Discrepancy theory and harmonic analysis

Dmitriy Bilyk

Abstract. In the present survey we discuss various applications of methods and ideas of harmonic analysis in problems of geometric discrepancy theory and irregularities of distribution. A great number of analytic tools (exponential sums, Fourier series, Fourier transform, orthogonal expansions and wavelets, Riesz products, Littlewood–Paley theory, Carleson's theorem) have found applications in this area. Some of the methods have been used since the birth of the subject, while the more modern ideas are still paving their way into the field. We illustrate their applications by considering several standard topics in uniform distribution theory: Weyl's criterion, metric estimates for the discrepancy of sequences, discrepancy with respect to balls and rotated cubes, lower bounds for the discrepancy function. This exposition of Fourier-analytic techniques in discrepancy theory is intended for a broad mathematical readership.

Keywords. Uniform distribution, discrepancy, harmonic analysis.

AMS classification. 11K38, 42-02.

1 Introduction

Traditionally harmonic analysis played a pivotal role in the development of discrepancy theory. Nearly every important result used analytic techniques and objects. Exponential sums arise already in the famous Weyl's criterion [48] and the Erdős–Turán inequality [23]; lower bounds for the discrepancy with respect to balls or rotated boxes (independently, Beck [5] and Montgomery [34]) employ Fourier analysis; Roth's celebrated lower bound for the L^2-discrepancy relied on wavelet expansions [37]; Halász's proofs of the endpoint (L^1 and L^∞) lower bounds for the discrepancy function in dimension 2 [26] cleverly used Riesz products much in the same way as in Sidon's theorem on lacunary Fourier series [42]; Walsh analysis is a common tool in the study of the distributional properties of digital nets [18, 22] – this sample convincingly demonstrates the impact of harmonic analysis on discrepancy theory.

While the aforementioned methods have been exploited by experts in uniform distribution for a long time, some of the more modern methods (e.g., Littlewood–Paley theory) have been overlooked and only gained acclaim in discrepancy theory in the recent years. They have been heavily used in some of the important latest achievements (such as the lower bounds of the star-discrepancy in higher dimensions [10, 11] and deterministic constructions of distributions with optimal L^p discrepancy [43]).

This work is supported by the NSF grant DMS-1101519.

In the present expository paper we attempt to survey some of the applications of harmonic analysis to discrepancy theory and irregularities of distribution, although any such endeavor is doomed to be incomplete, and the choice of topics is rather eclectic.

While there is a number of good books on discrepancy theory [7, 15, 22, 29, 33] and nice surveys on subjects closely related to the present topic [9, 17, 31, 35, 47], in this article the author has tried to collect in one place and crystallize some of the most important ideas involving applications of harmonic analysis to discrepancy, thus creating a good first introduction to the subject. The exposition is very concise and mostly based on some of the colloquium talks that the author has given in the recent years, therefore it can be comprehended in a matter of a couple of hours.

The paper has three main parts (each at most 5 pages long). Section 2 deals with the relations between discrepancy and exponential sums (Weyl's criterion, Erdős–Turán inequality, metric results). Useful further references for this section include [2, 29, 35]. Section 3 explains some ideas of Fourier analysis in application to discrepancy. The arguments outlined in this section (as well as further results) are presented in full detail in, e.g., [6, 7, 15, 33, 35]. Finally, the last section discusses the use of dyadic analysis and wavelets in discrepancy function estimates – the reader is referred to [9] for a very detailed survey of this topic.

We do not assume any special prerequisites on the reader's part – familiarity with basic concepts of harmonic analysis (Fourier trasform/series, orthogonal bases etc.) would be sufficient. We often use the symbol $\lesssim$ to mean "less than a constant multiple of", where the implicit constant may depend on the dimension and other parameters, but *not on the number of points* N.

2 Exponential sums

Harmonic analysis, in one way or another, has been present in the theory of uniform distribution from its very dawn. In the very paper, where the concept of uniform distribution was introduced, it was immediately connected to the notion of exponential sums [48]. This statement became known as the Weyl criterion.

We recall the definition of uniform distribution introduced by Hermann Weyl. A sequence $(\omega_n) \subset [0,1]$ is called *uniformly distributed* if for any subinterval $I \subset [0,1]$ we have

$$\lim_{N\to\infty} \frac{\#\{n \leq N : \omega_n \in I\}}{N} = |I|,$$

i.e. the proportion of points in I is asymptotically equal to the length of I. One can rewrite it to say that the relation $\lim\limits_{N\to\infty} \frac{1}{N} \sum\limits_{n=1}^{N} f(\omega_n) = \int_0^1 f(x)dx$ holds for all indicators of intervals $f = \mathbf{1}_I$, which may be easily extended to all continuous 1-periodic functions f. The celebrated Weyl criterion then immediately follows from the Weierstrass theorem on approximating continuous functions by trigonometric polynomials.

Theorem 2.1 (Weyl Criterion). *A sequence* $(\omega_n) \subset [0,1]$ *is uniformly distributed in* $[0,1]$ *if and only if for all* $k \in \mathbb{Z}$, $k \neq 0$*:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k \omega_n} = 0. \tag{2.1}$$

This criterion has an interesting and intuitive geometric explanation: the expression $\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i \omega_n}$ is the center of mass of the first N points of (ω_n) put on the unit circle. If (ω_n) is uniformly distributed in $[0,1]$, this center of mass should be close to zero. However, this is not sufficient – e.g., it holds also if the points are concentrated in the vertices of a regular polygon. This obstacle is overcome by considering all integer dilations $k \neq 0$.

Among the countless applications of the Weyl criterion, two very important ones are the most immediate. First, it implies that the sequence $\{n\theta\}$ is uniformly distributed in $[0,1]$ if and only if θ is irrational (here $\{x\}$ stands for the fractional part of x). Second, it can be used to deduce the fact that for any subsequence (n_k) of integers, the sequence $\{n_k\theta\}$ is uniformly distributed for *a.e.* θ. The latter may be viewed as a less precise generalization of the former fact.

The concept of uniform distribution may be quantified using the notion of discrepancy. For a sequence $\omega = (\omega_n)_{n=1}^{\infty}$ and an interval $I \subset [0,1]$ consider the quantity $\Delta_{N,I} = \sharp\{\omega_n : \omega_n \in I; n \leq N\} - N|I|$. The discrepancy of ω is defined as

$$D_N = \sup_{I\subset[0,1]} \left|\Delta_{N,I}\right|, \tag{2.2}$$

and it is easy to show that a sequence $(\omega_n)_{n=1}^{\infty}$ is uniformly distributed in $[0,1]$ if and only if $\lim\limits_{N\to\infty} \dfrac{D_N}{N} = 0$.

