DAVENPORT'S THEOREM
IN GEOMETRIC DISCREPANCY THEORY

W.W.L. CHEN

In memory of Heini Halberstam

ABSTRACT. Davenport's theorem was established nearly a lifetime ago, but there has been some very interesting recent developments. The various proofs over the years bring in different ideas from number theory, probability theory, analysis and group theory. In this short survey, we shall not present complete proofs, but will describe instead some of these underlying ideas.

1. INTRODUCTION

Davenport's theorem in geometric discrepancy theory, or irregularities of point distribution, concerns the mean squares discrepancy of point distributions in the unit square with respect to anchored and aligned rectangles, and shows that Roth's astonishing result in 1954 is best possible in dimension 2.

More precisely, let $\mathcal{P}$ be a set of N points in the unit square $[0,1]^2$. For every $\mathbf{x} = (x_1, x_2) \in [0,1]^2$, we consider the discrepancy

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

where $\#(\mathcal{P} \cap B(\mathbf{x}))$ denotes the number of points of $\mathcal{P}$ that fall into the rectangle $B(\mathbf{x}) = [0, x_1) \times [0, x_2)$, and μ denotes the usual Lebesgue area measure. We are interesting in the L^2-norm

$$\|D[\mathcal{P}]\|_2 = \left(\int_{[0,1]^2} |D[\mathcal{P}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{x} \right)^{1/2}$$

of the discrepancy function.

In the groundbreaking paper of Roth [17] in 1954, it is shown that there exists a positive absolute constant C_1 such that for every set $\mathcal{P}$ of N points in $[0,1]^2$,

$$\|D[\mathcal{P}]\|_2 \geqslant C_1 (\log N)^{1/2}.$$

This lower bound is essentially best possible, in view of the following result in 1956.

Theorem (Davenport [10]). *There exists a positive absolute constant C_2 such that for every integer $N \geqslant 2$, there exists a set $\mathcal{P}$ of N points in $[0,1]^2$ such that*

$$\|D[\mathcal{P}]\|_2 \leqslant C_2 (\log N)^{1/2}. \tag{1}$$

2000 *Mathematics Subject Classification.* 11K38; 11K60, 42B05, 42C10, 60Fxx.
Key words and phrases. Discrepancy; irregularities of point distribution.

While Davenport's theorem seems nearly a lifetime ago, there has been some very interesting recent developments. The various proofs over the years bring in different ideas from number theory, probability theory, analysis and group theory. Here we shall describe some of these underlying ideas, and compare the similarities and differences of some of the proofs.

To prove Davenport's theorem, we simply need to construct point sets $\mathcal{P}$ of N points that satisfy the inequality (1). However, the construction of such sets turns out to be rather delicate, as illustrated by an observation of Lev [15] in 1996.

For any point set $\mathcal{P}$ and every vector $\mathbf{t} \in [0,1]^2$, let $\mathcal{P}+\mathbf{t}$ denote the image of $\mathcal{P}$ under translation by $\mathbf{t}$ modulo $[0,1)^2$. Then for every positive integer N and every point set $\mathcal{P}$ of N points,

$$\sup_{\mathbf{t}\in[0,1]^2} \|D[\mathcal{P}+\mathbf{t}]\|_2 \gg \log N.$$

Put simply, any effort in finding a point set $\mathcal{P}$ that satisfies the inequality (1) can be wasted through a simple translation.

In this short survey, we shall present ten of the many proofs of Davenport's theorem, numbered in chronological order. While some of these proofs are direct attempts at establishing the result, others have been discovered through efforts to establish generalizations and extensions of Davenport's theorem.

Notation. For any function f and any positive function g, we write $f = O(g)$ or $f \ll g$ to denote that there exists a positive absolute constant C such that $|f| \leqslant Cg$, and write $f = O_\delta(g)$ or $f \ll_\delta g$ to denote that there exists a positive constant $C(\delta)$, which may depend on the parameter δ, such that $|f| \leqslant C(\delta)g$. For any two positive functions f and g, we write $f \gg g$ and $f \gg_\delta g$ to denote respectively $g \ll f$ and $g \ll_\delta f$. We also write $f \asymp g$ if $f \ll g$ and $g \ll f$.

