Mathematika 59 (2013) 365–379 © University College London
doi:10.1112/S0025579312001180

LOWER BOUNDS FOR L_1 DISCREPANCY

ARMEN VAGHARSHAKYAN

Abstract. We find the best asymptotic lower bounds for the coefficient of the leading term of the L_1 norm of the two-dimensional axis-parallel discrepancy that can be obtained by Roth's orthogonal function method among a large class of test functions. We use methods of combinatorics, probability, and complex and harmonic analysis.

§1. *Introduction.* Let P be a finite set in the two-dimensional unit square ($P \subset [0,1]^2$). Define the (two-dimensional axis-parallel) discrepancy function as follows:

$$D(x,y) = D_P(x,y) = \sharp(P \cap ([0,x] \times [0,y])) - \sharp(P) \cdot xy, \quad (x,y) \in [0,1]^2,$$

where $\sharp(Q)$ denotes the number of elements (cardinality) of a set Q. The discrepancy function measures the difference between the actual number of points of the set P in an axis-parallel rectangle $[0,x] \times [0,y]$ and the "expected" number of points in that rectangle. Thus, the discrepancy function quantifies the "closeness" of the distribution generated by a finite set P to the uniform distribution.

The L_∞ norm of the discrepancy function $\|D_P\|_{L_\infty([0,1]^2)}$ (also known as the star-discrepancy) has been studied for a while and is known to be related to numerical integration, metric entropy, the small ball inequality, etc. We refer the reader to [**1**, **3**] or [**10**] for a comprehensive discussion of the star-discrepancy.

In this paper we will be dealing with lower bounds for the L_1 norm of the discrepancy function. Quoting a recent essay on the discrepancy theory [**1**]: "the other endpoint of the L^p scale, $p = 1$, is not any less (and perhaps even more) difficult than the star-discrepancy estimates. The only information that is available is the two-dimensional inequality (proved in the paper of Halasz [**6**])".

When discussing the L_1 norm of the discrepancy function, it is convenient to define

$$d_N = \frac{1}{\sqrt{\ln N}} \cdot \inf_{\sharp(P)=N} \int_0^1 \int_0^1 |D_P(x,y)|\,dx\,dy. \tag{1.1}$$

It can be deduced from Davenport's argument [**4**] that

$$\limsup_{N\to\infty} d_N < +\infty. \tag{1.2}$$

Received 15 September 2012, published online 18 March 2013.
MSC (2010): 11K38 (primary).

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

On the other hand (see Remark 4.3) Halasz [6] proved the following inequality:

$$\liminf_{N\to\infty} d_N \geqslant c_H > 0. \tag{1.3}$$

A careful analysis of his proof shows that

$$c_H = \frac{1}{1152(\sqrt{\mathrm{e}}+1)\sqrt{\ln 2}} \approx 0.00039.$$

The inequalities (1.2) and (1.3) show that d_N is the coefficient of the leading term of the L_1 norm of the two-dimensional axis-parallel discrepancy function.

In this paper we improve the asymptotic lower bounds for d_N, namely the following.

THEOREM 1.4. *For the coefficient of the leading term of the L_1 norm of the two-dimensional discrepancy defined by (1.1) the following asymptotic lower bounds hold:*

$$\liminf_{N\to\infty} d_N \geqslant \frac{3}{256\sqrt{\mathrm{e}\ln 2}} \approx 0.00854$$

and

$$\limsup_{N\to\infty} d_N \geqslant \frac{1}{64\sqrt{\mathrm{e}\ln 2}} \approx 0.01138.$$

For comparison, the best-known lower estimates of the L_2 norm of the discrepancy were obtained in [7].

Remark 1.5. In [5] it was proved that

$$\liminf_{N\to\infty} \frac{1}{\sqrt{\ln N}} \cdot \inf_{\sharp(P)=N} \|D_P\|_{L_2([0,1]^2)} \leqslant 0.17905.$$

Numerical experiments for the sets introduced in [2] make the following estimate plausible:

$$\liminf_{N\to\infty} \frac{1}{\sqrt{\ln N}} \cdot \inf_{\sharp(P)=N} \|D_P\|_{L_2([0,1]^2)} \leqslant 0.17601.$$

Comparison of these results with Theorem 1.4 shows that for sets with minimal L_2 discrepancy their L_2 discrepancy is close to the L_1 discrepancy.

In this paper we also show that Theorem 1.4 provides the best bounds that can be obtained by Roth's orthogonal function method for a large class of test functions (see Theorems 4.7 and 7.1).

Theorem 1.4, as well as the fact that the best-known lower bounds for the lim sup and lim inf of the coefficient of the leading term of the L_2 norm of the two-dimensional discrepancy are also different (this fact can be derived from [7]), leads us to formulate the following conjecture.

