1-1-2000

A LeVeque-type Lower Bound for Discrepancy

Francis E. Su
Harvey Mudd College

reprinted from *Monte Carlo and Quasi-Monte Carlo Methods 1998*,
H. Niederreiter and J. Spanier, eds., Springer-Verlag, 2000, pp. 448-458.

A LeVeque-type lower bound for discrepancy

Francis Edward Su

Department of Mathematics, Harvey Mudd College, Claremont CA 91711, USA
su@math.hmc.edu

Abstract. A sharp lower bound for discrepancy on $\mathbf{R}/\mathbf{Z}$ is derived that resembles the upper bound due to LeVeque. An analogous bound is proved for discrepancy on $\mathbf{R}^k/\mathbf{Z}^k$. These are discussed in the more general context of the discrepancy of probablity measures. As applications, the bounds are applied to Kronecker sequences and to a random walk on the torus.

1 Introduction

Consider a sequence $\{\mathbf{x}_n\}_{n\geq 1}$ of points in $\mathbf{R}^k$ modulo 1, i.e., in the k-dimensional torus $\mathbf{T}^k = \mathbf{R}^k/\mathbf{Z}^k$. Let $J = [a_1, b_1) \times ... \times [a_k, b_k)$ be an *interval* in $\mathbf{R}^k$ such that $0 < b_i - a_i \leq 1$ for $i = 1, ..., k$. Let $I = J/\mathbf{Z}^k$ be the corresponding set on the torus, also called an *interval*. The *volume* λ_k of I is inherited from Lebesgue measure in $\mathbf{R}^k$: $\lambda_k(I) = \prod_{i=1}^{k}(b_i - a_i)$. Let $A_N(I)$ denote the number of points of the first N elements of the sequence $\{\mathbf{x}_n\}$ that fall in I, i.e., $A_N(I) = \sum_{n=1}^{N} \delta_I(\mathbf{x}_n)$ where δ_I denotes the indicator function of the set I. The *discrepancy* $D_N(\mathbf{x}_n)$ of the sequence $\{\mathbf{x}_n\}$ is defined by

$$D_N(\mathbf{x}_n) = \sup_{I\subseteq \mathbf{T}^k} \left| \frac{A_N(I)}{N} - \lambda_k(I) \right|. \tag{1}$$

Loosely speaking, this is the largest possible difference between the proportion of points in an interval I and the expected proportion if the points were uniformly distributed. Hence discrepancy is a measure of how far the sequence is from being uniformly distributed in $\mathbf{T}^k$.

The purpose of this paper is to establish new lower bounds for the discrepancy of a sequence in $\mathbf{T}^k$ using Fourier methods. These are similar in form to the upper bound established by LeVeque [7]. We illustrate the use of such bounds in several examples at the end of this paper.

Our results will be cast in the more general form of the discrepancy of probability measures, which we define presently. If P is a probability measure on $\mathbf{T}^k$, define the discrepancy of P to be

$$D(P) = \sup_{I\subseteq \mathbf{T}^k} |P(I) - \lambda_k(I)|.$$

This notion of discrepancy measures the uniformity of the distribution P. Hence discrepancy of sequences is the special case where $P = \frac{A_N}{N}$.

We work with the more general notion of discrepancy of probability measures for two reasons: (1) our results are derived by viewing $A_N(I)$ as a convolution of measures, rather than counting points in a set I, and (2) discrepancy of probabilities is useful in its own right as a metric for quantifying rates of convergence of random walks on $\mathbf{T}^k$ to the uniform distribution. We give an example at the conclusion of this paper. See [9] and [10] for further work in this direction.

Let $e(z)$ denote the expression $\exp(2\pi i z)$. The Fourier coefficients of a measure Q on $\mathbf{R}/\mathbf{Z}$ are indexed by integers $m \in \mathbf{Z}$ and defined by $\widehat{Q}(m) = \int_0^1 e(mx)\, Q(dx)$, a Riemann-Stieltjes integral with respect to the distribution function of the measure Q. For a sequence $\{x_n\}$, the Fourier coefficients become $\widehat{Q}(m) = \frac{1}{N}\sum_{i=1}^{N} e(mx_i)$. In [7], LeVeque proved the following well-known upper bound for the discrepancy of sequences mod 1.

LeVeque's Inequality. *If Q is any probability measure on $\mathbf{R}/\mathbf{Z}$, then*

$$D(Q) \;\leq\; \left(\frac{6}{\pi^2}\sum_{m=1}^{\infty}\frac{|\widehat{Q}(m)|^2}{m^2}\right)^{1/3}$$

where $\widehat{Q}(m)$ denotes the m-th Fourier coefficient of Q. The constant $\frac{6}{\pi^2}$ and exponent $1/3$ are best possible.