This statement together with the Weyl criterion suggests that discrepancy D_N and the exponential sums (2.1) should exhibit similar behavior.

Indeed, the inequality $\left|\sum\limits_{n=1}^{N} e^{2\pi i \omega_n}\right| \lesssim D_N(\omega)$ is quite simple and can be viewed, in particular, as a partial case of Koksma's inequality connecting discrepancy and numerical integration

$$\left|\int_0^1 f(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(\omega_n)\right| \lesssim \frac{1}{N}\|f'\|_1 \cdot D_N(\omega)$$

with $f(x) = e^{2\pi i x}$. A relation in the opposite direction is given by the Erdős–Turán inequality, which we state in a slightly simplified form.

Theorem 2.2 (Erdős–Turán [23]). *For any sequence* $\omega \subset [0,1]$ *and for any* N, $m \in \mathbb{N}$

$$D_N(\omega) \lesssim \frac{N}{m} + \sum_{k=1}^{m} \frac{1}{k}\left|\sum_{n=1}^{N} e^{2\pi i k \omega_n}\right|.$$

It is folklore knowledge in the subject that the Erdős–Turán inequality *misses the optimal estimates by a logarithm*. For example, for a *badly approximable* irrational number θ, Erdős–Turán yields $D_N(\{n\theta\}) \lesssim \log^2 N$, while in fact the sharp bound is $D_N(\{n\theta\}) \lesssim \log N$.

We now return to the discussion of sequences $\{n_k\theta\}$. Since, for any $(n_k) \subset \mathbb{N}$, the sequence $\{n_k\theta\}$ is uniformly distributed for almost every θ, it is completely natural to pose the question about the behavior of the discrepancy of such sequences.

In 1981 Baker [3] proved that for any $(n_k) \subset \mathbb{N}$

$$D_N(\{n_k\theta\}) \lesssim \sqrt{N} \log^{\frac{3}{2}+\varepsilon} N \quad \text{for } a.e.\ \theta. \tag{2.3}$$

Berkes and Philipp [8] constructed an example demonstrating that the power of logarithm cannot be better than $\frac{1}{2}$. It is conjectured that $\frac{1}{2} + \varepsilon$ is indeed the sharp exponent in this estimate.

This conjecture is strongly supported by the fact that it holds for the exponential sums: for any $(n_k) \subset \mathbb{N}$ we have

$$\left| \sum_{k=1}^{N} e^{2\pi i n_k \theta} \right| \lesssim \sqrt{N} \log^{\frac{1}{2}+\varepsilon} N \quad \text{for } a.e.\ \theta. \tag{2.4}$$

Inequality (2.3) is, in fact, obtained using this estimate together with Erdős–Turán: hence an extra logarithm. The proof of (2.4) is the very reason I decided to include this topic into the discussion. It demonstrates the use of very non-trivial harmonic analysis in discrepancy estimates. The argument relies on one of the deepest and most important results of the twentieth-century harmonic analysis – Carleson's theorem on the almost everywhere convergence of the Fourier series of L^2 functions [14]. Even though one only needs this theorem in a very special, innocently looking case, the author is not aware of any simplifications.

Carleson's theorem states that the maximal operator of partial Fourier sums is bounded in L^2, i.e. $\left\| \sup_N \left| \sum_{-N}^{N} \widehat{f}_k e^{2\pi i k x} \right| \right\|_2 \lesssim \|f\|_2$ for each $f \in L^2[0,1]$, which in turn implies that the Fourier series of f converges *a.e.*

Applying it to the function $f(\theta) = \sum_{k=1}^{N} e^{2\pi i n_k \theta}$ we find that

$$\left\| \sup_{M \le N} \left| \sum_{k=1}^{M} e^{2\pi i n_k \theta} \right| \right\|_2 \lesssim \sqrt{N}.$$

Chebyshev's inequality implies that

$$\mu\left(\sup_{M \le 2^m} \left| \sum_{k=1}^{M} e^{2\pi i n_k \theta} \right| > 2^{m/2} m^{\frac{1}{2}+\varepsilon} \right) \lesssim \frac{1}{m^{1+2\varepsilon}},$$

where μ is the Lebesgue measure. (Originally Carleson proved a weak L^2 estimate for the maximal operator. For our purposes, however, the weak bound would suffice.)

Since $\sum \frac{1}{m^{1+2\varepsilon}} < \infty$, Borel–Cantelli lemma implies that the measure of the set of θ for which $\sup_{M\leq 2^m} \left| \sum_{k=1}^{M} e^{2\pi i n_k \theta} \right| > 2^{m/2} m^{\frac{1}{2}+\varepsilon}$ for infinitely many values of m is zero, which immediately yields (2.4). □

Remark. Incidentally, while completing this manuscript, the author learned about the paper of Aistleitner [2] in the same volume which strongly overlaps with and expands the topics discussed in this section. The reader is enthusiastically directed towards this article for further reading – it contains many interesting results in discrepancy and metric number theory, including some recent developments.

3 Fourier analysis methods

In this section we shall illustrate the use of Fourier methods in discrepancy problems. It has long been known that the presence of curvature or rotational invariance crucially influences the behavior of the Fourier transform, see e.g. [44]. Hence ideas of Fourier analysis naturally arise in discrepancy questions related to such geometrical situations.

We shall discuss two somewhat similar cases – discrepancy with respect to balls (disks) and with respect to rectangles rotated in arbitrary directions. For simplicity we restrict ourselves to the two-dimensional setting.

3.1 Rotated rectangles

We consider a point distribution $\mathcal{P}_N \subset [0,1]^2$ with $\#\mathcal{P}_N = N$, and let $\mathcal{A}$ be a translation invariant collection of sets in $\mathbb{R}^2$, e.g. the set of all rotated rectangles.

