MEASURING BIAS IN CYCLIC RANDOM WALKS

CLIFFORD BERGMAN AND SUNDER SETHURAMAN

Abstract. We define the notion of the bias of a Bernoulli random variable and demonstrate its relationship to the property that the mod-2 sum of independent variables converges to a fair coin-toss. We then explore generalizations of these ideas to random walks on a finite cyclic group.

Imagine that you and some friends are playing a version of roulette. The wheel is divided into 36 sectors, alternately colored red and black. Before spinning the wheel, the contestant chooses a color and then wins or loses depending on whether or not his color comes up.

You, the master player, have honed an ability to spin the wheel exactly 360° with high probability. Thus, if the wheel is initially on a red sector, then after your spin, it will again be on a red sector, and similarly for black. Of course, nobody's perfect, so let us say that 90% of your spins return the wheel to the same color on which they begin.

After you've cleaned out your friends a couple of times, they begin to wise up. One of them proposes a small change in the rules. Instead of a single spin, the contestant must spin the wheel 10 consecutive times (say, without looking at the colors obtained along the way). It is only if his guess matches the final outcome that he wins the game.

Is this fellow on to something? Will the new rule blunt your advantage? Let us assume that you continue to bet on the wheel's starting color, and think of each spin as a coin toss in which the probability of 'heads' is 0.9 (i.e., the wheel returns to its starting color after one spin). Then you will win the game if the number of tails after 10 tosses is an even number. The probability of this is easily computed to be $\sum_{k=0}^{5} \binom{10}{2k} (.1)^{2k} (.9)^{10-2k} \approx 0.55$. It seems clear from Figure 1 that as the required number of spins increases, your advantage diminishes. When used with a large number of spins, the game resembles a fair coin-toss, no matter how biased is a single spin.

[1]Research supported in part by the Barbara J. Janson Professorship.
[2]Research supported in part by NSF–DMS–1159026.

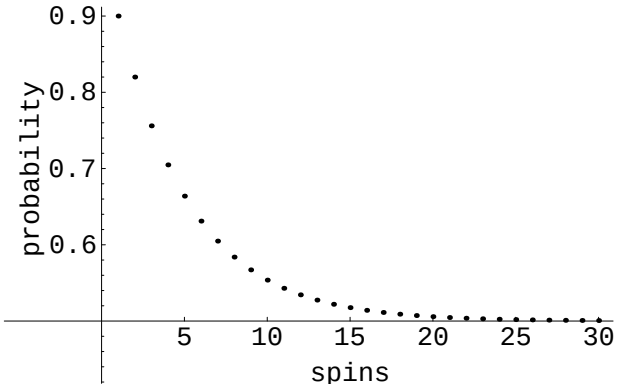

FIGURE 1. Probability of winning the roulette game as a function of the number of spins.

This 'equalizing' phenomenon has been understood at least since the 1950's in the context of cyclic random walks, Feller [7, section 16.2(d)]; random number generation, Horton and Smith [13] and Dvoretsky and Wolfowitz [6]; and card-shuffling, Aldous and Diaconis [1], among other areas. In recent times, this behavior has also been exploited in cryptanalysis, Matsui [18]. We had fun revisiting this old topic and the purpose of this note is to survey some of the interesting results and to give, in some cases, simpler and more probabilistic proofs than are found in the literature. In particular, we focus on ways to measure the 'bias' in the 'equalizing' phenomenon.

In Section 1, we consider the behavior of the 'bias' of sums of Bernoulli variables modulo 2. In Section 2, the discussion is generalized to sums modulo m, and in Section 3 several different 'biases' are compared and contrasted.

1. The Basic Problem

Let us formalize the above discussion as follows. Assume we have a coin for which 0 (Heads) occurs with probability p and 1 (Tails) occurs with chance $q = 1 - p$. Let $X_1, X_2, \ldots$ be a sequence of independent coin-tosses. Let the operation $\oplus$ represent the 'exclusive-or' operation. In other words, $X_1 \oplus X_2$ is nothing more than $X_1 + X_2 \bmod_2$,

$$X_1 \oplus X_2 = \begin{cases} 0, & \text{when } X_1 = X_2 = 0 \text{ or } X_1 = X_2 = 1; \\ 1, & \text{when } X_1 \neq X_2. \end{cases}$$

Our fundamental observation is that, when $0 < p < 1$, no matter how biased the individual tosses, their exclusive-or resembles a fair coin asymptotically.

Proposition 1. *Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed coin tosses with $P(X_i = 0) = p$, for every $i > 0$. Then $\lim_{n\to\infty} P(X_1 \oplus X_2 \oplus \cdots \oplus X_n = 0) = \frac{1}{2}$ if and only if $0 < p < 1$.*

If we use μ to denote the probability distribution of a fair coin, then the proposition says that, when $0 < p < 1$, as n tends to infinity, the sequence $X_1 \oplus X_2 \oplus \cdots \oplus X_n$, *converges in measure* to μ, in symbols, $X_1 \oplus X_2 \oplus \cdots \oplus X_n \Rightarrow \mu$.

