BULLETIN *of the* MALAYSIAN MATHEMATICAL SCIENCES SOCIETY

Bull. Malaysian Math. Sc. Soc. (Second Series) **24** (2001) 59-67

A Metric Discrepancy Estimate in Higher Dimensions Using L^2 Methods

HAILIZA KAMARUL HAILI
School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM Pulau Pinang, Malaysia
e-mail: hailiza@cs.usm.my

Abstact. We prove a general metrical result, which contains as a special case a discrepancy estimate, related to the uniform distribution of expressions of the form $(\lambda_{1,i,n}x_{1,i} + \cdots + \lambda_{k(i),i,n}x_{k(i)})_{n=1}^{\infty}$ $(i = 1, \cdots, d;\ 1 \le j \le k(i))$ for Lebesgue almost all points in $X = X_1 \times \cdots \times X_d$ with $X_i = [a_{1,i}, b_{1,i}] \times \cdots \times [a_{k(i),i}, b_{k(i),i}]$ where for each pair j, i there exist $c_{j,i} > 0$ such that $\left|\lambda_{j,i,(n+1)} - \lambda_{j,i,n}\right| \ge c_{j,i}$.

1. Introduction

For a positive integer *d*, fixed throughout this paper, let [*d*] denote the set of the first *d* natural numbers and let ***Z+*** denote the set of natural numbers. To each *i* in [*d*] we associate a natural number *k*(*i*) and a collection of *k*(*i*) intervals

$$[a_{1,i}, b_{1,i}], \cdots, [a_{k(i),i},\ b_{k(i),i}]$$

such that $-\infty \le a_{ji} < b_{ji} \le \infty$ and $j \in [k(i)]$. For $\{x\}$, we denotes the fractional part of *x*. To each interval $[a_{j,i}, b_{j,i}]$ we associate a sequence of continuously differentiable real functions $(g_{j,i,n}(x_{j,i}))_{n=1}^{\infty}$ defined for all *n* such that:

(i) for each pair m, n with $m \ne n$ the functions $g'_{j,i,m}(x_{j,i}) - g'_{j,i,n}(x_{j,i})$ are monotone on $[a_{j,i}, b_{j,i}]$; and

(ii) for each pair j, i there exist $c_{j,i} > 0$ such that

$$\inf_{\substack{m,n\in \mathbf{Z}+ \\ m\neq n}} \inf_{x_{j,i}\in[a_{j,i},b_{j,i}]} \left| g'_{j,i,n}(x_{j,i}) - g'_{j,i,m}(x_{j,i}) \right| \geq c_{j,i} .$$

Now to each $i \in [d]$ we associate the sequence of functions

$$g_{i,n}(x_{1,i}, \cdots, x_{k(i),i}) = \sum_{j=1}^{k(i)} g_{j,i,n}(x_{j,i})$$

Then for $x \in X$, where *X* denotes the Cartesian product $X_1 \times \cdots \times X_d$ with

$$X_i = [a_{1,i}, b_{1,i}] \times \cdots \times [a_{k(i),i}, b_{k(i),i}], \qquad (i \in [d])$$

we set

$$g_n(x) = \left(\left\{ g_{1,n}(x_{1,1,}, \cdots, x_{k(1),1}) \right\}, \cdots, \left\{ g_{d,n}(x_{1,d}, \cdots, x_{k(d),d}) \right\} \right) \qquad (n \in \mathbf{Z}^+)$$

In this paper we are interested in distribution of the sequence $(g_n(x))_{n=1}^{\infty}$ on $[0,1)^d$ for almost all x with respect to *d*-dimensional Lebesgue measure on *X*.
For a finite set of *d*-tuples $y_1, \cdots, y_N$ in $[0,1)^d$, let

$$D(y_1, \cdots, y_N) = \sup_{R\subseteq[0,1)^d} \left| \frac{1}{N} \sum_{i=1}^{N} \chi_R(y_i) - |R| \right|,$$

that is the *d*-dimensional *discrepancy* of $y_1, \cdots, y_N$ [7]. Here the supremum is taken over all *d*-dimensional rectangles *R*, contained in $[0,1)^d$ such that

$$R = J_1 \times \cdots \times J_d ,$$

where for each $i \in [d]$, J_i is a set interval of $[0,1)$, closed on the left and open on the right. For a set $B \subseteq [0,1)^d$, χ_B denotes its characteristic function and $|B|$ denotes its Lebesgue measure, assuming that is defined. In the case $d = 1$ this reduces to the one dimensional notion of discrepancy defined earlier. If for $x \in X$, we choose $y_j = g_j(x)$, then set

$$D(M, N, x) = D(yM = 1, \cdots, yM + n) \qquad (M \in \mathbf{Z}_{\geq 0};\ N \in \mathbf{Z}^+)$$

$$D(y_{M+1}, \cdots, y_{M+N})$$

and let

$$D(N, x) = D(0, N, x).$$

Here $\boldsymbol{Z}^+$, denotes the non-negative integers. Now we prove the following theorem

