LECTURES ON IRREGULARITIES OF POINT DISTRIBUTION

W W L CHEN

Contents

1. The Classical Problems

Suppose that $x_1, x_2, x_3, \ldots$ is an infinite sequence of real numbers in the interval $[0,1)$. For every natural number $n \in \mathbb{N}$ and every pair of real numbers $\alpha, \beta \in \mathbb{R}$ satisfying $0 \le \alpha < \beta \le 1$, let

$$Z[n,\alpha,\beta] = \#\{1 \le i \le n : x_i \in [\alpha,\beta)\}$$

denote the number of terms among the finite sequence $x_1, \ldots, x_n$ that fall into the interval $[\alpha,\beta)$. The infinite sequence $x_1, x_2, x_3, \ldots$ is said to be uniformly distributed in the interval $[0,1)$ if

$$\lim_{n\to\infty} \frac{Z[n,\alpha,\beta]}{n} = \beta - \alpha \qquad \text{for every } \alpha, \beta \in \mathbb{R} \text{ satisfying } 0 \le \alpha < \beta \le 1.$$

Note that for every $\alpha, \beta \in \mathbb{R}$ satisfying $0 \le \alpha < \beta \le 1$, the interval $[\alpha,\beta)$ can be described as the set difference $[0,\beta) \setminus [0,\alpha)$. It follows that the above setting can be simplified somewhat. For every natural number $n \in \mathbb{N}$ and every real number $\alpha \in \mathbb{R}$ satisfying $0 \le \alpha \le 1$, let

$$Z[n,\alpha] = \#\{1 \le i \le n : x_i \in [0,\alpha)\}$$

denote the number of terms among the finite sequence $x_1, \ldots, x_n$ that fall into the interval $[0,\alpha)$. Then it is easy to see that the infinite sequence $x_1, x_2, x_3, \ldots$ is uniformly distributed in the interval $[0,1)$ if

$$\lim_{n\to\infty} \frac{Z[n,\alpha]}{n} = \alpha \qquad \text{for every } \alpha \in \mathbb{R} \text{ satisfying } 0 \le \alpha \le 1.$$

† Lectures given at Macquarie University in 2000.

There are various criteria to characterize uniform distribution in the interval $[0,1)$. We mention here the famous Weyl criterion which reduces the problem to one of studying sums of exponential functions. Using this, it is easy to show that for every irrational number $\theta \in \mathbb{R}$, the infinite sequence $\{\theta\}, \{2\theta\}, \{3\theta\}, \ldots$ of fractional parts is uniformly distributed in the interval $[0,1)$.

For every natural number $n \in \mathbb{N}$ and every real number $\alpha \in \mathbb{R}$ satisfying $0 \le \alpha \le 1$, we now study the discrepancy

$$D[n,\alpha] = Z[n,\alpha] - n\alpha,$$

noting that $n\alpha$ represents the expected number of terms among the finite sequence $x_1, \ldots, x_n$ that fall into the interval $[0,\alpha)$. It is easy to show that the infinite sequence $x_1, x_2, x_3, \ldots$ is uniformly distributed in the interval $[0,1)$ if

$$D[n,\alpha] = o(n) \qquad \text{for every } \alpha \in \mathbb{R} \text{ satisfying } 0 \le \alpha \le 1.$$

This is a rather weak statement of a qualitative nature.

In fact, it can be shown that in the case when $\theta = \sqrt{2}$, the infinite sequence $\{\sqrt{2}\}, \{2\sqrt{2}\}, \{3\sqrt{2}\}, \ldots$ of fractional parts satisfies the bound

$$|D[n,\alpha]| \le C \log n \qquad \text{for every } n \in \mathbb{N} \text{ satisfying } n \ge 2 \text{ and every } \alpha \in \mathbb{R} \text{ satisfying } 0 \le \alpha \le 1,$$

where $C \in \mathbb{R}$ is a positive absolute constant. It can also be shown that the famous van der Corput sequence, which we shall define later, satisfies a similar bound.

These two examples raise the question whether such bounds of logarithmic order are best possible. More precisely, we can pose the following question.

Question. Does there exist a function $f(n) \to +\infty$ as $n \to \infty$ such that for every infinite sequence in the interval $[0,1)$, we have

$$\sup_{\alpha \in [0,1]} |D[n,\alpha]| \ge f(n) \qquad \text{for infinitely many } n \in \mathbb{N}? \tag{1.1}$$

Consider a distribution $\mathcal{P}$ of N points in the square $[0,1)^2$. For every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, let $B(\mathbf{x})$ denote the rectangle $[0,x_1) \times [0,x_2)$, and let

$$Z[\mathcal{P}; B(\mathbf{x})] = \#(\mathcal{P} \cap B(\mathbf{x}))$$

denote the number of points of $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x})$. We shall study the discrepancy

$$D[\mathcal{P}; B(\mathbf{x})] = Z[\mathcal{P}; B(\mathbf{x})] - N x_1 x_2,$$

noting that $N x_1 x_2$ represents the expected number of points of $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x})$.

We can pose the following question.

Question. Does there exist a function $g(N) \to +\infty$ as $N \to \infty$ such that for every distribution $\mathcal{P}$ of N points in the square $[0,1)^2$, we have

$$\sup_{\mathbf{x} \in [0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \ge g(N)? \tag{1.2}$$

In 1954, Roth showed that the above two problems are equivalent: If (1.1) holds with $f(n)$, then (1.2) holds with $g(N) = c_1 f(N)$ for a sufficiently small positive absolute constant c_1. If (1.2) holds with $g(N)$, then (1.1) holds with $f(n) = c_2 g(n)$ for a sufficiently small positive absolute constant c_2.

In these lectures, we shall establish the following results.

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

$$\int_0^1\int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^2\,\mathrm{d}x_1\mathrm{d}x_2 \gg \log N.$$

A simple consequence of Theorem 1 is the estimate

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

This can be sharpened as follows.

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

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

These lower bounds are complemented by the following upper bounds.

THEOREM 3. (Lerch 1904) *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the square $[0,1)^2$ such that*

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

THEOREM 4. (Davenport 1956) *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the square $[0,1)^2$ such that*

$$\int_0^1\int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^2\,\mathrm{d}x_1\mathrm{d}x_2 \ll \log N.$$

In other words, all the above results are best possible, apart from the implicit constants.

In fact, Roth (1954) established Theorem 1 in the setting of the K-dimensional cube $[0,1)^K$ for every positive integer $K \geq 2$, with a lower bound

$$\int_0^1 \cdots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^2\,\mathrm{d}x_1\ldots\mathrm{d}x_K \gg_K (\log N)^{K-1}. \tag{1.3}$$

This was further extended by Schmidt (1977), with a lower bound

$$\int_0^1 \cdots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^W\,\mathrm{d}x_1\ldots\mathrm{d}x_K \gg_{K,W} (\log N)^{\frac{1}{2}(K-1)W},$$

valid for every real number $W > 1$.

In the opposite direction, Roth (1980) extended Theorem 4 to the setting of the K-dimensional cube $[0,1)^K$ for every positive integer $K \geq 2$, with an upper bound

$$\int_0^1 \cdots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^2\,\mathrm{d}x_1\ldots\mathrm{d}x_K \ll_K (\log N)^{K-1}.$$

This was further extended by Chen (1980), with an upper bound

$$\int_0^1 \cdots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]|^W\,\mathrm{d}x_1\ldots\mathrm{d}x_K \ll_{K,W} (\log N)^{\frac{1}{2}(K-1)W},$$

valid for every real number $W > 0$.

Earlier, Halton (1960) extended Theorem 3 to the setting of the K-dimensional cube $[0,1)^K$ for every positive integer $K \geq 2$, with an upper bound

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

On the other hand, the lower bound (1.3) by Roth gives the lower bound

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

We therefore have the following intriguing question.

Great Open Problem. Suppose that $K \geq 3$ is an integer. Is it true that for every distribution $\mathcal{P}$ of N points in the cube $[0,1)^K$, we have

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

Halász (1981) studied this problem by a variation of the technique of Roth (1954). However, his ideas did not work when $K \geq 3$. Nevertheless, he obtained an alternative proof of Theorem 2 above. He was also able to establish the lower bound

$$\int_0^1 \ldots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}x_1 \ldots \mathrm{d}x_K \gg_K (\log N)^{\frac{1}{2}}. \tag{1.4}$$

Note that this is best possible when $K = 2$, but very weak when $K \geq 3$.

Second Great Open Problem. Suppose that $K \geq 3$ is an integer. Is it true that for every distribution $\mathcal{P}$ of N points in the cube $[0,1)^K$, we have

$$\int_0^1 \ldots \int_0^1 |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}x_1 \ldots \mathrm{d}x_K \gg_K (\log N)^{\frac{1}{2}(K-1)}?$$

More recently, by using the technique of Halász and ideas from graph theory, Beck (1989) establish the lower bound

$$\sup_{\mathbf{x}\in[0,1]^3} |D[\mathcal{P};B(\mathbf{x})]| \gg_\delta (\log N)(\log\log N)^{\frac{1}{8}-\delta}$$

for every $\delta > 0$. This was improved by Baker (1999) who adapted Beck's technique and established the lower bound

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

More importantly, Baker was able to remove the graph theory from Beck's argument.

We shall first study lower bounds. In Section 2, we shall establish Theorem 1 using the technique of Roth (1954) made more transparent by the treatment of Schmidt (1977). We then study the modification of the Roth technique by Halász (1981) in Section 3 and establish Theorem 2 as well as the lower bound (1.4) in the case $K = 2$.

We then turn our attention to upper bounds. In Section 4, we shall introduce the van der Corput sequence and use it to establish Theorem 3. In Section 5, we shall establish Theorem 4 by the technique of Davenport (1956) and using the sequence $\{\sqrt{2}\}, \{2\sqrt{2}\}, \{3\sqrt{2}\}, \ldots$ of fractional parts.

2. Roth's Orthogonal Function Method

Corresponding to every distribution $\mathcal{P}$ of N points in the square $[0,1)^2$, Roth constructed an auxiliary function $F(\mathbf{x}) = F[\mathcal{P};\mathbf{x}]$ such that, writing $D(\mathbf{x}) = D[\mathcal{P};B(\mathbf{x})]$ and $\mathrm{d}\mathbf{x} = \mathrm{d}x_1\mathrm{d}x_2$, we have

$$\left|\int_0^1\int_0^1 F(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| > c_1 \log N \tag{2.1}$$

and

$$\int_0^1\int_0^1 |F(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} < c_2 \log N. \tag{2.2}$$

Here, c_1 and c_2 are positive absolute constants. These, together with Schwarz's inequality

$$\left|\int_0^1\int_0^1 F(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|^2 \le \left(\int_0^1\int_0^1 |F(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}\right)\left(\int_0^1\int_0^1 |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x}\right),$$

give the inequality

$$\int_0^1\int_0^1 |D(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} > c_3 \log N,$$

where c_3 is a positive absolute constant.

We shall choose a non-negative integer n satisfying the condition

$$2^{n-1} < 2N \le 2^n. \tag{2.3}$$

Suppose that the integer i satisfies $0 \le i \le n$. We can partition the square $[0,1)^2$ into a union of 2^n rectangles with horizontal side length 2^{-i} and vertical side length 2^{i-n}, so that each such rectangle has area 2^{-n} and is of the form

$$[m_1 2^{-i}, (m_1+1)2^{-i}) \times [m_2 2^{i-n}, (m_2+1)2^{i-n}), \tag{2.4}$$

where $m_1, m_2 \in \mathbb{Z}$. Let B be such a rectangle. We shall define the function $R_i(\mathbf{x})$ for $\mathbf{x} \in B$ by writing $R_i(\mathbf{x}) = \pm 1$ according to the following picture:

-1	+1
+1	-1

We then define the function $f_i(\mathbf{x})$ for $\mathbf{x} \in B$ by writing

$$f_i(\mathbf{x}) = \begin{cases} 0 & \text{if } B \cap \mathcal{P} \ne \emptyset, \\ R_i(\mathbf{x}) & \text{if } B \cap \mathcal{P} = \emptyset. \end{cases}$$

We now consider the auxiliary function

$$F(\mathbf{x}) = \sum_{i=0}^{n} f_i(\mathbf{x}). \tag{2.5}$$

LEMMA 2A. *Suppose that the integers $i, j \in \mathbb{Z}$ satisfy $0 \le i < j \le n$. Then*

$$\int_0^1\int_0^1 f_i(\mathbf{x})f_j(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

Proof. We can partition the square $[0,1)^2$ into a union of 2^{n-i+j} rectangles of horizontal side length 2^{-j} and vertical side length 2^{i-n}. In any such rectangle S, we either have $f_i(\mathbf{x})f_j(\mathbf{x}) = 0$ for every $\mathbf{x} \in S$, or have $f_i(\mathbf{x})f_j(\mathbf{x}) = \pm 1$ according to one of the following two pictures:

-1	+1
+1	-1

+1	-1
-1	+1

In either case, we clearly have

$$\iint_S f_i(\mathbf{x})f_j(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

The result follows immediately. ♣

It now follows from Lemma 2A that

$$\int_0^1\int_0^1 |F(\mathbf{x})|^2\,\mathrm{d}\mathbf{x} = \sum_{i=0}^{n}\sum_{j=0}^{n}\int_0^1\int_0^1 f_i(\mathbf{x})f_j(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sum_{i=0}^{n}\int_0^1\int_0^1 f_i^2(\mathbf{x})\,\mathrm{d}\mathbf{x} \le n+1.$$

In view of (2.3), this establishes the inequality (2.2).

LEMMA 2B. *Suppose that the integer $i \in \mathbb{Z}$ satisfies $0 \le i \le n$. Then*

$$\int_0^1\int_0^1 f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} \le -N2^{-n-5}.$$

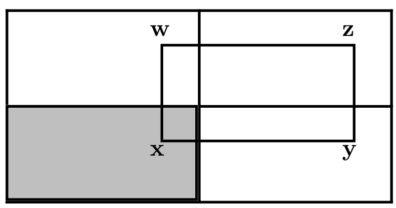
Proof. Suppose that B is a rectangle of the form (2.4), where $m_1, m_2 \in \mathbb{Z}$. Suppose first of all that $B \cap \mathcal{P} \ne \emptyset$. Then $f_i(\mathbf{x}) = 0$ for every $\mathbf{x} \in B$, and so

$$\iint_B f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

Suppose next that $B \cap \mathcal{P} = \emptyset$. We shall consider the rectangle

$$B' = [m_1 2^{-i}, (m_1 + \tfrac{1}{2})2^{-i}) \times [m_2 2^{i-n}, (m_2 + \tfrac{1}{2})2^{i-n}).$$

For every $\mathbf{x} \in B'$, we define $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{w}$ as vertices of a rectangle of horizontal side length 2^{-i-1} and vertical side length 2^{i-n-1} as shown in the picture below:



Here the shaded rectangle is B'. Observe now that

$$\iint_B f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} = \iint_{B'} (D(\mathbf{x}) - D(\mathbf{y}) + D(\mathbf{z}) - D(\mathbf{w}))\,\mathrm{d}\mathbf{x} = \iint_{B'} D(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w})\,\mathrm{d}\mathbf{x},$$

where it is easy to show that

$$D(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}) = D(\mathbf{x}) - D(\mathbf{y}) + D(\mathbf{z}) - D(\mathbf{w})$$

represents the discrepancy in the rectangle with vertices $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{w}$. Since $B \cap \mathcal{P} = \emptyset$, it follows that

$$D(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}) = 0 - N2^{-n-2}.$$

On the other hand, the rectangle B' has area 2^{-n-2}. Hence

$$\iint_B f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} = -N2^{-2n-4}. \tag{2.6}$$

In view of (2.3), there are at least $2^n - N \geq 2^{n-1}$ rectangles B of the type (2.4) where $B \cap \mathcal{P} = \emptyset$. It follows that

$$\int_0^1\int_0^1 f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq -N2^{-2n-4}2^{n-1} = -N2^{-n-5}$$

as required. ♣

It now follows from Lemma 2B that

$$\int_0^1\int_0^1 F(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sum_{i=0}^{n}\int_0^1\int_0^1 f_i(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq -N2^{-n-5}(n+1). \tag{2.7}$$

In view of (2.3), this establishes the inequality (2.1).

The proof of Theorem 1 is now complete.

3. Halász's Modification of Roth's Method

Corresponding to every distribution $\mathcal{P}$ of N points in the square $[0,1)^2$, Halász constructed an auxiliary function $H(\mathbf{x}) = H[\mathcal{P};\mathbf{x}]$ such that, writing $D(\mathbf{x}) = D[\mathcal{P};B(\mathbf{x})]$ and $\mathrm{d}\mathbf{x} = \mathrm{d}x_1\mathrm{d}x_2$, we have

$$\left|\int_0^1\int_0^1 H(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| > c_1 \log N \tag{3.1}$$

and

$$\int_0^1\int_0^1 |H(\mathbf{x})|\,\mathrm{d}\mathbf{x} < c_2. \tag{3.2}$$

Here, c_1 and c_2 are positive absolute constants. These, together with the trivial inequality

$$\left|\int_0^1\int_0^1 H(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \leq \left(\sup_{\mathbf{x}\in[0,1]^2}|D(\mathbf{x})|\right)\left(\int_0^1\int_0^1 |H(\mathbf{x})|\,\mathrm{d}\mathbf{x}\right),$$

give the inequality

$$\sup_{\mathbf{x}\in[0,1]^2}|D(\mathbf{x})| > c_3 \log N,$$

where c_3 is a positive absolute constant.

As before, we shall choose a non-negative integer n satisfying the condition

$$2^{n-1} < 2N \leq 2^n. \tag{3.3}$$

For every integer i satisfying $0 \le i \le n$, we define the function $f_i(\mathbf{x})$ as in Section 2. We then consider the auxiliary function

$$H(\mathbf{x}) = \prod_{i=0}^{n}(1+\alpha f_i(\mathbf{x})) - 1,$$

where α is a constant satisfying $0 < \alpha < 1/2$, to be determined later.

We shall first establish the following simple extension of Lemma 2A.

LEMMA 3A. *Suppose that the integers $i_1, \ldots, i_k \in \mathbb{Z}$ satisfy $0 \le i_1 < \ldots < i_k \le n$. Then*

$$\int_0^1\int_0^1 f_{i_1}(\mathbf{x}) \ldots f_{i_k}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

PROOF. We can partition the square $[0,1)^2$ into a union of $2^{n-i_1+i_k}$ rectangles of horizontal side length 2^{-i_k} and vertical side length 2^{i_1-n}. In any such rectangle S, we either have $f_{i_1}(\mathbf{x}) \ldots f_{i_k}(\mathbf{x}) = 0$ for every $\mathbf{x} \in S$, or have $f_{i_1}(\mathbf{x}) \ldots f_{i_k}(\mathbf{x}) = \pm 1$ according to one of the following two pictures:

-1	+1
+1	-1

+1	-1
-1	+1

In either case, we clearly have

$$\iint_S f_{i_1}(\mathbf{x}) \ldots f_{i_k}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

The result follows immediately. ♣

Note that

$$\prod_{i=0}^{n}(1+\alpha f_i(\mathbf{x})) = 1 + \alpha F(\mathbf{x}) + \sum_{k=2}^{n+1} \alpha^k F_k(\mathbf{x}),$$

where the function $F(\mathbf{x})$ is Roth's auxiliary function (2.5) and where, for every $k = 2, \ldots, n+1$, we have

$$F_k(\mathbf{x}) = \mathop{\sum_{i_1=0}^{n} \ldots \sum_{i_k=0}^{n}}_{i_1<\ldots<i_k} f_{i_1}(\mathbf{x}) \ldots f_{i_k}(\mathbf{x}).$$

In view of the restriction on α, it follows from the triangle inequality that

$$|H(\mathbf{x})| \le \prod_{i=0}^{n}(1+\alpha f_i(\mathbf{x})) + 1.$$

It then follows from Lemma 3A that

$$\int_0^1\int_0^1 |H(\mathbf{x})|\,\mathrm{d}\mathbf{x} \le 2.$$

This establishes the inequality (3.2).

Clearly

$$H(\mathbf{x}) = \alpha F(\mathbf{x}) + \sum_{k=2}^{n+1} \alpha^k F_k(\mathbf{x}). \tag{3.4}$$

LEMMA 3B. *Suppose that the integers $i_1,\ldots,i_k\in\mathbb{Z}$ satisfy $0\le i_1<\ldots<i_k\le n$. Then*

$$\left|\int_0^1\int_0^1 f_{i_1}(\mathbf{x})\ldots f_{i_k}(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\le N2^{-n+i_i-i_k-4}.$$

PROOF. Suppose that S is one of the $2^{n-i_1+i_k}$ rectangles in the proof of Lemma 3A, of horizontal side length 2^{-i_k} and vertical side length 2^{i_1-n}. For any such rectangle S, either $f_{i_1}(\mathbf{x})\ldots f_{i_k}(\mathbf{x})=\pm1$ for every $\mathbf{x}\in S$ or $f_{i_1}(\mathbf{x})\ldots f_{i_k}(\mathbf{x})=0$ for every $\mathbf{x}\in S$. In the first case, we must have $S\cap\mathcal{P}=\emptyset$. It can then be shown, as in the deduction of (2.6) in the proof of Lemma 2B, that

$$\iint_S f_{i_1}(\mathbf{x})\ldots f_{i_k}(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}=\pm N2^{-2(n-i_1+i_k)-4},$$

so that

$$\left|\iint_S f_{i_1}(\mathbf{x})\ldots f_{i_k}(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\le N2^{-2(n-i_1+i_k)-4}.$$

Note that the last inequality is trivially satisfied in the second case. The result now follows on applying the triangle inequality. ♣

LEMMA 3C. *For every $k=2,\ldots,n+1$, we have*

$$\left|\int_0^1\int_0^1 F_k(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\le\sum_{i=0}^{n-k+1}\sum_{h=1}^{n-i}N2^{-n-h-4}\binom{h-1}{k-2}.$$

PROOF. Suppose that $i_1=i$ and $i_k=i+h$, where i and h are fixed. Then there are precisely $\binom{h-1}{k-2}$ choices for integers $i_2,\ldots,i_{k-1}$ that satisfy the inequalities $i_1<i_2<\ldots<i_{k-1}<i_k$. The result follows immediately. ♣

It now follows from Lemma 3C that

$$\begin{aligned}\left|\sum_{k=2}^{n+1}\alpha^k\int_0^1\int_0^1 F_k(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|&\le\sum_{k=2}^{n+1}\sum_{i=0}^{n-k+1}\sum_{h=1}^{n-i}\alpha^kN2^{-n-h-4}\binom{h-1}{k-2}\\&=\sum_{i=0}^{n-1}\sum_{h=1}^{n-i}\sum_{k=2}^{h+1}\alpha^2N2^{-n-h-4}\binom{h-1}{k-2}\alpha^{k-2}\le N\sum_{i=0}^{n-1}\sum_{h=1}^{\infty}2^{-n-h-4}\alpha^2(1+\alpha)^h\\&\le\alpha^2N2^{-n-4}n\sum_{h=0}^{\infty}\left(\frac{1+\alpha}{2}\right)^h\le\alpha^2N2^{-n-2}n,\end{aligned}\tag{3.5}$$

since $0<\alpha<1/2$. In view of (3.3), the inequality (2.7) is valid. Combining this with (3.4) and (3.5) gives

$$\begin{aligned}\left|\int_0^1\int_0^1 H(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|&\ge\alpha\left|\int_0^1\int_0^1 F(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|-\left|\sum_{k=2}^{n+1}\alpha^k\int_0^1\int_0^1 F_k(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|\\&\ge\alpha N2^{-n-5}(n+1)-\alpha^2N2^{-n-2}n.\end{aligned}$$

This establishes the inequality (3.1) if we choose $\alpha=2^{-6}$. The proof of Theorem 2 is now complete.

We conclude this section by indicating the proof of the inequality (1.4) in the special case $K=2$. Corresponding to every distribution $\mathcal{P}$ of N points in the square $[0,1)^2$, Halász constructed another auxiliary function $T(\mathbf{x})=T[\mathcal{P};\mathbf{x}]$ such that, writing $D(\mathbf{x})=D[\mathcal{P};B(\mathbf{x})]$ and $\mathrm{d}\mathbf{x}=\mathrm{d}x_1\mathrm{d}x_2$, we have

$$\left|\int_0^1\int_0^1 T(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right|>c_4(\log N)^{\frac12}\tag{3.6}$$

and

$$\sup_{\mathbf{x}\in[0,1]^2} |T(\mathbf{x})| < c_5.$$

Here, c_4 and c_5 are positive absolute constants. These, together with the trivial inequality

$$\left|\int_0^1\int_0^1 T(\mathbf{x})D(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \le \left(\sup_{\mathbf{x}\in[0,1]^2} |T(\mathbf{x})|\right)\left(\int_0^1\int_0^1 |D(\mathbf{x})|\,\mathrm{d}\mathbf{x}\right),$$

give the inequality

$$\int_0^1\int_0^1 |D(\mathbf{x})|\,\mathrm{d}\mathbf{x} > c_6(\log N)^{\frac{1}{2}},$$

where c_6 is a positive absolute constant.

As before, we shall choose a non-negative integer n satisfying the condition (3.3). We then consider the auxiliary function

$$T(\mathbf{x}) = \prod_{i=0}^{n}(1 + \mathrm{i}n^{-1/2} f_i(\mathbf{x})) - 1,$$

where $\mathrm{i} = \sqrt{-1}$ and the functions $f_i(\mathbf{x})$ are defined as in Section 2. Here it is crucial to use complex numbers. Note that

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

is bounded, a result we are not able to achieve if the imaginary factor i is not present in the auxiliary function $T(\mathbf{x})$. The proof of the inequality (3.6) is now rather similar to that of the inequality (3.1).

