COLLOQUIA MATHEMATICA SOCIETATIS JÁNOS BOLYAI
60. SETS, GRAPHS AND NUMBERS, BUDAPEST (HUNGARY), 1991

On an Inequality of Erdős and Turán Concerning Uniform Distribution Modulo One I.

I. Z. RUZSA

1. Introduction

Let x_1, ..., x_N be a sequence of real numbers. We define the *discrepancy* Δ of this sequence by

$$\Delta = \sup_{0 \leq u \leq v < 1} \left| \frac{1}{N} \#\{1 \leq j \leq N : u \leq \{x_j\} < v\} - (v-u) \right|, \tag{1.1}$$

where $\{x\}$ denotes the fractional part of x. (Some authors use this term for the quantity $N\Delta$.) Put

$$\alpha_k = \frac{1}{N} \sum e^{2\pi i k x_j}, \tag{1.2}$$

the *Fourier coefficients* of this sequence. A famous theorem of Erdős and Turán ([1-2], reprinted in [8, pp. 432-449]) asserts that

$$\Delta \ll B = \min_k \frac{1}{k} + \sum_{j=1}^{k-1} \frac{|\alpha_j|}{j}. \tag{1.3}$$

This result found wide response in the literature. Koksma [4] gave a generalization to several dimensions. Niederreiter and Philipp [6] discuss its connections with the Berry-Esseen type inequalities of probability theory. They also give sharp values for the constant implicit in (1.3)

A different (quadratic) inequality was found by LeVeque [5]. However, H. Momtgomery observed that LeVeque's inequality can be deduced from Erdős and Turán's up to a constant factor. This led him to ask (on the CBMS regional conference on harmonic analysis and number theory, Kansas, Manhattan, May 1990) whether (1.3) is the best possible estimate of Δ if only upper estimates of $|\alpha_j|$ are known. Our aim is to show that this is essentially the case.

In Section 2 we show that the analog of (1.3) is (up to a constant factor) the best possible estimate for the discrepancy of a measure in terms of its Fourier coefficients. This does not immediately exclude the possibility that another estimate may hold for discrete measures. In Sections 3-4 we treat the possibility of approximating a continuous measure by a discrete one and prove the the non-improvability of (1.3), though in a somewhat less sharp form than for general measures. As a byproduct we solve a problem raised in [1].

Acknowledgement. I profited much from discussions with H. Montgomery.

2. The Erdős-Turán inequality for measures

Let μ and ν be distributions (probability measures) on $[0,1)$. We define their *distance* by

$$\varrho(\mu,\nu) = \sup_{0\le u\le v<1} |\mu([u,v)) - \nu([u,v))|. \tag{2.1}$$

Let Λ denote the uniform distribution (Lebesgue measure) on $[0,1)$. By the *discrepancy* of a distribution μ we mean

$$\Delta(\mu) = \varrho(\mu,\Lambda). \tag{2.2}$$

Let

$$\hat{\mu}(k) = \int e^{2\pi ikx}\, d\mu(x)$$

be the Fourier transform of μ. We have

$$\Delta(\mu) \ll B = \min_k \frac{1}{k} + \sum_{j=1}^{k-1} \frac{|\hat{\mu}(j)|}{j}. \tag{2.3}$$

(1.3) is equivalent to a special case of (2.3), namely that for

$$\mu = \frac{1}{N}\sum_{j=1}^{N}\delta(x_j),$$

where $\delta(x)$ denotes the point mass at x. Any proof of (1.3) yields also (2.3), or even it can be deduced from (1.3) by a limiting argument.

Our first result is the nonimprovability of (2.3).

