TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 350, Number 9, September 1998, Pages 3717–3741
S 0002-9947(98)02152-7

# CONVERGENCE OF RANDOM WALKS ON THE CIRCLE GENERATED BY AN IRRATIONAL ROTATION

FRANCIS EDWARD SU

ABSTRACT. Fix $\alpha \in [0,1)$. Consider the random walk on the circle $S^1$ which proceeds by repeatedly rotating points forward or backward, with probability $\frac{1}{2}$, by an angle $2\pi\alpha$. This paper analyzes the rate of convergence of this walk to the uniform distribution under "discrepancy" distance. The rate depends on the continued fraction properties of the number $\xi = 2\alpha$. We obtain bounds for rates when $\xi$ is any irrational, and a sharp rate when $\xi$ is a quadratic irrational. In that case the discrepancy falls as $k^{-\frac{1}{2}}$ (up to constant factors), where $k$ is the number of steps in the walk. This is the first example of a sharp rate for a discrete walk on a continuous state space. It is obtained by establishing an interesting recurrence relation for the distribution of multiples of $\xi$ which allows for tighter bounds on terms which appear in the Erdős-Turán inequality.

## 1. INTRODUCTION

Although it is well known that random walks on groups converge to the uniform distribution (with mild restrictions), it has only been relatively recently that probabilists have begun to ask how quickly that convergence takes place. Much work has been done to obtain rates of convergence in the finite group context (e.g., see Diaconis [2], Diaconis and Saloff-Coste [3]), and some rates have been obtained for walks on infinite compact groups (e.g., see Rosenthal [18], Porod [15]), but far less is known for *discrete* walks on infinite compact groups, primarily because few methods currently exist to attack problems of this nature. This paper analyzes walks on the continuous circle $S^1$ generated by a fixed irrational rotation, which are prototypical examples of a discrete walk on a compact group. Even in these basic examples the analysis required to obtain a precise rate of convergence is fairly involved.

Consider the following random walk on the continuous circle $S^1$. Fix a number $\alpha \in [0,1)$. Start at any point, and at each step, rotate the point forward or backward by an angle $2\pi\alpha$ with probability $\frac{1}{2}$. This corresponds to convolving the measure $Q_\alpha$ on $S^1$ which assigns mass $\frac{1}{2}$ to the angles $\pm 2\pi\alpha$. We shall henceforth refer to this walk as the $\pm\alpha$ *random walk on* $S^1$.

We investigate the convergence of this walk under the "discrepancy" distance on $S^1$: for any probability measures $P, Q$ on $S^1$, the discrepancy $D(P,Q)$ is defined to

Received by the editors October 18, 1996.
1991 *Mathematics Subject Classification.* Primary 60J15, 60B15; Secondary 11K38, 11J70.
*Key words and phrases.* Random walk, rate of convergence, discrepancy, Erdős-Turán inequality, continued fractions, irrational rotation, uniform distribution of sequences.
Supported in part by an NSF Graduate Fellowship.

be

$$D(P,Q) = \sup_{J \subset S^1} |P(J) - Q(J)| \tag{1.1}$$

where $J$ is any interval in $S^1$. Thus for $U$ the uniform distribution, the discrepancy $D(P,U)$ measures how well-distributed a measure $P$ is.

Note that when $\alpha$ is rational, the walk remains supported on a finite subgroup of $S^1$, and in such cases it suffices to understand the corresponding random walk on the finite circle $\mathbf{Z}/p\mathbf{Z}$. See Diaconis [2] and Su [21] for a treatment of walks on $\mathbf{Z}/p\mathbf{Z}$ and the development of upper and lower bounds for discrepancy on the finite circle.

The interesting case for the $\pm\alpha$ walk is when $\alpha$ is irrational. In this case the $k$-th step probability distribution of the walk converges in the weak-* topology to the uniform distribution (Haar measure) on the circle $S^1$, at a rate depending on the nature of $\alpha$. For instance, if $\alpha$ is very close to some rational, then one might expect the convergence to be slow, at least initially. If the multiples of $\alpha$ are in some sense "well-distributed", then perhaps the convergence will be faster. The rate of convergence will be shown to depend on the continued fraction properties of the number $\xi = 2\alpha$.

The main result of this paper is Theorem 4.9, which shows that when $\alpha$ is a quadratic irrational, the $k$-th step probability distribution $Q_\alpha^{*k}$ of this walk satisfies

$$\frac{C_1}{\sqrt{k}} \leq \ D(Q_\alpha^{*k}, U) \ \leq \frac{C_2}{\sqrt{k}} \tag{1.2}$$

where the constants $C_1, C_2$ depend on $\xi$ in a manner specified in the paper.

This result is interesting for several reasons: (1) It is the first sharp rate of convergence established for a discrete walk on a continuous group. Prior to this, the best upper bound known for this walk was order $\log k/\sqrt{k}$, due to Diaconis [2]. (2) Note that the rate is not exponential, but of polynomial decay. By contrast, walks on finite groups and walks on compact groups with absolutely continuous generating measure (with respect to Haar measure) must, in the long run, converge exponentially to their limiting distributions (in total variation distance), due to a theorem of Kloss [5]. (3) This result resembles results in the theory of uniform distribution of sequences mod 1; however, the analogous results in that theory concern *equally weighted* measures on sequences. For comparison, it is known that the discrepancy between the sequence $\{\alpha, 2\alpha, 3\alpha, ..., k\alpha\}$ and the uniform distribution falls as $\log k/k$, up to constant factors (see Kuipers and Niederreiter [7]). It is not surprising that when compared to equally weighted sequences, the exponent for random walk on the multiples of $\alpha$, as in (1.2), is halved, but it is somewhat surprising that $\log k$ term disappears.

The crucial ingredient in establishing the upper bound in (1.2) is a very interesting recurrence relation for the distribution of multiples of any given irrational (mod 1), which is used to bound the terms the sum obtained by applying Erdős-Turán's inequality. This recurrence, found in Theorem 3.16, is a major result of this paper and is of interest in its own right, because of its usefulness in bounding sums containing irrational multiples. For an illustration of the recurrence, refer to Tables 1–4 in Section 3.

When $\alpha$ is an arbitrary irrational, we obtain less sharp bounds on the convergence of the walk (Theorems 5.5 and 5.8) which depend on a "type" classification of irrationals by the closeness of their best rational approximations. When $\xi = 2\alpha$ is

of type $\eta$, we find that the discrepancy of the walk falls off "roughly" like $k^{-1/2\eta}$ where $k$ is the number of steps in the walk. In particular, since irrationals of "type 1" have full measure in the circle, we see that for almost all $\alpha$ that the discrepancy of the walk falls off roughly as the square root of the number of steps.

This paper is organized as follows. Section 2 discusses techniques for bounding discrepancy on the circle. Section 3 gives background on continued fractions and establishes the recurrence relation for the distribution of multiples of any irrational $\xi$. Section 4 applies this result in the case $\xi$ is a quadratic irrational, and establishes the main theorem of this paper. Section 5 derives less sharp bounds for $\xi$ any arbitrary irrational, using probabilistic considerations. Section 6 discusses motivation for studying discrete walks on compact groups and directions for future work.

## 2. Discrepancy

One may observe from equation (1.1) that discrepancy distance $D(P, U)$ expresses how well-distributed the measure $P$ is; in fact, if $B_1$ and $B_2$ are intervals in $S^1$ having the same length, then $|P(B_1) - P(B_2)| \leq 2\, D(P, U)$. That is, the likelihood of finding the walk in $B_1$ versus $B_2$ can differ by at most twice the discrepancy distance from the uniform distribution.

Discrepancy has been used extensively in analytic number theory to study the uniform distribution of sequences (mod 1). In that context, discrepancy of sequences is a special case of (1.1) in which one measure is taken to be equally weighted on all elements of a given sequence, and the other measure is the uniform distribution on $S^1$. See Kuipers and Niederreiter [7] for a survey of this field.

Convergence in discrepancy implies weak-* convergence of measures. This may be seen by noting that discrepancy bounds the Prokhorov metric, which metrizes weak-* convergence. A proof is contained in Su [21]. The converse is not true (consider delta measures on a convergent sequence of points).

A fact which will be of use to us later is that on $S^1$ (in fact, on any compact group), discrepancy decreases with convolution.

**Theorem 2.1.** *If $P, Q, R$ are arbitrary probability measures on a compact group $G$, then*

$$D(P * R, Q * R) \;\leq\; D(P, Q).$$

*Hence when $Q = U$, the uniform distribution, we have*

$$D(P * R, U) \;\leq\; D(P, U). \tag{2.1}$$

*Proof.* If $B$ is an arbitrary ball in $G$,

$$\begin{aligned}
|P * R(B) - Q * R(B)| &= |\, (P - Q) * R(B)\, | \\
&= \left| \int_{x \in B} \int_{y \in G} (P - Q)(xy^{-1})\; R(y)\; dy \quad dx \right| \\
&= \left| \int_{y \in G} \left( \int_{x \in B} (P - Q)(xy^{-1})\; dx \right) R(y)\; dy \right| \\
&\leq \int_{y \in G} \left|\, (P - Q)(By^{-1}) \,\right|\; R(y)\; dy \\
&\leq D(P, Q) \int_{y \in G} R(y)\; dy \;\leq\; D(P, Q)
\end{aligned}$$

where the change of the order of integration is allowed since the integrand is in $L^1$. Taking the supremum on the left side over all balls $B$ gives the desired inequality. The inequality (2.1) follows by noting that $U * R = U$. □

The following proposition gives a lower bound for the convergence in discrepancy of the $\pm\alpha$ random walk on $S^1$:

**Proposition 2.2.** *For any irrational $\alpha$, the discrepancy of the $\pm\alpha$ random walk satisfies, for $k \geq 1$,*

$$D(Q_\alpha^{*k}, U) > \frac{1}{2\sqrt{k}}.$$

*Proof.* The statement is a consequence of the fact that the discrepancy is bounded below by the size of any atom in the measure. The weights of the atoms are distributed binomially, so that weight of the largest atom is, for $k = 2m$ even, $\binom{2m}{m}/2^{2m}$, and for $k = 2m-1$ odd, $\binom{2m-1}{m-1}/2^{2m-1}$.

We may lower bound these expressions by using the fact that for even $n$, $n!! < \sqrt{2n}(n-1)!!$, where $n!!$ denotes the product of the even or odd numbers from 1 to $n$, according as $n$ is even or odd. (This fact may be proved by induction, noting that $\sqrt{(z-1)(z+1)} < z$.)

Hence for even $k = 2m$,

$$\frac{\binom{2m}{m}}{2^{2m}} = \frac{(2m)!}{2^{2m}m!m!} = \frac{(2m-1)!!}{2^m m!} \geq \frac{(2m)!!}{\sqrt{4m}}\frac{1}{2^m m!} = \frac{1}{\sqrt{2k}}.$$

For odd $k = 2m-1$,

$$\frac{\binom{2m-1}{m-1}}{2^{2m-1}} = \frac{(2m-1)!}{2^{2m-1}m!(m-1)!} = \frac{(2m-1)!!}{2^m m!} \geq \frac{(2m)!!}{\sqrt{4m}}\frac{1}{2^m m!} = \frac{1}{\sqrt{2(k+1)}} \geq \frac{1}{2\sqrt{k}}.$$

□

*Remark* 2.3. From (1.1), lower bounds on discrepancy may be obtained by choosing an interval $J$ (in the previous example it was the largest atom) on which to evaluate the difference between two measures on that set. Another possible method is the following lower bound for discrepancy derived by Su [21]:

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

where $Q$ is any probability distribution on $S^1$, and $\widehat{Q}(m)$ denotes the $m$-th Fourier coefficient of $Q$. The constant $\frac{2}{\pi^2}$ is best possible.

