Fourier techniques in the theory of irregularities of point distribution

W.W.L. Chen

Department of Mathematics
Macquarie University
Sydney NSW 2109
Australia

Email: wchen@maths.mq.edu.au

Abstract

By the use of two examples, we discuss the techniques of Fourier analysis in the study of problems in irregularities of point distribution. Such techniques include classical Fourier series and transforms, as well as Fourier-Walsh analysis and wavelet analysis. We show also that often the Fourier analysis can be combined with ideas and techniques in number theory, geometry, probability theory, group theory, characters and duality.

1 Introduction

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

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

denote the number of points of the distribution $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, and consider the corresponding discrepancy function

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

Theorem 1.

(i) There exists a positive absolute constant c_1 such that for every positive integer N and every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} > c_1 \log N.$$

(ii) *There exists a positive absolute constant C_1 such that for every integer $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} < C_1 \log N.$$

The lower bound was established by Roth [20] in 1954, while the upper bound was established by Davenport [12] in 1956.

Indeed, the lower bound of Theorem 1 can be extended to point distributions in the k-dimensional unit cube for arbitrary $k \geq 2$ without any extra difficulty, as shown in Roth [20] with lower bound $c_1(k)(\log N)^{k-1}$. However, ideas different from those of Davenport are necessary to extend the upper bound of Theorem 1 to the k-dimensional unit cube for arbitrary $k \geq 2$. Some of these ideas will be discussed in this article.

Suppose that $\mathcal{Q}$ is a distribution of N points in the unit square $[-\frac{1}{2}, \frac{1}{2}]^2$. For real numbers $r \geq 0$ and $\theta \in [0, 2\pi]$, let $A(r,\theta)$ denote the square $[-r,r]^2$ rotated anticlockwise by an angle θ. Furthermore, for every vector $\mathbf{x} \in \mathbf{R}^2$, let

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

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

$$Z[\mathcal{Q}; A(r,\theta,\mathbf{x})] = |\mathcal{Q} \cap A(r,\theta,\mathbf{x})|$$

denote the number of points of the distribution $\mathcal{Q}$ that fall into the similar square $A(r,\theta,\mathbf{x})$, and consider the corresponding discrepancy function

$$D[\mathcal{Q}; A(r,\theta,\mathbf{x})] = Z[\mathcal{Q}; A(r,\theta,\mathbf{x})] - N\mu(A(r,\theta,\mathbf{x}) \cap [-\tfrac{1}{2}, \tfrac{1}{2}]^2),$$

where μ denotes the usual Lebesgue area measure on $\mathbf{R}^2$.

Theorem 2.

(i) *There exists a positive absolute constant c_2 such that for every positive integer N and every distribution $\mathcal{Q}$ of N points in the unit square $[-\frac{1}{2}, \frac{1}{2}]^2$, we have*

$$\int_0^{1/4} \int_0^{2\pi} \int_{\mathbf{R}^2} |D[\mathcal{Q}; A(r,\theta,\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r > c_2 N^{1/2}.$$

(ii) *There exists a positive absolute constant C_2 such that for every positive integer N, there exists a distribution $\mathcal{Q}$ of N points in the unit square $[-\frac{1}{2}, \frac{1}{2}]^2$ such that*

$$\int_0^{1/4} \int_0^{2\pi} \int_{\mathbf{R}^2} |D[\mathcal{Q}; A(r,\theta,\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r < C_2 N^{1/2}.$$

The lower bound was established by Beck [1] in 1987, while the corresponding upper bound was established by Beck and Chen [2] in 1990, both as special cases of more general results in arbitrary dimensions $k \geq 2$.

The purpose of this article is to discuss the ideas behind some of the proofs of Theorems 1 and 2, paying special attention to the Fourier techniques involved. Such Fourier techniques include classical Fourier series and transforms, as well as Fourier-Walsh analysis and wavelet analysis. We show also that often the Fourier analysis can be combined with ideas and techniques in number theory, geometry, probability theory, group theory, characters and duality.

The paper is organized as follows. In Sections 2 – 3, we shall use Fourier transform techniques to establish Theorem 2. The basic techniques and the lower bound will be discussed in Section 2, while the upper bound will be discussed in Section 3. In Sections 4 – 6, we study the upper bound of Theorem 1. We briefly discuss Davenport's ideas in Section 4, together with a different approach by Beck and Chen [3]. In Section 5, we make use of the periodicity property of some point sets and study the same problem using classical Fourier series. Then in Section 6, we make use of the group structure of the same point sets and revisit the problem using Fourier-Walsh techniques. We then turn our attention to the lower bound of Theorem 1. We briefly discuss Roth's ideas in Section 7, and demonstrate a wavelet approach by Pollington [18] in Section 8.

Notation. As usual, $\mathbf{Z}$, $\mathbf{N}$, $\mathbf{Q}$ and $\mathbf{R}$ denote respectively the set of all integers, the set of all positive integers, the set of all rational numbers and the set of all real numbers. For the sake of convenience, we shall use $\mathbf{N}_0$ to denote the set of all non-negative integers. Suppose that $x \in \mathbf{R}$. We denote by $[x]$ the integer part of x, so that $[x]$ is equal to the unique integer $n \in \mathbf{Z}$ satisfying $n \leq x < n+1$. We denote by $\{x\} = x - [x]$ the fractional part of x. Furthermore, $\psi(x)$ denotes the sawtooth function, defined by $\psi(x) = x - [x] - 1/2$ when $x \notin \mathbf{Z}$ and by $\psi(x) = 0$ when $x \in \mathbf{Z}$. Throughout, for any functions f and h and any non-negative real valued function g, we use $f = O(g)$ to denote the existence of a positive constant c such that $|f| \leq cg$, and $f = h + O(g)$ to denote the existence of a positive constant c such that $|f - h| \leq cg$. The constant c may depend on some parameters in the argument, but will never depend on the number of points of the distribution under discussion. Finally, we shall also use the Vinogradov notation $f \ll g$ to represent the inequality $f = O(g)$, and the notation $f \asymp g$ to indicate that $f = O(g)$ and $g = O(f)$ both hold.

2 Beck's Fourier Transform Approach

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

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

denotes the number of points of $\mathcal{Q}$ 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 $\mathbf{R}^2$, restricted to the square $[-\frac{1}{2},\frac{1}{2}]^2$, so that for every measurable set $B \subseteq \mathbf{R}^2$,

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

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

$$D_0(B) = Z_0(B) - N\mu_0(B) = |\mathcal{Q} \cap B| - N\mu(B \cap [-\tfrac{1}{2},\tfrac{1}{2}]^2)$$

represents the discrepancy of the part of B which lies in $[-\frac{1}{2},\frac{1}{2}]^2$.

