THERE IS NO INTEGER BETWEEN 0 AND 1 – A VERY BRIEF INTRODUCTION TO GEOMETRIC DISCREPANCY THEORY

WILLIAM CHEN

Y.C. Wong Lecture, University of Hong Kong, 18 October 2012

Dedicated to the memory of my mother
Dr Doris Chen (30 July 1929 – 3 June 2012),
a member of the Department of Mathematics
at the University of Hong Kong from 1953 to 1985,
and who gifted me the pleasure of mathematics

It is a great privilege for me to have the opportunity to hold the prestigious Y.C. Wong Visiting Lectureship at the University of Hong Kong this month, and to deliver a Y.C. Wong Lecture.

In the late 1940s, when my mother was an undergraduate in mathematics at the Sun Yat-Sen University in Guangzhou (then Canton), Professor Wong was already a member of the Department of Mathematics at the University of Hong Kong. Those were days of much turmoil, and my mother was asked by her professor to bring a collection of important and valuable university documents to Hong Kong to give to Professor Wong for safe keeping. From this, I deduce that my association with Professor Wong goes back to when I was aged –7 or –8.

1. Introduction

In this lecture, we give a very brief introduction to geometric discrepancy theory. We highlight some of the major contributions to the subject and the powerful ideas involved. After setting up the questions, we explain the concept of *trivial errors*, a consequence of the fact that there is no integer between 0 and 1, and then give a simple example to show how such trivial errors can be exploited. We then briefly illustrate some of the techniques of the subject and the relationship with other areas of mathematics. We also discuss some applications.

Geometric discrepancy theory owes its existence to the fundamental work of the legendary mathematician Klaus Roth. He was awarded the Fields Medal in 1958 for his groundbreaking work on diophantine approximation and on arithmetic progressions, work that has greatly influenced such luminaries as Wolfgang Schmidt, Endre Szemerédi, Timothy Gowers, Ben Green and Terence Tao, to name just a few. Yet he considers his 1954 paper [31] on discrepancy theory to be his best work. As he explains, *"But I started a subject!"*

Discrepancy theory owes its existence to the original conjecture of Johannes van der Corput [18, 19] in 1935, that no sequence of real numbers between 0 and 1 can, in some sense, be too evenly distributed. However, the conjecture, as well as its proofs by Tatyana van Aardenne-Ehrenfest [1, 2] in 1945 and 1949, are in 1-dimensional setting, and therefore not geometric in nature. Indeed, Roth's paper [31] in 1954 is the first instance of such problems posed with a geometric framework.

Roth's early work is followed soon by a beautiful paper of Harold Davenport and then a remarkable series of ten papers by Wolfgang Schmidt between 1968 and

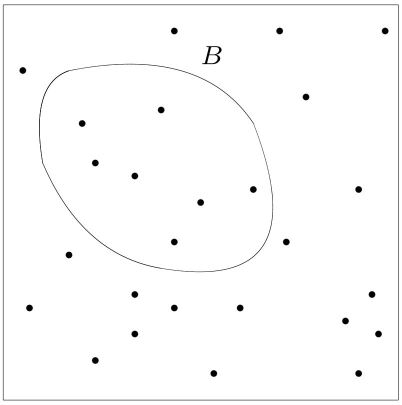
1977 and a monograph [37] in 1977, as well as three further papers by Roth himself. Gábor Halász also contributed a vital paper in 1981. These contributions form the solid foundation from which further work follows.

Much of the development of the subject in the 1980s has been carried out by József Beck and the author, separately as well as in collaboration, resulting in many breakthroughs as well as a monograph [7]. They are joined in the 1990s by Ralph Alexander, Jiří Matoušek and Maxim Skriganov. For a beginner in the subject, the beautifully written monograph by Matoušek [30] is highly recommended.

This century sees many more contributors to the subject, notably Dmitriy Bilyk, Michael Lacey and Armen Vagharshakyan. There are also many colleagues who take a keen interest in applications of discrepancy theory to problems on numerical integration.

2. Formulation of the Problems

Let N be a fixed positive integer. Suppose that $\mathcal{P}$ is a collection of N points in the unit square $[0,1]^2$.



For every measurable subset B of $[0,1]^2$, let $Z[\mathcal{P};B]$ denote the number of points of $\mathcal{P}$ that fall into B. If B is to have its fair share of points, then the quantity $N\mu(B)$, where $\mu(B)$ denotes the area of B, represents the corresponding expectation. Then the difference $D[\mathcal{P};B]=Z[\mathcal{P};B]-N\mu(B)$ is called the *discrepancy* of the collection $\mathcal{P}$ with respect to the subset B.

For every collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we now consider the discrepancy $D[\mathcal{P};B]$ over all subsets B of $[0,1]^2$ in an infinite collection $\mathcal{B}$, and not just for one subset B of $[0,1]^2$.

Lower bound results in discrepancy theory say that all collections $\mathcal{P}$ are *bad*, whereas upper bound results say that some collections $\mathcal{P}$ are *not too bad*.

For instance, a lower bound result may be of the form

for every $\mathcal{P}$, there exists B in $\mathcal{B}$ such that $|D[\mathcal{P};B]|$ is *large*.

Then a corresponding upper bound result may be of the form

there exists $\mathcal{P}$ such that for every B in $\mathcal{B}$, $|D[\mathcal{P};B]|$ is *not too large*.

Of course, we need to quantify *large* and *not too large*.

Alternatively, a lower bound result may be of the form

for every $\mathcal{P}$, $|D[\mathcal{P};B]|$ is *large on average* in $\mathcal{B}$.

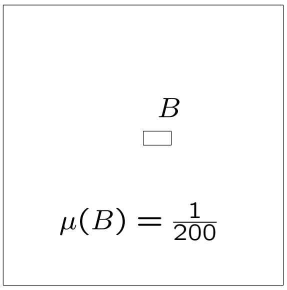
Then a corresponding upper bound result may be of the form

there exists $\mathcal{P}$ such that $|D[\mathcal{P};B]|$ is *not too large on average* in $\mathcal{B}$.

In this case, we have to further quantify *on average*.

3. Trivial Errors and Blowing Them Up

Suppose that $\mathcal{P}$ is a collection of $N = 100$ points in the unit square $[0,1]^2$.



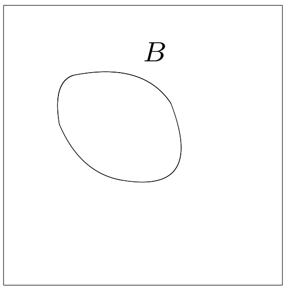
If B is a subset of $[0,1]^2$ of area $1/200$, then $N\mu(B) = 1/2$. Since there is no integer between 0 and 1, we must have $Z[\mathcal{P};B] = 0$ or $Z[\mathcal{P};B] \geqslant 1$. Then

$$D[\mathcal{P};B] = Z[\mathcal{P};B] - N\mu(B) \begin{cases} = 0 - \frac{1}{2}, & \text{if } Z[\mathcal{P};B] = 0, \\ \geqslant 1 - \frac{1}{2}, & \text{if } Z[\mathcal{P};B] \geqslant 1, \end{cases}$$

so that $|D[\mathcal{P};B]| \geqslant 1/2$. These are known as *trivial errors*, and we need to find ways to *blow them up*.

We illustrate how one may blow up the trivial errors by discussing the following result of Schmidt [35]. The result is essentially best possible, in the sense that the exponent $1/3$ cannot be improved.

