Upper bounds in irregularities of distribution

W.W.L. Chen

§0. Introduction

The subject of irregularities of distribution arises from uniform distribution, but is of independent interest, and owes its current prominance to the fundamental contribution of K.F. Roth [**23**–**27**] and W.M. Schmidt [**28**–**38**]. While the theory of uniform distribution may be described as qualitative, the theory of irregularities of distribution is definitely quantitative in nature, as one seeks to measure (with great precision in many instances) the actual discrepancy (in a certain sense) incurred by a finite set of points distributed within a finite region. There are lower bound results which say that the discrepancy of a set of points cannot be less than a certain minimum value which only depends on the number of points in question, and not where they are placed within the finite region. On the other hand, there are upper bound results which say that if the points are placed carefully, then the discrepancy cannot exceed a certain maximum value which again only depends on the number of points in question. In many instances, it has been shown that this upper bound is a constant multiple of the lower bound.

The tools in this subject are diverse, and involve ideas in harmonic analysis, number theory, geometry, combinatorics and probability theory.

The purpose of this paper is to discuss some of the central ideas in the study of upper bounds in the theory of irregularities of distribution. This paper is not intended as a survey, and many results have been omitted. Also, only a few proofs are given in detail; in many other instances, we shall discuss briefly the main ideas and omit the (often very complicated) details.

In §1, we shall give an overview of the subject as a whole, and illustrate its development from its infancy to the present day by mentioning some of the key results. We also mention many extremely difficult problems which remain unsolved. In §§2–6, we shall discuss the main ideas in the study of upper bound questions. We conclude this paper by proving in the appendix the famous lower bound result of Roth [**23**] which laid the foundations of the subject.

The material in this paper is the subject of a series of lectures given at Macquarie University in the first half of 1992. I would like to express my thanks to Grigori

Kolesnik, Gerry Myerson, Peter Pleasants and Tom Schmidt for their continuing interest and patience.

§1. Summary of Main Results

§1.1. The Classical Problem

Let $U_0 = [0,1)$ and $U_1 = (0,1]$. Suppose that $\mathcal{P}$ is a distribution of N points in U_0^K, where $K, N \in \mathbb{N}$ with $K \geq 2$. For every $\mathbf{x} = (x_1, \ldots, x_K) \in U_1^K$, let

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

in other words, $B(\mathbf{x})$ denotes a K–dimensional aligned rectangular box with one corner at the origin $\mathbf{0}$ and another corner at $\mathbf{x}$. Furthermore, let

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

and write

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

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

We are interested in the behaviour of the discrepancy function $D[\mathcal{P}; B(\mathbf{x})]$. For every $W > 0$, write

$$D(K, W, N) = \inf_{|\mathcal{P}|=N} \left(\int_{U_1^K} |D[\mathcal{P}; B(\mathbf{x})]|^W \mathrm{d}\mathbf{x} \right)^{1/W},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U_0^K. Also, write

$$D(K, \infty, N) = \inf_{|\mathcal{P}|=N} \sup_{\mathbf{x} \in U_1^K} |D[\mathcal{P}; B(\mathbf{x})]|,$$

where, again, the infimum is taken over all distributions $\mathcal{P}$ of N points in U_0^K.

The presently known lower bound results are as follows. They can all be proved by Roth's orthogonal function method or modifications of it.

THEOREM 1A. (Roth [**23**]) *We have*

$$D(K, 2, N) \gg_K (\log N)^{\frac{K-1}{2}}.$$

THEOREM 1B. (Roth [**23**]) *We have*

$$D(K,\infty,N) \gg_K (\log N)^{\frac{K-1}{2}}.$$

THEOREM 1C. (Schmidt [**34**]) (Halász [**18**]) *We have*

$$D(2,\infty,N) \gg \log N.$$

THEOREM 1D. (Schmidt [**37**]) *For every $W>1$, we have*

$$D(K,W,N) \gg_{K,W} (\log N)^{\frac{K-1}{2}}.$$

THEOREM 1E. (Halász [**18**]) *We have*

$$D(K,1,N) \gg_K (\log N)^{\frac{1}{2}}.$$

THEOREM 1F. (Beck [**6**]) *We have*

$$D(3,\infty,N) \gg_\epsilon (\log N)(\log\log N)^{\frac{1}{8}-\epsilon}.$$

These are complemented by the following presently known upper bound results. For simplicity, assume that $N \geq 2$ always.

THEOREM 2A. (Davenport [**16**]) *We have*

$$D(2,2,N) \ll (\log N)^{\frac{1}{2}}.$$

THEOREM 2B. (Halton [**19**]) (Hammersley [**20**]) *We have*

$$D(K,\infty,N) \ll_K (\log N)^{K-1}.$$

THEOREM 2C. (Roth [**26**]) *We have*

$$D(3,2,N) \ll \log N.$$

THEOREM 2D. (Roth [**27**]) *We have*

$$D(K,2,N) \ll_K (\log N)^{\frac{K-1}{2}}.$$

THEOREM 2E. (Chen [**12**]) *For every $W > 0$, we have*

$$D(K, W, N) \ll_{K,W} (\log N)^{\frac{K-1}{2}}.$$

There remain the following very hard open questions.

QUESTION 1. *Is it true that*

$$D(K, \infty, N) \gg_K (\log N)^{K-1}$$

for every $K \geq 2$?

QUESTION 2. *Is it true that*

$$D(K, 1, N) \gg_K (\log N)^{\frac{K-1}{2}}$$

for every $K \geq 2$?

QUESTION 3. *What lower bound can one prove for $D(K, W, N)$ if $K \geq 2$ and $0 < W < 1$?*

Question 1 is referred to as the "Great Open Problem" in the literature. Question 2 is equally hard. Question 3 appears to be even harder.

§1.2. Some Questions Raised by Schmidt's Work

In the late 60's and early 70's, Schmidt developed his integral equations method and proved many new results. To understand some of these results, let us first of all rephrase Theorems 1B and 1C above. For simplicity, write $U = [0, 1]$, treated as a torus. Suppose that $\mathcal{P}$ is a distribution of N points in U^K, where $K, N \in \mathbb{N}$ with $K \geq 2$. For every measurable set $B \subseteq U^K$, let

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

and write

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

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

THEOREM 1B'. *There exists an aligned rectangular box $B \subseteq U^K$ such that*

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

THEOREM 1C'. *There exists an aligned rectangle $B \subseteq U^2$ such that*

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

Suppose now that we no longer require our rectangular boxes to be aligned. In other words, suppose that we may allow orthogonal transformations of our rectangular boxes. Then the situation is very different.

THEOREM 3A. (Schmidt [**31**]) *There exists a tilted rectangle $B \subseteq U^2$, of diameter less than* 1*, such that*

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

THEOREM 3B. (Schmidt [**31**]) *There exists a tilted rectangular box $B \subseteq U^3$, of diameter less than* 1*, such that*

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

Although Schmidt's method failed for $K \geq 4$ in the case of tilted rectangular boxes, it worked well for circular balls.

THEOREM 3C. (Schmidt [**31**]) *There exists a circular ball $C \subseteq U^K$, of diameter less than* 1*, such that*

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

It can be shown that the exponents in these results are essentially sharp.

THEOREM 4A. (Beck [**2**]) *For every $N \in \mathbb{N}$, there exists a distribution $\mathcal{P}$ of N points in U^K such that for every rectangular box $B \subseteq U^K$ of diameter less than* 1*,*

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

THEOREM 4B. (Beck [**2**]) *For every $N \in \mathbb{N}$, there exists a distribution $\mathcal{P}$ of N points in U^K such that for every circular ball $C \subseteq U^K$ of diameter less than* 1*,*

$$|D[\mathcal{P}; C]| \ll_K N^{\frac{1}{2}-\frac{1}{2K}} (\log N)^{O(1)}.$$

§**1.3. Some of Beck's Work**

Naturally, Schmidt's work raised the question of the connection between tilted rectangular boxes and circular balls. In the former case, we allow orthogonal transformation.

In the latter case, the sets are invariant under orthogonal transformation. If one compares Theorem 2B (in the case $K = 2$) and Theorem 3A, one might be tempted to blame the "discrepancy" in the estimates on orthogonal transformation.

In arguably the greatest contribution to the subject to date, Beck [**4**,**5**] showed essentially that the discrepancy arises from orthogonal transformation and/or from the shape of the boundary surface.

Consider first of all the case when orthogonal transformation is permitted.

Let $U = [0, 1]$, treated as a torus. Suppose that $\mathcal{P}$ is a distribution of N points in U^K, where $K, N \in \mathbb{N}$ with $K \geq 2$. Let A be a compact and convex body in U^K. For any real number $\lambda \in (0, 1]$, any proper orthogonal transformation τ in $\mathbb{R}^K$ and any vector $\mathbf{u} \in U^K$, let

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

(note that $A(\lambda, \tau, \mathbf{u})$ and A are similar to each other), and let

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

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 volume in U^K.

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

$$D_1(A, 2, N) = \inf_{|\mathcal{P}|=N} \left(\int_0^1 \int_{\mathcal{T}} \int_{U^K} |D[\mathcal{P}; A(\lambda, \tau, \mathbf{u})]|^2 \mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda \right)^{1/2},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K. Also, write

$$D_1(A, \infty, N) = \inf_{|\mathcal{P}|=N} \sup_{\substack{\lambda \in (0,1] \\ \tau \in \mathcal{T} \\ \mathbf{u} \in U^K}} |D[\mathcal{P}; A(\lambda, \tau, \mathbf{u})]|,$$

where, again, the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K.

Theorem 5A. (Beck [**4**]) *Suppose that $r(A) \geq N^{-1/K}$, where $r(A)$ denotes the radius of the largest inscribed ball of A. Then*

$$D_1(A, 2, N) \gg_A N^{\frac{1}{2} - \frac{1}{2K}}.$$

Theorem 5B. (Beck [**4**]) *Suppose that $r(A) \geq N^{-1/K}$, where $r(A)$ denotes the radius of the largest inscribed ball of A. Then*

$$D_1(A, \infty, N) \gg_A N^{\frac{1}{2} - \frac{1}{2K}}.$$

These lower bound results are complemented by the following upper bound results. On the one hand, we have

Theorem 6A. (Beck–Chen [**9**]) *We have*

$$D_1(A,2,N) \ll_A N^{\frac{1}{2}-\frac{1}{2K}}.$$

Also, similar to Theorems 4A and 4B, we have

Theorem 6B. (Beck [**2**]) *We have*

$$D_1(A,\infty,N) \ll_A N^{\frac{1}{2}-\frac{1}{2K}}(\log N)^{O(1)}.$$

Consider next the case when orthogonal transformation is not permitted. There are technical difficulties when $K \geq 3$, so we shall concentrate on the case $K=2$.

Let $U=[0,1]$, treated as a torus. Suppose that $\mathcal{P}$ is a distribution of N points in U^2, where $N \in \mathbb{N}$. Let A be a compact and convex body in U^2. For any real number $\lambda \in (0,1]$ and any vector $\mathbf{u} \in U^K$, let

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

(note that $A(\lambda,\mathbf{u})$ and A are homothetic to each other), and let

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

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 volume in U^2.

Let

$$D_0(A,2,N) = \inf_{|\mathcal{P}|=N} \left(\int_0^1 \int_{U^2} |D[\mathcal{P};A(\lambda,\mathbf{u})]|^2 \mathrm{d}\mathbf{u}\mathrm{d}\lambda \right)^{1/2},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K. Also, write

$$D_0(A,\infty,N) = \inf_{|\mathcal{P}|=N} \sup_{\substack{\lambda\in(0,1] \\ \mathbf{u}\in U^2}} |D[\mathcal{P};A(\lambda,\mathbf{u})]|,$$

where, again, the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K.

Theorem 7A. (Beck [**5**]) *We have*

$$D_0(A,2,N) \gg_A \max\{(\log N)^{\frac{1}{2}}, \xi_N(A)\},$$

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

Theorem 7B. (Beck [**5**]) *We have*

$$D_0(A,\infty,N) \gg_A \max\{(\log N)^{\frac{1}{2}}, \xi_N(A)\}.$$

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,

Theorem 8A. (Beck [**5**]) *We have*

$$D_0(A,\infty,N) \ll_A \max\{\log N, \xi_N^2(A)\}.$$

Theorem 8B. (Beck–Chen [**8**]) *Suppose that A is a convex polygon. Then*

$$D_0(A,\infty,N) \ll_{A,\epsilon} (\log N)^{5+\epsilon}.$$

We comment here that Theorem 8B is far from being best possible. In fact, it is shown in [**5**] that the exponent can be replaced by $4+\epsilon$. However, the argument is much more complicated.

There are the following open questions.

QUESTION 4. *Close tha gap between the lower estimate in Theorem 5B and the upper estimate of Theorem 6B.*

QUESTION 5. *Is it true that*

$$D_0(A,\infty,N) \gg_A \log N?$$

QUESTION 6. *Study the higher–dimensional analogues of Theorem 7B. Is it true that*

$$D_0(A,\infty,N) \gg_A (\log N)^{K-1}?$$

Question 6 is sometimes referred to as the "Greater Open Problem".

§1.4. The Effect of Dimension on Discrepancy

Consider Theorems 1A and 2D. It is quite clear that the order of magnitude of the function $D(K,2,N)$ depends on the dimension K. Consider also Theorems 5A and

6A. It is again clear that the order of magnitude of the function $D_1(A,2,N)$ depends on the dimension K.

Let us consider now a combination of these two problems. Again let $U=[0,1]$, treated as a torus. Suppose that $\mathcal{P}$ is a distribution of N points in U^{K+L}, where $K,L,N\in\mathbb{N}$ with $K\geq 2$. Let A be a compact and convex body in U^K. For any real number $\lambda\in(0,1]$, any proper orthogonal transformation τ in $\mathbb{R}^K$ and any vector $\mathbf{u}\in U^K$, let

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

as as §1.3. Also, for every $\mathbf{y}=(y_1,\ldots,y_L)\in U^L$, let

$$B(\mathbf{y})=[0,y_1)\times\ldots\times[0,y_L)$$

as in §1.1. We now consider the cartesian product

$$A(\lambda,\tau,\mathbf{u})\times B(\mathbf{y})\in U^{K+L},$$

and write

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

We are interested in the discrepancy function

$$D[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times B(\mathbf{y})]=Z[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times B(\mathbf{y})]-N\mu_K(A(\lambda,\tau,\mathbf{u}))\mu_L(B(\mathbf{y})),$$

where μ_K and μ_L denote respectively the usual volume in U^K and U^L.

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

$$D(A,L,2,N)=\inf_{|\mathcal{P}|=N}\left(\int_0^1\int_{\mathcal{T}}\int_{U^K}\int_{U^L}|D[\mathcal{P};A(\lambda,\tau,\mathbf{u})\times B(\mathbf{y})]|^2\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda\right)^{1/2},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^{K+L}.

It follows easily from Theorem 5A that

Theorem 9. *Suppose that $r(A)\geq N^{-1/K}$, where $r(A)$ denotes the radius of the largest inscribed ball of A. Then*

$$D(A,L,2,N)\gg_{A,L}N^{\frac{1}{2}-\frac{1}{2K}}.$$

The natural question is whether this trivial lower estimate is best possible. Note that the order of magnitude of this estimate, while naturally dependent on the dimension K, is independent of the dimension L. This raised the question of whether the order of magnitude of $D(A,L,2,N)$ is independent of L. Rather surprisingly, this was shown to be the case.

Theorem 10. (Beck–Chen [**9**]) *We have*

$$D(A,L,2,N) \ll_{A,L} N^{\frac{1}{2}-\frac{1}{2K}}.$$

There are the following open questions.

QUESTION 7. *Repeat the investigation on the L^∞–norms.*

QUESTION 8. *Find other situations when dimension has interesting effect (or lack of it) on the discrepancy.*

Question 7 involves possibly first solving Question 4, but one might get away with not first having to solve Question 1.

§**1.5. Roth's Disc–Segment Problem**

Suppose that $\mathcal{P}$ is a distribution of N points in U_0, the closed disc of unit area and centred at the origin $\mathbf{0}$. For every measurable set B in $\mathbb{R}^2$, let

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

and write

$$D[\mathcal{P};B] = Z[\mathcal{P};B] - N\mu(B \cap U_0),$$

where μ denotes the usual 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 half-plane

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

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

Let

$$D(\infty,N) = \inf_{|\mathcal{P}|=N} \sup_{\substack{0 \leq r \leq \pi^{-1/2} \\ 0 \leq \theta \leq 2\pi}} |D[\mathcal{P};S(r,\theta)]|,$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U_0.

Roth asked the question of whether $D(\infty,N) \to +\infty$ as $N \to \infty$. This question was answered in the affirmative by Beck.

Theorem 11A. (Beck [**3**]) *We have*

$$D(\infty,N) \gg N^{\frac{1}{4}}(\log N)^{-\frac{7}{2}}.$$

More recently, using integral–geometric ideas, Alexander proved the following sharper result.

Theorem 11B. (Alexander [**1**]) *We have*

$$D(\infty,N) \gg N^{\frac{1}{4}}.$$

In fact, both Beck and Alexander studied the L^2–norm of the discrepancy function. For every $W>0$, write

$$D(W,N)=\inf_{|\mathcal{P}|=N}\left(\int_0^{2\pi}\int_0^{\pi^{-1/2}}|D[\mathcal{P};S(r,\theta)]|^W\,\mathrm{d}r\mathrm{d}\theta\right)^{1/W},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U_0.

Theorem 11C. (Alexander [**1**]) *We have*

$$D(2,N) \gg N^{\frac{1}{4}}.$$

Make the important observation that

$$\frac{1}{2}-\frac{1}{2K}=\frac{1}{4} \qquad \text{if} \qquad K=2.$$

This provides a link between the disc–segment problem and the questions in §1.3.

Theorem 11C is complemented by the result below, which can be proved using the methods for proving Theorems 6A and 10.

Theorem 12A. *We have*

$$D(2,N) \ll N^{\frac{1}{4}}.$$

The situation is drastically different when $W=1$. We shall prove the following rather surprising result.

Theorem 12B. (Beck–Chen [**10**]) *We have*

$$D(1,N) \ll (\log N)^2.$$

Simply compare the lower estimate in Theorem 11C and the upper estimate in Theorem 12B.

On the other hand, there are the following open questions.

QUESTION 9. *Is it true that*

$$(\log N)^{\frac{1}{2}} \ll D(1,N) \ll (\log N)^{\frac{1}{2}}?$$

QUESTION 10. *Study the behaviour of $D(W,N)$ for $1 < W < 2$.*

Question 10 is due to Schmidt at Oberwolfach 1990.

§1.6. Convex Polygons

Let us return to the questions in §1.3, but restrict ourselves to the special case when A is a convex polygon in U^2. More precisely, let $U = [0,1]$, treated as a torus. Suppose that $\mathcal{P}$ is a distribution of N points in U^2, where $N \in \mathbb{N}$. Let A be a convex polygon in U^2. For any real number $\lambda \in (0,1]$, any rotation $\tau \in [0,2\pi)$ and any vector $\mathbf{u} \in U^2$, let

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

and let

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

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 volume in U^2.

For every $W > 0$, write

$$D(A,W,N) = \inf_{|\mathcal{P}|=N} \left(\int_0^1 \int_0^{2\pi} \int_{U^2} |D[\mathcal{P}; A(\lambda,\tau,\mathbf{u})]|^W \mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda \right)^{1/W},$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^2.

As a special case of Theorem 5A and Theorem 6A, we have respectively

Theorem 13. *Suppose that $r(A) \geq N^{-1/2}$, where $r(A)$ denotes the radius of the largest inscribed ball of A. Then*

$$D(A,2,N) \gg_A N^{\frac{1}{4}}.$$

Theorem 14A. *We have*

$$D(A,2,N) \ll_A N^{\frac{1}{4}}.$$

Note now that a convex polygon is the interesction of a finite number of half–planes. We can therefore adapt the ideas in the proof of Theorem 12B to prove the following result.

Theorem 14B. (Beck–Chen [**11**]) *We have*

$$D(A,1,N) \ll_A (\log N)^2.$$

There are the following open questions.