We shall present several proofs of the sufficiency part of the proposition: a combinatorial proof, one using Markov chains, and later an 'eigenvalue' argument which provides some additional insight into the phenomenon.

Proof of Proposition 1. To argue necessity, observe that when $p = 0$, the variable $X_1 \oplus X_2 \oplus \cdots \oplus X_n$ alternates between 0 and 1, and hence cannot converge, while when $p = 1$, it has constant value 0.

We now prove sufficiency of $0 < p < 1$ for the limit to hold. For a combinatorial proof, note that the variable $X_1 \oplus X_2 \oplus \cdots \oplus X_n$ is equal to 0 precisely when an even number of the individual tosses return Tails. From the binomial theorem,

$$\begin{aligned}(p+q)^n + (p-q)^n &= \sum_{k=0}^{n} \binom{n}{k} q^k p^{n-k} + \sum_{k=0}^{n} (-1)^k \binom{n}{k} q^k p^{n-k} \\ &= 2 \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} q^{2k} p^{n-2k} \\ &= 2P(X_1 \oplus \cdots \oplus X_n = 0).\end{aligned}$$

Of course, $p + q = 1$ and $0 < |p - q| < 1$. Thus, for large n, the left-hand side of the last equation is close to 1. Hence,

$$P(X_1 \oplus X_2 \oplus \cdots \oplus X_n = 0) \;=\; \frac{1 + (p-q)^n}{2} \;\approx\; 1/2.$$

Moreover, when $1/2 < p < 1$, the probability decreases asymptotically to $1/2$ as in Figure 1. The reader is invited to explore the asymptotics for other values of p.

For a 'Markov chain' proof, we begin with the sequence $Z_n = X_1 \oplus \cdots \oplus X_n$, for $n = 1, 2, 3, \ldots$. One computes that $P(Z_1 = 0) = p = 1 - P(Z_1 = 1)$, and $P(Z_{n+1} = 0 \mid Z_n = 0) = p$, $P(Z_{n+1} = 0 \mid Z_n = 1) = q$, $P(Z_{n+1} = 1 \mid Z_n = 0) = q$, and $P(Z_{n+1} = 1 \mid Z_n = 1) = p$. Thus,

$$\begin{aligned}P(Z_n = 0) &= pP(Z_{n-1} = 0) + qP(Z_{n-1} = 1) \\ &= q + (p-q)P(Z_{n-1} = 0).\end{aligned}$$

The last equality follows from the fact that $P(Z_{n-1} = 0) + P(Z_{n-1} = 1) = 1$. We now invite the reader to confirm that the answer from the first suffiency proof indeed satisfies this recurrence system. □

We remark that the sequence $\{Z_n\}_{n\geq 1}$, used in the second proof, is a *Markov chain* in that the distribution of Z_n depends only on the previous state Z_{n-1}. In fact, $Z_n = Z_{n-1} \oplus X_n$ is a function of Z_{n-1} and an independent coin-toss X_n. The recurrence equation above can be expressed in terms of the 'transition matrix' R whose entries $R(i,j) = P(Z_n = j \mid Z_{n-1} = i)$ for $0 \leq i, j \leq 1$ and $n \geq 2$. That is,

$$\langle P(Z_n = 0),\ P(Z_n = 1)\rangle = \langle P(Z_{n-1} = 0),\ P(Z_{n-1} = 1)\rangle R$$

where

$$R = \begin{pmatrix} p & q \\ q & p \end{pmatrix}.$$

In Proposition 1, we assumed that the coin tosses were both independent and identically distributed. In fact, a far weaker condition than identical distribution is sufficient for the proposition to hold. To make this precise, we introduce a quantity which measures the bias of a Bernoulli variable.

Definition 2. *Let X be a random variable taking on two values with probabilities p and $q = 1-p$. Then the* bias of X, *denoted $B(X)$, is the quantity $|p - q|$.*

The bias $B(X)$ has several nice properties. First, it is symmetric in p and q. Second, we always have $0 \leq B(X) \leq 1$, attaining the lower bound precisely when X is a fair coin and the upper bound when X is a constant (i.e., $p = 1$ or $q = 1$). Third, $B(X)$ is the magnitude of the smallest eigenvalue of the matrix R. Indeed, 1 and $p - q$ are eigenvalues with respective eigenvectors $(1, 1)$ and $(1, -1)$. The utility of the bias is seen from the following identity.