4. The Method of van der Corput

In this section, we shall construct point sets with low supremum discrepancy. To do this, it is notationally convenient to consider infinite sets of points in $[0,1) \times [0,\infty)$ such that there is an average of one point per unit area. We then consider those N points contained in the rectangle $[0,1) \times [0,N)$, and rescale the second coordinate to obtain a set of N points in the square $[0,1)^2$.

We shall be concerned with rectangles in $[0,1) \times [0,\infty)$ of the form

$$I \times I_0, \tag{4.1}$$

where I is an interval of the form $[\alpha,\beta) \subseteq [0,1)$, while I_0 is an interval of the form $[\alpha_0,\beta_0)$ satisfying the conditions $0 \le \alpha_0 < \beta_0$. We shall look for distributions such that many rectangles of the type (4.1) will contain the right number of points.

Definition. Let s be a non-negative integer.

(1) By an elementary interval of order s, we mean an interval of the type

$$[m2^{-s}, (m+1)2^{-s}) \subseteq [0,1),$$

where m is a non-negative integer.

(2) By an elementary rectangle of order s, we mean a rectangle of area 1 and of the form

$$[m2^{-s}, (m+1)2^{-s}) \times [k2^s, (k+1)2^s) \subseteq [0,1) \times [0,\infty),$$

where m and k are non-negative integers.

For every non-negative integer n, we can consider the dyadic expansion

$$n = \sum_{\nu=1}^{\infty} a_\nu 2^{\nu-1},$$

where the integers $a_\nu \in \{0,1\}$ are uniquely determined by n. Let

$$x(n) = \sum_{\nu=1}^{\infty} a_\nu 2^{-\nu}.$$

We now consider the infinite set

$$\mathcal{Q} = \{(x(n), n) : n \in \mathbb{N} \cup \{0\}\}.$$

LEMMA 4A. *For every non-negative integer s, every elementary rectangle of order s contains exactly one point of the set $\mathcal{Q}$.*

PROOF. The condition $x(n) \in [m2^{-s}, (m+1)2^{-s})$ determines the coefficients $a_1, \ldots, a_s$ uniquely, and the condition $n \in [k2^s, (k+1)2^s)$ determines the coefficients $a_{s+1}, a_{s+2}, \ldots$ uniquely. ♣

For any rectangle B of the type (4.1) in $[0,1) \times [0,\infty)$, let $Z[\mathcal{Q}; B]$ denote the number of points of $\mathcal{Q}$ in B, and write

$$E[\mathcal{Q}; B] = Z[\mathcal{Q}; B] - \mu(B),$$

where $\mu(B)$ denotes the area of B. Note that if $B = B_1 \cup B_2$, where $B_1 \cap B_2 = \emptyset$, then

$$E[\mathcal{Q}; B] = E[\mathcal{Q}; B_1] + E[\mathcal{Q}; B_2].$$

For any natural number $N \geq 2$, we now choose a positive integer h to satisfy the inequalities

$$2^{h-1} < N \leq 2^h. \tag{4.2}$$

LEMMA 4B. *For any rectangle of the form $B(x,y) = [0,x) \times [0,y)$, where $x \in [0,1]$ and $y \in [0,N]$, we have*

$$|E[\mathcal{Q}; B(x,y)]| \leq h+1.$$

PROOF. The cases when $x=0$ or $x=1$ are trivial, so we assume that $0 < x < 1$. For every $s = 1, \ldots, h$, let $x_s = 2^{-s}[2^s x]$ denote the greatest integer multiple of 2^{-s} not exceeding x. For convenience, write $x_0 = 0$. Then

$$B(x,y) = \left(\bigcup_{s=1}^{h} ([x_{s-1}, x_s) \times [0,y))\right) \cup ([x_h, x) \times [0,y)),$$

where the union is clearly disjoint, so that

$$|E[\mathcal{Q}; B(x,y)]| \leq \left(\sum_{s=1}^{h} |E[\mathcal{Q}; [x_{s-1}, x_s) \times [0,y)]|\right) + |E[\mathcal{Q}; [x_h, x) \times [0,y)]|. \tag{4.3}$$

Suppose that $x_{s-1} \neq x_s$. Then $x_s - x_{s-1} = 2^{-s}$. Let $y_s = 2^s[2^{-s}y]$ denote the greatest integer multiple of 2^s not exceeding y. Then

$$E[\mathcal{Q}; [x_{s-1}, x_s) \times [0,y)] = E[\mathcal{Q}; [x_{s-1}, x_s) \times [0,y_s)] + E[\mathcal{Q}; [x_{s-1}, x_s) \times [y_s, y)].$$

Clearly the rectangle $[x_{s-1}, x_s) \times [0, y_s)$ is the union of a finite number of elementary rectangles of order s, so that $E[\mathcal{Q}; [x_{s-1}, x_s) \times [0, y_s)] = 0$. On the other hand, the rectangle $[x_{s-1}, x_s) \times [y_s, y)$ is contained in an elementary rectangle of order s, and so $|E[\mathcal{Q}; [x_{s-1}, x_s) \times [y_s, y)]| \leq 1$. Hence

$$|E[\mathcal{Q}; [x_{s-1}, x_s) \times [0,y)]| \leq 1 \tag{4.4}$$

for every $s = 1, \ldots, h$. On the other hand, since $y \le N \le 2^h$, the rectangle $[x_h, x) \times [0, y)$ is contained in an elementary rectangle of order h, and so

$$|E[\mathcal{Q}; [x_h, x) \times [0, y)]| \le 1. \tag{4.5}$$

The desired inequality now follows on combining (4.3)–(4.5). ♣

To deduce Theorem 3, we now consider the finite set

$$\mathcal{Q} \cap ([0,1) \times [0, N)) = \{(x(n), n) : n = 0, 1, \ldots, N-1\}.$$

Rescaling in the vertical direction, we obtain the set

$$\mathcal{P} = \{(x(n), n/N) : n = 0, 1, \ldots, N-1\}$$

in the square $[0,1)^2$. Note that for every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, we have

$$D[\mathcal{P}; B(\mathbf{x})] = E[\mathcal{Q}; B(x_1, N x_2)].$$

It follows from Lemma 4B that

$$\sup_{\mathbf{x} \in [0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \le h + 1. \tag{4.6}$$

This completes the proof of Theorem 3, in view of (4.2).

Consider next the special case when $N = 2^h$ for some positive integer h. Then the van der Corput point set is given by

$$\mathcal{P} = \left\{ \left(\sum_{\nu=1}^{h} a_\nu 2^{-\nu}, \sum_{\nu=1}^{h} a_\nu 2^{\nu-1-h} \right) : a_1, \ldots, a_h \in \{0,1\} \right\}.$$

Of course, the upper bound (4.6) holds. However, this set does not satisfy the conclusion of Theorem 4, for Chen and Skriganov (1999) have shown that for this point set, we have

$$\int_0^1 \int_0^1 |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}x_1 \mathrm{d}x_2 = 2^{-6} h^2 + O(h).$$

We remark that a lower bound without the specific constant 2^{-6} was given by Matoušek (1999).

On the other hand, this point set is the basis of Roth's work in connection with Theorem 4. Roth's idea is to introduce a translation variable t. In this special case when $N = 2^h$, we consider instead the translated point set

$$\mathcal{P}(t) = \left\{ \left(\sum_{\nu=1}^{h} a_\nu 2^{-\nu}, \left\{ t + \sum_{\nu=1}^{h} a_\nu 2^{\nu-1-h} \right\} \right) : a_1, \ldots, a_h \in \{0,1\} \right\},$$

where the second coordinates of the points have been translated by t modulo 1. One can then show that

$$\int_0^1 \int_0^1 \int_0^1 |D[\mathcal{P}(t); B(\mathbf{x})]|^2 \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}t \ll h,$$

so that there exists some $t^* \in [0,1]$ such that the point set $\mathcal{P}(t^*)$ satisfies the conclusion of Theorem 4. This is essentially the birth of probabilistic techniques in the theory of irregularities of point distribution.

5. The Method of Davenport

Consider a lattice Λ on the plane generated by the two vectors $(1,0)$ and $(\theta,1)$, where θ is an irrational number. Suppose that M is a positive integer. We are interested in the set $\mathcal{Q}$ which contains precisely the M points of Λ that fall into the rectangle $[0,1) \times [0, M)$. Clearly

$$\mathcal{Q} = \{(\{\theta n\}, n) : 0 \le n \le M-1\}.$$

For any rectangle of the form

$$B(y_1, y_2) = [0, y_1) \times [0, y_2) \subseteq [0, 1) \times [0, M),$$

let $Z[\mathcal{Q}; B(y_1, y_2)]$ denote the number of points of $\mathcal{Q}$ in $B(y_1, y_2)$, and write

$$E[\mathcal{Q}; B(y_1, y_2)] = Z[\mathcal{Q}; B(y_1, y_2)] - y_1 y_2.$$

We shall use the sawtooth function $\phi(x)$. This is defined by $\phi(x) = x - [x] - 1/2$ when $x \notin \mathbb{Z}$ and by $\phi(x) = 0$ when $x \in \mathbb{Z}$.

LEMMA 5A. *Suppose that y_2 is an integer in the interval $[0, M]$. Then*

$$E[\mathcal{Q}; B(y_1, y_2)] = \sum_{n=0}^{y_2-1} (\phi(\theta n - y_1) - \phi(\theta n))$$

for all but a finite number of values of y_1 in the interval $[0, 1]$.

PROOF. Suppose that $0 < y_1 \le 1$. Then it is easy to check that

$$y_1 + \phi(x - y_1) - \phi(x) = \begin{cases} 1 & \text{if } 0 < \{x\} < y_1, \\ 0 & \text{if } \{x\} > y_1. \end{cases}$$

Hence

$$Z[\mathcal{Q}; B(y_1, y_2)] = \sum_{n=0}^{y_2-1} (y_1 + \phi(\theta n - y_1) - \phi(\theta n)) = y_1 y_2 + \sum_{n=0}^{y_2-1} (\phi(\theta n - y_1) - \phi(\theta n)).$$

The result follows immediately. ♣

REMARK. The use of the function $\phi(x)$ is a technical device. We really want to study the characteristic function.

The function $\phi(x)$ has the Fourier expansion

$$\phi(x) \sim -\sum_{m \ne 0} \frac{e(xm)}{2\pi \mathrm{i} m},$$

where $e(\beta) = \mathrm{e}^{2\pi \mathrm{i}\beta}$ for every $\beta \in \mathbb{R}$. It follows that if y_2 is an integer in the interval $[0, M]$, then the discrepancy function $E[\mathcal{Q}; B(y_1, y_2)]$ has Fourier expansion

$$E[\mathcal{Q}; B(y_1, y_2)] \sim \sum_{m \ne 0} \left(\frac{1 - e(-y_1 m)}{2\pi \mathrm{i} m} \right) \left(\sum_{n=0}^{y_2-1} e(\theta n m) \right). \tag{5.1}$$

Ideally, we would like to square the expression (5.1) and integrate with respect to y_1 over the interval $[0, 1]$. Unfortunately, the term 1 in the numerator $1 - e(-y_1 m)$ proves to be a nuisance.

In order to overcome this difficulty, we consider an extra lattice Λ' on the plane generated by the two vectors $(1, 0)$ and $(-\theta, 1)$. Then

$$\mathcal{Q}' = \{(\{-\theta n\}, n) : 0 \le n \le M - 1\}$$

is the set which contains precisely the M points of Λ' that fall into the rectangle $[0, 1) \times [0, M)$.

We now consider the $2M$ points of $\mathcal{Q} \cup \mathcal{Q}'$ that fall into this rectangle $[0,1) \times [0,M)$. For every rectangle of the form $B(y_1,y_2)$ discussed earlier, let $Z[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)]$ denote the number of points of $\mathcal{Q} \cup \mathcal{Q}'$ in $B(y_1,y_2)$, and write

$$F[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)] = Z[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)] - 2y_1y_2 = E[\mathcal{Q}; B(y_1,y_2)] + E[\mathcal{Q}'; B(y_1,y_2)].$$

Then it is easy to see that if y_2 is an integer in the interval $[0,M]$, then

$$F[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)] = \sum_{n=0}^{y_2-1} (\phi(\theta n - y_1) - \phi(\theta n + y_1))$$

for all but a finite number of values of y_1 in the interval $[0,1]$. Furthermore, this has Fourier expansion

$$F[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)] \sim \sum_{m\neq 0} \left(\frac{e(y_1m) - e(-y_1m)}{2\pi \mathrm{i} m} \right) \left(\sum_{n=0}^{y_2-1} e(\theta nm) \right). \tag{5.2}$$

We now square the expression (5.2) and integrate with respect to y_1 over the interval $[0,1]$. By Parseval's theorem, we have

$$\int_0^1 |F[\mathcal{Q} \cup \mathcal{Q}'; B(y_1,y_2)]|^2 \,\mathrm{d}y_1 \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{n=0}^{y_2-1} e(\theta nm) \right|^2. \tag{5.3}$$

To estimate the sum on the right hand side of (5.3), we need to make some assumptions on the number θ. Suppose that θ has a continued fraction expansion with bounded partial quotients. Appealing to the theory of diophantine approximation, we know that there is a constant $c = c(\theta)$, depending only on θ, such that

$$m\|m\theta\| > c > 0 \tag{5.4}$$

for every natural number $m \in \mathbb{N}$, where $\|\cdot\|$ denotes the distance to the nearest integer. We may take $\theta = \sqrt{2}$ throughout if we wish.

LEMMA 5B. *Suppose that y_2 is a positive integer. Then*

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{n=0}^{y_2-1} e(\theta nm) \right|^2 \ll \log(2y_2).$$

PROOF. It is well known that

$$\left| \sum_{n=0}^{y_2-1} e(\theta nm) \right| \ll \min\{y_2, \|m\theta\|^{-1}\},$$

so that

$$S = \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{n=0}^{y_2-1} e(\theta nm) \right|^2 \ll \sum_{h=1}^{\infty} 2^{-2h} \sum_{2^{h-1} \leq m < 2^h} \min\{y_2^2, \|m\theta\|^{-1}\}.$$

The condition (5.4) implies that if $2^{h-1} \leq m < 2^h$, then

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

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

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

for otherwise the difference $(m_1 - m_2)$ of two of them would contradict (5.4). It follows that

$$\begin{aligned} S &\ll \sum_{h=1}^{\infty}\sum_{p=1}^{\infty} \min\{2^{-2h}y_2^2, p^{-2}\} \\ &= \sum_{2^h \le y_2}\sum_{p=1}^{\infty} \min\{2^{-2h}y_2^2, p^{-2}\} + \sum_{2^h > y_2}\sum_{p=1}^{\infty} \min\{2^{-2h}y_2^2, p^{-2}\} \\ &\ll \sum_{2^h \le y_2}\sum_{p=1}^{\infty} p^{-2} + \sum_{2^h > y_2}\left(2^{-2h}y_2^2 2^h y_2^{-1} + \sum_{p > 2^h y_2^{-1}} p^{-2}\right) \\ &\ll \sum_{2^h \le y_2} 1 + \sum_{2^h > y_2} 2^{-h}y_2 \ll \log(2y_2) \end{aligned}$$

as required. ♣

Relaxing the restriction that y_2 is an integer introduces an error of $O(1)$, where the implicit constant depends at most on θ. Integrating trivially with respect to y_2 over the interval $[0, M]$, we obtain

$$\int_0^M \int_0^1 |F[\mathcal{Q} \cup \mathcal{Q}'; B(y_1, y_2)]|^2 \, \mathrm{d}y_1 \mathrm{d}y_2 \ll M \log(2M),$$

where the implicit constant depends at most on θ. Rescaling in the y_2-direction by a factor $1/M$, we see that the set

$$\mathcal{P} = \{(\{\pm\theta n\}, n/M) : 0 \le n \le M - 1\}$$

of $N = 2M$ points in $[0, 1)^2$ satisfies the requirements of Theorem 4.

More recently, Roth (1979) devised an ingenious variation of Davenport's argument. For any real number $t \in \mathbb{R}$, we consider the translated lattice

$$t\mathbf{i} + \Lambda = \{t\mathbf{i} + \mathbf{v} : \mathbf{v} \in \Lambda\}.$$

We are interested in the set $\mathcal{Q}(t)$ which contains precisely the M points of $t\mathbf{i} + \Lambda$ that fall into the rectangle $[0, 1) \times [0, M)$. Clearly

$$\mathcal{Q}(t) = \{(\{t + \theta n\}, n) : 0 \le n \le M - 1\}.$$

For any rectangle of the form

$$B(y_1, y_2) = [0, y_1) \times [0, y_2) \subseteq [0, 1) \times [0, M),$$

let $Z[\mathcal{Q}(t); B(y_1, y_2)]$ denote the number of points of $\mathcal{Q}(t)$ in $B(y_1, y_2)$, and write

$$E[\mathcal{Q}(t); B(y_1, y_2)] = Z[\mathcal{Q}(t); B(y_1, y_2)] - y_1 y_2.$$

A similar argument as above will show that if y_2 is an integer in the interval $[0, M]$, then the discrepancy function $E[\mathcal{Q}(t); B(y_1, y_2)]$ has Fourier expansion

$$E[\mathcal{Q}(t); B(y_1, y_2)] \sim \sum_{m \ne 0} \left(\frac{1 - e(-y_1 m)}{2\pi \mathrm{i} m}\right) \left(\sum_{n=0}^{y_2 - 1} e(\theta n m)\right) e(tm). \tag{5.5}$$

We now square the expression (5.5) and integrate with respect to t over the interval $[0, 1]$. By Parseval's theorem, we have

$$\int_0^1 |E[\mathcal{Q}(t); B(y_1, y_2)]|^2 \, \mathrm{d}t \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left|\sum_{n=0}^{y_2 - 1} e(\theta n m)\right|^2.$$

Furthermore, if θ has a continued fraction expansion with bounded partial quotients, then integrating trivially with respect to y_1 over the interval $[0,1]$ and with respect to y_2 over the interval $[0,M]$, we have

$$\int_0^1\int_0^M\int_0^1 |E[\mathcal{Q}(t);B(y_1,y_2)]|^2\,\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}t \ll M\log(2M),$$

where the implicit constant depends at most on θ. Hence there exists $t^*\in[0,1]$ such that the set $\mathcal{Q}(t^*)$ satisfies

$$\int_0^M\int_0^1 |E[\mathcal{Q}(t^*);B(y_1,y_2)]|^2\,\mathrm{d}y_1\mathrm{d}y_2 \ll M\log(2M).$$

Rescaling in the y_2-direction by a factor $1/M$, we see that the set

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

of $N=M$ points in $[0,1)^2$ satisfies the requirements of Theorem 4.

6. Generalizations of the Problem

For convenience, we shall consider the torus $[0,1)^2$. For every measurable set $B\subseteq[0,1)^2$, let

$$Z[\mathcal{P};B]=\#(\mathcal{P}\cap B)$$

denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

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

where μ denotes the usual measure in $\mathbb{R}^2$ restricted to $[0,1)^2$.

In the late 1960's and early 1970's, Schmidt developed an integral equation method and proved many new results. To understand some of these results, let us first of all observe that Theorem 2 can be rephrased as follows.

THEOREM 2. (Schmidt 1972) *For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$, there exists an aligned rectangle B in $[0,1)^2$ such that*

$$|D[\mathcal{P};B]|\gg\log N.$$

Suppose now that we no longer require the rectangles to be aligned to the coordinate axes. In other words, suppose that we allow rotations of the rectangles. Then the situation is very different.

THEOREM 5. (Schmidt 1968) *For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$, there exists a tilted rectangle B in $[0,1)^2$, of diameter less than 1, such that*

$$|D[\mathcal{P};B]|\gg_\epsilon N^{\frac{1}{4}-\epsilon}.$$

On the other hand, discs are invariant under rotation.

THEOREM 6. (Schmidt 1968) *For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$, there exists a disc B in $[0,1)^2$, of diameter less than 1, such that*

$$|D[\mathcal{P};B]|\gg_\epsilon N^{\frac{1}{4}-\epsilon}.$$

The above two results are not restricted to dimension $K = 2$. Theorem 5 can be generalized to $[0,1)^3$, and there exists a tilted rectangular box B in $[0,1)^3$, of diameter less than 1, such that

$$|D[\mathcal{P};B]| \gg_\epsilon N^{\frac{1}{3}-\epsilon}.$$

Unfortunately, Schmidt's method fails for dimension $K \geq 4$. On the other hand, Theorem 6 can be generalized to $[0,1)^K$ for any dimension $K \geq 2$, and there exists a ball B in $[0,1)^K$, of diameter less than 1, such that

$$|D[\mathcal{P};B]| \gg_\epsilon N^{\frac{1}{2}-\frac{1}{2K}-\epsilon}.$$

Note that the two estimates agree when $K = 3$.

It can be shown that the exponents in Schmidt's estimates above are essentially sharp. The following two results can be established by using probabilistic techniques.

THEOREM 7. (Beck 1981) *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$ such that for every rectangle B in $[0,1)^2$, of diameter less than 1, we have*

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

THEOREM 8. (Beck 1981) *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$ such that for every disc B in $[0,1)^2$, of diameter less than 1, we have*

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

Naturally, the work of Schmidt raised the question of the connection between tilted rectangles and discs. In the former case, we allow rotation, or orthogonal transformation. In the latter case, the sets are invariant under rotation. If we compare Theorem 3 and Theorem 5, we might be tempted to blame the "discrepancy" in the estimates on rotation. But then we also have Theorem 6, where rotation is not present, or is it?

In arguably the greatest contribution to the subject to date, Beck showed in the mid to late 1980's that the discrepancy arises from rotation and/or the geometry of the boundary curve.

Consider first of all the case when rotation is permitted.

Suppose that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. Suppose further that A is a compact and convex region in $[0,1)^2$. For any real number $\lambda \in [0,1]$, any rotation τ in $\mathbb{R}^2$ and any vector $\mathbf{u} \in [0,1)^2$, consider the similar copy

$$A(\lambda,\tau,\mathbf{u}) = \{\tau(\lambda\mathbf{x}) + \mathbf{u} : \mathbf{x} \in A\}$$

of A, and let

$$Z[\mathcal{P};A(\lambda,\tau,\mathbf{u})] = \#(\mathcal{P} \cap A(\lambda,\tau,\mathbf{u}))$$

denote the number of points of $\mathcal{P}$ that fall into $A(\lambda,\tau,\mathbf{u})$. We are interested in the discrepancy function

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

where μ denotes the usual measure in $\mathbb{R}^2$ restricted to $[0,1)^2$.

Let $\mathcal{T}$ be the group of all rotations in $\mathbb{R}^2$, and let $\mathrm{d}\tau$ be the volume element of the invariant measure on $\mathcal{T}$, normalized such that $\int_{\mathcal{T}} \mathrm{d}\tau = 1$.

THEOREM 9. (Beck 1987) *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A) \geq N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every distribution $\mathcal{P}$ of N points in $[0,1)^2$, we have*

$$\int_0^1 \int_{\mathcal{T}} \int_{[0,1)^2} |D[\mathcal{P};A(\lambda,\tau,\mathbf{u})]|^2 \,\mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda \gg_A N^{\frac{1}{2}}.$$

A simple corollary is the following generalization and improvement of Theorems 5 and 6.

THEOREM 10. (Beck 1987) *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A)\ge N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every distribution $\mathcal{P}$ of N points in $[0,1)^2$, there exists a similar copy B of A in $[0,1)^2$ such that*

$$|D[\mathcal{P};B]|\gg_A N^{\frac{1}{4}}.$$

These results are complemented by the following upper bound results. The first is essentially a generalization of Theorems 7 and 8.

THEOREM 11. (Beck 1981) *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A)\ge N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every natural number $N\ge 2$, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that for every similar copy B of A in $[0,1)^2$, we have*

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

THEOREM 12. (Beck and Chen 1990) *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A)\ge N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every natural number N, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that*

$$\int_0^1\int_{\mathcal{T}}\int_{[0,1)^2}|D[\mathcal{P};A(\lambda,\tau,\mathbf{u})]|^2\,\mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda\ll_A N^{\frac{1}{2}}.$$

We remark that the four results above can be generalized to any dimension $K\ge 2$ without any extra difficulties. Here we require a technical condition that $r(A)\ge N^{-1/K}$, where $r(A)$ denotes the radius of the largest inscribed ball of A. The estimates $N^{\frac{1}{2}}$ in Theorems 9 and 12 should then be replaced by $N^{1-\frac{1}{K}}$, while the estimates $N^{\frac{1}{4}}$ in Theorems 10 and 11 should then be replaced by $N^{\frac{1}{2}-\frac{1}{2K}}$.

The situation is very different when rotation is not allowed.

Suppose that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. Suppose further that A is a compact and convex region in $[0,1)^2$. For any real number $\lambda\in[0,1]$ and any vector $\mathbf{u}\in[0,1)^2$, consider the homothetic copy

$$A(\lambda,\mathbf{u})=\{\lambda\mathbf{x}+\mathbf{u}:\mathbf{x}\in A\}$$

of A, and let

$$Z[\mathcal{P};A(\lambda,\mathbf{u})]=\#(\mathcal{P}\cap A(\lambda,\mathbf{u}))$$

denote the number of points of $\mathcal{P}$ that fall into $A(\lambda,\mathbf{u})$. We are interested in the discrepancy function

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

where μ denotes the usual measure in $\mathbb{R}^2$ restricted to $[0,1)^2$.