For real numbers $r \geq 0$ and $\theta \in [0, 2\pi]$, let $\chi_{r,\theta}$ denote the characteristic function of the rotated square $A(r,\theta)$. Consider the function

$$F_{r,\theta} = \chi_{r,\theta} * (\mathrm{d}Z_0 - N\mathrm{d}\mu_0), \tag{1}$$

where $f * g$ denotes the convolution of the functions f and g, so that for every $\mathbf{x} \in \mathbf{R}^2$,

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

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

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

It follows that

$$\begin{aligned}&\int_{\mathbf{R}^2} \chi_{r,\theta}(\mathbf{x}-\mathbf{y})(\mathrm{d}Z_0(\mathbf{y}) - N\mathrm{d}\mu_0(\mathbf{y}))\\ &\qquad = |\mathcal{Q} \cap A(r,\theta,\mathbf{x})| - N\mu(A(r,\theta,\mathbf{x}) \cap [-\tfrac{1}{2},\tfrac{1}{2}]^2),\end{aligned}$$

and therefore

$$F_{r,\theta}(\mathbf{x}) = Z_0(A(r,\theta,\mathbf{x})) - N\mu_0(A(r,\theta,\mathbf{x})) = D_0(A(r,\theta,\mathbf{x})) \tag{2}$$

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

We now appeal to the theory of Fourier transforms. Let $L_1(\mathbf{R}^2)$ denote the set of all measurable complex valued functions f that are absolutely integrable over $\mathbf{R}^2$, with Fourier transform $\widehat{f}$ defined for every $\mathbf{t} \in \mathbf{R}^2$ by

$$\widehat{f}(\mathbf{t}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} f(\mathbf{x}) \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}}\,\mathrm{d}\mathbf{x}.$$

It is well known that for any two functions $f, g \in L_1(\mathbf{R}^2)$, we have $f*g \in L_1(\mathbf{R}^2)$ and the Fourier transforms $\widehat{f}$ and $\widehat{g}$ satisfy

$$\widehat{f*g} = \widehat{f}\widehat{g}. \tag{3}$$

Let $L_2(\mathbf{R}^2)$ denote the set of all measurable complex valued functions f that are square integrable over $\mathbf{R}^2$. Then the Parseval-Plancherel theorem states that for every function $f \in L_1(\mathbf{R}^2) \cap L_2(\mathbf{R}^2)$, the Fourier transform $\widehat{f} \in L_2(\mathbf{R}^2)$ and satisfies

$$\int_{\mathbf{R}^2} |f(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} = \int_{\mathbf{R}^2} |\widehat{f}(\mathbf{t})|^2 \, \mathrm{d}\mathbf{t}. \tag{4}$$

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

$$\phi(\mathbf{t}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}} \, \mathrm{d}D_0(\mathbf{x}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}} (\mathrm{d}Z_0(\mathbf{x}) - N \mathrm{d}\mu_0(\mathbf{x})). \tag{5}$$

Then it follows from (1) and (3)–(5) that

$$\int_{\mathbf{R}^2} |F_{r,\theta}(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} = \int_{\mathbf{R}^2} |\widehat{F}_{r,\theta}(\mathbf{t})|^2 \, \mathrm{d}\mathbf{t} = \int_{\mathbf{R}^2} |\widehat{\chi}_{r,\theta}(\mathbf{t})|^2 |\phi(\mathbf{t})|^2 \, \mathrm{d}\mathbf{t}. \tag{6}$$

Note that the measure $D_0 = Z_0 - N\mu_0$, and hence the function ϕ, is determined by the point distribution $\mathcal{Q}$ and has nothing to do with the rotated squares $A(r, \theta)$. On the other hand, the characteristic function $\chi_{r,\theta}$ is determined by the rotated square $A(r, \theta)$ and has nothing to do with the point distribution $\mathcal{Q}$. In other words, the identity (6) represents a separation of measure and geometry as a result and at the expense of passing over to the corresponding Fourier transforms.

In lower bound proofs, the point distributions $\mathcal{Q}$ are arbitrary, so we have very little control over the measure $D_0 = Z_0 - N\mu_0$. However, we need only the following estimate on the trivial error arising from the gaps between successive integers.

Lemma 2.1. *Suppose that a measurable set $B \subseteq [-\frac{1}{2}, \frac{1}{2}]^2$ satisfies*

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

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

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

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

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

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 $Z_0(B+\mathbf{x}) = 0$. It follows that on writing $\mathbf{p} - B = \{\mathbf{p} - \mathbf{y} : \mathbf{y} \in B\}$ and $\chi_{\mathbf{p}-B}$ for its characteristic function, we have

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

as required.

□

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

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

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

$$\omega_q(\mathbf{t}) = \frac{1}{q} \int_{q/2}^{q} \int_{-\pi/4}^{\pi/4} |\widehat{\chi}_{r,\theta}(\mathbf{t})|^2 \,\mathrm{d}\theta \mathrm{d}r. \tag{7}$$

We have the following amplification result which we shall use to blow up the trivial error obtained in Lemma 2.1.

Lemma 2.2. *Suppose that $0 < p < q$. Then uniformly for all $\mathbf{t} \in \mathbf{R}^2$, we have*

$$\frac{\omega_q(\mathbf{t})}{\omega_p(\mathbf{t})} \gg \frac{q}{p}. \tag{8}$$

We shall split the proof of Lemma 2.2 into a number of steps. Throughout, we suppose that $r > 0$ and $-\pi/4 \leq \theta \leq \pi/4$.

First of all, it is easy to show that for every $\mathbf{t} = (t_1, t_2) \in \mathbf{R}^2$, we have

$$\widehat{\chi}_{r,\theta}(\mathbf{t}) = \widehat{\chi}_r(t_1\cos\theta + t_2\sin\theta, -t_1\sin\theta + t_2\cos\theta), \tag{9}$$

where χ_r denotes the characteristic function of the square $A(r) = A(r,0) = [-r,r]^2$. Furthermore, for every $\mathbf{u} = (u_1, u_2) \in \mathbf{R}^2$, we have

$$\widehat{\chi}_r(\mathbf{u}) = \frac{2\sin(ru_1)\sin(ru_2)}{\pi u_1 u_2}. \tag{10}$$

Lemma 2.2 follows easily from the result below.

Lemma 2.3. *Uniformly for all non-zero* $\mathbf{t} \in \mathbf{R}^2$*, we have*

$$\omega_q(\mathbf{t}) \asymp \min\left\{q^4, \frac{q}{|\mathbf{t}|^3}\right\}. \tag{11}$$

Proof. Note that in view of the integration over θ in the definition of $\omega_q(\mathbf{t})$, it suffices to show that uniformly for all $\mathbf{t} = (t_1, t_2) \in \mathbf{R}^2$ satisfying $t_1 > 0$ and $t_2 = 0$, we have

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

Using (9) and (10), we have

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

Since $-\pi/4 \le \theta \le \pi/4$, we have $\sin\theta \asymp \theta$ and $\cos\theta \asymp 1$, and so

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

We consider two cases. If $t_1 \le 4/\pi q$, then for all r and θ satisfying $q/2 \le r \le q$ and $-\pi/4 \le \theta \le \pi/4$, we have $\sin(rt_1\cos\theta) \asymp qt_1$ and $\sin(rt_1\sin\theta) \asymp qt_1\theta$. Hence

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

On the other hand, if $t_1 > 4/\pi q$, we then split the interval $[-\pi/4, \pi/4]$ into three intervals at the points $\theta = \pm 1/qt_1$. Clearly, we have the crude estimate

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

On the other hand, if $-1/qt_1 \le \theta \le 1/qt_1$, then we have

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

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

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

This completes the proof.

□

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 $-\pi/4 \le \theta \le \pi/4$, we have

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

Using Lemma 2.1 with $\delta = \frac{1}{9}$, we have

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

It follows from (2), (6) and (7) that

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

Using Lemma 2.2, we conclude that

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

Combining this with (2), (6) and (7), we conclude that

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

The lower bound of Theorem 2 follows immediately.

3 Upper Bounds via Fourier Transforms

In upper bound proofs, we work with specific point distributions $\mathcal{Q}$, and so have very good control over the measure $D_0 = Z_0 - N\mu_0$. Here we shall illustrate this point by sketching a proof of the upper bound of Theorem 2 in the special case when the number of points of the distribution is equal to an odd square $N = (2M+1)^2$.