Theorem S. *Suppose that $\mathcal{P}$ is a collection of N points in the unit square* $[0,1]^2$.

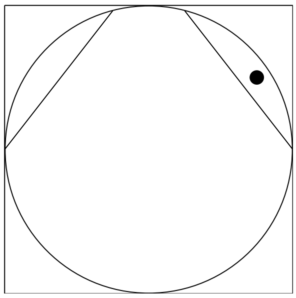


Then there exists a convex set B such that $|D[\mathcal{P};B]| \geqslant cN^{1/3}$.

Remark. This result is sometimes known as Schmidt's chocolate theorem. Here is an extract from an email received by the author a few years ago.
Dear William,
Recently I came upon some old writing of yours about me and chocolate. Actually my son had found it someplace on the internet and forwarded it to me. It is the note which contains two lemmas.
Lemma 1. Wolfgang Schmidt loves chocolate.
Lemma 2. Pat Schmidt makes lovely chocolate cake.
I am very touched by your kind comments. Am I forgetful or what, but I don't remember hearing you talk about this at a conference or reading it before. My son talked about your writing to my grandson (8 years) who then wrote about it in a school project, saying he liked me because I like chocolate and I am funny. Unfortunately I now have to eat less chocolate. I had kidney stones and nutritionists (they are bad people) say I should avoid chocolate and some other food to prevent kidney stones from recurring . . .
Best wishes, Wolfgang.

So let us prove Schmidt's chocolate theorem.

Proof. Consider the unit square $[0,1]^2$ (a square plate). It is possible to place a disc A of diameter 1 (cake) within $[0,1]^2$. Let $\mathcal{P}$ be any collection of N points (chocolates) in $[0,1]^2$. Consider disc segments of area $1/2N$ (tiny pieces of cake).



Elementary calculations show that there are at least $cN^{1/3}$ non-overlapping disc segments of area $1/2N$, where c is a positive constant. Some disc segments contain no points of $\mathcal{P}$, and we denote them by $S_1,\ldots,S_k$. The other disc segments each contains at least one point of $\mathcal{P}$, and we denote them by $T_1,\ldots,T_m$. Clearly we have $k+m\geqslant cN^{1/3}$. Furthermore

$$D[\mathcal{P};S_i]=-\tfrac{1}{2},\quad i=1,\ldots,k, \tag{1}$$

and

$$D[\mathcal{P};T_j]\geqslant\tfrac{1}{2},\quad j=1,\ldots,m, \tag{2}$$

Suppose first that we remove from A all those disc segments that contain no points of $\mathcal{P}$ (eat all those tiny pieces of cake that do not contain chocolates). The remainder is a convex set $A\setminus(S_1\cup\ldots\cup S_k)$, and

$$D[\mathcal{P};A\setminus(S_1\cup\ldots\cup S_k)]=D[\mathcal{P};A]-\sum_{i=1}^{k}D[\mathcal{P};S_i]. \tag{3}$$

Suppose instead that we remove from A all those disc segments that contain points of $\mathcal{P}$ (eat all those tiny pieces of cake that contain chocolates). The remainder is a convex set $A\setminus(T_1\cup\ldots\cup T_m)$, and

$$D[\mathcal{P};A\setminus(T_1\cup\ldots\cup T_m)]=D[\mathcal{P};A]-\sum_{j=1}^{m}D[\mathcal{P};T_j]. \tag{4}$$

Combining (3) and (4), and noting (1) and (2), we have

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

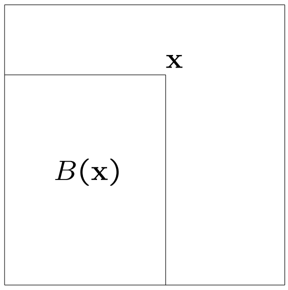
and so

$$|D[\mathcal{P};A\setminus(S_1\cup\ldots\cup S_k)]|\geqslant\tfrac{1}{4}cN^{1/3}\quad\text{or}\quad|D[\mathcal{P};A\setminus(T_1\cup\ldots\cup T_m)]|\geqslant\tfrac{1}{4}cN^{1/3},$$

or both. $\square$

4. The Classical Problem

The classical problem in discrepancy theory concerns the discrepancy of point collections with respect to aligned rectangles anchored at the origin. For every $\mathbf{x} = (x_1, x_2)$ in $[0,1]^2$, we denote by $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$ the rectangle with sides parallel to the coordinate axes, bottom left vertex anchored at the origin and top right vertex at $\mathbf{x}$.



Suppose that $\mathcal{P}$ is a collection of N points in unit square $[0,1]^2$. Then the discrepancy function is given by

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

The following is the pioneering result of Roth [31] in 1954.

Theorem R. *We have the following lower bound results.*

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

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

(ii) *There exists a constant $c_2 > 0$ such that for every collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\sup_{\mathbf{x} \in [0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \geqslant c_2 (\log N)^{1/2}. \tag{5}$$

Remark. For those readers not familiar with the notion of the supremum

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

think of this as the maximum value of $|D[\mathcal{P}; B(\mathbf{x})]|$ as $\mathbf{x}$ varies over $[0,1]^2$. This is not quite correct, but will suffice for the moment.

Note that part (ii) follows immediately from part (i). We comment that the estimate is not best possible. In fact, we can say a lot more.

Theorem RSHLDC. *We have the following lower bound results.*

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

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

(ii) *There exists a constant $c_2 > 0$ such that for every collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\sup_{\mathbf{x} \in [0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \geqslant c_2 \log N.$$

(iii) *For every fixed $q > 1$, there exists a constant $c_3(q) > 0$, depending at most on q, such that for every collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^q \, \mathrm{d}\mathbf{x} \geqslant c_3(q)(\log N)^{q/2}.$$

(iv) *There exists a constant $c_4 > 0$ such that for every collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$, we have*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \, \mathrm{d}\mathbf{x} \geqslant c_4(\log N)^{1/2}.$$

These are complemented by the following upper bound results.

(v) *There exists a constant $c_5 > 0$ such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\sup_{\mathbf{x} \in [0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]| \leqslant c_5 \log N.$$

(vi) *There exists a constant $c_6 > 0$ such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \leqslant c_6 \log N.$$

(vii) *For every fixed $q \geqslant 1$, there exists a constant $c_7(q) > 0$, depending at most on q, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit square $[0,1]^2$ such that*

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^q \, \mathrm{d}\mathbf{x} \leqslant c_7(q)(\log N)^{q/2}.$$

Here part (i) is the original result of Roth [31], while part (ii) is an improvement of Roth's bound (5) by Schmidt [34] in 1972. Part (iii), due to Schmidt [36] in 1977, is a generalization of part (i). Part (iv), due to Halász [23] in 1981, is stronger than both parts (i) and (iii).

Part (v), due to Lerch [29] in 1904, shows that part (ii) is best possible. Part (vi), due to Davenport [20] in 1956, shows that parts (i) and (iv) are best possible. Finally, part (vii), due to Chen [11] in 1980, is a generalization of part (vi) and shows that parts (i), (iii) and (iv) are best possible.

In other words, all of these bounds are best possible, apart from the values of the constants.

The situation is very different if we study the corresponding problems in higher dimensions. Let $K \geqslant 2$ be fixed. For every $\mathbf{x} = (x_1, \ldots, x_K)$ in $[0,1]^K$, consider aligned rectangular boxes $B(\mathbf{x}) = [0, x_1) \times \ldots \times [0, x_k)$. For every collection $\mathcal{P}$ of N points in $[0,1]^K$, we can study the corresponding discrepancy function $D[\mathcal{P}; B(\mathbf{x})]$.

