_{K-1}'' \in \mathcal{I}_{K-1}} M^h \prod_{i=1}^{K-1} p_i^{-|s_i' - s_i''|} \\
&\quad \ll_K M^h h^{K-1}.
\end{aligned}$$

Taking the bound (7.44) into account and integrating trivially with respect to $x_1, \ldots, x_{K-1}$, each over the interval $[0, 1)$, and with respect to y over the interval $[0, N)$, we conclude that

$$\begin{aligned}
&\int_0^N \int_0^1 \cdots \int_0^1 \int_0^{M^h} |E[\mathcal{H}_h(t); B(x_1, \ldots, x_{K-1}, y)]|^2 \,\mathrm{d}t \,\mathrm{d}x_1 \ldots \mathrm{d}x_{K-1} \,\mathrm{d}y \\
&\quad = \int_0^{M^h} \left(\int_0^N \int_0^1 \cdots \int_0^1 |E[\mathcal{H}_h(t); B(x_1, \ldots, x_{K-1}, y)]|^2 \,\mathrm{d}x_1 \ldots \mathrm{d}x_{K-1} \,\mathrm{d}y \right) \mathrm{d}t \\
&\quad \ll_K M^h h^{K-1} N.
\end{aligned}$$

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

$$\begin{aligned}
(7.45) \qquad &\int_0^N \int_0^1 \cdots \int_0^1 |E[\mathcal{H}_h(t^*); B(x_1, \ldots, x_{K-1}, y)]|^2 \,\mathrm{d}x_1 \ldots \mathrm{d}x_{K-1} \,\mathrm{d}y \\
&\quad \ll_K h^{K-1} N.
\end{aligned}$$

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

$$\mathcal{P}^* = \{(z_1, \ldots, z_{K-1}, N^{-1} z_K) : (z_1, \ldots, z_K) \in \mathcal{H}_h(t^*)\}$$

contains precisely N points in $[0,1)^K$, and the estimate (7.45) now translates to

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

in view of (7.43). This completes our brief sketch of the proof of Theorem 2.14 in the case $q = 2$.

7.6.2. Faure Point Sets. We now discuss Faure's contribution. Suppose again that we wish to study the higher dimensional analogues of Theorem 2.10 or 2.11. Let p denote a prime such that $p \geqslant K - 1$, a condition that cannot be relaxed. For every non-negative integer $n \in \mathbb{N}_0$, we write

$$(7.46) \qquad n = \sum_{j=1}^{\infty} a_j^{(1)} p^{j-1}$$

as a p-adic expansion. Then we write

$$(7.47) \qquad c_n^{(1)} = \sum_{j=1}^{\infty} a_j^{(1)} p^{-j}.$$

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

$$c_n^{(i)} = \sum_{j=1}^{\infty} a_j^{(i)} p^{-j}, \tag{7.48}$$

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

$$\mathcal{B} = \begin{bmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \binom{3}{0} & \cdots \\ & \binom{1}{1} & \binom{2}{1} & \binom{3}{1} & \cdots \\ & & \binom{2}{2} & \binom{3}{2} & \cdots \\ & & & \binom{3}{3} & \cdots \\ & & & & \ddots \end{bmatrix} \tag{7.49}$$

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

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

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 $n \in \mathbb{N}_0$, write

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

The set

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

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

Analogous to Lemma 7.10, we have the following result.

LEMMA 7.13. *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*

$$\prod_{i=1}^{K-1}[\ell_i\,p^{-s_i},(\ell_i+1)p^{-s_i})\times[mp^{s_1+\ldots+s_{K-1}},(m+1)p^{s_1+\ldots+s_{K-1}}) \tag{7.50}$$

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

To prove Lemma 7.13, we need a simple result concerning the matrix $\mathcal{B}$.

LEMMA 7.14. *Let the matrix $\mathcal{B}$ be given by (7.49). For every $i=1,\ldots,K-1$, we have*

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

PROOF OF LEMMA 7.13. Suppose that the 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 (7.50), we must have

$$c_n^{(i)}\in[\ell_i\,p^{-s_i},(\ell_i+1)p^{-s_i}) \tag{7.51}$$

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

$$n\in[mp^{s_1+\ldots+s_{K-1}},(m+1)p^{s_1+\ldots+s_{K-1}}). \tag{7.52}$$

Comparing (7.46) and (7.52), 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 (7.51) 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 (7.51) holds, then in view of (7.47) or (7.48), we must have $a_j^{(i)} = \beta_j^{(i)}$ for every $j = 1, \ldots, s_i$. This can be summarized by writing

$$\mathcal{W}_i \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \\ a_3^{(1)} \\ a_4^{(1)} \\ \vdots \end{bmatrix} \equiv \begin{bmatrix} \beta_1^{(i)} \\ \beta_2^{(i)} \\ \beta_3^{(i)} \\ \vdots \\ \beta_{s_i}^{(i)} \end{bmatrix} \bmod p, \tag{7.53}$$

where the matrix $\mathcal{W}_i$ contains precisely the first s_i rows of the matrix $\mathcal{B}^{i-1}$. Now recall that $a_j^{(1)}$ are already uniquely determined for every $j > S = s_1 + \ldots + s_{K-1}$ by (7.52), and clearly there are at most finitely many non-zero terms among these. The system (7.53) can therefore be simplified to one of the form

$$\mathcal{V}_i \begin{bmatrix} a_1^{(1)} \\ a_2^{(1)} \\ a_3^{(1)} \\ \vdots \\ a_S^{(1)} \end{bmatrix} \equiv \begin{bmatrix} \gamma_1^{(i)} \\ \gamma_2^{(i)} \\ \gamma_3^{(i)} \\ \vdots \\ \gamma_{s_i}^{(i)} \end{bmatrix} \bmod p, \tag{7.54}$$

where the matrix $\mathcal{V}_i$ contains precisely the first S columns of the matrix $\mathcal{W}_i$. On combining (7.54) 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

$$\mathcal{V} = \begin{bmatrix} \mathcal{V}_1 \\ \vdots \\ \mathcal{V}_{K-1} \end{bmatrix}.$$

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

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

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

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

in view of the assumption that $p \geqslant K-1$. 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. ◯

Analogous to Lemma 7.11, we have the following straightforward consequence of Lemma 7.14.

LEMMA 7.15. *For all non-negative integers $s_1, \ldots, s_{K-1}$ and $\ell_1, \ldots, \ell_{K-1}$ such that $\ell_i < p^{s_i}$ holds for every $i = 1, \ldots, K-1$, and for every real number $y > 0$, we have*

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

To study Theorem 2.15, let $N \geqslant 2$ be a given integer. It follows at once from the definition of $\mathcal{F}$ that the set

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

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

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

We can now deduce Theorem 2.15 from Lemmas 7.13 and 7.15 in a way similar to our deduction of the same result from Lemmas 7.10 and 7.11 in Section 7.6.1, noting that Lemma 7.12 remains valid with p_i replaced by p. Indeed, rescaling the second coordinate of the points of $\mathcal{F}_0$ by a factor N^{-1}, we obtain a set

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

of precisely N points in $[0,1)^K$ and which satisfies the conclusion of Theorem 2.15.

7.6.3. A General Point Set and a Digit Shift Argument. In this section, we briefly describe a rather general digit shift argument developed by Chen which enables us to establish Theorem 2.14 using Halton point sets discussed in Section 7.6.1 or Faure point sets discussed in Section 7.6.2. Recall that these point sets satisfy Lemma 7.10 and Lemma 7.13 respectively.

We shall only discuss the special case $q = 2$.

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

$$\mathcal{Z} = \{(\mathbf{c}_n, n) : n = 0,1,2,\ldots\} \tag{7.55}$$

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 such that $s_i \leqslant h$ and $\ell_i < p_i^{s_i}$ hold 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 $\mathcal{Z}$.

If the primes $p_1, \ldots, p_{K-1}$ are distinct, then the Halton set $\mathcal{H}$ 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 $\mathcal{F}$ is 1-set of every non-negative order with respect to $p, \ldots, p$, provided that $p \geqslant K-1$.

The property below follows almost immediately from the definition.

LEMMA 7.16. *Suppose that h be a non-negative integer, and that $\mathcal{Z}$ 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 $\ell_1, \ldots, \ell_{K-1}$ such that $s_i \leqslant h$ and $\ell_i < p_i^{s_i}$ hold for every $i = 1, \ldots, K-1$, and for every real number $y > 0$, we have*

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

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

$$p_1^{h-1} < N \leqslant p_1^h. \tag{7.56}$$

For any 1-set (7.55) of order h with respect to the primes $p_1, \ldots, p_{K-1}$, the set

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

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

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

of precisely N points in $[0,1)^K$ and which satisfies the conclusion of Theorem 2.15.

To study Theorem 2.14 in the case $q = 2$, we again choose the integer h to satisfy (7.56). However, we need to modify the 1-set $\mathcal{Z}$.

Let $\mathcal{M}$ 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 $\mathcal{M}$ 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 $\mathbf{T} \in \mathcal{M}$ 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

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

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 2.14 in the case $q = 2$, 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 $x_i^{(h)} = p_i^{-h}[p_i^h x_i]$, the greatest integer multiple of p_i^{-h} not exceeding x_i. Then for every $\mathbf{T} \in \mathcal{M}$, we have

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

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

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

from which it follows that

$$\begin{aligned}
&\int_0^N \int_0^1 \dots \int_0^1 \left(\sum_{\mathbf{T}\in\mathcal{M}} |E[\mathcal{Z}^{\mathbf{T}}; B(x_1^{(h)}, \dots, x_{K-1}^{(h)}, y)]|^2 \right) \mathrm{d}x_1 \dots \mathrm{d}x_{K-1}\, \mathrm{d}y \\
&\quad = \sum_{\mathbf{T}\in\mathcal{M}} \left(\int_0^N \int_0^1 \dots \int_0^1 |E[\mathcal{Z}^{\mathbf{T}}; B(x_1^{(h)}, \dots, x_{K-1}^{(h)}, y)]|^2 \, \mathrm{d}x_1 \dots \mathrm{d}x_{K-1}\, \mathrm{d}y \right) \\
&\quad \ll_K (p_1 \dots p_{K-1})^h h^{K-1} N.
\end{aligned}$$

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

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

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

$$\mathcal{P}^* = \{(z_1, \dots, z_{K-1}, N^{-1} z_K) : (z_1, \dots, z_K) \in \mathcal{Z}^{\mathbf{T}^*}\}$$

contains precisely N points in $[0,1)^K$, and satisfies the conclusion of Theorem 2.14 in the case $q = 2$.

CHAPTER 8

The Disc Segment Problem

8.1. Alexander's Technique

In this section, we introduce the integral geometric approach of Alexander and study Roth's disc segment problem first described in Section 3.2.

Suppose that U is the closed disc of unit area in $\mathbb{R}^2$, centred at the origin. Then any disc segment S is simply the intersection of U with a half plane. Theorem 3.9 says that for any distribution $\mathcal{P}$ of N points in U, there always exists a disc segment S with discrepancy $|D[\mathcal{P};S]| \gg N^{\frac{1}{4}}$.

We comment here that a slightly weaker estimate given in Theorem 3.8 can be established using a variant of the Fourier transform technique described earlier in Chapter 6, applied to $\mathbb{R}^2$ rather than the torus $\mathbb{T}^2$, and also involving smoothing arguments. The idea of Beck is to show that there is a thin rectangle in $\mathbb{R}^2$ that has large discrepancy and cuts U, so that the complement of this rectangle in U is a union of two disc segments. Then at least one of these two disc segments must have large discrepancy.