Note from (5) that

$$\phi(\mathbf{t}) = \frac{1}{2\pi}\left(\sum_{\mathbf{q}\in\mathcal{Q}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{t}} - N\int_{\mathbf{R}^2} \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}}\mathrm{d}\mu_0(\mathbf{x})\right).$$

It is easy to see that

$$\int_{\mathbf{R}^2} \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}} \mathrm{d}\mu_0(\mathbf{x}) = \int_{[-1/2,1/2]^2} \mathrm{e}^{-\mathrm{i}\mathbf{x}\cdot\mathbf{t}} \mathrm{d}\mathbf{x} = \frac{\sin\frac{t_1}{2}}{\frac{t_1}{2}} \frac{\sin\frac{t_2}{2}}{\frac{t_2}{2}}.$$

Furthermore, if we take

$$\mathcal{Q} = \left\{ \left(\frac{m_1}{2M+1}, \frac{m_2}{2M+1} \right) : m_1, m_2 \in \{-M, \ldots, 0, \ldots, M\} \right\}, \tag{12}$$

then simple calculation gives

$$\sum_{\mathbf{q}\in\mathcal{Q}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{t}} = \frac{\sin\frac{t_1}{2}}{\sin\frac{t_1}{2(2M+1)}} \frac{\sin\frac{t_2}{2}}{\sin\frac{t_2}{2(2M+1)}}.$$

It follows that for the point distribution $\mathcal{Q}$ given by (12), we have

$$\phi(\mathbf{t}) = \frac{1}{2\pi} \frac{\sin\frac{t_1}{2}}{\sin\frac{t_1}{2(2M+1)}} \frac{\sin\frac{t_2}{2}}{\sin\frac{t_2}{2(2M+1)}} \left(1 - \frac{\sin\frac{t_1}{2(2M+1)}}{\frac{t_1}{2(2M+1)}} \frac{\sin\frac{t_2}{2(2M+1)}}{\frac{t_2}{2(2M+1)}} \right). \tag{13}$$

Combining (2), (6) and (7), and in view of symmetry of the set $\mathcal{Q}$, we have

$$\begin{aligned} &\int_{q/2}^{q} \int_0^{2\pi} \int_{\mathbf{R}^2} |Z_0(A(r,\theta,\mathbf{x})) - N\mu_0(A(r,\theta,\mathbf{x}))|^2 \,\mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r \\ &\quad = 4q \int_{\mathbf{R}^2} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t}. \end{aligned} \tag{14}$$

Note that the major contribution to the integral on the right hand side comes from those points in $\mathbf{R}^2$ close to points of the form $\mathbf{t} = 2(2M+1)\pi\mathbf{k}$, where $\mathbf{k} = (k_1, k_2) \in \mathbf{Z}^2$ is non-zero. Accordingly, we partition the plane into a union

$$\mathbf{R}^2 = \bigcup_{\mathbf{k}\in\mathbf{Z}^2} S(\mathbf{k}),$$

where for every $\mathbf{k} = (k_1, k_2) \in \mathbf{Z}^2$, the aligned square $S(\mathbf{k})$ is centred at the point $2(2M+1)\pi\mathbf{k}$ and has side length $2(2M+1)\pi$. Then

$$\int_{\mathbf{R}^2} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t} = \sum_{\mathbf{k}\in\mathbf{Z}^2} \int_{S(\mathbf{k})} \omega_q(\mathbf{t}) |\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t}. \tag{15}$$

Next, note from Lemma 2.3 that for all $q \in (0,1)$, we have

$$\omega_q(\mathbf{t}) \ll \min\left\{ 1, \frac{1}{|\mathbf{t}|^3} \right\}.$$

Suppose that $\mathbf{k} \in \mathbf{Z}^2$ is non-zero. Then simple calculation gives

$$\begin{aligned}
\int_{S(\mathbf{k})} \omega_q(\mathbf{t})|\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t} &\ll \int_{S(\mathbf{k})} \frac{1}{|\mathbf{t}|^3} \frac{\sin^2 \frac{t_1}{2}}{\sin^2 \frac{t_1}{2(2M+1)}} \frac{\sin^2 \frac{t_2}{2}}{\sin^2 \frac{t_2}{2(2M+1)}} \,\mathrm{d}t_1 \mathrm{d}t_2 \\
&\ll \frac{1}{(2M+1)^3|\mathbf{k}|^3} \int_{S(\mathbf{k})} \frac{\sin^2 \frac{t_1}{2}}{\sin^2 \frac{t_1}{2(2M+1)}} \frac{\sin^2 \frac{t_2}{2}}{\sin^2 \frac{t_2}{2(2M+1)}} \,\mathrm{d}t_1 \mathrm{d}t_2 \\
&= \frac{1}{(2M+1)^3|\mathbf{k}|^3} \int_{S(\mathbf{0})} \frac{\sin^2 \frac{t_1}{2}}{\sin^2 \frac{t_1}{2(2M+1)}} \frac{\sin^2 \frac{t_2}{2}}{\sin^2 \frac{t_2}{2(2M+1)}} \,\mathrm{d}t_1 \mathrm{d}t_2 \\
&\ll \frac{2M+1}{|\mathbf{k}|^3} = \frac{N^{1/2}}{|\mathbf{k}|^3}. \qquad (16)
\end{aligned}$$

On the other hand, using the identity (13), one can show that the integral

$$\int_{S(\mathbf{0})} \omega_q(\mathbf{t})|\phi(\mathbf{t})|^2 \,\mathrm{d}\mathbf{t} = O(1). \qquad (17)$$

Combining (14)–(17), we obtain

$$\int_{q/2}^{q} \int_0^{2\pi} \int_{\mathbf{R}^2} |Z_0(A(r,\theta,\mathbf{x})) - N\mu_0(A(r,\theta,\mathbf{x}))|^2 \,\mathrm{d}\mathbf{x}\mathrm{d}\theta\mathrm{d}r \ll qN^{1/2}.$$

The upper bound of Theorem 2 follows immediately.

Note that by the separation of geometry and measure, the information concerning the geometric objects, namely the squares in this case, is contained in the Fourier transform $\widehat{\chi}_{r,\theta}(\mathbf{t})$ of the characteristic functions $\chi_{r,\theta}(\mathbf{x})$, and that Lemma 2.3 gives information about the decay of this Fourier transform in some average sense. Indeed, one can calculate the decay of the Fourier transform of a number of geometric objects, and use such information to obtain good bounds for discrepancy problems. For more detailed discussion on such problems, the reader is referred to the paper of Brandolini, Colzani and Travaglini [4], the paper of Brandolini, Rigoli and Travaglini [6], as well as the paper of Brandolini, Iosevich and Travaglini [5].

We conclude this section by discussing a probabilistic technique to provide some comparison to our Fourier techniques, and establish stronger results than the upper bound of Theorem 2, under the assumption that the number of points of the distribution is equal to an odd square $N = (2M+1)^2$, and that the unit square $[-\frac{1}{2}, \frac{1}{2}]^2$ is treated as a torus.

Consider a random point set $\widetilde{\mathcal{Q}}$ as follows: First we split the unit square into $N = (2M+1)^2$ small squares of area $1/N$ in the usual way. In each small square we place a random point, uniformly distributed in the small square and independently of the distribution of all the other random points in the other small squares.