QUESTION 11. *Is it true that*

$$(\log N)^{\frac{1}{2}} \ll_A D(A,1,N) \ll_A (\log N)^{\frac{1}{2}}?$$

QUESTION 12. *Study the behaviour of $D(A,W,N)$ for $1 < W < 2$.*

QUESTION 13. *Study the corresponding problem when rotation is omitted.*

QUESTION 14. *Investigate the problem in higher dimensions.*

Some progress is being made on Question 13. However, study of Question 14 appears to be severely hindered by our lack of knowledge on exponential sums, unless, of course, the answer is of an unexpected nature.

§**1.7. Comments**

Naturally, the above represent only a selection of results in the subject. Progress up to the mid–1980's is covered in the monograph of Beck–Chen [**7**]. Progress since is expected to be covered in the second half of a forthcoming monograph.

We shall discuss Theorems 2A and 2C in §2, Theorems 2B and 2D in §3, Theorem 8B in §4, Theorems 6B, 6A and 10 in §5 and Theorems 12B and 14B in §6. We shall also prove Theorem 1A in the appendix.

§**2. Davenport's Method**

We shall follow the notation in §1.1. We state Davenport's theorem [**16**] as follows.

THEOREM 2A. *For every even natural number N, there exists a distribution $\mathcal{P}$ of N points in U_0^2 such that*

$$\int_{U_1^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \mathrm{d}\mathbf{x} \ll \log N.$$

More than 20 years later, Roth [**26**] was able to extend Davenport's ideas to prove an analogue in U_0^3.

THEOREM 2C. *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in U_0^3 such that*

$$\int_{U_1^3} |D[\mathcal{P}; B(\mathbf{x})]|^2 \mathrm{d}\mathbf{x} \ll (\log N)^2.$$

§2.1. Davenport's Ideas

Let θ be any irrational number having a continued fraction with bounded partial quotients. A well–known result on diophantine approximation states that there exists a positive constant $c = c(\theta)$, depending on θ, such that

$$\nu \|\nu\theta\| > c > 0 \tag{2.1}$$

for all positive integers ν, where $\|\cdot\|$ denotes the distance from the nearest integer. For the remainder of this section, we assume that such a number θ has been chosen and fixed, and constants in the subsequent argument may depend on this choice of θ.

Lemma 2.1. *Let $W_1 \in \mathbb{Z}$ and $W \in \mathbb{N}$. Then*

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{n=W_1}^{W_1+W-1} e(\theta n \nu) \right|^2 \ll \log(2W).$$

Proof. It is well–known that

$$\left| \sum_{n=W_1}^{W_1+W-1} e(\theta n \nu) \right| \ll \min\{W, \|\nu\theta\|^{-1}\},$$

so that

$$S = \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{n=W_1}^{W_1+W-1} e(\theta n \nu) \right|^2 \ll \sum_{m=1}^{\infty} 2^{-2m} \sum_{2^{m-1} \leq \nu < 2^m} \min\{W^2, \|\nu\theta\|^{-1}\}.$$

For any pair $m, p \in \mathbb{N}$, there are at most two values of ν in the interval $2^{m-1} \leq \nu < 2^m$ for which

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

for otherwise the difference $(\nu_1 - \nu_2)$ of two of them would contradict (2.1). Hence

$$\begin{aligned}
S &\ll \sum_{m=1}^{\infty}\sum_{p=1}^{\infty}\min\{2^{-2m}W^2, p^{-2}\} \\
&= \sum_{2^m \leq W}\sum_{p=1}^{\infty}\min\{2^{-2m}W^2, p^{-2}\} + \sum_{2^m > W}\sum_{p=1}^{\infty}\min\{2^{-2m}W^2, p^{-2}\} \\
&\ll \sum_{2^m \leq W}\sum_{p=1}^{\infty}p^{-2} + \sum_{2^m > W}\left(2^{-2m}W^2 2^m W^{-1} + \sum_{p > 2^m W^{-1}} p^{-2}\right) \\
&\ll \sum_{2^m \leq W} 1 + \sum_{2^m > W} 2^{-m}W \ll \log(2W). \qquad \clubsuit
\end{aligned}$$

We shall be concerned with (2– and) 3–dimensional euclidean space, and denote a typical point by (x, y, z). The vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ denote respectively $(1,0,0), (0,1,0), (0,0,1)$. The symbol Λ is reserved for non–degenerate lattices in the xy–plane. We also let

$$\Lambda_0 = \Lambda_0(\theta\mathbf{i} + \mathbf{j}, \mathbf{i})$$

denote the lattice generated by $\theta\mathbf{i} + \mathbf{j}$ and $\mathbf{i}$.

For any lattice Λ and any rectangle of the form $R = [0, X) \times [Y_1, Y_2)$, let $Z[\Lambda; R]$ denote the number of points of Λ that fall into R, and write

$$E[\Lambda; R] = Z[\Lambda; R] - |d(\Lambda)|^{-1}A(R),$$

where $d(\Lambda)$ is the determinant of the lattice Λ and $A(R)$ is the area of R.

It is clear that $|d(\Lambda_0)| = 1$.

Let $M \in \mathbb{N}$. We are interested in the M points of Λ_0 that fall into $[0,1) \times [0, M)$. Let $R^* = [0, X) \times [Y_1, Y_2)$, where $0 < X \leq 1$ and where the integers Y_1, Y_2 satisfy $0 \leq Y_1 < Y_2 \leq M$.

Let $\psi(x) = x - [x] - 1/2$ when $x \notin \mathbb{Z}$ and $\psi(x) = 0$ when $x \in \mathbb{Z}$. Then since $0 < X \leq 1$, we have

$$\psi(x - X) - \psi(x) = \begin{cases} 1 - X & (0 < \{x\} < X), \\ -X & (\{x\} > X), \end{cases}$$

so that

$$Z[\Lambda_0; R^*] = \sum_{n=Y_1}^{Y_2 - 1}(X + \psi(\theta n - X) - \psi(\theta n))$$

for all but a finite number of values of X in the interval $0 < X \le 1$. We comment here that the use of the function ψ is a technical device. One really wants to study the characteristic function.

It follows that

$$E[\Lambda_0; R^*] = \sum_{n=Y_1}^{Y_2-1} (\psi(\theta n - X) - \psi(\theta n)) \tag{2.2}$$

for all but a finite number of values of X in the interval $0 < X \le 1$. Note now that $\psi(x)$ has the Fourier expansion

$$\psi(x) = -\sum_{\nu \neq 0} \frac{e(x\nu)}{2\pi \mathrm{i}\nu},$$

so that the right–hand side of (2.2) has the expansion

$$\sum_{\nu \neq 0} \left(\frac{1 - e(-\nu X)}{2\pi \mathrm{i}\nu} \right) \left(\sum_{n=Y_1}^{Y_2-1} e(\theta n \nu) \right). \tag{2.3}$$

We would like to square the expression (2.3) and integrate with respect to X over the interval $(0, 1]$. Unfortunately, the term 1 in $(1 - e(-\nu X))$ proves to be a nuisance.

In order to overcome this difficulty, Davenport introduced another lattice $\Lambda_0' = \Lambda_0'(-\theta\mathbf{i} + \mathbf{j}, \mathbf{i})$ and considered the $2M$ points of $\Lambda_0 \cup \Lambda_0'$ in $[0, 1) \times [0, M)$. Then, since $\psi(x)$ is an odd function,

$$\begin{aligned} & Z[\Lambda_0 \cup \Lambda_0'; R^*] - 2A(R^*) \\ &\quad = \sum_{n=Y_1}^{Y_2-1} (\psi(\theta n - X) - \psi(\theta n) + \psi(-\theta n - X) - \psi(-\theta n)) \\ &\quad = \sum_{n=Y_1}^{Y_2-1} (\psi(\theta n - X) - \psi(\theta n + X)) \end{aligned} \tag{2.4}$$

for all but a finite number of values of X in the interval $0 < X \le 1$. Now the right–hand side of (2.4) has the expansion

$$\sum_{\nu \neq 0} \left(\frac{e(\nu X) - e(-\nu X)}{2\pi \mathrm{i}\nu} \right) \left(\sum_{n=Y_1}^{Y_2-1} e(\theta n \nu) \right),$$

so that by Parseval's theorem and Lemma 2.1,

$$\begin{aligned} \int_0^1 |Z[\Lambda_0 \cup \Lambda_0'; R^*] - 2A(R^*)|^2 \mathrm{d}X &\ll \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{n=Y_1}^{Y_2-1} e(\theta n \nu) \right|^2 \\ &\ll \log(2(Y_2 - Y_1)) \ll \log(2M). \end{aligned}$$

If Y is any real number satisfying $0 < Y \leq M$, then for $R = [0, X) \times [0, Y)$, we take $R^* = [0, X) \times [0, -[-Y])$, where $-[-Y]$ is the least integer not less than Y. Then

$$Z[\Lambda_0 \cup \Lambda_0'; R] = Z[\Lambda_0 \cup \Lambda_0'; R^*]$$

and

$$A(R) - A(R^*) \ll 1,$$

so that

$$\int_0^1 |Z[\Lambda_0 \cup \Lambda_0'; R] - 2A(R)|^2 \mathrm{d}X \ll \log(2M).$$

It follows that

$$\int_0^M \int_0^1 |Z[\Lambda_0 \cup \Lambda_0'; R] - 2A(R)|^2 \mathrm{d}X \mathrm{d}Y \ll M \log(2M).$$

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

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

of $2M$ points in U_0^2 satisfies the requirements of Theorem 2A.

§**2.2. Roth's Averaging Argument**

Instead of introducing the extra lattice Λ_0' to overcome the difficulty in (2.3), Roth [**26**] devised an ingenious variation of the argument. This new idea, in its various different forms and disguises, proved to be extremely important in later work on upper bound theorems, as will be evident in the rest of this section and in §3, §5 and §6.

For any $t \in \mathbb{R}$, let $t\mathbf{i} + \Lambda_0$ be the lattice given by

$$t\mathbf{i} + \Lambda_0 = \{t\mathbf{i} + \mathbf{v} : \mathbf{v} \in \Lambda_0\};$$

in other words, $t\mathbf{i} + \Lambda_0$ is a translation in the x–direction of the lattice Λ_0. Then

$$E[t\mathbf{i} + \Lambda_0; R^*] = \sum_{n=Y_1}^{Y_2-1} (\psi(t + \theta n - X) - \psi(t + \theta n))$$

has the expansion

$$\sum_{\nu \neq 0} \left(\frac{1 - e(-\nu X)}{2\pi \mathrm{i}\nu}\right) \left(\sum_{n=Y_1}^{Y_2-1} e(\theta n \nu)\right) e(\nu t),$$

so that by Parseval's theorem,

$$\int_0^1 |E[t\mathbf{i} + \Lambda_0; R^*]|^2 \mathrm{d}t \ll \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left|\sum_{n=Y_1}^{Y_2-1} e(\theta n \nu)\right|^2.$$

Hence Roth was able to prove

Lemma 2.2. *Suppose that* $M \in \mathbb{N}$. *Suppose further that* $0 < X_2' - X_1' \leq 1$ *and* $0 < Y_2' - Y_1' \leq M$. *If* $R = [X_1', X_2') \times [Y_1', Y_2')$, *then*

$$\int_0^1 |E[t\mathbf{i} + \Lambda_0; R]|^2 \mathrm{d}t \ll \log(2M).$$

Note that the function $E[t\mathbf{i} + \Lambda_0; R]$ is periodic in t with period 1, so that we can assume that $X_1' = 0$ when proving Lemma 2.2. On the other hand, if we take the special case $R = [0, X) \times [0, Y)$ where $0 < X \leq 1$ and $0 < Y \leq M$, then on integrating trivially with respect to X and Y, we obtain

$$\int_0^1 \int_0^M \int_0^1 |E[t\mathbf{i} + \Lambda_0; R]|^2 \mathrm{d}X \mathrm{d}Y \mathrm{d}t \ll M \log(2M).$$

Rescaling in the Y–direction by a factor $1/M$, we see that for some $t \in [0, 1]$, the set

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

of M points in U_0^2 satisfies the requirements of Theorem 2A.

§2.3. Layers of Lattices

The following form of Lemma 2.2 is better suited for the proof of Theorem 2C.

Lemma 2.2'. *Suppose that* $M \in \mathbb{N}$. *Suppose further that* $0 < X_2'' - X_1'' \leq M^{-1}$ *and* $0 < Y_2'' - Y_1'' \leq 1$. *If* $R = [X_1'', X_2'') \times [Y_1'', Y_2'')$, *then*

$$\int_0^1 |E[M^{-1}t\mathbf{i} + M^{-1}\Lambda_0; R]|^2 \mathrm{d}t \ll \log(2M);$$

in other words,

$$\int_0^1 |E[t\mathbf{i} + M^{-1}\Lambda_0; R]|^2 \mathrm{d}t \ll \log(2M).$$

Note that the first inequality is obtained from Lemma 2.2 by a change of scale. On the other hand, the second inequality follows from the first, as we have, in view of periodicity,

$$\int_0^M |E[M^{-1}t\mathbf{i} + M^{-1}\Lambda_0; R]|^2 \mathrm{d}t \ll M \log(2M).$$

The idea of Roth is to consider layers of 2–dimensional lattices in 3–dimensional euclidean space.

If $\mathcal{S}$ is any subset of the 3–dimensional euclidean space, we define, for any vector $\mathbf{v}^*$,

$$\mathbf{v}^* + \mathcal{S} = \{\mathbf{v}^* + \mathbf{v} : \mathbf{v} \in \mathcal{S}\};$$

in other words, $\mathbf{v}^* + \mathcal{S}$ is a translation by $\mathbf{v}^*$ of the set $\mathcal{S}$.

Roth considered sets Ω of the type

$$\bigcup_{\nu=p_1}^{p_2} (\nu\mathbf{k} + \mathbf{w}_\nu + \Lambda), \tag{2.5}$$

where p_1, p_2 are non–negative integers, $\mathbf{w}_\nu = (x_\nu, y_\nu, 0)$ is a vector in the xy–plane for each ν, and Λ is a lattice in the xy–plane. He also considered boxes B of the type

$$[X', X'') \times [Y', Y'') \times [Z', Z''). \tag{2.6}$$

If Ω is a set of the type (2.5) and B is a box of the type (2.6) with $p_1 \le Z' < Z'' \le p_2+1$, we write $Z[\Omega; B]$ for the number of points of Ω that fall into B, and write

$$E[\Omega; B] = Z[\Omega; B] - |d(\Lambda)|^{-1} V(B),$$

where $V(B)$ is the volume of B.

The sets Ω that Roth constructed are obtained from Λ_0 as follows: Recall that $\Lambda_0 = \Lambda_0(\mathbf{u}, \mathbf{i})$, where $\mathbf{u} = \theta\mathbf{i} + \mathbf{j}$. For any non–negative integer m, write

$$\Lambda_m = 2^{-m}\Lambda_0 = \Lambda(2^{-m}\mathbf{u}, 2^{-m}\mathbf{i}).$$

Define

$$\mathbf{q}_0 = \mathbf{0}, \quad \mathbf{q}_1 = \tfrac{1}{2}\mathbf{u}, \quad \mathbf{q}_2 = \tfrac{1}{2}\mathbf{i}, \quad \mathbf{q}_3 = \tfrac{1}{2}\mathbf{u} + \tfrac{1}{2}\mathbf{i}.$$

Then it is not difficult to see that for every non–negative integer m,

$$\Lambda_{m+1} = \bigcup_{\tau=0}^{3} (2^{-m}\mathbf{q}_\tau + \Lambda_m).$$

Roth defined $\Omega_0, \Omega_1, \ldots$ successively by $\Omega_0 = \Lambda_0$ and

$$\Omega_{m+1} = \bigcup_{\tau=0}^{3} (\tau 4^m \mathbf{k} + 2^{-m}\mathbf{q}_\tau + \Omega_m).$$

Then it is easily seen that

Lemma 2.3. *Ω_m has a representation of the type* (2.5) *with $p_1 = 0$, $p_2 = 4^m - 1$ and $\Lambda = \Lambda_0$. Furthermore, the projection of Ω_m onto the xy–plane is Λ_m.*

The special boxes that Roth considered are defined as follows.

Definition. A box of the type $[0, X) \times [0, Y) \times [0, Z)$ is said to be admissible with respect to m if

$$0 < X \leq 2^{-m}, \quad 0 < Y \leq 1, \quad 0 < Z \leq 4^m.$$

Theorem 2C can then be easily be deduced from the following lemma.

Lemma 2.4. *There exists a constant $c_0 = c_0(\theta)$ such that for any non–negative integer m,*

$$\int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Omega_m; B]|^2 \mathrm{d}s\mathrm{d}t \leq c_0(m+1)^2 \tag{2.7}$$

for every box B that is admissible with respect to m.

For let the natural number $N \geq 2$ be given. Choose m so that $2^{m-1} < N \leq 2^m$. Writing $B = B(X, Y, Z) = [0, X) \times [0, Y) \times [0, Z)$, we have

$$\int_0^1 \int_0^1 \int_0^{4^m} \int_0^{N2^{-m}} \int_0^{2^{-m}} |E[s\mathbf{u} + t\mathbf{i} + \Omega_m; B(X, Y, Z)]|^2 \mathrm{d}X\mathrm{d}Y\mathrm{d}Z\mathrm{d}s\mathrm{d}t$$
$$\leq c_0(m+1)^2 N.$$

Hence there exist s^*, t^* satisfying $0 \leq s^*, t^* < 1$ such that

$$\int_0^{4^m} \int_0^{N2^{-m}} \int_0^{2^{-m}} |E[s^*\mathbf{u} + t^*\mathbf{i} + \Omega_m; B(X, Y, Z)]|^2 \mathrm{d}X\mathrm{d}Y\mathrm{d}Z \leq c_0(m+1)^2 N.$$

By Lemma 2.3, there are exactly N points of $s^*\mathbf{u} + t^*\mathbf{i} + \Omega_m$ in the region $[0, 2^{-m}) \times [0, N2^{-m}) \times [0, 4^m)$. If these are the points

$$(2^{-m}x_\nu, N2^{-m}y_\nu, 4^m z_\nu) \quad (\nu = 1, \ldots, N),$$

then the set

$$\mathcal{P} = \{(x_\nu, y_\nu, z_\nu) : \nu = 1, \ldots, N\}$$

of N points in U_0^3 satisfies the requirements of Theorem 2C.

We shall prove Lemma 2.4 by induction on m. For $m = 0$, the result is trivial if the constant c_0 is chosen to be large enough. Suppose now that $m \geq 0$ and that (2.7) holds for all boxes admissible with respect to m. Suppose now that the box

$$B^* = [0, X^*) \times [0, Y^*) \times [0, Z^*)$$

is admissible with respect to $(m+1)$. Let the integer μ satisfy $\mu 4^m < Z^* \leq (\mu+1)4^m$. Then $0 \leq \mu \leq 3$. If $\mu = 0$, then B^* is admissible with respect to m. Hence we may assume that $0 < \mu \leq 3$. Then

$$B^* = \left(\bigcup_{\tau=0}^{\mu-1} B^{(\tau)} \right) \cup B^{**},$$

where, for $0 \leq \tau \leq \mu - 1$,

$$B^{(\tau)} = [0, X^*) \times [0, Y^*) \times [\tau 4^m, (\tau+1)4^m)$$

and where

$$B^{**} = [0, X^*) \times [0, Y^*) \times [\mu 4^m, Z^*).$$

The idea is to use Lemma 2.2' on the projection of $B^{(\tau)}$ onto the xy–plane, with $M = 2^m$, and to note that $-\mu 4^m \mathbf{k} + B^{**}$ is admissible with respect to m.