We can view the discrepancy of the distribution $\mathcal{P}_N$ with respect to sets in $\mathcal{A}$ as a measure defined on $\mathcal{A}$. Namely for $A \in \mathcal{A}$

$$\mathcal{D}(A) = D(\mathcal{P}_N, A) = \sum_{p\in\mathcal{P}_N} \delta_p(A) - N \cdot \text{vol}\left(A \cap [0,1]^2\right),$$

where δ_p is a Dirac delta mass concentrated at the point p. We are then interested in the quantity $D(\mathcal{P}_N, \mathcal{A}) = \sup_{A\in\mathcal{A}} |\mathcal{D}(A)|$, however we look at the L^2 averages rather than the supremum. Due to translation invariance, for a fixed set A we see that

$$\Delta_A(x) := \mathcal{D}(A+x) = \int_{\mathbb{R}^2} \mathbf{1}_{A+x}(y) d\mathcal{D}(y) = \left(\mathbf{1}_A * \mathcal{D}\right)(x),$$

i.e. the discrepancy with respect to the translate of A is a convolution of the characteristic function of A with the discrepancy measure (we made a minor technical assumption that the set A is symmetric, i.e. $-A = A$). This fact suggests the use of the Fourier transform, which "diagonalizes" the convolution operator: $\widehat{\Delta}_A(\xi) = \widehat{\mathbf{1}}_A(\xi) \cdot \widehat{D}(\xi)$. Plancherel's theorem implies that

$$\|\Delta_A\|_2^2 = \|\widehat{\Delta}_A\|_2^2 = \int_{\mathbb{R}^2} |\widehat{\mathbf{1}}_A(\xi)|^2 \cdot |\widehat{\mathcal{D}}(\xi)|^2 d\xi.$$

In order to show that this integral is large for any choice of $\mathcal{P}_N$, most of the work shall be done on the factor $|\widehat{\mathbf{1}}_A(\xi)|$ which depends on the geometry of A, since we have almost no control over $|\widehat{\mathcal{D}}(\xi)|^2$.

Let $A_r = \left[-\frac{r}{2}, \frac{r}{2}\right]^2$, $r \leq 1$ be the indicator of an axis-parallel square. The proof starts with a *trivial observation* that if $r_0 \approx \frac{1}{2\sqrt{N}}$, then $\|\Delta_{A_{r_0}}\|_2^2 \gtrsim 1$. Indeed, in this case the counting part of the discrepancy is integer, while the area term yields $Nr^2 \approx \frac{1}{4}$ for most squares, therefore the difference is at least a constant.

If only we had the pointwise linear growth bound $|\widehat{\mathbf{1}}_{A_r}(\xi)|^2 \gtrsim \frac{r}{r_0} \cdot |\widehat{\mathbf{1}}_{A_{r_0}}(\xi)|^2$ for $r > r_0$, we would immediately blow up the trivial bound to the desired result by taking $r_0 \approx \frac{1}{2\sqrt{N}}$ and $r \approx 1$:

$$\|\Delta_{A_1}\|_2 \gtrsim \sqrt{\frac{1}{1/2\sqrt{N}}}\|\Delta_{A_{1/2\sqrt{N}}}\|_2 \approx N^{1/4}.$$

However, the linear growth doesn't hold, since $|\widehat{\mathbf{1}}_{A_r}(\xi)|^2 = \left(\frac{\sin(\pi\xi_1 r)}{\pi\xi_1}\right)^2 \cdot \left(\frac{\sin(\pi\xi_2 r)}{\pi\xi_2}\right)^2$.

Intuitively, in the expression above one can observe the dependence on r only very close to the coordinate axes, i.e. where either $\pi\xi_1 r$ or $\pi\xi_2 r$ is very small. In addition, the expression turns into zero periodically along the axes, thus annihilating the dependence on r. The graph for $r = 0.5$ depicted below illustrates exactly this behavior.

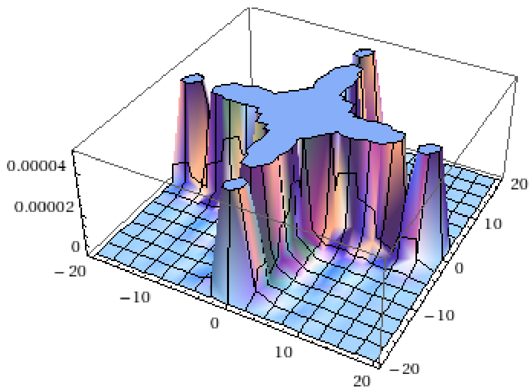

Figure 1. $|\widehat{\mathbf{1}}_{A_r}(\xi)|^2$ for $r = 0.5$

The remedy suggests itself: one should average over dilations to eliminate the zeros on the axes and then average over rotations in order to smear the dependence on r from the axes to the whole plane. More precisely, if $A_{r,\theta}$ denotes the rotation of A_r by θ, then a straightforward but technical calculation shows that

$$\omega_r(\xi) := \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{r}\int_r^{2r} |\widehat{\mathbf{1}}_{A_{\rho,\theta}}(\xi)|^2\, d\rho\, d\theta \approx \min\left\{r^4, \frac{r}{|\xi|^3}\right\},$$

therefore $\frac{\omega_r(\xi)}{\omega_{r_0}(\xi)} \gtrsim \frac{r}{r_0}$ and we have the desired linear growth.

This argument, due to Beck [5], proves the two-dimensional variant of the following theorem (higher-dimensional extensions are not much harder):

Theorem 3.1 (Beck 1988 [6]). *Let $\mathcal{A}$ be the family of arbitrarily rotated cubes in $\mathbb{R}^d$. For any point distribution $\mathcal{P}_N \subset [0,1]^d$ with $\#\mathcal{P}_N = N$ we have*

$$D(\mathcal{P}_N, \mathcal{A}) = \sup_{A \in \mathcal{A}} |\mathcal{D}(A)| \gtrsim N^{\frac{1}{2} - \frac{1}{2d}}.$$

3.2 The lower bound for circles

In this case, for the sake of simplicity, we shall treat the unit square $[0,1]^2$ as a torus $\mathbb{T}^2$. We present a very similar, albeit different in some details, method developed by Montgomery [34]. Start with a set $\mathcal{P}_N \subset \mathbb{T}^2$, with $\#\mathcal{P}_N = N$, and let C_r denote a circle of radius r centered at the origin. As before, if we write the discrepancy measure $\mathcal{D} = \sum_{p \in \mathcal{P}_N} \delta_p - \mathrm{vol}_d$ and consider the discrepancy of $\mathcal{P}_N$ with respect to the translate of the circle $D_r(q) = \#\{\mathcal{P}_N \cap (C_r + q)\} - N\pi r^2 = (\mathbf{1}_{C_r} * \mathcal{D})(q)$ for $q \in \mathbb{T}^2$, Parseval's identity for the Fourier series again yields the separation of factors depending on the geometries of C_r and $\mathcal{P}_N$:

$$\|D_r\|_2^2 = \sum_{\mathbf{t} \in \mathbb{Z}^2} |\widehat{\mathbf{1}}_{C_r}(\mathbf{t})|^2 \cdot |\widehat{\mathcal{D}}(\mathbf{t})|^2.$$