Theorem 1. *Let a_1, a_2, ... be a sequence of real numbers, $0 \le a_j \le 1$. Write*

$$B = \inf_k \frac{1}{k} + \sum_{j=1}^{k-1} a_j/j.$$

There is a distribution μ on $[0,1)$ such that $|\hat{\mu}(k)| \le a_k/2$ for all $k = 1$, 2, ..., while

$$\Delta(\mu) \ge B/\pi^2.$$

Proof. We define another sequence (b_j) of numbers as follows. If $\sum_{j=1}^{\infty} a_j \le 1$, then we put $b_j = a_j$ for all j. If

$$\sum_{j=1}^{m-1} a_j \le 1 < \sum_{j=1}^{m} a_j,$$

then we put $b_j = a_j$ for $j = 1, ..., m-1$,

$$b_m = 1 - \sum_{j=1}^{m-1} a_j,$$

and $b_j = 0$ for $j > m$. In any case we have $0 \le b_j \le a_j \le 1$ for all j, and also $\sum b_j \le 1$. Next we show that

$$B_1 = \sum_{j=1}^{\infty} b_j/j \ge B/2. \tag{2.4}$$

Indeed, if $\sum a_j \le 1$, then

$$\sum b_j/j = \sum a_j/j = \lim \frac{1}{k} + \sum_{j=1}^{k-1} a_j/j \ge B.$$

If $\sum a_j > 1$, then we have $\sum_{j=1}^{m} b_j = 1$, hence $B_1 = \sum b_j/j \geq 1/m$, and also

$$\sum_{j=1}^{m-1} b_j/j = \sum_{j=1}^{m-1} a_j/j;$$

adding these inequalities we obtain

$$B \leq \frac{1}{m} + \sum_{j=1}^{m-1} a_j/j \leq 2B_1.$$

Now consider a measure μ whose density function (with respect to the Lebesgue measure) is

$$f(x) = \frac{d\mu(x)}{dx} = 1 + \sum_{j=1}^{\infty} b_j \cos(2\pi jx + t_j)$$

with certain numbers t_j, to be defined below. The inequality $\sum b_j \leq 1$ guarantees that $f(x) \geq 0$, and the Fourier transform satisfies

$$\hat{\mu}(j) = \frac{1}{2} b_j e^{it_j}, \quad |\hat{\mu}(j)| = \frac{b_j}{2} \leq \frac{a_j}{2}.$$

For any $0 \leq v < 1$ we have

$$\begin{aligned} \mu([0,v)) - v &= \int_0^v \sum b_j \cos(2\pi jx + t_j)\, dx \\ &= \frac{1}{2\pi} \sum \frac{b_j}{j} (\sin(2\pi jv + t_j) - \sin t_j) \qquad (2.5) \\ &= \frac{1}{\pi} \sum \frac{b_j}{j} \sin \pi jv \cos(\pi jv + t_j). \end{aligned}$$

Now consider the function

$$g(v) = \sum \frac{b_j}{j} |\sin \pi jv|.$$

Since

$$\int_0^1 g(v)\, dv = \frac{2}{\pi} \sum b_j/j = \frac{2}{\pi} B_1,$$

there must be a number $0 < v < 1$ such that $g(v) \geq (2/\pi)B_1$. Take this value of v and choose t_j so that

$$\cos(\pi jv + t_j) = \operatorname{sgn} \sin \pi jv.$$

With this choice of the parameters (2.5) yields

$$\mu([0,v)) - v = \frac{1}{\pi} g(v) \geq \frac{2}{\pi^2} B_1,$$

consequently $\Delta(\mu) \geq (2/\pi^2)B_1 \geq B/\pi^2$ as claimed. ∎

3. Approximating a distribution by a discrete one

The distribution we constructed in Section 2 is continuous. Here we find for any distribution μ a sequence of N points whose distribution is near to μ in the distance ϱ and whose Fourier coefficients are also near to those of μ.