CONJECTURE 1.6. *The following strict inequality is true:*

$$\limsup_{N\to\infty} d_N > \liminf_{N\to\infty} d_N.$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

§2. *Auxiliary functions.* The existing proofs of lower bounds for discrepancy are based on a certain construction of functions called Roth's auxiliary functions (see [11]). In this section we will modify that construction so that now the auxiliary functions do not take the value zero, but from all other perspectives they are as good as Roth's original auxiliary functions. The fact that they do not take the value zero will play a crucial role later on.

For a finite set $P \subset [0, 1]^2$ denote by $N = \sharp(P)$ the number of elements of the set P. Let n be the only integer satisfying the condition

$$2^{n-1} < 2\sharp(P) = 2N \leqslant 2^n. \tag{2.1}$$

For every $i = 0, 1, \ldots, n$ we will define E_i to be a family of dyadic rectangles in $[0, 1]^2$:

$$E_i = \bigcup_{l=0}^{+\infty} E_i^l,$$

where the sets E_i^l will be defined recursively in the next paragraph, and we will define C_i to be a family of dyadic rectangles in $[0, 1]^2$:

$$C_i = \bigcup_{l=0}^{+\infty} C_i^l,$$

where the sets C_i^l will be defined recursively in the next paragraph.

Recursive definition. *Step* 0. In this step we define the families E_i^0 and C_i^0. Consider the dyadic rectangles $R = R_x \times R_y \in [0, 1]^2$ whose sidelengths satisfy $|R_x| = 2^{-i}$, $|R_y| = 2^{i-n}$. Denote by C_i^0 the family of those sets R that contain at least a point of the set P and by E_i^0 the family of those sets R that do not contain any point of the set P. We have

$$\sharp(C_i^0) + \sharp(E_i^0) = 2^n$$

and

$$\sharp(C_i^0) \leqslant N.$$

Step 1. Partition each of the sets $R \in C_i^0$ into $2^{2(n+1)}$ dyadic rectangles which have the form $R^1 = R_x^1 \times R_y^1$, where $|R_x^1| = |R_x|/2^{n+1}$, $|R_y^1| = |R_y|/2^{n+1}$. Define E_i^1 to be the family of those sets R^1 which do not contain any point of the set P and define C_i^1 to be the family of those sets R^1 which contain at least a point of the set P. We have

$$\sharp(C_i^1) \leqslant N.$$

Step $l + 1$. In step $l + 1$ partition each of the sets $R^l = R_x^l \times R_y^l \in C_i^l$ into $2^{2(n+1)}$ dyadic rectangles which have the form $R^{l+1} = R_x^{l+1} \times R_y^{l+1}$, where

$$|R_x^{l+1}| = |R_x^l|/2^{n+1}, \qquad |R_y^{l+1}| = |R_y^l|/2^{n+1}.$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

Define E_i^{l+1} to be the family of those sets R^{l+1} which do not contain any point of the set P and define C_i^{l+1} to be the family of those sets R^{l+1} which contain at least a point of the set P. We have

$$\sharp(C_i^{l+1}) \leqslant N.$$

Definition 2.2. Let $R = R_x \times R_y$ be a dyadic rectangle in $[0, 1]^2$. By h_R we will denote the L_∞ normalized Haar function adapted to the dyadic rectangle R, i.e.

$$h_R(x, y) = h_{R_x}(x) \cdot h_{R_y}(y),$$

where the function h_{R_x} equals 1 on the left half of the dyadic interval R_x, -1 on the right half of the dyadic interval R_x and 0 outside of the dyadic interval R_x. Similarly, the function h_{R_y} equals 1 on the left half of the dyadic interval R_y, -1 on the right half of the dyadic interval R_y and 0 outside of the dyadic interval R_y.

Definition 2.3. Define the auxiliary functions f_i by

$$f_i = \sum_{R \in E_i} h_R, \tag{2.4}$$

where h_R is the L_∞ normalized Haar function adapted to the rectangle R (see Definition 2.2).

Remark 2.5. The function

$$f_i^0 = \sum_{R \in E_i^0} h_R$$

is the original Roth's auxiliary function.

The following lemma shows that the auxiliary function f_i is defined almost everywhere on $[0, 1]^2$.

LEMMA 2.6. *The set*

$$M = [0, 1]^2 \Big\backslash \bigcup_{R \in E_i} R$$

has measure zero.

Proof.

$$\sum_{R \in C_i^l} |R| = \sharp(C_i^l) 2^{-n-2(n+1)l} = N 2^{-n-2(n+1)l} \to 0 \quad \text{as } l \to +\infty. \qquad \square$$

Remark 2.7. Due to Lemma 2.6 the function f_i is defined almost everywhere on $[0, 1]^2$. For definiteness, we will additionally take f_i to be equal to 1 on the remaining set of measure zero.