The first factor here is classical: $\widehat{\mathbf{1}}_{C_r}(\mathbf{t}) = \frac{r}{|\mathbf{t}|} J_1(2\pi|\mathbf{t}|r)$, where J_1 is the Bessel function of the first kind whose asymptotics is well-known

$$J_1(x) = \sqrt{\frac{2}{\pi x}} \cos(x - 3\pi/4) + \mathcal{O}(x^{-3/2}).$$

While this factor has the decay of the right order, an obstacle to obtaining a lower bound is the same as in the previous case – the leading term vanishes infinitely often. This can be dealt with by averaging over dilations just as before. However, a much simpler averaging procedure happens to work in this situation: one may average over only two points: $\frac{1}{2}$ and $\frac{1}{4}$. It turns out that $J_1^2(x) + J_1^2(2x) \gtrsim \frac{1}{x}$. (We skip the calculations and only present the graphs below, see Figure 2.) This implies, in particular,

$$|\widehat{\mathbf{1}}_{C_{1/4}}(\mathbf{t})|^2 + |\widehat{\mathbf{1}}_{C_{1/2}}(\mathbf{t})|^2 \gtrsim \frac{1}{|\mathbf{t}|^3}. \tag{3.1}$$

The Fourier coefficients of the discrepancy measure may be dealt with as follows. It is easy to see that $\widehat{\mathcal{D}}(\mathbf{0}) = 0$ and $\widehat{\mathcal{D}}(\mathbf{t}) = \sum_{p \in \mathcal{P}_N} e^{-2\pi i <p,\mathbf{t}>}$ for $\mathbf{t} \neq 0$. Simple algebraic manipulations then lead to

$$\begin{aligned}\sum_{|t_1|,|t_2| \leq M} \left| \sum_{p \in \mathcal{P}_N} e^{-2\pi i <p,\mathbf{t}>} \right|^2 &\geq \sum_{p,q \in \mathcal{P}_N} F_M(p_1 - q_1) F_M(p_2 - q_2) \\ &\geq N \cdot F_M^2(0) = M^2 N,\end{aligned}$$

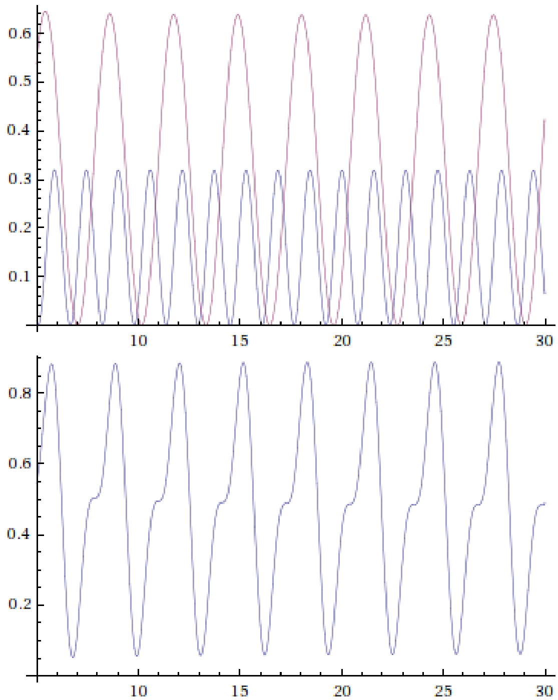

Figure 2. a) $x \cdot J_1^2(x)$ and $x \cdot J_1^2(2x)$; b) $x \cdot (J_1^2(x) + J_1^2(2x))$

where F_M is the Fejér kernel: $F_M(x) = \sum\limits_{m=-M}^{M} \Big(1 - \dfrac{|m|}{M}\Big) e^{-2\pi i m x} = \dfrac{1}{M}\left(\dfrac{\sin \pi M x}{\sin \pi x}\right)^2$.
The second inequality above is obtained by throwing away "non-diagonal" terms. Since for $\mathbf{t} = \mathbf{0}$ we have $\sum_{p\in\mathcal{P}_N} e^{-2\pi i <p,\mathbf{t}>} = N$, this implies that

$$\sum_{|t_1|,|t_2|\le M} |\widehat{\mathcal{D}}(\mathbf{t})|^2 \gtrsim M^2 N - N^2.$$

Putting this inequality together with the estimate $|\widehat{\mathbf{1}}_{C_{1/4}}(\mathbf{t})|^2 + |\widehat{\mathbf{1}}_{C_{1/2}}(\mathbf{t})|^2 \gtrsim \frac{1}{|\mathbf{t}|^3}$ and choosing $M \approx \sqrt{N}$ one arrives at

$$\|D_1\|_2^2 + \|D_{1/2}\|_2^2 \gtrsim \frac{1}{M^3}(M^2 N - N^2) \gtrsim N^{\frac{1}{2}},$$

which proves the following theorem due to Montgomery [34]:

Theorem 3.2. *For any distribution* $\mathcal{P}_N \subset \mathbb{T}^2$, *with* $\#\mathcal{P}_N = N$ *there exists a circle* C *of radius either* $\frac{1}{4}$ *or* $\frac{1}{2}$ *such that the discrepancy of* $\mathcal{P}_N$ *with respect to* C *satisfies*

$$D(\mathcal{P}_N, C) \gtrsim N^{1/4}.$$

3.3 Further remarks

While perfectly clean calculations are possible only in such special cases as disks or rotated boxes, more general situations can also be treated. In particular, for any closed C^1 curve with interior S of diameter at most one, if one defines $S(s,\theta)$ to be a copy of S compressed by a factor of $0 \le s \le 1$ and rotated by $\theta \in [0,2\pi)$, then it can be shown using an integration by parts argument that

$$\int_0^1 \int_0^{2\pi} \left|\widehat{\mathbf{1}}_{S(s,\theta)}(\mathbf{t})\right|^2 d\theta\, dt \approx \frac{L}{|\mathbf{t}|^3},$$

where L is the length of the curve. This is an analogue of (3.1) and similarly lets one prove that for any $\mathcal{P}_N \subset \mathbb{T}^2$ there exists a translated, scaled, and rotated copy of S with discrepancy $D(\mathcal{P}_N, S(s,\theta)+x_0) \gtrsim N^{1/4}$ (see Montgomery [35] for details).