Theorem 2. *Let μ be a distribution on $[0,1)$. For every positive integer N there are real numbers x_1, ..., x_N such that their distribution $\nu = N^{-1}\sum\delta(x_j)$ satisfies*

$$\varrho(\mu,\nu) \le 1/N \tag{3.1}$$

and

$$|\hat{\mu}(k)-\hat{\nu}(k)| \le \sqrt{\log(k+1)}\min\left(\frac{13\sqrt{k}}{N}, \frac{9}{\sqrt{N}}\right) \tag{3.2}$$

for all k. Moreover, if the measure μ satisfies

$$\mu([u,v)) \ge \beta(v-u) \tag{3.3}$$

for all $0 \le u \le v < 1$ with some $1/N < \beta \le 1$, then we also have

$$|\hat{\mu}(k)-\hat{\nu}(k)| \le \sqrt{\log(k+1)}\frac{22k}{\sqrt{\beta}N^{3/2}}. \tag{3.4}$$

Remark 3.1. $|\hat{\mu}(k)-\hat{\nu}(k)| \ll k/N$ follows from (3.1). (3.2) is much sharper for large values of k. (3.4) improves (3.2) for $k < \beta N/3$.

For the proof we need a probabilistic inequality.

Lemma 3.2. *Let ξ_1, ..., ξ_n be independent real random variables satisfying $\mathbf{E}\xi_j = 0$ and $|\xi_j| \le s_j$. Write $\eta = \sum\xi_j$, $S = \sum s_j^2$. For every $A > 0$ we have*

$$\mathbf{P}(|\eta| \ge A) \le 2\exp -\frac{A^2}{2S}. \tag{3.5}$$

See Hoefding [3].

Lemma 3.3. *Let ξ_1, ..., ξ_n be independent complex random variables satisfying $\mathbf{E}\xi_j = 0$ and $|\xi_j| \le s_j$. Write $\eta = \sum\xi_j$, $S = \sum s_j^2$. For every $A > 0$ we have*

$$\mathbf{P}(|\eta| \ge A) \le 4\exp -\frac{A^2}{4S}. \tag{3.6}$$

Proof. If $|\eta| \geq A$, then either $|\mathrm{Re}\,\eta| \geq A/\sqrt{2}$ or $|\mathrm{Im}\,\eta| \geq A/\sqrt{2}$. Applying the previous lemma for the real and imaginary parts of our variables we obtain (3.6). ■

Proof. of Theorem 2. Let $0 = t_0 \leq t_1 \leq \dots \leq t_N = 1$ be the N quantiles of μ, that is, t_j is a number such that

$$\mu([0,t_j)) \leq \frac{j}{N} \leq \mu([0,t_j]).$$

We shall select one x_j from each $[t_{j-1}, t_j]$; this already guarantees (3.1).

We can represent μ in the form

$$\mu = \frac{\mu_1 + \dots + \mu_N}{N},$$

where each μ_j is supported in the corresponding $[t_{j-1}, t_j]$. Let $X_1, \dots, X_N$ be independent real random variables, X_j with distribution μ_j. We show that with positive probability $x_j = X_j$ satisfies (3.2) and (3.4).

The Fourier coefficients of ν are

$$\hat{\nu}(k) = \frac{1}{N} \sum_{j=1}^{N} e^{2\pi i k X_j},$$

also random variables. Their expectations are

$$\mathbf{E}\hat{\nu}(k) = \frac{1}{N} \sum_j \mathbf{E} e^{2\pi i k X_j} = \frac{1}{N} \sum_j \int e^{2\pi i k x}\, d\mu_j(x)$$
$$= \int e^{2\pi i k x}\, d\mu(x) = \hat{\mu}(k).$$

To estimate the probability that $\hat{\nu}(k) - \hat{\mu}(k)$ is big we apply Lemma 3.3 to the variables

$$\xi_j = e^{2\pi i k X_j} - \mathbf{E} e^{2\pi i k X_j}.$$

Write $s_j = \max|\xi_j|$, $S = S_k = \sum s_j^2$. In the notations of Lemma 3.3 we have $\hat{\nu}(k) - \hat{\mu}(k) = \eta/N$, thus with arbitrary positive A_k (3.3) yields

$$\mathbf{P}(|\hat{\nu}(k) - \hat{\mu}(k)| \geq A_k) \leq 4 \exp - \frac{N^2 A_k^2}{4 S_k}.$$