The technique of Alexander is based on the following well known result in integral geometry. There is a motion invariant Borel measure λ_K on the hyperplanes h of euclidean space $\mathbb{R}^K$ such that

$$|\mathbf{u}-\mathbf{v}| = \frac{1}{2}\lambda_K(\{h : h \text{ cuts } \overline{\mathbf{uv}}\}), \tag{8.1}$$

where for every points $\mathbf{u},\mathbf{v} \in \mathbb{R}^K$, $|\mathbf{u}-\mathbf{v}|$ denotes the euclidean distance between $\mathbf{u}$ and $\mathbf{v}$, and $\overline{\mathbf{uv}}$ denotes the open line segment with endpoints $\mathbf{u}$ and $\mathbf{v}$.

Suppose that τ is a signed Borel measure with compact support in euclidean space $\mathbb{R}^K$. Consider the functional

$$I(\tau) = \int_{\mathbb{R}^K}\int_{\mathbb{R}^K} |\mathbf{u}-\mathbf{v}| \,\mathrm{d}\tau(\mathbf{u})\,\mathrm{d}\tau(\mathbf{v}).$$

The Crofton formula (8.1) leads to a representation of $I(\tau)$ as an integral with respect to the measure λ_K, of the form

$$I(\tau) = \int_{\mathcal{H}_K} \tau(h^+)\tau(h^-)\,\mathrm{d}\lambda_K(h), \tag{8.2}$$

where $\mathcal{H}_K$ represents the set of all hyperplanes of $\mathbb{R}^K$, and h^+, h^- denote the two open half spaces determined by the hyperplane h. To see this, note that in view of (8.1), we have

$$I(\tau) = \frac{1}{2}\int_{\mathcal{H}_K}\int_{\mathbb{R}^K}\int_{\mathbb{R}^K} \chi(\mathbf{u},\mathbf{v},h)\,\mathrm{d}\tau(\mathbf{u})\,\mathrm{d}\tau(\mathbf{v})\,\mathrm{d}\lambda_K(h), \tag{8.3}$$

where

$$\chi(\mathbf{u},\mathbf{v},h)=\begin{cases}1, & \text{if } h \text{ intersects } \overline{\mathbf{uv}} \text{ at precisely one point,}\\ 0, & \text{otherwise.}\end{cases}$$

Suppose that h is a given hyperplane in $\mathbb{R}^K$. Consider the inner integral

$$\int_{\mathbb{R}^K}\int_{\mathbb{R}^K}\chi(\mathbf{u},\mathbf{v},h)\,\mathrm{d}\tau(\mathbf{u})\,\mathrm{d}\tau(\mathbf{v}). \tag{8.4}$$

Clearly h intersects the open line segment $\overline{\mathbf{uv}}$ at precisely one point if and only if $\mathbf{u}$ and $\mathbf{v}$ are in different open half spaces determined by h. It follows that the integral (8.4) must be equal to $2\tau(h^+)\tau(h^-)$. Substituting this into (8.3) leads immediately to the formula (8.2).

Consider the special case when $\tau(\mathbb{R}^K)=0$. Clearly $\tau(h^+)+\tau(h^-)=\tau(\mathbb{R}^K)$ for almost all hyperplanes h in $\mathbb{R}^K$, and so it follows from (8.2) that

$$I(\tau)=-\int_{\mathcal{H}_K}|\tau(h^+)|^2\,\mathrm{d}\lambda_K(h)\leqslant 0.$$

Suppose that τ,τ' are signed Borel measures with compact support in euclidean space $\mathbb{R}^K$. We shall also consider the functional

$$J(\tau,\tau')=\int_{\mathbb{R}^K}\int_{\mathbb{R}^K}|\mathbf{u}-\mathbf{v}|\,\mathrm{d}\tau(\mathbf{u})\,\mathrm{d}\tau'(\mathbf{v}).$$

The need for this extra functional will be clear from (8.5) below.

Suppose now that U is the closed disc of unit area in $\mathbb{R}^2$, centred at the origin. Then any disc segment in U can be represented in the form $U\cap h^+$, where h is a line in $\mathbb{R}^2$. Suppose further that $\mathcal{P}$ is a distribution of N points in U. We consider the signed Borel measure $\sigma=\sigma_1-\sigma_2$, where σ_1 is the discrete measure with support $\mathcal{P}$, satisfying $\sigma_1(\mathbf{x})=1$ for every $\mathbf{x}\in\mathcal{P}$, and where $\sigma_2(S)=N\mu(U\cap S)$ for any Borel set S in $\mathbb{R}^2$. In other words, σ_2 is equal to N times the usual Lebesgue area measure μ in $\mathbb{R}^2$ restricted to U. It is easy to see that for every line h in $\mathbb{R}^2$, the quantity

$$\sigma(h^+)=\#(h^+\cap\mathcal{P})-N\mu(U\cap h^+)$$

represents the discrepancy of the disc segment $U\cap h^+$. We therefore need to find lower bounds for the quantity $|I(\sigma)|$, where

$$I(\sigma)=-\int_{\mathcal{H}_2}|\sigma(h^+)|^2\,\mathrm{d}\lambda_2(h).$$

To establish Theorem 3.9, it clearly suffices to show that $|I(\sigma)|>cN^{\frac{1}{2}}$, where $c>0$ is an absolute constant. Note that $\sigma(\mathbb{R}^2)=0$, and so we have $I(\sigma)\leqslant 0$.

Using Fubini's theorem, one can write

$$\begin{aligned}I(\sigma_1-\sigma_2)&=\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}(\sigma_1-\sigma_2)(\mathbf{x})\,\mathrm{d}(\sigma_1-\sigma_2)(\mathbf{y})\\&=\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\,\mathrm{d}\sigma_1(\mathbf{y})+\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_2(\mathbf{x})\,\mathrm{d}\sigma_2(\mathbf{y})\\&\quad-2\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\,\mathrm{d}\sigma_2(\mathbf{y})\\&=I(\sigma_1)+I(\sigma_2)-2\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\,\mathrm{d}\sigma_2(\mathbf{y}).\end{aligned} \tag{8.5}$$

Note that since $\sigma_1(\mathbb{R}^2) = \sigma_2(\mathbb{R}^2) = N \neq 0$, we do not have good control over the signs of $I(\sigma_1)$ and $I(\sigma_2)$. To handle this problem, we introduce a discrete measure Φ on the set $\mathbb{R}$ with support $\{r_1, \ldots, r_\ell\}$, to be specified later, such that

$$\sum_{t=1}^{\ell} |\Phi(r_t)| = 1, \tag{8.6}$$

and consider the product measure $\sigma \times \Phi$ on $\mathbb{R}^3$, defined by

$$\sigma \times \Phi = \sum_{t=1}^{\ell} \Phi(r_t)\sigma^{(t)}, \tag{8.7}$$

where, for every $t = 1, \ldots, \ell$, the measure $\sigma^{(t)}$ in $\mathbb{R}^3$ is supported by the set $U \times \{r_t\}$, with

$$\sigma^{(t)}(S \times \{r_t\}) = \sigma(S) \tag{8.8}$$

for every Borel set S in $\mathbb{R}^2$.

For every $t = 1, \ldots, \ell$, it is easy to see that

$$\sigma^{(t)}(\mathbb{R}^3) = \sigma^{(t)}(\mathbb{R}^2 \times \{r_t\}) = \sigma(\mathbb{R}^2) = 0 \tag{8.9}$$

and

$$\begin{aligned} I(\sigma^{(t)}) &= \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |(\mathbf{x}, r_t) - (\mathbf{y}, r_t)| \, \mathrm{d}\sigma^{(t)}(\mathbf{x}, r_t)\mathrm{d}\sigma^{(t)}(\mathbf{y}, r_t) \\ &= \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} |\mathbf{x} - \mathbf{y}| \, \mathrm{d}\sigma(\mathbf{x})\mathrm{d}\sigma(\mathbf{y}) = I(\sigma). \end{aligned} \tag{8.10}$$

LEMMA 8.1. *Suppose that* $|a_1| + \ldots + |a_\ell| = 1$. *Then*

$$-I\left(\sum_{t=1}^{\ell} a_t \sigma^{(t)}\right) \leqslant -\sum_{t=1}^{\ell} |a_t| I(\sigma^{(t)}).$$

PROOF. Suppose first of all that $a_1, \ldots, a_\ell$ are all non-negative. In view of (8.9), we have

$$-I\left(\sum_{t=1}^{\ell} a_t \sigma^{(t)}\right) = \int_{\mathcal{H}_2} \left(\sum_{t=1}^{\ell} a_t \sigma^{(t)}(h^+)\right)^2 \mathrm{d}\lambda_2(h).$$

Here we have used the fact that the measure $\sigma^{(t)}$ in $\mathbb{R}^3$ is concentrated on the set $\mathbb{R}^2 \times \{r_t\}$ and, in view of (8.8), is essentially the same as the measure σ in $\mathbb{R}^2$. Using the Cauchy-Schwarz inequality, we have

$$\left(\sum_{t=1}^{\ell} a_t \sigma^{(t)}(h^+)\right)^2 \leqslant \left(\sum_{t=1}^{\ell} a_t\right)\left(\sum_{t=1}^{\ell} a_t |\sigma^{(t)}(h^+)|^2\right),$$

so it follows that

$$-I\left(\sum_{t=1}^{\ell} a_t \sigma^{(t)}\right) \leqslant \left(\sum_{t=1}^{\ell} a_t\right) \sum_{t=1}^{\ell} a_t \int_{\mathcal{H}_2} |\sigma^{(t)}(h^+)|^2 \, \mathrm{d}\lambda_2(h) = -\sum_{t=1}^{\ell} a_t I(\sigma^{(t)}),$$

again in view of (8.9). The general case follows on noting that if $a_t < 0$, then $a_t\sigma^{(t)} = |a_t|(-\sigma^{(t)})$ and $I(-\sigma^{(t)}) = I(\sigma^{(t)})$. ○

Combining (8.6), (8.7), (8.10) and Lemma 8.1, we have

$$-I(\sigma \times \Phi) = -I\left(\sum_{t=1}^{\ell} \Phi(r_t)\sigma^{(t)}\right) \leqslant -\sum_{t=1}^{\ell} |\Phi(r_t)| I(\sigma^{(t)}) \tag{8.11}$$

$$\leqslant -\sum_{t=1}^{\ell} |\Phi(r_t)| I(\sigma) = -I(\sigma).$$

We therefore need to find a lower bound for $-I(\sigma \times \Phi)$.