We have the following generalization of Theorem 1.

THEOREM 13. (Beck 1988a) *Suppose that A is a compact and convex region in the torus $[0,1)^2$. For every distribution $\mathcal{P}$ of N points in $[0,1)^2$, we have*

$$\int_0^1\int_{[0,1)^2}|D[\mathcal{P};A(\lambda,\mathbf{u})]|^2\,\mathrm{d}\mathbf{u}\mathrm{d}\lambda\gg_A\max\{\log N,\xi_N^2(A)\},$$

where $\xi_N(A)$ depends on the boundary curve ∂A of A.

A simple corollary is the following result.

THEOREM 14. (Beck 1988a) *Suppose that A is a compact and convex region in the torus $[0,1)^2$. For every distribution $\mathcal{P}$ of N points in $[0,1)^2$, there exists a homothetic copy B of A in $[0,1)^2$ such that*

$$|D[\mathcal{P};B]| \gg_A \max\{(\log N)^{\frac{1}{2}}, \xi_N(A)\},$$

where $\xi_N(A)$ depends on the boundary curve ∂A of A.

Roughly speaking, the function $\xi_N(A)$ varies from being a constant, in the case when A is a convex polygon, to being a power of N, in the case when A is a circular disc. In fact, it is some sort of measure of how well A can be approximated by an inscribed polygon.

Here, upper bounds are harder to obtain. We have, for example, the following results.

THEOREM 15. (Beck 1988a) *Suppose that A is a compact and convex region in the torus $[0,1)^2$. For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that for every homothetic copy B of A in $[0,1)^2$, we have*

$$|D[\mathcal{P};B]| \ll_A \max\{\log N, \xi_N^2(A)\}.$$

THEOREM 16. (Beck 1988a) *Suppose that A is a convex polygon in the torus $[0,1)^2$. For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that for every homothetic copy B of A in $[0,1)^2$, we have*

$$|D[\mathcal{P};B]| \ll_{A,\epsilon} (\log N)^{4+\epsilon}.$$

Note that there is a significant gap between Theorems 14 and 15. The following is considered the analogous question to Theorem 2.

Open Problem. Suppose that A is a compact and convex region in the torus $[0,1)^2$. Is it true that for every distribution $\mathcal{P}$ of N points in $[0,1)^2$, there exists a homothetic copy B of A such that

$$|D[\mathcal{P};B]| \gg_A \log N?$$

Halász (unpublished) has established this when A is a square, and Beck and Chen (1989) have established this when A is a triangle. On the other hand, little is known in connection with homothetic copies in higher dimension, although some work on upper bounds have been carried out by Károlyi (1995ab). We have the following difficult problem which includes the previous problem as a special case.

Greater Open Problem. Suppose that $K \geq 2$ is an integer, and that A is a compact and convex region in the torus $[0,1)^K$. Is it true that for every distribution $\mathcal{P}$ of N points in $[0,1)^K$, there exists a homothetic copy B of A such that

$$|D[\mathcal{P};B]| \gg_A (\log N)^{K-1}?$$

Let us return to Theorems 9 and 12, and consider similar copies $A(\lambda,\tau,\mathbf{u})$ of a given compact and convex region A. Suppose now that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^3$, and that A is a compact and convex region in the torus $[0,1)^2$. For any real number $\lambda \in [0,1]$, any rotation τ in $\mathbb{R}^2$, any vector $\mathbf{u} \in [0,1)^2$ and any real number $y \in [0,1]$, we consider the cartesian product

$$A(\lambda,\tau,\mathbf{u}) \times [0,y]$$

of the similar copy of A with an interval $[0,y]$, and let

$$Z[\mathcal{P};A(\lambda,\tau,\mathbf{u}) \times [0,y]] = \#(\mathcal{P} \cap (A(\lambda,\tau,\mathbf{u}) \times [0,y]))$$

denote the number of points of $\mathcal{P}$ that fall into $A(\lambda,\tau,\mathbf{u})\times[0,y]$. We are interested in the discrepancy function

$$D[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times[0,y]]=Z[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times[0,y]]-Ny\mu(A(\lambda,\tau,\mathbf{u})),$$

where μ denotes the usual measure in $\mathbb{R}^2$ restricted to $[0,1)^2$.

We have the following results. The first is essentially a simple deduction from Theorem 9.

THEOREM 17. *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A)\geq N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^3$, we have*

$$\int_0^1\int_0^1\int_{\mathcal{T}}\int_{[0,1)^2}|D[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times[0,y]]|^2\,\mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda\mathrm{d}y\gg_A N^{\frac{1}{2}}.$$

THEOREM 18. (Beck and Chen 1990) *Suppose that A is a compact and convex region in the torus $[0,1)^2$ such that $r(A)\geq N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed disc of A. For every natural number N, there exists a distribution $\mathcal{P}$ of N points in the torus $[0,1)^3$ such that*

$$\int_0^1\int_0^1\int_{\mathcal{T}}\int_{[0,1)^2}|D[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times[0,y]]|^2\,\mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda\mathrm{d}y\ll_A N^{\frac{1}{2}}.$$

To understand the context of these two results, consider the special case when A is a circular disc. Then we are studying irregularities of point distribution with respect to circular cylinders. Comparing the results here to Theorems 1, 4, 9 and 12, we observe that the estimates have not increased in order of magnitude as functions of N. This is an illusrtarion that large discrepancy results from "curved" edges like the circular sides of the cylinders, whereas small discrepancy results from "flat" edges like the two flat ends of the cylinders. We should now relate this observation to Theorems 13 and 14 and the role played by the function $\xi_N(A)$.

To demonstrate further the role played by rotation, as well as the interplay of ideas from different settings, we consider the following problem suggested by Roth.

Suppose that $\mathcal{P}$ is a distribution of N points in the closed disc U of unit area and centred at the origin $\mathbf{0}$. For every measurable set B in $\mathbb{R}^2$, let

$$Z[\mathcal{P};B]=\#(\mathcal{P}\cap B)$$

denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

$$D^*[\mathcal{P};B]=Z[\mathcal{P};B]-N\mu(B\cap U),$$

where μ denotes the usual area measure in $\mathbb{R}^2$.

For every real number $r\geq 0$ and every angle θ satisfying $0\leq\theta\leq 2\pi$, let $S(r,\theta)$ denote the closed halfplane

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

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

Using Fourier transform techniques, Beck (1983) was the first to establish any lower bound. More recently, using ideas from integral geometry, Alexander established the following result.

THEOREM 19. (Alexander 1990) *For every distribution $\mathcal{P}$ of N points in the closed disc U, we have*

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

In fact, Beck's lower bound is only marginally weaker, with $N^{\frac{1}{2}}(\log N)^{-7}$ in place of $N^{\frac{1}{2}}$. On the other hand, a simple consequence of Theorem 19 is the estimate

$$|D^*[\mathcal{P}; S(r,\theta)]| \gg N^{\frac{1}{4}},$$

valid for some halfplane $S(r,\theta)$. Remarkably, this is best possible, in view of the following spectacular result.

THEOREM 20. (Matoušek 1995) *For every natural number N, there exists a distribution $\mathcal{P}$ of N points in the closed disc U such that for every halfplane $S(r,\theta)$, we have*

$$|D^*[\mathcal{P}; S(r,\theta)]| \ll N^{\frac{1}{4}}.$$

We remark that Beck's probabilistic technique can be used to establish the upper bound

$$|D^*[\mathcal{P}; S(r,\theta)]| \ll N^{\frac{1}{4}}(\log N)^{\frac{1}{2}},$$

valid for every halfplane $S(r,\theta)$. Furthermore, the method of Beck and Chen used to study Theorem 12 can be adapted to obtain the upper bound

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

There is therefore a close relationship between this problem and the problem of discrepancy with respect to similar copies of a compact and convex set, where rotation is allowed. Indeed, the following surprising result provides a further link.

THEOREM 21. (Beck and Chen 1993a) *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the closed disc U such that for every halfplane $S(r,\theta)$, we have*

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

Compare this to the lower estimate given in Theorem 19.

Let us return again to the problem of discrepancy with respect to similar copies of a compact and convex set, where rotation is allowed. Note that a convex polygon is the intersection of a finite number of halfplanes. One may take the alternative viewpoint that a halfplane is a convex monogon!

Corresponding to Theorems 9 and 12, we can establish the following result.

THEOREM 22. (Beck and Chen 1993b) *Suppose that A is a convex polygon in the torus $[0,1)^2$. For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that*

$$\int_0^1\int_{\mathcal{T}}\int_{[0,1)^2} |D[\mathcal{P}; A(\lambda,\tau,\mathbf{u})]|^2 \, \mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda \ll_A (\log N)^2.$$

Suppose now that rotation is not allowed. We then have the following analogue of Theorem 4.

THEOREM 23. (Beck and Chen 1997) *Suppose that A is a convex polygon in the torus $[0,1)^2$. For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that*

$$\int_0^1\int_{[0,1)^2} |D[\mathcal{P}; A(\lambda,\mathbf{u})]|^2 \, \mathrm{d}\mathbf{u}\mathrm{d}\lambda \ll_A \log N.$$

In the present lectures, it is not possible to prove all these results. We shall therefore only concentrate on a few to illustrate some of the ideas in the area.

In Section 7, we shall discuss Beck's probabilistic technique and indicate how it may be used to establish a special case of Theorem 11. Here we need some tools from probability theory; in particular, we need a classical large-deviation type inequality which we shall prove. In Section 8, we study more upper bounds and establish a weaker version of Theorem 16 by using a combinatorial and geometric approach.

We then turn to lower bounds. In Section 9, we shall illustrate Beck's Fourier transform technique by proving a weaker version of Theorem 10 in the special case when A is a square. In Section 10, we turn our attention to Alexander's integral geometric technique and establish a slightly different version of Theorem 19.

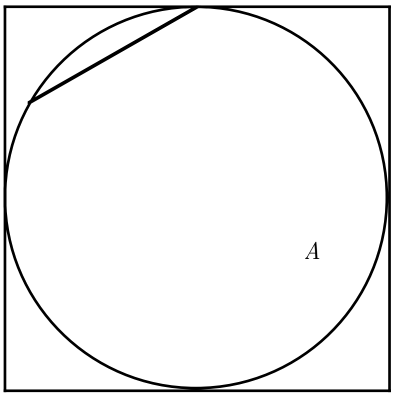
We conclude these lectures by returning to upper bounds in Section 11. We shall discuss the use of lattices in the study of Theorems 21, 22 and 23. In particular, we shall show that the proof of Theorem 23 requires nothing that Davenport and Roth do not know!

We conclude this section by describing Schmidt's ingenious proof of the following amazing result which is essentially best possible, as shown by Beck (1988b).

THEOREM 24. (Schmidt 1975) *For every distribution $\mathcal{P}$ of N points in the square $[0,1]^2$, there exists a convex set B in $[0,1]^2$ such that*

$$|D[\mathcal{P};B]| \gg N^{\frac{1}{3}}.$$

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



Any such disc segment S may contain no point of $\mathcal{P}$ or contain at least one point of $\mathcal{P}$. Corresponding to these two cases, it is easy to see that we have respectively

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

Simple geometric consideration will show that there are $\ge cN^{\frac{1}{3}}$ 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

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

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

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

It follows that

$$\max\{|D[\mathcal{P}; A \setminus (S_1 \cup \ldots \cup S_k)]|, |D[\mathcal{P}; A \setminus (T_1 \cup \ldots \cup T_m)]| \geq \frac{cN^{\frac{1}{3}}}{4}.$$

The result follows. ♣

Perhaps this is the way that Wolfgang Schmidt discovered the proof. The reader will need two lemmas that are axioms to Schmidt.

LEMMA 6A. *Wolfgang Schmidt loves chocolates.*

PROOF. This is part of the mathematical folklore. ♣

LEMMA 6B. *Pat Schmidt makes lovely chocolate cakes.*

PROOF. 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, he saw the cake. In view of Lemmas 6A and 6B, 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 $\gg N^{1/3}$ times without destroying the convexity of (what remained of) the cake.

Naturally, Lemma 6A 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 slightly 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 $\gg N^{1/3}$.

7. Beck's Probabilistic Method

In this section, we give a simple proof of a special case of Theorem 11 when $N = M^2$ with $M \geq 2$; in other words, when the number of points N is a perfect square greater than 1. Instead of considering a set of N points in the torus $[0,1)^2$, we shall consider a set of N points in the square $[0,M)^2$, treated as a torus. In particular, we shall be concerned with sets $\mathcal{Q}$ of $N = M^2$ points in $[0,M)^2$ such that every square of the form

$$S(\mathbf{l}) = [\ell_1, \ell_1 + 1) \times [\ell_2, \ell_2 + 1),$$

where $\mathbf{l} = (\ell_1, \ell_2) \in \mathbb{Z}^2 \cap [0,M)^2$, contains precisely one point of $\mathcal{Q}$.

Let A be a compact and convex set in $[0,M)^2$. Observe that the technical condition concerning the radius of the largest inscribed disc of A can now be described by $r(A) \geq 1$. We shall show that there exists sets $\mathcal{Q}$ with the above property and such that for any $\lambda \in [0,1]$, any rotation $\tau \in \mathcal{T}$ and any vector $\mathbf{u} \in \mathbb{R}^2$, we have

$$|\#(\mathcal{Q} \cap A(\lambda,\tau,\mathbf{u})) - \mu(A(\lambda,\tau,\mathbf{u}))| \ll (\sigma(\partial A))^{\frac{1}{2}} (\log M)^{\frac{1}{2}}, \tag{7.1}$$

where ∂A denotes the boundary curve of A, and where σ denotes the usual measure in $\mathbb{R}$. Theorem 11 in this special case now follows on noting that $\sigma(\partial A) \ll N^{\frac{1}{2}}$ and on rescaling.

The first important idea is to approximate every similar copy of A in the collection in question by the members of a finite subcollection of similar copies of A. The collection of all similar copies of A in question is given by

$$\mathcal{G} = \{A(\lambda,\tau,\mathbf{u}) : 0 \leq \lambda \leq 1,\ \tau \in \mathcal{T},\ \mathbf{u} \in \mathbb{R}^2\}.$$

We now slightly extend the restrictions on λ to obtain the bigger collection

$$\mathcal{G}_0 = \{A(\lambda,\tau,\mathbf{u}) : 0 \leq \lambda \leq 1.1,\ \tau \in \mathcal{T},\ \mathbf{u} \in \mathbb{R}^2\}.$$

Geometric consideration shows that there exists a finite subset $\mathcal{G}^*$ of $\mathcal{G}_0$ such that

$$\#\mathcal{G}^* \leq M^c, \tag{7.2}$$

where c is a positive absolute constant, and that for any $B \in \mathcal{G}$, there exist $B^-, B^+ \in \mathcal{G}^*$ such that $B^- \subseteq B \subseteq B^+$ and $\mu(B^+ \setminus B^-) \leq 1$. We then examine the set $\mathcal{G}^*$ more closely.

The second important idea is classical probability theory. Note that

$$[0,M)^2 = \sum_{\mathbf{l} \in \mathbb{Z}^2 \cap [0,M)^2} S(\mathbf{l}).$$

For each $\mathbf{l} \in \mathbb{Z}^2 \cap [0,M)^2$, let $\mathbf{q_l}$ be a random point in $S(\mathbf{l})$, uniformly distributed within $S(\mathbf{l})$. Assume further that the random variables $\mathbf{q_l}$, where $\mathbf{l} \in \mathbb{Z}^2 \cap [0,M)^2$, are independent of each other. Our aim is to show that the random set

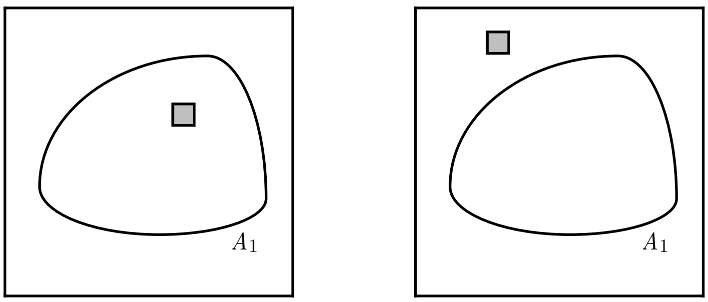
$$\mathcal{Q} = \{\mathbf{q_l} : \mathbf{l} \in \mathbb{Z}^2 \cap [0,M)^2\}$$

satisfies the inequality

$$|\#(\mathcal{Q} \cap A_1) - \mu(A_1)| \ll (\sigma(\partial A))^{\frac{1}{2}} (\log M)^{\frac{1}{2}} \tag{7.3}$$

simultaneously for all $A_1 \in \mathcal{G}^*$ with probability greater than half.

Let $A_1 \in \mathcal{G}^*$. Clearly any square $S(\mathbf{l})$ satisfying $S(\mathbf{l}) \subseteq A_1$ or $S(\mathbf{l}) \cap A_1 = \emptyset$ contributes nothing to the discrepancy of the set A_1.



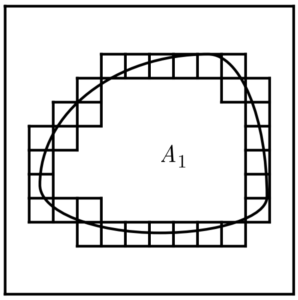
We therefore consider the set

$$\mathcal{L}(A_1) = \{\mathbf{l} \in \mathbb{Z}^2 \cap [0,M)^2 : S(\mathbf{l}) \cap A_1 \neq \emptyset \text{ and } S(\mathbf{l}) \not\subseteq A_1\}.$$

It is easy to see that

$$\#\mathcal{L}(A_1) \ll \sigma(\partial A_1) \ll \sigma(\partial A), \tag{7.4}$$

as shown in the picture below.



For every $\mathbf{l} \in \mathcal{L}(A_1)$, we define the random variable

$$\xi_{\mathbf{l}} = \begin{cases} 1 & \text{if } \mathbf{q_l} \in A_1, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\sum_{\mathbf{q_l} \in A_1} 1 - \mu(A_1) = \sum_{\mathbf{l} \in \mathcal{L}(A_1)} \xi_{\mathbf{l}} - \sum_{\mathbf{l} \in \mathcal{L}(A_1)} \mu(S(\mathbf{l}) \cap A_1) = \sum_{\mathbf{l} \in \mathcal{L}(A_1)} (\xi_{\mathbf{l}} - \mathbb{E}\xi_{\mathbf{l}}).$$

Note now that the random variables $\xi_{\mathbf{l}}$, where $\mathbf{l} \in \mathcal{L}(A_1)$, are independent of each other. We can therefore apply the classical large-deviation type inequality due to Bernstein and Chernoff.

LEMMA 7A. *Suppose that $\xi_1, \ldots, \xi_m$ are independent random variables satisfying $|\xi_i| \leq 1$ for every $i = 1, \ldots, m$. Suppose further that*

$$\beta = \sum_{i=1}^{m} \mathbb{E}(\xi_i - \mathbb{E}\xi_i)^2.$$

Then

$$\mathrm{Prob}\left(\left| \sum_{i=1}^{m} (\xi_i - \mathbb{E}\xi_i) \right| \geq \gamma \right) \leq \begin{cases} 2\mathrm{e}^{-\gamma/4} & \text{if } \gamma \geq \beta, \\ 2\mathrm{e}^{-\gamma^2/4\beta} & \text{if } \gamma \leq \beta. \end{cases}$$

In view of (7.4), we now take

$$\beta_1 = \sum_{\mathbf{l} \in \mathcal{L}(A_1)} \mathbb{E}(\xi_{\mathbf{l}} - \mathbb{E}\xi_{\mathbf{l}})^2 \leq \#\mathcal{L}(A_1) \leq c_1 \sigma(\partial A),$$

where c_1 is a positive absolute constant. We also let

$$\gamma_1 = c_2 (\sigma(\partial A))^{\frac{1}{2}} (\log M)^{\frac{1}{2}},$$

where c_2 is a sufficiently large positive constant, to be determined later. Elementary calculation now gives

$$-\frac{\gamma_1^2}{4\beta_1} \leq \log M^{-cc_3},$$

where c is given in (7.2) and $c_3 = c_2^2/4cc_1$. Then

$$4\mathrm{e}^{-\gamma_1^2/4\beta_1} \leq 4M^{-cc_3} \leq 4M^{-c(c_3-1)} M^{-c} \leq 2^{2-c(c_3-1)} M^{-c} \leq M^{-c},$$

provided that $c(c_3 - 1) \geq 2$; in other words, provided that

$$\frac{c_2^2}{4c_1} - c \geq 2. \tag{7.5}$$

On the other hand, we may assume that $\sigma(\partial A) \geq \log M$, for the result is trivial otherwise. Then $\gamma_1 \geq c_2 \log M$, and elementary calculation now gives

$$-\frac{\gamma_1}{4} \leq \log M^{-cc_4},$$

where c is given in (7.2) and $c_4 = c_2/4c$. Then

$$4\mathrm{e}^{-\gamma_1/4} \leq 4M^{-cc_4} \leq 4M^{-c(c_4-1)}M^{-c} \leq 2^{2-c(c_4-1)}M^{-c} \leq M^{-c},$$

provided that $c(c_4-1) \geq 2$; in other words, provided that

$$\frac{c_2}{4} - c \geq 2. \tag{7.6}$$

It follows that if c_2 is chosen sufficiently large to satisfy (7.5) and (7.6), then

$$\tfrac{1}{2}M^{-c} \geq \begin{cases} 2\mathrm{e}^{-\gamma_1/4} & \text{if } \gamma_1 \geq \beta_1, \\ 2\mathrm{e}^{-\gamma_1^2/4\beta_1} & \text{if } \gamma_1 \leq \beta_1. \end{cases}$$

Note that the conditions on γ_1 and β_1 are in fact irrelevant here. It now follows from Lemma 7A and (7.2) that

$$\mathrm{Prob}\left(\left|\sum_{\mathbf{l}\in\mathcal{L}(A_1)} (\xi_{\mathbf{l}} - \mathbb{E}\xi_{\mathbf{l}})\right| \geq \gamma_1\right) \leq \frac{1}{2}(\#\mathcal{G}^*)^{-1}.$$

Hence

$$\mathrm{Prob}\left(\left|\sum_{\mathbf{q}_1\in A_1} 1 - \mu(A_1)\right| \geq \gamma_1 \text{ for some } A_1 \in \mathcal{G}^*\right) \leq \frac{1}{2}.$$

This proves (7.3), and completes the proof of Theorem 11 in the special case.

It remains to prove Lemma 7A.

Proof of Lemma 7A. We assume, without loss of generality, that $\mathbb{E}(\xi_i) = 0$ for every $i = 1, \ldots, m$. Let

$$X = \sum_{i=1}^m \xi_i.$$

Clearly $\mathrm{Prob}(X \geq \gamma) = \mathrm{Prob}(\mathrm{e}^{yX} \geq \mathrm{e}^{y\gamma})$, where the parameter $y \in (0,1]$ will be fixed later. On the other hand, we have $\mathrm{e}^{yX} \geq \mathrm{e}^{y\gamma}\chi_\gamma(X)$, where

$$\chi_\gamma(X) = \begin{cases} 1 & \text{if } X \geq \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

Denoting the probability measure by ν, we have

$$\mathrm{Prob}(\mathrm{e}^{yX} \geq \mathrm{e}^{y\gamma}) = \int \chi_\gamma(X)\,\mathrm{d}\nu(X) \leq \mathrm{e}^{-y\gamma}\int \mathrm{e}^{yX}\,\mathrm{d}\nu(X) = \mathrm{e}^{-y\gamma}\mathbb{E}(\mathrm{e}^{yX}).$$

It follows that

$$\mathrm{Prob}(X \geq \gamma) \leq \mathrm{e}^{-y\gamma}\mathbb{E}(\mathrm{e}^{yX}). \tag{7.7}$$

Since X is the sum of independent random variables, we have

$$\mathbb{E}(\mathrm{e}^{yX}) = \prod_{i=1}^m \mathbb{E}(\mathrm{e}^{y\xi_i}).$$

We shall give an upper bound on each $\mathbb{E}(\mathrm{e}^{y\xi_i})$. Using the formula of the exponential series, we obtain after some easy calculation that

$$\mathbb{E}(\mathrm{e}^{y\xi_i}) = \sum_{n=0}^{\infty} \frac{y^n \mathbb{E}\xi_i^n}{n!} \leq 1 + \frac{y^2\mathbb{E}\xi_i^2}{2} + \frac{y^3\mathbb{E}\xi_i^2}{6-2y}.$$

Substituting this into (7.7), we obtain

$$\mathrm{Prob}(X \geq \gamma) \leq \exp\left(\frac{y^2\beta}{2}\left(1+\frac{y}{3-y}\right) - y\gamma\right). \tag{7.8}$$

We distinguish two cases. If $\gamma \geq \beta$, we let $y = 1$. Then it follows from (7.8) that

$$\mathrm{Prob}(X \geq \gamma) \leq \mathrm{e}^{-\gamma/4}.$$

If $\gamma \leq \beta$, we let $y = \gamma/\beta$. Then it follows from (7.8) that

$$\mathrm{Prob}(X \geq \gamma) \leq \mathrm{e}^{-\gamma^2/4\beta}.$$

Repeating the same calculation for $\mathrm{Prob}(X \leq -\gamma)$, we obtain the desired upper bounds. ♣

8. A Combinatorial and Geometric Approach

In this section, we shall establish a weaker version of Theorem 16. Suppose that A is a convex polygon in the torus $[0,1)^2$. For every natural number $N \geq 2$, we shall show that there exists a distribution $\mathcal{P}$ of N points in $[0,1)^2$ such that for every homothetic copy B of A in $[0,1)^2$, we have

$$|D[\mathcal{P};B]| \ll_{A,\epsilon} (\log N)^{5+\epsilon}.$$

In fact, we shall establish a slightly stronger result which will imply the above.