Writing

$$E_1(s,t) = \sum_{\tau=0}^{\mu-1} E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^{(\tau)}]$$

and

$$E_2(s,t) = E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^{**}],$$

we have

$$\int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^*]|^2 \mathrm{d}s\mathrm{d}t = I_1 + I_2 + 2J,$$

where, for $\beta = 1, 2$,

$$I_\beta = \int_0^1 \int_0^1 |E_\beta(s,t)|^2 \mathrm{d}s\mathrm{d}t,$$

and where

$$J = \int_0^1 \int_0^1 E_1(s,t) E_2(s,t) \mathrm{d}s\mathrm{d}t.$$

Consider first

$$I_1 \leq \mu \sum_{\tau=0}^{\mu-1} \int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^{(\tau)}]|^2 \mathrm{d}s\mathrm{d}t.$$

Then writing $R_0 = [0, X^*) \times [0, Y^*)$, we see that

$$E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^{(\tau)}] = E[s\mathbf{u} + t\mathbf{i} + 2^{-m}\mathbf{q}_\tau + \Lambda_m; R_0]. \tag{2.8}$$

In view of periodicity in s and t, we have

$$I_1 \leq \mu^2 \int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Lambda_m; R_0]|^2 \mathrm{d}s\mathrm{d}t \ll m + 1 \tag{2.9}$$

by Lemma 2.2' with $M = 2^m$. On the other hand,

$$\begin{aligned} I_2 &= \int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Omega_{m+1}; B^{**}]|^2 \mathrm{d}s\mathrm{d}t \\ &= \int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + 2^{-m}\mathbf{q}_\mu + \Omega_m; -\mu 4^m \mathbf{k} + B^{**}]|^2 \mathrm{d}s\mathrm{d}t \\ &= \int_0^1 \int_0^1 |E[s\mathbf{u} + t\mathbf{i} + \Omega_m; -\mu 4^m \mathbf{k} + B^{**}]|^2 \mathrm{d}s\mathrm{d}t, \end{aligned}$$

in view of periodicity. Since $-\mu 4^m \mathbf{k} + B^{**}$ is admissible with respect to m, we have, by the induction hypothesis, that

$$I_2 \le c_0 (m+1)^2.$$

To complete the proof of Lemma 2.4, it remains to show that $J \ll m+1$. However, by applying Schwarz's inequality to J and using the estimates for I_1 and I_2, we only get $J \ll (m+1)^{3/2}$. We therefore need extra ideas.

Note that by periodicity in s and t, we have

$$J = \int_0^1 \int_0^1 E_1(s + a2^{-m}, t + b2^{-m}) E_2(s + a2^{-m}, t + b2^{-m}) \mathrm{d}s\mathrm{d}t \tag{2.10}$$

for every pair of integers a, b. Furthermore, by (2.8), $E_1(s,t)$ is concerned with the lattice Λ_m, and so it is in fact periodic in both s and t with period 2^{-m}. Hence, summing (2.10) over the ranges $0 \le a, b \le 2^m - 1$, we have

$$4^m J = \int_0^1 \int_0^1 E_1(s,t) D(s,t) \mathrm{d}s\mathrm{d}t,$$

where

$$D(s,t) = \sum_{a=0}^{2^m-1} \sum_{b=0}^{2^m-1} E_2(s + a2^{-m}, t + b2^{-m}).$$

By Schwarz's inequality and (2.9),

$$(4^m J)^2 \ll (m+1) \int_0^1 \int_0^1 |D(s,t)|^2 \mathrm{d}s\mathrm{d}t. \tag{2.11}$$

We next show that $D(s,t)$ is concerned with a very special Ω, where "the sheets Λ are exactly on top of each other". More precisely,

$$D(s,t) = Z[s\mathbf{u} + t\mathbf{i} + \Omega'; B^{**}] - 4^m V(B^{**}),$$

where Ω' is obtained from

$$\bigcup_{a=0}^{2^m-1} \bigcup_{b=0}^{2^m-1} (a2^{-m}\mathbf{u} + b2^{-m}\mathbf{i} + \Omega_{m+1})$$

by restricting the last coordinate to the interval $[\mu 4^m, (\mu+1)4^m)$; in other words,

$$\begin{aligned} \Omega' &= \mu 4^m \mathbf{k} + 2^{-m}\mathbf{q}_\mu + \bigcup_{a=0}^{2^m-1} \bigcup_{b=0}^{2^m-1} (a2^{-m}\mathbf{u} + b2^{-m}\mathbf{i} + \Omega_m) \\ &= \mu 4^m \mathbf{k} + 2^{-m}\mathbf{q}_\mu + \Omega'', \end{aligned}$$

say. In view of Lemma 2.3, it is not difficult to see that Ω'' is of the form

$$\Omega'' = \bigcup_{\nu=0}^{4^m-1} (\nu\mathbf{k} + \Lambda_m).$$

Let

$$B^{***} = [0, X^*) \times [0, Y^*) \times [\mu 4^m, -[-Z^*]).$$

Then

$$Z[s\mathbf{u} + t\mathbf{i} + \Omega'; B^{**}] = Z[s\mathbf{u} + t\mathbf{i} + \Omega'; B^{***}]$$

and

$$|4^m V(B^{**}) - 4^m V(B^{**})| \leq 2^m.$$

Then using the projection onto the xy–plane, we have

$$D(s,t) = (-[-Z^*] - \mu 4^m) E[s\mathbf{u} + t\mathbf{i} + 2^{-m}\mathbf{q}_\mu + \Lambda_m; R_0] + O(2^m).$$

Since $0 \leq -[-Z^*] - \mu 4^m \leq 4^m$, we have, by (2.9), that

$$\int_0^1 \int_0^1 |D(s,t)|^2 \mathrm{d}s\mathrm{d}t \ll 4^{2m}(m+1). \tag{2.12}$$

Combining (2.11) and (2.12), we have $J \ll m+1$. Lemma 2.4 follows.

§**3. The Classical Method**

We shall again follow the notation in §1.1. Let $K \geq 2$. We shall first prove the theorem of Halton–Hammersley [**19**,**20**].

THEOREM 2B. *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in U_0^K such that*

$$\sup_{\mathbf{x} \in U_1^K} |D[\mathcal{P}; B(\mathbf{x})]| \ll_K (\log N)^{K-1}.$$

Some 20 years later, Roth [**27**] was able to introduce extra probabilistic ideas to prove the following stronger result.

THEOREM 2D. *For every natural number $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in U_0^K such that*

$$\int_{U_1^K} |D[\mathcal{P}; B(\mathbf{x})]|^2 \mathrm{d}\mathbf{x} \ll_K (\log N)^{K-1}.$$

§3.1. Some Old Ideas

Theorem 2B was proved in the case $K = 2$ by Lerch [**22**] in 1904, some 31 years before the subject of irregularities of distribution was born. In 1935, van der Corput [**14**,**15**] gave an alternative proof of this special case, using what is known nowadays as the van der Corput sequence. The argument of Halton–Hammersley is essentially a generalization of this idea to higher dimensions.

For notational convenience, we shall write $K = k + 1$ throughout this section.

The first idea is to consider, instead of sets of N points in U_0^{k+1}, infinite sets of points in $U_0^k \times [0, \infty)$ such that there is an average of one point per unit volume. We then consider those N points contained in $U_0^k \times [0, N)$, and rescale the last coordinate to obtain a set of N points in U_0^{k+1}.

We shall be concerned with boxes in $U_0^k \times [0, \infty)$ of the form

$$I_1 \times \ldots \times I_k \times I_0, \tag{3.1}$$

where, for each $j = 1, \ldots, k$, I_j is an interval of the form $[\alpha_j, \beta_j)$ and contained in U_0, while I_0 is an interval of the form $[\alpha_0, \beta_0)$ satisfying $0 \leq \alpha_0 < \beta_0$.

The second idea is to look for distributions such that many boxes of the type (3.1) contain the right number of points (*i.e.* equal to the volume of the boxes), while making sure that all other boxes can be approximated to a finite union of these boxes.

Definition. Let p be a prime and s be a non–negative integer. By an elementary p–type interval of order s, we mean an interval of the type $[\alpha, \beta)$, contained in U_0, where α and β are consecutive integer multiples of p^{-s}.

Let h be a non–negative integer, and let $p_1, \ldots, p_k$ be primes, not necessarily distinct.

Definition. By an elementary box of order h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$, we mean a set of the form (3.1), where, for each $j = 1, \ldots, k$, I_j is an elementary p_j–type interval of order s_j $(0 \leq s_j \leq h)$, and where I_0 is an interval of the form $[\alpha_0, \beta_0)$, with α_0, β_0 being consecutive non–negative integer multiples of $p_1^{s_1} \ldots p_k^{s_k}$.

It is clear that any elementary box of order h in $U_0^k \times [0, \infty)$ has volume 1.

Definition. By a special set of class h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$, we mean an infinite set $\mathcal{Q}$ of points in $U_0^k \times [0, \infty)$ which has the property that every elementary box of order h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$ contains exactly one point of $\mathcal{Q}$.

In 1935, van der Corput proved the following.

Lemma 3.1. *For each $n \in \mathbb{N} \cup \{0\}$, write*

$$n = \sum_{\nu=1}^{\infty} a_\nu 2^{\nu-1} \qquad (0 \le a_\nu < 2),$$

where the integers a_ν are uniquely determined by n, and write

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

Then the set

$$\{(x(n), n) : n \in \mathbb{N} \cup \{0\}\}$$

is a special set of class h with respect to the prime 2 in $U_0 \times [0, \infty)$ for any non–negative integer h.

To prove Lemma 3.1, simply note that if I is an elementary 2–type interval of order s, then the relation $x(n) \in I$ is satisfied by precisely all the non–negative integers of a residue class modulo 2^s. Using this idea on different primes and using the Chinese remainder theorem, we can prove the following generalization by Halton–Hammersley.

Lemma 3.2. *Suppose that $p_1, \ldots, p_k$ are distinct primes. For each $n \in \mathbb{N} \cup \{0\}$ and each $j = 1, \ldots, k$, write*

$$n = \sum_{\nu=1}^{\infty} a_{j,\nu} p_j^{\nu-1} \qquad (0 \le a_{j,\nu} < p_j),$$

where the integers $a_{j,\nu}$ are uniquely determined by n, and write

$$x_j(n) = \sum_{\nu=1}^{\infty} a_{j,\nu} p_j^{-\nu}.$$

Then the set

$$\{(x_1(n), \ldots, x_k(n), n) : n \in \mathbb{N} \cup \{0\}\} \tag{3.2}$$

is a special set of class h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$ for any non–negative integer h.

For any special set $\mathcal{Q}$ and any box B of the type (3.1) in $U_0^k \times [0, \infty)$, 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 volume of B. Note that if $B = B_1 \cup B_2$, where $B_1 \cap B_2 = \emptyset$, then

$$E[\mathcal{Q}; B] = E[\mathcal{Q}; B_1] + E[\mathcal{Q}; B_2]. \tag{3.3}$$

Definition. Let s be a non–negative integer. Suppose that $B^* = I_1 \times \ldots \times I_k \times I^*$, where for each $j = 1, \ldots, k$, $I_j = [0, \eta_j)$, where $0 < \eta_j \leq 1$ and η_j is an integer multiple of p_j^{-s}. Suppose further that $I^* = [0, Y)$, where Y is positive but otherwise unrestricted. Then we say that B^* is a box of class s with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$.

Theorem 2B can be deduced from Lemma 3.2 and the following lemma.

Lemma 3.3. *For any special set of class h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$ and for any box B^* of class h with respect to the same primes in $U_0^k \times [0, \infty)$, we have*

$$|E[\mathcal{Q}; B^*]| < (p_1 \ldots p_k) h^k. \tag{3.4}$$

Proof. Note that since each of the $I_j = [0, \eta_j)$ is a disjoint union of at most $p_j h$ elementary p_j–type intervals of order at most h, B^* is a disjoint union of at most $(p_1 \ldots p_k) h^k$ boxes of the type

$$I_1 \times \ldots \times I_k \times [0, Y), \tag{3.5}$$

where, for each $j = 1, \ldots, k$, I_j is an elementary p_j–type interval of order s_j where $0 \leq s_j \leq h$. To prove (3.4), it suffices to prove, in view of (3.3), that for each box B of the type (3.5), $|E[\mathcal{Q}; B]| < 1$. Let Y_0 denote the greatest integer multiple of $p_1^{s_1} \ldots p_k^{s_k}$ not exceeding Y. Then $[0, Y) = [0, Y_0) \cup [Y_0, Y)$. The box $I_1 \times \ldots \times I_k \times [0, Y_0)$ is the union of a finite number of elementary boxes of order h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$, and so contains the expected number of points of $\mathcal{Q}$. Hence $E[\mathcal{Q}; I_1 \times \ldots \times I_k \times [0, Y_0)] = 0$. On the other hand, the remainder $I_1 \times \ldots \times I_k \times [Y_0, Y)$ is contained in an elementary box of order h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$, and so contains at most one point of $\mathcal{Q}$ and has volume less than 1. It follows that $|E[\mathcal{Q}; I_1 \times \ldots \times I_k \times [Y_0, Y)]| < 1$, and the proof of the lemma is complete. ♣

We now deduce Theorem 2B. Let the natural number $N \geq 2$ be given. Let $p_1, \ldots, p_k$ be the first k primes, and let $\mathcal{Q}$ be the set (3.2) in Lemma 3.2. We choose integer h to satisfy

$$2^{h-1} < N \leq 2^h. \tag{3.6}$$

For any $\mathbf{x} = (x_1, \ldots, x_k) \in U_1^k$ and for any Y satisfying

$$0 < Y \leq N, \tag{3.7}$$

let $B(\mathbf{x}, Y)$ denote the box

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

Let $\mathbf{y} = \mathbf{y}(\mathbf{x}) = (y_1, \ldots, y_k)$ be defined such that for every $j = 1, \ldots, k$,

$$y_j = y_j(x_j) = -p_j^{-h}[-p_j^h x_j];$$

in other words, y_j is the least integer multiple of p_j^{-h} not less than x_j.

Lemma 3.4. *For any* $\mathbf{x} \in U_1^k$ *and any* Y *satisfying* (3.7), *we have*

$$|E[\mathcal{Q}; B(\mathbf{x}; Y)] - E[\mathcal{Q}; B^*(\mathbf{y}(\mathbf{x}); Y)]| \leq k.$$

Proof. For $j = 0, 1, \ldots, k$, for any fixed $\mathbf{x} \in U_1^k$ and for any fixed Y satisfying (3.7), let

$$B^{(j)} = [0, y_1) \times \ldots \times [0, y_j) \times [0, x_{j+1}) \times \ldots \times [0, x_k) \times [0, Y).$$

Then clearly $B^{(0)} = B(\mathbf{x}, Y)$ and $B^{(k)} = B^*(\mathbf{y}(\mathbf{x}), Y)$. To prove Lemma 3.4, it suffices to show that for every $j = 1, \ldots, k$, we have

$$|E[\mathcal{Q}; B^{(j)}] - E[\mathcal{Q}; B^{(j-1)}]| \leq 1.$$

This follows from observing that $B^{(j-1)} \subseteq B^{(j)}$ and that the complement of $B^{(j-1)}$ in $B^{(j)}$ is contained in an elementary box of order h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$. ♣

It now follows from Lemmas 3.3 and 3.4 that for any box $B(\mathbf{x}, Y)$ with $\mathbf{x} \in U_1^k$ and $0 < Y \leq N$, we have

$$|E[\mathcal{Q}; B(\mathbf{x}, Y)]| < (p_1 \ldots p_k) h^k + k. \tag{3.8}$$

The box $[0, 1)^k \times [0, N)$ contains precisely the points

$$(x_1(n), \ldots, x_k(n), n) \qquad (n = 0, 1, \ldots, N - 1).$$

Hence the set

$$\mathcal{P} = \{(x_1(n), \ldots, x_k(n), n/N) : 0 \leq n < N\}$$

satisfies, in view of (3.6) and (3.8),

$$\sup_{\mathbf{x} \in U_1^{k+1}} |D[\mathcal{P}; B(\mathbf{x})]| \ll_k (\log N)^k.$$

This completes the proof of Theorem 2B.

§3.2. Some Simple Truths

The existence of special sets of any class with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$ has been studied more closely recently. Halton–Hammersley showed in Lemma 3.2 that if $p_1, \ldots, p_k$ are distinct, then such sets exist. Faure [**17**] showed in 1982 that

- if the primes $p_1, \dots, p_k$ are all equal and if their common value is at least k, then such sets exist; and
- if the primes $p_1, \dots, p_k$ are all equal and if their common value is less than k, then such sets do not exist unless $h = 0$.

I [**13**] observed in 1983 that
- if $k = 3$ and $p_1 = p_2 = 2$ and $p_3 = 3$, then such sets do not exist unless $h \le 1$; and
- if $k = 3$ and $p_1 = 2$ and $p_2 = p_3 = 3$, then such sets do not exist unless $h = 0$.

So perhaps the following is true. However, any counterexample will be very interesting.

Conjecture. *Suppose that $p_1, \dots, p_k$ are primes, not all distinct but not all equal. Then there exists a positive integer h_0, depending at most on $p_1, \dots, p_k$, such that for every $h \ge h_0$, there are no special sets of class h with respect to the primes $p_1, \dots, p_k$ in $U_0^k \times [0, \infty)$.*

§3.3. More Old Ideas

The oldest idea in probability theory concerns taking an average. Accordingly, we shall modify the set $\mathcal{Q}$ used in §3.1 and average the discrepancy over such modifications. For the sake of convenience, we make a few technical refinements along the way.

Again, let $p_1, \dots, p_k$ denote the first k primes. Consider the subset

$$\{(x_1(n), \dots, x_k(n), n) : 0 \le n < (p_1 \dots p_k)^h\} \tag{3.9}$$

of the set (3.2). We then extend the set (3.9) by periodicity as follows. Writing $\mathbf{X}(n) = (X_1(n), \dots, X_k(n))$, we define $X_j(n) = x_j(n)$ for all $j = 1, \dots, k$ if $0 \le n < (p_1 \dots p_k)^h$; and let

$$\mathbf{X}(n) = \mathbf{X}(n + (p_1 \dots p_k)^h)$$

otherwise. Let

$$\Omega = \{(\mathbf{X}(n), n) : n \in \mathbb{Z}\}. \tag{3.10}$$

For any $t \in \mathbb{R}$, we define the translation $\Omega(t)$ by

$$\Omega(t) = \{(\mathbf{X}(n), n + t) : n \in \mathbb{Z}\}. \tag{3.11}$$

Clearly $\Omega(0) = \Omega$. Also, note that the subset $\{(\mathbf{X}(n), n + t) : n + t \ge 0\}$ of $\Omega(t)$ is a special set of class h with respect to the primes $p_1, \dots, p_k$ in $U_0^k \times [0, \infty)$.

For any box B in $U_0^k \times [0, \infty)$, let $Z[\Omega(t); B]$ denote the number of points of $\Omega(t)$ in B, and write

$$E[\Omega(t); B] = Z[\Omega(t); B] - \mu(B).$$

Then Theorem 2D can easily be deduced from the following lemma. Let

$$M = p_1 \dots p_k. \tag{3.12}$$

Lemma 3.5. *For any box B^* of class h with respect to the primes $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$, we have*

$$\int_0^{M^h} |E[\Omega(t); B^*]|^2 \mathrm{d}t < (4h)^k (p_1 \ldots p_k)^2 M^h.$$

Proof of Theorem 2D. For any natural number $N \geq 2$, let the integer h satisfy (3.6). For any $\mathbf{x} = (x_1, \ldots, x_k) \in U_1^k$ and for any Y satisfying $0 < Y \leq N$, let $B(\mathbf{x}, Y) = [0, x_1) \times \ldots \times [0, x_k) \times [0, Y)$. Then in view of Lemmas 3.4 and 3.5, we have

$$\int_0^{M^h} |E[\Omega(t); B^*]|^2 \mathrm{d}t \ll_k M^h h^k$$

for every $B(\mathbf{x}, Y)$. Hence

$$\int_0^{M^h} \int_0^N \int_{U^K} |E[\Omega(t); B^*]|^2 \mathrm{d}\mathbf{x} \mathrm{d}Y \mathrm{d}t \ll_k N M^h h^k,$$

so that there exists a real number t^*, satisfying $0 \leq t^* < M^h$, such that

$$\int_0^N \int_{U^K} |E[\Omega(t^*); B^*]|^2 \mathrm{d}\mathbf{x} \mathrm{d}Y \ll_k N (\log N)^k.$$

It follows that the set

$$\mathcal{P} = \{(X_1(n), \ldots, X_k(n), N^{-1}(n + t^*)) : 0 \leq n + t^* < N\}$$

of N points in U_0^{k+1} gives the desired result. ♣

It remains to prove Lemma 3.5. To do this, it is convenient to express $E[\Omega(t); B^*]$ as a sum of a finite number of 1–dimensional discrepancy functions.