Suppose that $A = A(r,\theta,\mathbf{x})$ is a square in $[-\frac{1}{2}, \frac{1}{2}]^2$. Let $\mathcal{A}$ denote the set of all small squares that intersect the boundary ∂A of A. Then it is easy to see that $|\mathcal{A}| = O(M)$. For each square $S \in \mathcal{A}$, let $\widetilde{\mathbf{p}}_S$ denote the random point in S, write

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

and let $\eta_S = \xi_S - \mathsf{E}\xi_S$. Then it is easy to see that $|\eta_S| \leq 1$ and $\mathsf{E}\eta_S = 0$. Furthermore,

$$D[\widetilde{\mathcal{Q}}; A] = \sum_{S \in \mathcal{A}} \eta_S.$$

We now want to estimate $\mathsf{E}(|D[\widetilde{\mathcal{Q}}; A]|^W)$ from above, where W is an even positive integer. Note first that

$$|D[\widetilde{\mathcal{Q}}; A]|^W = \sum_{S_1 \in \mathcal{A}} \ldots \sum_{S_W \in \mathcal{A}} \eta_{S_1} \ldots \eta_{S_W},$$

and so

$$\mathsf{E}(|D[\widetilde{\mathcal{Q}}; A]|^W) = \sum_{S_1 \in \mathcal{A}} \ldots \sum_{S_W \in \mathcal{A}} \mathsf{E}(\eta_{S_1} \ldots \eta_{S_W}). \tag{18}$$

The random variables η_S, where $S \in \mathcal{A}$, are independent because the distribution of the random points are independent of each other. If one of $S_1, \ldots, S_W$, say S_i, is different from all of the others, then

$$\mathsf{E}(\eta_{S_1} \ldots \eta_{S_W}) = \mathsf{E}(\eta_{S_i})\mathsf{E}(\eta_{S_1} \ldots \eta_{S_{i-1}} \eta_{S_{i+1}} \ldots \eta_{S_W}) = 0.$$

It follows that the only non-zero contribution to the sum (18) comes from those terms where each of $S_1, \ldots, S_W$ appear more than once. The major contribution comes when they occur in pairs, of which there are

$$O_W\left(\binom{|\mathcal{A}|}{W/2}\right) = O_W\left(|\mathcal{A}|^{W/2}\right) = O_W(M^{W/2}) = O_W(N^{W/4})$$

such pairs. Here the subscript W denotes that the implicit constants in the inequalities may depend on the parameter W. Bounding each of such terms $\mathsf{E}(\eta_{S_1} \ldots \eta_{S_W})$ trivially by $O(1)$, we obtain the estimate

$$\mathsf{E}(|D[\widetilde{\mathcal{Q}}; A(r, \theta, \mathbf{x})]|^W) = O_W(N^{W/4}).$$

The special case $W = 2$ leads to the upper bound of Theorem 2 on integrating trivially with respect to the variables r, θ and $\mathbf{x}$.

4 Davenport's Ideas

In 1956, Davenport studied the upper bound aspects of Theorem 1. To construct a distribution of N points, consider a lattice Λ on the plane generated by the two vectors $(1, 0)$ and (θ, N^{-1}), where θ is an irrational number. We are interested in the set $\mathcal{P}^*$ which contains precisely the N points of Λ that fall into the square $[0,1)^2$. It is easy to see that

$$\mathcal{P}^* = \{(\{\theta n\}, N^{-1}n) : 0 \leq n < N\}.$$

Then it can be shown that for every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, we have

$$\begin{aligned} D[\mathcal{P}^*; B(\mathbf{x})] &= Z[\mathcal{P}^*; B(\mathbf{x})] - Nx_1x_2 \\ &= \sum_{0 \leq n < Nx_2} (\psi(\theta n - x_1) - \psi(\theta n)) + O(1) \end{aligned} \tag{19}$$

for all but a finite number of values of x_1 in the interval $[0,1]$. One can then show that the Fourier expansion of $D[\mathcal{P}^*; B(\mathbf{x})]$, apart from an error of the form $O(1)$, can be written in the form

$$D[\mathcal{P}^*; B(\mathbf{x})] \sim \sum_{m \neq 0} \left(\frac{1 - e(-x_1 m)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leq n < Nx_2} e(\theta n m) \right), \tag{20}$$

where $e(\beta) = \mathrm{e}^{2\pi \mathrm{i} \beta}$ for every $\beta \in \mathbf{R}$. However, this does not allow one to use Parseval's theorem by integrating with respect to the variable x_1 over the interval $[0,1]$. Furthermore, it is clear that the difficulty is caused by the fact that the term $\psi(\theta n)$ in (19) does not depend on the variable x_1.

This Fourier approach suggests an extra lattice to enable us to replace the term $\psi(\theta n)$ in (19) by something that depends on the variable x_1. Consequently, we consider an extra lattice Λ' on the plane generated by the two vectors $(1,0)$ and $(-\theta, N^{-1})$. More precisely, we see that the set

$$\mathcal{P}^{**} = \{(\{-\theta n\}, N^{-1}n) : 0 \leq n < N\}$$

is also a set of N points in the square $[0,1]^2$. Hence the set $\mathcal{P} = \mathcal{P}^* \cup \mathcal{P}^{**}$ contains $2N$ points in $[0,1]^2$, with the convention that points are counted with multiplicity, and

$$\begin{aligned} D[\mathcal{P}; B(\mathbf{x})] &= Z[\mathcal{P}; B(\mathbf{x})] - 2Nx_1x_2 \\ &= \sum_{0 \leq n < Nx_2} (\psi(\theta n - x_1) - \psi(\theta n + x_1)) + O(1) \end{aligned}$$

for all but a finite number of values of x_1 in the interval $[0,1]$. Then the Fourier expansion of $D[\mathcal{P}; B(\mathbf{x})]$, apart from an error of the form $O(1)$, is now of the form

$$D[\mathcal{P}; B(\mathbf{x})] \sim \sum_{m \neq 0} \left(\frac{e(x_1 m) - e(-x_1 m)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leq n < Nx_2} e(\theta n m) \right).$$

We can therefore square this expression and integrate with respect to the variable x_1 over the interval $[0,1]$. By Parseval's theorem, we have

$$\int_0^1 |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}x_1 \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leq n < Nx_2} e(\theta n m) \right|^2.$$

To estimate the sum on the right hand side, 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 $m\|m\theta\| > c > 0$ for every natural number $m \in \mathbf{N}$, where $\|\cdot\|$ denotes the distance to the nearest integer. For such a badly approximable number θ, we have the estimate

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leq n < Nx_2} e(\theta nm) \right|^2 \ll \log(2Nx_2).$$

Using this and integrating trivially with respect to the variable x_2 over the interval $[0,1]$, we obtain the desired upper bound of Theorem 1.