For any real number z, we denote by $[z]$ the greatest integer not exceeding z, write $\{z\} = z - [z]$ to denote the fractional part of z, and write $\Psi(z) = \{z\} - \frac{1}{2}$ to denote the sawtooth function. We also use the notation $e(z) = \mathrm{e}^{2\pi \mathrm{i} z}$.

For any finite set S, we denote by $\#S$ the number of elements of S, counted with multiplicity.

2. Diophantine Approximation Approach

We begin by making a seriously flawed attempt. For simplicity, assume that N is the square of a positive integer. One is then tempted to partition the unit square $[0,1]^2$ into N little squares of area $1/N$ in the natural way, and place a point at the center of each little square, as shown in Figure 1.

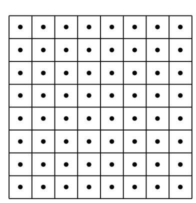

Fig.1. A seriously flawed attempt

As one moves the top boundary of $B(\mathbf{x})$ across a row of points, the discrepancy function $D[\mathcal{P};B(\mathbf{x})]$ jumps by an amount which can be as large as $N^{1/2}$, and so the estimate $\|D[\mathcal{P}]\|_2 \ll N^{1/2}$ is as much as we can deduce from this construction.

However, if we rotate a suitably sized square lattice by a suitable angle, then our task is not as hopeless if it may seem. Here we appeal to the famous results of Hardy and Littlewood [13, 14] concerning lattice points in a right angled triangle. We place a right angled triangle on the square lattice $\mathbb{Z}^2$ in such a way that the horizontal edge is precisely halfway between two consecutive rows of lattice points and the vertical edge is precisely halfway between two consecutive columns of lattice points, as shown in Figure 2.

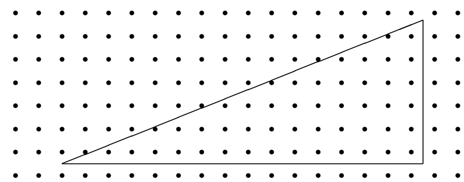

Fig.2. Lattice points in a right angled triangle

We now approximate the number of lattice points in the triangle by the area of the triangle, and the estimate is good when the slope of the hypothenuse is a badly approximable number.

Our seventh proof of Davenport's theorem, by Beck and Chen [4], makes use of this idea. We consider the lattice $L = (N^{-1/2}\mathbb{Z})^2$ which contains N points per unit area. Let $L(\theta)$ denote the image of L under anticlockwise rotation by an angle θ, where $\tan\theta$ is a badly approximable number. Then the point set

$$\mathcal{P} = L(\theta) \cap [0,1)^2 \tag{2}$$

contains roughly N points. Unfortunately, this is insufficient to give an estimate of the form (1). We shall return to this approach later.

This seventh proof, given only implicitly in [4], essentially follows the same ideas as the original proof by Davenport. The paper [4] concerns the general problem of point distributions with respect to homothetic copies of a given convex polygon, and uses a result of Davenport [11] in 1964 on simultaneous diophantine approximation. Davenport's theorem can essentially be considered a special case of this study.

We now describe Davenport's original approach in [10].

Consider the set

$$\mathcal{Q} = \Lambda \cap ([0,1) \times [0,N)) \tag{3}$$

of N points in $[0,1) \times [0,N)$, where Λ is the lattice generated by the vectors $(1,0)$ and $(\phi,1)$. The corresponding discrepancy function is given by

$$E[\mathcal{Q}; B(x,y)] = \#(\mathcal{Q} \cap B(x,y)) - xy,$$

where $B(x,y) = [0,x) \times [0,y) \subseteq [0,1) \times [0,N)$. It can then be shown that

$$E[\mathcal{Q}; B(x,y)] = \sum_{0 \neq m \in \mathbb{Z}} \left(\frac{1 - e(-mx)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leqslant n < y} e(\phi n m) \right) + O(1). \tag{4}$$

Here the term 1 arises from the hypothesis that $B(x,y)$ is anchored at the origin, and causes technical difficulties.