This lower bound complements in form an upper bound for discrepancy due to LeVeque [9]. However, in the case of the $\pm\alpha$ walk, this bound does not provide a significantly better lower bound than Proposition 2.2.

To obtain upper bounds for discrepancy, we shall use the following inequality, due to Erdős and Turán [4], which was formulated originally with unspecified constants for the discrepancy of sequences. Niederreiter and Philipp [14] established constants and generalized the result for arbitrary probability measures.

**Theorem 2.4** (Erdős-Turán)**.** *Let $Q$ be any probability distribution on $S^1$ and $U$ the uniform distribution. Then for any integer $m$,*

$$D(Q,U) \leq \frac{4}{m+1} + \frac{4}{\pi}\sum_{h=1}^{m}\left(\frac{1}{h} - \frac{1}{m+1}\right)\left|\widehat{Q}(h)\right|$$

*where $\widehat{Q}$ represents the Fourier transform of $Q$.*

Note that one may choose $m$ in the Erdős-Turán inequality so as to optimize the bound obtained.

The Fourier coefficients for the $\pm\alpha$ walk are $\widehat{Q_\alpha}(m) = \frac{1}{2}e^{2\pi i m\alpha} + \frac{1}{2}e^{-2\pi i m\alpha} = \cos(2\pi m\alpha)$. Applying Theorem 2.4 to the $k$-th step probability distribution $Q_\alpha^{*k}$, and allowing $m$ to grow with $k$, one finds that it is necessary to understand the behavior of a sum containing the Fourier coefficients

$$\widehat{Q_\alpha^{*k}}(m) = \cos^k(2\pi m\alpha).$$

Evidently, when $2m\alpha$ is close to an integer, $|\cos^k(2\pi m\alpha)|$ is close to 1, and the corresponding terms in the Erdős-Turán inequality tend to zero slowly. Hence the closest rational approximations to $\xi = 2\alpha$ will offer some control on the rate of convergence of the walk.

These may be determined by the continued fraction expansion of $\xi = 2\alpha$. The next section contains a brief summary of background needed from the theory of continued fractions, and derives a recurrence relation for irrational multiples to control the growth of each term in the Erdős-Turán inequality.

## 3. A Recurrence Relation for the Distribution of Irrational Multiples

3.1. **Continued Fractions.** We now recall a few facts about continued fractions. An excellent basic reference for this subject is the concise little book by Khinchin [6].

**Definition 3.1.** Given a positive irrational number $\xi$, let

$$\begin{array}{rclcrcl} a_0 &=& \lfloor \xi \rfloor & , & b_0 &=& \xi - a_0, \\ a_1 &=& \lfloor \frac{1}{b_0} \rfloor & , & b_1 &=& \frac{1}{b_0} - a_1, \\ a_i &=& \lfloor \frac{1}{b_{i-1}} \rfloor & , & b_i &=& \frac{1}{b_{i-1}} - a_i, \end{array}$$

where $\lfloor x \rfloor$ denotes the floor of $x$, i.e., the greatest integer less than or equal to $x$. This defines the $a_i$ such that

$$\xi = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

and $a_i \geq 1$ for $i \geq 1$. We shall denote this representation in the sequel by

$$\xi = [a_0; a_1, a_2, ...].$$

This representation is called the *continued fraction expansion* of $\xi$. The $a_i$ are called the *elements*, or *partial quotients* of the number $\xi$. The number $r_i = [a_i, a_{i+1}, a_{i+2}, ...]$ is called the $i$-th *remainder* of $\xi$ and is equal by construction to $\frac{1}{b_{i-1}}$.

**Example 3.2.** Let $\Phi = \frac{\sqrt{5}+1}{2} \approx 1.618...$ denote the *golden mean*. Let $\phi = 1/\Phi \approx 0.618...$, a number whose continued fraction expansion will be important to us later. One finds that $\phi = [0; 1, 1, 1, ...]$. Its remainders $r_i$, $i \geq 1$, are all equal to $\Phi$.

As other examples

$$\sqrt{2} = [1; 2, 2, 2, ...], \quad \sqrt{3} = [1; 1, 2, 1, 2, ...], \quad \text{and} \quad e = [2; 1, 2, 1, 1, 4, 1, 1, 6, ...].$$

**Definition 3.3.** Given $\xi = [a_0; a_1, a_2, ...]$, define $p_i, q_i$ by

$$\begin{array}{rclcrcl} p_{-1} & = & 1 & , & q_{-1} & = & 0, \\ p_0 & = & a_0 & , & q_0 & = & 1, \\ p_1 & = & a_1 a_0 + 1 & , & q_1 & = & a_1, \\ p_k & = & a_k p_{k-1} + p_{k-2} & , & q_k & = & a_k q_{k-1} + q_{k-2}. \end{array} \tag{3.1}$$

The fractions $C_i = \frac{p_i}{q_i}$ are called the *convergents* of $\xi$, and they provide the best rational approximations to $\xi$ in the sense specified in the next proposition. For $i \geq -1$, let $R_i$ denote the error term $|q_i\xi - p_i|$.

**Proposition 3.4.** *For $\xi$ a positive irrational, and any positive integers $a$ and $b$ such that $1 \leq b < q_{i+1}$,*

$$R_i = |q_i\xi - p_i| \ \leq \ |b\xi - a|.$$

*Proof.* According to Khinchin [6, Thm.17], for any $\xi \neq \frac{1}{2}$, every convergent is a best approximation in the sense that for integers $a$ and $1 \leq b < q_i$, $R_i \leq |b\xi - a|$. According to Khinchin [6, Thm.16], every best approximation in the above sense is a convergent, which shows that the condition on $b$ may be extended to all $b < q_{i+1}$. □

The preceding proposition shows that the $q_i$-th multiple of $\xi$ comes closer to an integer than any multiple of $\xi$ before it, and so the $q_i$ may be thought of as best successive "record-breakers" in distance to the nearest integer as we run through the multiples of $\xi$. The $R_i$ are then the best successive records.

The following proposition will tell us just how large the $R_i$ are. Though it may be a classical result, a proof is included here for completeness.

**Proposition 3.5.** *For any positive irrational $\xi$, and $p_i, q_i, b_i, r_i$ defined as in Definitions 3.1 and 3.3, we have for $n \geq -1$,*

$$q_n\xi - p_n = (-1)^n \ b_0 b_1 \cdots b_n = \frac{(-1)^n}{r_1 r_2 \cdots r_{n+1}}.$$

*(For $n = -1$, we interpret $q_{-1}\xi - p_{-1} = -1$.)*

*Proof.* We use induction. Trivially, for $n = -1$, $0 \cdot \xi - 1 = -1$, and for $n = 0$, $1 \cdot \xi - a_0 = b_0$.

Now suppose

$$q_{n-1}\xi - p_{n-1} = (-1)^{n-1} \ b_0 b_1 \cdots b_{n-1}.$$

Then

$$\begin{array}{rcl} q_n\xi - p_n & = & (a_n q_{n-1} + q_{n-2})\xi - (a_n p_{n-1} + p_{n-2}) \\ & = & (q_{n-2}\xi - p_{n-2}) + a_n(q_{n-1}\xi - p_{n-1}) \\ & = & (-1)^{n-2}(b_0 b_1 \cdots b_{n-2}) + a_n(-1)^{n-1}(b_0 b_1 \cdots b_{n-1}) \\ & = & b_0 \cdots b_{n-2}(1 - b_{n-1}a_n)(-1)^n \\ & = & b_0 \cdots b_n(-1)^n. \end{array}$$

The second equality follows trivially from $r_i = \frac{1}{b_{i-1}}$. □

Note the alternation in sign of $(q_n\xi - p_n)$. Also note that taking absolute values gives the value of $R_n$, and since the $b_n < 1$, the $R_n$ are strictly decreasing as $n$ increases. The following standard fact shows that they decrease no faster than the $q_i$ increase:

**Proposition 3.6.** *For any positive irrational $\xi$, and $q_i = q_i(\xi), R_i = R_i(\xi)$ as defined above,*

$$\frac{1}{q_k + q_{k+1}} < R_k < \frac{1}{q_{k+1}} \tag{3.2}$$

*for any integer $k \geq 0$.*

For proofs of these bounds, see Khinchin [6, Thms.9,13]. Since from (3.1), $q_k + q_{k+1} \leq q_{k+2}$, this yields, conversely,

**Proposition 3.7.** *For any integer $k \geq 2$,*

$$\frac{1}{R_{k-2}} < q_k < \frac{1}{R_{k-1}}. \tag{3.3}$$

For later use note that the $R_i$, like the $q_i$, satisfy a recurrence relation:

**Proposition 3.8.** *For any integer $i \geq 1$,*

$$R_{i-2} = a_i R_{i-1} + R_i. \tag{3.4}$$

*Proof.* The recurrence relations for $p_i, q_i$ show that adding $q_{i-2}\xi - p_{i-2}$ to the quantity $a_i(q_{i-1}\xi - p_{i-1})$ gives $q_i\xi - p_i$, which gives the assertion we are trying to prove by noting that $(q_{i-1}\xi - p_{i-1})$ is of the opposite sign as the other two terms. □

*Remark* 3.9. Observe that for $i = 1$, we obtain $1 = a_1 R_0 + R_1$, which will be used later.

**Example 3.10.** For $\xi = \phi \approx 0.618$, (3.1) shows that the $p_i$ and $q_i$ are the same, up to a shift of index, and are given by the well known *Fibonacci sequence*: $0, 1, 1, 2, 3, 5, 8, 13, ...$ in which each term is the sum of the two preceding ones. Successive ratios provide better and better approximations to $\phi$. Proposition 3.5 shows that $R_i = \phi^{i+1}$.

3.2. **The Recurrence Relation.** We now establish notation and prove a few facts that will be used to establish the main results of this section, Theorems 3.16 and 3.19. The recurrence relation in Theorem 3.16 answers the question: how are the multiples of an irrational $\xi$ distributed with respect to the nearest integer?

Theorem 3.16 is used to show Theorem 3.19, which is used in Section 4 to prove Theorem 4.9, the main result of this paper.

**Definition 3.11.** Given a real number $x \in \mathbf{R}$, let $\langle x \rangle$ denote the distance from $x$ to the nearest integer, i.e., $\langle x \rangle = \min_{n \in \mathbf{Z}} |x - n|$. Let $d_{\mathbf{Z}}(x)$ denoted the *signed* distance from $x$ to the nearest integer, i.e., writing $x = n + x'$ for $n \in \mathbf{Z}$ and $-.5 < x' \leq .5$, let $d_{\mathbf{Z}}(x) = x'$.