The attempt (20) to obtain a Fourier series also suggests the possibility of studying the problem via a Fourier series in terms of a new variable. This was achieved by Roth [21], who introduced a probabilistic approach to the problem in 1979. The idea is to consider a translation variable $t \in \mathbf{R}$, and consider translated copies

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

where $\mathbf{i} = (1,0)$, of the original lattice Λ. In other words, one studies point sets of the form

$$\mathcal{P}(t) = \{(\{t + \theta n\}, N^{-1}n) : 0 \leq n < N\}.$$

Then for every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, we have

$$D[\mathcal{P}(t); B(\mathbf{x})] = \sum_{0 \leq n < Nx_2} (\psi(t + \theta n - x_1) - \psi(t + \theta n)) + O(1)$$

for all but a finite number of values of x_1 in the interval $[0,1]$. Furthermore, the Fourier expansion of $D[\mathcal{P}(t); B(\mathbf{x})]$, apart from an error of the form $O(1)$, is now of the form

$$D[\mathcal{P}(t); B(\mathbf{x})] \sim \sum_{m \neq 0} \left(\frac{1 - e(-x_1 m)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leq n < Nx_2} e(\theta nm) \right) e(tm).$$

We can now square this expression and integrate with respect to the translation variable t over the interval $[0,1]$. By Parseval's theorem, we have

$$\int_0^1 |D[\mathcal{P}(t); B(\mathbf{x})]|^2 \, \mathrm{d}t \ll \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leq n < Nx_2} e(\theta nm) \right|^2 .$$

Integrating trivially with respect to the variables x_1 and x_2 over the interval $[0,1]$, and using the earlier assumption concerning the number θ, we obtain an existence proof of the upper bound of Theorem 1.

An approach using the lattice $\mathbf{Z}^2$ in the spirit of Davenport and Roth was made by Beck and Chen [3] in connection with their work on discrepancy relative to convex polygons. This uses a result of Davenport [13] on diophantine approximation which shows the existence of real numbers α such that both $\tan\alpha$ and $\tan(\alpha+\pi/2)$ are finite and badly approximable. The idea is to rotate the lattice $\mathbf{Z}^2$ anticlockwise by such an angle α to obtain a rotated lattice Λ_α, and then normalize Λ_α to obtain a square lattice $N^{-1/2}\Lambda_\alpha = \{N^{-1/2}\mathbf{u} : \mathbf{u} \in \Lambda_\alpha\}$, with an average of N points per unit area. Now let $\mathcal{P} = N^{-1/2}\Lambda_\alpha \cap [0,1)^2$. The study of the integral

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

is hindered by a difficulty which is similar in nature to the one that Davenport encountered with the Fourier expansion (20). However, the Roth technique enables one to introduce a probabilistic variable $\mathbf{w}$ which runs over all the values of a fundamental region S of the normalized lattice $N^{-1/2}\Lambda_\alpha$, and consider translated sets of the form

$$\mathcal{P}(\mathbf{w}) = (\mathbf{w} + N^{-1/2}\Lambda_\alpha) \cap [0,1)^2 = \{\mathbf{w}+\mathbf{u} : \mathbf{u} \in N^{-1/2}\Lambda_\alpha\} \cap [0,1)^2.$$

Then integrating first with respect to the variable $\mathbf{w}$ over the set S, and then integrating trivially with respect to the variables x_1 and x_2 over the interval $[0,1]$, we obtain again an existence proof of the upper bound of Theorem 1.

5 Further Use of Fourier Series

Much work in connection with the upper bound of Theorem 1 involves the van der Corput point sets and their generalizations. Every non-negative integer n has a unique representation of the form

$$n = \sum_{i=1}^{\infty} a_i 2^{i-1}, \quad a_i \in \{0,1\},$$

noting that the series has only finitely many non-zero terms. The sequence

$$x(n) = \sum_{i=1}^{\infty} a_i 2^{-i}, \quad n = 0,1,2,\ldots,$$

is the well known van der Corput sequence, giving rise to the van der Corput point set

$$\mathcal{V} = \{(x(n), n) : n \in \mathbf{N}_0\}.$$

The van der Corput sequence possesses a very nice periodicity property, underpinned by the fact that for any non-negative integers s and m satisfying $0 \le m < 2^s$, the set

$$\{n \in \mathbf{N}_0 : x(n) \in [m2^{-s}, (m+1)2^{-s})\}$$

is precisely the set of all non-negative integers in a residue class modulo 2^s. This nice periodicity property invites the use of Fourier series.

Let us truncate and normalize the van der Corput point set $\mathcal{V}$ to obtain the set

$$\mathcal{P}(2^h) = \{(x(n), 2^{-h}n) : 0 \leq n < 2^h\} \tag{21}$$

of $N = 2^h$ points in the unit square $[0,1]^2$. Then

$$\int_{[0,1]^2} |D[\mathcal{P}(2^h); B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = 2^{-6}h^2 + O(h), \tag{22}$$

as shown by Halton and Zaremba [15], and so this does not give a proof of the upper bound of Theorem 1.

This difficulty was studied in detail by Chen and Skriganov [8]. Denote by $\mathbf{x} = (x_1, x_2)$ the top right vertex of the rectangle $B(\mathbf{x})$. Suppose that $x_1 \neq 1$. Then it can be shown that there exists a finite set $\mathcal{I}(x_1) \subseteq \{1, \ldots, h\}$ such that

$$D[\mathcal{P}(2^h); B(\mathbf{x})] = \sum_{s \in \mathcal{I}(x_1)} \left(c_s - \psi\left(\frac{x_2 + z_s}{2^{s-h}}\right)\right) + O(1).$$

One therefore needs to study sums of the form

$$\sum_{s' \in \mathcal{I}(x_1)} \sum_{s'' \in \mathcal{I}(x_1)} \left(c_{s'} - \psi\left(\frac{x_2 + z_{s'}}{2^{s'-h}}\right)\right)\left(c_{s''} - \psi\left(\frac{x_2 + z_{s''}}{2^{s''-h}}\right)\right).$$

Using Fourier analysis and integrating with respect to the variable x_2 over the interval $[0,1]$, one can show that each of the summands above gives rise to an integral

$$\begin{aligned}\int_0^1 & \left(c_{s'} - \psi\left(\frac{x_2 + z_{s'}}{2^{s'-h}}\right)\right)\left(c_{s''} - \psi\left(\frac{x_2 + z_{s''}}{2^{s''-h}}\right)\right) \mathrm{d}x_2 \\ &= c_{s'}c_{s''} + O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).\end{aligned}$$

Unfortunately, the difficulty of handling the sum

$$\sum_{s' \in \mathcal{I}(x_1)} \sum_{s'' \in \mathcal{I}(x_1)} c_{s'}c_{s''}$$

is similar to the difficulty facing Davenport in (20).

There are a number of ways of overcoming this difficulty. In Roth [22], one uses a translation variable t and translates the point set $\mathcal{P}(2^h)$ vertically modulo 1 to obtain the point set $\mathcal{P}(2^h; t)$ and a corresponding discrepancy function

$$D[\mathcal{P}(2^h; t); B(\mathbf{x})] = \sum_{s \in \mathcal{I}(x_1)} \left(\psi\left(\frac{z_s + t}{2^{s-h}}\right) - \psi\left(\frac{w_s + t}{2^{s-h}}\right)\right) + O(1),$$

where z_2 and w_2 are constants that depend on x_2. Squaring and integrating with respect to the variable t over the interval $[0,1]$, we now handle integrals of the form

$$\int_0^1 \psi\left(\frac{z_{s'}+t}{2^{s'-h}}\right)\psi\left(\frac{z_{s''}+t}{2^{s''-h}}\right)\mathrm{d}t = O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).$$

In Chen [7], one uses digit translations to modify the point set $\mathcal{P}(2^h)$ horizontally to obtain the point set $\mathcal{P}(2^h;\chi)$ and a corresponding discrepancy function

$$D[\mathcal{P}(2^h;\chi);B(\mathbf{x})] = \sum_{s\in\mathcal{I}(x_1)}\left(c_s(\chi)+\psi\left(\frac{x_2+z_s(\chi)}{2^{s-h}}\right)\right)+O(1).$$

Squaring and integrating with respect to the variable x_2 over the interval $[0,1]$ and being economical with the truth, we now essentially handle integrals of the form

$$\begin{aligned}\int_0^1 &\left(c_{s'}(\chi)+\psi\left(\frac{x_2+z_{s'}(\chi)}{2^{s'-h}}\right)\right)\left(c_{s''}(\chi)+\psi\left(\frac{x_2+z_{s''}(\chi)}{2^{s''-h}}\right)\right)\mathrm{d}x_2\\ &= c_{s'}(\chi)c_{s''}(\chi)+O\left(\frac{2^{2\min\{s',s''\}}}{2^{s'+s''}}\right).\end{aligned}$$

Furthermore, over a large collection of digit translations χ, the sum

$$\sum_{s'\in\mathcal{I}(x_1)}\sum_{s''\in\mathcal{I}(x_1)} c_{s'}(\chi)c_{s''}(\chi)$$

has a small average.