Let $\ell \geq 2$, and let $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_\ell)$ satisfy $0 \leq \theta_1 < \ldots < \theta_\ell < \pi$. For each $i = 1,\ldots,\ell$, let $\mathbf{e}_i = (\cos\theta_i, \sin\theta_i)$, and denote by $\mathrm{POL}^\infty(\boldsymbol{\theta})$ the family of convex polygons $A \subseteq \mathbb{R}^2$ such that each side of A is parallel to one of the given directions $\mathbf{e}_i$.

THEOREM 16W. *For every $\epsilon > 0$, there exists an infinite discrete set $\mathcal{Q} \subseteq \mathbb{R}^2$ such that for every $A \in \mathrm{POL}^\infty(\boldsymbol{\theta})$ with $d(A) \geq 2$, we have*

$$|\#(\mathcal{Q}\cap A) - \mu(A)| \ll_{\ell,\epsilon} (\log d(A))^{5+\epsilon},$$

where $d(A)$ denotes the diameter of A.

The proof of Theorem 16W is based on a combination of combinatorial and geometric arguments. The combinatorial part is summarized by the following integer-making lemma.

LEMMA 8A. *Suppose that $X = \{x_1,\ldots,x_p\}$ is a finite set. For $i = 1,2,\ldots$, let*

$$\mathcal{Y}^{(i)} = Y_1^{(i)} \cup Y_2^{(i)} \cup \ldots$$

be a partition of X; in other words,

$$X = \bigcup_{j\geq 1} Y_j^{(i)}$$

is a union of mutually disjoint sets $Y_j^{(i)}$*. For every* $k=1,\ldots,p$*, let* $\alpha_k\in[0,1]$*. Then for every* $\epsilon>0$*, there exists a positive constant* $c(\epsilon)$*, depending at most on* ϵ*, and integers* $a_1,\ldots,a_p\in\{0,1\}$ *such that*

$$\left|\sum_{x_k\in Y_j^{(i)}}(a_k-\alpha_k)\right|<c(\epsilon)i^{1+\epsilon} \tag{8.1}$$

for every $i\geq 1$ *and* $j\geq 1$*.*

Proof. The construction of the integers a_k is based on the well-known result in linear algebra that a system of homogeneous linear equations with more variables than equations has a non-trivial solution. Let $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_p)\in[0,1]^p$. We shall construct inductively a sequence

$$\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_\nu=(\alpha_{1,\nu},\ldots,\alpha_{p,\nu}),\ldots \tag{8.2}$$

of vectors in $[0,1]^p$ with the following properties: For every $\nu\in\mathbb{N}\cup\{0\}$, let

$$X_\nu=\{x_k\in X:\alpha_{k,\nu}\not\in\{0,1\}\}.$$

Then we need

$$X_{\nu+1}\subsetneqq X_\nu, \tag{8.3}$$

$$\alpha_{k,\nu}\in\{0,1\}\Rightarrow\alpha_{k,\nu}=\alpha_{k,\nu+1}, \tag{8.4}$$

and

$$\sum_{x_k\in Y_j^{(i)}}\alpha_{k,\nu}=\sum_{x_k\in Y_j^{(i)}}\alpha_{k,\nu+1} \tag{8.5}$$

for all i and j with $\#(Y_j^{(i)}\cap X_\nu)\geq c(\epsilon)i^{1+\epsilon}$. Note that conditions (8.3) and (8.4) ensure that every term in the sequence (8.2) has more integer entries than the previous term, whereas condition (8.5) gives some control to the change at each step.

Let $\boldsymbol{\alpha}_0=\boldsymbol{\alpha}$. Suppose that $\boldsymbol{\alpha}_\nu$ has been defined and X_ν is non-empty. Let

$$\mathcal{Z}_\nu=\{Y_j^{(i)}:i,j\geq 1\text{ and }\#(Y_j^{(i)}\cap X_\nu)\geq c(\epsilon)i^{1+\epsilon}\}.$$

Since $Y_j^{(i)}\cap Y_k^{(i)}=\emptyset$ whenever $j\neq k$, it follows that

$$\#\mathcal{Z}_\nu=\sum_{i=1}^\infty\#\{j:\#(Y_j^{(i)}\cap X_\nu)\geq c(\epsilon)i^{1+\epsilon}\}\leq\sum_{i=1}^\infty\frac{\#X_\nu}{c(\epsilon)i^{1+\epsilon}}<\#X_\nu \tag{8.6}$$

if we choose

$$c(\epsilon)=2\sum_{i=1}^\infty\frac{1}{i^{1+\epsilon}}<\infty.$$

For $k=1,\ldots,p$, let y_k be a real variable, and consider the system of linear equations

$$\sum_{x_k\in Y_j^{(i)}\cap X_\nu}y_k=0,\quad Y_j^{(i)}\in\mathcal{Z}_\nu,$$

together with $y_k=0$ whenever $x_k\in X\setminus X_\nu$. In view of (8.6), this system has more variables than equations, and so has a non-trivial solution $\mathbf{y}=(y_1,\ldots,y_p)$. Suppose that t_0 is the largest positive real value for which the inequalities

$$0\leq\alpha_{k,\nu}+t_0y_k\leq 1,\quad x_k\in X_\nu,$$

hold. For $k=1,\ldots,p$, we now let

$$\alpha_{k,\nu+1}=\alpha_{k,\nu}+t_0y_k.$$

Then (8.3) clearly holds, in view of the maximality of t_0. On the other hand, (8.4) follows on noting that if $\alpha_{k,\nu} \in \{0,1\}$, then $x_k \in X \setminus X_\nu$ and so $y_k = 0$. Furthermore, (8.5) follows on noting that if $\#(Y_j^{(i)} \cap X_\nu) \geq c(\epsilon) i^{1+\epsilon}$, then $Y_j^{(i)} \in \mathcal{Z}_\nu$ and so

$$\sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu+1} = \sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu} + t_0 \sum_{x_k \in Y_j^{(i)}} y_k = \sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu} + t_0 \sum_{x_k \in Y_j^{(i)} \cap X_\nu} y_k = \sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu}.$$

In view of (8.3), the sequence (8.2) will remain constant after a finite number of steps, say s steps. Then $X_s = \emptyset$ and the vector $\boldsymbol{\alpha}_s$ has coordinates 0 and 1 only. For every $k = 1, \ldots, p$, we now let $a_k = \alpha_{k,s}$. For any $Y_j^{(i)}$, let t be the smallest integer value for which $\#(Y_j^{(i)} \cap X_t) < c(\epsilon) i^{1+\epsilon}$. Then

$$\begin{aligned}
\sum_{x_k \in Y_j^{(i)}} (a_k - \alpha_k) &= \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,s} - \alpha_{k,0}) = \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,s} - \alpha_{k,t}) + \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,t} - \alpha_{k,0}) \\
&= \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,s} - \alpha_{k,t}) + \sum_{\nu=0}^{t-1} \left(\sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu+1} - \sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu} \right) \\
&= \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,s} - \alpha_{k,t}),
\end{aligned}$$

in view of (8.5). Note now that in view of (8.4), we have

$$\left| \sum_{x_k \in Y_j^{(i)}} (a_k - \alpha_k) \right| = \left| \sum_{x_k \in Y_j^{(i)}} (\alpha_{k,s} - \alpha_{k,t}) \right| = \left| \sum_{x_k \in Y_j^{(i)} \cap X_t} (\alpha_{k,s} - \alpha_{k,t}) \right| \leq \#(Y_j^{(i)} \cap X_t) < c(\epsilon) i^{1+\epsilon}$$

as required. ♣

The geometric part of the argument is rather lengthy and involved. We shall consider the family $\mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$ of convex polygons $A \subseteq \mathbb{R}^2$ such that each side of A is parallel to one of the given directions $\mathbf{e}_i$ or parallel to one of the coordinate axes x_1 or x_2. Our ultimate aim is to approximate the characteristic function of an arbitrary polygon in $\mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$ by linear combinations of those of some special geometric objects. We shall therefore need to define these special objects first.

Definition. Suppose that $\mathbf{n} = (n_1, n_2) \in \mathbb{Z}^2$. By a special rectangle of order $\mathbf{n}$, we mean a rectangle of the form

$$[m_1 2^{n_1}, (m_1 + 1) 2^{n_1}] \times [m_2 2^{n_2}, (m_2 + 1) 2^{n_2}], \tag{8.7}$$

where $\mathbf{m} = (m_1, m_2) \in \mathbb{Z}^2$. We denote by $\mathrm{SR}(\mathbf{n})$ the family of all special rectangles of order $\mathbf{n}$.

Definition. Suppose that $1 \leq i \leq \ell$. By a triangle of type i, we mean a triangle with sides parallel to x_1, x_2 and $\mathbf{e}_i$.

Suppose that Δ_i is a triangle of type i, where $1 \leq i \leq \ell$. Suppose further that $t_i^{(1)}$ and $t_i^{(2)}$ denote respectively the lengths of the sides of Δ_i parallel to x_1 and x_2. Let

$$\lambda_i = \frac{t_i^{(1)}}{t_i^{(2)}},$$

and note that the value of λ_i is independent of the choice of the triangle Δ_i. Also, for $i = 1, \ldots, \ell$, write

$$\delta_i = \begin{cases} -1 & \text{if } \theta_i < \pi/2, \\ 1 & \text{if } \theta_i > \pi/2. \end{cases}$$

Naturally, we may assume without loss of generality that $\theta_i \neq \pi/2$ for any $i = 1, \ldots, \ell$.

For any $i = 1, \ldots, \ell$ and any $n \in \mathbb{Z}$, let $\Lambda(i,n)$ denote the rectangular lattice generated by $(2^n\lambda_i^{1/2}, 0)$ and $(0, 2^n\lambda_i^{-1/2})$; in other words, the lattice of points

$$\mathbf{u}(i,n,\mathbf{m}) = (m_1 2^n\lambda_i^{1/2}, m_2 2^n\lambda_i^{-1/2}), \quad \mathbf{m} = (m_1, m_2) \in \mathbb{Z}^2.$$

For convenience of notation, let $\mathbf{E}_1 = (1,0)$ and $\mathbf{E}_2 = (0,1)$.

Definition. Suppose that $1 \leq i \leq \ell$ and $n \in \mathbb{Z}$. By a special triangle of type i and order n, we mean a triangle with vertices

$$\mathbf{u}(i,n,\mathbf{m}), \qquad \mathbf{u}(i,n,\mathbf{m}+\delta_i\mathbf{E}_1), \qquad \mathbf{u}(i,n,\mathbf{m}+\mathbf{E}_2),$$

or a triangle with vertices

$$\mathbf{u}(i,n,\mathbf{m}), \qquad \mathbf{u}(i,n,\mathbf{m}-\delta_i\mathbf{E}_1), \qquad \mathbf{u}(i,n,\mathbf{m}-\mathbf{E}_2),$$

where $\mathbf{m} \in \mathbb{Z}^2$. We denote by $\mathrm{ST}(i,n)$ the family of all special triangles of type i and order n.

Definition. Suppose that $1 \leq i \leq \ell$ and $j = 1,2$. By a parallelogram of type (i,j), we mean a parallelogram with sides parallel to $\mathbf{e}_i$ and x_j.

For $i = 1, \ldots, \ell$, let ψ_i^* denote the linear transformation of determinant 1 represented in matrix notation by

$$\psi_i^* \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda_i^{1/2} & -\delta_i\lambda_i^{1/2} \\ 0 & \lambda_i^{-1/2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It is not difficult to see that $P_i^* = \{\psi_i^*(\mathbf{x}) : \mathbf{x} \in [0,1]^2\}$ is a parallelogram with vertices

$$\mathbf{u}(i,0,\mathbf{0}), \qquad \mathbf{u}(i,0,\mathbf{E}_1), \qquad \mathbf{u}(i,0,-\delta_i\mathbf{E}_1+\mathbf{E}_2), \qquad \mathbf{u}(i,0,(1-\delta_i)\mathbf{E}_1+\mathbf{E}_2).$$

Definition. Suppose that $1 \leq i \leq \ell$ and $\mathbf{n} \in \mathbb{Z}^2$. By a special parallelogram of type $(i,1)$ and order $\mathbf{n}$, we mean the image under ψ_i^* of a special rectangle of the form (8.7), where $\mathbf{m} = (m_1, m_2) \in \mathbb{Z}^2$. We denote by $\mathrm{SP}(i,1,\mathbf{n})$ the family of all special parallelograms of type $(i,1)$ and order $\mathbf{n}$.

Similarly, for $i = 1, \ldots, \ell$, let ψ_i^{**} denote the linear transformation of determinant 1 represented in matrix notation by

$$\psi_i^{**} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda_i^{1/2} & 0 \\ -\delta_i\lambda_i^{-1/2} & \lambda_i^{-1/2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Again, it is not difficult to see that $P_i^{**} = \{\psi_i^{**}(\mathbf{x}) : \mathbf{x} \in [0,1]^2\}$ is a parallelogram with vertices

$$\mathbf{u}(i,0,\mathbf{0}), \qquad \mathbf{u}(i,0,\mathbf{E}_2), \qquad \mathbf{u}(i,0,\mathbf{E}_1-\delta_i\mathbf{E}_2), \qquad \mathbf{u}(i,0,\mathbf{E}_1+(1-\delta_i)\mathbf{E}_2).$$

Definition. Suppose that $1 \leq i \leq \ell$ and $\mathbf{n} \in \mathbb{Z}^2$. By a special parallelogram of type $(i,2)$ and order $\mathbf{n}$, we mean the image under ψ_i^{**} of a special rectangle of the form (8.7), where $\mathbf{m} = (m_1, m_2) \in \mathbb{Z}^2$. We denote by $\mathrm{SP}(i,2,\mathbf{n})$ the family of all special parallelograms of type $(i,2)$ and order $\mathbf{n}$.

We shall also frequently refer to special rectangles as special parallelograms of type $(0,0)$. Also, for any set $B \subseteq \mathbb{R}^2$, let χ_B denote the characteristic function of B. We shall prove the following result.

LEMMA 8B. *Suppose that $A \in \mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$. Then there exist special triangles $T_1', \ldots, T_m'$ and $T_1'', \ldots, T_M''$ of types $\in \{1, \ldots, \ell\}$, and special parallelograms $P_1', \ldots, P_n'$ and $P_1'', \ldots, P_N''$ of types $\in \{(0,0)\} \cup \{(i,j) : i = 1, \ldots, \ell \text{ and } j = 1,2\}$, together with signs*

$$\epsilon_1', \ldots, \epsilon_m', \epsilon_1'', \ldots, \epsilon_M'', \delta_1', \ldots, \delta_n', \delta_1'', \ldots, \delta_N'' \in \{\pm 1\}$$

such that

$$\sum_{\nu=1}^{m}\epsilon'_\nu\chi_{T'_\nu}+\sum_{\beta=1}^{n}\delta'_\beta\chi_{P'_\beta}\le\chi_A\le\sum_{\nu=1}^{M}\epsilon''_\nu\chi_{T''_\nu}+\sum_{\beta=1}^{N}\delta''_\beta\chi_{P''_\beta} \tag{8.8}$$

and

$$\sum_{\nu=1}^{M}\epsilon''_\nu\mu(T''_\nu)+\sum_{\beta=1}^{N}\delta''_\beta\mu(P''_\beta)-\sum_{\nu=1}^{m}\epsilon'_\nu\mu(T'_\nu)-\sum_{\beta=1}^{n}\delta'_\beta\mu(P'_\beta)\ll\ell\log(d(A)+2).$$

Furthermore, these special objects can be chosen in such a way that

$$\max_{\nu,\beta}\{d(T'_\nu),d(P'_\beta),d(T''_\nu),d(P''_\beta)\}\ll d(A)$$

and the numbers m,M,n,N *satisfy*

$$\max\{m,M\}\ll\ell\log(d(A)+2)\qquad\text{and}\qquad\max\{n,N\}\ll\ell(\log(d(A)+2))^3.$$

Strictly speaking, the inequalities (8.8) are only accurate if one adopts a suitable convention concerning the boundaries of all the sets in question, as we adopt throughout the convention that two sets are disjoint if their intersection has measure zero.

The first step in the proof of Lemma 8B is to reduce the problem to one of investigating rectangles and triangles.

LEMMA 8C. *Every* $A\in\mathrm{POL}^\infty(\boldsymbol{\theta};x_1,x_2)$ *is representable in the form*

$$A=(P_1\cup P_2\cup P_3\cup P_4)\setminus\left(\left(\bigcup_{\beta=1}^{q_1}R_\beta\right)\cup\left(\bigcup_{\nu=1}^{q_2}T_\nu\right)\right),$$

where
(a) $P_1,\ldots,P_4$ *are special rectangles of the same order, and* $d(P_\alpha)<3d(A)$ *for every* $\alpha=1,\ldots,4$*;*
(b) R_β *is an aligned rectangle and* $d(R_\beta)<5d(A)$ *for every* $\beta=1,\ldots,q_1$*;*
(c) T_ν *is a triangle of type* $\in\{1,\ldots,l\}$ *and* $d(T_\nu)\le d(A)$ *for every* $\nu=1,\ldots,q_2$*;*
(d) $q_1\le4\ell+8$ *and* $q_2\le4\ell+6$*; and*
(e) $R_1,\ldots,R_{q_1}$ *and* $T_1,\ldots,T_{q_2}$ *are pairwise disjoint.*

PROOF. For $j=1,2$, let $A^{(j)}$ denote the projection of A onto the x_j-axis, and let $L^{(j)}$ denote the length of the interval $A^{(j)}$. Suppose that $n_j\in\mathbb{Z}$ satisfies $2^{n_j-1}<L^{(j)}\le2^{n_j}$. Then the interval $A^{(j)}$ is contained in the union of at most two intervals of the type $[m_j2^{n_j},(m_j+1)2^{n_j}]$, where $m_j\in\mathbb{Z}$. Let $\mathbf{n}=(n_1,n_2)$. Then A is contained in the union of at most four special rectangles of order $\mathbf{n}$. Denote these rectangles by P_1,P_2,P_3,P_4, with the convention that they may not be distinct, and note that

$$d(P_\alpha)=(2^{2n_1}+2^{2n_2})^{1/2}<(4d^2+4d^2)^{1/2}<3d,$$

where $d=d(A)$. Suppose now that $P=P_1\cup\ldots\cup P_4$. For $j=1,2$, denote by $P^{(j)}$ the projection of P onto the x_j-axis. Since A is convex, it has at most $(2\ell+4)$ vertices. It follows that if we draw a straight line parallel to the x_1-axis through each of these vertices, these lines will give a decomposition of A into at most two triangles and at most $(2\ell+1)$ trapeziums. Let B denote one of these triangles or trapeziums. For $j=1,2$, let $B^{(j)}$ denote the projection of B onto the x_j-axis. Clearly

$$B^{(1)}\times B^{(2)}=B\cup T'\cup T'',$$

where T' and T'' are disjoint triangles of types $\in\{1,\ldots,\ell\}$ and with diameters not exceeding $d(A)$. Furthermore,

$$P^{(1)}\times B^{(2)}=(B^{(1)}\times B^{(2)})\cup R'\cup R'',$$

where R' and R'' are disjoint rectangles with diameters not exceeding $((4d)^2+d^2)^{1/2}$. It is clear that $A \subseteq P^{(1)} \times A^{(2)}$, and $(P^{(1)} \times A^{(2)}) \setminus A$ is a (pairwise disjoint) union of at most $(4\ell+6)$ triangles of type $\in \{1,\ldots,\ell\}$ and $(4\ell+6)$ aligned rectangles. Finally, observe that $P \setminus (P^{(1)} \times A^{(2)})$ is a union of at most two disjoint rectangles of diameter not exceeding $((4d)^2+(2d)^2)^{1/2}$. ♣

Our next step is clearly to investigate these rectangles and triangles obtained from Lemma 8C. We first of all study the rectangles.

LEMMA 8D. *Suppose that R is an aligned rectangle.*
(a) There exist an integer $s \ll (\log(\mu(R)+2))^2$ and mutually disjoint special rectangles $R'_1,\ldots,R'_s$ such that

$$\bigcup_{\beta=1}^{s} R'_\beta \subseteq R \qquad \text{and} \qquad \mu\left(R \setminus \left(\bigcup_{\beta=1}^{s} R'_\beta\right)\right) \le 1.$$

(b) There exist mutually disjoint special rectangles $R''_1,\ldots,R''_4$, satisfying $\mu(R''_\beta) < 4\mu(R)$ for every $\beta=1,\ldots,4$, an integer $t \ll (\log(\mu(R)+2))^2$ and mutually disjoint special rectangles $R''_5,\ldots,R''_t$ such that

$$R \subseteq (R''_1 \cup \ldots \cup R''_4) \setminus \left(\bigcup_{\beta=5}^{t} R''_\beta\right) \qquad \text{and} \qquad \mu\left(\left((R''_1 \cup \ldots \cup R''_4) \setminus \left(\bigcup_{\beta=5}^{t} R''_\beta\right)\right) \setminus R\right) \le 1.$$

The proof of Lemma 8D is based on the following simple one-dimensional result. By a special interval, we mean an interval of the type $[m2^n,(m+1)2^n]$, where $m,n \in \mathbb{Z}$. Clearly, special rectangles are simply the cartesian product of two special intervals.

LEMMA 8E. *Suppose that $[a,b]$ is an interval in $\mathbb{R}$. Then for every natural number D, there exist special intervals $I_1,\ldots,I_D$ such that*

$$\bigcup_{\alpha=1}^{D} I_\alpha \subseteq [a,b] \qquad \text{and} \qquad \mu_0\left([a,b] \setminus \left(\bigcup_{\alpha=1}^{D} I_\alpha\right)\right) \le \left(\frac{7}{8}\right)^D (b-a).$$

Here μ_0 denotes the usual measure on $\mathbb{R}$.

Proof. Let I_1 denote the longest special interval in $[a,b]$. We then define I_α for $\alpha \ge 2$ inductively such that
(a) I_α is the longest special interval in the closure of $[a,b] \setminus (I_1 \cup \ldots \cup I_{\alpha-1})$;
(b) $I_1 \cup \ldots \cup I_\alpha$ is an interval; and
(c) if the closure of $[a,b] \setminus (I_1 \cup \ldots \cup I_{\alpha-1})$ is a union of two disjoint intervals, then I_α belongs to the longer of the two intervals, or to any one of them if they are of equal length.

Clearly $\mu_0(I_1) \ge (b-a)/4$. Indeed, if $n \in \mathbb{Z}$ satisfies $2^{n+1} \le b-a < 2^{n+2}$, then $2^n > (b-a)/4$ and there exists $m \in \mathbb{Z}$ such that $[m2^n,(m+1)2^n] \subseteq [a,b]$. A similar argument will give the inequality $\mu_0(I_\alpha) \ge \mu_0([a,b] \setminus (I_1 \cup \ldots \cup I_{\alpha-1}))/8$. The lemma follows easily. ♣

Proof of Lemma 8D. Suppose that $R=[a_1,b_1] \times [a_2,b_2]$. For $j=1,2$, we now apply Lemma 8E to the interval $[a_j,b_j]$ and obtain special intervals $I^{(j)}_1,\ldots,I^{(j)}_{D_j}$, with $D_j \ll \log(\mu(R)+2)$, such that

$$\bigcup_{\alpha_j=1}^{D_j} I^{(j)}_{\alpha_j} \subseteq [a_j,b_j] \qquad \text{and} \qquad \mu_0\left([a_j,b_j] \setminus \left(\bigcup_{\alpha_j=1}^{D_j} I^{(j)}_{\alpha_j}\right)\right) \le \frac{b_j-a_j}{2\mu(R)}.$$

The family of special rectangles

$$I^{(1)}_{\alpha_1} \times I^{(2)}_{\alpha_2}, \quad 1 \le \alpha_1 \le D_1 \text{ and } 1 \le \alpha_2 \le D_2,$$

clearly satisfies the requirements of (a). To prove (b), note first of all that for $j = 1, 2$, if $n_j \in \mathbb{Z}$ satisfies $2^{n_j - 1} < a_j \le 2^{n_j}$, then

$$[a_j, b_j] \subseteq [m_j 2^{n_j}, (m_j + 2)2^{n_j}]$$

for some $m_j \in \mathbb{Z}$. It follows that there exist four mutually disjoint special rectangles $R''_1, \ldots, R''_4$ such that $R \subseteq R''_1 \cup \ldots \cup R''_4$. Obviously, for every $\beta = 1, \ldots, 4$, $\mu(R''_\beta) < 4\mu(R)$. Furthermore, the set $(R''_1 \cup \ldots \cup R''_4) \setminus R$ is the disjoint union of at most four aligned rectangles. Applying (a) to each of these completes the proof. ♣

Next we study the triangles arising from Lemma 8C. Note that they are of types $\in \{1, \ldots, \ell\}$.

Definition. Suppose that $1 \le i \le \ell$. By a nice triangle of type i, we mean a triangle which is the intersection of a special triangle T^* of type i and a half-plane with the boundary parallel to one of the sides of T^*.

Suppose that $1 \le i \le \ell$, and that T is a triangle of type i. Let $T_0 \subseteq T$ be the largest inscribed special triangle of type i. Extending the edges of T_0 to the boundary of T, we see that T is the disjoint union of T_0 and at most three trapeziums and three parallelograms. Each of these trapeziums is clearly the disjoint union of a nice triangle of type i and a parallelogram. Note also that all the parallelograms are of types $\in \{(0, 0), (i, 1), (i, 2)\}$. To summarize, we have the following result.