In this paper we establish the following lower bound for discrepancy on $\mathbf{R}/\mathbf{Z}$, which resembles LeVeque's inequality in form, with sharp constant.

Theorem 1. *If Q is any probability measure on $\mathbf{R}/\mathbf{Z}$, then*

$$D(Q) \;\geq\; \left(\frac{2}{\pi^2}\sum_{m=1}^{\infty}\frac{|\widehat{Q}(m)|^2}{m^2}\right)^{1/2}$$

where $\widehat{Q}(m)$ denotes the m-th Fourier coefficient of Q. The inequality is sharp and the constant $\frac{2}{\pi^2}$ is best possible.

A similar inequality is implied in [6, Lemma 2.8] for the related star discrepancy (which is smaller); with a little extra effort we have bounded the usual discrepancy and sharpened up the constant.

Using a related, though slightly different approach, we can establish an analogous result in higher dimensions. On $\mathbf{T}^k$, the Fourier coefficients are indexed by $\mathbf{h} \in \mathbf{Z}^k$ and are given by $\widehat{Q}(\mathbf{h}) = \int_{\mathbf{T}^k} e(\mathbf{h}\cdot\mathbf{x})\, Q(d\mathbf{x})$. For sequences, this becomes $\widehat{Q}(\mathbf{h}) = \frac{1}{N}\sum_{i=1}^{N} e(\mathbf{h}\cdot\mathbf{x}_i)$.

Theorem 2. *If Q is any probability measure on $\mathbf{T}^k$, then*

$$D(Q) \;\geq\; \sup_{\mathbf{r}} \left[\sum_{\mathbf{0}\neq\mathbf{h}\in\mathbf{Z}^k} |\widehat{Q}(\mathbf{h})|^2 \prod_{i=1}^{k} \begin{Bmatrix} \frac{\sin^2(2\pi h_i r_i)}{\pi^2 h_i^2} & \text{if } h_i \neq 0 \\ 4r_i^2 & \text{if } h_i = 0 \end{Bmatrix}\right]^{1/2}$$

where $\mathbf{r} = (r_1, ..., r_k)$ *is any* k*-tuple such that each* $r_i \in (0, 1/2]$.

Note that $\mathbf{r}$ may be chosen so as to optimize the bound in relation to $\widehat{Q}$, i.e., one may choose the r_i so that $\sin(2\pi h_i r_i)$ is large when $\widehat{Q}(\mathbf{h})$ is large. To see how Theorem 2 compares with Theorem 1, set each $r_i = \frac{1}{4}$ in Theorem 2. Then terms for all lattice points $\mathbf{h}$ that have any non-zero even coordinate vanish. This yields

Corollary 3. *If* Q *is any probability measure on* $\mathbf{T}^k$*, then*

$$D(Q) \ \geq \ \frac{1}{\pi^k} \left(\sum_{\mathbf{h} \neq \mathbf{0}, \ h_i \ odd \ or \ 0} \frac{|\widehat{Q}(\mathbf{h})|^2}{R(\mathbf{h})^2} \right)^{1/2}$$

where $R(\mathbf{h}) = \prod_{i=1}^k \max\{1, |h_i|\}$ *for* $\mathbf{h} = (h_1, ..., h_k) \in \mathbf{Z}^k$.

This corollary shows that the bound of Theorem 2 is not as sharp as Theorem 1 when $k = 1$. In particular, Corollary 3 for $k = 1$ yields a bound much like Theorem 1, except that it omits all the even terms in the sum. (However, by choosing $\mathbf{r}$ appropriately in Theorem 2, one can obtain a bound like Corollary 3 with fewer terms omitted, at the expense of having a smaller multiplicative constant.)

In [4, p. 25], Drmota and Tichy prove a lower bound for discrepancy of sequences on $\mathbf{T}^k$ via the Koksma-Hlawka inequality. When $k = 1$, their bound is much like Theorem 1 but weaker (it has exponent 1 instead of 1/2, and a smaller constant). When $k > 1$, the Drmota-Tichy bound is in general not comparable to our Theorem 2.

In practice, since all the terms of the sum in the bounds of Theorems 1 and 2 and Corollary 3 are positive, one need only use as many terms as one needs for a lower bound. In fact, using only the dominant term(s) may be quite sufficient in some cases to yield good bounds. This occurs because, unlike the Drmota-Tichy bound, the dominant term in our bounds “matches” the dominant term in the upper bound of Erdös-Turán-Koksma.