The following proposition shows how bracketing meshes with addition:

**Lemma 3.12.** *If $d_{\mathbf{Z}}(a)$ and $d_{\mathbf{Z}}(b)$ are of opposite sign, then*

$$\langle a+b\rangle = |\langle a\rangle - \langle b\rangle|. \tag{3.5}$$

*Else, if $d_{\mathbf{Z}}(a)$ and $d_{\mathbf{Z}}(b)$ are of the same sign,*

$$\langle a+b\rangle = \begin{cases} \langle a\rangle + \langle b\rangle & \text{if } \langle a\rangle + \langle b\rangle \le .5, \\ 1 - (\langle a\rangle + \langle b\rangle) & \text{if } \langle a\rangle + \langle b\rangle > .5. \end{cases} \tag{3.6}$$

*Hence, for any $a$ and $b$,*

$$\langle a+b\rangle \ge |\langle a\rangle - \langle b\rangle|. \tag{3.7}$$

*Proof.* Write $a = n_1 + a'$ and $b = n_2 + b'$ where $n_1, n_2 \in \mathbf{Z}$ and $.5 < a', b' \le .5$.

If $a', b'$ are of opposite sign, then $|a' + b'| \le .5$. Hence

$$\langle a+b\rangle = \langle (n_1+n_2) + (a'+b')\rangle = |a'+b'| = |\langle a\rangle - \langle b\rangle|.$$

If $a', b'$ are of the same sign, then $\langle a+b\rangle = \langle (n_1+n_2)+(a'+b')\rangle = |a'+b'| = |\langle a\rangle + \langle b\rangle|$, unless $|a'+b'| > .5$ in which case, since $|a'+b'| \le 1$, we have that $\langle a+b\rangle = 1 - (\langle a\rangle + \langle b\rangle)$, as was to be shown.

To show the final inequality, it suffices to show that $1 - (\langle a\rangle + \langle b\rangle) \ge |\langle a\rangle - \langle b\rangle|$. This follows by noting that $1-(a'+b') \ge a'-b'$ since $a' \le .5$, and $1-(a'+b') \ge b'-a'$ since $b' \le .5$. □

**Proposition 3.13.** *For $0 < h < q_i + q_{i-1}$ and $h \ne q_i$,*

$$\langle h\xi\rangle \;>\; R_{i-1}.$$

Noting that for $h = q_i$, $\langle h\xi\rangle = R_i$, this proposition may be regarded as an extension of Proposition 3.4. It says that with one exception the $q_{i-1}$-th multiple of $\xi$ comes closer to an integer than any multiple of $\xi$ before it or after it, up to the $(q_i + q_{i-1})$-th multiple.

*Proof.* The statement is by Proposition 3.4 true for all $h < q_i$.

For $q_i < h < q_i + q_{i+1}$, rewrite $h = q_i + h'$. Note that since $h' < q_{i-1}$, we have $\langle h'\xi\rangle > R_{i-2}$. Then $\langle h\xi\rangle = \langle (q_i+h')\xi\rangle \ge \langle h'\xi\rangle - \langle q_i\xi\rangle > R_{i-2} - R_i = a_i R_{i-1}$, using (3.7) and (3.4) above. □

We shall use the $R_i$ as error bounds on the terms which involve the multiples of $\xi$. Namely, partition the unit interval into bins such that the partitions lie at the $R_i$. Drop the multiples of $\xi$ in the bins according to their distance from the nearest integer.

**Definition 3.14.** Call the *$i$-th bin* the interval $[R_i, R_{i-1})$. Let $s$ denote the number of the bin containing 0.5, i.e., $0.5 \in [R_s, R_{s-1})$. Note that either $s = 0$ or $s = 1$.

Let $N_\xi(m,n)$ denote how many of the first $m$ multiples of $\xi$ are in the $n$-th error bin:

$$N_\xi(m,n) = |\{\ q\ \mid 1 \le q \le m \text{ and } \langle q\xi\rangle \in [R_n, R_{n-1})\ \}|\ .$$

(For $n = s$, the restraint on $R_{n-1}$ is ignored.)

The following lemma shows that the multiples of $\xi$ cannot be too close to the partitions $R_j$ (unless they are at the $R_j$ already):

**Lemma 3.15.** *Let $1 \le h < q_i$ and $s \le j \le i-1$. If $h \ne q_j$ or $h \ne q_i - q_j$, then*

$$|\langle h\xi\rangle - R_j| > R_{i-1}\ .$$

*Proof.* For $j \leq i-1$, we have $|\langle h\xi\rangle - R_j| = |\langle h\xi\rangle - \langle q_j\xi\rangle| = \langle g\xi\rangle$, where either $g = h + q_j$ or $g = h - q_j$. This follows from (3.5). But since $h \pm q_j < q_i + q_{i-1}$, and $h + q_j \neq q_i$, and $h - q_j \neq 0$, we may apply Proposition 3.13 to obtain $\langle g\xi\rangle > R_{i-1}$ as was to be shown. □

We will use the lemma to prove the following theorem, which gives a recurrence relation for the $N_\xi(q_i, n)$. It holds for *any* positive irrational, although in the subsequent section we shall apply it when $\xi$ is a quadratic irrational.

**Theorem 3.16.** *Let $\xi = [a_0, a_1, a_2, ...]$ be any positive irrational, and let $q_i = q_i(\xi), R_i = R_i(\xi), N_\xi(m, n)$, and $s = s(\xi)$ be defined as in Definitions 3.3 and 3.14.*

*Set $N_\xi(q_{n-1}, n) = -1 + s\delta_n(s)$. Then for all $n \geq s$, we have $N_\xi(q_n, n) = 1$, and for all $i \geq n$,*

$$N_\xi(q_{i+1}, n) \;=\; a_{i+1}[\; N_\xi(q_i, n) + c(i, n)\;] \;+\; N_\xi(q_{i-1}, n)$$

*where for $i \geq n$*

$$c(i,n) = \begin{cases} 0, & \text{for } n = s \text{ and } i \equiv n \mod 2, \\ -1, & \text{for } n = s \text{ and } i \not\equiv n \mod 2, \\ (-1)^{i-n}, & \text{for } n > s\ . \end{cases}$$

It may be helpful to summarize the conditions in Theorem 3.16: let $\xi = [a_0, a_1, a_2, ...]$ denote the continued fraction expansion of an irrational $\xi$. Running through the multiples of $\xi$ sequentially, drop them on the unit interval one by one at a mark corresponding to their distance from the nearest integer. Every so often, we get a "recordbreaker", a multiple $q_i\xi$ that is closer than all multiples before it. Recall that $R_i$ denoted that distance. Place partitions at the $R_i$, dividing the unit interval into bins. Number the bins so that the $n$-th bin has $R_n$ at the left edge. Let $s$ denote the bin containing $\frac{1}{2}$. Let $N_\xi(m, n)$ denote the number of multiples in the $n$-th bin after $m$ multiples of $\xi$ are thrown down. Then the above recurrence relation holds.

Tables 1–4 exhibit the recurrence for various values of $\xi$. The entries represent $N_\xi(q_i, n)$, which is the count in each bin (columns) up to the $q_i$-th multiple of $\xi$ (rows). The recurrence works down each column. Note that the bins are arranged in decreasing order; this is to remind the reader that the bins are arranged that way on the unit interval. The sum across the $i$-th row should total $q_i$, being the total number of multiples in all the bins. When $\xi > \frac{1}{2}$, the zero-th bin is empty, since $R_0 = \xi$ and $s = 1$. In such cases also we have $q_0 = q_1 = 1$ so that $N(q_0, 1) = N(q_1, 1) = 1$. The data for the tables was generated using Mathematica [12] by actually performing the counting process on the multiples of $\xi$.

The idea of the proof of Theorem 3.16 is as follows. Since the $q_i$ are given by the recurrence $q_{i+1} = a_{i+1}q_i + q_{i-1}$, the second set of $q_i$ multiples of $\xi$ is just a shift of the first set by a distance $R_i$. The above lemma guarantees that most of the first $q_i$-th multiples remain in the same error bins when shifted by $R_i$ up to $a_{i+1}$ times in either direction, since they are far from the partitions. In the remaining cases, the shift occurs in a favorable direction, so the shifts still stay in the same bins. The presence of the $c(i, n)$ term accounts for the multiples which lie at the partitions moving either in or out of the bin when shifted.

*Proof of Theorem 3.16.* Now any $H \leq q_{i+1}$ may be written as

$$H = h + kq_i$$

where $0 \le k \le a_{i+1}$ and $0 \le h < q_i$. If $h > q_{i-1}$, then $k$ need only be $\le a_{i+1} - 1$. Thus

$$\langle H\xi \rangle = \langle (h + kq_i)\xi \rangle = \langle\ \langle h\xi \rangle \pm kR_i\ \rangle.$$

TABLE 1. $N_\xi(q_i, \text{bin}_n)$ for $\xi = \phi = \frac{\sqrt{5}-1}{2} = [0; 1, 1, 1, \ldots] \approx .6180$.

| $\cdots$ | $bin_8$ | $bin_7$ | $bin_6$ | $bin_5$ | $bin_4$ | $bin_3$ | $bin_2$ | $bin_1$ | $bin_0$ | $a_i$ | $i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 3 |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 0 | 1 | 4 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 3 | 2 | 0 | 1 | 5 |
| 0 | 0 | 0 | 1 | 1 | 1 | 3 | 3 | 4 | 0 | 1 | 6 |
| 0 | 0 | 1 | 1 | 1 | 3 | 3 | 7 | 5 | 0 | 1 | 7 |
| 0 | 1 | 1 | 1 | 3 | 3 | 7 | 9 | 9 | 0 | 1 | 8 |
| 1 | 1 | 1 | 3 | 3 | 7 | 9 | 17 | 13 | 0 | 1 | 9 |

TABLE 2. $N_\xi(q_i, \text{bin}_n)$ for $\xi = \sqrt{2} - 1 = [0; 2, 2, 2, ...] \approx .4142$.

| $\cdots$ | $bin_8$ | $bin_7$ | $bin_6$ | $bin_5$ | $bin_4$ | $bin_3$ | $bin_2$ | $bin_1$ | $bin_0$ | $a_i$ | $i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 1 | 2 | 2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 5 | 3 | 2 | 3 |
| 0 | 0 | 0 | 0 | 0 | 1 | 3 | 5 | 15 | 5 | 2 | 4 |
| 0 | 0 | 0 | 0 | 1 | 3 | 5 | 15 | 33 | 13 | 2 | 5 |
| 0 | 0 | 0 | 1 | 3 | 5 | 15 | 33 | 83 | 29 | 2 | 6 |
| 0 | 0 | 1 | 3 | 5 | 15 | 33 | 83 | 197 | 71 | 2 | 7 |
| 0 | 1 | 3 | 5 | 15 | 33 | 83 | 197 | 479 | 169 | 2 | 8 |
| 1 | 3 | 5 | 15 | 33 | 83 | 197 | 479 | 1153 | 409 | 2 | 9 |

TABLE 3. $N_\xi(q_i, \text{bin}_n)$ for $\xi = \frac{\sqrt{3}+5}{22} = [0; 3, 3, 1, 2, \overline{1, 2}, ...] \approx .3060$.

| $\cdots$ | $bin_8$ | $bin_7$ | $bin_6$ | $bin_5$ | $bin_4$ | $bin_3$ | $bin_2$ | $bin_1$ | $bin_0$ | $a_i$ | $i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 5 | 4 | 3 | 2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 5 | 6 | 1 | 3 |
| 0 | 0 | 0 | 0 | 0 | 1 | 3 | 1 | 17 | 14 | 2 | 4 |
| 0 | 0 | 0 | 0 | 1 | 1 | 3 | 3 | 21 | 20 | 1 | 5 |
| 0 | 0 | 0 | 1 | 3 | 1 | 11 | 5 | 61 | 52 | 2 | 6 |
| 0 | 0 | 1 | 1 | 3 | 3 | 13 | 9 | 81 | 72 | 1 | 7 |
| 0 | 1 | 3 | 1 | 11 | 5 | 39 | 21 | 225 | 194 | 2 | 8 |
| 1 | 1 | 3 | 3 | 13 | 9 | 51 | 31 | 305 | 266 | 1 | 9 |