We now investigate the properties of the functions f_i.

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

The following lemma shows that the auxiliary functions f_i do not take the value zero, unlike Roth's original auxiliary functions.

LEMMA 2.8.

$$f_i : [0, 1]^2 \to \{-1, 1\}$$

or equivalently

$$|f_i(x, y)| \equiv 1.$$

Proof. The proof follows from the definition of the functions f_i (see Definition 2.3) , from the Remark 2.7 and the fact that the elements of E_i are disjoint. □

The following lemma estimates the inner product of the discrepancy function D_P and the auxiliary function f_i.

LEMMA 2.9.

$$\int_0^1 \int_0^1 D_P f_i \leqslant -(2^n - N)\frac{N \cdot 2^{-2n}}{16}. \tag{2.10}$$

Proof. We use the following well-known fact about the interrelation between discrepancy and Haar functions (see e.g. [3]): if the coordinate rectangle R does not contain any point of the finite set P then

$$\int_0^1 \int_0^1 D_P h_R = -\frac{\sharp(P)|R|^2}{16}.$$

Let E_i^0 be as in the construction of the function f_i, then we have

$$\begin{aligned}\int_0^1 \int_0^1 D_P \cdot f_i &\leqslant \sum_{R \in E_i^0} \int_0^1 \int_0^1 D_P \cdot h_R \\ &= -\sum_{R \in E_i^0} \frac{N|R|^2}{16} \leqslant -(2^n - N)\frac{N \cdot 2^{-2n}}{16}.\end{aligned}$$

□

LEMMA 2.11. *Let* $0 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$. *Then*

$$\text{(a)} \quad \int_0^1 \int_0^1 f_{i_1} f_{i_2} \cdots f_{i_p} = 0;$$

and

$$\text{(b)} \quad \left| \int_0^1 \int_0^1 D_P \cdot f_{i_1} f_{i_2} \cdots f_{i_p} \right| \leqslant 2^{i_1 - i_p} \frac{N 2^{-n}}{16}.$$

Proof. Let

$$R_{i_1} \in E_{i_1},\ R_{i_2} \in E_{i_2},\ \ldots,\ R_{i_p} \in E_{i_p}.$$

Denote

$$R_{i_1} = R_{i_1,x} \times R_{i_1,y},\ R_{i_2} = R_{i_2,x} \times R_{i_2,y},\ \ldots,\ R_{i_p} = R_{i_p,x} \times R_{i_p,y}.$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

For definiteness assume

$$|R_{i_1,x}| = 2^{-i_1-(n+1)l_1},\ |R_{i_1,y}| = 2^{i_1-n-(n+1)l_1},$$
$$|R_{i_2,x}| = 2^{-i_2-(n+1)l_2},\ |R_{i_2,y}| = 2^{i_2-n-(n+1)l_2},$$
$$\vdots$$
$$|R_{i_p,x}| = 2^{-i_p-(n+1)l_p},\ |R_{i_p,y}| = 2^{i_p-n-(n+1)l_p}.$$

Note that the repeated subdivisions into $2^{2(n+1)}$ dyadic rectangles in the construction in §2 guarantees that the lengths of the dyadic intervals $R_{i_1,x}, \dots, R_{i_p,x}$ are mutually distinct, similarly the lengths of the dyadic intervals $R_{i_1,y}, \dots, R_{i_p,y}$ are mutually distinct. Hence the product

$$h_{R_1} h_{R_2} \cdots h_{R_p}$$

is either identically zero or (up to a sign) the Haar function on the intersection

$$R = \bigcap_{t=1}^{p} R_{i_t}.$$

Thus the first claim of the theorem follows.

We now estimate the sidelengths of the dyadic rectangle R. Write $R = R_x \times R_y$, then

$$|R_x| \leqslant 2^{-i_p}, \qquad |R_y| \leqslant 2^{i_1-n},$$

and therefore

$$|R| = |R_x \times R_y| \leqslant 2^{-i_p+i_1-n}. \tag{2.12}$$

Now denote

$$E = \left\{ R = \bigcap_{t=1}^{p} R_{i_t} : R_{i_1} \in E_{i_1}, \dots, R_{i_p} \in E_{i_p} \right\}.$$

Note that the elements of E are mutually disjoint (as the elements of each E_{i_t} are mutually disjoint). We also have

$$f_{i_1} f_{i_2} \cdots f_{i_p} = \sum_{R \in E} \pm h_R.$$

Using (2.12) we obtain

$$\left| \int_0^1 \int_0^1 D_P \cdot f_{i_1} f_{i_2} \cdots f_{i_p} \right| \leqslant \sum_{R \in E} \left| \int_0^1 \int_0^1 D_P \cdot h_R \right|$$
$$= \sum_{R \in E} \frac{N|R|^2}{16} \leqslant \frac{N 2^{-i_p+i_1-n}}{16} \sum_{R \in E} |R| = 2^{i_1-i_p} \frac{N 2^{-n}}{16}. \qquad \square$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