We can say quite a lot. In the next two collections of results, the various parts correspond to those in Theorem RSHLDC.

Theorem RSC. *We have the following lower bound results.*

(i) *There exists a constant $c_1'(K) > 0$, depending at most on K, such that for every collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

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

(iii) *For every fixed $q > 1$, there exists a constant $c_3'(K,q) > 0$, depending at most on K and q, such that for every collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]|^q \,\mathrm{d}\mathbf{x} \geqslant c_3'(K,q)(\log N)^{(K-1)q/2}.$$

These are complemented by the following upper bound results.

(vi) *There exists a constant $c_6'(K) > 0$, depending at most on K, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \leqslant c_6'(K)(\log N)^{K-1}.$$

(vii) *For every fixed $q \geqslant 1$, there exists a constant $c_7'(K,q) > 0$, depending at most on K and q, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]|^q \,\mathrm{d}\mathbf{x} \leqslant c_7'(K,q)(\log N)^{(K-1)q/2}.$$

In part (vi), the generalization of Davenport's result to higher dimensions is due to Roth [33]. On the other hand, the techniques of Roth, Schmidt and Chen in parts (i), (iii) and (vii) work well in all dimensions.

We also have the following partial results.

Theorem RHH. *We have the following lower bound results.*

(ii) *There exists a constant $c_2'(K) > 0$, depending at most on K, such that for every collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\sup_{\mathbf{x}\in[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \geqslant c_2'(K)(\log N)^{(K-1)/2}.$$

(iv) *There exists a constant $c_4'(K) > 0$, depending at most on K, such that for every collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$, we have*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}\mathbf{x} \geqslant c_4'(K)(\log N)^{1/2}.$$

These are complemented by the following upper bound results.

(v) *There exists a constant $c_5'(K) > 0$, depending at most on K, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\sup_{\mathbf{x}\in[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \leqslant c_5'(K)(\log N)^{K-1}.$$

(vii) *There exists a constant $c_7'(K) > 0$, depending at most on K, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit cube $[0,1]^K$ such that*

$$\int_{[0,1]^K} |D[\mathcal{P};B(\mathbf{x})]| \,\mathrm{d}\mathbf{x} \leqslant c_7'(K)(\log N)^{(K-1)/2}.$$

Part (ii) is a simple consequence of part (i), whereas part (iv) is due to Halász [23]. In part (v), the generalization of Lerch's result to higher dimensions is due to Halton [24] in 1960. Part (vii) is the special case $q = 1$ of part (vii) earlier.

Note that there is a substantial gap between the bounds in parts (ii) and (v), and likewise between the bounds in parts (iv) and (vii). The question of reducing these gaps is known as the *Great Open Problem.*

In groundbreaking work, Bilyk, Lacey and Vagharshakyan have shown in 2008 that the lower bound in part (ii) can be improved somewhat. This follows a smaller improvement by Beck [6] in 1989 in the special case $K = 3$.

Theorem BLV. *For every* $K \geqslant 3$*, there exist constants* $c(K) > 0$ *and* $\delta(K)$*, satisfying* $0 < \delta(K) < 1$ *and depending at most on* K*, such that for every collection* $\mathcal{P}$ *of* N *points in the unit cube* $[0,1]^K$*, we have*

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

The case $K = 3$ is due to Bilyk and Lacey [9], while the generalization to $K \geqslant 4$ is due to Bilyk, Lacey and Vagharshakyan [10].

5. Trivial Errors and Blowing Them Up – Again

In this section, we briefly discuss the ideas of Roth in Theorem R. While the ideas easily extend to higher dimensions, we shall confine our discussion here to dimension 2.

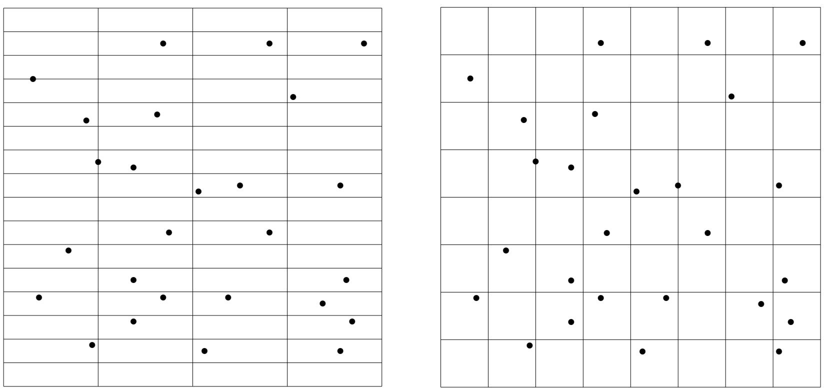
Suppose that $\mathcal{P}$ has N points.

If we partition $[0,1]^2$ into $2N$ or more parts, then at least half of these parts contain no points of $\mathcal{P}$.

More precisely, choose an integer n such that $2^{n-1} < 2N \leqslant 2^n$. If we partition $[0,1]^2$ into 2^n congruent rectangles, then at least half of these rectangles contain no points of $\mathcal{P}$.

Of course, we have to be very careful when points occasionally fall on the edges of the rectangles. We therefore adopt the convention that all rectangles are closed on the left and bottom and open on the right and top.

For instance, if $\mathcal{P}$ contains $N = 26$ points, then $n = 6$, and $2^n = 64 > 52 = 2N$. Below, the picture on the left shows 64 rectangles of size $2^{-2} \times 2^{-4}$, whereas the picture on the right shows 64 rectangles of size $2^{-3} \times 2^{-3}$. It is not difficult to check that in either case, there are at least 32 rectangles with no points of $\mathcal{P}$.



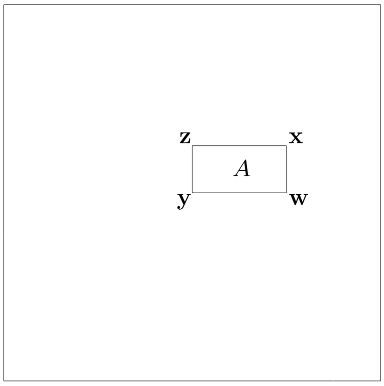
Suppose that a rectangle B is one of 2^n congruent rectangles of area 2^{-n}, and contains no points of $\mathcal{P}$. Then we have the trivial error

$$D[\mathcal{P}; B] = Z[\mathcal{P}; B] - N2^{-n} = -N2^{-n} < -\tfrac{1}{4}. \tag{6}$$

We wish to blow up such trivial errors.

In particular, we wish to make assertions concerning the discrepancy $D[\mathcal{P}; B(\mathbf{x})]$ with respect to rectangles of the form $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, where $\mathbf{x} = (x_1, x_2)$. We thus need to find some way of deducing such assertions from the information concerning the discrepancy $D[\mathcal{P}; A]$ with respect to rectangles A not anchored at the origin.

Suppose that a rectangle A, with vertices $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{w}$ as shown below, does not contain any points of $\mathcal{P}$.