We demonstrate the use of such bounds in Section 4, where they are applied to Kronecker sequences and to a random walk on the torus. First, we provide proofs of Theorems 1 and 2 in the next two sections of this paper.

2 Discrepancy on R/Z

Proof (of Theorem 1). We first prove two lemmas. Set $F(x) = Q([0, x]) - x + K$, where $[0, x]$ is an interval in $\mathbf{T}^1 = \mathbf{R}/\mathbf{Z}$, and K is chosen such that $\int_{[0,1]} F(x)\ dx = 0$.

Lemma 4.

$$\int_{[0,1]} F^2(x)\ dx \ = \ \frac{1}{2\pi^2} \sum_{m=1}^{\infty} \frac{|\widehat{Q}(m)|^2}{m^2} \tag{2}$$

where $F(x)$, defined as above, depends on Q.

Proof (of lemma). Compute the Fourier coefficents of F for $m \neq 0$:

$$\widehat{F}(m) = \int_0^1 (Q([0,x]) - x + K) e^{2\pi i m x} \, dx \ .$$

Integration by parts yields:

$$\widehat{F}(m) = -\int_0^1 \frac{e^{2\pi i m x}}{2\pi i m} (Q(dx) - dx) = \frac{\widehat{Q}(m)}{-2\pi i m} \ .$$

By our choice of K, $\widehat{F}(0) = 0$. Also note that $\widehat{Q}(-m) = \overline{\widehat{Q}(m)}$. These facts and Parseval's identity now give the conclusion of Lemma 4. □

An upper bound can be derived for $\int_{[0,1]} F^2(x) \, dx$ in terms of the discrepancy. In fact, we claim

Lemma 5.

$$\int_{[0,1]} F^2(x) \, dx \leq \frac{D^2}{4}$$

where $D = D(Q)$ is the discrepancy of Q.

We remark that it is easy to derive a bound of the form D^2; the non-trivial part is to improve it to $D^2/4$.

Proof (of lemma). Let $B_1 = \{x : F(x) \geq 0\}$ and $B_2 = \{x : F(x) < 0\}$. Define $W(x)$ by

$$W(x) = \begin{cases} \sup_{[0,1]} F & \text{for } x \in B_1 \\ \inf_{[0,1]} F & \text{for } x \in B_2. \end{cases}$$

By construction, $|W(x)| \geq |F(x)|$. Set $s = \lambda_1(B_1)$, $t = \sup F$. Hence $1 - s = \lambda_1(B_2)$, and it follows from the definition of the discrepancy that $D - t = |\inf F|$.

We claim that either $\int_{B_1} |W(x)| \, dx$ or $\int_{B_2} |W(x)| \, dx$ is bounded above by $\frac{D}{4}$. For suppose that $\int_{B_1} |W(x)| \, dx \geq \frac{D}{4}$. It follows that $st \geq \frac{D}{4}$. Then

$$\int_{B_2} |W(x)| \, dx = (1-s)(D-t) < (1 - \frac{D}{4t})(D-t) = \frac{5D}{4} - \frac{D^2}{4t} - t \ .$$

Elementary calculus shows that the maximum value of the expression on the right is $\frac{D}{4}$, which demonstrates this claim.

Therefore, set $A = \int_{B_1} |F(x)| \, dx$, and note by construction of F that $A = \int_{B_2} |F(x)| \, dx$, too. Since $|F(x)| \leq |W(x)|$, using the previous claim we see that $A \leq \frac{D}{4}$. Hence

$$\begin{aligned} \int_{[0,1]} F^2(x) \, dx &\leq t \int_{B_1} |F(x)| \, dx + (D-t) \int_{B_2} |F(x)| \, dx \\ &= (D-t)A + tA = DA \leq \frac{D^2}{4} \ . \end{aligned}$$

This is the conclusion of Lemma 5. □

The bound of Theorem 1 follows as an immediate consequence of the previous two lemmas.

Moreover, the bound is sharp in the sense that the ratio of the two sides can be made to approach 1. Given $\epsilon > 0$, let Q_ϵ be the measure that is uniform on $(0, \frac{1}{2} - \epsilon)$ and $(\frac{1}{2} + \epsilon, 1)$ and has point mass 2ϵ at 0. This is a family of measures for which the ratio of the two sides in Lemma 5 approaches 1 as $\epsilon \to 0$. (Calculations show the left side is $\epsilon^2 - \frac{4}{3}\epsilon^3$, and the right side is ϵ^2). Together with Lemma 4, this shows that the constant $\frac{2}{\pi^2}$ in the theorem is best possible.