§3. *Combinatorial properties of the auxiliary functions.* We now explore the combinatorial properties of the auxiliary functions f_i defined by 2.3. We adapt the following notation from the theory of symmetric functions (see e.g. [9]):

$$e_1 = e_1^n = \sum_{i=0}^{n} f_i, \tag{3.1}$$

$$e_2 = e_2^n = \sum_{0 \leqslant i_1 < i_2 \leqslant n} f_{i_1} f_{i_2},$$

$$e_3 = e_3^n = \sum_{0 \leqslant i_1 < i_2 < i_3 \leqslant n} f_{i_1} f_{i_2} f_{i_3},$$

and so on.

Let the number k be odd ($k = 1, 3, 5, \dots$). Then using the fact that $f_i^2 \equiv 1$ (see Lemma 2.8) we can write

$$(f_0 + f_1 + \cdots + f_n)^k = A_1^n(k) e_1^n + A_3^n(k) e_3^n + A_5^n(k) e_5^n + \cdots \tag{3.2}$$

(here we assume that $A_p^n(k) = 0$ once $p > \min(k, n+1)$).

Remark 3.3. For example, we compute

$$(f_0 + f_1)^3 = f_0^3 + 3 f_0 f_1^2 + 3 f_0^2 f_1 + f_1^3 = 4 f_0 + 4 f_1,$$

so that $A_1^1(3) = 4$.

Remark 3.4. It is important to point out that the coefficient of the term of

$$(f_0 + f_1 + \cdots + f_n)^k$$

that is linear with respect to $f_0 + f_1 + \cdots + f_n$ is $A_1^n(k)$.

§4. *Test functions.* We start this section by sketching Roth's orthogonal function method (as modified by Halasz [6]).

Let $P \subset [0, 1]^2$ be a finite set and let $f_0, f_1, \dots, f_n$ be the corresponding functions defined by 2.3. Let $H_n = H_n(z_0, z_1, \dots, z_n)$ be an odd, symmetric function that is entire in each of its $n+1$ variables.

By $\mathrm{LIN}(H_n)$ we will denote the coefficient of the term of the composite function $H_n(f_0, f_1, \dots, f_n)$ which is linear with respect to $\sum_{i=0}^n f_i$. For example, according to Remark 3.4, $\mathrm{LIN}((f_0 + \cdots + f_n)^k) = A_1^n(k)$.

Assume that

$$\| H_n(z_0, z_1, \dots, z_n) \|_{L_\infty([-1,1]^{n+1})} = 1. \tag{4.1}$$

Then we can estimate the L_1 norm of the discrepancy function $D_P(x, y)$ from below as follows:

$$\| D_P \|_1 \geqslant \left| \int_0^1 \int_0^1 D_P \cdot H_n(f_0, f_1, \dots, f_n) \right|$$

$$\geqslant \mathrm{LIN}(H_n) \left| \sum_{i=0}^{n} \int_0^1 \int_0^1 D_P f_i \right| + \mathit{Error},$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

where in Roth's orthogonal function method it is further proved that the term *Error* is negligible (estimating the term *Error* from above by means of Lemma 2.11). Then Lemma 2.9 is applied to obtain the following lower bound for the L_1 norm of the discrepancy function:

$$\|D_P\|_1 \geqslant \mathrm{LIN}(H_n)(n+1)(2^n - N)\frac{N2^{-2n}}{16} + Error, \tag{4.2}$$

where the integers N and n are defined by (2.1).

Informally, Roth's orthogonal function method (as modified by Halasz [**6**]) consists of finding appropriate functions H_n (called test functions) for which $\mathrm{LIN}(H_n)$ grows as fast as possible, whereas the error term *Error* is negligible, so that an appropriate lower estimate for $\|D_P\|_1$ follows.

Remark 4.3. In his proof Halasz [**6**] used the following test function:

$$G = \prod_{i=0}^{n}\left(1 + \frac{i\gamma}{\sqrt{\log N}} f_i^0\right) - 1,$$

where the auxiliary functions f_i^0 are as in Remark 2.5 and $\gamma > 0$ is a certain small constant. This test function involves complex numbers. The choice of the test function G was motivated by Riesz products. We will use a different approach by choosing our test function to be a solution of an extremal problem in Fourier analysis (see Theorem 7.1).