LEMMA 8F. *Suppose that $1 \le i \le \ell$, and that T is a triangle of type i. Then T is the disjoint union of one special triangle of type i and at most three nice triangles of type i and six parallelograms of types $\in \{(0, 0), (i, 1), (i, 2)\}$.*

It follows that to handle the triangles arising from Lemma 8C, we need to investigate parallelograms of various types as well as nice triangles. Recall now that special parallelograms of type (i, j) and order $\mathbf{n}$ are obtained from special rectangles of order $\mathbf{n}$ by a linear transformation of determinant 1. The following analogue of Lemma 8D is therefore obvious.

LEMMA 8G. *Suppose that $1 \le i \le \ell$, and that $j = 1, 2$. Suppose further that P is a parallelogram of type (i, j).*

(a) There exist an integer $s \ll (\log(\mu(P) + 2))^2$ and mutually disjoint special parallelograms $P'_1, \ldots, P'_s$ of type (i, j) such that

$$\bigcup_{\beta=1}^{s} P'_\beta \subseteq P \qquad \text{and} \qquad \mu\left(P \setminus \left(\bigcup_{\beta=1}^{s} P'_\beta\right)\right) \le 1.$$

(b) There exist mutually disjoint special parallelograms $P''_1, \ldots, P''_4$ of type (i, j), with $\mu(P''_\beta) < 4\mu(P)$ for every $\beta = 1, \ldots, 4$, an integer $t \ll (\log(\mu(P) + 2))^2$ and mutually disjoint special parallelograms $P''_5, \ldots, P''_t$ of type (i, j) such that

$$P \subseteq (P''_1 \cup \ldots \cup P''_4) \setminus \left(\bigcup_{\beta=5}^{t} P''_\beta\right) \qquad \text{and} \qquad \mu\left(\left((P''_1 \cup \ldots \cup P''_4) \setminus \left(\bigcup_{\beta=5}^{t} P''_\beta\right)\right) \setminus P\right) \le 1.$$

It remains to investigate nice triangles.

LEMMA 8H. *Suppose that $1 \le i \le \ell$, and that T is a nice triangle of type i.*

(a) There exist an integer $s \ll (\log(\mu(T) + 2))$ and mutually disjoint special triangles $T'_1, \ldots, T'_s$ of type i and parallelograms $P'_1, \ldots, P'_s$ of types $\in \{(0, 0), (i, 1), (i, 2)\}$ such that

$$\left(\bigcup_{\nu=1}^{s} T'_\nu\right) \cup \left(\bigcup_{\nu=1}^{s} P'_\nu\right) \subseteq T \qquad \text{and} \qquad \mu\left(T \setminus \left(\left(\bigcup_{\nu=1}^{s} T'_\nu\right) \cup \left(\bigcup_{\nu=1}^{s} P'_\nu\right)\right)\right) \le 1.$$

(b) *There exist a special triangle T_0'' of type i, with $d(T_0'') < 4d(T)$, integers $t,q \ll (\log(\mu(T)+2))$ and mutually disjoint special triangles $T_1'',\ldots,T_t''$ of type i and parallelograms $P_1'',\ldots,P_q''$ of types $\in \{(0,0),(i,1),(i,2)\}$ such that*

$$T \subseteq T_0'' \setminus \left(\left(\bigcup_{\nu=1}^{t} T_\nu''\right) \cup \left(\bigcup_{\nu=1}^{q} P_\nu''\right)\right) \qquad \text{and} \qquad \mu\left(\left(T_0'' \setminus \left(\left(\bigcup_{\nu=1}^{t} T_\nu''\right) \cup \left(\bigcup_{\nu=1}^{q} P_\nu''\right)\right)\right) \setminus T\right) \le 1.$$

Proof. Part (a) will follows if we can prove that for every natural number D, there exist mutually disjoint special triangles $T_1,\ldots,T_D$ of type i and parallelograms $P_1,\ldots,P_D$ of types $\in \{(0,0),(i,1),(i,2)\}$ such that

$$\left(\bigcup_{\nu=1}^{D} T_\nu\right) \cup \left(\bigcup_{\nu=1}^{D} P_\nu\right) \subseteq T \tag{8.9}$$

and

$$\mu\left(T \setminus \left(\left(\bigcup_{\nu=1}^{D} T_\nu\right) \cup \left(\bigcup_{\nu=1}^{D} P_\nu\right)\right)\right) \le 4^{-D}\mu(T). \tag{8.10}$$

To prove (8.9) and (8.10), note that T is the intersection of a special triangle T^* of type i and a halfplane H with boundary parallel to one of the sides of T. Let $\mathbf{v}'$ and $\mathbf{v}''$ denote the vertices of T on the boundary of H, and let $\mathbf{c}$ denote the third vertex of T. Suppose that $T_1 \subseteq T$ is the largest inscribed special triangle of type i. Then $\mathbf{c}$ is a vertex of T_1 and $\mu(T_1) \ge \mu(T)/4$. Let $\mathbf{v}_1'$ and $\mathbf{v}_1''$ denote the two other vertices of T_1. The trapezium with vertices $\mathbf{v}'$, $\mathbf{v}''$, $\mathbf{v}_1'$ and $\mathbf{v}_1''$ is then clearly the disjoint union of a nice triangle T_1' of type i and a parallelogram P_1 of type $\in \{(0,0),(i,1),(i,2)\}$. Obviously $\mu(T_1') \le \mu(T)/4$. We now repeat the argument to T_1' and obtain a special triangle T_2 of type i, a nice triangle T_2' of type i and a parallelogram P_2, mutually disjoint and such that $T_1' = T_2 \cup T_2' \cup P_2$ and $\mu(T_2') \le \mu(T_1')/4$. After D steps, we obtain (8.9) and (8.10). Part (a) now follows from a suitable choice of D. To prove part (b), denote by T_0'' the smallest special triangle of type i containing T. Then $d(T_0'') < 2d(T)$. Furthermore, $T_0'' \setminus T$ is the disjoint union of a nice triangle of type i and a parellelogram of type $\in \{(0,0),(i,1),(i,2)\}$. Part (b) now follows on applying part (a) to this latter nice triangle. ♣

Lemma 8B now follows on combining Lemmas 8C, 8D, 8F, 8G and 8H.

We now combine the combinatorial Lemma 8A and the geometric Lemma 8B to give a proof of Theorem 16W. Our strategy is as follows. On the one hand, Lemma 8B enables us to obtain information on the discrepancy function of any given convex polygon in $\mathrm{POL}^\infty(\boldsymbol{\theta};x_1,x_2)$ in terms of the discrepancy functions of members of the various families of special objects. On the other hand, suppose that $\mathcal{P}$ is a discrete and finite subset of $\mathbb{R}^2$, containing many more points than we need. We now consider partitions of $\mathcal{P}$ given by these various families of special objects. Lemma 8A then enables us to choose a suitable subset of $\mathcal{P}$ to use in our construction of the desired infinite discrete set $\mathcal{Q}$ in Theorem 16W.

Given any discrete subset $\mathcal{P} \subseteq \mathbb{R}^2$ and any compact subset $B \subseteq \mathbb{R}^2$, we shall consider the discrepancy function

$$E[\mathcal{P};B] = \#(\mathcal{P} \cap B) - \mu(B).$$

Suppose that $A \in \mathrm{POL}^\infty(\boldsymbol{\theta};x_1,x_2)$ is arbitrary. We shall first of all use Lemma 8B to investigate the discrepancy function of A. The following lemma is in a more general form than needed.

LEMMA 8J. *Suppose that $A,B_1',\ldots,B_q',B_1'',\ldots,B_r''$ are compact subsets of $\mathbb{R}^2$. Suppose further that there exist $\epsilon_1',\ldots,\epsilon_q',\epsilon_1'',\ldots,\epsilon_r'' \in \{\pm 1\}$ such that*

$$\sum_{\tau=1}^{q} \epsilon_\tau' \chi_{B_\tau'} \le \chi_A \le \sum_{\tau=1}^{r} \epsilon_\tau'' \chi_{B_\tau''}$$

and

$$\sum_{\tau=1}^{r} \epsilon_\tau'' \mu(B_\tau'') - \sum_{\tau=1}^{q} \epsilon_\tau' \mu(B_\tau') \le D_1.$$

Let $\mathcal{P} \subseteq \mathbb{R}^2$ be a discrete set such that $|E[\mathcal{P}; B'_\tau]| \le D_2$ for every $\tau = 1, \ldots, q$ and $|E[\mathcal{P}; B''_\tau]| \le D_2$ for every $\tau = 1, \ldots, r$. Then

$$|E[\mathcal{P}; A]| \le D_1 + D_2 \max\{q, r\}.$$

PROOF. Clearly

$$\begin{aligned}
E[\mathcal{P}; A] &= \sum_{\mathbf{p} \in A \cap \mathcal{P}} 1 - \mu(A) \le \sum_{\tau=1}^{r} \epsilon''_\tau \sum_{\mathbf{p} \in B''_\tau \cap \mathcal{P}} 1 - \mu(A) \\
&= \sum_{\tau=1}^{r} \epsilon''_\tau \left(\sum_{\mathbf{p} \in B''_\tau \cap \mathcal{P}} 1 - \mu(B''_\tau) \right) + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \mu(A) \right) \\
&= \sum_{\tau=1}^{r} \epsilon''_\tau E[\mathcal{P}; B''_\tau] + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \mu(A) \right) \\
&\le \sum_{\tau=1}^{r} |E[\mathcal{P}; B''_\tau]| + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \sum_{\tau=1}^{r} \epsilon'_\tau \mu(B'_\tau) \right) \\
&\le D_2 r + D_1. \qquad (8.11)
\end{aligned}$$

A similar argument gives

$$-E[\mathcal{P}; A] \le D_2 q + D_1. \qquad (8.12)$$

The result now follows on combining (8.11) and (8.12). ♣

Let $\mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ denote the big family of all special triangles, special parallelograms and special rectangles defined in this section; in other words,

$$\mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2) = \left(\bigcup_{\substack{1 \le i \le \ell \\ n \in \mathbb{Z}}} \mathrm{ST}(i, n) \right) \cup \left(\bigcup_{\substack{1 \le i \le \ell \\ 1 \le j \le 2 \\ \mathbf{n} \in \mathbb{Z}^2}} \mathrm{SP}(i, j, \mathbf{n}) \right) \cup \left(\bigcup_{\mathbf{n} \in \mathbb{Z}^2} \mathrm{SR}(\mathbf{n}) \right).$$

We now make use of the combinatorial information derived from Lemma 8A.

LEMMA 8K. *Suppose that $\mathcal{P} \subseteq \mathbb{R}^2$ is a finite set, and that $\alpha \in [0, 1]$ is fixed. Then there exists a function $f : \mathcal{P} \to \{-\alpha, 1 - \alpha\}$ such that for every polygon $B \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfying $d(B) \ge 1$, we have*

$$\left| \sum_{\mathbf{p} \in B \cap \mathcal{P}} f(\mathbf{p}) \right| \ll_\epsilon \left(\ell (\log(d(B) + 2))^2 \right)^{1+\epsilon}. \qquad (8.13)$$

PROOF. We apply Lemma 8A with $X = \mathcal{P}$, and so have to introduce a sequence of partitions of $\mathcal{P}$. Let

$$\begin{aligned}
\mathrm{SET}^\infty(\boldsymbol{\theta}; x_1, x_2) = &\{\mathrm{ST}(i, n) : 1 \le i \le \ell \text{ and } n \in \mathbb{Z}\} \\
&\cup \{\mathrm{SP}(i, j, \mathbf{n}) : 1 \le i \le \ell \text{ and } 1 \le j \le 2 \text{ and } \mathbf{n} \in \mathbb{Z}^2\} \\
&\cup \{\mathrm{SR}(\mathbf{n}) : \mathbf{n} \in \mathbb{Z}^2\}.
\end{aligned}$$

For every $C \in \mathrm{SET}^\infty(\boldsymbol{\theta}; x_1, x_2)$, denote by $d(C)$ the common diameter of all the elements of C. We now define a linear ordering on the subset

$$\{C \in \mathrm{SET}^\infty(\boldsymbol{\theta}; x_1, x_2) : d(C) \ge 1\}$$

according to the size of $d(C)$, with the convention that this ordering is defined arbitrarily in the case of equal diameters. Observe that for any real number $y \geq 1$,

$$\begin{aligned}
&\#\{C \in \mathrm{SET}^\infty(\boldsymbol{\theta}; x_1, x_2) : 1 \leq d(C) \leq y\} \\
&\quad = \sum_{i=1}^{\ell} \#\{n \in \mathbb{Z} : 1 \leq d(\mathrm{ST}(i,n)) \leq y\} + \sum_{i=1}^{\ell}\sum_{j=1}^{2} \#\{\mathbf{n} \in \mathbb{Z}^2 : 1 \leq d(\mathrm{SP}(i,j,\mathbf{n})) \leq y\} \\
&\qquad + \#\{\mathbf{n} \in \mathbb{Z}^2 : 1 \leq d(\mathrm{SR}(\mathbf{n})) \leq y\} \\
&\quad \ll \ell \log(y+2) + \ell(\log(y+2))^2 \ll \ell(\log(y+2))^2. \qquad (8.14)
\end{aligned}$$

Suppose that $\mathcal{P}$ is fixed. We now let $\mathcal{Y}^{(1)}, \mathcal{Y}^{(2)}, \mathcal{Y}^{(3)}, \ldots$ be the partitions of $\mathcal{P}$ defined by the families in $\{C \in \mathrm{SET}^\infty(\boldsymbol{\theta}; x_1, x_2) : 1 \leq d(C) \leq d(B)\}$ ordered in the way described. The result now follows from Lemma 8A and (8.14). ♣

We now begin our construction of the desired set $\mathcal{Q}$. Let $\kappa = 2^k$, where $k \in \mathbb{N}$, and consider the set

$$\mathcal{P} = \{(a/\kappa, b/\kappa) : a, b \in \mathbb{Z} \text{ and } -\kappa^2 \leq a, b < \kappa^2\}$$

in the square $[-\kappa,\kappa)^2$. Clearly $\#\mathcal{P} = 4\kappa^4$. Let $\alpha = \kappa^{-2}$. Then $\alpha\#\mathcal{P} = 4\kappa^2$, the expected number of points of the desired set $\mathcal{Q}$ in $[-\kappa,\kappa)^2$. By Lemma 8K, there exists a function $f : \mathcal{P} \to \{-\alpha, 1-\alpha\}$ such that the inequality (8.13) holds for all polygons $B \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfying $B \subseteq [-\kappa,\kappa)^2$ and $d(B) \geq 1$. Writing $\mathcal{P}_k = \{\mathbf{p} \in \mathcal{P} : f(\mathbf{p}) = 1-\alpha\}$, we have

$$\sum_{\mathbf{p} \in B \cap \mathcal{P}} f(\mathbf{p}) = \sum_{\mathbf{p} \in B \cap \mathcal{P}_k} 1 - \kappa^{-2} \sum_{\mathbf{p} \in B \cap \mathcal{P}} 1. \qquad (8.15)$$

Furthermore, it is easy to see that for any convex polygon $B \subseteq [-\kappa,\kappa)^2$, we have

$$\left| \sum_{\mathbf{p} \in B \cap \mathcal{P}} 1 - \kappa^2 \mu(B) \right| \ll \kappa\sigma(\partial B) \ll \kappa^2, \qquad (8.16)$$

where $\sigma(\partial B)$ denotes the length of the perimeter of B. It follows on combining (8.13), (8.15) and (8.16) that

$$\begin{aligned}
|E[\mathcal{P}_k; B]| &= \left| \sum_{\mathbf{p} \in B \cap \mathcal{P}_k} 1 - \mu(B) \right| \\
&\leq \left| \sum_{\mathbf{p} \in B \cap \mathcal{P}_k} 1 - \kappa^{-2} \sum_{\mathbf{p} \in B \cap \mathcal{P}} 1 \right| + \left| \kappa^{-2} \sum_{\mathbf{p} \in B \cap \mathcal{P}} 1 - \mu(B) \right| \\
&\ll_\epsilon \left(\ell(\log(d(B)+2))^2\right)^{1+\epsilon}
\end{aligned}$$

for all polygons $B \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfying $B \subseteq [-\kappa,\kappa)^2$ and $d(B) \geq 1$.

Suppose now that the polygon $B \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfies $B \subseteq [-\kappa,\kappa)^2$ and $d(B) < 1$. Then $B \subseteq B_0$ for some $B_0 \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ with $1 \leq d(B_0) < 2$. Applying (8.13) and (8.15) to B_0, we have

$$\sum_{\mathbf{p} \in B_0 \cap \mathcal{P}_k} 1 = \kappa^{-2} \sum_{\mathbf{p} \in B_0 \cap \mathcal{P}} 1 + \sum_{\mathbf{p} \in B_0 \cap \mathcal{P}} f(\mathbf{p}) \leq 4 + \sum_{\mathbf{p} \in B_0 \cap \mathcal{P}} f(\mathbf{p}) \ll_\epsilon \ell^{1+\epsilon},$$

noting that $\mu(B_0) \leq (d(B_0))^2 < 4$. Hence

$$\sum_{\mathbf{p} \in B \cap \mathcal{P}_k} 1 \leq \sum_{\mathbf{p} \in B_0 \cap \mathcal{P}_k} 1 \ll_\epsilon \left(\ell(\log(d(B)+2))^2\right)^{1+\epsilon}.$$

Using $\mu(B) < \mu(B_0) < 4$, we have

$$|E[\mathcal{P}_k; B]| \ll_\epsilon (\ell(\log(d(B)+2))^2)^{1+\epsilon}. \tag{8.17}$$

It now follows that the inequality (8.17) holds for all polygons $B \in \mathrm{SPEC}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfying $B \subseteq [-\kappa, \kappa)^2$. Combining this with Lemmas 8B and 8J, we conclude that

$$|E[\mathcal{P}_k; C]| \ll_\epsilon \ell^{2+\epsilon}(\log(d(C)+2))^{5+\epsilon} \tag{8.18}$$

for every $C \in \mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$ satisfying $C \subseteq [-\kappa, \kappa)^2$.

We now construct the set $\mathcal{Q}$ in terms of the sets $\mathcal{P}_k$ of some selected integer values of k. Note first of all that

$$\bigcup_{n\in\mathbb{N}} \left(\left[-2^{2^n}, 2^{2^n}\right)^2 \setminus \left[-2^{2^{n-1}}, 2^{2^{n-1}}\right)^2 \right) = \mathbb{R}^2 \setminus [-2,2)^2,$$

and that any set in this union is the disjoint union of four aligned rectangles. We shall show that the set

$$\mathcal{Q} = \mathcal{P}_1 \cup \left(\bigcup_{\substack{k=2^n\\ n\in\mathbb{N}}} \left(\mathcal{P}_k \cap \left([-2^k, 2^k)^2 \setminus [-2^{k/2}, 2^{k/2})^2 \right) \right) \right)$$

satisfies the requirements of Theorem 16W.

Consider any arbitrary $A \in \mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$. For every $k = 2^n$ with $n \in \mathbb{N}$, the intersection

$$A_k = A \cap \left([-2^k, 2^k)^2 \setminus [-2^{k/2}, 2^{k/2})^2 \right)$$

is the disjoint union of at most four sets in $\mathrm{POL}^\infty(\boldsymbol{\theta}; x_1, x_2)$. It follows from (8.18) that

$$\begin{aligned} |E[\mathcal{Q}; A]| &= \left| \sum_{\mathbf{q}\in A\cap\mathcal{Q}} 1 - \mu(A) \right| \\ &= \left| \#(A\cap\mathcal{P}_1) - \mu(A\cap[-2,2)^2) + \sum_{\substack{k=2^n\\ n\in\mathbb{N}}} (\#(A_k\cap\mathcal{P}_k) - \mu(A_k)) \right| \\ &\ll_\epsilon \sum\nolimits^* \ell^{2+\epsilon}(\min\{\log(d(A)+2), k\})^{5+\epsilon}. \end{aligned} \tag{8.19}$$

Here the summation $\sum^*$ is extended over all $k = 2^n$, where $n \in \mathbb{N}$, and for which A_k is non-empty. Simple calculation gives

$$\sum\nolimits^* (\min\{\log(d(A)+2), k\})^{5+\epsilon} \ll_\epsilon (\log(d(A)+2))^{5+\epsilon}. \tag{8.20}$$

Theorem 16W now follows from (8.19) and (8.20).

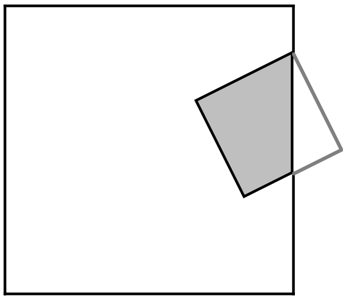
9. A Fourier Transform Approach

In this section, we shall first establish the following weaker form of Theorem 10 in the case when A is a square.

THEOREM 10W. *For every distribution $\mathcal{P}$ of N points in the square $[0,1]^2$, there exists a rotated square B in $\mathbb{R}^2$ of side length at most $\frac{1}{2}$ and such that*

$$|Z[\mathcal{P}; B] - N\mu(B \cap [0,1]^2)| \gg N^{\frac{1}{4}}.$$

Note that we are not insisting that the square B lies entirely in $[0,1]^2$, and we are only studying the discrepancy of the part of B which lies in $[0,1]^2$.



Suppose that $\mathcal{P}$ is a distribution of N points in the square $[0,1]^2$. We shall introduce two measures. The discrete measure Z_0 is the counting measure of the distribution $\mathcal{P}$, so that for every set $B \subseteq \mathbb{R}^2$,

$$Z_0(B) = \int_B \mathrm{d}Z_0(\mathbf{x}) = \int_{\mathbb{R}^2} \chi_B(\mathbf{x})\,\mathrm{d}Z_0(\mathbf{x}) = \#(\mathcal{P} \cap B)$$

denotes the number of points of $\mathcal{P}$ that fall into B. Here χ_B denotes the characteristic function of the set B. We also let μ_0 denote the Lebesgue area measure μ in $\mathbb{R}^2$, restricted to the square $[0,1]^2$, so that for every measurable set $B \subseteq \mathbb{R}^2$,

$$\mu_0(B) = \int_B \mathrm{d}\mu_0(\mathbf{x}) = \int_{\mathbb{R}^2} \chi_B(\mathbf{x})\,\mathrm{d}\mu_0(\mathbf{x}) = \mu(B \cap [0,1]^2).$$

With these two measures, it is then appropriate to consider the discrepancy measure $D_0 = Z_0 - N\mu_0$ of the point set $\mathcal{P}$, so that for every measurable set $B \subseteq \mathbb{R}^2$,

$$D_0(B) = Z_0(B) - N\mu_0(B) = \#(\mathcal{P} \cap B) - N\mu(B \cap [0,1]^2)$$

represents the discrepancy of the part of B which lies in $[0,1]^2$.

For real numbers $r \geq 0$ and $\theta \in [0,2\pi]$, let $B(r,\theta)$ denote the square $[-r,r]^2$ rotated anticlockwise by an angle θ, and let $\chi_{r,\theta}$ denote the characteristic function of $B(r,\theta)$. Furthermore, for every vector $\mathbf{x} \in \mathbb{R}^2$, let

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

denote the image of $B(r,\theta)$ under translation by $\mathbf{x}$. Consider the function

$$F_{r,\theta} = \chi_{r,\theta} * (\mathrm{d}Z_0 - N\mathrm{d}\mu_0), \tag{9.1}$$

where $f * g$ denotes the convolution of the functions f and g. More precisely, we consider

$$F_{r,\theta}(\mathbf{x}) = \int_{\mathbb{R}^2} \chi_{r,\theta}(\mathbf{x}-\mathbf{y})(\mathrm{d}Z_0(\mathbf{y}) - N\mathrm{d}\mu_0(\mathbf{y})).$$

Note that the rotated square $B(r,\theta)$ is symmetric across the origin, and so

$$\mathbf{x}-\mathbf{y} \in B(r,\theta) \quad \Leftrightarrow \quad \mathbf{y}-\mathbf{x} \in B(r,\theta) \quad \Leftrightarrow \quad \mathbf{y} \in B(r,\theta,\mathbf{x}).$$

It follows that

$$\int_{\mathbb{R}^2} \chi_{r,\theta}(\mathbf{x}-\mathbf{y})\,\mathrm{d}Z_0(\mathbf{y}) = \#(P \cap B(r,\theta,\mathbf{x})) = Z_0(B(r,\theta,\mathbf{x}))$$

and

$$\int_{\mathbb{R}^2} \chi_{r,\theta}(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mu_0(\mathbf{y}) = \mu(P \cap B(r,\theta,\mathbf{x}) \cap [0,1]^2) = \mu_0(B(r,\theta,\mathbf{x})),$$

and so

$$F_{r,\theta}(\mathbf{x}) = Z_0(B(r,\theta,\mathbf{x})) - N\mu_0(B(r,\theta,\mathbf{x})) = D_0(B(r,\theta,\mathbf{x})) \tag{9.2}$$

represents the discrepancy of the part of $B(r,\theta,\mathbf{x})$ in the square $[0,1]^2$.

We now appeal to the theory of Fourier transforms.