The bound is also sharp for sequences, since any of the above measures Q_ϵ can be approximated a closely as desired by the counting measure $\frac{A_N}{N}$ of an appropriate sequence. For instance, let N be even. Consider N^2 points evenly placed on $\mathbf{R}/\mathbf{Z}$, at the points $0, \frac{1}{N^2}, \frac{2}{N^2}, ...$ etc. Pick up the $(2N-1)$ points nearest the point $1/2$ and place them at 0. Calculations for this sequence of the quantities in Lemma 5 show the left side is $\frac{1}{N^2} + O(\frac{1}{N^3})$ and the right side is exactly $\frac{1}{N^2}$. As $N \to \infty$ we see that the bound in the theorem becomes sharp. □

The proof of Theorem 1 was motivated by the proof of LeVeque's upper bound. In that proof the expression (2) was bounded *below* by an expression involving the discrepancy. This was accomplished by noting the existence of certain "triangles" under the graph of $F(x)$ whose height could be expressed in terms of the discrepancy. The main idea behind Theorem 1 is that the expression (2) also can be bounded sharply from *above* by an expression involving the discrepancy.

3 Discrepancy on $\mathbf{R}^k/\mathbf{Z}^k$

We shall take a slightly different approach to prove Theorem 2. First we require a lemma. Let $I_{\mathbf{x},\mathbf{r}}$ denote the interval in $\mathbf{T}^k$ centered at $\mathbf{x} = (x_1, ..., x_k)$ with radius vector $\mathbf{r} = (r_1, ..., r_k)$, i.e.,

$$I_{\mathbf{x},\mathbf{r}} = [x_1 - r_1, x_1 + r_1) \times ... \times [x_k - r_k, x_k + r_k).$$

Every interval can be written in this form, with $0 < r_i < 1/2$. As shorthand we let $I_{\mathbf{r}} = I_{\mathbf{0},\mathbf{r}}$ denote the interval centered at $\mathbf{0}$. Let $\delta_{I_{\mathbf{r}}}$ denote the indicator function of the set $I_{\mathbf{r}}$.

Lemma 6. *If Q is any probability measure on $\mathbf{T}^k$, then for any $\mathbf{r}$, $0 < r_i < 1/2$,*

$$D(Q)^2 \ \geq \sum_{\mathbf{0} \neq \mathbf{h} \in \mathbf{Z}^k} |\ \widehat{Q}(\mathbf{h})\ \widehat{\delta}_{I_{\mathbf{r}}}(\mathbf{h})\ |^2.$$

We prove this by constructing a function, parametrized by a radius vector **r**, that measures the discrepancy only on a set of radius **r** centered at each point; this is bounded above by the total discrepancy. It can be rewritten in terms of a convolution of the original measure and the indicator function of the set (this was observed by J. Beck, see [2] and [3]). An appeal to Parseval's identity then gives the desired conclusion.

Proof. Define

$$\Delta_{\mathbf{r}}(\mathbf{x}) = Q(I_{\mathbf{x},\mathbf{r}}) - \lambda_k(I_{\mathbf{x},\mathbf{r}}).$$

By the definition of $D(Q)$, we see that $\Delta_{\mathbf{r}}(\mathbf{x}) \leq D(Q)$, and hence for all r,

$$\int_{\mathbf{T}^k} \Delta_{\mathbf{r}}^2(\mathbf{x})\ d\mathbf{x} \leq D(Q)^2. \tag{3}$$

On the other hand, Parseval's identity yields

$$\int_{\mathbf{T}^k} \Delta_{\mathbf{r}}^2(\mathbf{x})\ d\mathbf{x}\ =\ \sum_{\mathbf{h}\in\mathbf{Z}^k} |\widehat{\Delta}_{\mathbf{r}}(\mathbf{h})|^2. \tag{4}$$

Notice that $\Delta_{\mathbf{r}}$ may be rewritten as:

$$\Delta_{\mathbf{r}}(\mathbf{x}) = Q * \delta_{I_{\mathbf{r}}}(\mathbf{x}) - \lambda_k * \delta_{I_{\mathbf{r}}}(\mathbf{x}) = (Q - \lambda_k) * \delta_{I_{\mathbf{r}}}(\mathbf{x})$$

where δ_I is the indicator function of the set I, and $*$ indicates convolution of a measure with a function: $Q * f(\mathbf{x}) = \int f(\mathbf{x}-\mathbf{y})Q(d\mathbf{y})$.