6 Groups and Fourier-Walsh Series

The van der Corput point sets also possess nice group structure. To see this, note that the van der Corput point set $\mathcal{P}(2^h)$ defined by (21) can also be represented by

$$\mathcal{P}(2^h) = \{(0.a_1\ldots a_h, 0.a_h\ldots a_1) : a_1,\ldots,a_h\in\{0,1\}\},$$

where we have used digit expansion base 2 on the right hand side. Clearly $\mathcal{P}(2^h)$ forms a group under coordinatewise and digitwise addition modulo 2, and is isomorphic to the direct product $\mathbf{Z}_2^h$. This observation immediately invites the use of Fourier-Walsh functions and series.

Any $x\in[0,1)$ can be represented in the form

$$x = \sum_{i=1}^{\infty}\eta_i(x)2^{-i}, \quad \eta_i(x)\in\{0,1\}, \tag{23}$$

uniquely if we agree that the series above is finite whenever x is a binary rational number. If one looks at the subset $\mathcal{B}$ of all binary rational numbers in $[0,1)$, and defines the sum $x \oplus y \in \mathcal{B}$ of any two elements $x, y \in \mathcal{B}$ by setting

$$\eta_i(x \oplus y) = \eta_i(x) + \eta_i(y), \quad i = 1, 2, 3, \ldots,$$

where addition of the digits $\eta_i(x)$ and $\eta_i(y)$ is performed modulo 2, then it is easy to see that $\mathcal{B}$ forms an infinite abelian group. The characters of $\mathcal{B}$ are the Walsh functions.

For any $\ell \in \mathbf{N}_0$, the Walsh function $w_\ell(x)$, where

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell) 2^{i-1}, \quad \lambda_i(\ell) \in \{0,1\},$$

is defined for every real number $x \in [0,1)$ of the form (23) by

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

It is well known that the collection $\{w_\ell : \ell \in \mathbf{N}_0\}$ of Walsh functions gives rise to an orthonormal basis of the Hilbert space $L^2([0,1])$, and so there is a theory of Fourier-Walsh series. In particular, for any fixed $x \in [0,1)$, the characteristic function $\chi_{[0,x)}(y)$ of the interval $[0,x)$ has Fourier-Walsh expansion

$$\chi_{[0,x)}(y) \sim \sum_{\ell=0}^{\infty} \widetilde{\chi}_\ell(x) w_\ell(y), \tag{24}$$

where the Fourier-Walsh coefficients are given by

$$\widetilde{\chi}_\ell(x) = \int_0^1 \chi_{[0,x)}(y) w_\ell(y) \, \mathrm{d}y = \int_0^x w_\ell(y) \, \mathrm{d}y, \quad \ell = 0, 1, 2, \ldots.$$

Note that $\widetilde{\chi}_0(x) = x$. Furthermore, Fine [14] has shown that for every $\ell \in \mathbf{N}$, we have

$$\widetilde{\chi}_\ell(x) = \frac{1}{4} 2^{-\nu(\ell)} \left(w_{\ell \oplus 2^{\nu(\ell)}}(x) - \sum_{j=1}^{\infty} 2^{-j} w_{\ell \oplus 2^{\nu(\ell)+j}}(x) \right), \tag{25}$$

where $\nu(\ell) \in \mathbf{N}_0$ denotes the unique integer satisfying $2^{\nu(\ell)} \leq \ell < 2^{\nu(\ell)+1}$, and where for every $\ell, m \in \mathbf{N}_0$, the sum $\ell \oplus m$ is obtained by digitwise addition modulo 2 of the dyadic expansions of ℓ and m.

The Fourier-Walsh expansion of the characteristic function in turn leads to the Fourier-Walsh expansion of the discrepancy function

$$D[\mathcal{P}(2^h); B(\mathbf{x})] = \sum_{\mathbf{p} \in \mathcal{P}(2^h)} \chi_{B(\mathbf{x})}(\mathbf{p}) - 2^h \int_{[0,1]^2} \chi_{B(\mathbf{x})}(\mathbf{y}) \, \mathrm{d}\mathbf{y},$$

where for every $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ in $[0, 1]^2$,

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

To understand the estimate (22) and why we fail to establish the upper bound of Theorem 1, we observe that the term $\widetilde{\chi}_0(x) = x$ in the Fourier-Walsh expansion (24) corresponding to $\ell = 0$ is the expectation of the characteristic function $\chi_{[0,x)}(y)$. The remaining Fourier-Walsh coefficients, unfortunately, do not all have zero mean over the interval $[0, 1]$. To see this, note simply that when $\ell = 2^i$, where $i \in \mathbf{N}_0$, we have $\ell \oplus 2^{\nu(\ell)} = 0$. This lack of symmetry can be overcome by averaging techniques.

The discussion in this section so far and in Section 5 can be conducted in general in base p, where p is a fixed prime. In other words, we consider the generalization of the classical van der Corput point sets $\mathcal{P}(2^h)$ to sets of the form

$$\mathcal{P}(p^h) = \{(0.a_1 \ldots a_h, 0.a_h \ldots a_1) : a_1, \ldots, a_h \in \{0, 1, \ldots, p-1\}\},$$

where we now use digit expansion base p on the right hand side. Clearly $\mathcal{P}(p^h)$ forms a group of p^h elements under coordinatewise and digitwise addition modulo p, and is isomorphic to the direct product $\mathbf{Z}_p^h$. This suggests the use of Fourier-Walsh functions and series base p.

Any $x \in [0, 1)$ can be represented in the form

$$x = \sum_{i=1}^{\infty} \eta_i(x)p^{-i}, \quad \eta_i(x) \in \{0, 1, \ldots, p-1\}, \tag{26}$$

uniquely if we agree that the series above is finite whenever x is a p-ary rational number. If one looks at the subset $\mathcal{B}_p$ of all p-ary rational numbers in $[0, 1)$, and defines the sum $x \oplus y \in \mathcal{B}_p$ of any two elements $x, y \in \mathcal{B}_p$ by setting

$$\eta_i(x \oplus y) = \eta_i(x) + \eta_i(y), \quad i = 1, 2, 3, \ldots,$$

where addition of the digits $\eta_i(x)$ and $\eta_i(y)$ is performed modulo p, then it is easy to see that $\mathcal{B}_p$ forms an infinite abelian group. The characters of $\mathcal{B}_p$ are known as the Chrestenson-Levy functions if $p > 2$. We shall refer to them here as Walsh functions.