Theorem 2.1. *Suppose X,* $(g_n(x))_{n=1}^{\infty}$ and $D(N, x)$ $(N \in \boldsymbol{Z}^+)$ *are defined as above. Then given* $\varepsilon > 0$

$$D(N, x) \ = \ o(N^{-\frac{1}{2}} (\log N)^{\frac{3}{2}+d+\varepsilon}) \ a.e.$$

A proof of this result in the case $d = 1$ and $k(i) = 1$ appears in [1] though by a different method. This is the culmination of work going back to [6] and [5]. If we set $k(i) = k$ $(i = 1, \cdots, d)$ and $g_{j,i,n}(x_{j,i}) = \lambda_{j,i,n} x_{j,i}$ where

$$\left| \lambda_{j,i,n+1} - \lambda_{j,i,n} \right| \ \geq \ c_{j,i} \ .$$

Then Theorem 2.1 covers the case

$$g_n(x) = \Lambda_n x,$$

where Λ_n is the matrix $d \times k$ matrix $(\lambda_{j,i,n})$

Proof of Theorem 2.1.

The proof of Theorem 2.1 needs a number of lemmas some of which we now state. The first is due to I.S. Gál and J.F. Koksma [4].

Lemma 2.2. *For each* M *and* N $(M, N \in \boldsymbol{Z}^+)$, *let* $F(M, N, x)$ *denote a Lebesgue measurable function on X. Suppose for all integers* $L \in [0, N]$ *that*

$$\left| F(M, N, x) \right| \ \leq \ \left| F(M, L.x) \right| \ + \ \left| F(M + L, N - L, x) \right|.$$

Suppose further for each $\theta > \frac{1}{2}$ and $\phi \geq 0$ that we have

$$\left\| F(M, N, x) \right\| \ = \ O(N^{\theta} (M + N)^{\phi}).$$

Then given $\varepsilon > 0$

$$F(0, N, x) = o(N^{\theta+\phi} (\log N)^{\frac{1}{2}+\varepsilon}) \ \ a.e.$$

For real valued functions *f*, we use $\|f\|$ to denote $(\int_X |f|^2 dx)^{\frac{1}{2}}$. Lemma 2.2 reduces the proof of Theorem 2.1 to proving L^2 estimates for $D(M, N, x)$. To provide these we use the next lemma due in the case $d = 1$ to P. Erdös and P. Turan [2] and to J.F. Koksma [5] for general *d*. Henceforth, for real *x*, let $e(x)$ denote $e^{2\pi ix}$.

Lemma 2.3. *For* $h = (h_1, \cdots, h_d) \in \boldsymbol{Z}^d$ (*d-tuple of integers*), *let*

$$r(h) = \prod_{i=1}^{d} \max(1, |h_i|)$$

and let

$$M(h) = \max_{1 \le i \le d} |h_i|.$$

There exists a constant $C > 0$ *such that given* $y_1, \cdots, y_N \in [0, 1)^d$, *then for any positive integer L,*

$$N\ D(y_1, \cdots, y_N) \le C \left(\frac{N}{L} + \sum_{\substack{0 < M(h) \le L \\ h \in \boldsymbol{Z}^d}} \frac{1}{r(h)} \left| \sum_{n=1}^{N} e(< h, y_n >) \right| \right),$$

where $<,>$ *denotes the standard inner product on* $\Re^d$.

To use Lemma 2.2 to deduce Theorem 2.1 we need an estimate in terms of *M* and *N* for the L^2 norm of $D(M, N, x)$.

Let

$$S_h(M, N, x) = \sum_{n=M+1}^{M+N} (< h, g_n(x) >).$$

Then, using the L^2 triangle inequality

$$\| N\, D(M,N,x) \| \le C\left(\frac{N}{L} + \sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d}} \frac{1}{r(h)} \| S_h(M,N,x) \| \right).$$

So we need to estimate $\| S_h(M,N,x) \|$. Now

$$\| S_h(M,N,x) \|^2 = \int_X \left(\sum_{n=M+1}^{M+N} \sum_{m=M+1}^{M+N} e(< h, g_n(x) - g_m(x) >) \right) dx$$

$$= \sum_{n=M+1}^{M+N} \sum_{m=M+1}^{M+N} \int_X e(< h, g_n(x) - g_m(x) >)\, dx.$$