Then $\widehat{\Delta}_{\mathbf{r}}$ may be computed as $\widehat{\Delta}_{\mathbf{r}}(\mathbf{h}) = (\ \widehat{Q}(\mathbf{h}) - \widehat{\lambda}_k(\mathbf{h})\ )\ \widehat{\delta}_{I_{\mathbf{r}}}(\mathbf{h})$.

For $\mathbf{h} = \mathbf{0}$, $\widehat{\Delta}_{\mathbf{r}}(\mathbf{h}) = 0$ since

$$\widehat{Q}(\mathbf{0}) - \widehat{\lambda}_k(\mathbf{0}) = \int_{\mathbf{T}^k} Q(d\mathbf{x})\ - \int_{\mathbf{T}^k} \lambda_k(d\mathbf{x}) = 0.$$

For $\mathbf{h} \neq \mathbf{0}$, a trivial computation shows $\widehat{\lambda}_k(\mathbf{h}) = 0$, and thus

$$\widehat{\Delta}_{\mathbf{r}}(\mathbf{h}) = \widehat{Q}(\mathbf{h})\ \widehat{\delta}_{I_{\mathbf{r}}}(\mathbf{h}).$$

Combining this with (3) and (4), we obtain the desired bound. □

Now we can prove Theorem 2:

Proof (of Theorem 2). We compute $\widehat{\delta}_{I_{\mathbf{r}}}$:

$$\begin{aligned}
\widehat{\delta}_{I_{\mathbf{r}}}(\mathbf{h}) &= \int_{\mathbf{T}^k} \delta_{I_{\mathbf{r}}}(\mathbf{x})\ e(\mathbf{h}\cdot\mathbf{x})\ d\mathbf{x} \\
&= \int_{x_1=-r_1}^{r_1} \cdots \int_{x_k=-r_k}^{r_k} e^{2\pi i h_1 x_1} \cdots e^{2\pi i h_k x_k}\ d\mathbf{x} \\
&= \prod_{i=1}^{k} \left[\int_{x_i=-r_i}^{r_i} e^{2\pi i h_i x_i}\ dx_i\right].
\end{aligned} \tag{5}$$

Each of the factors in the last expression evaluates to $\frac{\sin(2\pi h_i r_i)}{\pi h_i}$ if $h_i \neq 0$, and $2r_i$ if $h_i = 0$. This gives

$$D(Q)^2 \;\geq \sum_{\mathbf{0}\neq\mathbf{h}\in\mathbf{Z}^k} |\widehat{Q}(\mathbf{h})|^2 \prod_{i=1}^{k} \left\{ \begin{array}{ll} \frac{\sin^2(2\pi h_i r_i)}{\pi^2 h_i^2} & \text{if } h_i \neq 0 \\ 4r_i^2 & \text{if } h_i = 0 \end{array} \right\}.$$

Since we are able to choose any $\mathbf{r}$ for a lower bound, this yields the conclusion of the theorem. □

Lemma 6 is a useful result in its own right, since it allows us to derive similar bounds for alternate variants of discrepancy, which arise from using sets other than intervals to evaluate the supremum in (1).

For example, *ball*-discrepancy is defined over balls under the usual Euclidean metric on $\mathbf{T}^k$. Let $B_{\mathbf{x},r} = \{\mathbf{x} \in \mathbf{R}^k : d(\mathbf{x}, \mathbf{c}) \leq r\}$ denote the k-dimensional ball in $\mathbf{R}^k$ of radius r centered at $\mathbf{x}$. (Here, d is the usual Euclidean metric.) This reduces to a ball in $\mathbf{T}^k$ which for simplicity we also denote by $B_{\mathbf{x},r}$. If B denotes any ball in $\mathbf{T}^k$, the ball discrepancy is defined by

$$D_{\text{ball}}(Q) = \sup_{B\subseteq\mathbf{T}^k} |P(B) - \lambda_k(B)|.$$

The Fourier coefficients of $\delta_{B_{\mathbf{x},r}}$ on $\mathbf{R}^k$ are known (see [5]) and can be expressed in terms of Bessel functions $J_{k/2}$:

$$\widehat{\delta}_{B_{\mathbf{x},r}}(\mathbf{h}) = \int_{\mathbf{R}^k} \delta_{B_{\mathbf{x},r}}(\mathbf{x}) e(\mathbf{h}\cdot\mathbf{x})\; d\mathbf{x} = \left(\frac{r}{|\mathbf{h}|}\right)^{k/2} J_{k/2}(2\pi r|\mathbf{h}|).$$

On $\mathbf{T}^k$ this result is still valid as long as $r < \frac{1}{2}$ (otherwise, the ball overlaps itself). Therefore from Lemma 6 we obtain:

$$D_{\text{ball}}(Q) \;\geq\; \sup_{r<1/2} \left[\sum_{\mathbf{0}\neq\mathbf{h}\in\mathbf{Z}^k} |\widehat{Q}(\mathbf{h})|^2 \;\left(\frac{r}{|\mathbf{h}|}\right)^{k} \left(J_{k/2}(2\pi r|\mathbf{h}|)\right)^2 \right]^{1/2}.$$

See Holt [5] for a related upper bound.