Remark 4.4. The rest of the paper is organized as follows. We will first show that the problem of finding appropriate test functions H_n reduces to finding a certain function T (see Theorem 4.7), so that the quantity corresponding to $\mathrm{LIN}(H_n)$ is

$$\mathrm{LIN}_n(T) = \mathrm{LIN}\left(T\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right)\right). \tag{4.5}$$

Then we will derive a precise formula for $\mathrm{LIN}_n(T)$ (see §6). After that we will find the function T that maximizes the quantity

$$\limsup \sqrt{n} \cdot \mathrm{LIN}_n(T) \tag{4.6}$$

in a certain natural class of functions (see Theorem 7.1) and show that for that function the error term *Error* appearing in (4.2) can be neglected (see §8).

THEOREM 4.7. *Without loss of generality one can assume that the test functions H_n have the following form:*

$$H_n(f_0, f_1, \ldots, f_n) = T\left(\frac{f_0 + f_1 + \cdots + f_n}{\sqrt{n}}\right).$$

Proof. According to the "Main theorem on symmetric functions" (see e.g. [**9**]) the symmetric function $H_n(f_0, f_1, \ldots, f_n)$ can be written as a function

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

depending only on the functions e_i (introduced by (3.1)). We borrow the following notation from the theory of symmetric functions:

$$p_k = \sum_{0 \leqslant j \leqslant n} f_j^k, \quad 1 \leqslant k \leqslant n+1.$$

Recalling Newton's identities

$$k e_k = \sum_{i=1}^{k} (-1)^{i-1} e_{k-i}\, p_i,$$

we obtain that a symmetric function can be written as an analytic function depending only on the functions p_i. Taking into account that $f_i^2 = 1$ (see Lemma 2.8) we have

$$p_k = p_1, \quad 1 \leqslant k \leqslant n+1,\ k \text{ is odd},$$

and

$$p_k = n+1 = p_0, \quad 1 \leqslant k \leqslant n+1,\ k \text{ is even}.$$

Thus, the function H_n depends only on p_0 and p_1, where p_0 is a constant. Recalling that H_n is odd, it turns out that H_n has to depend only on

$$p_1 = f_0 + \cdots + f_n,$$

i.e. we can write the function H_n in the following form:

$$H_n(f_0, f_1, \ldots, f_n) = F_n(f_0 + \cdots + f_n).$$

Introduce functions T_n as follows:

$$F_n(f_0 + \cdots + f_n) = T_n\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right).$$

Then the condition (4.1) becomes

$$\sup_{(-\sqrt{n}, \sqrt{n})} |T_n(x)| = 1.$$

Recall that $\{T_n(z)\}_{n=1}^{\infty}$ is a family of odd, entire functions. Either it is normal or it goes locally uniformly to 0, or to ∞. The latter two cases are not interesting, so let us assume that T_n has a subsequence that tends to a certain odd, entire function T locally uniformly in the complex plane. For clarity of notation we will assume that T_n itself tends to T.

Note that the functions f_i are non-correlated on the unit square and take the values 1 and -1 with equal probability, thus

$$\frac{f_0 + \cdots + f_n}{\sqrt{n}} \to N(0, 1).$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

So that the conditions

$$\sup_{(-\sqrt{n},\sqrt{n})} |T_n(x)| = 1$$

imposed on the functions T_n induce the following condition on the limiting function T:

$$\sup_{-\infty<x<+\infty} |T(x)| = 1,$$

and instead of discussing the functions

$$T_n\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right)$$

we can discuss the functions

$$T\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right). \qquad \square$$

§5. *Generating function of* $A_1^n(k)$. In this section we derive the formula (5.1) for the generating function of the sequence $\{A_1^n(k)\}_{k=1,3,\ldots}$ introduced by the identity (3.2).

Denote by B_n the binomial distribution, i.e. the distribution of the sum of n independent random variables which take values 1 and 0 with probability $1/2$. Denote by S_n the distribution of n independent random variables which take values 1 and -1 with probability $1/2$. Then

$$S_n = 2B_n - n.$$

Denote by M_{B_n} the moment-generating function of the binomial distribution B_n, that is

$$M_{B_n}(t) = E(\mathrm{e}^{tB_n}) = \sum_{k=0}^{\infty} \frac{1}{k!} E((B_n)^k) t^k,$$

where $E(X)$ is the mathematical expectation of a random variable with distribution X.