We use R to denote a residue class. In particular, $R(m, q)$ denotes the residue class of integers congruent to m modulo q. For any $t \in \mathbb{R}$, we denote by $t + R$ the set $\{t + n : n \in R\}$, and let $Z[t + R; I^*]$ denote the number of elements of $t + R$ that fall into the interval I^*, and write

$$F[t + R; I^*] = Z[t + R; I^*] - q^{-1} l(I^*), \tag{3.13}$$

where q is the modulus of the residue class R, and where $l(I^*)$ denotes the length of I^*. It is obvious that

$$|F[t + R; I^*]| \leq 1 \tag{3.14}$$

always.

Suppose that $B^* = [0, \eta_1) \times \ldots \times [0, \eta_k) \times I^*$, where $I^* = [0, Y)$ and where, for each $j = 1, \ldots, k$, $[0, \eta_j)$ is a union of $L_j < p_j h$ elementary p_j–type intervals I_j of order at most h and such that there are at most $(p_j - 1)$ elementary p_j–type intervals of any one order in the union.

Lemma 3.6. *Suppose that for each $j = 1, \ldots, k$, I_j is an elementary p_j–type interval of order s_j, where $0 \le s_j \le h$. Then there is precisely one residue class R modulo $p_1^{s_1} \ldots p_k^{s_k}$ such that*

$$E[\Omega(t); I_1 \times \ldots \times I_k \times I^*] = F[t + R; I^*].$$

Proof. Suppose that $0 \le n < (p_1 \ldots p_k)^h$ and $X_j(n) \in I_j$. Then since $X_j(n) = \sum_{\nu=1}^{\infty} a_{j,\nu} p_j^{-\nu}$ where $n = \sum_{\nu=1}^{\infty} a_{j,\nu} p_j^{\nu-1}$, the numbers $a_{j,1}, \ldots, a_{j,s_j}$ are determined uniquely, but the remaining $a_{j,\nu}$ are left arbitrary. Since $p_j^{s_j}$ divides $(p_1 \ldots p_k)^h$, it follows from periodicity that those integers $n \in \mathbb{Z}$ for which $X_j(n) \in I_j$ constitute precisely a residue class modulo $p_j^{s_j}$. By the Chinese remainder theorem, there exists a unique residue class R modulo $p_1^{s_1} \ldots p_k^{s_k}$ such that

$$\mathbf{X}(n) \in I_1 \times \ldots \times I_k \qquad \text{if and only if} \qquad n \in R.$$

The lemma follows immediately. ♣

Suppose that for each $j = 1, \ldots, k$, we have

$$[0, \eta_j) = \bigcup_{l_j=1}^{L_j} I_{j,l_j},$$

where for every $l_j = 1, \ldots, L_j$, I_{j,l_j} is an elementary p_j–type interval. Then

$$B^* = \bigcup_{l_1=1}^{L_1} \ldots \bigcup_{l_k=1}^{L_k} (I_{1,l_1} \times \ldots \times I_{k,l_k} \times I^*).$$

Writing $\mathbf{l} = (l_1, \ldots, l_k)$ and writing $R(\mathbf{l})$ for the residue class such that

$$E[\Omega(t); I_{1,l_1} \times \ldots \times I_{k,l_k} \times I^*] = F[t + R(\mathbf{l}); I^*],$$

we have, since E is additive,

$$E[\Omega(t); B^*] = \sum_{l_1=1}^{L_1} \ldots \sum_{l_k=1}^{L_k} F[t + R(\mathbf{l}); I^*],$$

so that, omitting reference to I^*, we have

$$\int_0^{M^h} |E[\Omega(t); B^*]|^2 \mathrm{d}t$$
$$= \sum_{l_1'=1}^{L_1} \ldots \sum_{l_k'=1}^{L_k} \sum_{l_1''=1}^{L_1} \ldots \sum_{l_k''=1}^{L_k} \int_0^{M^h} F[t + R(\mathbf{l}')] F[t + R(\mathbf{l}'')] \mathrm{d}t \qquad (3.15)$$

Clearly, application of (3.14) alone is not sufficient to give the desired result. Roth proved the following.

Lemma 3.7. *Suppose that for each $j = 1, \ldots, k$, we have $0 \leq s'_j, s''_j \leq h$. Suppose further that $m', m'' \in \mathbb{Z}$. Then, writing, for each $j = 1, \ldots, k$, $u_j = \min\{s'_j, s''_j\}$ and $d_j = |s'_j - s''_j|$, we have*

$$\begin{aligned}\int_0^{M^h} & F[t + R(m', p_1^{s'_1} \ldots p_k^{s'_k})]F[t + R(m'', p_1^{s''_1} \ldots p_k^{s''_k})]\mathrm{d}t \\ &= p_1^{-d_1} \ldots p_k^{-d_k} \int_0^{M^h} F[t + R(m', p_1^{u_1} \ldots p_k^{u_k})]F[t + R(m'', p_1^{u_1} \ldots p_k^{u_k})]\mathrm{d}t. \quad (3.16)\end{aligned}$$

Proof. We shall only show that

$$\begin{aligned} I &= \int_0^{M^h} F[t + R(m', p_1^{s'_1} P')]F[t + R(m'', p_1^{s''_1} P'')]\mathrm{d}t \\ &= p_1^{-d_1} \int_0^{M^h} F[t + R(m', p_1^{u_1} P')]F[t + R(m'', p_1^{u_1} P'')]\mathrm{d}t, \quad (3.17)\end{aligned}$$

where $P' = p_2^{s'_2} \ldots p_k^{s'_k}$ and $P'' = p_2^{s''_2} \ldots p_k^{s''_k}$. (3.16) then follows by repeating the argument on the other primes. To prove (3.17), we may assume, without loss of generality, that $s'_1 \leq s''_1$, so that $u_1 = s'_1$. Then the function $F[t + R(m', p_1^{s'_1} P')]$ is periodic in t with period $p_1^{u_1}(p_2 \ldots p_k)^h$, so that

$$F[t + ap_1^{u_1}(p_2 \ldots p_k)^h + R(m', p_1^{s'_1} P')]$$

is independent of the choice of the integer a. Furthermore, since $p_1^{u_1}(p_2 \ldots p_k)^h$ divides M^h, the period of the integrand in (3.17), we have

$$I = \int_0^{M^h} F[t + R(m', p_1^{u_1} P')]F[t + ap_1^{u_1}(p_2 \ldots p_k)^h + R(m'', p_1^{s''_1} P'')]\mathrm{d}t$$

for every integer a. Hence

$$\begin{aligned} p^{d_1} I &= \sum_{a=1}^{p^{d_1}} \int_0^{M^h} F[t + R(m', p_1^{u_1} P')]F[t + ap_1^{u_1}(p_2 \ldots p_k)^h + R(m'', p_1^{s''_1} P'')]\mathrm{d}t \\ &= \int_0^{M^h} F[t + R(m', p_1^{u_1} P')]F[t + R(m'', p_1^{u_1} P'')]\mathrm{d}t. \end{aligned}$$

♣

It now follows from (3.15) and Lemma 3.7 that if for each $j = 1, \ldots, k$ and $l_j = 1, \ldots, L_j$, the interval I_{j,l_j} is of order $s(j, l_j)$, then in view of (3.14),

$$M^{-h} \int_0^{M^h} |E[\Omega(t); B^*]|^2 \mathrm{d}t \leq \sigma_1 \ldots \sigma_k,$$

where, for $j = 1, \ldots, k$,

$$\sigma_j = \sum_{l'_j=1}^{L_j} \sum_{l''_j=1}^{L_j} p_j^{-|s(j,l'_j)-s(j,l''_j)|}.$$

We can write

$$\sigma_j = \sum_{b=0}^{h} \sum_{\substack{1 \leq l'_j \leq L_j \\ 1 \leq l''_j \leq L_j \\ \min\{s(j,l'_j), s(j,l''_j)\}=b}} p_j^{-|s(j,l'_j)-s(j,l''_j)|}.$$

If $L_j = 1$, we have $\sigma_j = 1$. If $L_j > 1$, then since there are at most $(p_j - 1)$ elementary p_j–type intervals of any fixed order, we have

$$\sigma_j < 2 \sum_{b=1}^{h} \sum_{\substack{1 \leq l'_j \leq L_j \\ 1 \leq l''_j \leq L_j \\ b=s(j,l'_j) \leq s(j,l''_j)}} p_j^{-|s(j,l'_j)-s(j,l''_j)|} \leq 2h(p_j - 1)^2 \sum_{d=0}^{\infty} p_j^{-d} < 4hp_j^2.$$

This completes the proof of Lemma 3.5.

§3.4. More Simple Truths

The same construction by Roth is sufficient to give Theorem 2E, but not sufficient to prove it. To prove Theorem 2E, we need to consider a whole class of distributions similar to that constructed by Roth, and use induction on both h and the dimension k.

On the other hand, recall §3.2. If $p_1 = \ldots = p_k \geq k$, then Faure showed that special sets of any class with respect to $p_1, \ldots, p_k$ in $U_0^k \times [0, \infty)$ exist. However, Roth's method using the variable t fails to give an alterntaive proof of Theorem 2D. In fact, I [**13**] showed that there is a way to give a proof of Theorem 2E, the stronger version of Theorem 2D, using special sets described in this section, as long as such sets can be shown to exist.

§4. A Combinatorial and Geometric Approach

In this section, we consider the problem of convex polygons, and prove a result which will imply Theorem 8B. Let $l \geq 2$, and let $\underline{\theta} = (\theta_1, \ldots, \theta_l)$ satisfy $0 \leq \theta_1 < \ldots < \theta_l < \pi$. For each $i = 1, \ldots, l$, let $\mathbf{e}_i = (\cos\theta_i, \sin\theta_i)$, and denote by $\mathrm{POL}^\infty(\underline{\theta})$ 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 8C. *For every $\epsilon > 0$, there exists an infinite discrete set $\mathcal{Q} \subseteq \mathbb{R}^2$ such that for every $A \in \mathrm{POL}^\infty(\underline{\theta})$ with $d(A) \geq 2$,*

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

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

The proof of Theorem 8C is based on a combination of combinatorial and geometric arguments. We shall duscuss the (short) combinatorial part in §4.1 and the (lengthy) geometric part in §4.2.

§4.1. A Combinatorial Lemma

The combinatorial part of the argument is summarized by the following lemma.

Lemma 4.1. *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 only on ϵ, and integers $a_1, \ldots, a_p \in \{0,1\}$ such that

$$\left| \sum_{x_k \in Y_j^{(i)}} (a_k - \alpha_k) \right| < c(\epsilon) i^{1+\epsilon} \tag{4.1}$$

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

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 $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p) \in [0,1]^p$. We shall construct a sequence

$$\underline{\alpha}_0, \underline{\alpha}_1, \underline{\alpha}_2, \ldots, \underline{\alpha}_\nu = (\alpha_{1,\nu}, \ldots, \alpha_{p,\nu}), \ldots \tag{4.2}$$

of vectors in $[0,1]^p$ with the following properties. Let

$$X_\nu = \{x_k \in X : \alpha_{k,\nu} \notin \{0,1\}\}.$$

Then we need

$$X_{\nu+1} \subsetneqq X_\nu, \tag{4.3}$$

$$\alpha_{k,\nu} \in \{0,1\} \Rightarrow \alpha_{k,\nu} = \alpha_{k,\nu+1}, \tag{4.4}$$

and

$$\sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu} = \sum_{x_k \in Y_j^{(i)}} \alpha_{k,\nu+1} \tag{4.5}$$

for all i and j with $\#(Y_j^{(i)} \cap X_\nu) \geq c(\epsilon) i^{1+\epsilon}$. We shall construct the sequence (4.2) by induction.

Let $\underline{\alpha}_0 = \underline{\alpha}$. Suppose that $\underline{\alpha}_\nu$ has been defined and X_ν in non–empty. Let

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

We claim that

$$\#\mathcal{Z}_\nu < \#X_\nu. \tag{4.6}$$

To see this, note that $Y_j^{(i)} \cap Y_k^{(i)} = \emptyset$ whenever $j \neq k$, so that

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

if we choose

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

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

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

and with $y_k = 0$ for all $x_k \in X \setminus X_\nu$. In view of (4.6), 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 value for which the inequalities

$$0 \leq \alpha_{k,\nu} + t_0 y_k \leq 1 \qquad (x_k \in X_\nu)$$

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

$$\alpha_{k,\nu+1} = \alpha_{k,\nu} + t_0 y_k.$$

Then (4.3) clearly holds, in view of the maximality of t_0. On the other hand, (4.4) follows on noting that if $\alpha_{k,\nu} \in \{0,1\}$, then $x_k \in X \setminus X_\nu$ and so $y_k = 0$. It now follows from (4.3) that the sequence $\underline{\alpha}_0, \underline{\alpha}_1, \underline{\alpha}_2, \ldots$ will remain constant after a finite number of steps (s steps, say). Then $X_s = \emptyset$ and the vector $\underline{\alpha}_s$ has coordinates 0 and 1 only. For every $k = 1, \ldots, p$, we now let $a_k = \alpha_{k,s}$. Then it follows from (4.4) and (4.5) that (4.1) holds for all $Y_j^{(i)} \in \mathcal{Y}^{(i)}$ satisfying $i \geq 1$ and $j \geq 1$. This completes the proof of Lemma 4.1.

§4.2. A Geometric Lemma

We shall consider the family $\mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$ 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 aim in this section is to approximate the characteristic function of an arbitrary polygon in $\mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$ by 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

$$[m_1 2^{n_1}, (m_1+1)2^{n_1}) \times [m_2 2^{n_2}, (m_2+1)2^{n_2}), \tag{4.7}$$

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 $1 \leq i \leq l$. 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 $1 \leq i \leq l$. 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 $i = 1, \ldots, l$, write

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

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

For any $i = 1, \ldots, l$ 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

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

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

Definition. Suppose that $1 \le i \le l$ 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 $1 \le i \le l$ 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 $i = 1, \ldots, l$, let ψ_i^* denote the linear transformation of determinant 1 represented in matrix notation by

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

Let U^2 denote the unit square $[0,1]^2$. Then it is not difficult to see that $P_i^* = \{\psi_i^*(\mathbf{x}) : \mathbf{x} \in U^2\}$ 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 $1 \le i \le l$ 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 (4.7), 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 $i = 1, \ldots, l$, 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 $P_i^{**} = \{\psi_i^{**}(\mathbf{x}) : \mathbf{x} \in U^2\}$ 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 $1 \le i \le l$ 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 (4.7), 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

Lemma 4.2. *Suppose that $A \in \mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$. Then there exist special triangles $T'_1, \ldots, T'_m$ and $T''_1, \ldots, T''_M$ of types $\in \{1, \ldots, l\}$, special parallelograms $P'_1, \ldots, P'_n$ and $P''_1, \ldots, P''_N$ of types $\in \{(0,0)\} \cup \{(i,j) : i = 1, \ldots, l$ and $j = 1, 2\}$ and 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*

$$\sum_{\nu=1}^{m} \epsilon'_\nu \chi_{T'_\nu} + \sum_{\beta=1}^{n} \delta'_\beta \chi_{P'_\beta} \le \chi_A \le \sum_{\nu=1}^{M} \epsilon''_\nu \chi_{T''_\nu} + \sum_{\beta=1}^{N} \delta''_\beta \chi_{P''_\beta} \tag{4.8}$$

and

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

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

$$\max\{m, M\} \ll l \log(d(A)+2)$$

and

$$\max\{n, N\} \ll l (\log(d(A)+2))^3.$$

The first step in the proof of Lemma 4.2 is to reduce the problem to one of investigating rectangles and triangles.

Lemma 4.3. *Every $A \in \mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$ 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

i) $P_1, \ldots, P_4$ are special rectangles of the same order and for every $\alpha = 1, \ldots, 4$, $d(P_\alpha) < 3d(A)$;
ii) for every $\beta = 1, \ldots, q_1$, R_β is an aligned rectangle and $d(R_\beta) < 5d(A)$;
iii) for every $\nu = 1, \ldots, q_2$, T_ν is a triangle of type $\in \{1, \ldots, l\}$ and $d(T_\nu) \leq d(A)$;
iv) $q_1 \leq 4l + 8$ and $q_2 \leq 4l + 6$; and
v) R_β ($\beta = 1, \ldots, q_1$) and T_ν ($\nu = 1, \ldots, q_2$) are pairwise disjoint (in the sense that the intersection has measure zero).

Proof. For $j = 1, 2$, denote the projection of A onto the x_j–axis by $A^{(j)}$, and let $L^{(j)}$ denote the length of the interval $A^{(j)}$. Suppose that $n_j \in \mathbb{Z}$ satisfies $2^{n_j - 1} < L^{(j)} \leq 2^{n_j}$. Then the interval $A^{(j)}$ is contained in the union of at most two intervals of the type $[m_j 2^{n_j}, (m_j + 1)2^{n_j}]$, 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_α ($\alpha = 1, \ldots, 4$) with the convention that they may not be distinct, and note that

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

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 $(2l + 4)$ 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 $(2l + 1)$ trapeziums. Let B denote one of these triangles or trapeziums, and 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 $\in \{1, \ldots, l\}$ and with diameters not exceeding $d(A)$. Furthermore,

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

where T' and T'' are disjoint rectangles with diameters not exceeding $\left((4d)^2 + d^2\right)^{1/2}$. Clearly $A \subseteq P^{(1)} \times A^{(2)}$, and $(P^{(1)} \times A^{(2)}) \setminus A$ is a (pairwise disjoint) union of at most $(4l + 6)$ triangles of type $\in \{1, \ldots, l\}$ and $(4l + 6)$ 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 $\left((4d)^2 + (2d)^2\right)^{1/2}$. ♣

Our next step is clearly to investigate these rectangles and triangles obtained from Lemma 4.3. We first of all study the rectangles.

Lemma 4.4. *Suppose that R is an aligned rectangle.*

i) There exist an integer $s \ll (\log(\mu(R)+2))^2$ and mutually disjoint special rectangles $R'_1, \ldots, R'_s$ such that

$$\bigcup_{\beta=1}^{s} R'_\beta \subseteq R$$

and

$$\mu\left(R \setminus \left(\bigcup_{\beta=1}^{s} R'_\beta\right)\right) \leq 1.$$

ii) There exist mutually disjoint special rectangles $R''_1, \ldots, R''_4$, with $\mu(R''_\beta) < 4\mu(R)$ 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

$$R \subseteq (R''_1 \cup \ldots \cup R''_4) \setminus \left(\bigcup_{\beta=5}^{t} R''_\beta\right)$$

and

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

The proof of Lemma 4.4 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 4.5. *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*

$$\bigcup_{\alpha=1}^{D} I_\alpha \subseteq [a,b)$$

and

$$\mu_0\left([a,b) \setminus \left(\bigcup_{\alpha=1}^{D} I_\alpha\right)\right) \leq \left(\frac{7}{8}\right)^D (b-a).$$

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 $\alpha \geq 2$ inductively such that

i) I_α is the longest special interval in $[a,b) \setminus (I_1 \cup \ldots \cup I_{\alpha-1})$;

ii) $I_1 \cup \ldots \cup I_\alpha$ is an interval; and

iii) if $[a,b) \setminus (I_1 \cup \ldots \cup I_{\alpha-1})$ is a union of two disjoint intervals, then I_α belongs to the longer of the two (any one if of equal length).

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

Proof of Lemma 4.4. Suppose that $R = [a_1, b_1) \times [a_2, b_2)$. For $j = 1, 2$, we now apply Lemma 4.5 to the interval $[a_j, b_j)$ and obtain special intervals $I_1^{(j)}, \ldots, I_{D_j}^{(j)}$, with $D_j \ll \log(\mu(R) + 2)$, such that

$$\bigcup_{\alpha_j=1}^{D_j} I_{\alpha_j}^{(j)} \subseteq [a_j, b_j)$$

and

$$\mu_0\left([a_j, b_j) \setminus \left(\bigcup_{\alpha_j=1}^{D_j} I_{\alpha_j}^{(j)}\right)\right) \leq \frac{b_j - a_j}{2\mu(R)}.$$

The family of special rectangles

$$I_{\alpha_1}^{(1)} \times I_{\alpha_2}^{(2)} \qquad (1 \leq \alpha_1 \leq D_1 \text{ and } 1 \leq \alpha_2 \leq D_2)$$

clearly satisfies the requirements of (i). To prove (ii), note first of all that for $j = 1, 2$, if $n_j \in \mathbb{Z}$ satisfies $2^{n_j - 1} < a_j \leq 2^{n_j}$, 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 (i) to each of these completes the proof. ♣

Next we study the triangles arising from Lemma 4.3. Note that they are of types $\in \{1, \ldots, l\}$.