4 Applications

As a demonstration, we compute lower bounds for Kronecker sequences. (A *Kronecker sequence* is the sequence of integer multiples of a given k-tuple of real numbers.) The results we derive compare well with known bounds. We conclude by using Corollary 3 to derive a rate of convergence for a random walk on the torus.

The Kronecker sequence for $k = 1$.

This is just the sequence $\{n\alpha\}$, $n = 0, 1, ..., N-1$, on $\mathbf{R}/\mathbf{Z}$.

Recall that an irrational α is said to be of *type* η if η is the infimum of all numbers τ for which there exists a positive constant C such that $\|q\alpha\| \geq \frac{C}{q^\tau}$ for all positive integers q. Here $\|x\|$ denotes the nearest distance from x to an integer.

Theorem 7. *Let α be an irrational of type η. The discrepancy of the sequence $\{x_n\} = \{n\alpha\}$, $n = 0, 1, 2, ..., (N-1)$, satisfies, for any $\epsilon > 0$,*

$$D_N(x_n) = \Omega(N^{-\frac{1}{\eta}-\epsilon}). \tag{6}$$

Here $f(x) = \Omega(g(x))$ means $f(x) \neq o(g(x))$. This result agrees with Behnke's lower bound, derived in [6] by a different method.

Proof. Given $\epsilon > 0$, there exists a δ such that $\frac{1}{\eta-\delta} = \frac{1}{\eta} + \epsilon$ where $0 < \delta < \eta$.

Since α is of type η, there exist infinitely many m such that $m^{\eta-\delta}\|m\alpha\| < \frac{1}{4}$. For such m, let $N = \lceil m^{\eta-\delta} \rceil$. Fix m large enough so that one still has $N\|m\alpha\| < \frac{1}{2}$. The measure Q for this sequence consists of weight $\frac{1}{N}$ on each member of the sequence. Hence

$$\left|\widehat{Q}(m)\right| = \left|\frac{1}{N}\sum_{n=0}^{N-1} e(nm\alpha)\right| = \frac{1}{N}\frac{|1-e(Nm\alpha)|}{|1-e(m\alpha)|} \geq \frac{2\|Nm\alpha\|}{N\pi\|m\alpha\|} \tag{7}$$

The inequality follows from noting that $4\|x\| \leq |1 - e(\|x\|)| \leq 2\pi\|x\|$. Since $N\|m\alpha\| < \frac{1}{2}$, $\|Nm\alpha\| = N\|m\alpha\|$ and so (7) reduces to $\widehat{Q}(m) \geq \frac{2}{\pi}$.

Using only the m-th term in the lower bound of Theorem 1, and using the fact that $N \geq m^{\eta-\delta}$, we obtain:

$$D_N(x_n) \geq \frac{2\sqrt{2}}{\pi^2 m} \geq \frac{2\sqrt{2}}{\pi^2} N^{-\frac{1}{\eta}-\epsilon}.$$

The statement holds for infinitely many choices of N (since N were derived from the choice of infinitely many m). This establishes (6). □

The Kronecker sequence for $k > 1$.

This is the sequence $\{n\mathbf{a}\}$, $n = 0, 1, ..., N-1$, for some $\mathbf{a} = (\alpha_1, \alpha_2, ..., \alpha_k)$ in $\mathbf{R}^k/\mathbf{Z}^k$. As in the case $k = 1$, we shall rely only on a single dominant term in the lower bound of Theorem 2.

Recall that $R(\mathbf{h}) = \prod_{i=1}^k \max\{1, |h_i|\}$. Fixing any $\mathbf{h}$, setting $r_i = \frac{1}{4h_i}$ for $h_i \neq 0$ and $r_i = \frac{1}{2\pi}$ for $h_i = 0$, and ignoring all other terms in Theorem 2, we obtain:

Lemma 8. *For any* $\mathbf{h} \in \mathbf{Z}^k$*,* $\mathbf{h} \neq \mathbf{0}$*,*

$$D(Q) \;\geq\; \frac{|\widehat{Q}(\mathbf{h})|}{\pi^k R(\mathbf{h})}.$$

We remark that a similar inequality (with weaker constant) can be derived via the Koksma-Hlawka inequality applied to the function $f(\mathbf{x}) = e(\mathbf{h} \cdot \mathbf{x})$ of bounded variation $V(f) \ll R(\mathbf{h})$.