From the definitions of X, h, $(g_n(x))_{n=1}^{\infty}$ and dx we see that

$$\int_X e(< h, g_n(x) - g_m(x) >)\, dx = \prod_{i=1}^{d} \prod_{j=1}^{k(i)} \left(\int_{a_{j,i}}^{b_{j,i}} e(h_i (g_{j,i,n}(x_{j,i}) - g_{j,i,m}(x_{j,i})))\, dx_{j,i} \right).$$

Let

$$d_{m,n,j,i} = \inf_{x_{j,i} \in X} \left| g'_{j,i,m}(x_{j,i}) - g'_{j,i,n}(x_{j,i}) \right|$$

for $m \ne n$ and let

$$d^*_{m,n,j,i,h} = \begin{cases} 1 & \text{if } \quad m = n \text{ or } h_i = 0 \\ |h_i|\, d_{m,n,j,i} & \text{otherwise} \end{cases} \tag{1}$$

Thus

$$\| S_h(M,N,x) \|^2 \le C \sum_{n=M+1}^{M+N} \sum_{m=M+1}^{M+N} \prod_{i=1}^{d} \prod_{j=1}^{k(i)} \left(\frac{1}{d^*_{m,n,j,i,h}} \right),$$

because

$$\left| \int_{a_{i,j}}^{b_{i,j}} e(h_i(g_{j,i,n}(y) - g_{j,i,m}(y)))\, dy \right| = \left| \frac{1}{2\pi i h_i} \int_{a_{j,i}}^{b_{j,i}} \frac{d(e(h_i(g_{j,i,n}(y) - g_{j,i,m}(y))))}{(g'_{j,i,n}(y) - g'_{j,i,m}(y))} \right|$$

the right hand side of which by the second mean value theorem is

$$\leq \left(\frac{C}{d^*_{m,n,j,i,h}} \right)$$

Using the arithmetic-geometric mean inequality

$$\prod_{i=1}^{d} \prod_{j=1}^{k(i)} \left(\frac{1}{d^*_{m,n,j,i,h}} \right) \leq \sum_{i=1}^{d} \sum_{j=1}^{k(i)} \frac{1}{d^*_{m,n,j,i,h}},$$

and hence there exists $C > 0$ such that

$$\| S_h(M, N, x) \|^2 \leq C \sum_{n=M+1}^{M+N} \sum_{m=M+1}^{M+N} \sum_{i=1}^{d} \sum_{j=1}^{k(i)} \left(\frac{1}{d^*_{m,n,j,i,h}} \right).$$

Also

$$\sum_{n=M+1}^{M+N} \sum_{m=M+1}^{M+N} \left(\frac{1}{d^*_{m,n,j,i,h}} \right) \leq \sum_{i=1}^{d} \sum_{j=1}^{k(i)} \left(N + \frac{2}{|h_i|} \sum_{M+1 \leq n < m \leq M+N} \left(\frac{1}{d_{m,n,j,i,h}} \right) \right).$$

Because of the conditions (i) and (ii) on the functions $g'_{j,i,m}(x) - g'_{j,i,n}(x)$ we can find a permutation $\pi_1, \cdots, \pi_N$ of the integers $M+1, \cdots, M+N$ such that $g'_{j,i,\pi_a}(x) \leq g'_{j,i,\pi_b}(x)$ if and only if $a \leq b$. This means that if $b > a$ then

$$g'_{j,i,\pi_b}(x) - g'_{j,i,\pi_a}(x) = \sum_{k=a}^{b-1} (g'_{j,i,\pi_{k+1}(x)} - g'_{j,i,\pi_k}(x)) \geq c_{j,i}(b-a).$$

Thus

$$\sum_{M+1\le n<m\le M+N}\left(\frac{1}{d_{m,n,j,i,h}}\right)\le\frac{1}{c_{i,j}|h_i|}\sum_{1\le a<b\le N}\left(\frac{1}{b-a}\right)$$

and as

$$\sum_{1\le a<b\le N}\frac{1}{b-a}\le C\,N\log N$$

we know that

$$\|S_h(M,N,x)\|^2\le C\,N\left(1+\frac{\log N}{|h_i|}\right).\qquad(2)$$

In consequence

$$\|N\,D\,(M,N,x)\|\le C\left(\frac{N}{L}+\sum_{\substack{0<M\,(h)\le L\\ h\in\mathbf{Z}^d}}\frac{1}{r(h)}N^{\frac{1}{2}}\left(1+(\log N)\sum_{i=1}^{d}|h_i|^{-\frac{1}{2}}\right)\right),$$

so estimating $\|N\,D\,(M,N,x)\|$ reduces to estimating

$$\text{(i)}\qquad\sum_{\substack{0<M\,(h)\le L\\ h\in\mathbf{Z}^d}}\frac{1}{r(h)}\qquad(3)$$

and

$$\text{(ii)}\qquad\sum_{\substack{0<M\,(h)\le L\\ h\in\mathbf{Z}^d}}\frac{1}{r(h)}\left(\sum_{i=1}^{d}|h_i|^{-\frac{1}{2}}\right)\qquad(4)$$