Definition. Suppose that $1 \leq i \leq l$. 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 $1 \leq i \leq l$, 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 (in the sense of measure) 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

Lemma 4.6. *Suppose that $1 \le i \le l$, 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 4.3, 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 4.4 is therefore obvious.

Lemma 4.7. *Suppose that $1 \le i \le l$, and that $j = 1, 2$. Suppose further that P is a parallelogram of type (i,j).*

i) 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

$$\bigcup_{\beta=1}^{s} P'_\beta \subseteq P$$

and

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

ii) 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

$$P \subseteq (P''_1 \cup \ldots \cup P''_4) \setminus \left(\bigcup_{\beta=5}^{t} P''_\beta\right)$$

and

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

It remains to investigate nice triangles.

Lemma 4.8. *Suppose that $1 \le i \le l$, and that T is a nice triangle of type i.*

i) 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

$$\left(\bigcup_{\nu=1}^{s} T'_\nu\right) \cup \left(\bigcup_{\nu=1}^{s} P'_\nu\right) \subseteq T$$

and

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

ii) 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

$$T \subseteq T''_0 \setminus \left(\left(\bigcup_{\nu=1}^{t} T''_\nu\right) \cup \left(\bigcup_{\nu=1}^{q} P''_\nu\right)\right)$$

and

$$\mu\left(\left(T''_0 \setminus \left(\left(\bigcup_{\nu=1}^{t} T''_\nu\right) \cup \left(\bigcup_{\nu=1}^{q} P''_\nu\right)\right)\right) \setminus T\right) \le 1.$$

Proof. (i) 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

$$\left(\bigcup_{\nu=1}^{D} T_\nu\right) \cup \left(\bigcup_{\nu=1}^{D} P_\nu\right) \subseteq T \tag{4.9}$$

and

$$\mu\left(T \setminus \left(\left(\bigcup_{\nu=1}^{D} T_\nu\right) \cup \left(\bigcup_{\nu=1}^{D} P_\nu\right)\right)\right) \le 4^{-D}\mu(T). \tag{4.10}$$

To prove (4.9) and (4.10), note that T, being a nice triangle of type i, is the intersection of a special triangle T^* of type i and a half–plane 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 $\mathbf{v}', \mathbf{v}'', \mathbf{v}'_1, \mathbf{v}''_1$ 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 (4.9) and (4.10). (i) now follows from a suitable choice of D. To prove (ii), 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 $\in \{(0,0),(i,1),(i,2)\}$. (ii) now follows on applying (i) to this latter nice triangle. ♣

Lemma 4.2 now follows on combining Lemmas 4.3, 4.4, 4.6, 4.7 and 4.8.

§4.3. Completion of the Proof

In this section, we combine the combinatorial Lemma 4.1 and the geometric Lemma 4.2 to give a proof of Theorem 8C. Our strategy is as follows. Lemma 4.2 enables us to obtain information on the discrepancy function of any given convex polygon in $\mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$. 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. Let $X = \mathcal{P}$, and let the partitions be given by the various families of special objects. We shall use Lemma 4.1 to choose a suitable subset of $\mathcal{P}$ to use in our construction of the desired infinite discrete set $\mathcal{Q}$ in Theorem 8C.

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

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

Suppose that $A \in \mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$ is arbitrary. We shall first of all use Lemma 4.2 to investigate the discrepancy function of A. The following lemma is in a more general form than needed.

Lemma 4.9. *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 for every $\tau = 1, \ldots, q$,

$$|E[\mathcal{P}; B'_\tau]| \le D_2,$$

and that for every $\tau = 1, \ldots, r$,

$$|E[\mathcal{P}; B''_\tau]| \le D_2.$$

Then

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

Proof. Clearly

$$E[\mathcal{P}; A] = \sum_{\mathbf{p} \in A \cap \mathcal{P}} 1 - \mu(A) \le \sum_{\tau=1}^{r} \epsilon''_\tau \sum_{\mathbf{p} \in B''_\tau \cap \mathcal{P}} 1 - \mu(A)$$

$$
\begin{aligned}
&= \sum_{\tau=1}^{r} \epsilon''_\tau \left(\sum_{\mathbf{p} \in B''_\tau \cap \mathcal{P}} 1 - \mu(B''_\tau) \right) + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \mu(A) \right) \\
&= \sum_{\tau=1}^{r} \epsilon''_\tau E[\mathcal{P}; B''_\tau] + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \mu(A) \right) \\
&\le \sum_{\tau=1}^{r} |E[\mathcal{P}; B''_\tau]| + \left(\sum_{\tau=1}^{r} \epsilon''_\tau \mu(B''_\tau) - \sum_{\tau=1}^{r} \epsilon'_\tau \mu(B'_\tau) \right) \\
&\le D_2 r + D_1. \qquad (4.11)
\end{aligned}
$$

A similar argument gives

$$
-E[\mathcal{P}; A] \le D_2 q + D_1. \qquad (4.12)
$$

The result now follows on combining (4.11) and (4.12). ♣

Let $\mathrm{SPEC}^\infty(\underline{\theta}; x_1, x_2)$ denote the big family of all special triangles, special parallelograms and special rectangles defined in §4.2; in other words,

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

We now make use of the combinatorial information derived from Lemma 4.1.

Lemma 4.10. *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 $B \in \mathrm{SPEC}^\infty(\underline{\theta}; x_1, x_2)$ satisfying $d(B) \ge 1$, we have*

$$
\left| \sum_{\mathbf{p} \in B \cap \mathcal{P}} f(\mathbf{p}) \right| \ll_\epsilon \left(l(\log(d(B) + 2))^2 \right)^{1+\epsilon}. \qquad (4.13)
$$

Proof. We apply Lemma 4.1 with $X = \mathcal{P}$, and so have to introduce a sequence of partitions of $\mathcal{P}$. Let

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

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

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

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 $y \geq 1$,

$$\#\{C \in \mathrm{SET}^{\infty}(\underline{\theta}; x_1, x_2) : 1 \leq d(C) \leq y\} = \sum_{i=1}^{l} \#\{n \in \mathbb{Z} : 1 \leq d(\mathrm{ST}(i,n)) \leq y\}$$
$$+ \sum_{i=1}^{l}\sum_{j=1}^{2} \#\{\mathbf{n} \in \mathbb{Z}^2 : 1 \leq d(\mathrm{SP}(i,j,\mathbf{n})) \leq y\} + \#\{\mathbf{n} \in \mathbb{Z}^2 : 1 \leq d(\mathrm{SR}(\mathbf{n})) \leq y\}$$
$$\ll l\log(y+2) + l(\log(y+2))^2 \ll l(\log(y+2))^2. \tag{4.14}$$

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 $\{C \in \mathrm{SET}^{\infty}(\underline{\theta}; x_1, x_2) : 1 \leq d(C) \leq d(B)\}$ ordered in the way described. Lemma 4.10 now follows from Lemma 4.1 and (4.14). ♣

We now use Lemma 4.10 to construct 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 $[-\kappa,\kappa)^2$. We now apply Lemma 4.10. There exists a function $f : \mathcal{P} \to \{-\alpha, 1-\alpha\}$ such that for all polygons $B \in \mathrm{SPEC}^{\infty}(\underline{\theta}; x_1, x_2)$ satisfying $B \subseteq [-\kappa,\kappa)^2$ and $d(B) \geq 1$, we have (4.13). Writing $\mathcal{P}_k = \{\mathbf{p} \in \mathcal{P} : f(\mathbf{p}) = 1-\alpha\}$, we have

$$\sum_{\mathbf{p}\in B\cap\mathcal{P}} f(\mathbf{p}) = \sum_{\mathbf{p}\in B\cap\mathcal{P}_k} 1 - \kappa^{-2} \sum_{\mathbf{p}\in B\cap\mathcal{P}} 1. \tag{4.15}$$

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

$$\left|\sum_{\mathbf{p}\in B\cap\mathcal{P}} 1 - \kappa^2\mu(B)\right| \ll \kappa\sigma(\partial B) \ll \kappa^2, \tag{4.16}$$

where $\sigma(\partial B)$ denotes the length of the perimeter of B. It follows, on combining (4.13), (4.15) and (4.16), that

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

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

Suppose now that the polygon $B \in \mathrm{SPEC}^\infty(\underline{\theta}; x_1, x_2)$ satisfies $B \subseteq [-\kappa, \kappa)^2$ and $d(B) < 1$. Then $B \subseteq B_0$ for some $B_0 \in \mathrm{SPEC}^\infty(\underline{\theta}; x_1, x_2)$ with $1 \leq d(B_0) < 2$. Applying (4.13) and (4.15) to B_0, we have

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

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

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

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

$$|E[\mathcal{P}_k; B]| \ll_\epsilon \left(l(\log(d(B)+2))^2\right)^{1+\epsilon}. \tag{4.17}$$

It now follows that (4.17) holds for all $B \in \mathrm{SPEC}^\infty(\underline{\theta}; x_1, x_2)$ satisfying $B \subseteq [-\kappa, \kappa)^2$. Combining this with Lemmas 4.2 and 4.9, we conclude that

$$|E[\mathcal{P}_k; C]| \ll_\epsilon l^{2+\epsilon} (\log(d(C)+2))^{5+\epsilon} \tag{4.18}$$

for all $C \in \mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$ 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 8C.

Consider any arbitrary $A \in \mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$. 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 $\mathrm{POL}^\infty(\underline{\theta}; x_1, x_2)$. It follows from (4.18) that

$$\begin{aligned} |E[\mathcal{Q}; A]| &= \left| \sum_{\mathbf{q} \in A \cap \mathcal{Q}} 1 - \mu(A) \right| \\ &= \left| \#(A \cap \mathcal{P}_1) - \mu(A \cap [-2,2)^2) + \sum_{\substack{k=2^n \\ n \in \mathbb{N}}} (\#(A_k \cap \mathcal{P}_k) - \mu(A_k)) \right| \\ &\ll_\epsilon {\sum}^* l^{2+\epsilon} (\min\{\log(d(A)+2), k\})^{5+\epsilon}. \end{aligned} \tag{4.19}$$

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

$$\sum^* (\min\{\log(d(A) + 2), k\})^{5+\epsilon} \ll_\epsilon (\log(d(A) + 2))^{5+\epsilon}. \tag{4.20}$$

Theorem 8C now follows from (4.19) and (4.20).

§5. Another Probabilistic Approach

§5.1. A Simple Argument

In this section, we give a simple proof of a special case of a renormalized version of Theorem 6B. Suppose that $N = M^K$, where $M \in \mathbb{N}$. Consider the cube $[0, M)^K$, treated as a torus. Let A be a compact, convex set in $[0, M)^K$. If the radius $r(A)$ of the largest inscribed ball of A satisfies $r(A) \ll 1$, then the result is trivial. We therefore assume, without loss of generality, that $r(A) \gg 1$. We shall prove the following result.

Theorem 6C. *There exists an $N = M^K$–element set $\mathcal{Q}$ with the following properties:*

i) For every $\mathbf{l} = (l_1, \ldots, l_K) \in \mathbb{Z}^K \cap [0, M)^K$, the cube

$$Q(\mathbf{l}) = [l_1, l_1 + 1) \times \ldots \times [l_K, l_K + 1)$$

contains precisely one point of $\mathcal{Q}$.

ii) For any $\lambda \in (0, 1]$, any proper orthogonal transformation $\tau \in \mathcal{T}$ and any vector $\mathbf{u} \in \mathbb{R}^K$, we have

$$|\#(\mathcal{Q} \cap A(\lambda, \tau, \mathbf{u})) - \mu(A(\lambda, \tau, \mathbf{u}))| \ll_K (\sigma(\partial A))^{1/2} (\log M)^{1/2}, \tag{5.1}$$

where ∂A denotes the boundary surface of A, and where σ denotes the usual measure in $\mathbb{R}^{K-1}$.

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

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

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

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

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

$$\#\mathcal{G}^* \leq M^{c(K)}, \tag{5.2}$$

where $c(K)$ is a positive constant depending at most on K, and that for any $B \in \mathcal{G}$, there exists $B^-, B^+ \in \mathcal{G}^*$ 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 idea is classical probability theory. Note that

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

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

$$\{\mathbf{q}_\mathbf{l} : \mathbf{l} \in \mathbb{Z}^K \cap [0,M)^K\}$$

satisfies (5.1) simultaneously for all $A_1 \in \mathcal{G}^*$ with probability greater than 1/2.

Let $A_1 \in \mathcal{G}^*$. Clearly the subset

$$\widetilde{A_1} = \bigcup_{Q(\mathbf{l}) \subseteq A_1} Q(\mathbf{l})$$

has no discrepancy. It remains to consider $A_1 \setminus \widetilde{A_1}$. Let

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

Then clearly

$$\mathcal{L}(A_1) \ll_K \sigma(\partial A_1) \ll_K \sigma(\partial A). \tag{5.3}$$

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

$$\xi_\mathbf{l} = \begin{cases} 1 & (\mathbf{q}_\mathbf{l} \in A_1), \\ 0 & (\text{otherwise}). \end{cases}$$

Then

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

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

Lemma 5.1. *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*

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

In view of (5.3), we now take

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

and take γ_1 to be a sufficiently large absolute constant (depending on K) multiple of $(\sigma(\partial A))^{1/2}(\log M)^{1/2}$. Then elementary calculation gives

$$\tfrac{1}{2}M^{-c(K)} \ge \begin{cases} 2\mathrm{e}^{-\gamma_1/4} & (\gamma_1 \ge \beta_1), \\ 2\mathrm{e}^{-\gamma_1^2/4\beta_1} & (\gamma_1 \le \beta_1). \end{cases}$$

It follows from Lemma 5.1 and (5.2) that

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

It now follows that

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

This completes the proof of Theorem 6C.

§5.2. Preliminary Discussion on Theorem 10

In §§5.3–5.5, we shall indicate a proof of the case $L = 1$ of Theorem 10. The argument will also give a proof of Theorem 6A.

The spirit of the proof is similar to that in the previous section, although we need extra combinatorial and probabilistic ideas. Let A be given and fixed. Given any natural number N, we shall first of all construct a sequence

$$\mathbf{q}_0, \ldots, \mathbf{q}_{N-1} \tag{5.4}$$

of N points in U^K, and shall eventually consider some random version of the set

$$\{(\mathbf{q}_0, 0), (\mathbf{q}_1, \tfrac{1}{N}), \ldots, (\mathbf{q}_{N-1}, \tfrac{N-1}{N})\}$$

of N points in U^{K+1}. However, as we shall also discuss Theorem 6A at the same time, it is convenient to note that Theorem 10 will follow if we can show that the sequence (5.4) of N points in U^K satisfies

$$\frac{1}{N}\sum_{M=1}^{N} \int_0^1 \int_{\mathcal{T}} \int_{U^K} |D[\mathcal{Q}_M; A(\lambda, \tau, \mathbf{u})]|^2 \mathrm{d}\mathbf{u}\mathrm{d}\tau\mathrm{d}\lambda \ll_{A,W} N^{1-1/K}, \tag{5.5}$$

where $\mathcal{Q}_M = \{\mathbf{q}_0, \ldots, \mathbf{q}_{M-1}\}$ for $1 \le M \le N$, and where

$$D[\mathcal{Q}_M; A(\lambda, \tau, \mathbf{u})] = Z[\mathcal{Q}_M; A(\lambda, \tau, \mathbf{u})] - M\mu(A(\lambda, \tau, \mathbf{u})).$$

Here, and in §§5.3–5.5, our discrepancy functions $D[\mathcal{P}; A]$ always denote discrepancy functions in U^K and not U^{K+1}.

Before we can construct the sequence, we need some notation and terminology.

Let h be a natural number, to be fixed later. For every $s = 0, 1, \ldots, h$ and for every $c \in \mathbb{Z}$, let

$$I(s, c) = [c2^{-s}, (c+1)2^{-s}). \tag{5.6}$$

In other words, $I(s, c)$ is an interval of length 2^{-s} and whose endpoints are consecutive integer multiples of 2^{-s}.

We shall construct a finite sequence $\mathbf{q}_n$ $(0 \le n < 2^{Kh})$ of $2^{Kh} \ge N$ points in U^K such that the following is satisfied. For every $s = 0, 1, \ldots, h$ and for every non–negative integer c satisfying $c < 2^{K(h-s)}$, every set of the form

$$I(s, a_1) \times \ldots \times I(s, a_K)$$

in U^K, where $a_1, \ldots, a_K \in \mathbb{Z}$, contains exactly one point of

$$\{\mathbf{q}_n : c2^{Ks} \le n < (c+1)2^{Ks}\}.$$

We shall describe the combinatorial part of the argument in §5.3 and the probabilistic part of the argument in §5.4.

§5.3. A Combinatorial Approach

For every integer s satisfying $1 \le s \le h$, integers $\tau_1, \ldots, \tau_{s-1} \in \{0, 1, \ldots, 2^K - 1\}$ and vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{s-1} \in \{0, 1\}^K$, let

$$G[\tau_1, \ldots, \tau_{s-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{s-1}] : \{0, 1, \ldots, 2^K - 1\} \to \{0, 1\}^K$$

be a bijective mapping, with the convention that the mapping in the case $s = 1$ is denoted by $G[\emptyset]$. Given these mappings, we can define a bijective mapping

$$F : \{0, 1, \ldots, 2^{Kh} - 1\} \to \{0, 1, \ldots, 2^h - 1\}^K$$

as follows. Suppose that n is an integer satisfying $0 \le n < 2^{Kh}$. Write

$$n = \tau_h 2^{K(h-1)} + \tau_{h-1} 2^{K(h-2)} + \ldots + \tau_1, \tag{5.7}$$

where $\tau_1, \ldots, \tau_h \in \{0, 1, \ldots, 2^K - 1\}$. We now let $\mathbf{a}_1, \ldots, \mathbf{a}_h \in \{0, 1\}^K$ be the solution of the system of equations

$$\begin{cases} G[\emptyset](\tau_1) = \mathbf{a}_1, \\ G[\tau_1; \mathbf{a}_1](\tau_2) = \mathbf{a}_2, \\ G[\tau_1, \tau_2; \mathbf{a}_1, \mathbf{a}_2](\tau_3) = \mathbf{a}_3, \\ \qquad \vdots \\ G[\tau_1, \ldots, \tau_{s-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{s-1}](\tau_s) = \mathbf{a}_s, \\ \qquad \vdots \\ G[\tau_1, \ldots, \tau_{h-2}; \mathbf{a}_1, \ldots, \mathbf{a}_{h-2}](\tau_{h-1}) = \mathbf{a}_{h-1}, \\ G[\tau_1, \ldots, \tau_{h-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{h-1}](\tau_h) = \mathbf{a}_h. \end{cases} \tag{5.8}$$

Suppose now that for each integer $t = 1, \ldots, h$,

$$\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,K}) \in \{0, 1\}^K. \tag{5.9}$$

We now write

$$F_j(n) = a_{1,j} 2^{h-1} + a_{2,j} 2^{h-2} + \ldots + a_{h,j} \tag{5.10}$$

and let

$$F(n) = (F_1(n), \ldots, F_k(n)). \tag{5.11}$$

We next partition U^K into a sequence of 2^{Kh} smaller cubes

$$S(n) = I(h, F_1(n)) \times \ldots \times I(h, F_k(n)),$$

where, for every $j = 1, \ldots, K$ and every $n = 0, 1, ..., 2^{Kh} - 1$, the interval $I(h, F_j(n))$ is defined by (5.6)–(5.10).

Lemma 5.2. *Suppose that s is an integer satisfying $0 \leq s \leq h$. Then for every integer n_0, the set*

$$\bigcup_{\substack{0 \leq n < 2^{Kh} \\ n \equiv n_0 \pmod{2^{Ks}}}} S(n) \tag{5.12}$$

is a cube of the form

$$C(s, \mathbf{c}) = I(s, c_1) \times \ldots \times I(s, c_K) \subseteq U^K, \tag{5.13}$$

where $\mathbf{c} = (c_1, \ldots, c_K) \in \{0, 1, \ldots, 2^s - 1\}^K$. On the other hand, every cube of the form (5.13), where $\mathbf{c} = (c_1, \ldots, c_K) \in \{0, 1, \ldots, 2^s - 1\}^K$, is a union of the form (5.12) for some integer n_0.