Lemma 8 can be used to establish a lower bound for Kronecker sequences that depends on a type classification of k-tuples, similar to the type classification of irrationals defined earlier. A k-tuple $\mathbf{a} = (\alpha_1, ..., \alpha_k)$ of irrationals is said to be of *type* η if η is the infimum of all numbers τ for which there exists a positive constant C such that $R(\mathbf{h})^\tau \|\mathbf{h} \cdot \mathbf{a}\| \geq C$ holds for all non-zero lattice points $\mathbf{h}$ in $\mathbf{Z}^k$.

Theorem 9. *Let* $\mathbf{a} \in \mathbf{R}^k$ *be a k-tuple of irrationals, of type* η*. The discrepancy of the sequence* $\{x_n\} = \{n\mathbf{a}\}$*,* $n = 0, 1, 2, ..., (N-1)$*, satisfies, for any* $\epsilon > 0$*,*

$$D_N(x_n) = \Omega(N^{-\frac{1}{\eta}-\epsilon}). \tag{8}$$

The exponent is nearly sharp; an upper bound of $O(N^{-\frac{1}{\eta}+\epsilon})$ may be established by the method of Niederreiter [8] (who carries out the case for $\eta = 1$).

Proof. Given $\epsilon > 0$, there exists a δ such that $\frac{1}{\eta-\delta} = \frac{1}{\eta} + \epsilon$ where $0 < \delta < \eta$. Since α is of type η, there exist infinitely many $\mathbf{h} \in \mathbf{Z}^k$ such that $R(\mathbf{h})^{\eta-\delta}\|\mathbf{h} \cdot \mathbf{a}\| < \frac{1}{4}$. For such $\mathbf{h}$, let $N = \lceil R(\mathbf{h})^{\eta-\delta} \rceil$. Fix $\mathbf{h}$ large enough so that one still has $N\|\mathbf{h} \cdot \mathbf{a}\| < \frac{1}{2}$. It can now be shown that $\widehat{Q}(\mathbf{h}) \geq \frac{2}{\pi}$ by a calculation analogous to (7) in which $m\alpha$ is replaced by $\mathbf{h} \cdot \mathbf{a}$. Using this in Lemma 8, we obtain

$$D_N(x_n) \geq \frac{2}{\pi^{k+1}} N^{-\frac{1}{\eta}-\epsilon}.$$

Since this holds for infinitely many N (derived from infinitely many $\mathbf{h}$), (8) follows. □

A random walk on the torus.

Consider the random walk on $\mathbf{T}^k = \mathbf{R}^k/\mathbf{Z}^k$ which starts at the origin, and at each step moves to a point uniformly chosen within a box of radius A about the current point. If Q is a probability measure uniformly supported on $[-A, A]^k/\mathbf{Z}^k$, then Q^{*n} represents the n-th step probability distribution of this random walk, which evidently converges to the uniform distribution on the torus. In fact, using Corollary 3, we shall show that the discrepancy $D(Q^{*n})$ converges to zero at a geometric rate in n, by showing that both upper and lower bounds converge geometrically:

Theorem 10. *The discrepancy of the above random walk on the torus satisfies*

$$C_1\ \beta^n\ <\ D(Q^{*n})\ <\ C_2\ \beta^n$$

where n is the number of steps taken by the random walk, $\beta = |\frac{\sin(2\pi A)}{2\pi A}|$, and $C_1 = C_1(k)$ and $C_2 = C_2(k, A)$ are constants specified in the proof.

Proof. Since the density of Q is $(2A)^{-k}$ on its support, by a derivation similar to (5) we obtain the Fourier coefficients of Q:

$$\widehat{Q}(\mathbf{h}) = \prod_{i=1}^{k} \left\{ \begin{array}{ll} \frac{\sin(2\pi h_i A)}{2\pi h_i A} & \text{if } h_i \neq 0 \\ 1 & \text{if } h_i = 0 \end{array} \right\}. \tag{9}$$

Recall that $\widehat{Q^{*n}}(\mathbf{h}) = \widehat{Q}^n(\mathbf{h})$. Now for positive integers m, $\frac{\sin(2\pi m A)}{\sin(2\pi A)} = U_{m-1}(\cos(2\pi A))$, where U_{m-1} is the Chebyshev polynomial of the second kind. It satisfies the inequality $|U_{m-1}(x)| \leq m$ for $x \in [-1, 1]$ (see [1]). Hence for any $h_i \neq 0$, $|\frac{\sin(2\pi h_i A)}{2\pi h_i A}| \leq |\frac{\sin(2\pi A)}{2\pi A}|$ which is less than 1 for $A > 0$.