Let $L_1(\mathbb{R}^2)$ denote the set of all measurable real or complex valued functions f defined on $\mathbb{R}^2$ such that the integral

$$\int_{\mathbb{R}^2} |f(\mathbf{x})| \,\mathrm{d}\mathbf{x}$$

is finite. For such a function $f \in L_1(\mathbb{R}^2)$, the Fourier transform $\hat{f}$ is a complex valued function defined on $\mathbb{R}^2$ satisfying

$$\hat{f}(\mathbf{t}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} f(\mathbf{x}) \mathrm{e}^{-\mathrm{i}\langle \mathbf{x},\mathbf{t}\rangle} \,\mathrm{d}\mathbf{x}$$

for every $\mathbf{t} \in \mathbb{R}^2$. Here $\langle \mathbf{x},\mathbf{t}\rangle$ denotes the inner product of $\mathbf{x}$ and $\mathbf{t}$. It is not too difficult to check that for any two functions $f,g \in L_1(\mathbb{R}^2)$, the Fourier transforms $\hat{f}$ and $\hat{g}$ satisfy

$$\widehat{f*g} = \hat{f}\hat{g}. \tag{9.3}$$

Let $L_2(\mathbb{R}^2)$ denote the set of all measurable real or complex valued functions f defined on $\mathbb{R}^2$ such that the integral

$$\int_{\mathbb{R}^2} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x}$$

is finite. Then the Parseval-Plancherel theorem states that for every function $f \in L_1(\mathbb{R}^2) \cap L_2(\mathbb{R}^2)$, the Fourier transform $\hat{f} \in L_2(\mathbb{R}^2)$ and satisfies

$$\int_{\mathbb{R}^2} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^2} |\hat{f}(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t}. \tag{9.4}$$

For every $\mathbf{t} \in \mathbb{R}^2$, we write

$$\phi(\mathbf{t}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \mathrm{e}^{-\mathrm{i}\langle \mathbf{x},\mathbf{t}\rangle} \,\mathrm{d}D_0(\mathbf{x}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \mathrm{e}^{-\mathrm{i}\langle \mathbf{x},\mathbf{t}\rangle} (\mathrm{d}Z_0(\mathbf{x}) - N\mathrm{d}\mu_0(\mathbf{x})). \tag{9.5}$$

Then it follows from (9.1) and (9.3)–(9.5) that

$$\int_{\mathbb{R}^2} |F_{r,\theta}(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^2} |\hat{F}_{r,\theta}(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t} = \int_{\mathbb{R}^2} |\hat{\chi}_{r,\theta}(\mathbf{t})|^2 |\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t}. \tag{9.6}$$

Note that the measure $Z_0 - N\mu_0$, and hence the function ϕ, is determined by the point distribution $\mathcal{P}$ and has nothing to do with the squares $B(r,\theta)$. On the other hand, the characteristic function $\chi_{r,\theta}$ is determined by the square $B(r,\theta)$ and has nothing to do with the point distribution $\mathcal{P}$. In other words, the identity (9.6) represents a separation of measure and squares as a result and at the expense of passing over to the corresponding Fourier transforms.

In lower bound proofs, the point distributions $\mathcal{P}$ are arbitrary, so we have very little control over the measure $Z_0 - N\mu_0$. However, we have a lot of information to study the characteristic functions $\chi_{r,\theta}$ and their Fourier transforms $\hat{\chi}_{r,\theta}$. Indeed, for the measure part of the argument here, we need only the following estimate on the trivial error.

LEMMA 9A. *Suppose that a measurable set $B \subseteq [0,1]^2$ satisfies the inequalities*

$$0 < \frac{\delta}{N} \le \mu(B) \le \frac{1-\delta}{N}$$

for some real number $\delta > 0$. Then

$$\int_{\mathbb{R}^2} |Z_0(B+\mathbf{x}) - N\mu_0(B+\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \ge \delta^3.$$

Here $B+\mathbf{x} = \{\mathbf{x}+\mathbf{y} : \mathbf{y} \in B\}$ represents the image of the set B under translation by the vector $\mathbf{x}$.

Proof. Suppose first of all that $Z_0(B+\mathbf{x})\ge 1$. Then

$$Z_0(B+\mathbf{x})-N\mu_0(B+\mathbf{x})\ge Z_0(B+\mathbf{x})-N\mu(B)\ge Z_0(B+\mathbf{x})+\delta-1\ge \delta Z_0(B+\mathbf{x}),$$

so that

$$|Z_0(B+\mathbf{x})-N\mu_0(B+\mathbf{x})|\ge \delta Z_0(B+\mathbf{x}).$$

Note that this last inequality is trivial if $Z_0(B+\mathbf{x})=0$. Hence, writing $\mathbf{p}-B=\{\mathbf{p}-\mathbf{y}:\mathbf{y}\in B\}$ and $\chi_{\mathbf{p}-B}$ for its characteristic function, we have

$$\begin{aligned}\int_{\mathbb{R}^2}|Z_0(B+\mathbf{x})-N\mu_0(B+\mathbf{x})|^2\,\mathrm{d}\mathbf{x}&\ge \delta^2\int_{\mathbb{R}^2}Z_0^2(B+\mathbf{x})\,\mathrm{d}\mathbf{x}\ge \delta^2\int_{\mathbb{R}^2}Z_0(B+\mathbf{x})\,\mathrm{d}\mathbf{x}\\&=\delta^2\sum_{\mathbf{p}\in\mathcal{P}}\int_{\mathbb{R}^2}\chi_{\mathbf{p}-B}(\mathbf{x})\,\mathrm{d}\mathbf{x}=\delta^2\sum_{\mathbf{p}\in\mathcal{P}}\mu(\mathbf{p}-B)=\delta^2N\mu(B)\ge\delta^3.\end{aligned}$$ ♣

The main part of the proof is therefore to study the characteristic functions $\chi_{r,\theta}$ and their Fourier transforms $\hat{\chi}_{r,\theta}$. Ideally, we would like an inequality of the type

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

However, this makes use of only one square $B(r,\theta)$, with no rotation or contraction. For any parameter $q>0$, we consider instead an average

$$\omega_q(\mathbf{t})=\frac{1}{q}\int_{q/2}^{q}\int_0^{\pi/4}|\hat{\chi}_{r,\theta}(\mathbf{t})|^2\,\mathrm{d}\theta\mathrm{d}r. \tag{9.7}$$

We have the following amplification result which we shall use to blow up the trivial error obtained in Lemma 9A.

LEMMA 9B. *Suppose that $0<p<q$. Then we have*

$$\frac{\omega_q(\mathbf{t})}{\omega_p(\mathbf{t})}\gg\frac{q}{p}, \tag{9.8}$$

uniformly for all vectors $\mathbf{t}\in\mathbb{R}^2$.

We shall split the proof of Lemma 9B into a number of steps. Throughout, we suppose that $r>0$ and $0\le\theta\le\pi/4$.

LEMMA 9C. *For every $\mathbf{t}=(t_1,t_2)\in\mathbb{R}^2$, we have*

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

Proof. For every $\mathbf{x}\in\mathbb{R}^2$, we shall denote by $\theta\mathbf{x}$ the image of $\mathbf{x}$ under anticlockwise rotation by an angle θ, and denote by $\theta^{-1}\mathbf{x}$ the image of $\mathbf{x}$ under clockwise rotation by an angle θ. Note also that the transpose of a rotation matrix is the inverse rotation matrix. Hence

$$\begin{aligned}\hat{\chi}_{r,\theta}(\mathbf{t})&=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}\langle\mathbf{x},\mathbf{t}\rangle}\chi_{r,\theta}(\mathbf{x})\,\mathrm{d}\mathbf{x}=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}\langle\mathbf{x},\mathbf{t}\rangle}\chi_r(\theta^{-1}\mathbf{x})\,\mathrm{d}\mathbf{x}\\&=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}\langle\theta\mathbf{y},\mathbf{t}\rangle}\chi_r(\mathbf{y})\,\mathrm{d}\mathbf{y}=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}\langle\mathbf{y},\theta^{-1}\mathbf{t}\rangle}\chi_r(\mathbf{y})\,\mathrm{d}\mathbf{y}=\hat{\chi}_r(\theta^{-1}\mathbf{t}).\end{aligned}$$

The result follows on noting that $\theta^{-1}\mathbf{t}=(t_1\cos\theta+t_2\sin\theta,-t_1\sin\theta+t_2\cos\theta)$. ♣

LEMMA 9D. *For every $\mathbf{u}=(u_1,u_2)\in\mathbb{R}^2$, we have*

$$\hat{\chi}_r(\mathbf{u})=\frac{2\sin(ru_1)\sin(ru_2)}{\pi u_1u_2}.$$

PROOF. Note that

$$\hat{\chi}_r(\mathbf{u})=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{e}^{-\mathrm{i}(x_1u_1+x_2u_2)}\chi_r(x_1,x_2)\,\mathrm{d}\mathbf{x}=\frac{1}{2\pi}\left(\int_{-r}^{r}\mathrm{e}^{-\mathrm{i}x_1u_1}\,\mathrm{d}x_1\right)\left(\int_{-r}^{r}\mathrm{e}^{-\mathrm{i}x_2u_2}\,\mathrm{d}x_2\right).$$

The result follows easily. ♣

We note next that in view of the integration over θ in the definition of $\omega_q(\mathbf{t})$, it suffices to establish the inequality (9.8) for those $\mathbf{t}=(t_1,t_2)\in\mathbb{R}^2$ satisfying $t_1>0$ and $t_2=0$. Lemma 9B then follows easily from the result below.

LEMMA 9E. *Suppose that $t_1>0$. Then*

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

PROOF. Using Lemmas 9C and 9D, we have

$$\omega_q(t_1,0)\asymp\frac{1}{q}\int_{q/2}^{q}\int_0^{\pi/4}\frac{\sin^2(rt_1\cos\theta)\sin^2(rt_1\sin\theta)}{t_1^4\cos^2\theta\sin^2\theta}\,\mathrm{d}\theta\mathrm{d}r.$$

Since $0\le\theta\le\pi/4$, we have $\sin\theta\asymp\theta$ and $\cos\theta\asymp1$, and so

$$\omega_q(t_1,0)\asymp\frac{1}{q}\int_{q/2}^{q}\int_0^{\pi/4}\frac{\sin^2(rt_1\cos\theta)\sin^2(rt_1\sin\theta)}{t_1^4\theta^2}\,\mathrm{d}\theta\mathrm{d}r.$$

We consider two cases.

Suppose first of all that $t_1\le4/\pi q$. Then for all r and θ satisfying $q/2\le r\le q$ and $0\le\theta\le\pi/4$, we have $\sin(rt_1\cos\theta)\asymp qt_1$ and $\sin(rt_1\sin\theta)\asymp qt_1\theta$. Hence

$$\omega_q(t_1,0)\asymp\frac{1}{q}\int_{q/2}^{q}\int_0^{\pi/4}\frac{(qt_1)^2(qt_1\theta)^2}{t_1^4\theta^2}\,\mathrm{d}\theta\mathrm{d}r\asymp q^4\asymp\min\left\{q^4,\frac{q}{t_1^3}\right\}.$$

Suppose next that $t_1>4/\pi q$. We then split the interval $[0,\pi/4]$ into two intervals at the point $\theta=1/qt_1$. On the one hand, we have the crude estimate

$$\int_{1/qt_1}^{\pi/4}\frac{\sin^2(rt_1\cos\theta)\sin^2(rt_1\sin\theta)}{t_1^4\theta^2}\,\mathrm{d}\theta\le\int_{1/qt_1}^{\pi/4}\frac{\mathrm{d}\theta}{t_1^4\theta^2}=\frac{1}{t_1^4}\left(qt_1-\frac{4}{\pi}\right).$$

On the other hand, if $0\le\theta\le1/qt_1$, then we have

$$\sin(rt_1\sin\theta)\asymp qt_1\theta\qquad\text{and}\qquad\frac{1}{q}\int_{q/2}^{q}\sin^2(rt_1\cos\theta)\,\mathrm{d}r\asymp1.$$

For the inequalities on the right hand side, the upper bound is obvious. For the lower bound, note that as r runs through the interval $[q/2,q]$, the quantity $rt_1\cos\theta$ runs through an interval of length

$$\frac{qt_1\cos\theta}{2}>\frac{2}{\pi}\cos\frac{\pi}{4}.$$

It now follows that

$$\omega_q(t_1,0) \asymp \int_0^{1/qt_1} \frac{q^2}{t_1^2}\,\mathrm{d}\theta + O\left(\frac{1}{t_1^4}\left(qt_1 - \frac{4}{\pi}\right)\right) \asymp \frac{q}{t_1^3} \asymp \min\left\{q^4, \frac{q}{t_1^3}\right\}. \qquad \clubsuit$$

We now make the choice $p = \frac{1}{3}N^{-\frac{1}{2}}$ and $q = \frac{1}{4}$. Note that for every r and θ satisfying $p/2 \le r \le p$ and $0 \le \theta \le \pi/4$, we have

$$\frac{1}{9N} \le \mu(B(r,\theta)) \le \frac{4}{9N}.$$

Using Lemma 9A with $\delta = \frac{1}{9}$, we have

$$\int_{\mathbb{R}^2} |Z_0(B(r,\theta,\mathbf{x})) - N\mu_0(B(r,\theta,\mathbf{x}))|^2\,\mathrm{d}\mathbf{x} \gg 1.$$

It follows from (9.2), (9.6) and (9.7) that

$$\int_{\mathbb{R}^2} \omega_p(\mathbf{t})|\phi(\mathbf{t})|^2\,\mathrm{d}\mathbf{t} \gg 1.$$

Using Lemma 9B, we conclude that

$$\int_{\mathbb{R}^2} \omega_q(\mathbf{t})|\phi(\mathbf{t})|^2\,\mathrm{d}\mathbf{t} \gg \frac{q}{p} \gg N^{\frac{1}{2}}.$$

Combining this with (9.2), (9.6) and (9.7), we conclude that

$$\int_{1/8}^{1/4}\int_0^{\pi/4}\int_{\mathbb{R}^2} |Z_0(B(r,\theta,\mathbf{x})) - N\mu_0(B(r,\theta,\mathbf{x}))|^2\,\mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r \gg N^{\frac{1}{2}}.$$

Theorem 10W follows on noting that $Z_0(B(r,\theta,\mathbf{x})) - N\mu_0(B(r,\theta,\mathbf{x})) = 0$ if $\mathbf{x} \notin [-1-r\sqrt{2}, 1+r\sqrt{2}]^2$.

Next, we compare Beck's technique to a method due to Montgomery. Montgomery's analysis is not as clear as that of Beck, but certain parts of his argument has great similarity with the approach of Beck. In simple cases, Montgomery uses Fourier series. Then the corresponding Fourier coefficients behave like the Fourier transforms studied by Beck.

Let us return to the formulation where we consider the torus $[0,1)^2$. Suppose that S is a measurable set in the torus $[0,1)^2$. We study the discrepancy function

$$D[\mathcal{P};S+\mathbf{x}] = Z[\mathcal{P};S+\mathbf{x}] - N\mu(S+\mathbf{x}),$$

where $S+\mathbf{x}$ denotes the image of S under translation by the vector $\mathbf{x}$. Suppose further that the set S is symmetric across the origin, so that

$$\mathbf{p} \in S+\mathbf{x} \quad\Leftrightarrow\quad \mathbf{p}-\mathbf{x} \in S \quad\Leftrightarrow\quad \mathbf{x}-\mathbf{p} \in S.$$

Then

$$D[\mathcal{P};S+\mathbf{x}] = \sum_{\mathbf{p}\in\mathcal{P}} \chi_S(\mathbf{x}-\mathbf{p}) - N\mu(S).$$

We now appeal to the theory of Fourier series.

Let $L_1([0,1)^2)$ denote the set of all measurable real or complex valued functions f defined on $[0,1)^2$ such that the integral

$$\int_{[0,1)^2} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x}$$

is finite. For such a function $f \in L_1([0,1)^2)$, we can define the Fourier coefficient $\hat{f}(\mathbf{k})$ for every $\mathbf{k} \in \mathbb{Z}^2$ by

$$\hat{f}(\mathbf{k}) = \int_{[0,1)^2} f(\mathbf{x}) \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{x},\mathbf{k} \rangle} \,\mathrm{d}\mathbf{x}.$$

Let $L_2([0,1)^2)$ denote the set of all measurable real or complex valued functions f defined on $[0,1)^2$ such that the integral

$$\int_{[0,1)^2} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x}$$

is finite. Then the Parseval theorem states that for every function $f \in L_1([0,1)^2) \cap L_2([0,1)^2)$, we have

$$\int_{[0,1)^2} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} |\hat{f}(\mathbf{k})|^2.$$

The $\mathbf{k}$-th Fourier coefficient of the function $D[\mathcal{P}; S+\mathbf{x}]$ is given by

$$\hat{D}[\mathcal{P}; S; \mathbf{k}] = \sum_{\mathbf{p} \in \mathcal{P}} \int_{[0,1)^2} \chi_S(\mathbf{x}-\mathbf{p}) \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{x},\mathbf{k} \rangle} \,\mathrm{d}\mathbf{x} - N\mu(S) \int_{[0,1)^2} \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{x},\mathbf{k} \rangle} \,\mathrm{d}\mathbf{x}.$$

Note that for $\mathbf{k} \neq \mathbf{0}$, the second integral on the right hand side vanishes, and so

$$\hat{D}[\mathcal{P}; S; \mathbf{k}] = \sum_{\mathbf{p} \in \mathcal{P}} \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{p},\mathbf{k} \rangle} \int_{[0,1)^2} \chi_S(\mathbf{x}-\mathbf{p}) \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{x}-\mathbf{p},\mathbf{k} \rangle} \,\mathrm{d}\mathbf{x} = \hat{\chi}_S(\mathbf{k}) \hat{U}(\mathbf{k}),$$

where

$$\hat{U}(\mathbf{k}) = \sum_{\mathbf{p} \in \mathcal{P}} \mathrm{e}^{-2\pi \mathrm{i} \langle \mathbf{p},\mathbf{k} \rangle}$$

is determined by the point distribution $\mathcal{P}$ and has nothing to do with the set S. On the other hand, the characteristic function χ_S, and hence its Fourier coefficients, is determined by the set S and has nothing to do with the point distribution $\mathcal{P}$.

We shall first illustrate this method by sketching a proof of the following variant of Theorem 1.

THEOREM 1V. *For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$, there exists an aligned square B in $[0,1)^2$ such that*

$$|D[\mathcal{P}; B]| \gg (\log N)^{\frac{1}{2}}.$$

For any real number r satisfying $0 \le r \le \frac{1}{2}$, let $B(r)$ denote the square $[-r,r]^2$, and let χ_r denote the characteristic function of $B(r)$. Furthermore, for every vector $\mathbf{x} \in [0,1)^2$, let

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

denote the image of $B(r)$ under translation by $\mathbf{x}$. Then

$$\int_{[0,1)^2} |D[\mathcal{P}; B(r,\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} |\hat{\chi}_r(\mathbf{k})|^2 |\hat{U}(\mathbf{k})|^2.$$

It is easy to check that writing $\mathbf{k} = (k_1, k_2)$, we have

$$\hat{\chi}_r(\mathbf{k}) = \frac{\sin(2\pi r k_1)}{\pi k_1} \cdot \frac{\sin(2\pi r k_2)}{\pi k_2},$$

with the understanding that if $k_i = 0$, then the corresponding factor on the right hand side is replaced by the term $2r$. It follows that if $k_1 k_2 \neq 0$, then

$$\int_0^{1/2} |\hat{\chi}_r(\mathbf{k})|^2 \,\mathrm{d}r = \frac{1}{4\pi^4 k_1^2 k_2^2} \int_0^{1/2} (1-\cos(4\pi r k_1))(1-\cos(4\pi r k_2)) \,\mathrm{d}r \ge \frac{1}{8\pi^4 k_1^2 k_2^2}.$$

On the other hand, if $k_1\ne0$, then

$$\int_0^{1/2}|\hat{\chi}_r(k_1,0)|^2\,\mathrm{d}r=\frac{2}{\pi^2k_1^2}\int_0^{1/2}r^2(1-\cos(4\pi rk_1))\,\mathrm{d}r=\frac{2}{\pi^2k_1^2}\left(\frac{1}{24}-\frac{1}{16\pi^2k_1^2}\right)\ge\frac{1}{8\pi^4k_1^2}.$$

Similarly, if $k_2\ne0$, then

$$\int_0^{1/2}|\hat{\chi}_r(0,k_2)|^2\,\mathrm{d}r\ge\frac{1}{8\pi^4k_2^2}.$$

It follows that

$$\int_0^{1/2}\int_{[0,1)^2}|D[\mathcal{P};B(r,\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}\mathrm{d}r\ge\sum_{\mathbf{0}\ne\mathbf{k}\in\mathbb{Z}^2}a(\mathbf{k})|\hat{U}(\mathbf{k})|^2,\tag{9.9}$$

where

$$a(\mathbf{k})=\frac{1}{8\pi^4\max\{1,k_1^2\}\max\{1,k_2^2\}}.$$

Unfortunately, we need to do a little analysis on the terms $\hat{U}(\mathbf{k})$.

LEMMA 9F. *Suppose that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. Then for any positive real numbers X_1 and X_2, we have*

$$\sum_{\substack{\mathbf{0}\ne\mathbf{k}\in\mathbb{Z}^2\\|k_1|\le X_1\\|k_2|\le X_2}}|\hat{U}(\mathbf{k})|^2\ge NX_1X_2-N^2.$$

PROOF. Let $K_1=[X_1]$ and $K_2=[X_2]$. Then it clearly suffices to show that

$$\sum_{\substack{\mathbf{k}\in\mathbb{Z}^2\\|k_1|\le K_1\\|k_2|\le K_2}}|\hat{U}(\mathbf{k})|^2\ge N(K_1+1)(K_2+1).$$

Note first of all that

$$\sum_{\substack{\mathbf{k}\in\mathbb{Z}^2\\|k_1|\le K_1\\|k_2|\le K_2}}|\hat{U}(\mathbf{k})|^2\ge\sum_{\substack{\mathbf{k}\in\mathbb{Z}^2\\|k_1|\le K_1\\|k_2|\le K_2}}\left(1-\frac{|k_1|}{K_1+1}\right)\left(1-\frac{|k_2|}{K_2+1}\right)|\hat{U}(\mathbf{k})|^2.\tag{9.10}$$

We next write $\mathcal{P}=\{\mathbf{p}_1,\ldots,\mathbf{p}_N\}$, so that

$$\hat{U}(\mathbf{k})=\sum_{n=1}^N\mathrm{e}^{-2\pi\mathrm{i}\langle\mathbf{p}_n,\mathbf{k}\rangle}\qquad\text{and}\qquad|\hat{U}(\mathbf{k})|^2=\sum_{m=1}^N\sum_{n=1}^N\mathrm{e}^{2\pi\mathrm{i}\langle\mathbf{p}_m-\mathbf{p}_n,\mathbf{k}\rangle}.$$

Substituting into (9.10) and changing the order of summation, we obtain

$$\sum_{\substack{\mathbf{k}\in\mathbb{Z}^2\\|k_1|\le K_1\\|k_2|\le K_2}}|\hat{U}(\mathbf{k})|^2\ge\sum_{m=1}^N\sum_{n=1}^N\Delta_{K_1+1}(p_{m1}-p_{n1})\Delta_{K_2+1}(p_{m2}-p_{n2}),\tag{9.11}$$

where $\mathbf{p}_m=(p_{m1},p_{m2})$, $\mathbf{p}_n=(p_{n1},p_{n2})$ and

$$\Delta_K(x)=\sum_{\substack{k\in\mathbb{Z}\\|k|\le K}}\left(1-\frac{|k|}{K}\right)\mathrm{e}^{2\pi\mathrm{i}xk}=\frac{1}{K}\left(\frac{\sin(\pi Kx)}{\sin(\pi x)}\right)^2$$

denotes the Fejér kernel. Note that the summands on the right hand side of (9.11) are non-negative, and the diagonal terms with $m = n$ contribute the amount $N(K_1+1)(K_2+1)$. ♣

Let X be a parameter to be chosen later. For every real number $x \in [1, X]$, let

$$\mathcal{R}(x) = [-x, x] \times \left[-\frac{X}{x}, \frac{X}{x}\right]$$

be a rectangle centred at the origin and with area $4X$, and let $\chi_{\mathcal{R}(x)}$ denote its characteristic function. For every $\mathbf{k} \in \mathbb{Z}^2$, let

$$b(\mathbf{k}) = \frac{\mathrm{e}}{4\pi^4 X^2} \int_1^X \chi_{\mathcal{R}(x)}(\mathbf{k}) \,\frac{\mathrm{d}x}{x}.$$

By Lemma 9F, we have

$$\sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{R}(x)} |\hat{U}(\mathbf{k})|^2 \geq NX - N^2.$$

It follows that

$$\begin{aligned}\sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} b(\mathbf{k}) |\hat{U}(\mathbf{k})|^2 &= \frac{\mathrm{e}}{4\pi^4 X^2} \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} |\hat{U}(\mathbf{k})|^2 \int_1^X \chi_{\mathcal{R}(x)}(\mathbf{k}) \,\frac{\mathrm{d}x}{x} \\ &= \frac{\mathrm{e}}{4\pi^4 X^2} \int_1^X \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{R}(x)} |\hat{U}(\mathbf{k})|^2 \,\frac{\mathrm{d}x}{x} \geq \frac{\mathrm{e}}{4\pi^4 X^2} (NX - N^2) \log X. \end{aligned} \tag{9.12}$$

Elementary calculation gives $b(\mathbf{k}) \leq a(\mathbf{k})$ for every $\mathbf{k} \in \mathbb{Z}^2$. Noting this, taking $X = 2N$ and combining (9.9) and (9.12), we obtain

$$\int_0^{1/2} \int_{[0,1)^2} |D[\mathcal{P}; B(r, \mathbf{x})]|^2 \,\mathrm{d}\mathbf{x}\mathrm{d}r \gg \log N.$$

Theorem 1V follows immediately.