Proof. Note that the condition $n \equiv n_0 \pmod{2^{Ks}}$ determines precisely the values of $\tau_1, \ldots, \tau_s$ in (5.7). We can therefore solve the system of equations

$$\begin{cases} G[\emptyset](\tau_1) = \mathbf{a}_1 \\ G[\tau_1; \mathbf{a}_1](\tau_2) = \mathbf{a}_2 \\ \qquad \vdots \\ G[\tau_1, \ldots, \tau_{s-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{s-1}](\tau_s) = \mathbf{a}_s \end{cases} \tag{5.14}$$

for $\mathbf{a}_1, \ldots, \mathbf{a}_s$. On the other hand, $\tau_{s+1}, \ldots, \tau_h$ in (5.7) can take all possible values. It follows from

$$\begin{cases} G[\tau_1, \ldots, \tau_s; \mathbf{a}_1, \ldots, \mathbf{a}_s](\tau_{s+1}) = \mathbf{a}_{s+1} \\ \quad\vdots \\ G[\tau_1, \ldots, \tau_{h-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{h-1}](\tau_h) = \mathbf{a}_h \end{cases} \tag{5.15}$$

that $\mathbf{a}_{s+1}, \ldots, \mathbf{a}_h$ can take all possible values. The first assertion follows. To prove the second assertion, simply note that $\tau_1, \ldots, \tau_s$ are determined uniquely with given $\mathbf{a}_1, \ldots, \mathbf{a}_s$ by (5.14), and that if $\mathbf{a}_{s+1}, \ldots, \mathbf{a}_h$ take all possible values, then $\tau_{s+1}, \ldots, \tau_h$ take all possible values in view of (5.15). ♣

For every $\mathbf{c} = (c_1, \ldots, c_K) \in \{0, 1, \ldots, 2^h - 1\}^K$, let $\mathbf{q}(\mathbf{c})$ be a point in the cube

$$C(h; \mathbf{c}) = I(h, c_1) \times \ldots \times I(h, c_K) \subseteq U^K.$$

Using F, we can define a permutation $\mathbf{q}_n$ $(0 \le n < 2^{Kh})$ of the $\mathbf{q}(\mathbf{c})$ as follows. For $n = 0, 1, \ldots, 2^{Kh} - 1$, let

$$\mathbf{q}_n = \mathbf{q}(F(n)) = \mathbf{q}(F_1(n), \ldots, F_K(n)).$$

Clearly $\mathbf{q}_n \in S(n)$ for every $n = 0, 1, \ldots, 2^{Kh} - 1$. Then it follows from Lemma 5.2 that

Lemma 5.3. *Suppose that s and H are integers satisfying $0 \le s \le h$ and $0 \le H \le 2^{K(h-s)}$. Suppose further that $\mathbf{c} = (c_1, \ldots, c_K) \in \{0, 1, \ldots, 2^s - 1\}^K$. Then the cube* (5.13) *contains exactly one element of the set*

$$\{\mathbf{q}_n : H2^{Ks} \le n < (H+1)2^{Ks}\}.$$

Proof. The restriction $H2^{Ks} \le n < (H+1)2^{Ks}$ determines precisely the values of $\tau_{s+1}, \ldots, \tau_h$ in (5.7) with no restriction on $\tau_1, \ldots, \tau_s$. On the other hand, the restriction $\mathbf{q}_n \in C(s; \mathbf{c})$ for a given $\mathbf{c}$ determines precisely the values of $\mathbf{a}_1, \ldots, \mathbf{a}_s$ with no restriction on $\mathbf{a}_{s+1}, \ldots, \mathbf{a}_h$. The system of equations (5.14) now determines precisely the values of $\tau_1, \ldots, \tau_s$. Hence n is uniquely determined. ♣

We denote this element obtained by Lemma 5.3 by $\mathbf{q}(s; \mathbf{c}; H)$. In other words, for integers $s, c_1, \ldots, c_K, H$ satisfying the hypotheses of Lemma 5.3,

$$\mathbf{q}(s; \mathbf{c}; H) = \{\mathbf{q}_n : H2^{Ks} \le n < (H+1)2^{Ks}\} \cap C(s; \mathbf{c}).$$

§5.4. Some Probabilistic Lemmas

We now use some elementary concepts and facts from probability theory and define a "randomization" of the deterministic points $\mathbf{q}(\mathbf{c}) = \mathbf{q}(c_1, \ldots, c_K)$, mappings $G[\tau_1, \ldots, \tau_{s-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{s-1}]$ and F, and the sequence $\mathbf{q}_n$ as follows.

(A) For $\mathbf{c} = (c_1, \dots, c_K) \in \{0, 1, \dots, 2^h - 1\}^K$, let $\widetilde{\mathbf{q}}(\mathbf{c})$ be a random point uniformly distributed in the cube $C(h; \mathbf{c})$. More precisely,

$$\operatorname{Prob}(\widetilde{\mathbf{q}}(\mathbf{c}) \in \mathcal{S}) = \frac{\mu(C(h; \mathbf{c}) \cap \mathcal{S})}{\mu(C(h; \mathbf{c}))}$$

for all Borel sets $\mathcal{S} \subseteq \mathbb{R}^K$.

(B) For every integer s satisfying $1 \le s \le h$, integers $\tau_1, \dots, \tau_{s-1} \in \{0, 1, \dots, 2^K - 1\}$ and vectors $\mathbf{a}_1, \dots, \mathbf{a}_{s-1} \in \{0, 1\}^K$, let $\widetilde{G}[\tau_1, \dots, \tau_{s-1}; \mathbf{a}_1, \dots, \mathbf{a}_{s-1}]$ be a uniformly distributed random bijective mapping from $\{0, 1, ..., 2^K - 1\}$ to $\{0, 1\}^K$. More precisely, if $\pi : \{0, 1, ..., 2^K - 1\} \to \{0, 1\}^K$ is one of the $(2^K)!$ different (deterministic) bijective mappings, then

$$\operatorname{Prob}(\widetilde{G}[\tau_1, \dots, \tau_{s-1}; \mathbf{a}_1, \dots, \mathbf{a}_{s-1}] = \pi) = \frac{1}{(2^K)!}.$$

(C) Let $\widetilde{F}$ be the random bijective mapping from $\{0, 1, \dots, 2^{Kh} - 1\}$ to $\{0, 1, \dots, 2^h - 1\}^K$ defined by (5.7), $(\widetilde{5.8})$ and (5.9)–(5.11), where $(\widetilde{5.8})$ denotes that in the system (5.8) of equations, we replace each deterministic mapping by its corresponding random mapping.

(D) Let $\widetilde{\mathbf{q}}_n$ $(0 \le n < 2^{Kh})$ denote the random sequence defined by $\widetilde{F}$; in other words, for $n = 0, 1, \dots, 2^{Kh} - 1$,

$$\widetilde{\mathbf{q}}_n = \mathbf{q}(\widetilde{F}(n)).$$

(E) Let $\widetilde{\mathbf{q}}(s; \mathbf{c}; H)$ denote the randomization of $\mathbf{q}(s; \mathbf{c}; H)$; in other words, for integers $s, c_1, \dots, c_K, H$ satisfying the hypotheses of Lemma 5.3,

$$\widetilde{\mathbf{q}}(s; \mathbf{c}; H) = \{\widetilde{\mathbf{q}}_n : H 2^{Ks} \le n < (H+1) 2^{Ks}\} \cap C(s; \mathbf{c}). \tag{5.16}$$

(F) Finally, we may assume that the random variables

$$\widetilde{\mathbf{q}}(\mathbf{c}) \qquad (\mathbf{c} = (c_1, \dots, c_K) \in \{0, 1, \dots, 2^h - 1\}^K)$$

and

$$\widetilde{G}[\tau_1, \dots, \tau_{s-1}; \mathbf{a}_1, \dots, \mathbf{a}_{s-1}] \qquad \begin{aligned} (1 \le s \le h \text{ and } \tau_1, \dots, \tau_{s-1} \in \{0, 1, \dots, 2^K - 1\} \\ \text{and } \mathbf{a}_1, \dots, \mathbf{a}_{s-1} \in \{0, 1\}^K) \end{aligned}$$

are independent of each other. In fact, the existence of such a set of random variables follows immediately from the Kolmogorov extension theorem in probability theory.

Let $(\Omega, \mathcal{F}, \operatorname{Prob})$ denote the underlying probability measure space.

We have

Lemma 5.4. *Suppose that s and H are integers satisfying $0 \le s \le h$ and $0 \le H < 2^{K(h-s)}$. Then for every $\mathbf{c} = (c_1, \ldots, c_K) \in \{0, 1, \ldots, 2^s - 1\}^K$, the random point $\widetilde{\mathbf{q}}(s; \mathbf{c}; H)$ is uniformly distributed in the cube $C(s; \mathbf{c})$.*

Proof. Suppose that for $j = 1, \ldots, K$,

$$c_j = a_{1,j} 2^{s-1} + a_{2,j} 2^{s-2} + \ldots + a_{s,j}.$$

For $t = 1, \ldots, s$, let

$$\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,K}).$$

Since the random mapping $\widetilde{G}[\emptyset]$ is uniformly distributed, it follows that the (random) solution $\widetilde{\tau}_1$ of the equation

$$\widetilde{G}[\emptyset](\widetilde{\tau}_1) = \mathbf{a}_1$$

has the property that for any $\delta \in \{0, 1, \ldots, 2^K - 1\}$,

$$\text{Prob}(\widetilde{\tau}_1 = \delta) = 2^{-K}.$$

Now let $\widetilde{\tau}_1 = \tau_1$ (*i.e.* fix the value of this random variable), and consider the (random) equation

$$\widetilde{G}[\tau_1; \mathbf{a}_1](\widetilde{\tau}_2) = \mathbf{a}_2.$$

Since $\widetilde{G}[\tau_1; \mathbf{a}_1]$ is also uniformly distributed, we have, for any $\delta = \{0, 1, \ldots, 2^K - 1\}$, that

$$\text{Prob}(\widetilde{\tau}_2 = \delta | \tau_1 = \tau) = 2^{-K}.$$

In other words, the random variables $\widetilde{\tau}_1$ and $\widetilde{\tau}_2$ are independent of each other. Repeating this argument, we conclude that $\widetilde{\tau}_1, \ldots, \widetilde{\tau}_s$, obtained from

$$\begin{cases} \widetilde{G}[\emptyset](\widetilde{\tau}_1) = \mathbf{a}_1, \\ \widetilde{G}[\tau_1; \mathbf{a}_1](\widetilde{\tau}_2) = \mathbf{a}_2, \\ \qquad \vdots \\ \widetilde{G}[\tau_1, \ldots, \tau_{s-1}; \mathbf{a}_1, \ldots, \mathbf{a}_{s-1}](\widetilde{\tau}_s) = \mathbf{a}_s, \end{cases}$$

are independent random variables with common distribution function

$$\text{Prob}(\widetilde{\tau}_t = \delta) = 2^{-K}$$

for every $t = 1, \ldots, s$ and $\delta \in \{0, 1, \ldots, 2^K - 1\}$. Let

$$\widetilde{n}_0 = \widetilde{\tau}_s 2^{K(s-1)} + \widetilde{\tau}_{s-1} 2^{K(s-2)} + \ldots + \widetilde{\tau}_1.$$

Then $\widetilde{n}_0$ is uniformly distributed in the set $\{0, 1, \ldots, 2^{Ks} - 1\}$. Write

$$\widetilde{n} = \tau_h 2^{K(h-1)} + \ldots + \tau_{s+1} 2^{Ks} + \widetilde{n}_0,$$

where

$$H2^{Ks} = \tau_h 2^{K(h-1)} + \ldots + \tau_{s+1} 2^{Ks}.$$

Then

$$\widetilde{\mathbf{q}}(s;\mathbf{c};H) = \widetilde{\mathbf{q}}_{\widetilde{n}}.$$

Suppose now that $H2^{Ks} \le n < (H+1)2^{Ks}$. Then

$$\mathrm{Prob}(\widetilde{\mathbf{q}}(s;\mathbf{c};H) = \widetilde{\mathbf{q}}_n) = \mathrm{Prob}(\widetilde{n} = n) = 2^{-Ks}.$$

Since $\widetilde{\mathbf{q}}_n$ is uniformly distributed in $S(n)$ for every n satisfying $H2^{Ks} \le n < (H+1)2^{Ks}$, the result follows from the independence of $\widetilde{n}$ and $\widetilde{\mathbf{q}}_n$. ♣

Let $\mathcal{S}$ be a fixed compact and convex set in U^K. For integers s and H satisfying $0 \le s \le h$ and $0 \le H < 2^{K(h-s)}$, consider the random set

$$\widetilde{\mathcal{P}}(s,H) = \{\widetilde{\mathbf{q}}(s;\mathbf{c};H) : \mathbf{c} = (c_1,\ldots,c_K) \in \{0,1,\ldots,2^s-1\}^K\}, \tag{5.17}$$

and write

$$Z[\widetilde{\mathcal{P}}(s,H);\mathcal{S}] = \#(\widetilde{\mathcal{P}}(s,H) \cap \mathcal{S})$$

and

$$\widetilde{D}(s,H) = Z[\widetilde{\mathcal{P}}(s,H);\mathcal{S}] - 2^{Ks}\mu(\mathcal{S}). \tag{5.18}$$

Note that $\widetilde{D}(s,H)$ depends on $\mathcal{S}$. Let

$$T(s,H) = \{\mathbf{c} \in \{0,1,\ldots,2^s-1\}^K : C(s;\mathbf{c}) \cap \mathcal{S} \ne \emptyset \text{ and } C(s;\mathbf{c}) \setminus \mathcal{S} \ne \emptyset\}.$$

It is easy to see that

$$\#T(s,H) \le 2K2^{(K-1)s}. \tag{5.19}$$

Since every cube $C(s;\mathbf{c})$ contains exactly one element (namely $\widetilde{\mathbf{q}}(s;\mathbf{c};H)$) of the (random) set $\widetilde{\mathcal{P}}(s,H)$, we have

$$\widetilde{D}(s,H) = \sum_{\substack{\mathbf{c}\in T(s,H)\\ \widetilde{\mathbf{q}}(s;\mathbf{c};H)\in\mathcal{S}}} 1 - 2^{Ks} \sum_{\mathbf{c}\in T(s,H)} \mu(C(s;\mathbf{c}) \cap \mathcal{S}).$$

For every $\mathbf{c} \in T(s,H)$, let

$$\xi(s;\mathbf{c};H) = \begin{cases} 1 & (\widetilde{\mathbf{q}}(s;\mathbf{c};H) \in \mathcal{S}), \\ 0 & (\text{otherwise}). \end{cases} \tag{5.20}$$

By Lemma 5.4, we have

$$\mathbb{E}\xi(s;\mathbf{c};H) = \frac{\mu(C(s;\mathbf{c}) \cap \mathcal{S})}{\mu(C(s;\mathbf{c}))} = 2^{Ks}\mu(C(s;\mathbf{c}) \cap \mathcal{S}),$$

so that writing

$$\eta(s;\mathbf{c};H)=\xi(s;\mathbf{c};H)-\mathbb{E}\xi(s;\mathbf{c};H), \tag{5.21}$$

we have

$$\widetilde{D}(s,H)=\sum_{\mathbf{c}\in T(s,H)}\eta(s;\mathbf{c};H). \tag{5.22}$$

Note that $\mathbb{E}\eta=0$ and $|\eta|\le 1$.

We need the following result, but we omit its lengthy proof.

Lemma 5.5. *Suppose that $s',s''\in\{0,1,\ldots,h\}$. Suppose further that H' and H'' are integers satisfying $0\le H'<2^{K(h-s')}$ and $0\le H''<2^{K(h-s'')}$ and that $\mathbf{c}'\in\{0,1,\ldots,2^{s'}-1\}^K$ and $\mathbf{c}''\in\{0,1,\ldots,2^{s''}-1\}^K$. Suppose further that either*
i) $s'=s''$ and $\mathbf{c}'\neq\mathbf{c}''$; or
ii) $s'>s''$.
Then

$$\mathbb{E}(\eta(s';\mathbf{c}';H')\eta(s'';\mathbf{c}'';H''))\le\frac{\mu(C(s';\mathbf{c}')\cap C(s'';\mathbf{c}''))}{\mu(C(s'';\mathbf{c}''))}.$$

§5.5. Continuation of the Proof

For every natural number M satisfying $1\le M\le 2^{Kh}$, let

$$\widetilde{\mathcal{Q}}_M=\{\widetilde{\mathbf{q}}_0,\widetilde{\mathbf{q}}_1,\ldots,\widetilde{\mathbf{q}}_{M-1}\} \tag{5.23}$$

and, for every compact and convex set $\mathcal{S}\subseteq U^K$, let

$$Z[\widetilde{\mathcal{Q}}_M;\mathcal{S}]=\#(\widetilde{\mathcal{Q}}_M\cap\mathcal{S}),$$

and write

$$D[\widetilde{\mathcal{Q}}_M;\mathcal{S}]=Z[\widetilde{\mathcal{Q}}_M;\mathcal{S}]-M\mu(\mathcal{S}). \tag{5.24}$$

Lemma 5.6. *For every natural number M satisfying $1\le M\le 2^{Kh}$, we have*

$$\mathbb{E}\Big(D[\widetilde{\mathcal{Q}}_M;\mathcal{S}]\Big)^2\le K2^{4K}M^{1-1/K}.$$

Proof. Write

$$M-1=\tau_h 2^{K(h-1)}+\tau_{h-1}2^{K(h-2)}+\ldots+\tau_1,$$

where $\tau_1,\ldots,\tau_h\in\{0,1,\ldots,2^K-1\}$. Suppose that $\tau_{k+1}=\ldots=\tau_h=0$ and $\tau_k\neq 0$. Then

$$\widetilde{\mathcal{Q}}_M=\bigcup_{s=1}^{k}\bigcup_{m_s=0}^{\tau_s-1}\Big\{\widetilde{\mathbf{q}}_n:M_s+m_s2^{K(s-1)}\le n<M_s+(m_s+1)2^{K(s-1)}\Big\}, \tag{5.25}$$

where, for $1 \leq s \leq k$,

$$\begin{aligned} M_s &= \tau_h 2^{K(h-1)} + \tau_{h-1} 2^{K(h-2)} + \ldots + \tau_{s+1} 2^{Ks} \\ &= \tau_k 2^{K(k-1)} + \tau_{k-1} 2^{K(k-2)} + \ldots + \tau_{s+1} 2^{Ks}. \end{aligned} \tag{5.26}$$

Note that

$$M \geq 2^{K(k-1)}. \tag{5.27}$$

It now follows from (5.16), (5.17), (5.25) and (5.26) that

$$\widetilde{\mathcal{Q}}_M = \bigcup_{s=1}^{k} \bigcup_{m_s=0}^{\tau_s-1} \widetilde{\mathcal{P}}(s-1, H(s,m_s)), \tag{5.28}$$

where, for $1 \leq s \leq k$,

$$\begin{aligned} H(s,m_s) &= 2^{-K(s-1)} M_s + m_s \\ &= \tau_k 2^{K(k-s)} + \tau_{k-1} 2^{K(k-s-1)} + \ldots + \tau_{s+1} 2^{K} + m_s. \end{aligned}$$

Combining (5.18), (5.22), (5.24) and (5.28), we have

$$D[\widetilde{\mathcal{Q}}_M; \mathcal{S}] = \sum_{s=1}^{k} \sum_{m_s=0}^{\tau_s-1} \sum_{\mathbf{c} \in T(s-1,H(s,m_s))} \eta(s-1; \mathbf{c}; H(s,m_s)). \tag{5.29}$$