To overcome this handicap, Davenport considers the mirror image of Λ across vertical axis. More precisely, consider the set $\mathcal{Q}' = \Lambda' \cap ([0,1) \times [0,N))$ of N points in $[0,1) \times [0,N)$, where Λ' is the lattice generated by the vectors $(1,0)$ and $(-\phi,1)$.

For the set $\mathcal{Q}^* = \mathcal{Q} \cup \mathcal{Q}'$ of $2N$ points in $[0,1) \times [0,N)$, the discrepancy function is now given by

$$F[\mathcal{Q}^*; B(x,y)] = \#(\mathcal{Q}^* \cap B(x,y)) - 2xy.$$

It can then be shown that

$$F[\mathcal{Q}^*; B(x,y)] = \sum_{0 \neq m \in \mathbb{Z}} \left(\frac{e(mx) - e(-mx)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leqslant n < y} e(\phi n m) \right) + O(1).$$

Integrating with respect to $x \in [0,1]$ and applying Parseval's identity, the problem is then reduced to showing that

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leqslant n < y} e(\phi n m) \right|^2 \ll_{\phi} \log N \tag{5}$$

if ϕ is badly approximable. This is sufficient to deduce Davenport's theorem.

This technique is now known as Davenport's reflection principle.

Our second proof, by Roth [18] in 1979, has Davenport's construction (3) as the starting point. However, instead of using reflection, Roth considers horizontal translations of the lattice Λ through a period. For every $t \in [0,1]$, consider the translate $\Lambda(t) = \Lambda + (t,0)$ of Λ. For the set

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

of N points in $[0,1) \times [0,N)$, the discrepancy function is now given by

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

It can then be shown that

$$E[\mathcal{Q}(t); B(x,y)] = \sum_{0 \neq m \in \mathbb{Z}} \left(\frac{1 - e(-mx)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leqslant n < y} e(\phi n m) \right) e(tm) + O(1).$$

Integrating with respect to $t \in [0,1]$ and applying Parseval's identity, the problem is again reduced to the estimate (5) if ϕ is badly approximable. This enables us to deduce an average version of Davenport's theorem over the parameter t. Davenport's theorem is thus realized for some value of $t \in [0,1]$.

This is the first instance when probability is used in discrepancy theory, and is now known as Roth's probabilistic method. Note that using this technique, we can only show that sets $\mathcal{P}$ that satisfy the estimate (1) exist, but we cannot give them explicitly.

At this point, let us make a digression and return to our seventh proof which we described in part at the beginning of this section. To proceed with our proof, we need to consider translates $\mathcal{P} + \mathbf{t}$ of the set (2), and estimate the integral

$$\int_R |D[\mathcal{P} + \mathbf{t}; B(\mathbf{x})]|^2 \, \mathrm{d}\mathbf{t},$$

where R denotes a fundamental region of the lattice $L(\theta)$.

Let us now return to Davenport's approach using the set (3) for some badly approximable number ϕ. Recall that Davenport uses mirror reflection and Roth uses periodic translation. The question now arises as to whether either of these is necessary.

Our fourth proof, given by Sós and Zaremba [20] in a paper submitted in 1980 and published in 1982 in a volume dated 1979, gives a partial answer to this question. This involves the continued fractions expansion of ϕ, and we shall only describe this briefly. Write

$$\phi = [a_0; a_1, a_2, a_3, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}}$$

in continued fractions notation, where the partial quotients $a_1, a_2, a_3, \ldots$ are positive integers if ϕ is irrational. If the partial quotients are all equal, such as when

$$\phi = \frac{1+\sqrt{5}}{2} = [1; 1, 1, 1, 1, 1, 1, \ldots]$$

is the golden ratio, then the set $\mathcal{P}$ of N points in $[0,1)^2$, obtained by a linear contraction in the vertical direction from the set $\mathcal{Q} \subset [0,1) \times [0,N)$ given by (2), satisfies the estimate (1), and so gives a proof of Davenport's theorem without mirror reflection or periodic translation.

The complete solution to this question is given by Bilyk [6] recently.

For the remainder of this section, we shall assume that the number ϕ is badly approximable, so that the partial quotients are bounded.