Independently Beck showed that for any convex body $A \subset \mathbb{R}^d$, which is "not too thin" (contains a ball of radius $N^{-1/d}$), and a point distribution $\mathcal{P}_N \subset \mathbb{T}^d$, there exists A', a translated, scaled, and rotated image of A with $|D(\mathcal{P}_N, A')| \gtrsim N^{\frac{1}{2}-\frac{1}{2d}} \cdot \sqrt{\sigma(A)}$, where $\sigma(A)$ is the surface area of A. If one insists that the copy be completely contained in $[0,1]^d$, then the exponent above becomes $\frac{1}{2} - \frac{1}{2d} - \varepsilon$, see [5].

It is known that these results are almost sharp [4]: e.g., there exist distributions for which the discrepancy with respect to balls is of the order $N^{\frac{1}{2}-\frac{1}{2d}}\sqrt{\log N}$ (a random perturbation of a square lattice; $\sqrt{\log N}$ arises from the large deviation arguments), and the lower bound is indeed sharp for the L^2 average of the discrepancy over translations and dilations (in this case even a non-randomized lattice would suffice). Chen and Travaglini [19] noticed an interesting effect that in high dimensions the randomized lattice yields slightly smaller L^2-discrepancy with respect to balls than the standard lattice, while in low dimensions the latter beats the former (if $d \not\equiv 1 \bmod 4$).

The striking difference between the case of rotational invariance or curvature (rotated boxes, balls: discrepancy polynomial in N) and the absence of rotations/curvature (e.g., axis-parallel rectangles; discrepancy logarithmic in N, see the next section) was first studied by Schmidt [39]. Recently the author with Ma, Pipher, and Spencer [12, 13] studied some intermediate situations. Let $\Omega \subset [0,2\pi]$ be a set of directions and define the directional discrepancy $D_\Omega(N) = \inf_{\mathcal{P}_N} \sup_R \left|D(\mathcal{P}_N, R)\right|$, where the supremum is over rectangles R pointing in directions of Ω. Below is a corollary of more general results obtained in terms of the metric entropy properties of Ω, showing the difference between "thicker" and "thinner" rotation sets.

Theorem 3.3. *The following estimates hold:*

- *Lacunary directions, e.g.* $\Omega = \{2^{-n}\}$*:* $\quad D_\Omega(N) \lesssim \log^3 N$
- *Lacunary of order* M*, e.g.* $\Omega = \{2^{-n_1} + \ldots + 2^{-n_m}\}$*:* $\quad D_\Omega(N) \lesssim \log^{M+2} N$
- *"Superlacunary" sequence, e.g.* $\Omega = \{2^{-2^n}\}$*:* $\quad D_\Omega(N) \lesssim \log N \cdot (\log\log N)^2$
- Ω *has upper Minkowski dimension* $0 \le d < 1$*:* $\; D_\Omega(N) \lesssim N^{\frac{d}{d+1}+\varepsilon}$.

4 Dyadic harmonic analysis: discrepancy function estimates

Let $\mathcal{P}_N \subset [0,1]^d$ be an N-point subset of the unit cube. In accordance with (2.2) we define the discrepancy function of the set $\mathcal{P}_N$ as

$$D_N(x) = \sharp\{\mathcal{P}_N \cap [0,x)\} - Nx_1x_2\ldots x_d.$$

Different norms of this function give us information about equidistribution of the set $\mathcal{P}_N$ as well as the numerical integration errors for cubature formulas given by P_N in various function classes [47, 22]. The most interesting is the L^∞ norm of D_N, i.e. the "extreme" discrepancy. We shall start, however, by discussing the "average" case.

4.1 L^p-discrepancy, $1 < p < \infty$.

This situation is more or less fully understood. The following statement has been proved in a seminal paper of Roth [37] in 1954 for $p = 2$ and extended to all $p > 1$ by Schmidt in 1977 [41]: Let $1 < p < \infty$, for all $\mathcal{P}_N \subset [0,1]^d$ with $\#\mathcal{P}_N = N$ we have

$$\|D_N\|_p \gtrsim (\log N)^{\frac{d-1}{2}}. \tag{4.1}$$

This estimate has been shown to be sharp in the order of magnitude, i.e. there exist point distributions $\mathcal{P}_N \subset [0,1]^d$ with

$$\|D_N\|_p \lesssim (\log N)^{\frac{d-1}{2}}$$

(Davenport [21], 1956 ($d = 2$, $p = 2$); Roth [38], 1979 ($d \geq 3$, $p = 2$); Frolov [25], 1980 ($p > 2$, $d = 2$); Chen [16], 1983 ($p > 2$, $d \geq 3$)). In the last decade using harmonic analysis methods (e.g., Walsh analysis and Littlewood–Paley theory) Chen and Skriganov [18, 43] constructed the first non-probabilistic examples in $d \geq 3$.

Most standard proofs of discrepancy function estimates akin to (4.1) utilize *Haar functions*, which brings us to the realm of dyadic harmonic analysis and wavelets – see e.g. [36] for a nice exposition. In one dimension, we define the system of dyadic intervals $\mathcal{D}$ to consist of all intervals of the form $I = \left[k2^{-n}, (k+1)2^{-n}\right)$ for $n, k \in \mathbb{Z}_+$ with $k < 2^n$. For each such interval, an L^∞ normalized Haar function is defined as $h_I = -\mathbf{1}_{I_{\text{left}}} + \mathbf{1}_{I_{\text{right}}}$, where I_{left} and I_{right} are the left and right halves of I. It is well known that together with the constant function, this system forms an orthogonal basis of $L^2[0,1]$. In higher dimensions, for a rectangle $R = R_1 \times \ldots \times R_d \in \mathcal{D}^d$, the Haar function is defined as a tensor product: $h_R(x) = \prod_{k=1}^d h_{R_k}(x_k)$.

The main idea of the proof of (4.1), which propagated into many further results in the theory, is that the behavior of the discrepancy function is mostly determined by the portion of the Haar expansion corresponding to rectangles of volume roughly $1/N$:

$$D_N \approx \sum_{R\in\mathcal{D}^d:\, |R|\approx\frac{1}{N}} \frac{\langle D_N, h_R\rangle}{|R|} h_R \tag{4.2}$$

Heuristically, these are the scales that carry the most information about the distribution (large rectangles encode too much cancelation, while the smaller ones are too fine to "see" an N-point set). This idea is similar to the concept of hyperbolic cross approximation in approximation theory [47] (constant volume of R means that the product of frequencies of h_R is constant, i.e. frequencies live on a hyperbola $\tau_1 \cdot \ldots \cdot \tau_d \approx N$).