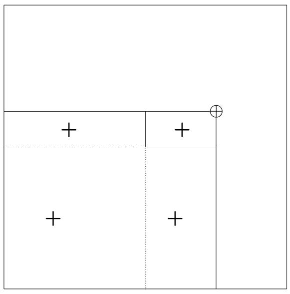


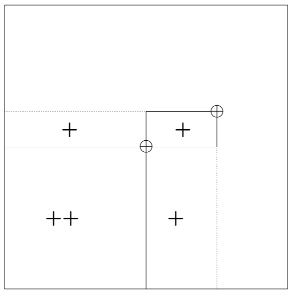
Then we have some trivial error for $D[\mathcal{P};A]$ similar to (6).

Below, the picture on the left shows the rectangle A and the rectangle $B(\mathbf{x})$, whereas the picture on the right shows the rectangle A and the rectangle $B(\mathbf{y})$. The $+$ signs indicate areas counted, so that we have noted

$$\mu(B(\mathbf{x})) \quad \text{and} \quad \mu(B(\mathbf{x}))+\mu(B(\mathbf{y}))$$

respectively.

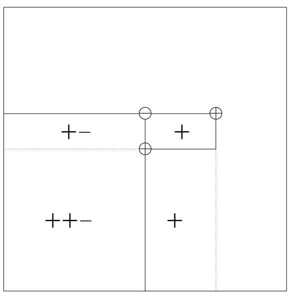


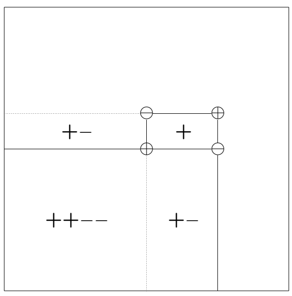


Below, the picture on the left shows the rectangle A and the rectangle $B(\mathbf{z})$, whereas the picture on the right shows the rectangle A and the rectangle $B(\mathbf{w})$. The $-$ signs indicate areas counted negatively, so that we have noted

$$\mu(B(\mathbf{x}))+\mu(B(\mathbf{y}))-\mu(B(\mathbf{z})) \quad \text{and} \quad \mu(B(\mathbf{x}))+\mu(B(\mathbf{y}))-\mu(B(\mathbf{z}))-\mu(B(\mathbf{w}))$$

respectively.





In particular, we have

$$\mu(B(\mathbf{x}))+\mu(B(\mathbf{y}))-\mu(B(\mathbf{z}))-\mu(B(\mathbf{w}))=\mu(A). \tag{7}$$

We can repeat this exercise, now counting points of $\mathcal{P}$ in the various areas instead of areas, and conclude that

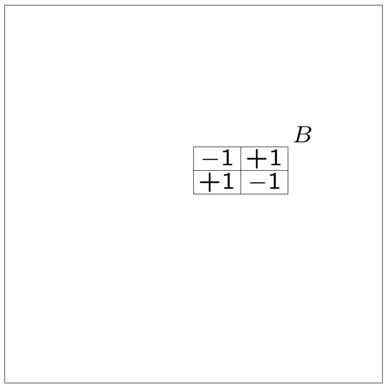
$$Z[\mathcal{P};B(\mathbf{x})]+Z[\mathcal{P};B(\mathbf{y})]-Z[\mathcal{P};B(\mathbf{z})]-Z[\mathcal{P};B(\mathbf{w})]=Z[\mathcal{P};A]. \tag{8}$$

Combining (7) and (8), we conclude that

$$D[\mathcal{P};B(\mathbf{x})]+D[\mathcal{P};B(\mathbf{y})]-D[\mathcal{P};B(\mathbf{z})]-D[\mathcal{P};B(\mathbf{w})]=D[\mathcal{P};A].$$

Note that we have given a weight of $+1$ to the points $\mathbf{x}$ and $\mathbf{y}$, and a weight of -1 to the points $\mathbf{z}$ and $\mathbf{w}$.

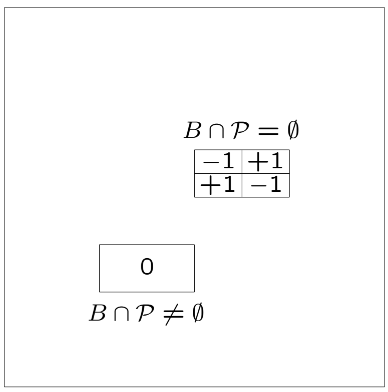
More precisely, suppose that a rectangle B is one of 2^n congruent rectangles of area 2^{-n}, and contains no points of $\mathcal{P}$. Then we define an auxiliary function f on rectangle B by writing $f(\mathbf{x})=\pm1$ as shown in the picture below.



This is an example of a Rademacher function, part of a system of well known orthogonal functions.

Among the 2^n congruent rectangles of area 2^{-n}, there will be some rectangles B which contain at least one point of $\mathcal{P}$. In order to ensure that these rectangles do not compromise the good work we have done so far, we make sure that they make no contribution at all by defining an auxiliary function f on B by writing $f(\mathbf{x})=0$.

Thus the auxiliary function f is defined on the unit square $[0,1]^2$ as shown in the picture below.



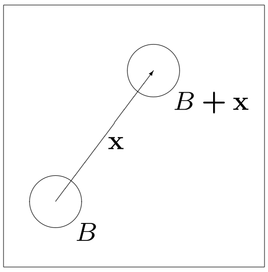
This is an example of a modified Rademacher function. The collection of such functions are easily shown to be remain orthogonal.

This idea is a fundamental contribution of Roth, and the idea of orthogonality plays an enormous role in the study of discrepancy theory, in both lower and upper bound questions.

6. A Fourier Transform Approach

Much of the progress in the 1980s is made through the use of a Fourier transform approach, introduced in the study of irregularities of integer sequences by Roth [32] in 1964; see also Beck and Chen [7, Section 9.1].

Suppose that B is a measurable subset of the unit torus $[0,1]^2$, symmetric about its centre. For every vector $\mathbf{x}$ in $[0,1]^2$, we can translate B by $\mathbf{x}$ to obtain the set $B+\mathbf{x}$.



Then

$$D[\mathcal{P};B+\mathbf{x}]=\int_{[0,1]^2}\chi_B(\mathbf{x}-\mathbf{y})(\mathrm{d}Z-N\mathrm{d}\mu)(\mathbf{y}).$$

In other words, under translation, the function

$$D=\chi_B*(\mathrm{d}Z-N\mathrm{d}\mu)$$

is a convolution.

Within this convolution, χ_B is the geometry part, depends on B but not on $\mathcal{P}$. On the other hand, $\mathrm{d}Z-N\mathrm{d}\mu$ is the measure part, depends on $\mathcal{P}$ but not on B. Ideally, we wish to study the two parts separately, but the convolution proves to be an obstacle.

The way to proceed is to consider Fourier transforms, and note that

$$\widehat{D}=\widehat{\chi_B}\cdot(\mathrm{d}\widehat{Z-N}\mathrm{d}\mu)$$

is an ordinary product. Writing $\phi=\widehat{Z-N}\mu$, we have the Parseval identity

$$\int_{[0,1]^2}|D[\mathcal{P};B+\mathbf{x}]|^2\,\mathrm{d}\mathbf{x}=\sum_{\mathbf{0}\neq\mathbf{t}\in\mathbb{Z}^2}|\widehat{\chi_B}(\mathbf{t})|^2|\phi(\mathbf{t})|^2.$$

Using this idea, Beck can establish many interesting results.

Suppose that A is a closed convex set in the unit torus $[0,1]^2$, and consider similar copies $A(\lambda,\tau,\mathbf{x})$ of A under contraction $\lambda\in[0,1]$, rotation $\tau\in\mathcal{T}$ and translation $\mathbf{x}\in[0,1]^2$. For every collection $\mathcal{P}$ of N points in $[0,1]^2$, the discrepancy function is given by

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

The following is a major result of Beck [4] in 1987.