Consider the main term in the expression (4), given by

$$M_y(x) = \sum_{0 \neq m \in \mathbb{Z}} \left(\frac{1 - e(-mx)}{2\pi \mathrm{i} m} \right) \left(\sum_{0 \leqslant n < y} e(\phi n m) \right).$$

For fixed y, we now perform Fourier analysis on the variable x. It then follows from Parseval's identity that

$$\|M_y\|_2^2 \leqslant |\widehat{M_y}(0)|^2 + C \sum_{m=1}^{\infty} \frac{1}{m^2} \left| \sum_{0 \leqslant n < y} e(\phi n m) \right|^2 \tag{6}$$

for some positive absolute constant C. Recall now the estimate (5) when ϕ is badly approximable. Thus our problem is reduced to studying the term

$$\widehat{M_y}(0) = \sum_{0 \neq m \in \mathbb{Z}} \frac{1}{2\pi \mathrm{i} m} \left(\sum_{0 \leqslant n < y} e(\phi n m) \right) = - \sum_{0 \leqslant n < y} \Psi(\phi n).$$

For convenience, we change notation once more, and write

$$S_y(\phi) = \sum_{0 \leqslant n < y} \Psi(\phi n). \tag{7}$$

Then the inequality (6) becomes

$$\|M_y\|_2^2 \leqslant S_y^2(\phi) + C_\phi \log N,$$

where C_ϕ is a positive constant depending at most on ϕ. We can now integrate over $y \in [0,N)$ and rescale the set $\mathcal{Q}$ to obtain a set $\mathcal{P}$ in $[0,1)^2$ such that

$$\|D[\mathcal{P}]\|_2^2 \leqslant \frac{1}{N} \sum_{y=0}^{N-1} S_y^2(\phi) + C_\phi \log N.$$

Thus Davenport's theorem will follow without the need for mirror reflection or periodic translation if and only if

$$\frac{1}{N}\sum_{y=0}^{N-1} S_y^2(\phi) \ll_\phi \log N.$$

Sums of the type (7) have been studied extensively by Beck. In [2, Theorem 3.2], it is shown that for the Cesaro mean

$$T_N(\phi) = \frac{1}{N}\sum_{y=0}^{N-1} S_y(\phi)$$

of these sums, we have the estimate

$$T_N(\phi) = \frac{1}{12}\sum_{k=1}^{n}(-1)^k a_k + O\left(\max_{1\leqslant i\leqslant n} a_i\right),$$

where n is smallest index for which $q_n \geqslant N$, where q_n is the denominator of the n-th convergent to ϕ. It is well known that $n \asymp \log N$. In the same paper, it is shown that the second moments $V_N(\phi)$ of these sums satisfy

$$V_N(\phi) = \frac{1}{N}\sum_{y=0}^{N-1}(S_y(\phi) - T_N(\phi))^2 \asymp \sum_{q_m\leqslant N} a_m^2 \ll_\phi \log N. \tag{8}$$

Furthermore, it is shown in [2, Theorem 4.1] that the central limit theorem holds for these sums, in the form

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{0 \leqslant y \leqslant N-1 : \frac{S_y(\phi) - T_N(\phi)}{\sqrt{V_N(\phi)}} \leqslant \lambda\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda} \mathrm{e}^{-t^2/2}\,\mathrm{d}t. \tag{9}$$

One can then deduce that

$$T_N(\phi) \leqslant \left(\sum_{y=0}^{N-1} S_y^2(\phi)\right)^{1/2} \ll_\phi T_N(\phi) + \sqrt{\log N}.$$

Here the first inequality is simply the Cauchy–Schwarz inequality, while the second inequality can be deduced from (8) and (9). We thus conclude that

$$\|D[\mathcal{P}]\|_2^2 \asymp_\phi \max\left\{\left|\sum_{k=1}^{n}(-1)^k a_k\right|^2, \log N\right\},$$

This means that the set $\mathcal{P}$ gives a proof of Davenport's theorem directly, without mirror reflection or periodic translation, if and only if the number ϕ satisfies

$$\left|\sum_{k=1}^{n}(-1)^k a_k\right| \ll_\phi \sqrt{n}.$$

Note that this last condition is satisfied by $\sqrt{2} = [1; 2, 2, 2, 2, 2, 2, \ldots]$, but not by $\sqrt{3} = [1; 1, 2, 1, 2, 1, 2, \ldots]$.