TABLE 4. $N_\xi(q_i, \text{bin}_n)$ for $\xi = \sqrt{14} - 3 = [0; \overline{1, 2, 1, 6}, ...] \approx .7417$.

| $\cdots$ | $bin_8$ | $bin_7$ | $bin_6$ | $bin_5$ | $bin_4$ | $bin_3$ | $bin_2$ | $bin_1$ | $bin_0$ | $a_i$ | $i$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 2 | 2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 1 | 3 |
| 0 | 0 | 0 | 0 | 0 | 1 | 11 | 1 | 14 | 0 | 6 | 4 |
| 0 | 0 | 0 | 0 | 1 | 1 | 11 | 3 | 15 | 0 | 1 | 5 |
| 0 | 0 | 0 | 1 | 3 | 1 | 35 | 5 | 44 | 0 | 2 | 6 |
| 0 | 0 | 1 | 1 | 3 | 3 | 45 | 9 | 58 | 0 | 1 | 7 |
| 0 | 1 | 11 | 1 | 27 | 13 | 311 | 53 | 392 | 0 | 6 | 8 |
| 1 | 1 | 11 | 3 | 29 | 17 | 355 | 63 | 449 | 0 | 1 | 9 |

We claim that

**Lemma 3.17.** *If $h \neq 0$ and $h \neq q_j$ for all $j \geq s$, then $\langle H\xi \rangle$ falls into the same bin as $\langle h\xi \rangle$.*

*Proof of Lemma 3.17.* Since $h \neq q_j$, Lemma 3.15 and equation (3.4) show that when $h \neq q_i - q_j$, we have

$$|\langle h\xi \rangle - R_j| > R_{i-1} > a_{i+1} R_i \ .$$

Thus the partitions are too far away for $\langle h\xi \rangle \pm kR_i$, $k \leq a_{i+1}$, to fall into a different bin than $\langle h\xi \rangle$. If $\langle h\xi \rangle$ falls in the first bin (the $s$-th), then we must verify that $\langle h\xi \rangle + kR_i < 1 - R_s$; this follows from $\langle h\xi \rangle - kR_i > R_s$ and $\langle h\xi \rangle < 1 - \langle h\xi \rangle$.

We are left to consider the case where $h = q_i - q_j$ for some $j \leq i - 1$.

Observe that we need not consider the case where $a_i = 1$ and $j = (i-1)$ or $(i-2)$, since otherwise $h = q_{i-2}$ or $q_{i-1}$, which is ruled out by assumption. Nor do we need consider the case when $i = 1$, $a_i = 2$, and $j = i - 1 = 0$, for then we would have $q_1 = 2q_0$, giving $h = q_1 - q_0 = q_0$, which is ruled out by assumption.

In all other cases, $h = q_i - q_j > q_{i-1}$, so that we may assume $k \leq a_{i+1} - 1$ as observed earlier. Also, hereafter we may assume that $a_i > 1$ if $j = (i-1)$ or $(i-2)$, and that $a_i > 2$ if $j = (i-1) = 0$.

Then $\langle H\xi \rangle = \langle ((k+1)q_i - q_j)\xi \rangle = \langle R_j - (-1)^{i-j}(k+1)R_i \rangle$. This expression is obtained by recalling that the $q_i\xi$ alternate signs with $i$.

For $i, j$ of the same parity, we have $j \leq i - 2$, and we wish to show that for $0 \leq k \leq a_{i+1} - 1$, all the $\langle H\xi \rangle = \langle (h + kq_i)\xi \rangle$ fall into the same bin, the $(j+1)$-st bin: $[R_{j+1}, R_j)$. ($\langle h\xi \rangle$ is the case when $k = 0$.)

This amounts to showing that $\langle H\xi \rangle = R_j - (k+1)R_i > R_{j+1}$, or, equivalently, that $R_j - R_{j+1} > (k+1)R_i$. Since $R_j = a_{j+2}R_{j+1} + R_{j+2}$ and $R_{i-1} = a_{i+1}R_i + R_{i+1} > a_{i+1}R_i \geq (k+1)R_i$, it suffices to show

$$(a_{j+2} - 1)R_{j+1} \ + \ R_{j+2} \ > \ R_{i-1}.$$

This will be demonstrated in two cases.

*Case 1.* If $j = i - 2$, then we may assume that $a_i = a_{j+2} > 1$, and we have

$$(a_{j+2} - 1)R_{j+1} \ + \ R_{j+2} \geq R_{j+1} + R_{j+2} > R_{j+1} \geq R_{i-1}.$$

*Case 2.* If $j \leq i - 4$, then

$$(a_{j+2} - 1)R_{j+1} \ + \ R_{j+2} \geq R_{j+2} \geq R_{i-2} > R_{i-1}.$$

For $i, j$ of different parity, we have $j \leq i-1$, and we wish to show that for $0 \leq k \leq a_{i+1}-1$, all the $\langle H\xi\rangle = \langle (h+kq_i)\xi\rangle$ fall into the same bin, the $j$-th bin: $[R_j, R_{j-1})$. ($\langle h\xi\rangle$ is the case when $k=0$.)

For $j > s$, this amounts to showing that $\langle H\xi\rangle = R_j + (k+1)R_i < R_{j-1}$, or, equivalently, that $R_{j-1} - R_j > (k+1)R_i$. The following argument works for all $j \geq 1$, and so includes the case $j > s$.

Since $R_{j-1} = a_{j+1}R_j + R_{j+1}$ and $R_{i-1} = a_{i+1}R_i + R_{i+1} > a_{i+1}R_i \geq (k+1)R_i$, it suffices to show

$$(a_{j+1}-1)R_j \ + \ R_{j+1} \ > \ R_{i-1}.$$

This will be demonstrated in two cases.

*Case 1.* If $j = i-1$, then we may assume that $a_i = a_{j+1} > 1$, and we have

$$(a_{j+1}-1)R_j + R_{j+1} \geq R_j + R_{j+1} > R_j \geq R_{i-1}.$$

*Case 2.* If $j \leq i-3$, then

$$(a_{j+1}-1)R_j + R_{j+1} \geq R_{j+1} \geq R_{i-2} > R_{i-1}.$$

We now treat the boundary case $j = s$.

If $j = s = 1$, then we wish to show that $\langle H\xi\rangle = R_1 + (k+1)R_i < 1 - R_1$ which would show that $\langle H\xi\rangle$ remained in the $s$-th bin. However, $s = 1$ implies $a_1 = 1$ and by (3.4), $R_0 = 1 - R_1$. Thus this reduces to the previous case of showing $R_j + (k+1)R_i < R_{j-1}$ for $j = 1$.

If $j = s = 0$, then we wish to show that $\langle H\xi\rangle = R_0 + (k+1)R_i < 1 - R_0$ which would show that $\langle H\xi\rangle$ remained in the $s$-th bin. However, (3.4) implies that $1 - R_0 = (a_1 - 1)R_0 + R_1$. So it suffices to show $(k+1)R_i < (a_1 - 2)R_0 + R_1$. If $i \geq 3$, (3.4) yields $(k+1)R_i < R_2$ which is less than $R_1$ and we are done. If $i = 1$, then recall that we noted earlier that we can assume that $a_1 > 2$ if $j = i - 1 = 0$. Again, (3.4) yields $(k+1)R_i < R_0$ which is less than $(a_1 - 2)R_0$ and we are done.

This completes the proof of Lemma 3.17. □

We return the proof of Theorem 3.16.

The preceding lemma showed that multiples before the $q_i$-th which are not equal to a $q_j$-th for some $j$ do not get shifted out of their current bins when multiples of $q_i\xi$ are added.

We now determine what happens to the $q_j$-th multiples, which fall at the partitions $R_j$. When adding multiples of $q_i\xi$, they get shifted into either the bin on the left or bin on the right, but no further, except for the last ($q_{i+1}$-th) multiple of $\xi$:

**Lemma 3.18.** *For $q_i < H \leq q_{i+1}$, write $H = h + kq_i$, where $1 \leq k \leq a_{i+1}$.*

*Suppose $h = 0$. Then for $1 \leq k \leq a_{i+1}$, $\langle H\xi\rangle$ falls in the $i$-th bin: $[R_i, R_{i-1})$, unless $i = s = 0$ and $k = a_{i+1}$, in which case, $\langle H\xi\rangle$ falls in the 2-nd bin: $[R_2, R_1)$.*

*Suppose $h = q_j$ for some $s \leq j < i$.*

*If $i, j$ are of the same parity, then $\langle H\xi\rangle$ falls in the $j$-th bin: $[R_j, R_{j-1})$.*

*If $i, j$ are of opposite parity, then $\langle H\xi\rangle$ falls in the $(j+1)$-st bin: $[R_{j+1}, R_j)$, unless $H = q_{i+1}$ (when $j = i-1$ and $k = a_{i+1}$). In that case $\langle H\xi\rangle$ falls in the $(j+2)$-nd bin: $[R_{j+2}, R_{j+1})$.*

*Proof of Lemma 3.18.* Assume $h = 0$. Then $\langle q_i\xi\rangle = R_i$, and we shall show that for $1 < k \leq a_{i+1}$, $\langle H\xi\rangle = \langle kq_i\xi\rangle$ also falls in the $i$-th bin: $[R_i, R_{i-1})$, barring one exception. This amounts to showing that $kR_i < R_{i-1}$ which is true by (3.4), unless $i = s$.

In that case, in light of (3.6) we must show that $kR_s < 1 - R_s$. For $s = 1$ this follows from $1 - R_1 = R_0$ and (3.4); for $s = 0$ we need to verify that $kR_0 < 1 - R_0 = (a_1 - 1)R_0 + R_1$, which is true for $k < a_1 - 1$, but false for $k = a_1$. When $k = a_1$, we have $H = q_2$, and $\langle H\xi \rangle$ falls in the 2-nd bin: $[R_2, R_1)$, as was to be shown.

Now assume for the remainder of the proof of this lemma that $h = q_j$ for some $s \le j < i$.

If $i, j$ are of the same parity, then $\langle H\xi \rangle = \langle (q_j + kq_i)\xi \rangle = \langle R_j + kR_i \rangle$. We wish to show that $R_j + kR_i$ is always less than $R_{j-1}$, or if $j = s$, less than $1 - R_s$. This would show that $\langle H\xi \rangle$ falls in the $j$-th bin.

For $j > s$ we have two cases:

*Case 1.* If $j = i$, then since $k < a_{i+1} - 1$,

$$R_j + kR_i \le a_{i+1}R_i < R_{i-1} = R_{j-1}.$$

*Case 2.* If $j \le i - 2$, then

$$R_j + kR_i \le R_j + a_{i+1}R_i = R_j + R_{i-1} - R_{i+1} < R_j + R_{i-1} \le R_j + R_{j+1} \le R_{j-1}.$$

For $j = s$ we note that if $j = s = 1$, then $a_1 = 1$, which implies $1 - R_1 = R_0$, and we remark that the previous arguments to show that $R_j + kR_i < R_{j-1}$ are valid for $j = 1$. If $j = s = 0$, then $a_1 > 1$ and $1 - R_0 = (a_1 - 1)R_0 + R_1$. We must show that $R_0 + kR_i < (a_1 - 1)R_0 + R_1$. This follows from $kR_i < R_1$ since $i \ge 2$.