Theorem B. *Suppose that A is a closed convex set in the unit torus $[0,1]^2$ that satisfies some minor technical condition. Then we have the following lower bound results.*

(i) *There exists a constant $c_1(A)>0$, depending at most on A, such that for every collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$, we have*

$$\int_{[0,1]^2}\int_{\mathcal{T}}\int_0^1|D[\mathcal{P};A(\lambda,\tau,\mathbf{x})]|^2\,\mathrm{d}\lambda\,\mathrm{d}\tau\,\mathrm{d}\mathbf{x}\geqslant c_1(A)N^{1/2}.$$

(ii) *There exists a constant $c_2(A) > 0$, depending at most on A, such that for every collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$, we have*

$$\sup_{\lambda,\tau,\mathbf{x}} |D[\mathcal{P}; A(\lambda,\tau,\mathbf{x})]| \geqslant c_2(A) N^{1/4}.$$

Here the minor technical condition says that A is *not too thin*, and is satisfied if A contains a small disc of some fixed small positive radius.

These results are very close to best possible.

Theorem BC. *Suppose that A is a closed convex set in the unit torus $[0,1]^2$. Then we have the following upper bound results.*

(iii) *There exists a constant $c_3(A) > 0$, depending at most on A, such that for every natural number N, there exists a collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$ such that*

$$\int_{[0,1]^2} \int_{\mathcal{T}} \int_0^1 |D[\mathcal{P}; A(\lambda,\tau,\mathbf{x})]|^2 \,\mathrm{d}\lambda \,\mathrm{d}\tau \,\mathrm{d}\mathbf{x} \leqslant c_3(A) N^{1/2}. \tag{9}$$

(iv) *There exists a constant $c_4(A) > 0$, depending at most on A, such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$ such that*

$$\sup_{\lambda,\tau,\mathbf{x}} |D[\mathcal{P}; A(\lambda,\tau,\mathbf{x})]| \leqslant c_4(A) N^{1/4} (\log N)^{1/2}. \tag{10}$$

Here part (iii) is due to Beck and Chen [8], while part (iv) is due to Beck [3].

The two theorems above extend in a straightforward manner to the unit torus $[0,1]^K$ in higher dimensions $K \geqslant 2$, with the exponents $1/2$ and $1/4$ for N replaced respectively by $1 - 1/K$ and $1/2 - 1/2K$.

Next we study the corresponding problem when rotation is not present.

Suppose that A is a closed convex set in the unit torus $[0,1]^2$, and consider now homothetic copies $A(\lambda,\mathbf{x})$ of A obtained under contraction $\lambda \in [0,1]$ and translation $\mathbf{x} \in [0,1]^2$, with no rotation. For every collection $\mathcal{P}$ of N points in $[0,1]^2$, the discrepancy function is given by

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

Suppose again that A satisfies some minor technical condition as before. What can one say about the quantities

$$\int_{[0,1]^2} \int_0^1 |D[\mathcal{P}; A(\lambda,\mathbf{x})]|^2 \,\mathrm{d}\lambda \,\mathrm{d}\mathbf{x} \quad \text{and} \quad \sup_{\lambda,\mathbf{x}} |D[\mathcal{P}; A(\lambda,\mathbf{x})]|$$

in this situation?

Beck [5] has given a partial answer to this question in 1988, and shown that the answers depend on the geometry of the boundary ∂A of the set A. We do not give Beck's precise result here, but simply illustrate the difficulty of the problem by highlighting two contrasting cases.

Suppose that A is a circular disc. Then there is *de facto* rotation, and the above two theorems apply, so there exist constants $c_1'(A) > 0$ and $c_2'(A) > 0$, depending at most on the circular disc A, such that for every collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$, we have

$$\int_{[0,1]^2} \int_0^1 |D[\mathcal{P}; A(\lambda,\mathbf{x})]|^2 \,\mathrm{d}\lambda \,\mathrm{d}\tau \,\mathrm{d}\mathbf{x} \geqslant c_1'(A) N^{1/2}$$

and

$$\sup_{\lambda,\mathbf{x}} |D[\mathcal{P}; A(\lambda,\mathbf{x})]| \geqslant c_2'(A) N^{1/4}.$$

Suppose next that A is an aligned rectangle. Then this becomes a problem very much like the classical one studied by Roth, so there exist constants $c_1'(A) > 0$ and $c_2'(A) > 0$, depending at most on the rectangle A, such that for every collection $\mathcal{P}$ of N points in the unit torus $[0,1]^2$, we have

$$\int_{[0,1]^2} \int_0^1 |D[\mathcal{P}; A(\lambda, \mathbf{x})]|^2 \,\mathrm{d}\lambda \,\mathrm{d}\tau \,\mathrm{d}\mathbf{x} \geqslant c_1'(A) \log N$$

and

$$\sup_{\lambda, \mathbf{x}} |D[\mathcal{P}; A(\lambda, \mathbf{x})]| \geqslant c_2'(A)(\log N)^{1/2}. \tag{11}$$

Corresponding to the Great Open Problem, we may ask whether the estimate in (11) can be improved to something of order of magnitude $\log N$.

Furthermore, no result on the corresponding problem in any dimension $K > 2$ has been established. Indeed, a complete solution to the order of magnitude of

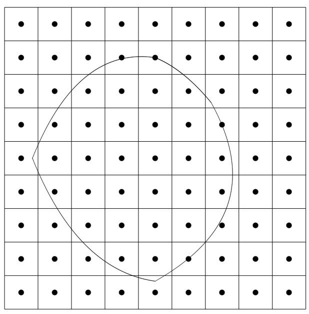
$$\sup_{\lambda, \mathbf{x}} |D[\mathcal{P}; A(\lambda, \mathbf{x})]|$$

for arbitrary closed convex sets A in the unit torus $[0,1]^K$ in arbitrary dimensions $K \geqslant 2$ is known as the *Greater Open Problem*.

For a simple introduction to the Fourier transform technique, see the survey by Chen [13].

7. Some Probabilistic Techniques

In this section, we briefly describe the basic ideas underpinning the upper bounds (9) and (10). For simplicity, let us assume that $\mathcal{P}$ is a set of $N = M^2$ points in the unit torus $[0,1]^2$. We can split the unit torus into N little squares S of area $1/N$ in the obvious way, and place a point in the middle of each such little square.



Suppose first of all that $B \cap S = \emptyset$. Then clearly

$$D[\mathcal{P}; B \cap S] = Z[\mathcal{P}; \emptyset] - N\mu(\emptyset) = 0 - 0 = 0.$$

Suppose next that $S \subseteq B$, so that $B \cap S = S$. Then clearly

$$D[\mathcal{P}; B \cap S] = Z[\mathcal{P}; S] - N\mu(S) = 1 - 1 = 0.$$

It follows that

$$D[\mathcal{P}; B] = \sum_{\substack{S \\ \partial B \cap S \neq \emptyset}} D[\mathcal{P}; B \cap S],$$

and so

$$|D[\mathcal{P}; B]| \leqslant \sum_{\substack{S \\ \partial B \cap S \neq \emptyset}} |D[\mathcal{P}; B \cap S]| \leqslant \#\{S : \partial B \cap S \neq \emptyset\},$$

since $|D[\mathcal{P}; B \cap S]| \leqslant 1$ trivially for every little square S.