With the choice

$$A_k = \frac{1}{N} \sqrt{12 S_k \log(k+1)} \tag{3.7}$$

the right side becomes $4/(k+1)^3$, thus we obtain

$$\sum_k \mathbf{P}(|\hat{\nu}(k) - \hat{\mu}(k)| \geq A_k) \leq \sum_{k=1}^{\infty} \frac{4}{(k+1)^3} < 1.$$

Consequently with positive probability all the inequalities

$$|\hat{\nu}(k) - \hat{\mu}(k)| \leq A_k, \quad k = 1, 2, \ldots, \tag{3.8}$$

hold with the A_k given by (3.7).

To deduce (3.2) and (3.4) we need to estimate S_k.

By definition we have obviously $s_j \leq 2$. If $k(t_j - t_{j-1}) < 1/2$, then $\exp 2\pi i k X_j$ is contained in an arc of length $2\pi k(t_j - t_{j-1})$ of the unit circle. The expectation must be in the convex hull of this arc, which is the corresponding segment. Since the diameter of this segment is

$$2 \sin \pi k(t_j - t_{j-1}) \leq 2\pi k(t_j - t_{j-1}),$$

we conclude that

$$s_j \leq 2 \min(1, \pi k(t_j - t_{j-1})).$$

Since a minimum is less than a geometric mean, we have

$$\sum s_j^2 \leq 4 \sum \pi k(t_j - t_{j-1}) = 4\pi k.$$

Consequently $S_k \leq \min(4\pi k, 2N)$, and after substituting this into (3.7), (3.8) yields (3.2).

Assume now that (3.3) holds. Then we have

$$1/N \geq \mu((t_{j-1}, t_j)) \geq \beta(t_j - t_{j-1}),$$

that is, $t_j - t_{j-1} \leq 1/(\beta N)$. This implies

$$s_j^2 \leq (2\pi k)^2 (t_j - t_{j-1})^2 \leq \frac{(2\pi k)^2}{\beta N}(t_j - t_{j-1}),$$

hence $S \leq (2\pi k)^2/(\beta N)$. Substituting this into (3.7), (3.8) yields (3.4). ■

4. The Erdős-Turán inequality for sequences

A formal analog of Theorem 1 for sequences would assert the existence of points $x_1, \ldots, x_N$ whose Fourier coefficients α_k defined in (1.2) satisfy $|\alpha_k| \le a_k$ while the discrepancy is not much less than the bound B given in (1.3). Such a result cannot hold, since the system of inequalities $|\alpha_k| \le a_k$ may not have a solution at all for very small a_k; among n consecutive values there must be one which is not too small. This observation is the starting point of Turán's theory of power sums and we refer to his book [7] for the details.

We give a result which is, however, in all practical cases as strong as Theorem 1.

Theorem 3. *Let $a_1, a_2, \ldots$ be a sequence of real numbers, $0 \le a_j \le 1$. Write*

$$B = \inf_k \frac{1}{k} + \sum_{j=1}^{k-1} a_j/j.$$

There are N numbers $x_1, \ldots, x_N$ in $[0,1)$ such that their Fourier coefficients α_k satisfy

$$|\alpha_k| \le \frac{a_k}{4} + 32\sqrt{\log(k+1)} \min\left(\frac{k}{N^{3/2}}, \frac{1}{\sqrt{N}}\right) \tag{4.1}$$

for all $k = 1, 2, \ldots$, while their discrepancy is

$$\Delta \ge B/40. \tag{4.2}$$

Proof. Let μ be the measure obtained by applying Theorem 1 to the sequence (a_k). Now we apply Theorem 2 to the measure $\mu' = (\mu + \Lambda)/2$ (recall that Λ is the Lebesgue measure on $[0,1)$). We have

$$\Delta(\mu') = \Delta(\mu)/2 \ge B/20$$

and, denoting the distribution of of the points $x_1, \ldots, x_N$ by ν we have $\varrho(\nu, \mu') \le 1/N$, hence

$$\Delta = \Delta(\nu) = \varrho(\nu, \Lambda) \ge \varrho(\mu', \Lambda) - \varrho(\mu', \nu) = \Delta(\mu') - \varrho(\mu', \nu) \ge \frac{B}{20} - \frac{1}{N}.$$

This implies (4.2) if $B \geq 40/N$, while if $B < 40/N$ then (4.2) holds trivially, since the minimal possible value of Δ is $1/N$.