If $i, j$ are of opposite parity, then $\langle H\xi \rangle = \langle (q_j + kq_i)\xi \rangle = R_j - kR_i$. We wish to show this quantity is always greater $R_{j+1}$, which would show that $\langle H\xi \rangle$ falls in the $(j + 1)$-st bin.

*Case 1.* If $j = i - 1$ and if $k \le a_{i+1} - 1$,

$$\begin{aligned} R_j - kR_i &\ge R_{i-1} - (a_{i+1} - 1)R_i = a_{i+1}R_i + R_{i+1} - (a_{i+1} - 1)R_i \\ &= R_i + R_{i+1} > R_{j+1}. \end{aligned}$$

If $j = i-1$ and $k = a_{i+1}$, then $R_j - kR_i = R_{i+1} = R_{j+2}$, which lies in the $(j+2)$-nd bin.

*Case 2.* If $j \le i - 3$, then

$$\begin{aligned} R_j - kR_i &\ge R_j - a_{i+1}R_i = R_j - R_{i-1} + R_{i+1} > R_j - R_{i-1} \ge R_j - R_{j+2} \\ &= a_{j+2}R_{j+1} > R_{j+1}. \end{aligned}$$

This concludes the proof of the Lemma 3.18. □

We proceed to use the above lemmas to derive a recurrence relation for the $N_\xi(q_i, n)$.

Since $s$ denotes the first bin that could possibly be occupied, we have $N_\xi(q_i, n) = 0$ for all $n < s$.

Consider the $n$-th error bin: $[R_n, R_{n-1})$. As multiples are thrown onto the bins, it will remain empty until $i = n$, when $R_i = \langle q_i\xi \rangle$ occupies the bin at the left edge. These statements imply that $N_\xi(q_i, n) = 0$ for all $s \le i < n$, and $N_\xi(q_n, n) = 1$ for all $n \ge s$.

We wish to count how many of the first $q_{i+1}$ multiples of $\xi$ are in the $n$-th bin. Since $q_{i+1} = a_{i+1}q_i + q_{i-1}$, we will do so by breaking the multiples into subsets, and considering the set of the first $q_i$ multiples of $\xi$, then the next $(a_{i+1} - 1)$ sets of $q_i$ multiples of $\xi$, then the last set of $q_{i-1}$ multiples of $\xi$.

For $i \ge n$, the $n$-th error bin contains $N_\xi(q_i, n)$ of the set of the first $q_i$ multiples of $\xi$.

The next $(a_{i+1}-1)$ sets containing $q_i$ multiples each are like the first set but shifted a bit, and the above lemmas show that they fall in the same bins as before, except possibly for the shifts of those multiples which lie at the endpoints of the bin. Hence each of these sets contributes $N_\xi(q_i,n)+c(i,n)$ number of multiples to the $n$-th bin, where $c(i,n)$ is some correcting factor to be determined now.

When $i \geq n$ and $n > s$, notice that multiples lie at both endpoints of $[R_n, R_{n-1})$. $R_n$ is originally in the bin, and $R_{n-1}$ is originally out. Lemma 3.18 shows that shifts of these endpoints are either both in or both out according to its parity with $i$, so that the correcting factor $c(i,n)$ should be $(-1)^{i-n}$.

For $i \geq n$ and $n = s$, there is a multiple at $R_n$ but not at the endpoint $R_{n-1}$, so the right-hand endpoint will contribute nothing when shifted, giving a correcting factor $c(i,n)$ of 0 or $-1$ according as the left endpoint $R_n$ stays in or is shifted out, i.e., as the parity of $i, n$ agree or disagree. One exception to this rule occurs when $i = n = s = 0$, in the very last shift $k = a_1$, when the multiple at $R_0$, the left-hand endpoint, is according to Lemma 3.18 shifted out of the 0-th bin (into the 2-nd bin). The counting procedure did not take that into account, so we must subtract 1 from the total contribution if $n = i = 0$. We must also add 1 to the total contribution if $n = i + 2 = 2$; however, since the recurrence will only be operative for $i \geq n$, we need not worry about this case.

Thus the total contribution from the $(a_{i+1}-1)$ sets containing $q_i$ multiples is $(a_{i+1}-1)[N_\xi(q_i,n)+c(i,n)] - \delta_n(i)\delta_n(0)$.

The last set of $q_{i-1}$ multiples is a shift of the first $q_{i-1}$ multiples of $\xi$, and again, shifting does not take them out of their original bins, except possibly for the endpoints *and* for the last shift of $q_{i-1}$. In that case, by Lemma 3.18, $R_{i-1}$ is shifted down not to the $i$-th bin, but to the $(i+1)$-st. We now investigate how this affects the count.

For $n = i$, the usual count $N_\xi(q_{n-1},n)+c(n,n)$ must be decreased by 1 since it should no longer include the shift of the right-hand endpoint $R_{n-1}$. (It is helpful to remember that the index $i$ in $c(i,n)$ refers to the shift increments and not to the size of the set being shifted.) The only possible exception is when the right-hand endpoint does not exist at $i = s$; if $n = i = s = 1$, then $N_\xi(q_0 = q_1, 1) = 1$, the left endpoint exists and must be eliminated so we still need to decrease the count by 1; if $i = n = s = 0$, the count is already 0 and there is no right-hand endpoint, so we do not need to decrease the count by 1.

For $n = (i+1)$, we have $i < n$, which will not affect the validity of the formula for $i \geq n$. (It is in fact this extra multiple that accounts for the start of the recurrence: $N_\xi(q_n,n) = 1$, as determined before.)

Thus the contribution from this last set is

$$N_\xi(q_{i-1},n) + c(i,n) - \delta_n(i)(1-\delta_n(0)).$$

To summarize the contributions, we have for $i \geq n \geq s$:

$$\begin{aligned} N_\xi(q_{i+1},n) &= N_\xi(q_i,n) + (a_{i+1}-1)[N_\xi(q_i,n)+c(i,n)] - \delta_n(i)\delta_n(0) \\ &\quad + N_\xi(q_{i-1},n) + c(i,n) - \delta_n(i)(1-\delta_n(0)) \\ &= a_{i+1}[\, N_\xi(q_i,n) + c(i,n)\,] + N_\xi(q_{i-1},n) - \delta_n(i) \end{aligned}$$

where

$$c(i,n) = \begin{cases} 0, & \text{for } n = s \text{ and } i \equiv j \mod 2, \\ -1 & \text{for } n = s \text{ and } i \not\equiv j \mod 2, \\ (-1)^{i-n}, & \text{for } n > s\ . \end{cases}$$

Note that for $i = n$, $N_\xi(q_{n-1}, n) = 0$ except if $n = s = 1$, when $N_\xi(q_0 = q_1, 1) = 1$. Since the recurrence will only be operative for $i \geq n$, we choose to subsume the $\delta_n(i)$ in the definition of $N_\xi(q_{n-1}, n)$, by setting $N_\xi(q_{n-1}, n) = -1 + s\delta_n(s)$.

The other initial case is given by $N_\xi(q_n, n) = 1$. This concludes the proof of Theorem 3.16. □

We can use this recurrence to obtain the following absolute bound on $N_\xi(q_i, n)$:

**Theorem 3.19.** *Let $\xi$ be any positive quadratic irrational. Let $q_{i,n} = q_i(r_n)$ denote the denominator of the $i$-th convergent of $r_n$, the $n$-th remainder of $\xi$. Then for $i \geq n$,*

$$N_\xi(q_i, n) \ \leq \ 2q_{i-n,n} - 1.$$

*Proof.* We shall induct on $i$. The initial cases are easily verified:

$$N_\xi(q_n, n) = 1 \leq 2q_0(r_n) - 1 \quad = \quad 1.$$

And by Theorem 3.16,

$$\begin{aligned}
\text{For } n > s, \qquad N_\xi(q_{n+1}, n) = 2a_{n+1} - 1 &\leq 2q_1(r_n) - 1 = 2a_{n+1} - 1.\\
\text{For } n = s = 0, \qquad N_\xi(q_{s+1}, s) = a_{s+1} - 1 &\leq 2q_1(r_s) - 1 = 2a_{s+1} - 1.\\
\text{For } n = s = 1, \qquad N_\xi(q_{s+1}, s) = a_{s+1} &\leq 2q_1(r_s) - 1 = 2a_{s+1} - 1.
\end{aligned}$$

Now assume the theorem is true for all $i \leq k$. Then by Theorem 3.16,

$$\begin{aligned}
N_\xi(q_{k+1}, n) &\leq a_{k+1}[N_\xi(q_k, n) + c(k, n)] + N_\xi(q_{k-1}, n)\\
&\leq a_{k+1}[2q_{k-n}(r_n)] + 2q_{k-n-1}(r_n) - 1\\
&\leq 2q_{k-n+1}(r_n) - 1
\end{aligned}$$

which shows the theorem is true for $i = k + 1$. □

## 4. Bounds for Quadratic Irrationals

As previously noted, for the $\pm\alpha$ walk on $S^1$ we shall be interested in the continued fraction expansion of $\xi = 2\alpha$. We shall first establish some notation and definitions for quadratic irrationals which are needed to prove the main theorem of this section, Theorem 4.9.

**Definition 4.1.** The continued fraction $\xi = [a_0, a_1, a_2, ...]$ is called *periodic* if its elements eventually repeat, i.e., if there exist a $k_0$ and $h$ such that, for $k \geq k_0$,

$$a_k = a_{k+h}.$$

**Example 4.2.** The number $\sqrt{3} = [1; 1, 2, \overline{1, 2}, ...]$ is periodic of period 2. The number $\frac{1}{3} + \sqrt{5} = [2; \overline{1, 1, 3, 9, 1, 3}, ...]$ is periodic of period 6. The number $\frac{14-\sqrt{7}}{27} = [0; 2, 2, \overline{1, 1, 1, 4}, ...]$ is periodic of period 4, although it does not start to repeat until place $k_0 = 3$.

**Definition 4.3.** A number $\xi$ is said to be a *quadratic irrational* if it is the root of a second-degree polynomial with integral coefficients.

The following theorem is due to Lagrange. For a proof, see Khinchin [6, Thm.28].

**Theorem 4.4** (Lagrange)**.** *A number is represented by a periodic continued fraction if and only if it is a quadratic irrational.*

We shall use the periodic nature of quadratic irrationals together with the recurrence relation derived in the last section to provide good estimates for the sum that occurs in the theorem of Erdős-Turán.

Recall that the $r_i$'s represent the remainders of a continued fraction. Note that if the continued fraction is periodic, then its remainders are also periodic.

**Definition 4.5.** Let $h$ be the period of the continued fraction expansion of a quadratic irrational $\xi$, and $k_0$ the element at which it begins repeating. Define

$$\overline{r} = (\prod_{i=k_0+1}^{k_0+h} r_i)^{1/h}.$$

It is useful to think of $\overline{r}$ as a kind of "average remainder". Notice that $\prod_{i=1}^{n} \frac{r_i}{\overline{r}}$ is bounded and eventually repeating in $n$. Let $J = J(\xi) \leq 1$ be the minimum value of this product, and $K = K(\xi) \geq 1$ be its maximum value. (1 is achieved for the vacuous product at $n = 0$.)