If B is convex, then $\#\{S : \partial B \cap S \neq \emptyset\} \leqslant CN^{1/2}$, and this gives a trivial upper bound $CN^{1/2}$ compared to that given in (10). We need to aim to take the square root of this trivial upper bound.

To achieve a better result, we randomise the N points by allowing each to move in a uniform fashion within its own little square, and independently of each other, and then apply large deviation type techniques in probability theory.

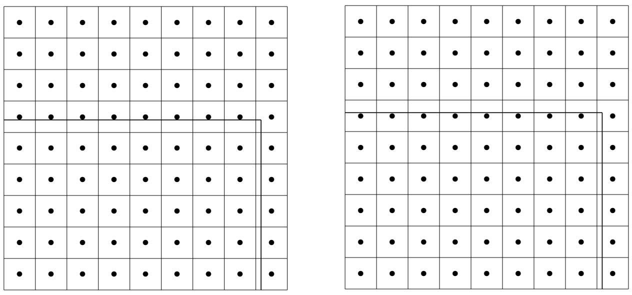
Fore more details, and a brief description on deducing the bound (9), see the survey by Chen [14].

8. A Lattice Point Approach to the Classical Problem

Let us return to the classical discrepancy problem, and attempt to understand the upper bounds in parts (v) and (vi) of Theorem RSHLDC.

The smarty pants may suggest the following *obvious* approach. For simplicity, let us again assume that $\mathcal{P}$ is a set of $N = M^2$ points in the unit square $[0,1]^2$. We can split the unit square into N little squares of area $1/N$ in the obvious way, and place a point in the middle of each such little square.

However, let us consider a rectangle B_1 as shown in the picture below on the left, and a slightly bigger rectangle B_2 as shown in the picture below on the right.

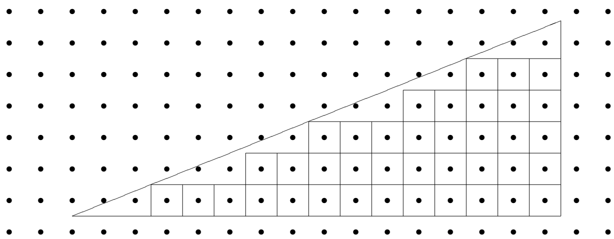


Clearly, $Z[\mathcal{P}; B_1]$ and $Z[\mathcal{P}; B_2]$ differ by nearly M, and since the two rectangles have almost identical areas, it follows that $D[\mathcal{P}; B_1]$ and $D[\mathcal{P}; B_2]$ differ by more than $M/2$, and so the approach is doomed, as no probabilistic technique can reduce the estimates to logarithmic size.

However, our smarty pants are not too stupid after all. A square lattice of points still works, but one needs to rotate it. The question of lattice points in right-angled triangles has been studied as long ago as the 1920s by Hardy and Littlewood [26, 27].

Consider the lattice $\mathbb{Z}^2$, and place a right angled triangle in such as way that the horizontal and vertical sides sit halfway between consecutive rows and columns respectively of lattice points.

We want to estimate the discrepancy between the actual number of lattice points that fall into the triangle and the area of the triangle. Here, the triangle has been placed to ensure that there is essentially no discrepancy arising from the horizontal and vertical edges, and so any discrepancy must arise from the hypothenuse, as shown in the picture below.



In fact, any estimate on the discrepancy depends on arithmetic properties of the slope of the hypothenuse. Thus we enter the area of diophantine approximation, and in particular, badly approximable numbers, such as $\sqrt{2}$.

In short, if our smarty pants had rotated the square lattice by an angle θ such that $\tan\theta = \sqrt{2}$, then they would not have looked so stupid after all.

Indeed, using a variant of this idea, as well as a very clever reflection principle, Davenport [20] has established part (vi) of Theorem RSHLDC.

9. More on the Classical Problem

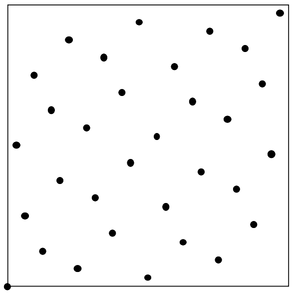
A very fruitful approach to upper bounds in the classical problem is given by van der Corput point sets and their variants.

The van der Corput point set $\mathcal{P}_h$ of 2^h points in $[0,1]^2$ is given by

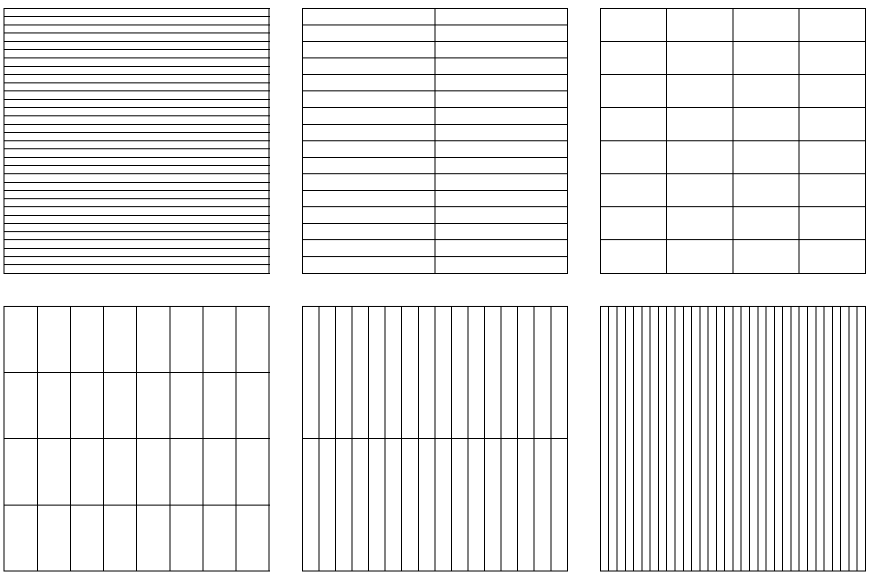
$$\mathcal{P}_h = \{(0.a_1 \ldots a_h, 0.a_h \ldots a_1) : a_1, \ldots, a_h \in \{0,1\}\},$$

using dyadic expansions.

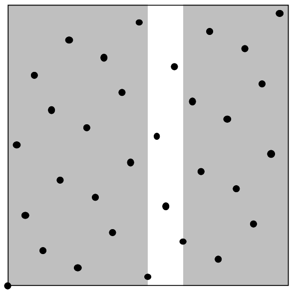
Below is a picture of the van der Corput point set $\mathcal{P}_5$ of $2^5 = 32$ points.



One crucial property of the van der Corput point sets $\mathcal{P}_h$ is that many rectangles in $[0,1]^2$ contains the expected number of points, under the convention that all rectangles are closed on the left and bottom and open on the right and top. For instance, we can partition the unit square $[0,1]^2$ into $2^5 = 32$ congruent rectangles in six different ways. In each case, each small rectangle contains precisely one of the 32 points of $\mathcal{P}_5$.



Another crucial property of the van der Corput point sets $\mathcal{P}_h$ is periodicity, best illustrated by the picture below concerning the van der Corput point set $\mathcal{P}_5$.



Note that the vertical distribution of the points in the white rectangle is periodic. This periodicity property permits the use of Fourier series.