It is easy to check that

$$\sigma \times \Phi = (\sigma_1 - \sigma_2) \times \Phi = (\sigma_1 \times \Phi) - (\sigma_2 \times \Phi).$$

Write $\nu_1 = \sigma_1 \times \Phi$ and $\nu_2 = \sigma_2 \times \Phi$. Then, corresponding to (8.5), in view of Fubini's theorem, we have

$$-I(\sigma \times \Phi) = -I(\nu_1) - I(\nu_2) + 2J(\nu_1, \nu_2). \tag{8.12}$$

Consider the product measure $\nu_2 = \sigma_2 \times \Phi$ in $\mathbb{R}^3$. Analogous to (8.7), we have

$$\sigma_2 \times \Phi = \sum_{t=1}^{\ell} \Phi(r_t)\sigma_2^{(t)},$$

where, for every $t = 1, \ldots, \ell$, the measure $\sigma_2^{(t)}$ in $\mathbb{R}^3$ is supported by the set $U \times \{r_t\}$, with $\sigma_2^{(t)}(S \times \{r_t\}) = \sigma_2(S)$ for every Borel set S in $\mathbb{R}^2$. Clearly

$$\nu_2(\mathbb{R}^3) = \sigma_2(\mathbb{R}^2) \sum_{t=1}^{\ell} \Phi(r_t).$$

It follows that if

$$\sum_{t=1}^{\ell} \Phi(r_t) = 0, \tag{8.13}$$

then $\nu_2(\mathbb{R}^3) = 0$, and so

$$-I(\nu_2) = \int_{\mathcal{H}_3} |\nu_2(h^+)|^2 \,\mathrm{d}\lambda_3(h) \geqslant 0. \tag{8.14}$$

Recall next that the measure σ_1 in $\mathbb{R}^2$ has support $\mathcal{P}$. Write

$$\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}.$$

Then the product measure $\nu_1 = \sigma_1 \times \Phi$ in $\mathbb{R}^3$ can be described by

$$\sigma_1 \times \Phi = \sum_{i=1}^{N} \sigma_1(\mathbf{p}_i)\Phi^{(i)},$$

where, for every $i = 1, \ldots, N$, the measure $\Phi^{(i)}$ in $\mathbb{R}^3$ is supported by the points $(\mathbf{p}_i, r_1), \ldots, (\mathbf{p}_i, r_\ell)$, with

$$\Phi^{(i)}(\mathbf{p}_i, r_t) = \Phi(r_t)$$

for every $t = 1, \ldots, \ell$.

LEMMA 8.2. *We have*

$$I(\nu_1) = \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} J(\Phi^{(i)}, \Phi^{(j)}) + NI(\Phi).$$

PROOF. Note that the measure ν_1 is supported by the points $(\mathbf{p}_i, r_t)$, where $i = 1, \dots, N$ and $t = 1, \dots, \ell$. Since

$$J(\Phi^{(i)}, \Phi^{(j)}) = \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t) \Phi^{(j)}(\mathbf{p}_j, r_u)$$

and $\sigma_1(\mathbf{p}_i) = \sigma_1(\mathbf{p}_j) = 1$, it follows that

$$\begin{aligned}
I(\nu_1) &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \nu_1(\mathbf{p}_i, r_t) \nu_1(\mathbf{p}_j, r_u) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \sigma_1(\mathbf{p}_i) \Phi^{(i)}(\mathbf{p}_i, r_t) \sigma_1(\mathbf{p}_j) \Phi^{(j)}(\mathbf{p}_j, r_u) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t) \Phi^{(j)}(\mathbf{p}_j, r_u) \right) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} J(\Phi^{(i)}, \Phi^{(j)}) = \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} J(\Phi^{(i)}, \Phi^{(j)}) + \sum_{i=1}^{N} I(\Phi^{(i)}).
\end{aligned}$$

On the other hand, for every $i = 1, \dots, N$, we have

$$\begin{aligned}
I(\Phi^{(i)}) &= \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_i, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t) \Phi^{(i)}(\mathbf{p}_i, r_u) \\
&= \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |r_t - r_u| \Phi(r_t) \Phi(r_u) = I(\Phi).
\end{aligned}$$

The result follows. ◯

At this point, we make the observation that

$$\begin{aligned}
J(\Phi^{(i)}, \Phi^{(j)}) &= \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t) \Phi^{(j)}(\mathbf{p}_j, r_u) \\
&= \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} \left(|\mathbf{p}_i - \mathbf{p}_j|^2 + |r_t - r_u|^2 \right)^{\frac{1}{2}} \Phi(r_t) \Phi(r_u)
\end{aligned}$$

depends only on the functional Φ and the euclidean distance $d = |\mathbf{p}_i - \mathbf{p}_j|$. We can therefore consider the function

$$J(\Phi, d) = \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} \left(d^2 + |r_t - r_u|^2 \right)^{\frac{1}{2}} \Phi(r_t) \Phi(r_u),$$

so that for every $i, j = 1, \dots, N$, we have

$$J(\Phi^{(i)}, \Phi^{(j)}) = J(\Phi, |\mathbf{p}_i - \mathbf{p}_j|).$$

We next consider the term $J(\nu_1, \nu_2)$. Note that the product measure $\nu_2 = \sigma_2 \times \Phi$ on $\mathbb{R}^3$ can be described by

$$\sigma_2 \times \Phi = \int_{\mathbb{R}^2} \Phi^{(\mathbf{y})} \, \mathrm{d}\sigma_2(\mathbf{y}),$$

where, for every $\mathbf{y} \in \mathbb{R}^2$, the measure $\Phi^{(\mathbf{y})}$ in $\mathbb{R}^3$ is supported by the points $(\mathbf{y}, r_1), \ldots, (\mathbf{y}, r_\ell)$, with

$$\Phi^{(\mathbf{y})}(\mathbf{y}, r_t) = \Phi(r_t)$$

for every $t = 1, \ldots, \ell$.

Note that $\sigma_1(\mathbf{p}_i) = 1$ for every $i = 1, \ldots, N$. It follows, similar to the proof of Lemma 8.2, that

$$\begin{aligned} J(\nu_1, \nu_2) &= \sum_{i=1}^{N} \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} \int_{\mathbb{R}^2} |(\mathbf{p}_i, r_t) - (\mathbf{y}, r_u)| \sigma_1(\mathbf{p}_i) \Phi(r_t) \Phi(r_u) \, \mathrm{d}\sigma_2(\mathbf{y}) \\ &= \sum_{i=1}^{N} \int_{\mathbb{R}^2} \left(\sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{y}, r_u)| \Phi(r_t) \Phi(r_u) \right) \mathrm{d}\sigma_2(\mathbf{y}) \\ &= \sum_{i=1}^{N} \int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|) \, \mathrm{d}\sigma_2(\mathbf{y}). \end{aligned} \tag{8.15}$$

We would like to ensure that $J(\Phi, d)$ is *small* when d is *large*. Using the series expansion

$$(d^2 + h^2)^{\frac{1}{2}} = d \left(1 + \left(\frac{h}{d} \right)^2 \right)^{\frac{1}{2}} = d + d \sum_{k=1}^{\infty} \binom{\frac{1}{2}}{k} \left(\frac{h}{d} \right)^{2k},$$

we can write

$$J(\Phi, d) = d \left(\sum_{t=1}^{\ell} \Phi(r_t) \right)^2 + \sum_{k=1}^{\infty} \binom{\frac{1}{2}}{k} I^{(2k)}(\Phi) d^{-2k+1},$$

where, for every $k = 1, 2, 3, \ldots$, we have

$$I^{(2k)}(\Phi) = \sum_{t=1}^{\ell} \sum_{u=1}^{\ell} |r_t - r_u|^{2k} \Phi(r_t) \Phi(r_u).$$

In view of (8.13), we have

$$J(\Phi, d) = \sum_{k=1}^{\infty} \binom{\frac{1}{2}}{k} I^{(2k)}(\Phi) d^{-2k+1}. \tag{8.16}$$

Let us summarize the various restrictions on the functional Φ so far. We have assumed that

$$\sum_{t=1}^{\ell} |\Phi(r_t)| = 1 \quad \text{and} \quad \sum_{t=1}^{\ell} \Phi(r_t) = 0.$$

On the other hand, it follows from (8.16) that $J(\Phi, d)$ will be small when d is large if we can ensure that $I^{(2)}(\Phi) = 0$. We note also that if $J(\Phi, d)$ is non-positive, then it follows from Lemma 8.2 that

$$-I(\nu_1) \geqslant -N I(\Phi). \tag{8.17}$$

LEMMA 8.3. *Suppose that*

$$\sum_{t=1}^{\ell} \Phi(r_t) = 0 \quad and \quad \sum_{t=1}^{\ell} r_t \Phi(r_t) = 0.$$

Then $I^{(2)}(\Phi) = 0$.

PROOF. Note that

$$\begin{aligned} I^{(2)}(\Phi) &= \sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |r_t - r_u|^2 \Phi(r_t)\Phi(r_u) = \sum_{t=1}^{\ell}\sum_{u=1}^{\ell} (r_t^2 - 2r_t r_u + r_u^2)\Phi(r_t)\Phi(r_u) \\ &= \sum_{t=1}^{\ell} \left(\sum_{u=1}^{\ell} \Phi(r_u)\right) r_t^2 \Phi(r_t) - 2\left(\sum_{t=1}^{\ell} r_t \Phi(r_t)\right)\left(\sum_{u=1}^{\ell} r_u \Phi(r_u)\right) \\ &\quad + \sum_{u=1}^{\ell}\left(\sum_{t=1}^{\ell} \Phi(r_t)\right) r_u^2 \Phi(r_u). \end{aligned}$$

The result follows immediately. ○

We therefore need

$$\sum_{t=1}^{\ell} |\Phi(r_t)| = 1 \quad \text{and} \quad \sum_{t=1}^{\ell} \Phi(r_t) = 0 \quad \text{and} \quad \sum_{t=1}^{\ell} r_t \Phi(r_t) = 0. \tag{8.18}$$

Then (8.14) holds. It follows from (8.11) and (8.12) that if (8.17) holds, then we need a bound of the form

$$-I(\Phi) \geqslant c_1 N^{-\frac{1}{2}}, \tag{8.19}$$

as well as a bound of the form

$$J(\nu_1, \nu_2) \geqslant -c_2 N^{\frac{1}{2}}, \tag{8.20}$$

where c_1 and c_2 are positive constants satisfying $c_1 > 2c_2$.

The conditions (8.18) require that the measure Φ in $\mathbb{R}$ is supported by at least three points. The measure $\widetilde{\Phi}$ in $\mathbb{R}$, defined by $\ell = 3$ and with support $\{0, \pm N^{-\frac{1}{2}}\}$, such that

$$\widetilde{\Phi}(0) = \frac{1}{2} \quad \text{and} \quad \widetilde{\Phi}(\pm N^{-\frac{1}{2}}) = -\frac{1}{4},$$

will satisfy (8.18) and give (8.19) for some constant $c_1 > 0$. Furthermore, it can be shown that $J(\widetilde{\Phi}, d) \leqslant 0$ for every real number $d \geqslant 0$, so that (8.17) holds. While we can also establish (8.20) for some constant $c_2 > 0$, it is not clear whether $c_1 > 2c_2$. We therefore consider instead a measure Φ in $\mathbb{R}$, defined by $\ell = 3$ and with support $\{0, \pm\alpha N^{-\frac{1}{2}}\}$, such that

$$\Phi(0) = \frac{1}{2} \quad \text{and} \quad \Phi(\pm\alpha N^{-\frac{1}{2}}) = -\frac{1}{4},$$

where α is a positive real number. Clearly the conditions (8.18) are satisfied. We shall determine a suitable value for α later.