Proposition 3.5 shows that

$$J \leq \overline{r}^{k+1} R_k \leq K$$

which may be rewritten

$$\frac{J}{\overline{r}^{k+1}} \leq R_k \leq \frac{K}{\overline{r}^{k+1}}. \tag{4.1}$$

Furthermore, by (3.3), we have

$$\frac{\overline{r}^{k-1}}{K} < q_k < \frac{\overline{r}^k}{J}. \tag{4.2}$$

We should also bound the $q_{i-n,n}$, which were defined in Theorem 3.19:

**Lemma 4.6.** *Given $\xi$, let $\overline{r}, J, K$ be as defined in Definition 4.5. Then*

$$\frac{J}{K}\overline{r}^{i-n-1} \leq q_{i-n,n} \leq \frac{K}{J}\overline{r}^{i-n}. \tag{4.3}$$

*Proof.* Using (3.3), Proposition 3.5, and (4.1) we have that

$$q_{i-n,n} \leq \frac{1}{R_{i-n-1}(r_n)} = r_{n+1}r_{n+2}\cdots r_i = \frac{R_{n-1}}{R_{i-1}} \leq \frac{K}{J}\overline{r}^{i-n}.$$

Similarly,

$$q_{i-n,n} \geq \frac{1}{R_{i-n-2}(r_n)} = r_{n+1}r_{n+2}\cdots r_{i-1} = \frac{R_{n-1}}{R_{i-2}} \geq \frac{J}{K}\overline{r}^{i-n-1}.$$

These inequalities yield the conclusion of the lemma. □

The average remainder $\overline{r} = \overline{r}(\xi)$ can be no smaller than $\overline{r}(\phi) = \Phi$. Recall that $\Phi = \frac{\sqrt{5}+1}{2}$ is the golden mean, and $\phi = 1/\Phi$.

**Proposition 4.7.** *For any irrational $\xi$, $\overline{r}(\xi) \geq \Phi$.*

*Proof.* One may show for any irrational $\xi$ that $q_k(\phi) \leq q_k(\xi)$, by using induction, noting that $\phi = [0; 1, 1, 1, ...]$ has the smallest possible elements, and using the recurrence (3.1) as the inductive step.

Note that for all $k > 0$, $r_k(\phi) = \Phi$. Now suppose for some $\xi$ that $\overline{r} = \overline{r}(\xi)$ were less than $\Phi$, say $\overline{r} = c\Phi$ for some $c < 1$. Then by equation (4.2) one would obtain, for all $k > 0$,

$$\frac{\Phi^{k-1}}{K} < q_k(\phi) \leq q_k(\xi) < \frac{\overline{r}^k}{J} = \frac{c^k \Phi^k}{J}.$$

This implies that $\frac{J}{K\Phi} < c^k$ for all $k > 0$, which clearly cannot be true for $k$ chosen large enough.

Hence for all $\xi$, $\overline{r}(\xi)$ must be greater than or equal to $\Phi$. □

**Example 4.8.** For $\xi = \sqrt{3} - 1 \approx .7321$ we have $\xi = [0, 1, 2, 1, 2, ...]$, which is period two. $r_1 = [1, 2, 1, 2, ...] \approx 1.366$ and $r_2 = [2, 1, 2, 1, ...] \approx 2.732$ and so $\overline{r} = (r_1 r_2)^{\frac{1}{2}} \approx 1.932$. $J = \frac{r_1}{\overline{r}} \approx .7071$, and $K = \frac{r_1 r_2}{\overline{r}^2} = 1$. Also $\frac{K}{J} = \sqrt{2}$.

4.1. **A Rate using Erdős-Turán's Inequality.** Recall that $Q_\alpha$ denotes the generating measure for the $\pm\alpha$-walk. We are interested bounding the discrepancy of the $k$-th step probability distribution from the uniform distribution. The following theorem gives matching upper and lower bounds for this walk, when $\alpha$ is a quadratic irrational.

**Theorem 4.9.** *Suppose $\xi = 2\alpha$ is a quadratic irrational, and let $\overline{r}, J, K$ be as defined in Definition 4.5 for $\xi$. Then the discrepancy of the $\pm\alpha$ random walk on the circle satisfies*

$$\frac{C_1}{\sqrt{k}} \leq \ D(Q_\alpha^{*k}, U) \ \leq \frac{C_2}{\sqrt{k}}$$

*where the constants depend on $\xi$ and can be taken to be $C_1 = \frac{1}{2}$ and $C_2 = \frac{K^4}{J^4}\overline{r}^9$.*

*Proof.* The lower bound follows from Theorem 2.2.

For the upper bound, we appeal to the theorem of Erdős-Turán (Theorem 2.4):

$$D(Q_\alpha^{*k}, U) \leq \frac{4}{m} + \frac{4}{\pi} \sum_{h=1}^{m} \frac{|\widehat{Q_\alpha}^k(h)|}{h}, \tag{4.4}$$

an inequality true for all $m$. We shall choose $m$ so that both terms in the above expression decrease like $\frac{1}{\sqrt{k}}$.

Recall that $\widehat{Q_\alpha}(h) = \cos(2\pi h\alpha) = \cos(\pi h \xi)$.

Choose $l$ such that $q_{l-1} \leq \sqrt{k} \leq q_l$. From Theorem 2.1 we have

$$D(Q_\alpha^k, U) \leq D(Q_\alpha^{q_{l-1}^2}, U) \leq \frac{4}{m} + \frac{4}{\pi} \underbrace{\sum_{h=1}^{m} \frac{|\cos^{q_{l-1}^2}(\pi h \xi)|}{h}}_{S}. \tag{4.5}$$

Let $m = q_l$. We wish to analyze the sum S in the above expression:

$$\begin{aligned} S &\le |\cos^{q_{l-1}^2}(\pi\xi)| \ + \sum_{i=s+1}^{l} \sum_{h=q_{i-1}+1}^{q_i} \frac{|\cos^{q_{l-1}^2}(\pi h\xi)|}{q_{i-1}+1} \\ &\le \exp\left(-q_{l-1}^2\pi^2 R_s^2/2\right) \\ &\qquad + \sum_{i=s+1}^{l} \sum_{n=s}^{i} \frac{[N_\xi(q_i,n) - N_\xi(q_{i-1},n)]}{q_{i-1}+1} \ \exp\left(-q_{l-1}^2\pi^2 R_n^2/2\right) \\ &\le \sum_{n=s}^{l} \sum_{i=n}^{l} \frac{2q_{i-n,n}}{q_{i-1}+1} \ \exp\left(-q_{l-1}^2\pi^2 R_n^2/2\right). \end{aligned}$$

The first inequality follows by splitting up the sum into groups by the $q_i$. Note that $q_s = 1$, but since $q_s$ may equal $q_{s-1}$ (it does when $s = 1$) we treat that term separately. The second inequality follows from grouping the multiples of $\xi$ by the bins they fall in, bounding them by the $R_n$, and using $\cos x \le \exp(-x^2/2)$ for $|x| \le \frac{\pi}{2}$. (Note that by definition $R_s < \frac{1}{2}$.) The last inequality follows from Theorem 3.19 and switching the order of summation. Note that the boundary case $n = s$ has been subsumed within this bound, because $q_{0,s} = 1$ so that the factor $\frac{2q_{i-n,n}}{q_{i-1}+1} \ge 1$.

Using (4.3) and (4.1), and noting from (4.2) that $\frac{1}{q_{i-1}+1} < \frac{K}{\overline{r}^{i-2}}$ even when $i = 0$, we obtain

$$\begin{aligned} S &\le \sum_{n=0}^{l} \sum_{i=n}^{l} 2 \ \frac{K}{\overline{r}^{i-2}} \ \frac{K}{J}\overline{r}^{i-n} \ \exp\left(-(\frac{\overline{r}^{l-2}}{K})^2 \ \frac{\pi^2}{2} \ (\frac{J}{\overline{r}^{n+1}})^2\right) \\ &= \sum_{n=0}^{l} \sum_{i=n}^{l} \frac{2K^2}{J}\overline{r}^{2-n} \ \exp\left(-\frac{(\overline{r}^{l-n-3}\pi J)^2}{2K^2}\right) \\ &= \frac{J}{\overline{r}^l} \sum_{n=0}^{l} (l-n+1)\frac{2K^2}{J^2}\overline{r}^{2-n+l} \ \exp\left(-\frac{(\overline{r}^{l-n-3}\pi J)^2}{2K^2}\right) \\ &\le \frac{1}{q_l} \sum_{N:=l-n+1=1}^{l+1} \frac{2K^2\overline{r}}{J^2} N \ \overline{r}^N \ \exp\left(-\frac{\overline{r}^{2N}\pi^2 J^2}{2K^2\overline{r}^8}\right) \\ &\le \frac{1}{\sqrt{k}} \ \frac{\pi^2}{W\overline{r}^7} \sum_{N=1}^{\infty} N \ \overline{r}^N \ \exp(-W\overline{r}^{2N}) \end{aligned}$$

where $W = \frac{\pi^2 J^2}{2K^2\overline{r}^8}$. We wish to bound the sum in the last expression, which is a constant independent of $k$ (but not of $\xi$). Proposition 4.7 shows $\overline{r} \ge \Phi$, which implies $n < \overline{r}^n$ and

$$\sum_{N=1}^{\infty} N \ \overline{r}^N \ e^{-W\overline{r}^{2N}} \le \sum_{N=1}^{\infty} \overline{r}^{2N} \ e^{-W\overline{r}^{2N}}.$$

Now for any $f(x)$ which is unimodal on $[a, b]$,

$$\sum_{i=a}^{b} f(i) \le \int_a^b f(x)dx + \max_{[a,b]} f(x).$$

Observe that $f(x) = \overline{r}^{2x} e^{-W\overline{r}^{2x}}$ is unimodal, and has maximum value $\frac{1}{eW}$. Then

$$\begin{aligned}\sum_{N=1}^{\infty} \overline{r}^{2N} e^{-W\overline{r}^{2N}} &\leq \int_1^{\infty} \overline{r}^{2x} e^{-W\overline{r}^{2x}} dx + \frac{1}{eW} \\ &\leq \frac{e^{-W\overline{r}^{2x}}}{-W2\ln\overline{r}} \Big|_1^{\infty} + \frac{1}{eW} \\ &= \frac{1}{W}\left(\frac{e^{-W\overline{r}^2}}{2\ln\overline{r}} + \frac{1}{e}\right) \leq \frac{1.5}{W}.\end{aligned}$$

Combining what we know we find that

$$\sum_{h=1}^{m} \frac{|\cos^{q_{l-1}^2}(\pi h\xi)|}{h} \leq \frac{1}{\sqrt{k}} \frac{1.5\pi^2}{W^2\overline{r}^7} = \frac{1}{\sqrt{k}} \frac{6K^4\overline{r}^9}{\pi^2 J^4}.$$

Hence we have from (4.5)

$$D(Q_\alpha^{*k}, U) \leq \frac{4}{q_l} + \frac{4}{\pi} \sum_{h=1}^{q_l} \frac{|\cos^{q_{l-1}^2}(\pi h\xi)|}{h} \leq \frac{1}{\sqrt{k}} \left(4 + \frac{24}{\pi^3} \frac{K^4\overline{r}^9}{J^4}\right) \leq \frac{1}{\sqrt{k}} \left(\frac{K^4\overline{r}^9}{J^4}\right).$$

The last inequality follows from Proposition 4.7. □

*Remark* 4.10. For $\xi = \phi \approx 0.618$ the constant $C_2$ in the upper bound can be improved to $\overline{r}^7$, since $N_\phi(q_i, n) - N_\phi(q_{i-1}, n) = N_\phi(q_{i-2}, n) + c(i-1, n) \leq 2q_{i-n-2,n}$ for $i \geq n+2$. Also $J = K = 1$ since all remainders are equal, and equal to $\overline{r} = \Phi \approx 1.618$. Thus $C_2 \approx 29.04$.