We complete this section by indicating a proof of Theorem 10 in the case when A is a disc.

THEOREM 10S. *For every distribution $\mathcal{P}$ of N points in the torus $[0,1)^2$, there exists a disc B in $[0,1)^2$ such that*

$$|D[\mathcal{P}; B]| \gg N^{\frac{1}{4}}.$$

For any real number r satisfying $0 \leq r \leq \frac{1}{2}$, let $B(r)$ denote the disc of radius r and centred at the origin, and let χ_r denote the characteristic function of $B(r)$. Furthermore, for every vector $\mathbf{x} \in [0,1)^2$, let

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

denote the image of $B(r)$ under translation by $\mathbf{x}$. Then

$$\int_{[0,1)^2} |D[\mathcal{P}; B(r, \mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^2} |\hat{\chi}_r(\mathbf{k})|^2 |\hat{U}(\mathbf{k})|^2.$$

It can be shown that

$$\hat{\chi}_r(\mathbf{k}) = \frac{r}{|\mathbf{k}|} J_1(2\pi r |\mathbf{k}|),$$

where J_1 is a Bessel function of the first kind. This Bessel function oscillates, but it can be shown that

$$\int_0^{1/2} |\hat{\chi}_r(\mathbf{k})|^2 \,\mathrm{d}r \gg |\mathbf{k}|^{-3},$$

so that

$$\int_0^{1/2}\int_{[0,1)^2}|D[\mathcal{P};B(r,\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}\mathrm{d}r \gg \sum_{\mathbf{0}\neq\mathbf{k}\in\mathbb{Z}^2}|\mathbf{k}|^{-3}|\hat{U}(\mathbf{k})|^2. \tag{9.13}$$

We need the following analogue of Lemma 9F which we state without proof.

LEMMA 9G. *Suppose that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. Then for any convex set $\mathcal{C}$ symmetric about the origin, we have*

$$\sum_{\substack{\mathbf{0}\neq\mathbf{k}\in\mathbb{Z}^2\\ \mathbf{k}\in\mathcal{C}}}|\hat{U}(\mathbf{k})|^2 \geq \frac{1}{4}N\mu(\mathcal{C})-N^2. \tag{9.14}$$

We now apply Lemma 9G with

$$\mathcal{C}=\{\mathbf{t}\in\mathbb{R}^2:|\mathbf{t}|\leq 2\sqrt{N}\}, \tag{9.15}$$

and note that

$$\sum_{\mathbf{0}\neq\mathbf{k}\in\mathbb{Z}^2}|\mathbf{k}|^{-3}|\hat{U}(\mathbf{k})|^2 \gg N^{-\frac{3}{2}}\sum_{\substack{\mathbf{0}\neq\mathbf{k}\in\mathbb{Z}^2\\ |\mathbf{k}|\leq 2\sqrt{N}}}|\hat{U}(\mathbf{k})|^2. \tag{9.16}$$

Note that $\mu(\mathcal{C})=4\pi N$. It follows on combining (9.13)–(9.16) that

$$\int_0^{1/2}\int_{[0,1)^2}|D[\mathcal{P};B(r,\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}\mathrm{d}r \gg N^{\frac{1}{2}}.$$

Theorem 10S follows immediately.

10. An Integral Geometric Approach

The technique of Alexander is based on the following well known result in integral geometry. There is a motion invariant Borel measure μ_K on the hyperplanes h of euclidean space $\mathbb{R}^K$ such that

$$|\mathbf{u}-\mathbf{v}|=\frac{1}{2}\mu_K(\{h:h\text{ cuts }\overline{\mathbf{uv}}\}), \tag{10.1}$$

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 (10.1) leads to a representation of $I(\tau)$ as an integral with respect to the measure μ_K, of the form

$$I(\tau)=\int_{\mathcal{H}_K}\tau(h^+)\tau(h^-)\,\mathrm{d}\mu_K(h), \tag{10.2}$$

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 (10.1), we have

$$I(\tau)=\frac{1}{2}\int_{\mathcal{H}_K}\int_{\mathbb{R}^K}\int_{\mathbb{R}^K}\chi(\mathbf{u},\mathbf{v},h)\,\mathrm{d}\tau(\mathbf{u})\mathrm{d}\tau(\mathbf{v})\mathrm{d}\mu_K(h), \tag{10.3}$$

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

$$\int_{\mathbb{R}^K}\int_{\mathbb{R}^K}\chi(\mathbf{u},\mathbf{v},h)\,\mathrm{d}\tau(\mathbf{u})\mathrm{d}\tau(\mathbf{v}). \tag{10.4}$$

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 (10.4) must be equal to $2\tau(h^+)\tau(h^-)$. Substituting this into (10.3) leads immediately to the formula (10.2).

Consider the special case when $\tau(\mathbb{R}^K)=0$. It is easy to see that $\tau(h^+)+\tau(h^-)=\tau(\mathbb{R}^K)$ for almost all hyperplanes h in $\mathbb{R}^K$, and so it follows from (10.2) that

$$-I(\tau)=\int_{\mathcal{H}_K}|\tau(h^+)|^2\,\mathrm{d}\mu_K(h)\geq 0.$$

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 shall prove the following variant of Theorem 19.

THEOREM 19V. *We have*

$$-I(\sigma)=\int_{\mathcal{H}_2}|\sigma(h^+)|^2\,\mathrm{d}\mu_2(h)\geq\frac{1}{128}N^{\frac{1}{2}}.$$

Using Fubini's theorem, we can write

$$\begin{aligned}I(\sigma_1-\sigma_2)&=\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}(\sigma_1-\sigma_2)(\mathbf{x})\mathrm{d}(\sigma_1-\sigma_2)(\mathbf{y})\\&=\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\mathrm{d}\sigma_1(\mathbf{y})+\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_2(\mathbf{x})\mathrm{d}\sigma_2(\mathbf{y})\\&\quad-2\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\mathrm{d}\sigma_2(\mathbf{y})\\&=I(\sigma_1)+I(\sigma_2)-2\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}|\mathbf{x}-\mathbf{y}|\,\mathrm{d}\sigma_1(\mathbf{x})\mathrm{d}\sigma_2(\mathbf{y}).\end{aligned} \tag{10.5}$$

Immediately, we have two problems. 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 $I(\sigma_2)$. Also, to study the last term on the right hand side of (10.5), we clearly need to introduce an extra functional.

Suppose that τ,τ' are signed Borel measures with compact support in euclidean space $\mathbb{R}^K$. We shall consider the functional

$$J(\tau,\tau')=\int_{\mathbb{R}^K}\int_{\mathbb{R}^K}|\mathbf{u}-\mathbf{v}|\,\mathrm{d}\tau(\mathbf{u})\mathrm{d}\tau'(\mathbf{v}).$$

To handle the first problem, we introduce a discrete measure Φ in the set $\mathbb{R}$ with support $\{r_1,\ldots,r_\ell\}$ such that

$$\sum_{t=1}^{\ell}|\Phi(r_t)|=1, \tag{10.6}$$

and consider the product measure $\sigma \times \Phi$ on $\mathbb{R}^3$, defined by

$$\sigma \times \Phi = \sum_{t=1}^{\ell} \Phi(r_t)\sigma^{(t)}, \tag{10.7}$$

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

$$\sigma^{(t)}(S, r_t) = \sigma(S) \tag{10.8}$$

for every Borel set S in $\mathbb{R}^2$.

LEMMA 10A. *For every $t = 1, \ldots, \ell$, we have $\sigma^{(t)}(\mathbb{R}^3) = 0$. Furthermore, we have $I(\sigma^{(t)}) = I(\sigma)$.*

PROOF. The first assertion follows immediately from

$$\sigma^{(t)}(\mathbb{R}^3) = \sigma^{(t)}(\mathbb{R}^2 \times \{r_t\}) = \sigma(\mathbb{R}^2).$$

On the other hand, we have

$$I(\sigma^{(t)}) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} |(\mathbf{x}, r_t) - (\mathbf{y}, r_t)| \, \mathrm{d}\sigma^{(t)}(\mathbf{x}, r_t)\mathrm{d}\sigma^{(t)}(\mathbf{y}, r_t) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} |\mathbf{x} - \mathbf{y}| \, \mathrm{d}\sigma(\mathbf{x})\mathrm{d}\sigma(\mathbf{y}) = I(\sigma). \quad \clubsuit$$

LEMMA 10B. *Suppose that $|a_1| + \ldots + |a_\ell| = 1$. Then*

$$-I\left(\sum_{t=1}^{\ell} a_t\sigma^{(t)}\right) \le -\sum_{t=1}^{\ell} |a_t| I(\sigma^{(t)}).$$

PROOF. Suppose first of all that $a_1, \ldots, a_\ell$ are all non-negative. Then it follows from the first assertion of Lemma 10A that

$$-I\left(\sum_{t=1}^{\ell} a_t\sigma^{(t)}\right) = \int_{\mathcal{H}_2} \left(\sum_{t=1}^{\ell} a_t\sigma^{(t)}(h^+)\right)^2 \mathrm{d}\mu_2(h).$$

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 (10.8), is essentially the same as the measure σ in $\mathbb{R}^2$. Since

$$\left(\sum_{t=1}^{\ell} a_t\sigma^{(t)}(h^+)\right)^2 \le \left(\sum_{t=1}^{\ell} a_t\right)\left(\sum_{t=1}^{\ell} a_t|\sigma^{(t)}(h^+)|^2\right)$$

by the Cauchy-Schwarz inequality, it follows that

$$-I\left(\sum_{t=1}^{\ell} a_t\sigma^{(t)}\right) \le \left(\sum_{t=1}^{\ell} a_t\right) \sum_{t=1}^{\ell} a_t \int_{\mathcal{H}_2} |\sigma^{(t)}(h^+)|^2 \, \mathrm{d}\mu_2(h) = -\sum_{t=1}^{\ell} a_t I(\sigma^{(t)}),$$

again as a consequence of the first assertion of Lemma 10A. The general case follows on noting that if $a_t < 0$, then $a_t\sigma^{(t)} = |a_t|(-\sigma^{(t)})$ and $I(-\sigma^{(t)}) = I(\sigma^{(t)})$. ♣

LEMMA 10C. *We have $-I(\sigma \times \Phi) \le -I(\sigma)$.*

PROOF. It follows from (10.7), (10.6), Lemma 10B and the second assertion of Lemma 10A that

$$-I(\sigma \times \Phi) = -I\left(\sum_{t=1}^{\ell} \Phi(r_t)\sigma^{(t)}\right) \le -\sum_{t=1}^{\ell} |\Phi(r_t)| I(\sigma^{(t)}) = -\sum_{t=1}^{\ell} |\Phi(r_t)| I(\sigma) = -I(\sigma),$$

where we use (10.6) again in the last step. ♣

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, as in (10.5), we have

$$I(\sigma\times\Phi)=I(\nu_1-\nu_2)=I(\nu_1)+I(\nu_2)-2J(\nu_1,\nu_2),$$

in view of Fubini's theorem. In other words,

$$-I(\sigma\times\Phi)=-I(\nu_1)-I(\nu_2)+2J(\nu_1,\nu_2). \tag{10.9}$$

Consider the product measure $\nu_2=\sigma_2\times\Phi$ in $\mathbb{R}^3$. Analogous to (10.7), 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 $U\times\{r_t\}$. Furthermore, $\sigma_2^{(t)}(S,r_t)=\sigma_2(S)$ 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

$$\sum_{t=1}^{\ell}\Phi(r_t)=0, \tag{10.10}$$

then $\nu_2(\mathbb{R}^3)=0$, and so

$$-I(\nu_2)=\int_{\mathcal{H}_3}|\nu_2(h^+)|^2\,\mathrm{d}\mu_3(h)\geq 0. \tag{10.11}$$

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 10D. *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. The measure ν_1 is supported by the points $(\mathbf{p}_i,r_t)$, where $i=1,\ldots,N$ and $t=1,\ldots,\ell$. 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

$$\begin{aligned}
I(\nu_1) &= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \nu_1(\mathbf{p}_i, r_t)\nu_1(\mathbf{p}_j, r_u) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \sigma_1(\mathbf{p}_i)\Phi^{(i)}(\mathbf{p}_i, r_t)\sigma_1(\mathbf{p}_j)\Phi^{(j)}(\mathbf{p}_j, r_u) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \left(\sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_j, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t)\Phi^{(j)}(\mathbf{p}_j, r_u) \right) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} J(\Phi^{(i)}, \Phi^{(j)}) = \sum_{\substack{i=1 \\ i\neq j}}^{N}\sum_{j=1}^{N} J(\Phi^{(i)}, \Phi^{(j)}) + \sum_{i=1}^{N} I(\Phi^{(i)}).
\end{aligned}$$

The result follows, since for every $i = 1, \ldots, N$, we have

$$I(\Phi^{(i)}) = \sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |(\mathbf{p}_i, r_t) - (\mathbf{p}_i, r_u)| \Phi^{(i)}(\mathbf{p}_i, r_t)\Phi^{(i)}(\mathbf{p}_i, r_u) = \sum_{t=1}^{\ell}\sum_{u=1}^{\ell} |r_t - r_u| \Phi(r_t)\Phi(r_u) = I(\Phi). \qquad \clubsuit$$

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

$$J(\Phi^{(i)}, \Phi^{(j)}) = J(\Phi, |\mathbf{p}_i - \mathbf{p}_j|)$$

for every $i, j = 1, \ldots, N$.

We next consider the term $J(\nu_1, \nu_2)$. 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 $t = 1, \ldots, \ell$.

LEMMA 10E. *We have*

$$J(\nu_1, \nu_2) = \sum_{i=1}^{N} \int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|) \, \mathrm{d}\sigma_2(\mathbf{y}).$$

PROOF. Note that $\sigma_1(\mathbf{p}_i)=1$ for every $i=1,\ldots,N$. It follows, similar to the proof of Lemma 10D, that

$$\begin{aligned}
J(\nu_1,\nu_2) &= \sum_{i=1}^{N}\sum_{t=1}^{\ell}\sum_{u=1}^{\ell}\int_{\mathbb{R}^2}|(\mathbf{p}_i,r_t)-(\mathbf{y},r_u)|\sigma_1(\mathbf{p}_i)\Phi(r_t)\Phi(r_u)\,\mathrm{d}\sigma_2(\mathbf{y})\\
&= \sum_{i=1}^{N}\int_{\mathbb{R}^2}\left(\sum_{t=1}^{\ell}\sum_{u=1}^{\ell}|(\mathbf{p}_i,r_t)-(\mathbf{y},r_u)|\Phi(r_t)\Phi(r_u)\right)\mathrm{d}\sigma_2(\mathbf{y})\\
&= \sum_{i=1}^{N}\int_{\mathbb{R}^2}J(\Phi,|\mathbf{p}_i-\mathbf{y}|)\,\mathrm{d}\sigma_2(\mathbf{y}). \qquad \clubsuit
\end{aligned}$$

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

$$\begin{aligned}
J(\Phi,d) &= d\sum_{t=1}^{\ell}\sum_{u=1}^{\ell}\Phi(r_t)\Phi(r_u)+\sum_{k=1}^{\infty}\binom{\frac{1}{2}}{k}\left(\sum_{t=1}^{\ell}\sum_{u=1}^{\ell}|r_t-r_u|^{2k}\Phi(r_t)\Phi(r_u)\right)d^{-2k+1}\\
&= d\left(\sum_{t=1}^{\ell}\Phi(r_t)\right)^2+\sum_{k=1}^{\infty}\binom{\frac{1}{2}}{k}I^{(2k)}(\Phi)d^{-2k+1},
\end{aligned}$$

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 (10.10), we have

$$J(\Phi,d)=\sum_{k=1}^{\infty}\binom{\frac{1}{2}}{k}I^{(2k)}(\Phi)d^{-2k+1}. \tag{10.12}$$

Let us summarize the various restrictions on the functional Φ so far. We have assumed that

$$\sum_{t=1}^{\ell}|\Phi(r_t)|=1 \qquad \text{and} \qquad \sum_{t=1}^{\ell}\Phi(r_t)=0.$$

On the other hand, it follows from (10.12) 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 10D that

$$-I(\nu_1)\geq -NI(\Phi). \tag{10.13}$$

LEMMA 10F. *Suppose that*

$$\sum_{t=1}^{\ell}\Phi(r_t)=0 \qquad \textit{and} \qquad \sum_{t=1}^{\ell}r_t\Phi(r_t)=0.$$

Then $I^{(2)}(\Phi)=0$.

PROOF. Note simply that

$$\begin{aligned} I^{(2)}(\Phi) &= \sum_{t=1}^{\ell}\sum_{u=1}^{\ell}|r_t-r_u|^2\Phi(r_t)\Phi(r_u) = \sum_{t=1}^{\ell}\sum_{u=1}^{\ell}(r_t^2-2r_tr_u+r_u^2)\Phi(r_t)\Phi(r_u) \\ &= \sum_{t=1}^{\ell}\left(\sum_{u=1}^{\ell}\Phi(r_u)\right)r_t^2\Phi(r_t) - 2\left(\sum_{t=1}^{\ell}r_t\Phi(r_t)\right)\left(\sum_{u=1}^{\ell}r_u\Phi(r_u)\right) + \sum_{u=1}^{\ell}\left(\sum_{t=1}^{\ell}\Phi(r_t)\right)r_u^2\Phi(r_u). \quad \clubsuit \end{aligned}$$

We therefore need

$$\sum_{t=1}^{\ell}|\Phi(r_t)| = 1 \qquad \text{and} \qquad \sum_{t=1}^{\ell}\Phi(r_t) = 0 \qquad \text{and} \qquad \sum_{t=1}^{\ell}r_t\Phi(r_t) = 0. \tag{10.14}$$

Then (10.11) holds. It follows from Lemma 10C and (10.9) that if (10.13) holds, then we need a bound of the form

$$-I(\Phi) \geq c_1 N^{-\frac{1}{2}}, \tag{10.15}$$

as well as a bound of the form

$$J(\nu_1,\nu_2) \geq -c_2 N^{\frac{1}{2}}, \tag{10.16}$$

where c_1 and c_2 are positive constants satisfying $c_1 > 2c_2$.

The conditions (10.14) require that the measure Φ in $\mathbb{R}$ is supported by at least three points. The measure $\tilde{\Phi}$ in $\mathbb{R}$, defined by $\ell = 3$ and with support $\{0, \pm N^{-\frac{1}{2}}\}$, such that

$$\tilde{\Phi}(0) = \frac{1}{2} \qquad \text{and} \qquad \tilde{\Phi}(\pm N^{-\frac{1}{2}}) = -\frac{1}{4},$$

will satisfy (10.14) and give (10.15) for some constant $c_1 > 0$. Furthermore, it can be shown that $J(\tilde{\Phi}, d) \leq 0$ for every real number $d \geq 0$, so that (10.13) holds. While we can also establish (10.16) 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

$$\Phi(0) = \frac{1}{2} \qquad \text{and} \qquad \Phi(\pm\alpha N^{-\frac{1}{2}}) = -\frac{1}{4},$$

where α is a positive real number. Clearly the conditions (10.14) are satisfied. We shall determine a suitable value for α later. In the mean time, elementary calculations will give the following result.

LEMMA 10G. *We have*

$$I(\Phi) = -\frac{\alpha}{4}N^{-\frac{1}{2}}.$$

Furthermore, for every $k = 1, 2, 3, \ldots$, we have

$$I^{(2k)}(\Phi) = \frac{1}{8}\alpha^{2k}(4^k - 4)N^{-k}.$$

LEMMA 10H. *Suppose that $d \geq 4\alpha N^{-\frac{1}{2}}$. Then*

$$|J(\Phi, d)| \leq \frac{3}{16}\alpha^4 N^{-2} d^{-3}.$$

PROOF. It is easy to check that if $d \geq 4\alpha N^{-\frac{1}{2}}$, then the series (10.12) for $J(\Phi, d)$ is a convergent alternating series, since by Lemma 10G, 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 $\binom{1/2}{k}$ is decreasing in magnitude and alternating in sign. Furthermore, we have $I^{(2)}(\Phi) = 0$, and so

$$|J(\Phi,d)| \le \left| \binom{\frac{1}{2}}{k} I^{(4)}(\Phi) d^{-3} \right| = \frac{3}{16}\alpha^4 N^{-2} d^{-3}. \qquad \clubsuit$$

LEMMA 10J. *The function $-J(\Phi,d)$ is positive and decreasing for $d \ge 0$, with*

$$-J(\Phi,0) = \frac{\alpha}{4} N^{-\frac{1}{2}}.$$

PROOF. It is easy to check that

$$16J(\Phi,d) = 6d - 8(d^2+\alpha^2N^{-1})^{\frac{1}{2}} + 2(d^2+4\alpha^2N^{-1})^{\frac{1}{2}}.$$

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 (10.16), we refer to Lemma 10E 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 10J that $-J(\Phi,|\mathbf{p}_i - \mathbf{y}|) \ge 0$ for every $\mathbf{y} \in \mathbb{R}^2$. Hence

$$-\int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\sigma_2(\mathbf{y}) \le -N\int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\mu(\mathbf{y}) = -2\pi N \int_0^\infty J(\Phi,r) r\, \mathrm{d}r.$$

By Lemma 10J, we have

$$-\int_0^{4\alpha N^{-1/2}} J(\Phi,r) r\, \mathrm{d}r \le \frac{\alpha}{4} N^{-\frac{1}{2}} \int_0^{4\alpha N^{-1/2}} r\, \mathrm{d}r = 2\alpha^3 N^{-\frac{3}{2}}.$$

By Lemma 10H, we have

$$-\int_{4\alpha N^{-1/2}}^\infty J(\Phi,r) r\, \mathrm{d}r \le \int_{4\alpha N^{-1/2}}^\infty \frac{3}{16}\alpha^4 N^{-2} r^{-2}\, \mathrm{d}r = \frac{3}{64}\alpha^3 N^{-\frac{3}{2}}.$$

It follows that

$$-\int_{\mathbb{R}^2} J(\Phi, |\mathbf{p}_i - \mathbf{y}|)\, \mathrm{d}\sigma_2(\mathbf{y}) \le \frac{131}{32}\pi\alpha^3 N^{-\frac{1}{2}}.$$

Combining this with Lemma 10E gives

$$J(\nu_1,\nu_2) \ge -\frac{131}{32}\pi\alpha^3 N^{\frac{1}{2}}. \tag{10.17}$$

Combining Lemma 10C, (10.9), (10.11), (10.13), Lemma 10G and (10.17), we conclude that

$$|I(\sigma)| \ge \frac{\alpha}{4} N^{\frac{1}{2}} - \frac{131}{16}\pi\alpha^3 N^{\frac{1}{2}}.$$

Choosing $\alpha = \frac{1}{16}$ gives

$$|I(\sigma)| \ge \frac{1}{128} N^{\frac{1}{2}}$$

and completes the proof of Theorem 19V.

11. 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 $\mathbf{0}$. Suppose that $\mathcal{P}$ is a distribution of N points in U. For every measurable set B in $\mathbb{R}^2$, let $Z[\mathcal{P};B]$ denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

$$D^*[\mathcal{P};B]=Z[\mathcal{P};B]-N\mu(B\cap U),$$

where μ denotes the usual measure in $\mathbb{R}^2$.

For every real number $r\in\mathbb{R}$ and every angle θ satisfying $0\le\theta\le 2\pi$, let $S(r,\theta)$ denote the closed halfplane

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

Here $\mathbf{e}(\theta)=(\cos\theta,\sin\theta)$ and $\mathbf{x}\cdot\mathbf{y}$ denotes the scalar product of $\mathbf{x}$ and $\mathbf{y}$. Furthermore, let

$$R(\theta)=\sup\{r\ge 0:S(r,\theta)\cap U\ne\emptyset\}.$$

The following result is more general than Theorem 21.

THEOREM 21S. *For every natural number $N\ge 2$, there exists a distribution $\mathcal{P}$ of N points in U such that*

$$\int_0^{2\pi}\int_0^{R(\theta)}|D^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\mathrm{d}\theta\ll_U(\log N)^2.$$

The proof of this result, with Theorem 21 as a special case, is in fact motivated by another special case where U is the square $[-\frac{1}{2},\frac{1}{2}]^2$. We shall therefore first show that for every natural number M, there exists a set $\mathcal{P}$ of $N=4M^2+4M+1$ 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 $\mathcal{P}$ of points in V and every measurable set B in $\mathbb{R}^2$, let $Z[\mathcal{P};B]$ denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

$$E^*[\mathcal{P};B]=Z[\mathcal{P};B]-\mu(B\cap V).$$

We shall show that the set

$$\mathcal{P}=\{-M,-M+1,\ldots,-1,0,1,\ldots,M-1,M\}^2$$

of $N=4M^2+4M+1$ integer lattice points in V satisfies

$$\int_0^{2\pi}\int_0^{M(\theta)}|E^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\mathrm{d}\theta\ll M(\log M)^2,\tag{11.1}$$

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 $S(r,\theta)$, and can be rewritten in the form

$$x_1\cos\theta+x_2\sin\theta=r,$$

where $\mathbf{x}=(x_1,x_2)\in\mathbb{R}^2$.

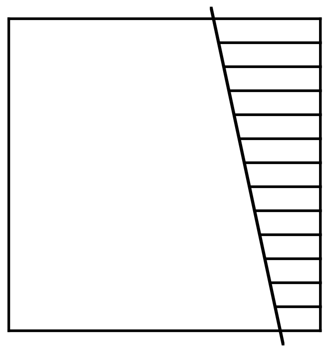
Suppose that $0\le\theta\le\pi/4$. Clearly $M(\theta)=(M+\frac{1}{2})(\cos\theta+\sin\theta)$. We distinguish two cases.