We know that

$$M_{B_n}(t) = (\tfrac{1}{2} + \tfrac{1}{2}\mathrm{e}^t)^n.$$

As a consequence we obtain

$$\begin{aligned} M_{2B_n-n+1}(t) &= E(\mathrm{e}^{(2B_n-n+1)t}) = \mathrm{e}^{(-n+1)t} M_{B_n}(2t) \\ &= \mathrm{e}^{(-n+1)t}\left(\frac{1}{2} + \frac{1}{2}\mathrm{e}^{2t}\right)^n = \mathrm{e}^t\left(\frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2}\right)^n. \end{aligned}$$

Let $x_0, x_1, \ldots, x_n$ be independent random variables, where $x_0 \equiv 1$ and $x_1, \ldots, x_n$ take values 1 and -1 with probability $1/2$. Let k be an odd number. We have

$$E((x_0 + \cdots + x_n)^k) = A_1^n(k).$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

But on the other hand, the sum $x_0 + \cdots + x_n$ has the distribution $1 + S_n$. We note that $1 + S_n = 2B_n - n + 1$, so that

$$\begin{aligned}\sum_{k=1,3,\ldots} \frac{1}{k!} A_1^n(k) t^k &= \sum_{k=1,3,\ldots} \frac{1}{k!} E((x_0 + x_1 + \cdots + x_n)^k) t^k \\ &= \frac{M_{2B_n-n+1}(t) - M_{2B_n-n+1}(-t)}{2} \\ &= \frac{e^t - e^{-t}}{2} \left(\frac{e^t + e^{-t}}{2} \right)^n. \end{aligned} \tag{5.1}$$

§6. *Precise formula for the linear part.* Recall (see equation (4.5)) that

$$\mathrm{LIN}_n(T) = \mathrm{LIN}\left(T\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}} \right) \right).$$

In this section we will derive the precise formula (6.2) for $\mathrm{LIN}_n(T)$.

We have

$$T\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}} \right) = T'(0) \frac{\sum_{i=0}^n f_i}{\sqrt{n}} + \frac{T^{(3)}(0)}{3!} \frac{(\sum_{i=0}^n f_i)^3}{(\sqrt{n})^3} + \cdots .$$

The coefficient of the term of this expansion that depends on $\sum_{i=0}^n f_i$ linearly is

$$\mathrm{LIN}_n(T) = \left[\frac{T'(0)}{\sqrt{n}} + \frac{T^{(3)}(0)}{3!(\sqrt{n})^3} A_1^n(3) + \frac{T^{(5)}(0)}{3!(\sqrt{n})^3} A_1^n(5) + \cdots \right],$$

or in terms of the differentiation operator ∂,

$$\mathrm{LIN}_n(T) = \left[\frac{\partial}{\sqrt{n}} + \frac{\partial^3}{3!(\sqrt{n})^3} A_1^n(3) + \frac{\partial^5}{3!(\sqrt{n})^3} A_1^n(5) + \cdots \right](T)(0).$$

Using the formula (5.1), we obtain

$$\mathrm{LIN}_n(e^{i\omega x}) = \left(\frac{e^{i\omega/\sqrt{n}} - e^{-i\omega/\sqrt{n}}}{2} \right) \left(\frac{e^{i\omega/\sqrt{n}} + e^{-i\omega/\sqrt{n}}}{2} \right)^n, \tag{6.1}$$

or in general,

$$\mathrm{LIN}_n(T) = \left[\left(\frac{e^{\partial/\sqrt{n}} - e^{-\partial/\sqrt{n}}}{2} \right) \left(\frac{e^{\partial/\sqrt{n}} + e^{-\partial/\sqrt{n}}}{2} \right)^n \right](T)(0). \tag{6.2}$$

§7. *Solution of the extremal problem.* Recall (see Remark 4.4) that we are looking for an odd function T (introduced in Theorem 4.7) satisfying

$$\sup_{-\infty < x < \infty} |T(x)| = 1,$$

for which the quantity

$$\limsup \sqrt{n} \cdot \mathrm{LIN}_n(T)$$

is maximal.

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

The following theorem proves that under some additional conditions on T this extremal problem has a solution and that solution is

$$T(x) = \sin(x).$$

THEOREM 7.1. *If the function T satisfies the following conditions:*

(a) *the support of the Fourier transform of T is countable, and the set $\{|\omega_j|\}_{j=1}^{\infty}$ is linearly independent over the set Z of all integers*

$$T(x) = \sum_{j=1}^{\infty} c_j e^{-i\omega_j x}; \tag{7.2}$$

(b) *T is an odd function;*

(c) *$\|T\|_{\infty} = 1$;*

then

$$\limsup \sqrt{n} \cdot LIN_n(T) \leqslant \frac{1}{\sqrt{e}} \tag{7.3}$$

and the maximum is obtained for the function

$$T(x) = \sin(x).$$

Remark 7.4. In Theorem 7.1 by $\mathrm{LIN}_n(T)$ we understand the operator LIN_n formally applied to the series (7.2) by formula (6.1).

Remark 7.5. In Theorem 7.1 the inequality (7.3) is understood in the sense that if the left-hand side is real then the inequality is satisfied.