Let $\beta = |\frac{\sin(2\pi A)}{2\pi A}|$. Then $|\widehat{Q}(\mathbf{h})| \leq \beta$ for all $\mathbf{h} \neq \mathbf{0}$. The dominant Fourier coefficients occur when exactly one of the $|h_i| = 1$ and all the other coordinates are 0. There are exactly $2k$ such coefficients. Using only these coefficients in the lower bound from Corollary 3, we see that

$$D(Q^{*n}) \geq \frac{1}{\pi^k}(2k\beta^{2n})^{1/2} = \frac{\sqrt{2k}}{\pi^k}\beta^n.$$

On the other hand, an upper bound for $D(Q^{*n})$ can be obtained via the Erdös-Turán-Koksma inequality (see [4]) which also involves the Fourier coefficients: for any positive integer H,

$$D(Q^{*n}) \leq (3/2)^k \left(\frac{2}{H+1} + \sum_{0<\|\mathbf{h}\|_\infty \leq H} \frac{|\widehat{Q}^n(\mathbf{h})|}{R(\mathbf{h})} \right). \tag{10}$$

Using $|\widehat{Q}^n(\mathbf{h})| \leq \beta^{n-1}|\widehat{Q}(\mathbf{h})|$ and equation (9),

$$\begin{aligned} \sum_{0<\|\mathbf{h}\|_\infty \leq H} \frac{|\widehat{Q}^n(\mathbf{h})|}{R(\mathbf{h})} &\leq \beta^{n-1} \sum_{0<\|\mathbf{h}\|_\infty \leq H} \prod_{i=1}^{k} \left\{ \begin{array}{ll} \frac{|\sin(2\pi h_i A)|}{2\pi h_i^2 A} & \text{if } h_i \neq 0 \\ 1 & \text{if } h_i = 0 \end{array} \right\} \\ &\leq \beta^{n-1} \left(\left(1 + 2\sum_{h=1}^{H} \frac{1}{2\pi h^2 A} \right)^k - 1 \right) \\ &\leq \beta^{n-1} \ \left(1 + \frac{\pi}{6A}\right)^k. \end{aligned}$$

Hence, letting $H \to \infty$ in (10), we have

$$D(Q^{*n}) \leq \left(\frac{6A\pi + \pi^2}{2\sin(2\pi A)} \right)^k \beta^n.$$

Therefore the discrepancy convergence of this random walk is geometric in n with rate $\beta = |\sin(2\pi A)/2\pi A|$, as claimed. $\square$

A more refined lower bound may be obtained by taking more terms in Corollary 3 or Theorem 2. Note, however, that the extra terms in Theorem 2 and in the Erdös-Turán-Koksma inequality involve $\widehat{Q}^n(\mathbf{h})$ and thus become negligible in comparison to their dominant term(s) when the number of steps n is large. This is true for any random walk on the torus. The matching dominant terms of these bounds therefore make discrepancy an effective tool to measure how quickly random walks on the torus converge to the uniform distribution.

Acknowledgements

I would like to thank Persi Diaconis and Jerry Spanier for helpful discussions, and Jeff Holt and Ken Ross for useful feedback.

References

1. G. Arfken, *Mathematical Methods for Physicists*, 3rd ed., Academic Press, New York, 1985.
2. J. Beck, *On a problem of K. F. Roth concerning irregularities of point distribution*, Invent. Math. 74(1983), 477-487.
3. J. Beck and W. Chen, *Irregularities of Distribution*, Cambridge University Press, 1987.
4. M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Math. 1651, Springer, 1997.
5. J. J. Holt, *On a form of the Erdös-Turán inequality*, Acta. Arith. 74 (1996), 61-66.
6. L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.
7. W. J. LeVeque, *An inequality connected with Weyl's criterion for uniform distribution*, Proc. Symp. Pure Math. 8(1965), 22-30.
8. H. Niederreiter, Application of diophantine approximations to numerical integration, in C. F. Osgood (Ed.), *Diophantine Approximation and its Applications*, Academic Press, New York, 1973, pp. 129-199.
9. F. E. Su, *Methods for Quantifying Rates of Convergence for Random Walks on Groups*, Ph.D. thesis, Harvard University, 1995.
10. F. E. Su, *Convergence of random walks on the circle generated by an irrational rotation*, Trans. Amer. Math. Soc., 350(1998), pp. 3717-3741.