It is easy to check that we have

$$I(\Phi) = -\frac{\alpha}{4} N^{-\frac{1}{2}}, \tag{8.21}$$

and that for every $k = 1, 2, 3, \ldots$, we have

$$I^{(2k)}(\Phi) = \frac{1}{8}\alpha^{2k}(4^k - 4)N^{-k}. \tag{8.22}$$

LEMMA 8.4. *Suppose that $d \geqslant 4\alpha N^{-\frac{1}{2}}$. Then*

$$|J(\Phi, d)| \leqslant \frac{3}{16}\alpha^4 N^{-2} d^{-3}.$$

PROOF. It is easy to check that if $d \geqslant 4\alpha N^{-\frac{1}{2}}$, then the series (8.16) for $J(\Phi, d)$ is a convergent alternating series, since by (8.22), the quantity

$$I^{(2k)}(\Phi) d^{-2k} = \frac{1}{8}\alpha^{2k}(4^k - 4)N^{-k} d^{-2k}$$

is positive and decreasing in k, and the binomial coefficient $\binom{\frac{1}{2}}{k}$ is decreasing in magnitude and alternating in sign. Furthermore, it also follows from (8.22) that $I^{(2)}(\Phi) = 0$, and so

$$|J(\Phi, d)| \leqslant \left| \binom{\frac{1}{2}}{k} I^{(4)}(\Phi) d^{-3} \right| = \frac{3}{16}\alpha^4 N^{-2} d^{-3},$$

as required. ◯

LEMMA 8.5. *The function $-J(\Phi, d)$ is positive and decreasing for $d \geqslant 0$, with*

$$-J(\Phi, 0) = \frac{\alpha}{4} N^{-\frac{1}{2}}.$$

PROOF. It is easy to check that

$$16 J(\Phi, d) = 6d - 8(d^2 + \alpha^2 N^{-1})^{\frac{1}{2}} + 2(d^2 + 4\alpha^2 N^{-1})^{\frac{1}{2}}.$$

Elementary calculus gives

$$\lim_{d \to +\infty} J(\Phi, d) = 0,$$

as well as $J'(\Phi, d) > 0$ for $d > 0$. The first assertion follows. The second assertion is trivial. ◯

To study the term $J(\nu_1, \nu_2)$ and obtain a bound of the type (8.20), we refer to (8.15) and study the integral

$$-\int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\sigma_2(\mathbf{y}).$$

For every $i = 1, \ldots, N$, we know from Lemma 8.4 that $-J(\Phi, |\mathbf{p}_i - \mathbf{y}|) \geqslant 0$ for every $\mathbf{y} \in \mathbb{R}^2$. Hence

$$\begin{aligned} -\int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\sigma_2(\mathbf{y}) &\leqslant -N \int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\mu(\mathbf{y}) \\ &= -2\pi N \int_0^\infty J(\Phi, r) r\, \mathrm{d}r. \end{aligned}$$

By Lemma 8.5, we have

$$-\int_0^{4\alpha N^{-\frac{1}{2}}} J(\Phi, r) r\, \mathrm{d}r \leqslant \frac{\alpha}{4} N^{-\frac{1}{2}} \int_0^{4\alpha N^{-\frac{1}{2}}} r\, \mathrm{d}r = 2\alpha^3 N^{-\frac{3}{2}}.$$

By Lemma 8.4, we have

$$-\int_{4\alpha N^{-\frac{1}{2}}}^{\infty} J(\Phi,r)r\,\mathrm{d}r \leqslant \int_{4\alpha N^{-\frac{1}{2}}}^{\infty} \frac{3}{16}\alpha^4 N^{-2} r^{-2}\,\mathrm{d}r = \frac{3}{64}\alpha^3 N^{-\frac{3}{2}}.$$

It follows that

$$-\int_{\mathbb{R}^2} J(\Phi,|\mathbf{p}_i-\mathbf{y}|)\,\mathrm{d}\sigma_2(\mathbf{y}) \leqslant \frac{131}{32}\pi\alpha^3 N^{-\frac{1}{2}}.$$

Combining this with (8.15) gives

$$J(\nu_1,\nu_2) \geqslant -\frac{131}{32}\pi\alpha^3 N^{\frac{1}{2}}. \tag{8.23}$$

Combining (8.11), (8.12), (8.14), (8.17) and (8.21)–(8.23), we conclude that

$$|I(\sigma)| \geqslant \frac{\alpha}{4}N^{\frac{1}{2}} - \frac{131}{16}\pi\alpha^3 N^{\frac{1}{2}}.$$

Choosing $\alpha = \frac{1}{16}$ gives

$$|I(\sigma)| \geqslant \frac{1}{128}N^{\frac{1}{2}}$$

and completes the proof.

8.2. The Davenport–Roth Method Revisited

Let U be a closed convex set in $\mathbb{R}^2$ of unit area, and with centre of gravity at the origin. For every non-negative real number $r \in \mathbb{R}$ and every angle $\theta \in [0,2\pi]$, let $H(r,\theta)$ denote the closed halfplane

$$H(r,\theta) = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}\cdot\mathbf{e}(\theta) \geqslant r\},$$

where $\mathbf{e}(\theta) = (\cos\theta,\sin\theta)$ and $\mathbf{x}\cdot\mathbf{y}$ denotes the scalar product of $\mathbf{x}$ and $\mathbf{y}$, and write

$$S(r,\theta) = H(r,\theta)\cap U \quad \text{and} \quad R(\theta) = \sup\{r \geqslant 0 : S(r,\theta) \neq \emptyset\}.$$

The following result is more general than Theorem 3.11.

Theorem 8.6 (Beck and Chen 1993). *For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in U such that*

$$\int_0^{2\pi}\int_0^{R(\theta)} |D[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\,\mathrm{d}\theta \ll_U (\log N)^2.$$

Note that Theorem 3.11 is the special case $U = U_0$, the closed disc of unit area in $\mathbb{R}^2$, centred at the origin.

The proof of Theorem 8.6 is in fact motivated by another special case where U is the square $[-\frac{1}{2},\frac{1}{2}]^2$. To illustrate the main ideas, we shall first show that for every natural number M, there exists a set $\mathcal{P}$ of $N = (2M+1)^2$ points in $[-\frac{1}{2},\frac{1}{2}]^2$ such that

$$\int_0^{2\pi}\int_0^{R(\theta)} |D[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\,\mathrm{d}\theta \ll (\log N)^2.$$

For ease of notation, we consider instead the following renormalized version of the problem. Let V be the square $[-M-\frac{1}{2},M+\frac{1}{2}]^2$. For every finite distribution $\mathcal{Q}$

of points in V and every measurable subset $B \subseteq V$, let $Z[\mathcal{Q}; B]$ denote the number of points of $\mathcal{Q}$ that fall into B, and consider the discrepancy function

$$E[\mathcal{Q}; B] = Z[\mathcal{Q}; B] - \mu(B).$$

We shall show that the set

$$\mathcal{Q} = \{-M, \ldots, 0, \ldots, M\}^2$$

of $N = (2M+1)^2$ integer lattice points in V satisfies

$$\int_0^{2\pi} \int_0^{M(\theta)} |E[\mathcal{Q}; S(r,\theta)]| \, \mathrm{d}r \, \mathrm{d}\theta \ll M(\log M)^2, \tag{8.24}$$

where, for every $\theta \in [0, 2\pi]$, we have $M(\theta) = (2M+1)R(\theta)$.

The line

$$T(r,\theta) = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \cdot \mathbf{e}(\theta) = r\}$$

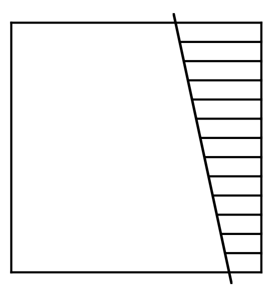
is the boundary of the halfplane $H(r,\theta)$, and can be rewritten in the form

$$x_1 \cos\theta + x_2 \sin\theta = r,$$

where $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$.

Suppose that $0 \leqslant \theta \leqslant \frac{1}{4}\pi$. Clearly $M(\theta) = (M + \frac{1}{2})(\cos\theta + \sin\theta)$. We distinguish two cases.

Case 1: If $0 \leqslant r \leqslant (M + \frac{1}{2})(\cos\theta - \sin\theta)$, then it is not difficult to see that $T(r,\theta)$ intersects the top edge $\{(x_1, M + \frac{1}{2}) : |x_1| \leqslant M + \frac{1}{2}\}$ and the bottom edge $\{(x_1, -M - \frac{1}{2}) : |x_1| \leqslant M + \frac{1}{2}\}$ of V.



Then

$$S(r,\theta) = \bigcup_{n=-M}^{M} S(n, V, r, \theta),$$

where, for every $n = -M, \ldots, 0, \ldots, M$,

$$S(n, V, r, \theta) = S(r,\theta) \cap \left(\mathbb{R} \times \left[n - \frac{1}{2}, n + \frac{1}{2}\right]\right).$$

Clearly

$$E[\mathcal{Q}; S(r,\theta)] = \sum_{n=-M}^{M} E[\mathcal{Q}; S(n, V, r, \theta)].$$

Now, for every $n = -M, \ldots, 0, \ldots, M$, it is easy to check that

$$Z[\mathcal{Q}; S(n, V, r, \theta)] = [M + n\tan\theta - r\sec\theta + 1]$$

and

$$\mu(S(n,V,r,\theta)) = M + n\tan\theta - r\sec\theta + \frac{1}{2},$$

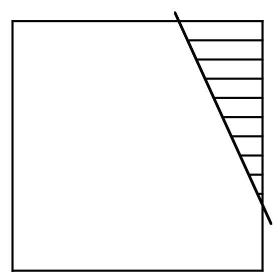
so that

$$E[\mathcal{Q};S(n,V,r,\theta)] = -\psi(n\tan\theta - r\sec\theta),$$

where $\psi(z) = z - [z] - \frac{1}{2}$ for every $z \in \mathbb{R}$. It follows that

$$E[\mathcal{Q};S(r,\theta)] = -\sum_{n=-M}^{M} \psi(n\tan\theta - r\sec\theta).$$

Case 2: If $(M+\frac{1}{2})(\cos\theta - \sin\theta) \leqslant r \leqslant (M+\frac{1}{2})(\cos\theta + \sin\theta)$, then it is not difficult to see that $T(r,\theta)$ intersects the top edge $\{(x_1, M+\frac{1}{2}) : |x_1| \leqslant M+\frac{1}{2}\}$ and the right edge $\{(M+\frac{1}{2}, x_2) : |x_2| \leqslant M+\frac{1}{2}\}$ of V.



In particular, the line $T(r,\theta)$ intersects the right edge of V at the point

$$\left(M+\frac{1}{2}, -\left(M+\frac{1}{2}\right)\cot\theta + r\csc\theta\right),$$

so that $S(n,V,r,\theta) = \emptyset$ for every $n < -(M+\frac{1}{2})\cot\theta + r\csc\theta - \frac{1}{2}$. On the other hand, it is trivial that $E[\mathcal{Q};S(n,V,r,\theta)] = O(1)$ always. It follows that

$$E[\mathcal{Q};S(r,\theta)] = -\sum_{\substack{n=-M \\ (8.25)}}^{M} \psi(n\tan\theta - r\sec\theta) + O(1),$$

where the summation is under the further restriction

$$n \geqslant -\left(M+\frac{1}{2}\right)\cot\theta + r\csc\theta. \tag{8.25}$$

Note that in Case 1, the restriction (8.25) is superfluous since it is weaker than the requirement that $n \geqslant -M$. It follows that for every $r \geqslant 0$, we have

$$E[\mathcal{Q};S(r,\theta)] - G[\mathcal{Q};r,\theta] \ll 1,$$

where

$$G[\mathcal{Q};r,\theta] = -\sum_{\substack{n=-M \\ (8.25)}}^{M} \psi(n\tan\theta - r\sec\theta).$$

Furthermore, it is easy to check that the Fourier expansion of $G[\mathcal{Q};r,\theta]$ is given by

$$\sum_{\nu\neq 0}\frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-M\\(8.25)}}^{M}e(n\nu\tan\theta).$$

However, the restriction (8.25) prevents us from applying Parseval's theorem.

We are in a similar situation to that encountered in Section 7.1. The restriction (8.25) is indeed an analogue of the unfortunate term 1 in the expression (7.4). However, Davenport's idea of using an extra lattice does not appear to help us here, as there is no obvious candidate for such an extra lattice. Unfortunately, Roth's idea of translating the lattice points creates large discrepancy near some of the edges of V far greater than we can comfortably accommodate. We therefore need a new idea.