We estimate (i) and (ii) by arguing as follows. Let $r'(h) = \Pi_{i=1}^{d} \max(1, |h_i|)$. Then by considerations of symmetry

$$\sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d}} \frac{1}{r(h)} \le 2^d \sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d_{\ge 0}\setminus\{0\}}} \frac{1}{r'(h)}.$$

Let $2_k^{[d]} (k = 1, \cdots, N)$ denote the collection of subsets of $[d]$ contain exactly k elements. Now note that

$$\sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d_{\ge 0}\setminus\{0\}}} \frac{1}{r'(h)} = \sum_{k=1}^{d} \sum_{\tau\in 2_k^{[d]}} \left(\sum_{\substack{i_1=1,\cdots,i_k=1 \\ \tau=\{i_1,\cdots,i_k\}}}^{L} \cdots \sum^{L} \frac{1}{h_{i_1} \cdots h_{i_k}} \right)$$

and that

$$\sum_{\substack{i_1=1,\cdots,i_k=1 \\ \tau=\{\tau_1,\cdots,\tau_k\}}}^{L} \cdots \sum (h_{i_1} \cdots h_{i_k})^{-1} = \left(\sum_{n=1}^{L} n^{-1} \right)^k.$$

Hence as $\#(2_k^{[d]} = \binom{d}{k}$ we see that

$$\sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d_{\ge 0}\setminus\{0\}}} \frac{1}{r'(h)} = \sum_{k=0}^{d} \binom{d}{k} \left(\sum_{n=1}^{L} n^{-1} \right)^k - 1$$

$$= \left(1 + \sum_{j=1}^{L} j^{-1} \right)^d - 1,$$

showing

$$\sum_{\substack{0<M(h)\le L \\ h\in \mathbf{Z}^d}} \frac{1}{r(h)} \le C(\log L)^d.$$

Now

$$\sum_{\substack{0<M(h)\le L\\ h\in \mathbf{Z}^d}} \frac{1}{r(h)}\left(\sum_{i:h_i\neq 0}|h_i|^{-\frac{1}{2}}\right)\le 2^d \sum_{\substack{0<M(h)\le L\\ h\in \mathbf{Z}^d}} \frac{1}{r'(h)}\left(\sum_{i:h_i\neq 0} h_i^{-\frac{1}{2}}\right)$$

$$= 2^d\left(\sum_{k=1}^{d}\sum_{\tau\in 2_k^{[d]}}\left(\sum_{\substack{i_1=1,\cdots,i_k=1\\ \tau=\{i_1,\cdots,i_k\}}}^{L}\cdots\sum^{L}(h_{i1}\cdots h_{ik})^{-1}\right)\left(\sum_{i:h_i\neq\{0\}}|h_i|^{-\frac{1}{2}}\right)\right)$$

Thus

$$\sum_{\substack{0<M(h)\le L\\ h\in \mathbf{Z}^d}} \frac{1}{r(h)}\left(\sum_{i=1}^{d}|h_i|^{-\frac{1}{2}}\right)\le C(\log L)^{d-1},$$

combining these estimates we obtain

$$\| N\,D(M,N,x)\| \le N^{\frac{1}{2}}(\log N)^d ,$$

which in view of Lemma 2.2 proves Theorem 2.1.

References

1. J.W.S. Cassels, Some metrical theorem in diophantine approximation III, *Proc. Camb. Phil. Soc.* **46** (1950), 214-225.
2. P. Erdös and P. Turan, On a problem in the theory of uniform distribution I, *Indag Math.* **10** (1948), 370-378.
3. P. Erdös and P. Turan, On a problem in the theory of uniform distribution II, *Indag Math.* **10** (1948), 406-413.
4. I.S. Gàl and J.F. Koksma, Sar l'order de grandeur des fonstions summable, *Proc. K. Ned. Akad. Wet.* **53** (1950), 638-639.
5. J.F. Koksma, En mengenttheoretischer Satz über de Gleichverteilung modulo Eins, *Compositio Math.* **2** (1935), 250-258.
6. H. Weyl, Über die Gleichverteilung von Zahler mod. Eins, *Maths. Anna*. **77** (1916), 313-352.
7. L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.