3. Use of van der Corput Point Sets

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

$$\mathcal{P}_h = \{(0.a_1 \dots a_h, 0.a_h \dots a_1) : a_1, \dots, a_h \in \{0,1\}\} \tag{10}$$

in binary notation.

There are $h+1$ ways of partitioning $[0,1)^2$ into congruent rectangles of area 2^{-h}. These will give rectangles with side lengths 2^{-h_1} and 2^{-h_2}, where the integers $h_1, h_2 \geqslant 0$ satisfy $h_1 + h_2 = h$. If we use the convention that all rectangles are closed on the bottom and left edges and open on the top and right edges, then each rectangle arising from any such partition contains precisely one point of $\mathcal{P}_h$.

The van der Corput point sets have nice periodicity properties that are very useful in our attempt to establish Davenport's theorem. This can be easily illustrated by studying $\mathcal{P}_5$ in Figure 3.

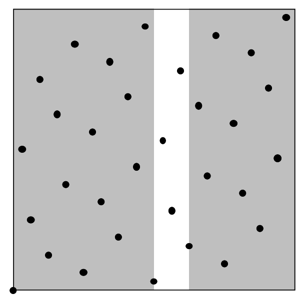

Fig.3. The van der Corput set $\mathcal{P}_5$ with the rectangle $[\frac{1}{2}, \frac{5}{8}) \times [0,1)$ highlighted

The vertical distribution of the points of $\mathcal{P}_5$ within the rectangle $[\frac{1}{2}, \frac{5}{8}) \times [0,1)$ is periodic.

Let $(x_1, x_2) \in [0,1]^2$ be given. Suppose first of all that $x_1 \in 2^{-h}\mathbb{Z}$. Then the periodicity property of the van der Corput set $\mathcal{P}_h$ leads to an expression of the form

$$D[\mathcal{P}_h; B(x_1,x_2)] = \sum_{i=1}^{h}{}^{*} \left(\alpha_i - \Psi\left(\frac{x_2 + \beta_i}{2^{i-h}} \right) \right), \tag{11}$$

where $*$ indicates that some terms are not present, and that the summation depends on x_1. In fact, the restriction $x_1 \in 2^{-h}\mathbb{Z}$ can be removed at the expense of an error $O(1)$ in (11). This error is insignificant.

The functions Ψ form a quasi-orthogonal system with respect to the variable x_2. Without the terms α_i, which arise from the hypothesis that $B(x_1,x_2)$ is anchored at the origin, Davenport's theorem will follow on squaring the expression (11) and integrating first with respect to x_2 and then trivially with respect to x_1.

Unfortunately, in Halton and Zaremba [12], it is shown that the term

$$\sum_{i=1}^{h}{}^{*} \sum_{j=1}^{h}{}^{*} \alpha_i \alpha_j$$

leads to the estimate

$$\|D[\mathcal{P}_h]\|_2^2 = 2^{-6}h^2 + O(h), \tag{12}$$

so that $\mathcal{P}_h$ fails to give Davenport's theorem.

Our sixth proof of Davenport's theorem is by Proinov [16] in 1988, and uses a variant of Davenport's reflection principle. Consider the image

$$\mathcal{P}_h' = \{(p_1, 1-p_2) : (p_1,p_2) \in \mathcal{P}_h\}$$

of the van der Corput set $\mathcal{P}_h$ under reflection across the line $x_2 = \frac{1}{2}$. Then apart from a negligible error of $O(1)$, we have an expression of the form

$$D[\mathcal{P}_h'; B(x_1,x_2)] = \sum_{i=1}^{h}{}^{*} \left(-\alpha_i - \Psi\left(\frac{x_2+\gamma_i}{2^{i-h}}\right)\right).$$

Combining this with (11), we conclude that

$$D[\mathcal{P}_h \cup \mathcal{P}_h'; B(x_1,x_2)] = -\sum_{i=1}^{h}{}^{*} \left(\Psi\left(\frac{x_2+\gamma_i}{2^{i-h}}\right) + \Psi\left(\frac{x_2+\beta_i}{2^{i-h}}\right)\right),$$

a sum of quasi-orthogonal functions in the variable x_2. Davenport's theorem now follows on squaring this expression and integrating first with respect to x_2 and then trivially with respect to x_1.