Recall that every closed halfplane $H(r,\theta)$ is described in terms of the variables r and θ relative to the origin. However, this is not necessary at all, as we can equally well describe such halfplanes in terms of variables relative to any point $\mathbf{y}$ in V. Accordingly, we introduce the following *probabilistic* argument which is somewhat analogous to Roth's idea of translation.

Let $\mathbf{y}=(y_1,y_2)\in[-\frac{1}{2},\frac{1}{2}]^2$. For every $\theta\in[0,\frac{1}{4}\pi]$ and every $r\geqslant 1$, let

$$T(\mathbf{y};r,\theta)=T(r+y_1\cos\theta+y_2\sin\theta,\theta) \tag{8.26}$$

and

$$S(\mathbf{y};r,\theta)=S(r+y_1\cos\theta+y_2\sin\theta,\theta), \tag{8.27}$$

noting here that $r+y_1\cos\theta+y_2\sin\theta\geqslant 0$ always. Then

$$E[\mathcal{Q};S(\mathbf{y};r,\theta)]=E[\mathcal{Q};S(r+y_1\cos\theta+y_2\sin\theta,\theta)].$$

It is not difficult to see that if we write

$$G[\mathcal{Q};\mathbf{y};r,\theta]=-\sum_{\substack{n=-M\\(8.25)}}^{M}\psi(n\tan\theta-(r+y_1\cos\theta+y_2\sin\theta)\sec\theta),$$

then

$$\begin{aligned}&E[\mathcal{Q};S(\mathbf{y};r,\theta)]-G[\mathcal{Q};\mathbf{y};r,\theta]\\&\qquad\ll\begin{cases}\cot\theta, & \text{if } M(\theta)-(2M+1)\sin\theta-1\leqslant r\leqslant M(\theta),\\ 1, & \text{otherwise},\\ M, & \text{trivially},\end{cases}\end{aligned}$$

where the first estimate $\cot\theta$ arises from the fact that we have not modified the extra restriction (8.25). Note also that $|y_1\cos\theta+y_2\sin\theta|\leqslant 1$. It follows that if $r\leqslant M(\theta)-(2M+1)\sin\theta-1$, then $T(\mathbf{y};r,\theta)$ intersects the top and bottom edges of V. Hence

$$\int_0^{\frac{1}{4}\pi}\int_1^{M(\theta)}|E[\mathcal{Q};S(\mathbf{y};r,\theta)]-G[\mathcal{Q};\mathbf{y};r,\theta]|\,\mathrm{d}r\,\mathrm{d}\theta\ll M. \tag{8.28}$$

Now $G[\mathcal{Q};\mathbf{y};r,\theta]$ has the Fourier expansion

$$\sum_{\nu\neq 0}\frac{e(-(r+y_1\cos\theta+y_2\sin\theta)\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-M\\(8.25)}}^{M}e(n\nu\tan\theta)$$
$$=\sum_{\nu\neq 0}\frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-M\\(8.25)}}^{M}e((n-y_2)\nu\tan\theta)e(-y_1\nu).$$

It follows that for every $y_2\in[-\frac{1}{2},\frac{1}{2}]$, we have, by Parseval's theorem, that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}|G[\mathcal{Q};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(8.25)}}^{M}e((n-y_2)\nu\tan\theta)\right|^2$$
$$=\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(8.25)}}^{M}e(n\nu\tan\theta)\right|^2,$$

so that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}}|G[\mathcal{Q};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\,\mathrm{d}y_2\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(8.25)}}^{M}e(n\nu\tan\theta)\right|^2 \tag{8.29}$$
$$\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\min\{M^2,\|\nu\tan\theta\|^{-2}\},$$

where $\|\beta\|=\min_{n\in\mathbb{Z}}|\beta-n|$ for every $\beta\in\mathbb{R}$.

We need the following crucial estimate. The short proof is due to Vaughan.

LEMMA 8.7. *We have*

$$\int_0^{\frac{1}{4}\pi}\left(\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\min\{M^2,\|\nu\tan\theta\|^{-2}\}\right)^{\frac{1}{2}}\mathrm{d}\theta\ll(\log M)^2.$$

PROOF. Since $\tan\theta\asymp\theta$ if $0\leqslant\theta\leqslant\frac{1}{4}\pi$, it suffices to show that

$$\int_0^{1}\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\right)^{\frac{1}{2}}\mathrm{d}\omega\ll(\log M)^2. \tag{8.30}$$

Clearly

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\leqslant\sum_{n=1}^{M^2}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}+1,$$

so that

$$\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\right)^{\frac{1}{2}}\leqslant\sum_{n=1}^{M^2}\frac{1}{n}\min\{M,\|n\omega\|^{-1}\}+1. \tag{8.31}$$

Now, for every $n = 1, \ldots, M^2$, we have

$$\int_0^1 \min\{M, \|n\omega\|^{-1}\}\,\mathrm{d}\omega = 2n\int_0^{1/2n} \min\{M, (n\omega)^{-1}\}\,\mathrm{d}\omega \ll \log M. \tag{8.32}$$

Inequality (8.30) now follows on combining (8.31) and (8.32). ◯

By the Cauchy-Schwarz inequality, we have

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} |G[\mathcal{Q};\mathbf{y};r,\theta]|\,\mathrm{d}y_1\,\mathrm{d}y_2 \ll \left(\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} |G[\mathcal{Q};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\,\mathrm{d}y_2\right)^{\frac{1}{2}}. \tag{8.33}$$

It follows from (8.28), (8.29), (8.33) and Lemma 8.7 that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_0^{\frac{1}{4}\pi}\int_1^{M(\theta)} |E[\mathcal{Q};S(\mathbf{y};r,\theta)]|\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}y_1\,\mathrm{d}y_2 \ll M(\log M)^2. \tag{8.34}$$

For every $\theta \in [0, \frac{1}{4}\pi]$, every $r \geqslant 1$ and every $\mathbf{y} \in [-\frac{1}{2}, \frac{1}{2}]^2$, let

$$s = r + y_1\cos\theta + y_2\sin\theta.$$

Then it is easy to see that $|r - s| < 1$. Since $S(\mathbf{y};r,\theta) = S(r + y_1\cos\theta + y_2\sin\theta, \theta)$, where $r + y_1\cos\theta + y_2\sin\theta \geqslant 0$, we must have

$$\int_2^{M(\theta)-1} |E[\mathcal{Q};S(r,\theta)]|\,\mathrm{d}r \leqslant \int_1^{M(\theta)} |E[\mathcal{Q};S(\mathbf{y};r,\theta)]|\,\mathrm{d}r. \tag{8.35}$$

On the other hand, we have the trivial estimate

$$\left(\int_0^2 + \int_{M(\theta)-1}^{M(\theta)}\right) |E[\mathcal{Q};S(r,\theta)]|\,\mathrm{d}r \ll M. \tag{8.36}$$

It now follows from (8.34)–(8.36) that

$$\int_0^{\frac{1}{4}\pi}\int_0^{M(\theta)} |E[\mathcal{Q};S(r,\theta)]|\,\mathrm{d}r\,\mathrm{d}\theta \ll M(\log M)^2.$$

The inequality (8.24) then follows from symmetry.

Remark. It is clear that our argument is probabilistic in nature. However, we manage at the end not to have to pay a price for using the probabilistic variable $\mathbf{y}$. This is a rare instance in the subject of irregularities of point distribution where we have used a probabilistic argument and still finish with an explicit point set. The reason for this is obvious – the probabilistic variable $\mathbf{y}$ does not modify the point set in question.

Next, we consider the case when $U = U_0$, the closed disc of unit area and centred at the origin.

Let N be any given natural number. Again we consider a renormalized version of the problem, and take V to be the closed disc of area N and centred at the origin. However, if we simply attempt to take all the integer lattice points in V as our set $\mathcal{Q}$, then by a famous theorem of Hardy on the number of lattice points in a disc, the number of points of $\mathcal{Q}$ can differ from N by an amount sufficiently large to make our task impossible.

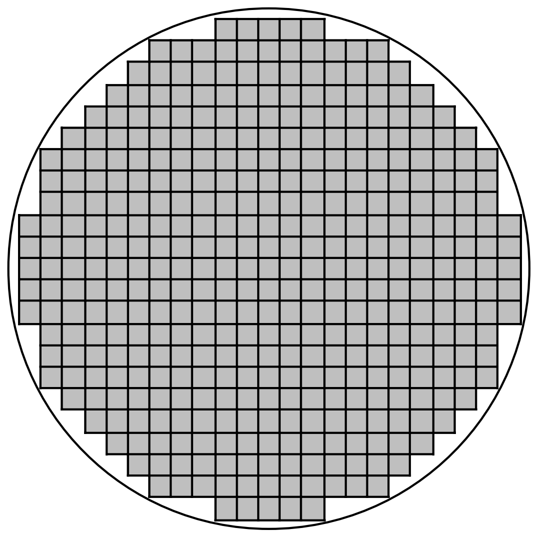
Our new idea is to introduce a set $\mathcal{Q}$ such that the majority of points of $\mathcal{Q}$ are integer lattice points in V, and that the remaining points give rise to a *one-dimensional discrepancy* along and near the boundary of V. More precisely, for any $\mathbf{x} = (x_1, x_2) \in \mathbb{Z}^2$, let

$$A(\mathbf{x}) = A(x_1, x_2) = \left[x_1 - \frac{1}{2}, x_1 + \frac{1}{2}\right] \times \left[x_2 - \frac{1}{2}, x_2 + \frac{1}{2}\right];$$

in other words, $A(\mathbf{x})$ is the aligned closed square of unit area and centred at $\mathbf{x}$. Let

$$\mathcal{Q}_1 = \{\mathbf{q} \in \mathbb{Z}^2 : A(\mathbf{q}) \subseteq V\} \quad \text{and} \quad V_1 = \bigcup_{\mathbf{q} \in \mathcal{Q}_1} A(\mathbf{q}).$$

The set V_1 is represented by the shaded part in the picture below.



Note that the points of $\mathcal{Q}_1$ form the majority of any point set $\mathcal{Q}$ of N points in V. For the remaining points, let $V_2 = V \setminus V_1$. Then it is easy to see, writing $\pi M^2 = N$, that $\mu(V_2) \in \mathbb{N}$ and $\mu(V_2) \ll M$. We partition V_2 as follows. Write $L = \mu(V_2)$, and let $0 = \theta_0 < \theta_1 < \ldots < \theta_{L-1} < \theta_L = 1$ be such that for every $j = 1, \ldots, L$, the set $R_j = \{\mathbf{x} \in V_2 : 2\pi\theta_{j-1} \leqslant \arg \mathbf{x} < 2\pi\theta_j\}$ satisfies $\mu(R_j) = 1$. For every $j = 1, \ldots, L$, let $\mathbf{q}_j \in R_j$, and write $\mathcal{Q}_2 = \{\mathbf{q}_1, \ldots, \mathbf{q}_L\}$. If we now take

$$\mathcal{Q} = \mathcal{Q}_1 \cup \mathcal{Q}_2, \tag{8.37}$$

then clearly $\mathcal{Q}$ contains exactly N points.