Let us describe the use of van der Corput point sets to deduce Davenport's theorem, that there is a constant $c_6 > 0$ such that for every natural number $N \geqslant 2$, there exists a collection $\mathcal{P}$ of N points such that

$$\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \leqslant c_6 \log N.$$

Suppose that $N = 2^h$, so that we are tempted to test the van der Corput point set $\mathcal{P}_h$. However, a result of Halton and Zaremba [25] says that

$$\int_{[0,1]^2} |D[\mathcal{P}_h; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} = \frac{h^2}{64} + \text{error of order } h,$$

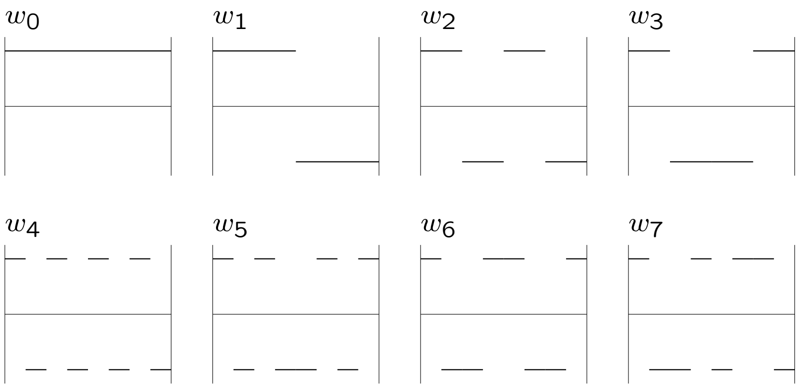
thus of order $(\log N)^2$, and so $\mathcal{P}_h$ will not give Davenport's theorem. We need to modify $\mathcal{P}_h$ by introducing a probabilistic parameter and taking an average.

The probabilistic variable in the argument of Roth [33] is translation. For this to work, one needs the periodicity property of the van der Corput point sets. An alternative approach by Chen [12] uses digit shifts as the probabilistic variable, and does not require periodicity, and so has the benefit of working for sets other than the van der Corput sets.

The van der Corput sets possess a third crucial property which is the basis of the more recent breakthrough by Chen and Skriganov [15, 16].

Take $\oplus$ to be coordinatewise and digitwise addition modulo 2. Then $(\mathcal{P}_h, \oplus)$ is a group isomorphic to the additive group $\mathbb{Z}_2^h$.

The characters of such groups are the Walsh functions with values ± 1. Below are pictures of the first eight Walsh functions.



The Walsh functions form an orthonormal basis for $L^2([0,1])$.

Furthermore, it is also well known that the 2-dimensional Walsh functions form an orthonormal basis for $L^2([0,1]^2)$, and this leads to Fourier–Walsh series and analysis on $L^2([0,1]^2)$, an analogue of the classical Fourier series and analysis on $L^2([0,1]^2)$ based on the orthonormal system of exponential functions. We thus have at our disposal a tool which is much better suited to understanding the intricacies of the van der Corput point sets. For a survey of this area, the reader is referred to the forthcoming survey by Chen and Skriganov [17].

10. Connections with Other Areas

The work of Bilyk, Lacey and Vagharshakyan [10] has highlighted connections between various areas, via the small ball inequality. The small ball inequality in probability theory and harmonic analysis in two dimensions is due to Talagrand [38]. An alternative proof, due to Temlyakov [40], follows a variant of Roth's original ideas by Halász [23] in discrepancy theory.

Work on the Great Open Problem in discrepancy theory is related also to the small ball conjecture for brownian sheets in probability, as studied by Kuelbs and Li [28], and by Dunker, Kühn, Lifshits and Linde [22], and also related to the question of the entropy of mixed smoothness classes in approximation theory, as studied by Temlyakov [39].

11. Applications to Numerical Integration

Suppose that $f(\mathbf{x})$ is a function defined on the unit square $[0,1]^2$. The *continuous average* of the function is given by the integral

$$I(f) = \int_{[0,1]^2} f(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{12}$$

However, there are many such functions that arise in the practical world for which we do not have enough technique to evaluate this integral. To overcome this, we often take a set $\mathcal{P}$ of N points in $[0,1]^2$, where N is reasonably large, and calculate instead the *discrete average*

$$S(f,\mathcal{P}) = \frac{1}{N}\sum_{\mathbf{p}\in\mathcal{P}} f(\mathbf{p}), \tag{13}$$

and then use this as an approximation to the integral (12). Then the error of the approximation is given by the difference

$$E(f,\mathcal{P}) = S(f,\mathcal{P}) - I(f). \tag{14}$$

Suppose that B is a measurable subset of the unit square $[0,1]^2$. Consider the special case when f is the characteristic function of B, so that

$$f(\mathbf{x}) = \chi_B(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x}\in B,\\ 0, & \text{if } \mathbf{x}\not\in B.\end{cases}$$

Then

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

clearly a discrepancy problem.

Thus the study of the error of approximation (14) is a generalization of the problem of discrepancy.

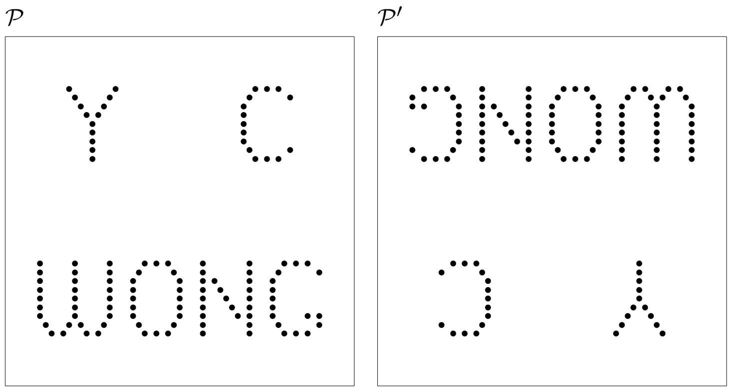
We also have the following remarkable observation of Woźniakowski [41] in 1991.

Theorem W. *Let $C(2)$ denote the class of continuous real valued functions in $[0,1]^2$, endowed with the Wiener sheet measure ν. For every function $f \in C(2)$, let $I(f)$, $S(f,\mathcal{P})$ and $E(f,\mathcal{P})$ be defined by* (12)–(14). *Then*

$$\int_{C(2)} |E(f,\mathcal{P})|^2 \,\mathrm{d}\nu = \frac{1}{N^2} \int_{[0,1]^2} |D[\mathcal{P}'; B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x},$$

where the set $\mathcal{P}'$ is closely related to the set $\mathcal{P}$, with $\mathcal{P}' = \{(1,1) - \mathbf{p} : \mathbf{p} \in \mathcal{P}\}$.

An illustration is given below.



On the other hand, many physical or economical phenomena are governed by many parameters, entailing functions of many real variables.

Suppose that $f(\mathbf{x})$ is a function defined on the unit cube $[0,1]^K$, where the dimension K is large, corresponding to the number of real variables that we require. Again, we often cannot evaluate the integral

$$I(f) = \int_{[0,1]^K} f(\mathbf{x}) \,\mathrm{d}\mathbf{x},$$

and instead take a set $\mathcal{P}$ of N points in $[0,1]^K$ and evaluate the sum

$$S(f,\mathcal{P}) = \frac{1}{N} \sum_{\mathbf{p} \in \mathcal{P}} f(\mathbf{p}).$$

The corresponding discrepancy problem involves the function $D[\mathcal{P}; B(\mathbf{x})]$ in $[0,1]^K$, concerning discrepancy with respect to rectangular boxes of the type

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

where $\mathbf{x} = (x_1, \ldots, x_K) \in [0,1]^K$.