Our third proof of Davenport's theorem arises from the work of Roth [19] on the generalization of Davenport's theorem to arbitrary dimensions, and builds on his probabilistic technique developed in [18] and described in the last section. We now consider vertical translation modulo 1. For every $t \in [0,1]$, consider the image

$$\mathcal{P}_h(t) = \mathcal{P}_h + (0,t)$$

of the van der Corput set $\mathcal{P}_h$ under translation modulo $[0,1)^2$. Then, apart from a negligible error of $O(1)$, the analog of the expression (11) is now given by

$$D[\mathcal{P}_h(t); B(x_1,x_2)] = \sum_{i=1}^{h}{}^{*} \left(\Psi\left(\frac{z_i+t}{2^{i-h}}\right) - \Psi\left(\frac{w_i+t}{2^{i-h}}\right)\right),$$

a sum of quasi-orthogonal functions in the probabilistic variable t. Squaring this expression and integrating first with respect to t and then trivially with respect to x_1 and x_2, we obtain an average version of Davenport's theorem over the parameter t. Davenport's theorem is thus realized for some value of $t \in [0,1]$.

Recently, this probabilistic approach of Roth has been derandomized through the work of Bilyk [5]. This constitutes our ninth proof.

The starting point is to obtain a better understanding of the estimate (12). For convenience, we introduce the notation $N = 2^h$ and $\mathbf{p} = (p_1,p_2) \in \mathcal{P}_h$. Then it is easy to deduce that

$$\sum_{\mathbf{p}\in\mathcal{P}_h} p_1p_2 = \frac{N}{4} - \frac{1}{2} + \frac{1}{4N} + \frac{h}{8}. \tag{13}$$

This leads to the identity

$$\int_{[0,1)^2} D[\mathcal{P}_h; B(\mathbf{x})]\,\mathrm{d}\mathbf{x} = \sum_{\mathbf{p}\in\mathcal{P}_h} p_1p_2 - \frac{N}{4} + 1 = \frac{h}{8} + \frac{1}{2} + \frac{1}{4N}.$$

Clearly

$$\left(\int_{[0,1)^2} |D[\mathcal{P}_h;B(\mathbf{x})]|^2\,\mathrm{d}\mathbf{x}\right)^{1/2} \geqslant \int_{[0,1)^2} |D[\mathcal{P}_h;B(\mathbf{x})]|\,\mathrm{d}\mathbf{x}$$
$$\geqslant \left|\int_{[0,1)^2} D[\mathcal{P}_h;B(\mathbf{x})]\,\mathrm{d}\mathbf{x}\right| \geqslant \frac{h}{8}.$$

Thus the term $h/8$ in (13) is solely responsible for the term $2^{-6}h^2$ in (12).

Recall the definition (10) of the van der Corput set $\mathcal{P}_h$, and now view the digits $a_1,\ldots,a_h$ as independent random variables, with values 0 or 1 equally likely. Then

$$\mathbb{E}(a_ia_j) = \begin{cases} \mathbb{E}(a_i)\mathbb{E}(a_j) = \frac{1}{4}, & \text{if } i \neq j;\\ \mathbb{E}(a_i^2) = \mathbb{E}(a_i) = \frac{1}{2}, & \text{if } i = j.\end{cases}$$

We now view the expression

$$\frac{1}{2^h}\sum_{\mathbf{p}\in\mathcal{P}_h} p_1p_2$$

as the expectation of the value of p_1p_2. Note that

$$\sum_{\mathbf{p}\in\mathcal{P}_h} p_1p_2 = \sum_{i=1}^{h}\sum_{j=1}^{h}\frac{1}{4}2^{j-i-1} + \sum_{i=1}^{h}\frac{1}{4}2^{-1} = \sum_{i=1}^{h}\sum_{j=1}^{h}\frac{1}{4}2^{j-i-1} + \frac{h}{8}.$$