For any $\ell \in \mathbf{N}_0$, the Walsh function $w_\ell(x)$, where

$$\ell = \sum_{i=1}^{\infty} \lambda_i(\ell)p^{i-1}, \quad \lambda_i(\ell) \in \{0, 1, \ldots, p-1\},$$

is defined for every real number $x \in [0, 1)$ of the form (26) by

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

where $e_p(z) = \mathrm{e}^{2\pi\mathrm{i}z/p}$ for every real number z. As in the case $p = 2$, the collection $\{w_\ell : \ell \in \mathbf{N}_0\}$ of Walsh functions gives rise to an orthonormal basis of the Hilbert space $L^2([0,1])$, and so there is a theory of Fourier-Walsh series base p. In particular, for any fixed $x \in [0,1)$, the characteristic function $\chi_{[0,x)}(y)$ of the interval $[0,x)$ has Fourier-Walsh expansion

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

where $\widetilde{\chi}_0(x) = x$ and the analogues of (25) are given by Price [19]. Using this and the abbreviation $\mathcal{P}$ for the point set $\mathcal{P}(p^h)$, one can show that the discrepancy function

$$D[\mathcal{P};B(\mathbf{x})] = \sum_{\mathbf{p}\in\mathcal{P}} \chi_{B(\mathbf{x})}(\mathbf{p}) - p^h \int_{[0,1]^2} \chi_{B(\mathbf{x})}(\mathbf{y})\,\mathrm{d}\mathbf{y}$$

has Fourier-Walsh expansion

$$D[\mathcal{P};B(\mathbf{x})] \sim \sum_{\mathbf{p}\in\mathcal{P}} \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{\infty} \sum_{\ell_2=0}^{\infty} \widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2)\overline{w_{\ell_1}(p_1)w_{\ell_2}(p_2)},$$

which can be approximated by the finite series

$$\begin{aligned} D_h[\mathcal{P};B(\mathbf{x})] &= \sum_{\mathbf{p}\in\mathcal{P}} \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{p^h-1} \sum_{\ell_2=0}^{p^h-1} \widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2)\overline{w_{\ell_1}(p_1)w_{\ell_2}(p_2)} \\ &= \sum_{\substack{\ell_1=0 \\ (\ell_1,\ell_2)\neq(0,0)}}^{p^h-1} \sum_{\ell_2=0}^{p^h-1} \left(\sum_{\mathbf{p}\in\mathcal{P}} \overline{w_{\ell_1}(p_1)w_{\ell_2}(p_2)} \right) \widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2). \end{aligned}$$

Recall that the Walsh functions are characters of the group $\mathcal{P}$, so that we have the orthogonality relationship

$$\sum_{\mathbf{p}\in\mathcal{P}} w_{\ell_1}(p_1)w_{\ell_2}(p_2) = \begin{cases} p^h & \text{if } (\ell_1,\ell_2) \in \mathcal{P}^\perp, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{P}^\perp \subseteq \mathbf{N}_0^2$ is the orthogonal dual to the group $\mathcal{P}$; see, for example, Lidl and Niederreiter [16]. Hence

$$D_h[\mathcal{P};B(\mathbf{x})] = p^h \sum_{(\ell_1,\ell_2)\in\mathcal{P}^\perp\setminus\{(0,0)\}} \widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2).$$

One would like to square this expression and then integrate with respect to $\mathbf{x} = (x_1,x_2)$ over the unit square $[0,1]^2$. Unfortunately, the Fourier-Walsh coefficients

$$\widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2), \quad (\ell_1,\ell_2) \in \mathcal{P}^\perp \setminus \{(0,0)\}, \tag{27}$$

are not orthogonal in $L^2([0,1]^2)$ in general.

In Chen and Skriganov [9], it is shown that as long as the prime p is chosen large enough, there exist groups $\mathcal{P}$ of p^h elements in the square $[0,1]^2$, in the spirit of van der Corput, such that the Fourier-Walsh coefficients (27) are *quasi-orthonormal* in $L^2([0,1]^2)$. Indeed, they are able to establish Theorem 1(ii) for arbitrary dimensions with explicitly constructed point sets. More recently, Chen and Skriganov [10] have shown that in fact, as long as the prime p is chosen large enough, there exist groups $\mathcal{P}$ of p^h elements in the square $[0,1]^2$, in the spirit of van der Corput, such that the Fourier-Walsh coefficients (27) are orthogonal in $L^2([0,1]^2)$, so that

$$\int_{[0,1]^2} |D_h[\mathcal{P};B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} = p^{2h} \sum_{(\ell_1,\ell_2)\in\mathcal{P}^\perp\setminus\{(0,0)\}} \int_{[0,1]^2} |\widetilde{\chi}_{\ell_1}(x_1)\widetilde{\chi}_{\ell_2}(x_2)|^2 \,\mathrm{d}\mathbf{x}.$$

7 Roth's Orthogonal Function Method

In this section, we sketch Roth's proof of Theorem 1(i). Corresponding to every distribution $\mathcal{P}$ of N points in the unit square $[0,1]^2$, Roth creates an auxiliary function $F(\mathbf{x}) = F[\mathcal{P};\mathbf{x}]$ such that, writing $D(\mathbf{x})$ for $D[\mathcal{P};B(\mathbf{x})]$, we have the inequalities

$$\int_{[0,1]^2} F(\mathbf{x})D(\mathbf{x}) \,\mathrm{d}\mathbf{x} \gg \log N \tag{28}$$

and

$$\int_{[0,1]^2} |F(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} \ll \log N. \tag{29}$$

These, together with Schwarz's inequality

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

give the desired inequality

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

In particular, the idea of Roth is to construct an auxiliary function $F(\mathbf{x})$ of the form

$$F(\mathbf{x}) = \sum_{i=0}^{n} f_i(\mathbf{x}), \tag{30}$$

where the non-negative integer n satisfies the condition

$$2^{n-1} < 2N \le 2^n, \tag{31}$$

and where the functions $f_i(\mathbf{x})$ are orthogonal to each other, in the sense that

$$\int_{[0,1]^2} f_{i'}(\mathbf{x})f_{i''}(\mathbf{x}) \,\mathrm{d}\mathbf{x} = 0 \quad \text{whenever } i' \ne i''. \tag{32}$$

For each integer i satisfying $0 \leq i \leq n$, we 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

$$B_i(m_1, m_2) = [m_1 2^{-i}, (m_1+1)2^{-i}) \times [m_2 2^{i-n}, (m_2+1)2^{i-n}), \tag{33}$$

where $m_1, m_2 \in \mathbf{Z}$. In the terminology of classical Walsh functions, we now define the function $f_i(\mathbf{x})$ for every $\mathbf{x} = (x_1, x_2) \in B_i(m_1, m_2)$ by writing

$$f_i(\mathbf{x}) = \begin{cases} 0 & \text{if } B_i(m_1,m_2) \cap \mathcal{P} \neq \emptyset, \\ -w_{2^i}(x_1) w_{2^{n-i}}(x_2) & \text{if } B_i(m_1,m_2) \cap \mathcal{P} = \emptyset, \end{cases}$$

so that $f_i(\mathbf{x})$ is either identically zero or equal to ± 1 in the rectangle $B_i(m_1, m_2)$. Furthermore, if one partitions the rectangle $B_i(m_1, m_2)$ into a union of four smaller rectangles with horizontal side length 2^{-i-1} and vertical side length 2^{i-n-1}, then $f_i(\mathbf{x})$ is constant in each such smaller rectangle. It is not difficult to see that (32) holds for every $i', i'' \in \{0, 1, \ldots, n\}$, so it follows from (30) that

$$\int_{[0,1]^2} |F(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} = \sum_{i=0}^{n} \int_{[0,1]^2} f_i^2(\mathbf{x}) \, \mathrm{d}\mathbf{x} \leq n+1.$$

In view of (31), this establishes the inequality (29).