## 5. Asymptotic Results for the Arbitrary Irrational

In this section we obtain bounds for the rate of convergence of the $\pm\alpha$ random walk for any irrational $\alpha$. They are less precise than the bounds we derived for quadratic irrationals but give estimates for arbitrary $\alpha$ based on a "type" classification of irrationals as described below. The main results of this section are Theorems 5.5 and 5.8, which give upper and lower bounds for the walk that differ only by an arbitrarily small $\epsilon$ in the exponent.

### 5.1. **Classification of Irrational Numbers.**

Irrational numbers may be classified by how quickly their multiples approach integers. Evidently this classification will affect how quickly the $\pm\alpha$ walk converges to uniform. Therefore, in this section we shall recall some definitions from the theory of diophantine approximation which will be useful to us later. Our notation and exposition follows that of Kuipers and Niederreiter [7].

**Definition 5.1.** An irrational $\xi$ is said to be of *type* $\eta$ if $\eta$ is the supremum of all $\gamma$ such that $\liminf_{q\to\infty} q^\gamma \langle q\xi\rangle = 0$.

*Remark* 5.2. (a) A standard result on continued fractions is the fact that for the convergents $\frac{p_n}{q_n}$ of an irrational $\xi$, $|\xi - \frac{p_n}{q_n}| < \frac{1}{q_n^2}$. From this it follows that the type of any irrational number will satisfy $\eta \geq 1$.

(b) Furthermore, it follows from the theorem of Thue-Siegel-Roth (see, for example, Schmidt [19, p.116]) that $\eta = 1$ for all algebraic irrationals (in particular, quadratic irrationals). Irrationals whose continued fractions have bounded elements are also known to be of type 1. Both these sets of irrationals are of measure 0 in

the unit interval. However, there are many more type 1 irrationals, since type 1 irrationals are of full measure in the unit interval (this follows from Khinchin [6, Thm.32]).

(c) There exist numbers of type $\eta = \infty$, called Liouville numbers. Liouville used them to show that transcendental numbers exist, since algebraic numbers cannot admit as good rational approximations as these, by Liouville's Theorem (see Khinchin [6, Thm.27]). One such number is $\xi = \sum_{i=1}^{\infty} 10^{-i!}$. For if $q = 10^{j!}$, then $\langle q\xi \rangle \leq 10^{-(j+1)!+j!+1}$, and we have, for any $d$,

$$q^d \langle q\xi \rangle \leq 10^{j!d} 10^{-(j+1)!+j!+1} = 10^{j!(d-j)+1}$$

which approaches zero for large enough $j$. This shows that $\liminf_{q\to\infty} q^d \langle q\xi \rangle = 0$ for any $d$.

**Definition 5.3.** Let $\psi$ be a nondecreasing positive function defined on the positive integers. Then an irrational number $\xi$ is said to be of *type* $< \psi$ if $q\langle q\alpha \rangle \geq 1/\psi(q)$ holds for all positive integers $q$. If $\psi$ is a constant function, then we say $\xi$ is *of constant type*.

These two definitions are related by the following lemma.

**Lemma 5.4.** *The irrational number $\xi$ is of type $\eta$ if and only if $\eta$ is the infimum of all real numbers $\tau$ for which there exists a positive constant $c = c(\tau, \xi)$ such that $\xi$ is of type $< \psi$ where $\psi(q) = cq^{\tau-1}$.*

The following proof, included here for completeness, is taken directly from Kuipers and Niederreiter [7, p.121].

*Proof.* Let $\eta$ be finite. $\xi$ of type $\eta$ implies that, for any $\epsilon > 0$,

$$\liminf_{q\to\infty} \; q^{\eta-\epsilon} \langle q\xi \rangle = 0 \tag{5.1}$$

and

$$\limsup_{q\to\infty} \; q^{\eta+\epsilon} \langle q\xi \rangle > 0. \tag{5.2}$$

(5.1) implies that for any positive $c$ there is a positive integer $q$ such that $q\langle q\xi \rangle < 1/cq^{\eta-1-\epsilon}$. Hence $\xi$ is not of type $< \psi$ for any $\psi$ of the form $\psi(q) = cq^{\eta-1-\epsilon}$. However, (5.2) implies that for any $\epsilon > 0$ there is a positive constant $a(\epsilon, \xi)$ such that for all $q$, $q^{\eta+\epsilon}\langle q\xi \rangle \geq a(\epsilon, \xi)$. Thus $\xi$ is of type $< \psi$ for $\psi(q) = (\frac{1}{a(\epsilon,\xi)}) q^{\eta-1+\epsilon}$. Reversing the arguments yields the converse.

If $\eta = \infty$, the same ideas work with obvious modifications, interpreting the statement of the theorem to mean that no such numbers $\tau$ with the indicated property exist. □

5.2. **Upper and Lower Bounds for the Arbitrary Irrational.** Theorems 5.5 and 5.8 give upper and lower bounds on the rate of convergence for the $\pm\alpha$ walk for arbitrary irrationals. These bounds resemble (with different exponents) those derived in Kuipers and Niederreiter [7] for the case of discrepancy of equally weighted sequence of $\alpha$-multiples, and the proofs owe a sizable debt to the ideas found there (see Kuipers and Niederreiter [7, Thms.3.2,3.3]).

**Theorem 5.5.** *Suppose $2\alpha$ is an irrational of type $\eta < \infty$. Then, for any fixed $\epsilon > 0$, the discrepancy of the $\pm\alpha$ random walk satisfies*

$$D(Q_\alpha^{*k}, U) = O(k^{-\frac{1}{2\eta}+\epsilon}).$$

*Proof.* Erdős-Turán's inequality (Theorem 2.4) shows that it will be necessary to understand the behavior of $\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h}$. We'll use the following lemma:

**Lemma 5.6.** *Suppose $2\alpha$ is of type $< \psi$. Then for any positive integer $m$,*

$$\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h} \leq \sqrt{\frac{2}{\pi}} \frac{\psi(2m)(\log m + 1)}{\sqrt{k}} \tag{5.3}$$

*Proof of Lemma 5.6.* Abel's summation formula (see Marsden [11, p.135]) gives

$$\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h} = \sum_{h=1}^{m} \frac{s_h}{h(h+1)} + \frac{s_m}{m+1} \tag{5.4}$$

where $s_h = \sum_{j=1}^{h} |\cos^k(2\pi j\alpha)|$. Using the fact that $\cos x \leq e^{-x^2/2}$ for $|x| \leq \frac{\pi}{2}$,

$$s_h \leq \sum_{j=1}^{h} \cos^k(\pi\langle 2j\alpha\rangle) \leq \sum_{j=1}^{h} e^{-k\frac{\pi^2}{2}\langle 2j\alpha\rangle^2}. \tag{5.5}$$

For $2\alpha$ of type $< \psi$ and for $0 \leq p < q \leq h$ we have that

$$|\langle q2\alpha\rangle - \langle p2\alpha\rangle| \geq \langle (q \pm p)2\alpha\rangle \geq \frac{1}{(q \pm p)\psi(q \pm p)} \geq \frac{1}{2h\psi(2h)} =: \frac{1}{J}.$$

This fact implies that each of the intervals $[0, \frac{1}{J}), [\frac{1}{J}, \frac{2}{J}), \ldots [\frac{h}{J}, \frac{h+1}{J})$ contains at most one $\langle j2\alpha\rangle$, for $1 \leq j \leq h$, with no such number lying in the first interval (since $j = 0$ does).

Then from (5.5),

$$s_h \leq \sum_{j=1}^{h} e^{-k\frac{\pi^2}{2}\left(\frac{j}{J}\right)^2} \leq \int_0^{\infty} e^{-\frac{1}{2}\left(\frac{\pi\sqrt{k}x}{J}\right)^2} dx \leq \frac{1}{2}\sqrt{2\pi}\frac{J}{\pi\sqrt{k}} \leq \sqrt{\frac{2}{\pi}}\frac{h\psi(2h)}{\sqrt{k}}.$$

Together with (5.4) this gives

$$\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h} \leq \sqrt{\frac{2}{\pi}} \sum_{h=1}^{m} \frac{\psi(2h)}{(h+1)\sqrt{k}} + \sqrt{\frac{2}{\pi}}\frac{\psi(2m)}{\sqrt{k}}.$$

Bounding $\psi(2h) \leq \psi(2m)$ (since $\psi$ is non-decreasing) and $\sum_{h=1}^{m} \frac{1}{h+1} < \log m$, we obtain (5.3), which concludes the proof of Lemma 5.6. □

We return now to the proof of Theorem 5.5.

Since $2\alpha$ is of type $\eta$, we may in light of Lemma 5.4 set

$$\psi(q) = cq^{\eta - 1 + \epsilon/2}$$

for any $\epsilon > 0$, where $c$ is some constant. Then Lemma 5.6 yields

$$\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h} = O\left(\frac{(2m)^{\eta-1+\epsilon/2}\log m}{\sqrt{k}}\right) = O\left(\frac{m^{\eta-1+\epsilon}}{\sqrt{k}}\right)$$

where the last equality arises from $\log m = o(m^{\epsilon/2})$. Erdős-Turán's inequality (Theorem 2.4) shows that

$$D(Q_\alpha^{*k}, U) = O\left(\frac{1}{m} + \frac{m^{\eta-1+\epsilon}}{\sqrt{k}}\right)$$

for any $m$. Setting $m = k^{\frac{1}{2\eta}}$ yields the conclusion of the theorem. □

**Corollary 5.7.** *Suppose $2\alpha$ is an irrational of constant type. Then the discrepancy of the $\pm\alpha$ random walk satisfies*

$$D(Q_\alpha^{*k}, U) = O(\frac{\log k}{\sqrt{k}}).$$

Irrationals of constant type are precisely those whose continued fraction expansion has bounded elements. (See, for instance, Lang [8, p.24].)

*Proof.* We have $\psi(q) = c$ for some constant $c$.

By the same argument as in Theorem 5.5, we resort to Lemma 5.6, which yields

$$\sum_{h=1}^{m} \frac{|\cos^k(2\pi h\alpha)|}{h} = O\left(\frac{\log m}{\sqrt{k}}\right)$$

Theorem 2.4 shows that

$$D(Q_\alpha^{*k}, U) = O\left(\frac{1}{m} + \frac{\log m}{\sqrt{k}}\right)$$

for any $m$. Setting $m = \sqrt{k}$ yields the conclusion of the theorem. □

We may also show the following lower bound in terms of the type of $\alpha$. Again, the proof closely follows the ideas in Kuipers and Niederreiter [7, Thm.3.3] for the case of discrepancy of an equally weighted sequence of $\alpha$-multiples.

**Theorem 5.8.** *Let $2\alpha$ be an irrational of type $\eta < \infty$. Then, for any fixed $\epsilon > 0$, the discrepancy of the $\pm\alpha$ random walk satisfies*

$$D(Q_\alpha^{*k}, U) = \Omega(k^{-\frac{1}{2\eta}-\epsilon}).$$

Here, $f(x) = \Omega(g(x))$ means $f(x) \neq o(g(x))$.