For simplicity, let us assume that the integer h is even. We now modify the van der Corpot set $\mathcal{P}_h$ to a set $\widetilde{\mathcal{P}}_h$ by shifting half of the digits in second coordinate. More precisely, write

$$\widetilde{\mathcal{P}}_h = \{(0.a_1\ldots a_h, 0.\widetilde{a}_h\ldots\widetilde{a}_1) : a_1,\ldots,a_h \in \{0,1\}\},$$

where

$$\widetilde{a}_i = \begin{cases} a_i, & \text{if } i \text{ is odd};\\ 1-a_i, & \text{if } i \text{ is even}.\end{cases}$$

It follows that

$$\mathbb{E}(a_i\widetilde{a}_j) = \begin{cases} \mathbb{E}(a_i)\mathbb{E}(\widetilde{a}_j) = \frac{1}{4}, & \text{if } i \neq j;\\ \mathbb{E}(a_i^2) = \mathbb{E}(a_i) = \frac{1}{2}, & \text{if } i = j \text{ is odd};\\ \mathbb{E}(a_i\widetilde{a}_i) = \mathbb{E}(a_i(1-a_i)) = 0, & \text{if } i = j \text{ is even}.\end{cases}$$

We thus conclude that

$$\int_{[0,1)^2} D[\widetilde{\mathcal{P}}_h;B(\mathbf{x})]\,\mathrm{d}\mathbf{x} = \frac{1}{2} + \frac{1}{4N}.$$

Using this as motivation, Bilyk can find a specific value $t^* \in [0,1]$ such that the translated van der Corput set $\mathcal{P}_h(t^*) = \mathcal{P}_h + (0,t^*)$ modulo $[0,1)^2$ gives a proof of Davenport's theorem. We have $t^* = 1 - 2^{-h}k$, where the integer k is given in binary notation in the form

$$k = \underbrace{0\ldots0}_{h_0}\underbrace{00001111\ldots00001111}_{h_1}\underbrace{000111\ldots000111}_{h_2},$$

where

$$h_0 < 568, \quad 8 \mid h_1, \quad 6 \mid h_2, \quad h_0+h_1+h_2 = h \quad \text{and} \quad \frac{h_2}{h_1} = \frac{54}{17}.$$

4. Group Structure and Orthogonality

Our fifth proof of Davenport's theorem by Chen [7] is motivated by the existence of sets other than van der Corput sets which have nice distribution properties but do not possess the necessary periodicity properties. Thus Roth's probabilistic technique by translation cannot be applied to these sets. We therefore wish to find an alternative to this where we can dispense with periodicity.

Our study here is motivated by a very simple observation on the van der Corput sets. Let us return to Figure 3, and look at the rectangle $[\frac{1}{2}, \frac{5}{8}) \times [0, 1)$. Take the left half of the white strip and translate the two points upwards by $\frac{1}{4}$ modulo 1. Alternatively, replace the left half of white strip, together with its two points, by the right half of the white strip, together with its two points. Clearly we have the same effect. Note now that the latter is achieved by shifting the digit a_3 in the first coordinates of the points. This suggests that the Roth translation variable t can be replaced by the collection of all possible digit shifts of the first coordinates of the points of $\mathcal{P}_h$. At the time of its discovery, this technique only serves to give alternative proofs of Davenport's theorem and its generalizations. Later, it serves an extremely important role as a catalyst to the solution of the explicit construction problem in generalizations of Davenport's theorem to higher dimensions, where explicit point sets are not known until 2002. We now proceed to describe these new ideas, and shall return to digit shifts at the end of this survey.

Our eighth proof of Davenport's theorem is due to Chen and Skriganov [8, 9] in 2002.

Consider the van der Corput set $\mathcal{P}_h$ given by (10). Let $\oplus$ denote 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 these groups are the Walsh functions, with values ± 1. It is well known that the collection of Walsh functions forms an orthonormal basis for $L^2([0,1])$. This suggests the use of Fourier–Walsh analysis and series.