The functions $f_i(\mathbf{x})$ are designed to blow up the trivial error. To see this, consider a rectangle $B_i(m_1, m_2)$ of the form (33). Suppose first of all that $B_i(m_1, m_2) \cap \mathcal{P} \neq \emptyset$. Then $f_i(\mathbf{x}) = 0$ for every $\mathbf{x} \in B_i(m_1, m_2)$, and so

$$\int_{B_i(m_1,m_2)} f_i(\mathbf{x}) D(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0.$$

Suppose next that $B_i(m_1, m_2) \cap \mathcal{P} = \emptyset$. We shall consider the rectangle

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

For every $\mathbf{x} \in B_i'(m_1, m_2)$, consider a rectangle of horizontal side length 2^{-i-1} and vertical side length 2^{i-n-1}, with bottom left vertex $\mathbf{x}$. Let $\mathbf{x}^*$ be the top right vertex, and let $\mathbf{x}'$ and $\mathbf{x}''$ be the other two vertices. Then it can be shown that

$$\int_{B_i(m_1,m_2)} f_i(\mathbf{x}) D(\mathbf{x}) \, \mathrm{d}\mathbf{x} = -\int_{B_i'(m_1,m_2)} (D(\mathbf{x}) - D(\mathbf{x}') - D(\mathbf{x}'') + D(\mathbf{x}^*)) \, \mathrm{d}\mathbf{x},$$

where $D(\mathbf{x}) - D(\mathbf{x}') - D(\mathbf{x}'') + D(\mathbf{x}^*)$ represents the discrepancy of this rectangle with bottom left vertex $\mathbf{x}$. Since $B \cap \mathcal{P} = \emptyset$, we must have the trivial error

$$D(\mathbf{x}) - D(\mathbf{x}') - D(\mathbf{x}'') + D(\mathbf{x}^*) = -N 2^{-n-2}.$$

On the other hand, the rectangle $B_i'(m_1, m_2)$ has area 2^{-n-2}. Hence

$$\int_{B_i(m_1,m_2)} f_i(\mathbf{x}) D(\mathbf{x}) \, \mathrm{d}\mathbf{x} = N 2^{-2n-4}.$$

Observe now that in view of (31), there are at least $2^n - N \geq 2^{n-1}$ rectangles $B_i(m_1, m_2)$ of the type (33) where $B_i(m_1, m_2) \cap \mathcal{P} = \emptyset$. It follows that

$$\int_{[0,1]^2} f_i(\mathbf{x}) D(\mathbf{x}) \, \mathrm{d}\mathbf{x} \geq N 2^{-2n-4} 2^{n-1} \gg 1.$$

The inequality (28) follows immediately, in view of (30) and (31).

8 A Haar Wavelet Approach

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

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

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

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

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

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

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

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

together with the function $\varphi(x)$. This is known as the wavelet basis for $L^2([0,1])$; see, for example, Daubechies [11] or Meyer[17].

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

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

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

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

together with the two collections of functions

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

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

We now give an alternative proof of Theorem 1(i), due to Pollington [18]. First of all, note that the discrepancy function $D(\mathbf{x}) = D[\mathcal{P}; B(\mathbf{x})]$ can be written in the form $D(\mathbf{x}) = Z(\mathbf{x}) - Nx_1x_2$, where

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

where $\chi_S(x)$ denotes the characteristic function of the set S. We now make use of the rectangular wavelet basis for $L^2([0,1]^2)$. For every $\mathbf{n} = (n_1, n_2) \in \mathbf{N}_0^2$ and every $\mathbf{k} = (k_1, k_2)$, where $k_1 = 0, 1, 2, \ldots, 2^{n_1} - 1$ and $k_2 = 0, 1, 2, \ldots, 2^{n_2} - 1$, consider the wavelet coefficients

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

It is easy to see that

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

Simple calculation gives

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

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

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

On the other hand, we have

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

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

$$\begin{aligned} \int_p^1 \vartheta_{n,k}(x)\,\mathrm{d}x &= 2^{n/2}\int_p^{2^{-n}(k+1)} \vartheta(2^nx-k)\,\mathrm{d}x \\ &= \frac{2^{n/2}}{2^n}\int_{\{2^np\}}^1 \vartheta(y)\,\mathrm{d}y = -\frac{2^{n/2}}{2^n}\|2^np\|, \end{aligned}$$

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

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

Combining the above, we then have the wavelet coefficients

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

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

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

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

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

It follows that

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

References

[1] J. Beck. Irregularities of distribution. *Acta Math.*, 159 (1987), 1–49.

[2] J. Beck, W.W.L. Chen. Note on irregularities of distribution II. *Proc. London Math. Soc.*, 61 (1990), 251–272.

[3] J. Beck, W.W.L. Chen. Irregularities of point distribution relative to convex polygons III. *J. London Math. Soc.*, 56 (1997), 222–230.

[4] L. Brandolini, L. Colzani, G. Travaglini. Average decay of Fourier transforms and integer points in polyhedra. *Ark. Mat.*, 35 (1997), 253–275.

[5] L. Brandolini, A. Iosevich, G. Travaglini. Planar convex bodies, Fourier transform, lattice points and irregularities of distribution (preprint).

[6] L. Brandolini, M. Rigoli, G. Travaglini. Average decay of Fourier transforms and geometry of convex sets. *Rev. Mat. Iberoam.*, 14 (1998), 519–560.

[7] W.W.L. Chen. On irregularities of distribution II. *Quart. J. Math. Oxford*, 34 (1983), 257–279.

[8] W.W.L. Chen, M.M. Skriganov. Davenport's theorem in the theory of irregularities of point distribution. *Zapiski Nauch. Sem. POMI*, 269 (2000), 339–353.

[9] W.W.L. Chen, M.M. Skriganov. Explicit constructions in the classical mean squares problem in irregularities of point distribution. *J. Reine Angew. Math.*, 545 (2002), 67–95.

[10] W.W.L. Chen, M.M. Skriganov. Explicit constructions in the classical mean squares problem in irregularities of point distribution II (preprint).

[11] I. Daubechies. *Ten Lectures on Wavelets* (SIAM, 1992).

[12] H. Davenport. Note on irregularities of distribution. *Mathematika*, 3 (1956), 131–135.

[13] H. Davenport. A note on diophantine approximation II. *Mathematika*, 11 (1964), 50–58.

[14] N.J. Fine. On the Walsh functions. *Trans. American Math. Soc.*, 65 (1949), 373–414.

[15] J.H. Halton, S.K. Zaremba. The extreme and L_2 discrepancies of some plane sets. *Monats. Math.*, 73 (1969), 316–328.

[16] R. Lidl, H. Niederreiter. *Finite Fields* (Addison-Wesley, 1983).

[17] Y. Meyer. *Wavelets and Operators* (Cambridge University Press, 1992).

[18] A.D. Pollington. Haar wavelets and irregularities of distribution (preprint).

[19] J.J. Price. Certain groups of orthonormal step functions. *Canadian J. Math.*, 9 (1957), 413–425.

[20] K.F. Roth. On irregularities of distribution. *Mathematika*, 1 (1954), 73–79.

[21] K.F. Roth. On irregularities of distribution III. *Acta Arith.*, 35 (1979), 373–384.

[22] K.F. Roth. On irregularities of distribution IV. *Acta Arith.*, 37 (1980), 67–75.