Proof. Write

$$T(x) = \sum_{j=1}^{\infty} c_j e^{-ix\omega_j}.$$

Note that according to [**8**, p. 183, paragraph 6.9, Theorem 9.3] we have

$$\sup_{-\infty<x<\infty} |T(x)| = \sum_{j=1}^{\infty} |c_j|.$$

Hence

$$\sum_{j=1}^{\infty} |c_j| = 1.$$

On the other hand, according to Remark 7.4 we obtain

$$\limsup \sqrt{n} \cdot \mathrm{LIN}_n(T) = \lim \sqrt{n} \cdot \mathrm{LIN}_n(T) = [\partial e^{\partial^2/2}](T)(0).$$

We calculate

$$[\partial e^{\partial^2/2}](T)(x) = -i \sum_{j=1}^{\infty} \omega_j e^{-\omega_j^2/2} e^{-ix\omega_j} \cdot c_j.$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

Plugging in $x = 0$, we obtain

$$[\partial e^{\partial^2/2}](T)(0) = -i \sum_{j=1}^{\infty} \omega_j e^{-\omega_j^2/2} \cdot c_j.$$

In order to continue with the proof note the following remark.

Remark 7.6. We have

$$\max_{-\infty<\omega<\infty} \omega e^{-\omega^2/2} = \frac{1}{\sqrt{e}},$$

and the maximum is obtained at the point $\omega = 1$. Similarly,

$$\min_{-\infty<\omega<\infty} \omega e^{-\omega^2/2} = -\frac{1}{\sqrt{e}},$$

and the minimum is obtained at the point $\omega = -1$.

Thus, in order that the maximum of the quantity

$$-i \sum_{j=1}^{\infty} \omega_j e^{-\omega_j^2/2} \cdot c_j$$

subject to the restriction

$$\sum_{j=1}^{\infty} |c_j| = 1$$

is obtained we need the sequence $\{\omega_j\}$ to consist of only two points, namely: 1 and -1. As a consequence we obtain that the odd function T that maximizes the quantity

$$[\partial e^{\partial^2/2}](T)(0)$$

subject to the conditions stated in the theorem is

$$T(x) = \frac{i}{2} e^{-ix} + \frac{(-i)}{2} e^{ix} = \sin(x),$$

and in that case

$$[\partial e^{\partial^2/2}](t)(0) = \frac{1}{\sqrt{e}}. \qquad \square$$

§8. *Derivation of the best constant.*

THEOREM 8.1. *We have*

$$\liminf_{N\to\infty} \frac{1}{\sqrt{\ln N}} \inf_{\sharp(P)=N} \int_0^1 \int_0^1 |D_P| \geqslant \frac{3}{256\sqrt{e \ln 2}}$$

and

$$\limsup_{N\to\infty} \frac{1}{\sqrt{\ln N}} \inf_{\sharp(P)=N} \int_0^1 \int_0^1 |D_P| \geqslant \frac{1}{64\sqrt{e \ln 2}}.$$

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

Proof. Taking $T(x) = \sin(x)$ we obtain

$$\begin{aligned} T\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right) &= \sin\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right) \\ &= \operatorname{Im}\left[\prod_{j=0}^{n}\left(\cos\left(\frac{1}{\sqrt{n}}\right) + \mathrm{i}\cdot f_j \cdot \sin\left(\frac{1}{\sqrt{n}}\right)\right)\right] \\ &= \cos^n\left(\frac{1}{\sqrt{n}}\right)\sin\left(\frac{1}{\sqrt{n}}\right)\left(\sum_{j=0}^{n} f_j\right) - 3!\cos^{n-2}\left(\frac{1}{\sqrt{n}}\right) \\ &\quad \times \sin^3\left(\frac{1}{\sqrt{n}}\right)\left(\sum_{0\leqslant j_1<j_2<j_3} f_{j_1} f_{j_2} f_{j_3}\right) + \cdots . \end{aligned}$$

Hence,

$$\begin{aligned} &\int_0^1\int_0^1 D_P \sin\left(\frac{f_0 + \cdots + f_n}{\sqrt{n}}\right) \\ &= \cos^n\left(\frac{1}{\sqrt{n}}\right)\sin\left(\frac{1}{\sqrt{n}}\right)\left(\sum_{j=0}^{n}\int_0^1\int_0^1 D_P f_j\right) - 3!\cos^{n-2}\left(\frac{1}{\sqrt{n}}\right) \\ &\quad \times \sin^3\left(\frac{1}{\sqrt{n}}\right)\left(\sum_{0\leqslant j_1<j_2<j_3}\int_0^1\int_0^1 D_P f_{j_1} f_{j_2} f_{j_3}\right) + \cdots . \end{aligned}$$

For the first term, using Lemma 2.9, we estimate

$$\begin{aligned} &\cos^n\left(\frac{1}{\sqrt{n}}\right)\sin\left(\frac{1}{\sqrt{n}}\right)\left(\sum_{j=0}^{n}\int_0^1\int_0^1 D_P f_j\right) \\ &\geqslant \cos^n\left(\frac{1}{\sqrt{n}}\right)\cdot\left(\frac{1}{\sqrt{n}} + o(1)\right)(n+1)(2^n - N)\frac{N2^{-2n}}{16} \\ &\geqslant \left(\frac{1}{\sqrt{\mathrm{e}}} + o(1)\right)\sqrt{n}(2^n - N)\frac{N2^{-2n}}{16}. \end{aligned}$$

Let us point out that the appearance of

$$\frac{1}{\sqrt{\mathrm{e}}}$$

here is not surprising as it corresponds to $\mathrm{LIN}_n(\sin)$ (see §7).