The measure μ' satisfies (4.3) with $\beta = 1/2$, hence (4.1) follows from (4.1) and (4.4), taking into account that

$$|\hat{\mu}'(k)| = \frac{1}{2}|\hat{\mu}(k)| \leq \frac{a_k}{4}. \qquad \blacksquare$$

The error term in (4.1) is much smaller than any conceivable estimate one can obtain for α_k, thus its seems safe to conclude that in any realistical situation Erdős and Turán's bound gives the correct order of magnitude.

In particular, in [1] Erdős and Turán ask the following. Assume that the unnormalized sums $N\alpha_k$ satisfy

$$|N\alpha_k| \leq k^\lambda \quad \text{for} \quad 1 \leq k \leq N^{1/(\lambda+1)} \tag{4.3}$$

with some $\lambda \geq 1$. Then (1.3) yields the estimate

$$\Delta \ll N^{-1/(\lambda+1)} ;$$

is it the best possible? Our Theorem 3 implies that it is, even for $\lambda > 1/3+\varepsilon$ with some $\varepsilon > 0$ if $N > N_0(\varepsilon)$, and if $\lambda > 1/2 + \varepsilon$, then the range of (4.3) can be extended to all positive integers.

The following case is not covered by our result. Assume that $\alpha_1 = \dots = \alpha_{k-1} = 0$ for some $k < N$; how small can the discrepancy be? (1.3) gives $\Delta \ll 1/k$, and this is best possible. To see this, consider the following construction. Let E_k denote the system of k points $0, 1/k, 2/k, \dots, (k-1)/k$. Write $N = kq + r$ with $0 \leq r < k$, and take $q - 1$ copies of E_k plus one copy of E_{k+r}. (This argument also solves the previous problem, even for $\lambda \leq 1/3$, but does not give the extended range for k.)

I cannot solve the following. Assume that $\alpha_2 = \dots = \alpha_k = 0$, but no assumption is made about α_1. (1.3) gives only the trivial estimate; can anything better be proved? Or assume that $\alpha_j = 0$ for all $j \leq N$ except when j is a prime; how large can Δ be?

References

[1] P. Erdős and P. Turán, On a problem in the theory of uniform distribution I., *Indag. Math.* **10**(1948), 370–378.

[2] P. Erdős and P. Turán, On a problem in the theory of uniform distribution II., *Indag. Math.* **10**(1948), 406–413.

[3] W. Hoefding, Probability inequalities for sums of bounded random variables, *J. Amer. Stat. Assoc.* **58**(1963), 13–30.

[4] J. F. Koksma, *Some theorems on Diophantine inequalities*, Math. Centrum Scriptum No. 5, Amsterdam 1950.

[5] W. J. LeVeque, An inequality connected with Weyl's criterion for uniform distribution, in: *Proc. Symp. Pure Math., Vol. VIII.*, Amer. Math. Soc. , Providence, R. I. (1965), 22–30.

[6] H. Niederreiter and W. Philipp, Berry-Esseen bounds and a theorem of Erdős and Turán on uniform distribution mod 1, *Duke Math. J.* **40**(1973), 633–649.

[7] P. Turán, *On a new method of analysis and its applications*, Wiley & Sons, New York – Chichester 1984.

[8] P. Turán, *Collected Papers*, (ed.: P. Erdős), Akadémiai Kiadó, Budapest 1990.

Imre Z. Ruzsa

Mathematical Institute of the
Hungarian Academy of Sciences
H-1364 Hungary
Budapest, P.O.B. 127