In particular, for $n \approx \log_2 N$, using Parseval's identity, we would get

$$\|D_N\|_2 \geq \left(\sum_{|R|=2^{-n}} \frac{|\langle D_N, h_R\rangle|^2}{|R|} \right)^{1/2} \gtrsim n^{\frac{d-1}{2}} \approx (\log N)^{\frac{d-1}{2}} \tag{4.3}$$

provided that $|\langle D_N, h_R\rangle| \gtrsim 2^{-n}$ for "many" R's (which is true if $2^n \approx 2N$, see [9]).

The factor n^{d-1} above is roughly the number of different shapes of dyadic rectangles of volume 2^{-n}, in other words $d-1$ is the number of free parameters (d dimensions minus one free parameter removed by the condition $|R| = 2^{-n}$).

Alternatively, one could prove the same bound by duality, considering the function

$$F = \sum_{\substack{R\colon |R|=2^{-n},\\ |\langle D_N, h_R\rangle| \gtrsim 2^{-n}}} \pm h_R \tag{4.4}$$

A similar argument shows that for appropriate choices of signs we have $\langle D_N, F\rangle \gtrsim n^{d-1}$, $\|F\| \approx n^{\frac{d-1}{2}}$, and hence by Cauchy–Schwartz $\|D_N\|_2 \gtrsim n^{\frac{d-1}{2}} \approx (\log N)^{\frac{d-1}{2}}$.

Both approaches can be extended from L^2 to L^p using Littlewood–Paley inequalities, which we describe here in a simplified form. Let $f = \sum_{I\in\mathcal{D}} a_I h_I$. The *dyadic* Littlewood–Paley square function of f is defined as

$$Sf = \left(\sum_{I\in\mathcal{D}} |a_I|^2 \mathbf{1}_I \right)^{1/2}.$$

The Littlewood–Paley inequalities state that its L^p norm is equivalent to that of f: for all $1 < p < \infty$, there exist A_p, $B_p > 0$ such that

$$A_p \|Sf\|_p \leq \|f\|_p \leq B_p \|Sf\|_p. \tag{4.5}$$

The product Littlewood–Paley inequalities, which extend (4.5) to our setting, were obtained by Fefferman and Pipher [24]. Let $f = \sum a_R h_R$ and define the *dyadic* product Littlewood–Paley square function

$$S_d f = \left(\sum_{R\in\mathcal{D}^d} |a_R|^2 \mathbf{1}_R \right)^{1/2}.$$

Then for all $1 < p < \infty$, one has

$$A_p^d \|S_d f\|_p \leq \|f\|_p \leq B_p^d \|S_d f\|_p. \tag{4.6}$$

Such inequalities are usually obtained by applying the one-dimensional (Hilbert-space valued) inequalities (4.5) iteratively in each coordinate, hence the constants are raised to the power d. In our setting however, it is enough to apply it only $d-1$ times, since we have $d-1$ free parameters.

One can easily see that, applying (4.6) to either (4.3) or (4.4), one immediately extends the proof to all values of $p \in (1, \infty)$ and obtains (4.1).

4.2 The L^∞ discrepancy estimates

It is natural to conjecture that the extreme discrepancy should be much larger than the average one, i.e. $\|D_N\|_\infty \gg (\log N)^{\frac{d-1}{2}}$. This is supported by the two-dimensional result of Schmidt [40], 1972 : for any $\mathcal{P}_N \subset [0,1]^2$

$$\|D_N\|_\infty \gtrsim \log N. \tag{4.7}$$

It was known for a long time that this bound is best possible, i.e. there exist sets in $[0,1]^2$ with $\|D_N\|_\infty \lesssim \log N$ (Lerch [32] 1904; van der Corput [20], 1934). The best known higher-dimensional constructions go back to Halton and Hammersley [27, 28] and have discrepancy of the order $(\log N)^{d-1}$.

There is no consensus among the experts today as to what the exact form of the conjecture should be. Motivated by the the best known examples, some believe that the correct power of the logarithm should be $d-1$; at the same time, the limitations of the wavelet method, cf. (4.10) below, convince others that the right bound should be

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}. \tag{4.8}$$

A result of the author, Lacey, and Vagharshakyan [11], 2008 states that in all dimensions one gets at least slightly better than the L^2 bound:

$$\|D_N\|_\infty \gtrsim (\log N)^{\frac{d-1}{2}+\eta}$$

for some $\eta = \eta(d) > 0$.

The heuristic (4.2) used in the proof of the L^2 bounds suggests that in order to understand the behavior of $\|D_N\|_\infty$ one should look at the L^∞ norm of linear combinations of h_R with $|R| \approx \frac{1}{N}$. The relevant conjecture (which arises in probability and approximation theory, cf. [9]) states that in dimensions $d \geq 2$, for all $\alpha_R \in \mathbb{R}$

$$n^{\frac{d-2}{2}} \Big\| \sum_{|R|=2^{-n}} \alpha_R h_R \Big\|_\infty \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|. \tag{4.9}$$

This became known as the *small ball conjecture*. If we restrict the choice of coefficients to $\alpha_R = \pm 1$, this form of the conjecture

$$\Big\| \sum_{|R|=2^{-n}} \alpha_R h_R \Big\|_\infty \gtrsim n^{d/2} \tag{4.10}$$

reveals striking similarity to the discrepancy conjecture (4.8). Moreover, an argument similar to (4.3) shows that $\big\| \sum_{|R|=2^{-n}} \alpha_R h_R \big\|_2 \gtrsim n^{\frac{d-1}{2}}$ in perfect unison with the L^2-discrepancy bound (4.1), i.e. in both conjectures one gains $\sqrt{n}$ over the L^2 estimate.

The two-dimensional version of (4.9) was proved by Talagrand [45], 1994 and Temlyakov [46], 1995. Choosing α_R to be random i.i.d. ± 1 one can show that (4.9) is best possible, which is the motivation behind the discrepancy conjecture (4.8).