For $s = 1, \ldots, k$, let

$$X_s = \{\eta(s-1; \mathbf{c}; H(s,m_s)) : 0 \leq m_s < \tau_s \text{ and } \mathbf{c} \in T(s-1, H(s,m_s))\},$$

and let

$$X = \bigcup_{s=1}^{k} X_s.$$

Then by (5.29), we have

$$\mathbb{E}\Big(D[\widetilde{\mathcal{Q}}_M; \mathcal{S}]\Big)^2 = \sum_{\eta_1 \in X} \sum_{\eta_2 \in X} \mathbb{E}(\eta_1 \eta_2) = I_1 + 2I_2, \tag{5.30}$$

where

$$I_1 = \sum_{s=1}^{k} \sum_{\eta_1,\eta_2 \in X_s} \mathbb{E}(\eta_1\eta_2) \qquad \text{and} \qquad I_2 = \sum_{1 \leq s < t \leq k} \sum_{\eta_1 \in X_s} \sum_{\eta_2 \in X_t} \mathbb{E}(\eta_1\eta_2).$$

Consider first I_1. By (5.19) and Lemma 5.5, and noting that $|\eta| \leq 1$, we have

$$\begin{aligned} |I_1| &\leq \sum_{s=1}^{k} \sum_{m_s'=0}^{\tau_s-1} \sum_{m_s''=0}^{\tau_s-1} \#(T(s-1, H(s,m_s')) \cap T(s-1, H(s,m_s''))) \\ &\leq \sum_{s=1}^{k} \tau_s^2 2K 2^{(K-1)s} \leq K 2^{2K+2} 2^{(K-1)k} = K 2^{3K+1} 2^{(K-1)(k-1)} \\ &\leq K 2^{3K+1} M^{1-1/K}, \end{aligned} \tag{5.31}$$

in view of (5.27). We now consider I_2. Suppose that $1 \le s < t \le k$ and $\mathbf{c} \in T(t-1, H(t, m_t))$. Then there is at most one $\mathbf{c}' \in T(s-1, H(s, m_s))$ such that $C(s-1;\mathbf{c}') \cap C(t-1;\mathbf{c}) \neq \emptyset$. In fact, we then have $C(t-1;\mathbf{c}) \subseteq C(s-1;\mathbf{c}')$, and so

$$\frac{\mu(C(s-1;\mathbf{c}') \cap C(t-1;\mathbf{c}))}{\mu(C(s-1;\mathbf{c}'))} = 2^{K(s-t)}.$$

It follows from (5.19) and Lemma 5.5 that

$$\begin{aligned} |I_2| &\le \sum_{t=1}^{k} \sum_{\eta \in X_t} \sum_{s=1}^{t-1} \sum_{m_s=0}^{\tau_s - 1} 2^{K(s-t)} \le 2 \sum_{t=1}^{k} \sum_{\eta \in X_t} 1 = 2 \sum_{t=1}^{k} \sum_{m_t=0}^{\tau_t - 1} \#(T(t-1, H(t, m_t)) \\ &\le 4K \sum_{t=1}^{k} \sum_{m_t=0}^{\tau_t - 1} 2^{(K-1)(t-1)} \le K2^{K+2} \sum_{t=1}^{k} 2^{(K-1)(t-1)} \\ &\le K2^{K+3} 2^{(K-1)(k-1)} \le K2^{K+3} M^{1-1/K}, \end{aligned} \tag{5.32}$$

in view of (5.27). The lemma now follows on combining (5.30)–(5.32). ♣

Let A be a given compact and convex body in U^K. It now follows from Lemma 5.6 that for any real number $\lambda \in (0,1]$, any proper orthogonal transformation τ in $\mathbb{R}^K$ and any vector $\mathbf{u} \in U^K$, we have

$$\mathbb{E}\Big(D[\widetilde{\mathcal{Q}}_M; A(\lambda, \tau, \mathbf{u})]\Big)^2 \ll_K M^{1-1/K}$$

for every M satisfying $1 \le M \le 2^{Kh}$. If we now choose h to satisfy

$$2^{K(h-1)} < N \le 2^{Kh},$$

then

$$\mathbb{E}\left(\frac{1}{N} \sum_{M=1}^{N} \int_0^1 \int_{\mathcal{T}} \int_{U^K} \left|D[\widetilde{\mathcal{Q}}_M; A(\lambda, \tau, \mathbf{u})]\right|^2 \mathrm{d}\mathbf{u}\,\mathrm{d}\tau\,\mathrm{d}\lambda\right) \ll_K N^{1-1/K}.$$

(5.5) follows immediately. This proves Theorem 10 in the case $L = 1$. Note also the simpler inequality

$$\mathbb{E}\left(\int_0^1 \int_{\mathcal{T}} \int_{U^K} \left|D[\widetilde{\mathcal{Q}}_N; A(\lambda, \tau, \mathbf{u})]\right|^2 \mathrm{d}\mathbf{u}\,\mathrm{d}\tau\,\mathrm{d}\lambda\right) \ll_K N^{1-1/K}.$$

Theorem 6A follows.

§5.6. The General Case

We now discuss briefly how we may expand on the argument in §§5.3–5.5 to give a proof of Theorem 10.

We shall in fact construct an infinite sequence of points in U^{K+L} and use only the first N terms of this sequence. The main ingredient in the construction of this sequence is the Chinese remainder theorem. This not only makes it possible for the determination of the first K coordinates of the points of the sequence to be carried out independently of the determination of the last L coordinates of these points, but also enables us to treat the discrepancy arising from $A(\lambda, \tau, \mathbf{u})$ quite separately from the discrepancy arising from the $B(\mathbf{y})$. Furthermore, it ensures that important properties of the sequence are also present in many subsequences that arise from our argument.

Let h be a natural number, to be fixed subsequently in the argument. Let $p_1, \ldots, p_L$ denote the first L odd primes.

For every $p = 2, p_1, \ldots, p_L$, for every $s = 0, 1, \ldots, h$ and for every $c \in \mathbb{Z}$, let

$$I(p, s, c) = [cp^{-s}, (c+1)p^{-s}).$$

In other words, $I(p, s, c)$ is an interval of length p^{-s} and whose endpoints are consecutive integer multiples of p^{-s}.

We shall construct an infinite sequence $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \ldots$ of points in U^{K+L} such that the following is satisfied. For every $s_0, s_1, \ldots, s_L \in \{0, 1, \ldots, h\}$ and for every non–negative integer c, every set of the form

$$I(2, s_0, a_1) \times \ldots \times I(2, s_0, a_K) \times I(p_1, s_1, b_1) \times \ldots \times I(p_L, s_L, b_L)$$

in U^{K+L}, where $a_1, \ldots, a_K, b_1, \ldots, b_L \in \mathbb{Z}$, contains exactly one point of

$$\{\mathbf{p}_n : c2^{Ks_0} p_1^{s_1} \ldots p_L^{s_L} \leq n < (c+1)2^{Ks_0} p_1^{s_1} \ldots p_L^{s_L}\}.$$

As before, the construction of such a sequence involves ideas in combinatorics and poses no real difficulty. The first K coordinates of the points are constructed in a similar fashion as in the special case discussed earlier, although we also use periodicity to obtain an infinite sequence. The last L coordinates of the points are constructed as in §3.3, using the Halton–Hammersley sequence. As before, such a sequence alone is insufficient to give a proof of Theorem 10, and we appeal again to tools in probability theory. Note, however, that the situation here is much more complicated than the situation when $L = 1$. Indeed, we need to apply probabilistic arguments in two quite different ways. One of these, to deal with the discrepancy arising from $A(\lambda, \tau, \mathbf{u})$, is essentially similar to the probabilistic arguments in §5.4, with only minor modifications. However, to deal with the discrepancy arising from $B(\mathbf{y})$, we appeal to my discrete version of the probabilistic idea of Roth in §3.3. This discrete version was first developed in [**13**] to show that Faure sets, as discussed in §3.2, give an alternative proof of Theorem 2D.

Needless to say, our combinatorial construction has to be carried out in such a way that our probabilistic arguments can be implemented with ease.

§6. Davenport's Method Revisited

In this section, we use the ideas of Davenport [**16**] and Roth [**26**] described in §§2.1–2.2

to prove two upper bound theorems on certain L^1–norms concerning irregularities of distribution relative to half–planes and relative to convex polygons. In the two parts of this section, we shall employ slightly different notation.

§**6.1. Roth's Disc–Segment Problem**

Let U be a 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 $\mathcal{P}$ in B, and write

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

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

For every real number $r \in \mathbb{R}$ and every angle θ satisfying $0 \leq \theta \leq 2\pi$, let $S(r,\theta)$ denote the closed half–plane

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

Here $\mathbf{e}(\theta) = (\cos\theta, \sin\theta)$ and $\mathbf{x} \cdot \mathbf{y}$ denotes the scalar product of $\mathbf{x}$ and $\mathbf{y}$. For any θ satisfying $0 \leq \theta \leq 2\pi$, let

$$R(\theta) = \sup\{r \geq 0 : S(r,\theta) \cap U \neq \emptyset\}.$$

The following theorem is more general than Theorem 12B.

THEOREM 12C. (Beck–Chen [**10**]) *For every natural number $N \geq 2$, 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 is motivated by the special case when U is the square $[-1/2,1/2]^2$. We shall therefore first of all show that for every natural number N, there exists a set $\mathcal{P}$ of $4N^2+4N+1$ points in U such that

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

For ease of notation, we consider the following renormalized version of the problem. Let V be the square $[-N-1/2,N+1/2]^2$. 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 $\mathcal{P}$ in B, and write

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

We shall show that the set

$$\mathcal{P} = \{-N, -N+1, \dots, -1, 0, 1, \dots, N-1, N\}^2$$

of $4N^2 + 4N + 1$ integer lattice points in V satisfies

$$\int_0^{2\pi} \int_0^{M(\theta)} |E[\mathcal{P}; S(r,\theta)]| \mathrm{d}r \mathrm{d}\theta \ll N(\log N)^2, \tag{6.1}$$

where, for every $\theta \in [0, 2\pi]$, we have $M(\theta) = (2N+1)R(\theta)$.

The line

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

is the boundary of the half–plane $S(r,\theta)$, 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 $M(\theta) = (N+1/2)(\cos\theta + \sin\theta)$. We distinguish two cases.

Case 1: $0 \le r \le (N+1/2)(\cos\theta - \sin\theta)$. It is not difficult to see that $T(r,\theta)$ intersects the edges $\{(x_1, N+1/2) : |x_1| \le N+1/2\}$ and $\{(x_1, -N-1/2) : |x_1| \le N+1/2\}$ of V, *i.e.*, the "top" and "bottom" edges of V. Then

$$S(r,\theta) \cap V = \bigcup_{n=-N}^{N} S(n, V, r, \theta),$$

where, for every $n = -N, \dots, 0, \dots, N$,

$$S(n, V, r, \theta) = S(r,\theta) \cap V \cap (\mathbb{R} \times [n-1/2, n+1/2]).$$

Clearly

$$E[\mathcal{P}; S(r,\theta)] = \sum_{n=-N}^{N} E[\mathcal{P}; S(n, V, r, \theta)].$$

Now, for every $n = -N, \dots, 0, \dots, N$, we have

$$Z[\mathcal{P}; S(n, V, r, \theta)] = [N + n\tan\theta - r\sec\theta + 1]$$

and

$$\mu(S(n, V, r, \theta)) = N + n\tan\theta - r\sec\theta + 1/2,$$

so that

$$E[\mathcal{P}; S(n, V, r, \theta)] = -\psi(n\tan\theta - r\sec\theta),$$

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

$$E[\mathcal{P}; S(r,\theta)] = -\sum_{n=-N}^{N} \psi(n\tan\theta - r\sec\theta).$$

Case 2: $(N+1/2)(\cos\theta - \sin\theta) \le r \le (N+1/2)(\cos\theta + \sin\theta)$. It is not difficult to see that $T(r,\theta)$ intersects the edges $\{(x_1, N+1/2) : |x_1| \le N+1/2\}$ and $\{(N+1/2, x_2) : |x_2| \le N+1/2\}$ of V, *i.e.*, the "top" and "right" edges of V. Furthermore,

$$\begin{aligned} &T(r,\theta) \cap \{(N+1/2, x_2) : |x_2| \le N+1/2\} \\ &\quad = \{(N+1/2, -(N+1/2)\cot\theta + r\operatorname{cosec}\theta)\}. \end{aligned}$$

Then $S(n,V,r,\theta) = \emptyset$ if $n < -(N+1/2)\cot\theta + r\operatorname{cosec}\theta - 1/2$. On the other hand, it is trivial that $E[\mathcal{P}; S(n,V,r,\theta)] = O(1)$ always. It follows that

$$E[\mathcal{P}; S(r,\theta)] = -\sum_{\substack{n=-N \\ (*)}}^{N} \psi(n\tan\theta - r\sec\theta) + O(1),$$

where the summation is under the further restriction

$$n \ge -(N+1/2)\cot\theta + r\operatorname{cosec}\theta. \tag{$*$}$$

Note that in Case 1, the restriction $(*)$ would become superfluous since it is weaker than the requirement $n \ge -N$. It follows that for all $r \ge 0$, we have

$$E[\mathcal{P}; S(r,\theta)] - G[\mathcal{P}; r,\theta] \ll 1,$$

where

$$G[\mathcal{P}; r,\theta] = -\sum_{\substack{n=-N \\ (*)}}^{N} \psi(n\tan\theta - r\sec\theta).$$

The function $\psi(z) = z - [z] - 1/2$ has the Fourier expansion

$$-\sum_{\nu \ne 0} \frac{e(z\nu)}{2\pi \mathrm{i}\nu},$$

so that $-\psi(n\tan\theta - r\sec\theta)$ has the Fourier expansion

$$\sum_{\nu \ne 0} \frac{e(-r\nu\sec\theta)}{2\pi \mathrm{i}\nu} e(n\nu\tan\theta).$$

It follows that the Fourier expansion of $G[\mathcal{P};r,\theta]$ is given by

$$\sum_{\nu\neq 0}\frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-N\\(*)}}^{N}e(n\nu\tan\theta).$$

However, the restriction $(*)$ prevents us from applying Parseval's theorem.

To overcome this difficulty, we introduce the following idea which is motivated by Roth's variation of Davenport's method in §§2.1–2.2.

Let $\mathbf{y}=(y_1,y_2)\in[-1/2,1/2]^2$. For every $\theta\in[0,\pi/4]$ and every $r\geq 1$, let

$$T(\mathbf{y};r,\theta)=T(r+y_1\cos\theta+y_2\sin\theta,\theta) \tag{6.2}$$

and

$$S(\mathbf{y};r,\theta)=S(r+y_1\cos\theta+y_2\sin\theta,\theta) \tag{6.3}$$

(note 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

$$G[\mathcal{P};\mathbf{y};r,\theta]=-\sum_{\substack{n=-N\\(*)}}^{N}\psi(n\tan\theta-(r+y_1\cos\theta+y_2\sin\theta)\sec\theta),$$

then

$$E[\mathcal{P};S(\mathbf{y};r,\theta)]-G[\mathcal{P};\mathbf{y};r,\theta]\ll\begin{cases}\cot\theta & (M(\theta)-(2N+1)\sin\theta-1\leq r\leq M(\theta)),\\ 1 & (\text{otherwise}),\\ N & (\text{trivially}),\end{cases}$$

so that

$$\int_0^{\pi/4}\int_1^{M(\theta)}|E[\mathcal{P};S(\mathbf{y};r,\theta)]-G[\mathcal{P};\mathbf{y};r,\theta]|\mathrm{d}r\mathrm{d}\theta\ll N \tag{6.4}$$

(note that $|y_1\cos\theta+y_2\sin\theta|\leq 1$, so that if $r\leq M(\theta)-(2N+1)\sin\theta-1$, then $T(\mathbf{y};r,\theta)$ intersects the top and bottom edges of V).

Now $G[\mathcal{P};\mathbf{y};r,\theta]$ has the Fourier expansion

$$\sum_{\nu\neq 0}\frac{e(-(r+y_1\cos\theta+y_2\sin\theta)\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-N\\(*)}}^{N}e(n\nu\tan\theta)$$

$$=\sum_{\nu\neq 0}\frac{e(-r\nu\sec\theta)}{2\pi\mathrm{i}\nu}\sum_{\substack{n=-N\\(*)}}^{N}e((n-y_2)\nu\tan\theta)e(-y_1\nu).$$

It follows that for every $y_2 \in [-1/2, 1/2]$, we have, by Parseval's theorem, that

$$\int_{-1/2}^{1/2} |G[\mathcal{P};\mathbf{y};r,\theta]|^2 \mathrm{d}y_1 \ll \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{\substack{n=-N \\ (*)}}^{N} e((n-y_2)\nu\tan\theta) \right|^2$$

$$= \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{\substack{n=-N \\ (*)}}^{N} e(n\nu\tan\theta) \right|^2,$$

so that

$$\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} |G[\mathcal{P};\mathbf{y};r,\theta]|^2 \mathrm{d}y_1\mathrm{d}y_2 \ll \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \left| \sum_{\substack{n=-N \\ (*)}}^{N} e(n\nu\tan\theta) \right|^2$$

$$\ll \sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \min\{N^2, \|\nu\tan\theta\|^{-2}\}, \tag{6.5}$$

where $\|\beta\| = \min_{n\in\mathbb{Z}} |\beta - n|$ for every $\beta \in \mathbb{R}$.

We need the following crucial estimate.

Lemma 6.1. *We have*

$$\int_0^{\pi/4} \left(\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \min\{N^2, \|\nu\tan\theta\|^{-2}\} \right)^{1/2} \mathrm{d}\theta \ll (\log N)^2.$$

Proof. Since $\tan\theta \asymp \theta$ if $0 \le \theta \le \pi/4$, it suffices to show that

$$\int_0^{1} \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \min\{N^2, \|n\omega\|^{-2}\} \right)^{1/2} \mathrm{d}\omega \ll (\log N)^2. \tag{6.6}$$

Clearly

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \min\{N^2, \|n\omega\|^{-2}\} \le \sum_{n=1}^{N^2} \frac{1}{n^2} \min\{N^2, \|n\omega\|^{-2}\} + 1,$$

so that

$$\left(\sum_{n=1}^{\infty} \frac{1}{n^2} \min\{N^2, \|n\omega\|^{-2}\} \right)^{1/2} \le \sum_{n=1}^{N^2} \frac{1}{n} \min\{N, \|n\omega\|^{-1}\} + 1. \tag{6.7}$$

Now, for every $n = 1, \ldots, N^2$, we have

$$\int_0^1 \min\{N, \|n\omega\|^{-1}\} \mathrm{d}\omega = 2n \int_0^{1/2n} \min\{N, (n\omega)^{-1}\} \mathrm{d}\omega \ll \log N. \tag{6.8}$$

Inequality (6.6) now follows on combining (6.7) and (6.8). ♣

By the Cauchy–Schwarz inequality, we have

$$\begin{aligned} &\int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |G[\mathcal{P}; \mathbf{y}; r, \theta]| \mathrm{d}y_1 \mathrm{d}y_2 \\ &\qquad \ll \left(\int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |G[\mathcal{P}; \mathbf{y}; r, \theta]|^2 \mathrm{d}y_1 \mathrm{d}y_2 \right)^{1/2}. \end{aligned} \tag{6.9}$$

It follows from (6.4), (6.5), (6.9) and Lemma 6.1 that

$$\int_{-1/2}^{1/2} \int_{-1/2}^{1/2} \int_0^{\pi/4} \int_1^{M(\theta)} |E[\mathcal{P}; S(\mathbf{y}; r, \theta)]| \mathrm{d}r \mathrm{d}\theta \mathrm{d}y_1 \mathrm{d}y_2 \ll N (\log N)^2. \tag{6.10}$$

Note now that for every $\theta \in [0, \pi/4]$, every $r \geq 1$ and every $\mathbf{y} \in [-1/2, 1/2]^2$, we have, writing $s = r + y_1 \cos\theta + y_2 \sin\theta$, that $|r - s| < 1$. It follows that since $S(\mathbf{y}; r, \theta) = S(r + y_1 \cos\theta + y_2 \sin\theta, \theta)$, where $r + y_1 \cos\theta + y_2 \sin\theta \geq 0$, we must have

$$\int_2^{M(\theta)-1} |E[\mathcal{P}; S(r, \theta)]| \mathrm{d}r \leq \int_1^{M(\theta)} |E[\mathcal{P}; S(\mathbf{y}; r, \theta)]| \mathrm{d}r. \tag{6.11}$$

On the other hand,

$$\left(\int_0^2 + \int_{M(\theta)-1}^{M(\theta)} \right) |E[\mathcal{P}; S(r, \theta)]| \mathrm{d}r \ll N. \tag{6.12}$$

It now follows from (6.10)–(6.12) that

$$\int_0^{\pi/4} \int_0^{M(\theta)} |E[\mathcal{P}; S(r, \theta)]| \mathrm{d}r \mathrm{d}\theta \ll N (\log N)^2.$$

Similarly, for $j = 1, \ldots, 7$, we have

$$\int_{j\pi/4}^{(j+1)\pi/4} \int_0^{M(\theta)} |E[\mathcal{P}; S(r, \theta)]| \mathrm{d}r \mathrm{d}\theta \ll N (\log N)^2.$$

Inequality (6.1) now follows.