For every measurable subset $B \subseteq V$, let $Z[\mathcal{Q}; B]$ denote the number of points of $\mathcal{Q}$ that fall into B, and consider the discrepancy function

$$E[\mathcal{Q}; B] = Z[\mathcal{Q}; B] - \mu(B).$$

For any disc segment $S(r, \theta)$, the analysis of the discrepancy function

$$E[\mathcal{Q}; S(r,\theta) \cap V_1] = E[\mathcal{Q}_1; S(r,\theta) \cap V_1]$$

is essentially similar to our earlier discussion, while the analysis of the discrepancy function

$$E[\mathcal{Q}; S(r,\theta) \cap V_2] = E[\mathcal{Q}_2; S(r,\theta) \cap V_2]$$

gives rise to an error term of smaller order of magnitude. Detailed calculations, using explicitly the equation of ∂V, the boundary of V, will show that the set (8.37) satisfies the inequality

$$\int_0^{2\pi} \int_0^M |E[\mathcal{Q}; S(r,\theta)]| \, \mathrm{d}r \, \mathrm{d}\theta \ll M(\log M)^2.$$

However, if we want to establish the full generality of Theorem 8.6, then we have no explicit information on the boundary of V. Extra geometric consideration is then required.

CHAPTER 9

Convex Polygons

9.1. Similar Copies of a Convex Polygon

Let us return to the unit torus $\mathbb{T}^2$.

Suppose now that B is a closed convex polygon in $\mathbb{T}^2$. For every real number $\lambda \in [0,1]$, every rotation $\theta \in [0,2\pi]$ and every translation $\mathbf{x} \in \mathbb{T}^2$, we can consider similar copies of B given by

$$B(\lambda,\theta,\mathbf{x}) = \{\theta(\lambda\mathbf{y}) + \mathbf{x} : \mathbf{y} \in B\}. \tag{9.1}$$

We now briefly indicate how our technique in Section 8.2 can be adapted to establish the following result.

Theorem 3.12 (Beck and Chen 1993). *Let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2}\int_0^{2\pi}\int_0^1 |D[\mathcal{P};B(\lambda,\theta,\mathbf{x})]| \,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\mathbf{x} \ll_B (\log N)^2.$$

Our study of Theorem 3.12 is motivated by our study of Theorem 3.11, and is based on the simple observation that a convex polygon is the intersection of a finite number of halfplanes. Indeed, a more striking way of putting this is to say that a half plane is a convex monogon!

We shall only briefly discuss the problem in the special case when $N = M^2$ is a perfect square. As before, it is convenient to consider a renormalized version of the problem. Let V be the square $[0,M]^2$, treated as a torus modulo M for each coordinate.

Suppose that B is a closed convex polygon in V, treated as a torus. For every real number $\lambda \in [0,1]$, every rotation $\theta \in [0,2\pi]$ and every translation $\mathbf{x} \in V$, we can again define similar copies of B by (9.1). For every finite distribution $\mathcal{Q}$ of points in V, we consider the corresponding discrepancy function

$$E[\mathcal{Q};B(\lambda,\theta,\mathbf{x})] = Z[\mathcal{Q};B(\lambda,\theta,\mathbf{x})] - \mu(B(\lambda,\theta,\mathbf{x})).$$

To establish Theorem 3.12 in our special case, it clearly suffices to show that for every natural number $M \geqslant 2$, the set

$$\mathcal{Q} = \left\{\left(m-\frac{1}{2}, n-\frac{1}{2}\right) : m,n \in \mathbb{N} \text{ and } 1 \leqslant m,n \leqslant M\right\} \tag{9.2}$$

of $N = M^2$ points in V satisfies the inequality

$$\int_V\int_0^{2\pi}\int_0^1 |E[\mathcal{Q};B(\lambda,\theta,\mathbf{x})]| \,\mathrm{d}\lambda\,\mathrm{d}\theta\,\mathrm{d}\mathbf{x} \ll_B N(\log N)^2. \tag{9.3}$$

The idea is roughly as follows. Consider a fixed similar copy $B(\lambda, \theta, \mathbf{x})$ of the convex polygon B. Then each edge of $B(\lambda, \theta, \mathbf{x})$ gives rise to a discrepancy of a similar nature to the discrepancy arising from the edge of the halfplane $S(r, \theta)$ in our discussion in Section 8.2, and can be handled in a similar manner. The only difference is that there are a few such edges rather than just one. This difference poses no real difficulty, since discrepancy is additive in a certain sense. The only difficulty is to find a suitable analogue of the probabilistic variable $\mathbf{y}$. However, we observe that the translation variable $\mathbf{x}$, handled with great care, plays this role. Indeed, the key idea in the proof of (9.3) is to split the integral over V into a sum of integrals over squares of unit area centred at the points of $\mathcal{Q}$. This will enable us to use the translation variable $\mathbf{x}$ in essentially the same way as the probabilistic variable $\mathbf{y}$ in our earlier discussion. It can then be shown that the set (9.2) satisfies the inequality (9.3).

9.2. Homothetic Copies of a Convex Polygon

Suppose again that B is a closed convex polygon in $\mathbb{T}^2$. For every real number $\lambda \in [0,1]$ and every translation $\mathbf{x} \in \mathbb{T}^2$, consider homothetic copies of B given by

$$B(\lambda, \mathbf{x}) = \{\lambda \mathbf{y} + \mathbf{x} : \mathbf{y} \in B\},$$

and denote by $\mathcal{A}^*(B)$ the collection of all homothetic copies of B obtained this way.

We shall briefly indicate the central idea behind the proof of the following results.

THEOREM 9.1 (Chen and Travaglini 2007). *Let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\sup_{A \in \mathcal{A}^*(B)} |D[\mathcal{P}; A]| \ll_B \log N.$$

THEOREM 9.2 (Beck and Chen 1997). *Let B denote a closed convex polygon in the unit torus $\mathbb{T}^2$. For every natural number $N \geqslant 2$, there exists a distribution $\mathcal{P}$ of N points in $\mathbb{T}^2$ such that*

$$\int_{\mathbb{T}^2} \int_0^1 |D[\mathcal{P}; B(\lambda, \mathbf{x})]|^2 \, \mathrm{d}\lambda \, \mathrm{d}\mathbf{x} \ll_B \log N.$$

It is easy to see that in the proof of Theorem 3.12 in the last section, the point set $\mathcal{P}$ is made up of a square lattice that is suitably scaled. It is clear that the resulting discrepancy function $D[\mathcal{P}; B(\lambda, \theta, \mathbf{x})]$ can be rather large in magnitude for some values of θ and rather small in magnitude for other values of θ, and our result follows since certain averages of the discrepancy function over θ is small. This observation leads us to consider, in the present case, the possibility of rotating a square lattice to a suitable angle, and then perhaps making some appropriate adjustments near the *edge* of the torus.

The square lattice $\Lambda = (N^{-\frac{1}{2}}\mathbb{Z})^2$ has roughly N points per unit area. Rotating this to a suitable angle presents no difficulties, and we appeal to the following result of Davenport on diophantine approximation.

LEMMA 9.3 (Davenport 1964). *Suppose that $f_1, \ldots, f_r$ are real-valued functions of a real variable, and have continuous first derivatives in some open interval I containing θ_0, where $f_1'(\theta_0), \ldots, f_r'(\theta_0)$ are all non-zero. Then there exists $\theta \in I$ such that $f_1(\theta), \ldots, f_r(\theta)$ are all badly approximable.*

Without loss of generality, assume that the convex polygon B has centre of gravity at the origin $\mathbf{0}$. Suppose further that B has k sides, with vertices $\mathbf{v}_1, \ldots, \mathbf{v}_k$, where

$$(\mathbf{v}_j - \mathbf{v}_{j-1}) \cdot \mathbf{e}\left(\theta_j + \frac{\pi}{2}\right) = |\mathbf{v}_j - \mathbf{v}_{j-1}|,$$

with $0 \leqslant \theta_1 < \ldots < \theta_k < 2\pi$ and $\mathbf{v}_0 = \mathbf{v}_k$. Here $\mathbf{e}(\theta) = (\cos\theta, \sin\theta)$ and $\mathbf{u} \cdot \mathbf{v}$ denotes the scalar product of $\mathbf{u}$ and $\mathbf{v}$. Let T_j denote the side of B with vertices $\mathbf{v}_{j-1}$ and $\mathbf{v}_j$, and note that the perpendicular from $\mathbf{0}$ to T_j makes an angle θ_j with the positive x_1-axis.

An immediate conseqeunce of Lemma 9.3 is that there exists a real number $\theta \in [0, 2\pi]$ such that the $k+2$ numbers

$$\tan\theta, \tan\left(\theta + \frac{\pi}{2}\right), \tan(\theta + \theta_1), \ldots, \tan(\theta + \theta_k)$$

are all finite and badly approximable. We now choose one such value of θ and keep it fixed. We then rotate the square lattice $\Lambda = (N^{-\frac{1}{2}}\mathbb{Z})^2$ anticlockwise by angle θ to obtain the lattice Λ_θ, and let $\mathcal{P}_0 = \Lambda_\theta \cap [0,1)^2$. Note that while the set $\mathcal{P}_0$ has roughly N points, it does not necessarily have precisely N points. This turns out not to be an issue, and arbitrarily adding or deleting a suitable number of points gives rise to a set $\mathcal{P}$ of precisely N points that satisfies the conclusion of Theorem 9.1.

However, the analysis of the adjusted point set appears to give rise to an error term too large for the method to lead to a proof of Theorem 9.2. To overcome this difficulty, we appeal to Roth's probabilistic method first discussed in Section 7.2, introduce an extra translation variable and consider some average of the discrepancy function over a suitable collection of translated copies of our basic construction.

More precisely, for every $\mathbf{w} \in \mathbb{R}^2$, write

$$\mathbf{w} + \Lambda_\theta = \{\mathbf{w} + \mathbf{v} : \mathbf{v} \in \Lambda_\theta\}.$$

In other words, the lattice $\mathbf{w} + \Lambda_\theta$ is obtained from the lattice Λ by first rotating anticlockwise by angle θ and then translating by $\mathbf{w}$. Note that $\mathbf{w} + \Lambda_\theta$ is a square lattice with determinant N^{-1}. We then study the discrepancy of the set

$$\mathcal{P}_0(\mathbf{w}) = (\mathbf{w} + \Lambda_\theta) \cap [0,1)^2 \tag{9.4}$$

in $[0,1)^2$, and show that there exists $\mathbf{w}^* \in (N^{-\frac{1}{2}}\mathbb{Z})^2$ such that

$$\int_{\mathbb{T}^2} \int_0^1 |D[\mathcal{P}_0(\mathbf{w}^*); B(\lambda, \mathbf{x})]|^2 \, \mathrm{d}\lambda \, \mathrm{d}\mathbf{x} \ll_B N \log N. \tag{9.5}$$

As before, the set (9.4) may not have precisely N points. However, it can be shown that arbitrarily adding or deleting a suitable number of points gives rise to a modification of (9.4) which does not jeopardize the estimate (9.4).

Remark. The method here can de adapted to give a proof of Theorems 2.10 and 2.11, and this explains the first footnote in Chapter 7.

9.3. Some Further Remarks

Throughout this section, $D[\mathcal{P}; A]$ denotes the discrepancy of a set $\mathcal{P}$ of N points in the unit torus $\mathbb{T}^2$ with respect to a measurable subset $A \subseteq \mathbb{T}^2$. We first describe the behaviour of the function

$$D(\mathcal{A}, N) = \inf_{|\mathcal{P}|=N} \sup_{A \in \mathcal{A}} |D[\mathcal{P}; A]|$$

with respect to three classes $\mathcal{A}$ of convex polygons in $\mathbb{T}^2$. Here the infimum is taken over all distributions $\mathcal{P}$ of N points in $\mathbb{T}^2$.