Case 1: If $0 \le r \le (M+\frac{1}{2})(\cos\theta - \sin\theta)$, then it is not difficult to see that $T(r,\theta)$ intersects the top edge $\{(x_1, M+\frac{1}{2}) : |x_1| \le M+\frac{1}{2}\}$ and the bottom edge $\{(x_1, -M-\frac{1}{2}) : |x_1| \le M+\frac{1}{2}\}$ of V. Then

$$S(r,\theta) \cap V = \bigcup_{n=-M}^{M} S(n,V,r,\theta),$$

where, for every $n = -M, \ldots, 0, \ldots, M$,

$$S(n,V,r,\theta) = S(r,\theta) \cap V \cap (\mathbb{R} \times [n-\tfrac{1}{2}, n+\tfrac{1}{2}]).$$



Clearly

$$E^*[\mathcal{P}; S(r,\theta)] = \sum_{n=-M}^{M} E^*[\mathcal{P}; S(n,V,r,\theta)].$$

Now, for every $n = -M, \ldots, 0, \ldots, M$, it is easy to check that

$$Z[\mathcal{P}; S(n,V,r,\theta)] = [M + n\tan\theta - r\sec\theta + 1] \qquad \text{and} \qquad \mu(S(n,V,r,\theta)) = M + n\tan\theta - r\sec\theta + \tfrac{1}{2},$$

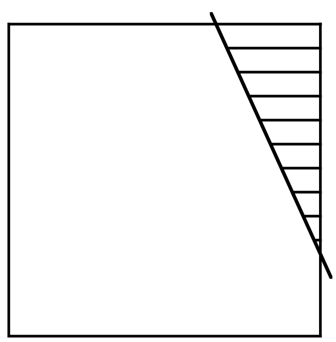
so that

$$E^*[\mathcal{P}; S(n,V,r,\theta)] = -\phi(n\tan\theta - r\sec\theta),$$

where $\phi(z) = z - [z] - \frac{1}{2}$ for every $z \in \mathbb{R}$. It follows that

$$E^*[\mathcal{P}; S(r,\theta)] = -\sum_{n=-M}^{M} \phi(n\tan\theta - r\sec\theta).$$

Case 2: If $(M+\frac{1}{2})(\cos\theta - \sin\theta) \le r \le (M+\frac{1}{2})(\cos\theta+\sin\theta)$, then it is not difficult to see that $T(r,\theta)$ intersects the top edge $\{(x_1, M+\frac{1}{2}) : |x_1| \le M+\frac{1}{2}\}$ and the right edge $\{(M+\frac{1}{2}, x_2) : |x_2| \le M+\frac{1}{2}\}$ of V.



In particular, the line $T(r,\theta)$ intersects the right edge of V at the point

$$(M+\tfrac{1}{2}, -(M+\tfrac{1}{2})\cot\theta + r\csc\theta),$$

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 $E^*[\mathcal{P};S(n,V,r,\theta)] = O(1)$ always. It follows that

$$E^*[\mathcal{P};S(r,\theta)] = -\sum_{\substack{n=-M \\ (*)}}^{M} \phi(n\tan\theta - r\sec\theta) + O(1),$$

where the summation is under the further restriction

$$n \geq -(M+\tfrac{1}{2})\cot\theta + r\csc\theta. \tag{$*$}$$

Note that in Case 1, the restriction $(*)$ is superfluous since it is weaker than the requirement that $n \geq -M$. It follows that for every $r \geq 0$, we have

$$E^*[\mathcal{P};S(r,\theta)] - G[\mathcal{P};r,\theta] \ll 1,$$

where

$$G[\mathcal{P};r,\theta] = -\sum_{\substack{n=-M \\ (*)}}^{M} \phi(n\tan\theta - r\sec\theta).$$

Furthermore, it is easy to check that the Fourier expansion of $G[\mathcal{P};r,\theta]$ is given by

$$\sum_{\nu\neq 0} \frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu} \sum_{\substack{n=-M \\ (*)}}^{M} e(n\nu\tan\theta).$$

However, the restriction $(*)$ prevents us from applying Parseval's theorem.

We are in a similar situation to that encountered in Section 5. 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 confortably accommodate.

Recall that every closed halfplane $S(r,\theta)$ is described in terms of the variables r and θ relative to the origin $\mathbf{0}$. 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 $\theta\in[0,\pi/4]$ and every $r\geq 1$, let

$$T(\mathbf{y};r,\theta) = T(r+y_1\cos\theta+y_2\sin\theta,\theta) \tag{11.2}$$

and

$$S(\mathbf{y};r,\theta) = S(r+y_1\cos\theta+y_2\sin\theta,\theta), \tag{11.3}$$

noting here that $r+y_1\cos\theta+y_2\sin\theta\geq 0$ always. Then

$$E^*[\mathcal{P};S(\mathbf{y};r,\theta)] = E^*[\mathcal{P};S(r+y_1\cos\theta+y_2\sin\theta,\theta)].$$

It is not difficult to see that if we write

$$G[\mathcal{P};\mathbf{y};r,\theta] = -\sum_{\substack{n=-M \\ (*)}}^{M} \phi(n\tan\theta - (r+y_1\cos\theta+y_2\sin\theta)\sec\theta),$$

then

$$E^*[\mathcal{P};S(\mathbf{y};r,\theta)]-G[\mathcal{P};\mathbf{y};r,\theta]\ll\begin{cases}\cot\theta & \text{if } M(\theta)-(2M+1)\sin\theta-1\le r\le M(\theta),\\ 1 & \text{otherwise,}\\ M & \text{trivially,}\end{cases}$$

where the first estimate $\cot\theta$ arises from the fact that we have not modified the extra restriction $(*)$. Note also that $|y_1\cos\theta+y_2\sin\theta|\le 1$, so that if $r\le M(\theta)-(2M+1)\sin\theta-1$, then $T(\mathbf{y};r,\theta)$ intersects the top and bottom edges of V. It follows that

$$\int_0^{\pi/4}\int_1^{M(\theta)}|E^*[\mathcal{P};S(\mathbf{y};r,\theta)]-G[\mathcal{P};\mathbf{y};r,\theta]|\,\mathrm{d}r\mathrm{d}\theta\ll M. \tag{11.4}$$

Now $G[\mathcal{P};\mathbf{y};r,\theta]$ has the Fourier expansion

$$\begin{aligned}&\sum_{\nu\ne 0}\frac{e(-(r+y_1\cos\theta+y_2\sin\theta)\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-M\\(*)}}^{M}e(n\nu\tan\theta)\\ &\qquad=\sum_{\nu\ne 0}\frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-M\\(*)}}^{M}e((n-y_2)\nu\tan\theta)e(-y_1\nu).\end{aligned}$$

It follows that for every $y_2\in[-\frac{1}{2},\frac{1}{2}]$, we have, by Parseval's theorem, that

$$\int_{-1/2}^{1/2}|G[\mathcal{P};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(*)}}^{M}e((n-y_2)\nu\tan\theta)\right|^2=\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(*)}}^{M}e(n\nu\tan\theta)\right|^2,$$

so that

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}|G[\mathcal{P};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\mathrm{d}y_2\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\left|\sum_{\substack{n=-M\\(*)}}^{M}e(n\nu\tan\theta)\right|^2\ll\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\min\{M^2,\|\nu\tan\theta\|^{-2}\}, \tag{11.5}$$

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 11A. *We have*

$$\int_0^{\pi/4}\left(\sum_{\nu=1}^{\infty}\frac{1}{\nu^2}\min\{M^2,\|\nu\tan\theta\|^{-2}\}\right)^{\frac{1}{2}}\mathrm{d}\theta\ll(\log M)^2.$$

PROOF. Since $\tan\theta\asymp\theta$ if $0\le\theta\le\pi/4$, it suffices to show that

$$\int_0^1\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\right)^{\frac{1}{2}}\mathrm{d}\omega\ll(\log M)^2. \tag{11.6}$$

Clearly

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\le\sum_{n=1}^{M^2}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}+1,$$

so that

$$\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\min\{M^2,\|n\omega\|^{-2}\}\right)^{\frac{1}{2}}\le\sum_{n=1}^{M^2}\frac{1}{n}\min\{M,\|n\omega\|^{-1}\}+1. \tag{11.7}$$

Now, for every $n=1,\ldots,M^2$, we have

$$\int_0^1\min\{M,\|n\omega\|^{-1}\}\,\mathrm{d}\omega=2n\int_0^{1/2n}\min\{M,(n\omega)^{-1}\}\,\mathrm{d}\omega\ll\log M. \tag{11.8}$$

Inequality (11.6) now follows on combining (11.7) and (11.8). ♣

By the Cauchy-Schwarz inequality, we have

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}|G[\mathcal{P};\mathbf{y};r,\theta]|\,\mathrm{d}y_1\mathrm{d}y_2\ll\left(\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}|G[\mathcal{P};\mathbf{y};r,\theta]|^2\,\mathrm{d}y_1\mathrm{d}y_2\right)^{\frac{1}{2}}. \tag{11.9}$$

It follows from (11.4), (11.5), (11.9) and Lemma 11A that

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}\int_0^{\pi/4}\int_1^{M(\theta)}|E^*[\mathcal{P};S(\mathbf{y};r,\theta)]|\,\mathrm{d}r\mathrm{d}\theta\mathrm{d}y_1\mathrm{d}y_2\ll M(\log M)^2. \tag{11.10}$$

For every $\theta\in[0,\pi/4]$, every $r\ge1$ 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 $r+y_1\cos\theta+y_2\sin\theta\ge0$, we must have

$$\int_2^{M(\theta)-1}|E^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\le\int_1^{M(\theta)}|E^*[\mathcal{P};S(\mathbf{y};r,\theta)]|\,\mathrm{d}r. \tag{11.11}$$

On the other hand, we have the trivial estimate

$$\left(\int_0^2+\int_{M(\theta)-1}^{M(\theta)}\right)|E^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\ll M. \tag{11.12}$$

It now follows from (11.10)–(11.12) that

$$\int_0^{\pi/4}\int_0^{M(\theta)}|E^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\mathrm{d}\theta\ll M(\log M)^2.$$

Similarly, for $j=1,\ldots,7$, we have

$$\int_{j\pi/4}^{(j+1)\pi/4}\int_0^{M(\theta)}|E^*[\mathcal{P};S(r,\theta)]|\,\mathrm{d}r\mathrm{d}\theta\ll M(\log M)^2.$$

Inequality (11.1) now follows.

Remark. Note 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 $\mathcal{P}$. 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 is the closed disc of unit area and centred at the origin $\mathbf{0}$.

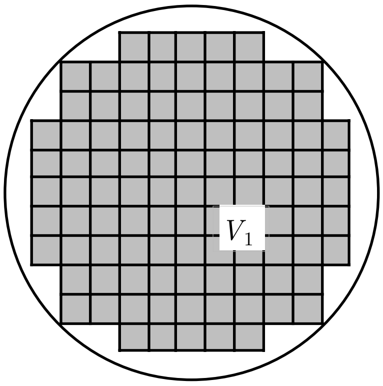
Let N be any given natural number. Again we consider a renormalized version of the problem, and take V to be the closed disc of area N and centred at the origin $\mathbf{0}$. However, if we simply attempt to take all the integer lattice points in V as our set $\mathcal{P}$, then by a famous theorem of Hardy on the number of lattice points in a disc, the number of points of $\mathcal{P}$ can differ from N by an amount sufficiently large to make our task impossible.

Our new idea is to introduce a set $\mathcal{P}$ such that the majority of points of $\mathcal{P}$ 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

$$A(\mathbf{x}) = A(x_1, x_2) = [x_1 - \tfrac{1}{2}, x_1 + \tfrac{1}{2}] \times [x_2 - \tfrac{1}{2}, x_2 + \tfrac{1}{2}];$$

in other words, $A(\mathbf{x})$ is the aligned closed square of unit area and centred at $\mathbf{x}$. Let

$$\mathcal{P}_1 = \{\mathbf{p} \in \mathbb{Z}^2 : A(\mathbf{p}) \subseteq V\} \qquad \text{and} \qquad V_1 = \bigcup_{\mathbf{p} \in \mathcal{P}_1} A(\mathbf{p}).$$



Note that the points of $\mathcal{P}_1$ form the majority of any point set $\mathcal{P}$ 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 $0 = \theta_0 < \theta_1 < \ldots < \theta_{L-1} < \theta_L = 1$ such that for every $j = 1, \ldots, L$, the set $R_j = \{\mathbf{x} \in V_2 : 2\pi\theta_{j-1} \le \arg \mathbf{x} < 2\pi\theta_j\}$ satisfies $\mu(R_j) = 1$. For every $j = 1, \ldots, L$, let $\mathbf{p}_j \in R_j$, and write $\mathcal{P}_2 = \{\mathbf{p}_1, \ldots, \mathbf{p}_L\}$. If we now take

$$\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2, \tag{11.13}$$

then clearly $\mathcal{P}$ contains exactly N points.

For every measurable set B in $\mathbb{R}^2$, let $Z[\mathcal{P}; B]$ denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

$$E^*[\mathcal{P}; B] = Z[\mathcal{P}; B] - \mu(B \cap V).$$

For any halfplane $S(r, \theta)$, the analysis of the discrepancy function

$$E^*[\mathcal{P}; S(r,\theta) \cap V_1] = E^*[\mathcal{P}_1; S(r,\theta) \cap V_1]$$

is essentially similar to our earlier discussion, while the analysis of the discrepancy function

$$E^*[\mathcal{P}; S(r,\theta) \cap V_2] = E^*[\mathcal{P}_2; S(r,\theta) \cap V_2]$$

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 (11.13) satisfies the inequality

$$\int_0^{2\pi} \int_0^M |E^*[\mathcal{P}; S(r,\theta)]| \, \mathrm{d}r \mathrm{d}\theta \ll M (\log M)^2. \tag{11.14}$$

However, if we want to establish the full generality of Theorem 21S, then we have no explicit information on the boundary of V. Extra geometric consideration is then required.

Suppose next that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. For every measurable set B in $[0,1)^2$, let $Z[\mathcal{P};B]$ denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

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

where μ denotes the usual measure in $\mathbb{R}^2$.

Let A be a closed convex polygon in $[0,1)^2$, of diameter less than 1 and with centre of gravity at the origin $\mathbf{0}$. For every real number r satisfying $0\le r\le 1$ and for every angle θ satisfying $0\le\theta\le 2\pi$, let $\mathbf{v}=\theta(\mathbf{u})$ denote

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{11.15}$$

where $\mathbf{v}=(v_1,v_2)$ and $\mathbf{u}=(u_1,u_2)$, and write

$$A(r,\theta)=\{r\mathbf{v}:\mathbf{v}=\theta(\mathbf{u})\text{ for some }\mathbf{u}\in A\}; \tag{11.16}$$

in other words, $A(r,\theta)$ is obtained from A by first rotating anticlockwise by angle θ and then contracting by factor r about the origin $\mathbf{0}$. For every $\mathbf{x}\in U$, let

$$A(\mathbf{x},r,\theta)=\{\mathbf{x}+\mathbf{v}:\mathbf{v}\in A(r,\theta)\}, \tag{11.17}$$

so that $A(\mathbf{x},r,\theta)$ is a similar copy of A, with centre of gravity at $\mathbf{x}$.

Our study of Theorem 22 is motivated by our study of Theorem 21, and is based on the simple observation that a convex polygon is the intersection of a finite number of halfplanes. We shall only briefly discuss the problem 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. For every finite distribution $\mathcal{P}$ of points in V and every measurable set B in V, let $Z[\mathcal{P};B]$ denote the number of points of $\mathcal{P}$ that fall into B, and consider the discrepancy function

$$E[\mathcal{P};B]=Z[\mathcal{P};B]-\mu(B).$$

Let $A\subseteq V$ be a closed convex polygon of diameter less than M and with centre of gravity at the origin $\mathbf{0}$. For every real number r satisfying $0\le r\le 1$, every angle θ satisfying $0\le\theta\le 2\pi$ and every $\mathbf{x}\in V$, we define $A(\mathbf{x},r,\theta)$ in terms of (11.15)–(11.17). To establish Theorem 22 in this special case, it clearly suffices to show that for every natural number $N\ge 2$, there exists a distribution $\mathcal{P}$ of N points in V such that

$$\int_0^1\int_0^{2\pi}\int_V |E[\mathcal{P};A(\mathbf{x},r,\theta)]|\,\mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r\ll_A N(\log N)^2. \tag{11.18}$$

The idea is roughly as follows. Consider a similar copy $A(\mathbf{x},r,\theta)$, where the contraction $r\in[0,1]$, the rotation $\theta\in[0,2\pi]$ and the centre of gravity $\mathbf{x}\in V$ are fixed. Then each edge of $A(\mathbf{x},r,\theta)$ gives rise to a discrepancy of a similar nature to the discrepancy arising from the edge of the halfplane $S(r,\theta)$ in our earlier discussion, 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 (11.18) is to split the integral over V in (11.18) into a sum of integrals over sets whose diameters are very small. This will enable us to use the translation variable $\mathbf{x}$ in the same way as the probabilistic variable $\mathbf{y}$ in our earlier discussion. It can then be shown that the set

$$\mathcal{P}=\{(m-\tfrac{1}{2},n-\tfrac{1}{2}):m,n\in\mathbb{N}\text{ and }1\le m,n\le M\}$$

of $N=M^2$ points in V satisfies the inequality (11.18).

Remark. Note that our argument is again probabilistic in nature. Again, the probabilistic variable $\mathbf{x}$ does not modify the point set in question.

Let us now study the case when rotation is not present. More precisely, suppose that $\mathcal{P}$ is a distribution of N points in the torus $[0,1)^2$. Suppose that A is a closed convex polygon of diameter less than 1 and with centre of gravity at the origin $\mathbf{0}$. For every real number r satisfying $0 \le r \le 1$ and every $\mathbf{x} \in U$, let

$$A(\mathbf{x}, r) = \{\mathbf{x} + r\mathbf{u} : \mathbf{u} \in A\},$$

so that $A(\mathbf{x}, r)$ is a homothetic copy of A, with centre of gravity at $\mathbf{x}$.

In the proof of the special case of Theorem 22 we just discussed, the point set $\mathcal{P}$ is made up of a square lattice. It is clear that the resulting discrepancy function $D[\mathcal{P}; A(\mathbf{x}, r, \theta)]$ can be rather large in magnitude for some values of θ and rather small in magnitude for other values of θ. This observation leads us to consider, in the case of Theorem 23, the possibility of rotating a square lattice to a suitable angle, and then perhaps make some appropriate adjustments near the edge of the square $[0,1)^2$. Rotating a square lattice to a suitable angle presents no difficulties, and we appeal to a result of Davenport on diophantine approximation. However, the analysis of the adjusted point set appears to give rise to an error term too large for the method to succeed.

To overcome this difficulty, we appeal to Roth's probabilistic method first discussed in Section 5, introduce an extra translation variable and consider some average of the discrepancy function over a collection of translated copies of our basic construction. There is still a complication. If the collection of translated copies of the basic construction is too small, then we cannot use Parseval's theorem and study the coefficients arising from the Fourier series of the discrepancy function. If the collection of translated copies of the basic construction is large enough to enable us to use Parseval's theorem in an appropriate way, then we may end up with a point set which does not contain the correct number of points. However, there is a simple technique to overcome this last difficulty.

As before, we rescale and consider instead a distribution $\mathcal{P}$ of N points in the square $V = [0, N^{\frac{1}{2}}]^2$, treated as a torus. Suppose that A is a closed convex polygon of diameter less than $N^{\frac{1}{2}}$ and with centre of gravity at the origin $\mathbf{0}$. Suppose further that A has k sides, with vertices $\mathbf{v}_1, \ldots, \mathbf{v}_k$, where

$$(\mathbf{v}_j - \mathbf{v}_{j-1}) \cdot \mathbf{e}(\theta_j + \pi/2) = |\mathbf{v}_j - \mathbf{v}_{j-1}|,$$

with $0 \le \theta_1 < \ldots < \theta_k < 2\pi$ 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 A 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.

Recall that a real number β is said to be badly approximable if there exists a constant $c = c(\beta) > 0$ such that $\nu\|\nu\beta\| > c(\beta)$ for every positive integer ν. Here $\|x\| = \min_{n\in\mathbb{Z}} |x - n|$ denotes the distance of x from the nearest integer. We need the following result of Davenport on diophantine approximation.

LEMMA 11B. *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.*

An immediate conseqeunce of Lemma 11B is that there exists a real number $\theta \in [0, 2\pi)$ such that the $k+2$ numbers

$$\tan\theta, \tan(\theta + \pi/2), \tan(\theta + \theta_1), \ldots, \tan(\theta + \theta_k)$$

are all finite and badly approximable. We now choose one such value of θ and keep it fixed. We then rotate the square lattice $\Lambda = \mathbb{Z}^2$ anticlockwise by angle θ to obtain the lattice

$$\Lambda_\theta = \{\mathbf{v} : \mathbf{v} = \theta(\mathbf{u}) \text{ for some } \mathbf{u} \in \Lambda\},$$

where $\mathbf{v} = \theta(\mathbf{u})$ is defined by (11.15). 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 1. We then study the discrepancy of the set

$$\mathcal{P}(\mathbf{w}) = (\mathbf{w} + \Lambda_\theta) \cap V \tag{11.19}$$

in V, and show that there exists $\mathbf{w}^* \in \mathbb{Z}^2$ such that

$$\int_0^1 \int_V |E[\mathcal{P}(\mathbf{w}^*); A(\mathbf{x}, r)]|^2 \, \mathrm{d}\mathbf{x}\mathrm{d}r \ll N \log N. \tag{11.20}$$

The set (11.19) may not have exactly N points. By a simple argument, it can be shown that points can be removed from or added to (11.19) in any way to ensure that the number of points is exactly N while not jeopardizing the estimate (11.20). We omit the details here.

References

JR Alexander (1990). Geometric methods in the study of irregularities of distribution, *Combinatorica*, **10**, 115–136.

RC Baker (1999). On irregularities of distribution II, *J. London Math. Soc.*, **59**, 50–64.

J Beck (1981). Balanced two-colourings of finite sets in the square I, *Combinatorica*, **1**, 327–335.

J Beck (1983). On a problem of K.F. Roth concerning irregularities of point distribution, *Invent. Math.*, **74**, 477–487.

J Beck (1987). Irregularities of point distribution I, *Acta Math.*, **159**, 1–49.

J Beck (1988a). Irregularities of point distribution II, *Proc. London Math. Soc.*, **56**, 1–50.

J Beck (1988b). On the discrepancy of convex plane sets, *Monatsh. Math.*, **105**, 91–106.

J Beck (1989). A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution, *Compositio Math.*, **72**, 269–339.

J Beck and WWL Chen (1989). Irregularities of point distribution relative to convex polygons, *Irregularities of Partitions*, pp. 1–22 (Algorithms and Combinatorics **8**, Springer Verlag).

J Beck and WWL Chen (1990). Note on irregularities of distribution II, *Proc. London Math. Soc.*, **61**, 251–272.

J Beck and WWL Chen (1993a). Irregularities of point distribution relative to half-planes I, *Mathematika*, **40**, 102–126.

J Beck and WWL Chen (1993b). Irregularities of point distribution relative to convex polygons II, *Mathematika*, **40**, 127–136.

J Beck and WWL Chen (1997). Irregularities of point distribution relative to convex polygons III, *J. London Math. Soc.*, **56**, 222–230.

WWL Chen (1980). On irregularities of distribution, *Mathematika*, **27**, 153–170.

WWL Chen and MM Skriganov (1999). Davenport's theorem in the theory of irregularities of point distribution (preprint).

H Davenport (1956). Note on irregularities of distribution, *Mathematika*, **3**, 131–135.

G Halász (1981). On Roth's method in the theory of irregularities of point distributions, *Recent Progress in Analytic Number Theory*, vol. 2, pp. 79–94 (Academic Press, London).

JH Halton (1960). On the efficiency of certain quasirandom sequences of points in evaluating multidimensional integrals, *Num. Math.*, **2**, 84–90.

G Károlyi (1995a). Geometric discrepancy theorems in higher dimensions, *Studia Sci. Math. Hungarica*, **30**, 59–94.

G Károlyi (1995b). Irregularities of point distributions relative to homothetic convex bodies, *Monatsh. Math.*, **120**, 247–279.

M Lerch (1904). Question 1547, *L'Intermediaire Math.*, **11**, 144–145.

J Matoušek (1995). Tight upper bounds for the discrepancy of halfspaces, *Discrete Comput. Geom.*, **13**, 593–601.

J Matoušek (1999). *Geometric Discrepancy* (Algorithms and Combinatorics **18**, Springer Verlag).

KF Roth (1954). On irregularities of distribution, *Mathematika*, **1**, 73–79.

KF Roth (1979). On irregularities of distribution III, *Acta Arith.*, **35**, 373–384.

KF Roth (1980). On irregularities of distribution IV, *Acta Arith.*, **37**, 67–75.

WM Schmidt (1968). Irregularities of distribution IV, *Invent. Math.*, **7**, 55–82.

WM Schmidt (1972). Irregularities of distribution VII, *Acta Arith.*, **21**, 45–50.

WM Schmidt (1975). Irregularities of distribution IX, *Acta Arith.*, **27**, 385–396.

WM Schmidt (1977). Irregularities of distribution X, *Number Theory and Algebra*, pp. 311–329 (Academic Press).