For the second term we apply Lemma 2.11 to obtain

$$\begin{aligned} &3!\cos^{n-2}\left(\frac{1}{\sqrt{n}}\right)\sin^3\left(\frac{1}{\sqrt{n}}\right)\sum_{0\leqslant j_1<j_2<j_3}\left|\int_0^1\int_0^1 D_P f_{j_1} f_{j_2} f_{j_3}\right| \\ &\leqslant 3!\frac{1}{(\sqrt{n})^3}\frac{N2^{-n}}{16}\sum_{0\leqslant j_1<j_2<j_3\leqslant n} 2^{j_1 - j_3} \\ &\leqslant 3!\frac{1}{(\sqrt{n})^3}\frac{N2^{-n}}{16}\cdot\sum_{d=1}^{n} 2^{-d} d(n-d) = O\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

and the higher order terms are estimated similarly.

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press

Hence we can write

$$\inf_{2^{n-1}<2\sharp(P)=2N\leqslant 2^n} \|D_P\|_1 \geqslant \inf_{2^{n-1}<2\sharp(P)=2N\leqslant 2^n} |D_P H_n|$$
$$\geqslant \left(\frac{1}{\sqrt{\mathrm{e}}}+o(1)\right)\sqrt{n}(2^n-N)\frac{N2^{-2n}}{16}+O\left(\frac{1}{\sqrt{n}}\right).$$

Thus, since

$$2^{n-1} < 2\sharp(P) = 2N \leqslant 2^n$$

then

$$\liminf d_N \geqslant \frac{3}{256\sqrt{\mathrm{e}\ln 2}}$$

and

$$\limsup d_N \geqslant \frac{1}{64\sqrt{\mathrm{e}\ln 2}}.$$

In fact,

$$d_{2^{n-1}} \geqslant \frac{1}{\sqrt{\ln(2^{n-1})}}\frac{1}{64\sqrt{\mathrm{e}}}\sqrt{n} \to \frac{1}{64\sqrt{\mathrm{e}\ln 2}}.$$ □

Acknowledgements. I would like to thank Dmitriy Bilyk and Jill Pipher for their valuable comments.

References

1. D. Bilyk, Roth's orthogonal function method in discrepancy theory and some new connections. In *Panorama of Discrepancy Theory* (eds W. W. L. Chen, A. Srivastav and G. Travaglini), Springer (2013).
2. D. Bilyk, V. Temlyakov and R. Yu, Fibonacci sets and symmetrization in discrepancy theory. *J. Complexity* **28** (2012), 18–36.
3. W. Chen, Lectures on irregularities of point distribution (2000), http://rutherglen.ics.mq.edu.au/wchen/researchfolder/iod00.pdf.
4. H. Davenport, Note on irregularities of distribution. *Mathematika* **3** (1956), 131–135.
5. H. Faure, F. Pillichshammer, G. Pirsic and W. Schmid, L_2 discrepancy of generalized two-dimensional Hammersley point sets scrambled with arbitrary permutations. *Acta Arith.* **141** (2010), 395–418.
6. G. Halasz, On Roth's method in the theory of irregularities of point distributions. In *Recent Progress in Analytic Number Theory,* Vol. 2 (Durham, 1979), Academic Press (London, 1981), 79–94.
7. A. Hinrichs and L. Markashin, On lower bounds for the L_2 discrepancy. *J. Complexity* **27**(2) (2011), 127–132.
8. Y. Katznelson, *An Introduction to Harmonic Analysis*, 2nd edition, Dover (New York, 1976).
9. I. G. Macdonald, *Symmetric Functions and Orthogonal Polynomials* (*University Lecture Series* **12**), American Mathematical Society (Providence, RI, 1998).
10. J. Matoušek, *Geometric Discrepancy* (*Algorithms and Combinatorics* **18**), Springer (Berlin, 2010).
11. K. F. Roth, On irregularities of distribution. II. *Comm. Pure Appl. Math.* **29**(6) (1976), 739–744.

Armen Vagharshakyan,
Mathematics Department, Brown University,
151 Thayer St, Providence, RI 02912, U.S.A.
E-mail: armen@math.brown.edu

https://doi.org/10.1112/S0025579312001180 Published online by Cambridge University Press