Let $\Theta = (\theta_1, \ldots, \theta_k)$, where $\theta_1, \ldots, \theta_k \in [0, \pi)$ are fixed, and denote by $\mathcal{A}(\Theta)$ the collection of all convex polygons A in $\mathbb{T}^2$ such that every side of A makes an angle θ_i for some $i = 1, \ldots, k$ with the positive horizontal axis. The proof of Theorem 9.1 can be adapted to show that for every integer $N \geqslant 2$, we have

$$D(\mathcal{A}(\Theta), N) \ll_\Theta \log N.$$

On the other hand, the Roth–Halász technique described in Section 4.3 has been adapted by Beck and Chen to show that for every integer N, we have

$$D(\mathcal{A}(\Theta), N) \gg_\Theta \log N.$$

Hence the problem is well understood for $\mathcal{A}(\Theta)$.

Next, we relax the restriction on the direction of the sides of the convex polygons, replace this with a restriction on the number of sides instead, and denote by $\mathcal{A}_k$ the collection of all convex polygons in $\mathbb{T}^2$ with at most k sides. Then a result of Beck implies that for every integer N, we have

$$D(\mathcal{A}_k, N) \gg_k N^{\frac{1}{4}}.$$

On the other hand, the large deviation technique in Section 5.2 has been adapted by Chen and Travaglini to show that for every integer $N \geqslant 2$, we have

$$D(\mathcal{A}_k, N) \ll_k N^{\frac{1}{4}} (\log N)^{\frac{1}{2}}.$$

There remains a small gap between the lower bound and the upper bound.

Finally, we relax all the restrictions on the direction and number of sides of the convex polygons, and denote by $\mathcal{A}^*$ the collection of all convex polygons in $\mathbb{T}^2$. Then the elegant argument of Schmidt in Section 4.1 has been extended by Chen and Travaglini to show that for every every integer N, we have

$$D(\mathcal{A}^*, N) \gg N^{\frac{1}{3}}.$$

On the other hand, Theorem 3.5 implies that for every integer $N \geqslant 2$, we have

$$D(\mathcal{A}^*, N) \ll N^{\frac{1}{3}} (\log N)^4.$$

Again, there remains a small gap between the lower bound and the upper bound.

We conclude this chapter by mentioning some recent work motivated by the above observations.

Let Ω be a set of directions in $\mathbb{R}^2$, and let $\mathcal{A}_{k,\Omega}$ denote the collection of all convex polygons in $\mathbb{T}^2$ with at most k sides with normals belonging to $\pm\Omega$.

THEOREM 9.4 (Bilyk, Ma, Pipher and Spencer 2011). *Let $k \geqslant 3$ be a fixed integer.*

(i) *Suppose that Ω is a lacunary sequence. Then*

$$D(\mathcal{A}_{k,\Omega}, N) \ll_{k,\Omega} (\log N)^3.$$

(ii) *Suppose that Ω is a finite union of lacunary sets of order at most M. Then*

$$D(\mathcal{A}_{k,\Omega}, N) \ll_{k,\Omega} (\log N)^{2M+1}.$$

(iii) *Suppose that Ω has upper Minkowski dimension $d \in (0, 1)$. Then*

$$D(\mathcal{A}_{k,\Omega}, N) \ll_{k,\Omega,\epsilon} N^{\frac{1}{2}\tau(\tau+1)^{-1}+\epsilon}$$

for any $\epsilon > 0$, where $\tau = 2(1-d)^{-2} - 2$.

CHAPTER 10

Fourier–Walsh Analysis

10.1. A Fourier–Walsh Approach to van der Corput Sets

In this section, we sketch yet another proof of Theorem 2.11 by highlighting the interesting group structure of the van der Corput point set

$$\mathcal{P}(2^h) = \{(0.a_1a_2a_3\ldots a_h, 0.a_h\ldots a_3a_2a_1) : a_1,\ldots,a_h \in \{0,1\}\}.$$

This is a finite abelian group isomorphic to the group $\mathbb{Z}_2^h$. 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 $\ell \in \mathbb{N}_0$, written uniquely in the form

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell) 2^{i-1}, \tag{10.1}$$

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

$$y = \sum_{i=1}^{\infty} \eta_i(y) 2^{-i}, \tag{10.2}$$

where the coefficient $\eta_i(y) \in \{0,1\}$ for every $i \in \mathbb{N}$. This representation is unique if we agree that the series in (10.2) is finite for every $y = m2^{-s}$ where $s \in \mathbb{N}_0$ and $m \in \{0,1,\ldots,2^s-1\}$.

For every $\ell \in \mathbb{N}_0$ of the form (10.1), we define the Walsh function $w_\ell : [0,1) \to \mathbb{R}$ by writing

$$w_\ell(y) = (-1)^{\sum\limits_{i=1}^{\infty} \lambda_i(\ell)\eta_i(y)}. \tag{10.3}$$

Since (10.1) 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 $\ell, k \in \mathbb{N}_0$, we can define $\ell \oplus k$ by setting

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

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

$$w_{\ell \oplus k}(y) = w_\ell(y) w_k(y). \tag{10.4}$$

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

$$w_\ell(x \oplus y) = w_\ell(x) w_\ell(y). \tag{10.5}$$

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

$$D[\mathcal{P}(2^h); B(\mathbf{x})] = \sum_{\mathbf{p} \in \mathcal{P}(2^h)} \chi_{B(\mathbf{x})}(\mathbf{p}) - 2^h x_1 x_2. \tag{10.6}$$

Clearly the characteristic function 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[1] to study characteristic functions of the form $\chi_{[0,x)}(y)$. We have the Fourier–Walsh series

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

where, for every $\ell \in \mathbb{N}_0$, 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 $x \in [0,1)$.

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

$$\chi_{[0,x)}^{(h)}(y) = \sum_{\ell=0}^{2^h - 1} \widetilde{\chi}_\ell(x) w_\ell(y). \tag{10.7}$$

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

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

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 $[m2^{-h}, (m+1)2^{-h})$.

---

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

The approximation (10.7) 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 $\mathbf{l} = (\ell_1, \ell_2)$,

$$\widetilde{\chi}_{\mathbf{l}}(\mathbf{x}) = \widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2) \quad \text{and} \quad W_{\mathbf{l}}(\mathbf{y}) = w_{\ell_1}(y_1)w_{\ell_2}(y_2). \tag{10.8}$$

Corresponding to this, we approximate the discrepancy function (10.6) by

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

noting that

$$\sum_{\mathbf{p}\in\mathcal{P}(2^h)} W_{\mathbf{0}}(\mathbf{p}) = \#\mathcal{P}(2^h) = 2^h. \tag{10.9}$$

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

$$\sum_{\mathbf{p}\in\mathcal{P}(2^h)} W_{\mathbf{l}}(\mathbf{p}) \in \{0, 2^h\}. \tag{10.10}$$

We therefore need to have some understanding on the set

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

Then

$$D^{(h)}[\mathcal{P}(2^h); B(\mathbf{x})] = 2^h \sum_{\mathbf{l}\in L(h)} \widetilde{\chi}_{\mathbf{l}}(\mathbf{x}). \tag{10.11}$$

Recall the discussion at the beginning of Section 7.4. The estimate (7.20) shows that the set $\mathcal{P}(2^h)$ is insufficient for us to establish Theorem 2.11. To overcome this problem, we use digit shifts in Section 7.5. Here, for every $\mathbf{t} \in \mathbb{Z}_2^{2h}$, we consider the set

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

where, for every

$$\mathbf{p} = (0.a_1\ldots a_h, 0.a_h\ldots a_1) \in \mathcal{P}(2^h) \quad \text{and} \quad \mathbf{t} = (t'_1,\ldots,t'_h,t''_h,\ldots,t''_1) \in \mathbb{Z}_2^{2h},$$

we have the shifted point[2]

$$\mathbf{p}\oplus\mathbf{t} = (0.b'_1\ldots b'_h, 0.b''_h\ldots b''_1),$$

[2]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.

with the digits $b_1',\ldots,b_h',b_1'',\ldots,b_h''\in\{0,1\}$ 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 $s=1,\ldots,h$. Then

$$\begin{aligned}D^{(h)}[\mathcal{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]&=\sum_{\mathbf{p}\in\mathcal{P}(2^h)}\chi^{(h)}_{B(\mathbf{x})}(\mathbf{p}\oplus\mathbf{t})-2^hx_1x_2\\&=\sum_{\substack{\ell_1=0\\(\ell_1,\ell_2)\neq(0,0)}}^{2^h-1}\sum_{\ell_2=0}^{2^h-1}\left(\sum_{\mathbf{p}\in\mathcal{P}(2^h)}W_{\mathbf{l}}(\mathbf{p}\oplus\mathbf{t})\right)\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})\\&=\sum_{\substack{\ell_1=0\\(\ell_1,\ell_2)\neq(0,0)}}^{2^h-1}\sum_{\ell_2=0}^{2^h-1}W_{\mathbf{l}}(\mathbf{t})\left(\sum_{\mathbf{p}\in\mathcal{P}(2^h)}W_{\mathbf{l}}(\mathbf{p})\right)\widetilde{\chi}_{\mathbf{l}}(\mathbf{x}),\end{aligned}$$

in view of (10.5) and the second identity in (10.8). It follows that

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

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

$$\begin{aligned}\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}|D^{(h)}[\mathcal{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2&=4^h\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}\left(\sum_{\mathbf{l}\in L(h)}W_{\mathbf{l}}(\mathbf{t})\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})\right)^2\\&=4^h\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}\sum_{\mathbf{l}',\mathbf{l}''\in L(h)}W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t})\widetilde{\chi}_{\mathbf{l}'}(\mathbf{x})\widetilde{\chi}_{\mathbf{l}''}(\mathbf{x})\\&=4^h\sum_{\mathbf{l}',\mathbf{l}''\in L(h)}\left(\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t})\right)\widetilde{\chi}_{\mathbf{l}'}(\mathbf{x})\widetilde{\chi}_{\mathbf{l}''}(\mathbf{x}).\end{aligned}\tag{10.12}$$

LEMMA 10.1. *For every $\mathbf{l}',\mathbf{l}''\in\mathbb{N}_0^2$, we have*

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

PROOF. Note first of all that in view of (10.4) and the second identity in (10.8), 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

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

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

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

in view of (10.5) and the second identity in (10.8). Clearly $S=0$ in this case. ○

Combining (10.12) and Lemma 10.1, we deduce that

$$\frac{1}{4^h}\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}|D^{(h)}[\mathcal{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2=4^h\sum_{\mathbf{l}\in L(h)}|\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2, \tag{10.13}$$

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

$$\frac{1}{4^h}\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}\int_{[0,1]^2}|D^{(h)}[\mathcal{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}=4^h\sum_{\mathbf{l}\in L(h)}\int_{[0,1]^2}|\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}. \tag{10.14}$$

To estimate the right hand side of (10.14), we need to use a formula of Fine on the Fourier–Walsh coefficients of the characteristic function $\chi_{[0,x)}(y)$.

Let $\rho(0)=0$. For any integer $\ell\in\mathbb{N}$ with representation (10.1), let

$$\rho(\ell)=\max\{i\in\mathbb{N}:\lambda_i(\ell)\neq 0\},\quad\text{so that}\quad 2^{\rho(\ell)-1}\leqslant\ell<2^{\rho(\ell)}. \tag{10.15}$$

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 (10.8), 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 (10.14) becomes

$$\frac{1}{4^h}\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}}\int_{[0,1]^2}|D^{(h)}[\mathcal{P}(2^h)\oplus\mathbf{t};B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}=\frac{4^h}{9}\sum_{\mathbf{l}\in L(h)}4^{-\rho(\mathbf{l})}. \tag{10.16}$$

To estimate the sum on the right hand side of (10.16), we need some reasonably precise information on the set $L(h)$. The following result is rather useful.