Next, we generalize the van der Corput sets to base p, where p is a prime. These more general van der Corput sets are of the form

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

in p-ary notation. Let $\oplus$ denote coordinatewise and digitwise addition modulo p. Then $(\mathcal{P}_h, \oplus)$ is a group isomorphic to the additive group $\mathbb{Z}_p^h$. The characters of these groups are the base p Walsh functions, with values p-th roots of unity. As in the binary case, the collection of base p Walsh functions forms an orthonormal basis for $L^2([0,1])$. This suggests the use of base p Fourier–Walsh analysis and series. For Davenport's theorem, we use 2-dimensional base p Fourier–Walsh series. It can be shown that there are explicitly constructed relatives $\mathcal{P}_h^*$ of $\mathcal{P}_h$ such that good approximations of the discrepancy function $D[\mathcal{P}_h^*; B(x_1, x_2)]$ can be expressed as a 2-dimensional base p Fourier–Walsh series with orthogonal coefficients, provided that the prime p is large enough, for instance, $p \geqslant 11$. Thus orthogonality leads to the best upper bounds for $\|D[\mathcal{P}_h^*]\|_2$, and therefore no probabilistic considerations are required.

Suppose now that we take $p = 2$ and the Fourier–Walsh coefficients are not orthogonal or even quasi-orthogonal. We return to our fifth proof and use digit shifts. For the van der Corput set $\mathcal{P}_h$ given by (10), consider digit shifts

$$(\mathbf{b}, \mathbf{c}) = (b_1, \ldots, b_h, c_h, \ldots, c_1) \in \mathbb{Z}_2^{2h},$$

and denote by $\mathcal{P}_h^{(\mathbf{b},\mathbf{c})}$ the set of points obtained from $\mathcal{P}_h$ where the digits a_i in the first coordinates of the points are replaced by $a_i \oplus b_i$ and the digits a_i in the second coordinates of the points are replaced by $a_i \oplus c_i$. The digit shifts form a group isomorphic to $\mathbb{Z}_2^{2h}$, and the characters are 2-dimensional Walsh functions, with orthogonality conditions of the form

$$\sum_{\mathbf{t}\in\mathbb{Z}_2^{2h}} W_{\mathbf{l}'}(\mathbf{t})W_{\mathbf{l}''}(\mathbf{t}) = \begin{cases} 4^h, & \text{if } \mathbf{l}'=\mathbf{l}''; \\ 0, & \text{otherwise.} \end{cases}$$

Through the digit shifts, we recover some orthogonality through the orthogonality of the Walsh functions.

References

[1] J. Beck, From probabilistic diophantine approximation to quadratic fields, in *Random and Quasi-Random Point Sets*, eds. P. Hellekalek and G. Larcher, Lecture Notes in Statistics, Vol. 138 (Springer, 1998), pp. 1–48.

[2] J. Beck, Randomness in lattice point problems, *Discrete Math.* **229** (2001) 29–55.

[3] J. Beck, Lattice point problems: crossroads of number theory, probability theory and Fourier analysis, in *Fourier Analysis and Convexity*, eds. L. Brandolini, L. Colzani, A. Iosevich and G. Travaglini, Applied and Numerical Harmonic Analysis (Birkhäuser, 2004), pp. 1–35.

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

[5] D. Bilyk, Cyclic shifts of the van der Corput set, *Proc. Amer. Math. Soc.* **137** (2009) 2591–2600.

[6] D. Bilyk, The L^2 discrepancy of irrational lattices, in *Monte Carlo and Quasi-Monte Carlo Methods 2012*, eds. J. Dick, F.Y. Kuo, G.W. Peters and I.H. Sloan, Springer Proceedings in Mathematics and Statistics, Vol. 65 (Springer, 2013), pp. 289–296.

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

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

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

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

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

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

[13] G.H. Hardy and J.E. Littlewood, The lattice points of a right angled triangle I, *Proc. London Math. Soc.* **3** (1920) 15–36.

[14] G.H. Hardy and J.E. Littlewood, The lattice points of a right angled triangle II, *Abh. Math. Sem. Hamburg* **1** (1922) 212–249.

[15] V.F. Lev, Translations of nets and relationship between supreme and L^κ-discrepancies, *Acta Math. Hungar.* **70** (1996) 1–12.

[16] P.D. Proinov, Symmetrization of the van der Corput generalised sequences, *Proc. Japan Acad. Ser. A Math. Sci.* **64** (1988) 159–162.

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

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

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

[20] V.T. Sós and S.K. Zaremba, The mean-squares discrepancies of some two-dimensional lattices, *Studia Sci. Math. Hungar.* **14** (1979) 255–271.

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