*Remark* 5.9. This theorem together with Theorem 5.5 answers the question: what is the rate of convergence for the $\pm\alpha$ random walk if $\alpha$ is picked uniformly at random? Recalling that type 1 irrationals are full measure in the unit interval, these two theorems show, in particular, that for almost all $\alpha$, the discrepancy roughly falls as the inverse square root of the number of steps. For type 1 irrationals we have seen that the lower bound can be improved to $k^{-\frac{1}{2}}$ (Theorem 2.2), and for quadratic irrationals we have shown, up to a constant, a matching upper bound as well (Theorem 4.9).

*Proof of Theorem 5.8.* Given $\epsilon > 0$, choose $0 < \delta < \eta$ such that $\frac{1}{2(\eta-\delta)} = \frac{1}{2\eta} + \epsilon$.

We set $\xi = 2\alpha$. Using a close fractional approximation of $\xi$, we shall see that, even after a moderately large number of steps, the walk cannot be very far from the finite subset generated by that fraction. This will give a lower bound on discrepancy by estimating the discrepancy on an interval between fractional multiples.

Since $\xi$ is of type $\eta$, we have that

$$\liminf_{q\to\infty} q^{\eta-\frac{\delta}{2}} \langle q\xi \rangle = 0.$$

In particular there are infinitely many $q$ such that $\langle q\xi \rangle < q^{-\eta+\frac{\delta}{2}}$. Now each $q$ corresponds to some $p$ such that

$$|\xi - \frac{p}{q}| < q^{-1-\eta+\frac{\delta}{2}}. \tag{5.6}$$

Choose one such $q$ and set $k = \lfloor q^{2(\eta-\delta)} \rfloor$. The above inequality (5.6) implies $\xi = \frac{p}{q} + \theta q^{-1-\eta+\frac{\delta}{2}}$ for some $|\theta| < 1$. Multiples of $\xi$ are of the form $n\xi = n\frac{p}{q} + n\theta q^{-1-\eta+\frac{\delta}{2}}$.

For $n < \sigma\sqrt{k}$, we have

$$|n\theta q^{-1-\eta+\frac{\delta}{2}}| < |\sigma\sqrt{k} q^{-1-\eta+\frac{\delta}{2}}| \leq \sigma q^{-1-\frac{\delta}{2}}.$$

The above inequalities show that none of the numbers $\{n\xi \mod 1\}$, $1 \leq n \leq \sigma\sqrt{k}$, lie in the interval $J_a = (\frac{a}{q} + \sigma q^{-1-\frac{\delta}{2}}, \frac{a+1}{q} - \sigma q^{-1-\frac{\delta}{2}})$, for any $a$. Choose $J = J_a$ such that $Q_\alpha^{*k}(J)$ is the smallest over all $a$.

Choose $\gamma = \frac{\delta}{8(\eta-\delta)}$, and let $\sigma = k^\gamma$. By definition of discrepancy,

$$D(Q_\alpha^{*k}, U) \geq |U(J) - Q_\alpha^{*k}(J)|.$$

For $q$ large enough, $U(J) \geq \frac{1}{2q}$, since

$$|\sigma q^{-\frac{\delta}{2}}| = |k^\gamma q^{-\frac{\delta}{2}}| \leq |q^{2\gamma(\eta-\delta)-\frac{\delta}{2}}| = |q^{-\frac{\delta}{4}}|$$

which tends to 0, for large $q$.

Also, for $q$ large enough, and therefore for large $k$, $Q_\alpha^{*k}$ remains supported near the multiples of $1/q$ and hence the support on $J$ tends to 0. More precisely, let $X$ be a random variable denoting the net number of steps taken by time $k$ in the walk, counted with sign, i.e., such that plus and minus steps cancel. Then $\mathrm{Var}(X) = k$. Chebyshev's inequality gives

$$\mathrm{Prob}\{\ |X| \geq \sigma\sqrt{k} = k^{\frac{1}{2}+\gamma}\ \} \leq \frac{k}{k^{1+2\gamma}} \leq \frac{1}{k^{2\gamma}}$$

which tends to 0 for $k$ large. This probability represents the total probability that $Q_\alpha^{*k}$ is in any one of the intervals $J_a$. Hence the definition of $J$ shows that $Q_\alpha^{*k}(J) \leq 1/(q\,k^{2\gamma})$. In particular for $k$ large enough, $Q_\alpha^{*k}(J) \leq \frac{1}{4q}$.

Combining these facts, we see that

$$U(J) - Q_\alpha^{*k}(J) \geq \frac{1}{2q} - \frac{1}{4q} = \frac{1}{4q} \geq c\,k^{-\frac{1}{2(\eta-\delta)}} = c\,k^{-\frac{1}{2\eta}-\epsilon}$$

where the last inequality follows from $q^{2(\eta-\delta)} \leq 2k$. Thus $c$ is a constant depending only on $\eta$ and $\epsilon$. This inequality is true for infinitely many $q$ (large enough), as remarked earlier, which proves the theorem. □

*Remark* 5.10. For irrationals of type $\infty$, the proof of Theorem 5.8 may be modified (by setting $\eta = \frac{1}{2\epsilon}$) to show that given any $\epsilon > 0$, the discrepancy of the $\pm\alpha$ random walk satisfies

$$D(Q_\alpha^{*k}, U) = \Omega(k^{-\epsilon}).$$

In fact, it is possible to construct Liouville numbers for which the $\pm\alpha$ random walk converges as slowly as desired, i.e., the discrepancy is $\Omega(g(x))$ for any fixed decreasing $g(x)$ such that $g(x) \to 0$.

## 6. Motivation and Directions for Further Work

Many practical problems suggest the importance of studying discrete random walks on compact groups. One such problem is the matter of placing points on a sphere uniformly. See Sloane [20] for a discussion of applications of this question to tomography. Such a placement might be achieved in practice by allowing a discrete random walk on the sphere to generate the points. In a related vein, Lubotsky, Phillips, and Sarnak [10] obtain bounds for the speed of equidistribution for a

method for placing points using the orbit of a point under the action of a certain finite set of rotations.

Another example comes from analyzing statistical data. Attributes of the data may be thought of as points in a space of high dimension. Patterns which emerge in projections onto one or two dimensional subspaces may indicate meaningful correlations; hence, one requires an algorithm for generating various projections to view. The "Grand Tour" is one such method (see Diaconis [2] and Asimov [1]). Statisticians view a projection and then modify it by one of a fixed set of rotations—this corresponds to a random walk on $O(n)$, the rotation group of $\mathbf{R}^n$. This is a discrete walk on an infinite compact group. For analysis of the covering time of this walk, i.e., the time it takes for the walk to come within a fixed distance of any view, see Matthews [13].

Variants of the $\pm\alpha$ walk warrant further study. One might wonder whether the biased $\pm\alpha$ walk, in which the probabilities for hopping left and right are not equal, converges any faster. It does not— note that the proof of Theorem 5.8 carries over without change. Moreover, the size of the largest atom in the measure still falls as the inverse square root of the number of steps, yielding a similar lower bound as in Proposition 2.2.

Another natural question is to ask whether the convergence of the $\pm\alpha$ walk could be speeded up by using more generators. Preliminary considerations suggest that with careful choices of generators, this may be achieved. Work in this direction involves the study of simultaneous approximation of irrationals, and will be discussed in a future paper.

E. Rains [17] has noticed a connection between a variant of the $\pm\alpha$ walk and the pinwheel tiling of the plane, an aperiodic tiling studied by Radin [16]. The pinwheel tiling of the plane has the interesting property that its tiles assume infinitely many orientations. One may ask: what is the rate at which new orientations enter as the tiling grows? This question is equivalent to asking for the rate of convergence of a certain random walk on $S^1$ with four generators, which fall into two rationally-related generator classes that differ by the transcendental angle $2\arccos(2/\sqrt{5})$. With some tweaking, this walk can be related to the biased $\pm\alpha$ walk for $\alpha = \arccos(2/\sqrt{5})$. Continued fraction properties of this irrational that would allow analysis of this walk are as yet unknown to the author.

## 7. Acknowledgements

This work is a version of one chapter of the author's Ph.D. thesis at Harvard University. The author wishes to thank his advisor, Persi Diaconis, for his encouragement and many helpful discussions, Brad Mann, who was a sounding board for many ideas contained herein, and Harald Niederreiter, who read a draft of this paper and made helpful suggestions.

## References

[1] D. Asimov, *The grand tour*, SIAM Jour. Sci. Statist. Comp. **6**(1983), 128-143. MR **86h:**62087

[2] P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes, Vol. 11, Hayward, CA, 1988. MR **90a:**60001

[3] P. Diaconis and L. Saloff-Coste, *Comparison Techniques for Random Walks on Finite Groups*, Ann. Probab. **21**(1993), 2131-2156. MR **95a:**60009

[4] P. Erdős and P. Turán, *On a problem in the theory of uniform distribution I*, Indag. Math. **10**(1948), 370-378. MR **10:**372c

[5] B.M. Kloss, *Limiting distributions on bicompact topological groups*, Th. Probab. Appl. **4**(1959), 237-270.
[6] A.Ya. Khinchin, *Continued Fractions*, Univ. of Chicago Press, 1964. MR **28:**5037
[7] L. Kuipers and H. Neiderreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974. MR **54:**7415
[8] S. Lang, *Introduction to Diophantine Approximations*, Springer-Verlag, 1995. MR **96h:**11067
[9] W.J. LeVeque, *An inequality connected with Weyl's criterion for uniform distribution*, Proc. Symp. Pure Math. **8**(1965), 22-30. MR **31:**3401
[10] A. Lubotsky, R. Phillips, and P. Sarnak, *Hecke Operators and Distributing Points on the Sphere, I*, Comm. Pure. Appl. Math. **39**(1986), S149-S186. MR **88m:**11025a
[11] J. Marsden, *Elementary Classical Analysis*, W.H. Freeman and Co., 1974. MR **50:**10161
[12] *Mathematica*, version 2. Wolfram Research, Inc., 1991.
[13] P. Matthews, *Covering problems for random walks on spheres and finite groups*, Ph.D. Thesis, Dept. of Statistics, Stanford Univ., 1985.
[14] 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. MR **49:**2642
[15] U. Porod, *The cut-off phenomenon for random reflections*, Ann. Probab. **24**(1996), 74-96. MR **97e:**60012
[16] C. Radin, *The pinwheel tilings of the plane*, Annals of Math. **139**(1994), 661-702. MR **95d:**52021
[17] E. Rains, personal communication, 1995.
[18] J.S. Rosenthal, *Random rotations: characters and random walks on $SO(n)$*, Ann. Probab. **22**(1994), 398-423. MR **95c:**60008
[19] W.M. Schmidt, *Diophantine Approximation*. Lecture Notes in Math. No. 785, Springer-Verlag, 1980. MR **81j:**10038
[20] N.J.A. Sloane, *Encrypting by Random Rotations*, in *Cryptography*, Lecture Notes in Computer Science, no. 149. T.Beth, editor. Berlin, 1983. MR **85i:**94017
[21] F.E. Su, *Methods for Quantifying Rates of Convergence for Random Walks on Groups*, Ph.D. Thesis, Harvard University, 1995.

Department of Mathematics, Harvey Mudd College, Claremont, California 91711
*E-mail address*: su@math.hmc.edu