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 [**21**] 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

$$A(\mathbf{x}) = A(x_1, x_2) = [x_1 - 1/2, x_1 + 1/2] \times [x_2 - 1/2, x_2 + 1/2];$$

in other words, $A(\mathbf{x})$ is the aligned closed square of unit area and centred at $\mathbf{x}$. Let

$$\mathcal{P}_1 = \{\mathbf{p} \in \mathbb{Z}^2 : A(\mathbf{p}) \subseteq V\},$$

and write

$$V_1 = \bigcup_{\mathbf{p} \in \mathcal{P}_1} A(\mathbf{p}).$$

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

$$\mu(V_2) \in \mathbb{N} \qquad \text{and} \qquad \mu(V_2) \ll M.$$

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

$$\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2, \tag{6.13}$$

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 $\mathcal{P}$ in B, and write

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

It can be shown that the set (6.13) satisfies

$$\int_0^{2\pi} \int_0^M |E[\mathcal{P}; S(r,\theta)]| \mathrm{d}r \mathrm{d}\theta \ll M(\log N)^2. \tag{6.14}$$

The inequality (6.14) can be proved using explicitly the equation of ∂V, the boundary of V. However, if we want to prove the full generality of Theorem 12C, then such information is clearly not available. Extra geometric consideration is then required.

§6.2. Convex Polygons

Suppose that $\mathcal{P}$ is a distribution of N points in the unit square $U = [0,1]^2$, treated as a torus. For every measurable set B in U, let $Z[\mathcal{P}; B]$ denote the number of points of $\mathcal{P}$ in B, and write

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

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

Let $A \subseteq U$ be a closed convex polygon of diameter not exceeding 1 and centred 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

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{6.15}$$

where $\mathbf{v} = (v_1, v_2)$ and $\mathbf{u} = (u_1, u_2)$, and write

$$A(r,\theta) = \{r\mathbf{v} : \mathbf{v} = \theta(\mathbf{u}) \text{ for some } \mathbf{u} \in A\}; \tag{6.16}$$

in other words, $A(r,\theta)$ is obtained from A by first rotating clockwise by angle θ and then contracting by factor r about the origin $\mathbf{0}$. For every $\mathbf{x} \in U$, let

$$A(\mathbf{x}, r, \theta) = \{\mathbf{x} + \mathbf{v} : \mathbf{v} \in A(r,\theta)\}, \tag{6.17}$$

so that $A(\mathbf{x}, r, \theta)$ is a similar copy of A, with centre of gravity at $\mathbf{x}$.

THEOREM 14B. *For every natural number $N \ge 2$, there exists a distribution $\mathcal{P}$ of N points in U such that*

$$\int_0^1 \int_0^{2\pi} \int_U |D[\mathcal{P}; A(\mathbf{x}, r, \theta)]| \mathrm{d}\mathbf{x} \mathrm{d}\theta \mathrm{d}r \ll_A (\log N)^2.$$

Our proof is motivated by our study of irregularities of point distribution relative to half–planes in §6.1. In fact, the analogy between the two problems becomes clear on noting that a convex polygon is the intersection of a finite number of half–planes (or, to put it in precisely the viewpoint held at the time, that a half–plane is a convex "monogon").

We shall only briefly discuss the problem when N is a perfect square. For ease of notation, we consider the following renormalized version of the problem. Let V be the square $[0, N^{1/2}]^2$, again treated as a torus (modulo $N^{1/2}$ 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 $\mathcal{P}$ in B, and write

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

Let $A \subseteq V$ be a closed convex polygon of diameter not exceeding $N^{1/2}$ and centred 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 (6.15)–(6.17). 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

$$\int_0^1 \int_0^{2\pi} \int_V |E[\mathcal{P}; A(\mathbf{x}, r, \theta)]| \mathrm{d}\mathbf{x} \mathrm{d}\theta \mathrm{d}r \ll_A N(\log N)^2. \tag{6.18}$$

The key idea in the proof of (6.18) is to split the integral over V in (6.18) into a sum of integrals over sets whose diameters are very small. We may then use the variable $\mathbf{x}$ in the same way as the probabilistic variable $\mathbf{y}$ in §6.1.

Suppose that $N = M^2$, where $M \in \mathbb{N}$. It can be shown that the set

$$\mathcal{P} = \{(m - 1/2, n - 1/2) : m, n \in \mathbb{N} \text{ and } 1 \le m, n \le M\}$$

of N points in V satisfies the inequality (6.18).

The idea is roughly as follows. Let $A \subseteq V$ be a closed convex polygon of k sides and of diameter not exceeding M. Consider the set $A(\mathbf{x}, r, \theta)$, 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 half–plane $S(r, \theta)$ in §6.1, 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 what is the analogue of the probabilistic variable $\mathbf{y}$. The answer to this is that the translation variable $\mathbf{x}$, handled with great care, plays a similar role. The result now follows modulo technical refinements.

Appendix. Roth's Classical Theorem

The method of Roth [**23**] to prove Theorem 1A is dependent on Schwarz's inequality. Corresponding to every distribution $\mathcal{P}$ of N points in U_0^K, Roth constructed an auxiliary function $F[\mathcal{P};\mathbf{x}]$ such that, writing $D(\mathbf{x})$ and $F(\mathbf{x})$ in place of $D[\mathcal{P};B(\mathbf{x})]$ and $F[\mathcal{P};\mathbf{x}]$ respectively, and abusing notation and writing U instead of U_0,

$$\int_{U^K} F(\mathbf{x})D(\mathbf{x})\mathrm{d}\mathbf{x} > c_1(K)(\log N)^{K-1} \tag{$A.1$}$$

and

$$\int_{U^K} F^2(\mathbf{x})\mathrm{d}\mathbf{x} < c_2(K)(\log N)^{K-1}. \tag{$A.2$}$$

These, together with Schwarz's inequality, give

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

so that Theorem 1A follows easily.

We remark here that to prove Theorem 1D, Schmidt [**37**] proved the analogue of $(A.2)$ for higher moments, and then used Hölder's inequality instead of Schwarz's inequality. However, Schmidt's auxiliary function is slightly different from Roth's original auxiliary function. Here, we shall use Schmidt's auxiliary function.

Any $x \in U_0$ can be written in the form

$$x = \sum_{j=0}^{\infty} \beta_j(x)2^{-j-1},$$

where $\beta_j(x) = 0$ or 1 and such that the sequence $\beta_j(x)$ does not end with $1, 1, \ldots$. For $r = 0, 1, 2, \ldots$, let

$$R_r(x) = (-1)^{\beta_r(x)}$$

(these are called the Rademacher functions).

Definition. By an r–interval, we mean an interval of the form $[m2^{-r}, (m+1)2^{-r})$, where the integer m satisfies $0 \le m < 2^r$.

Definition. By an r–function, we mean a function $f(x)$ defined on U_0 such that in every r–interval, $f(x) = R_r(x)$ or $f(x) = -R_r(x)$.

Clearly, if $f(x)$ is an r–function, then

$$\int_U f(x)\mathrm{d}x = 0.$$

Suppose now that $\mathbf{r} = (r_1, \ldots, r_K)$ is a K–tuple of non–negative integers. Let

$$|\mathbf{r}| = r_1 + \ldots + r_K;$$

and for any $\mathbf{x} = (x_1, \ldots, x_K) \in U_0^K$, let

$$R_{\mathbf{r}}(\mathbf{x}) = R_{r_1}(x_1) \ldots R_{r_K}(x_K).$$

Definition. By an $\mathbf{r}$–box, we mean a set of the form $I_1 \times \ldots \times I_K$, where, for every $j = 1, \ldots, K$, I_j is an r_j–interval.

Definition. By an $\mathbf{r}$–function, we mean a function $f(\mathbf{x})$ defined on U_0^K such that in every $\mathbf{r}$–box, $f(\mathbf{x}) = R_{\mathbf{r}}(\mathbf{x})$ or $f(\mathbf{x}) = -R_{\mathbf{r}}(\mathbf{x})$.

Let $n \gg\ll \log N$ be a suitably chosen natural number.

Lemma A.1. *Suppose that* $|\mathbf{r}| = |\mathbf{s}| = n$. *Then*

$$\int_{U^K} R_{\mathbf{r}}(\mathbf{x}) R_{\mathbf{s}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \begin{cases} 1 & (\mathbf{r} = \mathbf{s}), \\ 0 & (\mathbf{r} \neq \mathbf{s}). \end{cases}$$

Proof. The result is clear if $\mathbf{r} = \mathbf{s}$. If $\mathbf{r} \neq \mathbf{s}$, then there exists $j = 1, \ldots, K$ such that $r_j \neq s_j$. Assume, without loss of generality, that $r_j > s_j$. Then $R_{r_j}(x_j) R_{s_j}(x_j)$ is an r_j–function, so that

$$\int_U R_{r_j}(x_j) R_{s_j}(x_j) \mathrm{d}x_j = 0. \qquad \clubsuit$$

We consider the function

$$F(\mathbf{x}) = \sum_{|\mathbf{r}|=n} f_{\mathbf{r}}(\mathbf{x}), \qquad (A.3)$$

where, for each $\mathbf{r}$, $f_{\mathbf{r}}(\mathbf{x})$ is a suitably chosen $\mathbf{r}$–function. To establish $(A.2)$, we have

Lemma A.2. *Suppose that* $K \geq 2$ *and* $n \geq 0$. *Suppose further that for every* $\mathbf{r}$ *with* $|\mathbf{r}| = n$, $f_{\mathbf{r}}(\mathbf{x})$ *is an* $\mathbf{r}$*–function. Then the function* $(A.3)$ *satisfies*

$$\int_{U^K} F^2(\mathbf{x}) \mathrm{d}\mathbf{x} \ll_K n^{K-1}.$$

Proof. Clearly

$$\int_{U^K} F^2(\mathbf{x}) \mathrm{d}\mathbf{x} = \sum_{|\mathbf{r}|=n} \int_{U^K} f_{\mathbf{r}}^2(\mathbf{x}) \mathrm{d}\mathbf{x} + \sum_{\substack{|\mathbf{r}|=n \\ |\mathbf{s}|=n \\ \mathbf{r} \neq \mathbf{s}}} \int_{U^K} f_{\mathbf{r}}(\mathbf{x}) f_{\mathbf{s}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \Sigma_1 + \Sigma_2,$$

say. In view of Lemma A.1, $\Sigma_2 = 0$. On the other hand, $f_{\mathbf{r}}^2(\mathbf{x}) = 1$ for every $\mathbf{x} \in U_0^K$. It follows that

$$\Sigma_1 \leq \sum_{|\mathbf{r}|=n} 1 = \binom{n+K-1}{K-1} \ll_K n^{K-1}. \qquad \clubsuit$$

It remains to establish $(A.1)$. Let n be chosen to satisfy

$$2N \leq 2^n < 4N.$$

Lemma A.3. *Suppose that $2^n \geq 2N$. Then for every $\mathbf{r}$ satisfying $|\mathbf{r}| = n$, there is an $\mathbf{r}$–function $f_{\mathbf{r}}$ satisfying*

$$\int_{U^K} f_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x} \geq 2^{-n-2K-1} N. \qquad (A.4)$$

$(A.1)$ now follows, for we can construct $F(\mathbf{x})$ by $(A.3)$, where for every $|\mathbf{r}| = n$, $f_{\mathbf{r}}$ is chosen to satisfy $(A.4)$. Now the number of K–tuples satisfying $|\mathbf{r}| = n$ is

$$\binom{n+K-1}{K-1} \gg_K n^{K-1},$$

so that

$$\int_{U^K} F(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x} \gg_K n^{K-1} 2^{-n} N.$$

Proof of Lemma A.3. We decompose the integral in $(A.4)$ into integrals over $\mathbf{r}$–boxes, and choose $f_{\mathbf{r}}(\mathbf{x})$ such that the integral $\int f_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x}$ over every $\mathbf{r}$–box is non–negative. Let B be an $\mathbf{r}$–box given by

$$B = [m_1 2^{-r_1}, (m_1+1) 2^{-r_1}) \times \ldots \times [m_K 2^{-r_K}, (m_K+1) 2^{-r_K}),$$

and let B' be the box

$$B' = [m_1 2^{-r_1}, (m_1+1/2) 2^{-r_1}) \times \ldots \times [m_K 2^{-r_K}, (m_K+1/2) 2^{-r_K}).$$

Then it is not difficult to see that

$$\int_B R_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x}$$
$$= \int_{B'} \sum_{\alpha_1=0}^{1} \ldots \sum_{\alpha_K=0}^{1} (-1)^{\alpha_1+\ldots+\alpha_K} D((y_1 + \alpha_1 2^{-r_1-1}, \ldots, y_K + \alpha_K 2^{-r_K-1})) \mathrm{d}\mathbf{y}.$$

Note that the sum

$$\left| \sum_{\alpha_1=0}^{1} \ldots \sum_{\alpha_K=0}^{1} (-1)^{\alpha_1+\ldots+\alpha_K} Z[\mathcal{P}; B((y_1+\alpha_1 2^{-r_1-1}, \ldots, y_K+\alpha_K 2^{-r_K-1}))] \right|$$

is the number of points of $\mathcal{P}$ in $[y_1, y_1+2^{-r_1-1}) \times \ldots \times [y_K, y_K+2^{-r_K-1})$. This box is contained in B. Hence if B contains no points of $\mathcal{P}$, the sum is 0. Note also that

$$\sum_{\alpha_1=0}^{1} \ldots \sum_{\alpha_K=0}^{1} (-1)^{\alpha_1+\ldots+\alpha_K} (y_1+\alpha_1 2^{-r_1-1}) \ldots (y_K+\alpha_K 2^{-r_K-1}) = (-1)^K 2^{-|\mathbf{r}|-K}.$$

It follows from the definition of $D(\mathbf{x})$ that if B contains no points of $\mathcal{P}$, then since $|\mathbf{r}| = n$, we have

$$\int_B R_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x} = (-1)^{K+1} 2^{-2n-2K} N.$$

There are 2^n $\mathbf{r}$–boxes B with $|\mathbf{r}| = n$, but only $N \leq 2^{n-1}$ points. It follows that at least half of the $\mathbf{r}$–boxes contain no points of $\mathcal{P}$. Since $f_{\mathbf{r}}(\mathbf{x})$ is chosen to make the integral $\int f_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x}$ over any $\mathbf{r}$–box non–negative, it follows that

$$\int_{U^K} f_{\mathbf{r}}(\mathbf{x}) D(\mathbf{x}) \mathrm{d}\mathbf{x} \geq (2^n - N) 2^{-2n-2K} N \geq 2^{-n-2K-1} N. \qquad \clubsuit$$

References

1. J.R. Alexander, "Geometric methods in the study of irregularities of distribution", *Combinatorica*, **10** (1990), 115–136.

2. J. Beck, "Balanced two–colourings of finite sets in the square I", *Combinatorica*, **1** (1981), 327–335.

3. J. Beck, "On a problem of K.F. Roth concerning irregularities of point distribution", *Invent. Math.*, **74** (1983), 477–487.

4. J. Beck, "Irregularities of point distribution I", *Acta Math.*, **159** (1987), 1–49.

5. J. Beck, "Irregularities of point distribution II", *Proc. London Math. Soc. (3)*, **56** (1988), 1–50.

6. J. Beck, "A two–dimensional van Aardenne–Ehrenfest theorem in irregularities of distribution", *Compositio Math.*, **72** (1989), 269–339.

7. J. Beck and W.W.L. Chen, *Irregularities of Distribution* (Cambridge Tracts in Mathematics **89**, Cambridge University Press, Cambridge, 1987).

8. J. Beck and W.W.L. Chen, "Irregularities of point distribution relative to convex polygons", *Irregularities of Partitions*, pp. 1–22 (Algorithms and Combinatorics **8**, Springer Verlag, Berlin, 1989).

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

10. J. Beck and W.W.L. Chen, "Irregularities of point distribution relative to half–planes I", *Mathematika* (to appear).

11. J. Beck and W.W.L. Chen, "Irregularities of point distribution relative to convex polygons II", *Mathematika* (to appear).

12. W.W.L. Chen, "On irregularities of distribution", *Mathematika*, **27** (1980), 153–170.

13. W.W.L. Chen, "On irregularities of distribution II", *Quart. J. Math. Oxford (2)*, **34** (1983), 257–279.

14. J.G. van der Corput, "Verteilungsfunktionen I", *Proc. Kon. Ned. Akad. v. Wetensch.*, **38** (1935), 813–821.

15. J.G. van der Corput, "Verteilungsfunktionen II", *Proc. Kon. Ned. Akad. v. Wetensch.*, **38** (1935), 1058–1066.

16. H. Davenport, "Note on irregularities of distribution", *Mathematika*, **3** (1956), 131–135.

17. H. Faure, "Discrépance de suites associées à un système de numération (en dimension s)", *Acta Arith.*, **41** (1982), 337–351.

18. G. Halász, "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, 1981).

19. J.H. Halton, "On the efficiency of certain quasirandom sequences of points in evaluating multidimensional integrals", *Num. Math.*, **2** (1960), 84–90.

20. J.M. Hammersley, "Monte Carlo methods for solving multivariable problems", *Ann. New York Acad. Sci.*, **86** (1960), 844–874.

21. G.H. Hardy, "On the expression of a number as the sum of two squares", *Quart. J. Math.*, **46** (1915), 263–283.

22. M. Lerch, "Question 1547", *L'Intermediaire Math.*, **11** (1904), 144–145.

23. K.F. Roth, "On irregularities of distribution", *Mathematika*, **1** (1954), 73–79.

24. K.F. Roth, "Remark concerning integer sequences", *Acta Arith.*, **9** (1964), 257–260.

25. K.F. Roth, "On irregularities of distribution II", *Communications on Pure and Applied Math.*, **29** (1976), 749–754.

26. K.F. Roth, “On irregularities of distribution III”, *Acta Arith.*, **35** (1979), 373–384.

27. K.F. Roth, “On irregularities of distribution IV”, *Acta Arith.*, **37** (1980), 67–75.

28. W.M. Schmidt, “Irregularities of distribution”, *Quart. J. Math. Oxford (2)*, **19** (1968), 181–191.

29. W.M. Schmidt, “Irregularities of distribution II”, *Tran. American Math. Soc.*, **136** (1969), 347–360.

30. W.M. Schmidt, “Irregularities of distribution III”, *Pacific J. Math.*, **29** (1968), 225–234.

31. W.M. Schmidt, “Irregularities of distribution IV”, *Invent. Math.*, **7** (1968), 55–82.

32. W.M. Schmidt, “Irregularities of distribution V”, *Proc. American Math. Soc.*, **25** (1970), 608–614.

33. W.M. Schmidt, “Irregularities of distribution VI”, *Compositio Math.*, **24** (1972), 63–74.

34. W.M. Schmidt, “Irregularities of distribution VII”, *Acta Arith.*, **21** (1972), 45–50.

35. W.M. Schmidt, “Irregularities of distribution VIII”, *Tran. American Math. Soc.*, **198** (1974), 1–22.

36. W.M. Schmidt, “Irregularities of distribution IX”, *Acta Arith.*, **27** (1975), 385–396.

37. W.M. Schmidt, “Irregularities of distribution X”, *Number Theory and Algebra*, pp. 311–329 (Academic Press, New York, 1977).

38. W.M. Schmidt, *Lectures on Irregularities of Distribution* (Tata Institute of Fundamental Research, Bombay, 1977).

School of MPCE
Macquarie University
Sydney NSW 2109