Here the best upper bounds are given by Halton [24] and Roth [33], with

$$\sup_{\mathbf{x} \in [0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]| \leqslant C_1(K)(\log N)^{K-1} \tag{15}$$

and

$$\int_{[0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \leqslant C_2(K)(\log N)^{K-1}. \tag{16}$$

Now smarty pants will have us use square lattices. We can then obtain the theoretically weaker, and indeed trivial, upper bounds

$$\sup_{\mathbf{x} \in [0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]| \leqslant C_1^*(K) N^{1-1/K} \tag{17}$$

and

$$\int_{[0,1]^K} |D[\mathcal{P}; B(\mathbf{x})]|^2 \,\mathrm{d}\mathbf{x} \leqslant C_2^*(K) N^{2-2/K}. \tag{18}$$

For large values of K, the bounds (15) and (16) are better than the bounds (17) and (18) only when the number of points N is very large and, indeed, beyond the scope of computation at the current state of computer science. Woźniakowski has coined this the *curse of dimensionality*.

For further reading on numerical integration, the reader is referred to the recent monograph of Dick and Pillichshammer [21].

References

[1] T. van Aardenne-Ehrenfest. Proof of the impossibility of a just distribution of an infinite sequence of points over an interval. *Proc. Kon. Ned. Akad. v. Wetensch.*, 48 (1945), 266–271.

[2] T. van Aardenne-Ehrenfest. On the impossibility of a just distribution. *Proc. Kon. Ned. Akad. v. Wetensch.*, 52 (1949), 734–739.

[3] J. Beck. Balanced two-colorings of finite sets in the square I. *Combinatorica*, 1 (1981), 327–335.

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

[5] J. Beck. Irregularities of distribution II. *Proc. London Math. Soc.*, 56 (1988), 1–50.

[6] J. Beck. A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution. *Compos. Math.*, 72 (1989), 269–339.

[7] J. Beck, W.W.L. Chen. *Irregularities of Distribution* (Cambridge Tracts in Mathematics 89, Cambridge University Press, 1987; paperback reprint 2008).

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

[9] D. Bilyk, M.T. Lacey. On the small ball inequality in three dimensions. *Duke Math. J.*, 143 (2008), 81–115.

[10] D. Bilyk, M.T. Lacey, A. Vagharshakyan. On the small ball inequality in all dimensions. *J. Funct. Anal.*, 254 (2008), 2470–2502.

[11] W.W.L. Chen. On irregularities of distribution. *Mathematika*, 27 (1980), 153–170.

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

[13] W.W.L. Chen. Fourier techniques in the theory of irregularities of point distribution. *Fourier Analysis and Convexity* (L. Brandolini, L. Colzani, G. Travaglini, eds.), pp. 59–82 (Applied and Numerical Harmonic Analysis, Birkhuser, 2004).

[14] W.W.L. Chen. Upper bounds in discrepancy theory. *Monte Carlo and Quasi-Monte Carlo Methods 2010* (L. Plaskota, H. Woźniakowski, eds.), pp. 23–41 (Springer Proceedings in Mathematics and Statistics 23, Springer, 2012).

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

[16] W.W.L. Chen, M.M. Skriganov. Orthogonality and digit shifts in the classical mean squares problem in irregularities of point distribution. *Diophantine Approximation: Festschrift for Wolfgang Schmidt* (H.P. Schlickewei, K. Schmidt, R.F. Tichy, eds.), pp. 141–159 (Developments in Mathematics 16, Springer, 2008).

[17] W.W.L. Chen, M.M. Skriganov. Upper bounds in classical discrepancy theory. *A Panorama of Discrepancy Theory* (W.W.L. Chen, A. Srivastav, G. Travaglini, eds.), pp. 3–64 (Lecture Notes in Mathematics, Springer, to appear).
http://rutherglen.science.mq.edu.au/wchen/researchfolder/cb04.pdf

[18] J.G. van der Corput. Verteilungsfunktionen I. *Proc. Kon. Ned. Akad. v. Wetensch.*, 38 (1935), 813–821.

[19] J.G. van der Corput. Verteilungsfunktionen II. *Proc. Kon. Ned. Akad. v. Wetensch.*, 8 (1935), 1058–1066.

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

[21] J. Dick, F. Pillichshammer. *Digital Nets and Sequences* (Cambridge University Press, 2010).

[22] T. Dunker, T Kühn, M. Lifshits, W. Linde. Metric entropy of the integration operator and small ball probabilities for the Brownian sheet. C. R. Acad. Sci. Paris Sér. I Math., 326 (1998), 347–352.

[23] G. Halász. On Roth's method in the theory of irregularities of point distributions. *Recent Progress in Analytic Number Theory*, volume 2 (H. Halberstam, C. Hooley, eds.), pp. 79–94 (Academic Press, 1981).

[24] J.H. Halton. On the efficiency of certain quasirandom sequences of points in evaluating multidimensional integrals. *Numer. Math.*, 2 (1960), 84–90.

[25] J.H. Halton, S.K. Zaremba. The extreme and L^2 discrepancies of some plane sets. *Monatsh. Math.*, 73 (1969), 316–328.

[26] G.H. Hardy, J.E. Littlewood. The lattice points of a right angled triangle I. *Proc. London Math. Soc.*, 3 (1920), 15–36.
[27] G.H. Hardy, J.E. Littlewood. The lattice points of a right angled triangle II. *Abh. Math. Sem. Hamburg*, 1 (1922), 212–249.
[28] J Kuelbs, W.V. Li. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.*, 116 (1993), 133–157.
[29] M. Lerch. Question 1547. *L'Intermediaire Math.*, 11 (1904), 144–145.
[30] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide* (Algorithms and Combinatorics 18, Springer, 1999; paperback reprint 2010).
[31] K.F. Roth. On irregularities of distribution. *Mathematika*, 1 (1954), 73–79.
[32] K.F. Roth. Remark concerning integer sequences. *Acta Arith.*, 9 (1964), 257–260.
[33] K.F. Roth. On irregularities of distribution IV. *Acta Arith.*, 37 (1980), 67–75.
[34] W.M. Schmidt. Irregularities of distribution VII. *Acta Arith.*, 21 (1972), 45–50.
[35] W.M. Schmidt. Irregularities of distribution IX. *Acta Arith.*, 27 (1975), 385–396.
[36] W.M. Schmidt. Irregularities of distribution X. *Number Theory and Algebra* (H. Zassenhaus, ed.), pp. 311–329 (Academic Press, 1977).
[37] W.M. Schmidt. *Lectures on Irregularities of Distribution* (Tata Institute of Fundamental Research, 1977).
[38] M. Talagrand. The small ball problem for the Brownian sheet. *Ann. Probab.*, 22 (1994), 1331–1354.
[39] V.N. Temlyakov. *Approximation of Functions with Bounded Mixed Derivative* (Proceedings of the Steklov Institute of Mathematics, 1989).
[40] V.N. Temlyakov. An inequality for trigonometric polynomials and its application for estimating the entropy numbers. *J. Complexity*, 11 (1995), 293–307.
[41] H. Woźniakowski. Average case complexity of multivariate integration. *Bull. American Math. Soc.*, 24 (1991), 185–194.

Department of Mathematics, Macquarie University, Sydney NSW 2109, Australia
E-mail address: `william.chen@mq.edu.au`