Take a closer look at the two-dimensional small ball inequality, which has the form

$$\Big\| \sum_{|R|=2^{-n}} \alpha_R h_R \Big\|_\infty \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|. \tag{4.11}$$

One cannot help but notice the similarity to the classical Sidon's theorem [42] in Fourier analysis: if a bounded 2π-periodic function f has lacunary Fourier series $\sum_{k=1}^{\infty} a_k e^{in_k x}$, $n_{k+1}/n_k > \lambda > 1$, then

$$\big\| f \big\|_\infty \gtrsim \sum_{k=1}^{\infty} |a_k|.$$

This theorem is proved using Riesz products $P_K(x) = \prod_{k=1}^{K} (1 + \varepsilon_k \cos n_k x)$. Hence it is not surprising that a similar technique works for (4.11). This proof (due to Temlyakov [46]) is so short and elegant – it is indeed a "proof from the book" – that the author decided to reproduce it here.

Proof. For $k = 0, 1, \ldots, n$, consider the Rademacher-like functions

$$f_k = \sum_{\substack{R:\, |R|=2^{-n} \\ |R_1|=2^{-k}}} \operatorname{sign}(\alpha_R) h_R.$$

Notice that the rectangles in the sum above do not overlap. Construct the test function as a *Riesz product*:

$$\Psi := \prod_{k=1}^{n} (1 + f_k)\,.$$

As a product of non-negative terms, obviously $\Psi \geq 0$. Moreover, in two dimensions, if $R \neq R'$ and $|R| = |R'|$, then $h_R \cdot h_{R'} = \pm h_{R \cap R'}$. Therefore, $\int \Psi = 1$. (Indeed, multiply out the Riesz product – the leading term is one, and the rest are products of Haar functions, hence have integral zero.) Thus $\big\|\Psi\big\|_1 = 1$, and we have

$$\Big\| \sum_{|R|=2^{-n}} \alpha_R h_R \Big\|_\infty \geq \Big\langle \sum \alpha_R h_R, \Psi \Big\rangle = \sum |\alpha_R| \langle h_R, h_R \rangle = \sum_{|R|=2^{-n}} |\alpha_R| {\cdot} 2^{-n},$$

where the non-diagonal terms vanish by the same token. This finishes the proof. □

The proof of (4.7) in [26] is just a bit more technical than the argument above.

4.3 The other endpoint: L^1

The opposite endpoint of the L^p scale is even more mysterious. In 1981 Halász [26] proved that in dimension $d = 2$ for any collection of N points $\mathcal{P}_N \subset [0,1]^2$ one has

$$\|D_N\|_1 \gtrsim \sqrt{\log N}. \tag{4.12}$$

It can be shown that this bound, $\|D_N\|_1 \gtrsim \sqrt{\log N}$, continues to hold for $d \geq 3$, and nothing better is known in higher dimensions, i.e. it is not even known if the bound grows with the dimension! It is conjectured, of course, that the L^1 estimate should match the L^p bounds for $1 < p < \infty$:

$$\|D_N\|_1 \gtrsim (\log N)^{\frac{d-1}{2}}. \tag{4.13}$$

While this conjecture seems to be out of reach at the moment, some relevant work has been done recently. In 2010, Lacey [30] has obtained estimates in function spaces "near" L^1 . He proved that

$$\|D_N\|_{L(\log L)^{\frac{d-2}{2}}} \gtrsim (\log N)^{\frac{d-1}{2}}. \tag{4.14}$$

This result is non-trivial, but the same bound in $L(\log L)^{\frac{d-1}{2}}$ is quite easy – one would need to estimate the test function F (4.4) in the dual space $\exp\left(L^{\frac{2}{d-1}}\right)$, which is straightforward and aligns with the logic that the Littlewood–Paley inequalitites only need to be applied $d-1$ times. Lacey also proves the estimate in the (dyadic) d-parameter Hardy space H^p, $0 < p \leq 1$:

$$\|D_N\|_{H^p} \gtrsim (\log N)^{\frac{d-1}{2}}.$$

The classical Littlewood–Paley characterizations of H^p yields the norm equivalence $\|f\|_{H^p} \approx \|S_d f\|_{L^p}$, hence the proof essentially repeats the L^p argument for $p > 1$.

In addition, several L^1 "dichotomy" results have been obtained recently by the author, Amirkhanyan, and Lacey [1]. They say that if for some $\mathcal{P}_N$ the L^1-discrepancy is too small and violates the conjecture, then the discrepancy of $\mathcal{P}_N$ has to be large in some other norm. Below is a sampler of the results quantifying this phenomenon.

Theorem 4.1 ([1], 2013). *The following statements hold:*

(i) If $\mathcal{P}_N \subset [0,1]^d$ satisfies $\|D_N\|_p \lesssim (\log N)^{\frac{d-1}{2}}$ for some $1 < p < \infty$, then

$$\|D_N\|_1 \gtrsim (\log N)^{\frac{d-1}{2}}. \tag{4.15}$$

(ii) Every $\mathcal{P}_N \subset [0,1]^d$ satisfies either

$$\|D_N\|_1 \geq (\log N)^{(d-1)/2-\epsilon} \quad \text{or} \quad \|D_N\|_2 \geq \exp(c(\log N)^{\epsilon}). \tag{4.16}$$

(iii) For $d \geq 3$*, if* $\mathcal{P}_N \subset [0,1]^d$ *satisfies* $\|D_N\|_1 \lesssim \sqrt{\log N}$*, then*

$$\|D_N\|_2 \gtrsim N^C. \tag{4.17}$$

Some remarks are in order. The first result, (4.15), says essentially that if some point distribution $\mathcal{P}_N$ violates the L^1 conjecture (4.13), then it necessarily has L^p-discrepancy larger than the optimal $(\log N)^{\frac{d-1}{2}}$. This statement is very simple – it just requires an application of Hölder's inequality and the Roth–Schmidt bound (4.1).

The second inequality, (4.16), states that if the exponent in the L^1 conjecture is incorrect, then the L^2-discrepancy must be much larger. The idea of the proof is the following: while the "dual" function F (4.4) does not have good L^∞ bounds, it does satisfy certain exponential estimates (see [11]), hence it can only be large on very small sets. If in addition, $\|D_N\|_2$ is not too big, then throwing away these small sets is harmless, and one can prove an L^1 bound.

Finally, inequality (4.17) speculates what happens if Halász's bound (4.12) were indeed best possible in higher dimensions. In this case, the L^2 discrepancy of the optimal distribution has to be huge (polynomial in N). The rough idea is that while the L^1 norm is very small, the very close $L(\log L)^{\frac{d-2}{2}}$ norm is much larger (4.14), which implies very strong localization of D_N. Extrapolation arguments allow one to push this heuristic to prove the L^2 estimate (4.17). The details are contained in [1].