LEMMA 10.2. *For every $y\in[0,1)$ and every $s\in\mathbb{N}_0$, 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 (10.2) that $\eta_i(y)=0$ whenever $1\leqslant i\leqslant s$. On the other hand, for every $\ell=0,1,2,\ldots,2^s-1$, it follows from (10.1) 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 (10.2) 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 (10.4). The result follows immediately. ◯

LEMMA 10.3. *For every* $s_1, s_2 \in \{0, 1, \ldots, h\}$, *let*

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

Then

$$\Xi(s_1, s_2) = \begin{cases} 2^{s_1+s_2}, & \textit{if } s_1 + s_2 \geqslant h, \\ 2^h, & \textit{if } s_1 + s_2 \leqslant h. \end{cases}$$

PROOF. Writing $\mathbf{p} = (p_1, p_2)$ and $\mathbf{l} = (\ell_1, \ell_2)$ and noting the second identity in (10.8) and Lemma 10.2, we have

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

It is not difficult to deduce from Lemma 7.3 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 $m_1, m_2 \in \mathbb{N}_0$, and area 2^{-h}, contains precisely one point of $\mathcal{P}(2^h)$. Let us say that such a rectangle is an elementary rectangle. Suppose first of all that $s_1 + s_2 \geqslant h$. 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 $\mathcal{P}(2^h)$. Clearly it contains the point $\mathbf{0} \in \mathcal{P}(2^h)$, and so

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

Suppose then that $s_1 + s_2 \leqslant h$. 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 $\mathcal{P}(2^h)$, whence

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

This completes the proof. ◯

Note that with $s_1 = s_2 = h$, Lemma 10.3 gives

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

In view of (10.9) and (10.10), we conclude that $\#L(h) = 2^h - 1$. We now study the set $L(h)$ in greater detail.

LEMMA 10.4. *For every* $s_1, s_2 \in \{1, \ldots, h\}$, *let*

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

Then

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

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

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

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 (10.15). This establishes part (i). To prove part (ii), note that in view of (10.10), we have, in the notation of Lemma 10.3,

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

Part (ii) now follows easily from Lemma 10.3. Finally, it is easily checked that

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

The last assertion follows immediately. ○

Using Lemma 10.4, we deduce that

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

Combining this with (10.16), we obtain

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

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

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

essentially establishing Theorem 2.11, 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 (10.11) and make the hypothetical assumption that the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l}\in L(h)$, are orthogonal. Then

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

Note that the right hand side is exactly the same as the right hand side of (10.14), 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 $\mathbf{t}\in\mathbb{Z}_2^{2h}$, and bring into the argument, one may say through the back door, some orthogonality in the form of Lemma 10.1.

10.2. Group Structure and p-adic Fourier–Walsh Analysis

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

Let us again restrict our attention to Theorem 2.11. Let p be a prime, and consider the base p van der Corput point set

$$\mathcal{P}(p^h)=\{(0.a_1a_2a_3\ldots a_h,0.a_h\ldots a_3a_2a_1):a_1,\ldots,a_h\in\{0,1,\ldots,p-1\}\}.$$

This is a finite abelian group isomorphic to the group $\mathbb{Z}_p^h$. 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 $\ell\in\mathbb{N}_0$, written uniquely in the form

$$\ell=\sum_{i=1}^{\infty}\lambda_i(\ell)p^{i-1},\tag{10.17}$$

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

$$y=\sum_{i=1}^{\infty}\eta_i(y)p^{-i},\tag{10.18}$$

where the coefficient $\eta_i(y)\in\{0,1,\ldots,p-1\}$ for every $i\in\mathbb{N}$. This representation is unique if we agree that the series in (10.18) is finite for every $y=mp^{-s}$ where $s\in\mathbb{N}_0$ and $m\in\{0,1,\ldots,p^s-1\}$.

For every $\ell \in \mathbb{N}_0$ of the form (10.17), we define the Walsh function $w_\ell : [0,1) \to \mathbb{R}$ by writing

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

where $e_p(z) = \mathrm{e}^{2\pi \mathrm{i} z/p}$ for every real number z. Since (10.17) 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 (10.4) and (10.5) 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 $\ell \in \mathbb{N}_0$, 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

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

by

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

noting that

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

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

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

We therefore need to have some understanding on the set

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

Then

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

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

LEMMA 10.5. *Suppose that the prime p satisfies $p \geqslant 8$. Then the functions $\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})$, where $\mathbf{l} \in L(h)$, are orthogonal, so that*

$$\int_{[0,1]^2} |D^{(h)}[\mathcal{P}(p^h);B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = p^{2h} \sum_{\mathbf{l}\in L(h)} \int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x}. \tag{10.19}$$

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

$$\int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \left(\int_0^1 |\widetilde{\chi}_{\ell_1}(x_1)|^2 \,\mathrm{d}x_1 \right) \left(\int_0^1 |\widetilde{\chi}_{\ell_2}(x_2)|^2 \,\mathrm{d}x_2 \right) \tag{10.20}$$

on the right hand side of (10.19).

LEMMA 10.6. *We have*

$$\int_0^1 |\widetilde{\chi}_0(x)|^2 \,\mathrm{d}x = \frac{1}{4} + \frac{1}{4(p^2-1)} \sum_{j=1}^{p-1} \csc^2 \frac{\pi j}{p}. \tag{10.21}$$

Furthermore, for every $\ell \in \mathbb{N}$, we have

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2 \,\mathrm{d}x = p^{-2\rho(\ell)} \left(\frac{1}{2} \csc^2 \frac{\pi\lambda(\ell)}{p} - \frac{1}{4} + \frac{1}{4(p^2-1)} \sum_{j=1}^{p-1} \csc^2 \frac{\pi j}{p} \right), \tag{10.22}$$

where

$$\rho(\ell) = \begin{cases} 0, & \text{if } \ell = 0, \\ \max\{i \in \mathbb{N} : \lambda_i(\ell) \neq 0\}, & \text{if } \ell \in \mathbb{N}, \end{cases}$$

denotes the position of the leading coefficient of ℓ given by the representation (10.17) *and $\lambda(\ell) = \lambda_{\rho(\ell)}(\ell)$ denotes its value.*

SKETCH OF PROOF. We have the Fine–Price formula, that for every $\ell \in \mathbb{N}_0$,

$$\widetilde{\chi}_\ell(x) = p^{-\rho(\ell)} u_\ell(x), \tag{10.23}$$

where

$$u_0(x) = \frac{1}{2} w_0(x) + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} w_{jp^{i-1}}(x), \tag{10.24}$$

and where for every $\ell \in \mathbb{N}$,

$$\begin{aligned} u_\ell(x) &= (1-\zeta^{\lambda(\ell)})^{-1} w_{\tau(\ell)}(x) + \left(\frac{1}{2} - (1-\zeta^{\lambda(\ell)})^{-1}\right) w_\ell(x) \\ &\quad + \sum_{i=1}^{\infty} p^{-i} \sum_{j=1}^{p-1} \zeta^j (1-\zeta^j)^{-1} w_{\ell + jp^{\rho(\ell)+i-1}}(x). \end{aligned} \tag{10.25}$$

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. The right hand side of (10.25) is a linear combination of distinct Walsh functions. It follows that for every $\ell \in \mathbb{N}$, we have

$$\begin{aligned} &\int_0^1 |u_\ell(x)|^2 \, \mathrm{d}x \\ &\quad = \frac{1}{(1-\zeta^{\lambda(\ell)})(1-\zeta^{-\lambda(\ell)})} + \left(\frac{1}{2} - \frac{1}{1-\zeta^{\lambda(\ell)}}\right)\left(\frac{1}{2} - \frac{1}{1-\zeta^{-\lambda(\ell)}}\right) \\ &\quad\quad + \sum_{i=1}^{\infty} p^{-2i} \sum_{j=1}^{p-1} |1-\zeta^j|^{-2} \\ &\quad = 2|1-\zeta^{\lambda(\ell)}|^{-2} - \frac{1}{4} + \frac{1}{p^2-1} \sum_{j=1}^{p-1} |1-\zeta^j|^{-2}. \end{aligned} \tag{10.26}$$

The identity (10.22) follows on combining (10.23) and (10.26) with the observation

$$|1-\zeta^j|^2 = \left(1 - \cos\frac{2\pi j}{p}\right)^2 + \sin^2\frac{2\pi j}{p} = 4\sin^2\frac{\pi j}{p}. \tag{10.27}$$

Similarly, we have

$$\int_0^1 |u_0(x)|^2 \, \mathrm{d}x = \frac{1}{4} + \sum_{i=1}^{\infty} p^{-2i} \sum_{j=1}^{p-1} |1-\zeta^j|^{-2} = \frac{1}{4} + \frac{1}{p^2-1} \sum_{j=1}^{p-1} |1-\zeta^j|^{-2}. \tag{10.28}$$

The identity (10.21) follows on combining (10.23), (10.27) and (10.28). ◯

LEMMA 10.7. *For every $\ell \in \mathbb{N}_0$, we have*

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

Sketch of Proof. Suppose first of all that $\ell \neq 0$. Then using the inequality

$$\csc^2 \frac{\pi j}{p} \leqslant \frac{p^2}{4}$$

for every $j = 1, \ldots, p-1$, we see from (10.22) that

$$\int_0^1 |\widetilde{\chi}_\ell(x)|^2 \, \mathrm{d}x \leqslant p^{-2\rho(\ell)} \left(\frac{p^2}{8} + \frac{1}{4} + \frac{p^2(p-1)}{16(p^2-1)} \right) \leqslant \frac{p^{2-2\rho(\ell)}}{4}.$$

On the other hand, it follows similarly from (10.21) that

$$\int_0^1 |\widetilde{\chi}_0(x)|^2 \, \mathrm{d}x \leqslant \frac{1}{4} + \frac{p^2(p-1)}{16(p^2-1)} \leqslant \frac{p^2}{4} = \frac{p^{2-2\rho(0)}}{4}$$

as required. ◯

Combining (10.20) and Lemma 10.7, we conclude that

$$\int_{[0,1]^2} |\widetilde{\chi}_{\mathbf{l}}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \leqslant \frac{p^{4-2\rho(\mathbf{l})}}{16},$$

where $\rho(\mathbf{l}) = \rho(\ell_1) + \rho(\ell_2)$. Thus we need to estimate the sum

$$\sum_{\mathbf{l} \in L(h)} p^{-2\rho(\mathbf{l})}. \tag{10.29}$$

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 $\mathcal{P}(p^h)$ has sufficiently large Rosenblum–Tsfasman weight, then we can obtain a good estimate for the sum (10.29).

The details are beyond the scope of these lectures.

10.3. Explicit Constructions and Orthogonality

The first proof of the analogue of Theorem 2.11 in arbitrary dimensions by Roth is probabilistic in nature, as are the subsequent proofs by Chen and Skriganov. 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 of Theorem 2.11 is not probabilistic in nature, and one can describe the point set explicitly. However, finding explicit constructions in dimensions $K \geqslant 3$ turns out to be rather hard. Its eventual solution by Chen and Skriganov 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 10.5 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 $K \geqslant 3$. 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 $K \geqslant 2$ for some suitably defined Walsh function $W_{\mathbf{l}}(\mathbf{t})$.

Lemma 10.8. *For every $\mathbf{l}', \mathbf{l}'' \in \mathbb{N}_0^K$, we have*

$$\sum_{\mathbf{t} \in \mathbb{Z}_p^{Kh}} W_{\mathbf{l}'}(\mathbf{t}) W_{\mathbf{l}''}(\mathbf{t}) = \begin{cases} p^{Kh}, & \textit{if } \mathbf{l}' = \mathbf{l}'', \\ 0, & \textit{otherwise.} \end{cases}$$

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.