Bibliography

[1] G. Amirkhanyan, D. Bilyk, M. Lacey, Dichotomy results for the L^1 norm of the discrepancy function. *Journal of Math. Analysis and Appl.* **410** (2014), 1–6.

[2] C. Aistleitner, Metric number theory, lacunary series, and systems of dilated functions. *In: this volume*.

[3] R. C. Baker, Metric number theory and the large sieve. *J. London Math. Soc. (2)* **24** (1981), 34–40.

[4] J. Beck, Some upper bounds in the theory of irregularities of distribution. *Acta Arith.*, **43** (1984), 115–130.

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

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

[7] J. Beck and W. W. L. Chen, Irregularities of distribution. Cambridge University Press, Cambridge (1987)

[8] I. Berkes and W. Philipp, The size of trigonometric and Walsh series and uniform distribution mod 1. *J. London Math. Soc. (2)* **50** (1994), 454–464.

[9] D. Bilyk, On Roth's orthogonal function method in discrepancy theory. *Uniform Distribution Theory* **6** (2011), 143–184.

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

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

[12] D. Bilyk, X. Ma, J. Pipher, C. Spencer, Directional discrepancy in two dimensions. *Bull. London Math. Soc.* **43** (2011), 1151–1166.

[13] D. Bilyk, X. Ma, J. Pipher, C. Spencer, Diophantine approximation and directional discrepancy of rotated lattices. *available on www.arxiv.org*, (2013).

[14] L. Carleson, On convergence and growth of partial sums of Fourier series. *Acta Mathematica* **116** (1966), 135–157.

[15] B. Chazelle, The Discrepancy Method. Randomness and complexity. Cambridge University Press., Cambridge (2000)

[16] W. W. L. Chen, On irregularities of distribution. II. *Quart. J. Math. Oxford Ser. (2)* **34** (1983), 257–279.

[17] W. W. L. Chen, Fourier techniques in the theory of irregularities of point distribution. In: *Fourier analysis and convexity, Appl. Numer. Harmon. Anal.*, pp. 59–82. Birkhäuser, Boston (2004)

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

[19] W. W. L. Chen and G. Travaglini, Deterministic and probabilistic discrepancies. *Ark. Mat.* **247** (2009), 273–293.

[20] J. G. van der Corput, Verteilungsfunktionen I. *Akad. Wetensch. Amsterdam, Proc.* **38** (1935), 813–821.

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

[22] J. Dick and F. Pillichshammer, Digital nets and sequences. Discrepancy theory and quasi-Monte Carlo integration. Cambridge University Press, Cambridge (2010)

[23] P. Erdős and P. Turán, On a problem in the theory of uniform distribution. I. *Nederl. Akad. Wetensch.* **51** (1948), 1146–1154.

[24] R. Fefferman and J. Pipher, Multiparameter operators and sharp weighted inequalities. *Amer. J. Math.* **119** (1997), 337–369.

[25] K. K. Frolov, Upper bounds on the discrepancy in L_p, $2 \leq p < \infty$. *Dokl. Akad. Nauk USSR* **252** (1980), 805–807.

[26] G. Halász, On Roth's method in the theory of irregularities of point distributions. In: *Recent progress in analytic number theory*, Vol. 2, pp. 79–94. Academic Press, London (1981)

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

[28] J. M. Hammersley, Monte Carlo methods for solving multivariable problems. *Ann. New York Acad. Sci.* **86** (1960), 844–874.

[29] L. Kuipers and H. Niederreiter, Uniform distribution of sequences John Wiley & Sons, New York-London-Sydney (1974)

[30] M. Lacey, On the discrepancy function in arbitrary dimension, close to L^1. *Analysis Math.* **34** (2008), 119–136.

[31] M. Lacey, Small Ball and Discrepancy Inequalities (2008). *Available on www.arxiv.org.*

[32] M. Lerch, Question 1547. *L'Intermediaire Math.* **11** (1904), 144–145.

[33] J. Matoušek, Geometric Discrepancy: An Illustrated Guide. Springer-Verlag, Berlin (1999)

[34] H. L. Montgomery, On irregularities of distribution. *In: Congress of Number Theory (Zarautz, 1984)*, pp. 11–27. Univ. del País Vasco, Bilbao, 1989.

[35] H. L. Montgomery, Ten lectures on the interface between analytic number theory and harmonic analysis. AMS, Providence (1994)

[36] C. Pereyra, Lecture notes on dyadic harmonic analysis. *In: S. Pérez-Esteva, C. Villegas-Blas, Topics in analysis: harmonic, complex, nonlinear and quantization*, pp.1–61. Contemp. Math. **289** AMS (2001)

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

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

[39] W. M. Schmidt, Irregularities of distribution. IV. *Invent. Math.*, **7** (1969), 55–82.

[40] W. M. Schmidt, Irregularities of distribution. VII. *Acta Arith.* **21** (1972), 45–50.

[41] W. M. Schmidt, Irregularities of distribution. X. *In: Number theory and algebra*, pp. 311–329. Academic Press, New York (1977)

[42] S. Sidon, Verallgemeinerung eines Satzes über die absolute Konvergenz von Fourierreihen mit Lücken. *Math. Ann.* **97** (1927), 675–676.

[43] M. M. Skriganov, Harmonic analysis on totally disconnected groups and irregularities of point distributions. *J. Reine Angew.* **600** (2006), 25–49.

[44] E. M. Stein and S. Wainger, Problems in harmonic analysis related to curvature, *Bull. Amer. Math. Soc.* **84** (1978), 1239–1295.

[45] M. Talagrand, The small ball problem for the Brownian sheet. *Ann. Probab.* **22** (1994), 1331–1354.

[46] V. N. Temlyakov, Some Inequalities for Multivariate Haar Polynomials. *East Journal on Approximations* **1** (1995), 61–72.

[47] V. N. Temlyakov, Cubature formulas, discrepancy, and nonlinear approximation. *J. Complexity* **19** (2003), 352–391.

[48] H. Weyl. Über die Gleichverteilung von Zahlen mod Eins. *Math. Ann.* **77** (1916) 313–352.

Author information

Dmitriy Bilyk, School of Mathematics, University of Minnesota, Minneapolis MN 55455, USA.
E-mail: dbilyk@umn.edu