Proposition 3. *Let $X_1, \ldots, X_n$ be independent coin-flips, and let $Z_n = X_1 \oplus \cdots \oplus X_n$. Then*

$$B(Z_n) = \prod_{i=1}^{n} B(X_i).$$

Before proving this result, we show how it immediately implies our earlier claim (Proposition 1). Applying the above identity, we obtain

$$\big|P(Z_n = 0) - P(Z_n = 1)\big| = |p - q|^n$$

which vanishes (exponentially) as n tends to infinity if and only if $0 < p < 1$.

Proof of Proposition 3. It is enough to show $B(X_1 \oplus X_2) = B(X_1)B(X_2)$. This can certainly be checked by direct computation, but for further development we use the following argument. Since the particular values taken on

by X play no role in $B(X)$, let us identify Heads (0) with 1 and Tails (1) with -1. With this identification, the bias $B(X)$ is nothing but $|E[X]|$, the magnitude of the expected value of X. Furthermore, $X_1 \oplus X_2$ is simply the product X_1X_2. Since expected value respects the product of independent observations, $B(X_1X_2) = |E[X_1X_2]| = |E[X_1]| \cdot |E[X_2]| = B(X_1)B(X_2)$. □

In fact, this proof allows for not necessarily identically distributed coin-flips in Proposition 1.

Corollary 4. *Let $X_1, X_2, \ldots$ be a sequence of (non-identically distributed) independent coin-flips. Let $Z_n = X_1 \oplus \cdots \oplus X_n$. Then*

$$Z_n \Rightarrow \mu \iff \lim_{n\to\infty} \prod_{i=1}^{n} B(X_i) = 0.$$

It is interesting to note that the condition at right allows for the coin-flips to become increasingly biased at some rate. For example, if $B(X_i) = 1 - \frac{1}{i+1}$, then $\prod_{i=1}^{n} B(X_i) = (n+1)^{-1}$ which converges to 0 as $n \uparrow \infty$. Hence, in this case, $Z_n \Rightarrow \mu$ even though $\lim_{n\uparrow\infty} B(X_n) = 1$. On the other hand, if the bias of X_n converges too fast to 1, then the limit of Z_n will remain biased. For instance, if $B(X_n) = 1 - \frac{1}{(i+1)^2}$, then $\prod_{i=1}^{n} B(X_i) = \frac{n+2}{2n+2} \to 1/2$ in the limit. We also point out, what may be surprising at first glance, that if even a single X_i is unbiased, then so is Z_n for $n \geq i$.

2. Iterations Modulo m

As the exclusive-or operation can be identified with mod_2 addition, it is natural to ask about mod_m generalizations. Let $X_1, X_2, \ldots$ be a sequence of independent rolls of m-sided dice where the distribution of X_i is given by

$$X_i = l \qquad \text{with chance } p_i(l), \qquad \text{for } l = 0, 1, \ldots, m-1.$$

Does $Z_n = X_1 + \cdots + X_n \bmod_m$ converge to μ where μ is the uniform distribution on $\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$, i.e., $\mu(l) = 1/m$ for $0 \leq l \leq m-1$? In fact, this will be the case when 'degeneracies' are avoided.

Although the combinatorics when $m > 2$ becomes difficult, we can still interpret the asymptotics of Z_n through Markov chain analysis when the observations $\{X_i\}_{i\geq 1}$ are identically distributed with common distribution $\mathbf{p} = \langle p(l) : 0 \leq l \leq m-1 \rangle$. As before, $\{Z_n\}_{n\geq 1}$ forms a Markov sequence since $Z_n = Z_{n-1} + X_n \bmod_m$. A moment's thought reveals that the

transition matrix R, with entries $R(i,j) = P(Z_n = j | Z_n = i)$, takes form

$$R = \begin{pmatrix} p(0) & p(1) & \cdots & p(m-1) \\ p(m-1) & p(0) & \cdots & p(m-2) \\ & & \vdots & \\ p(1) & p(2) & \cdots & p(0) \end{pmatrix}. \tag{2.1}$$

It will be useful later to note that R is a $m \times m$ circulant matrix corresponding to the vector $\langle p(0), p(1), \ldots, p(m-1) \rangle$.

The Markov chain analysis when the distribution of Z_n converges to μ, however, is not as direct as for the $m = 2$ case, but relies on the ergodic theorem for finite-state space chains (cf. Levine, Peres, and Wilmer [17, section 4.4.3] for instance). We invite the reader to pursue this line of attack, but in the following we discuss methods which are simpler and more general.

Let us now consider the general case when the X_i's are not necessarily identically distributed. For this analysis, it is convenient to work with a multiplicative cyclic group, rather than the additive group of integers modulo m. Let $\omega_m = e^{2\pi i/m}$ denote a primitive mth root of unity. The group $\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ under addition mod m is isomorphic to $\{1, \omega_m, \omega_m^2, \ldots, \omega_m^{m-1}\}$ under ordinary multiplication of complex numbers. This isomorphism maps an element $j \in \mathbb{Z}_m$ to ω_m^j. As in Proposition 3, $Z_n = X_1 + X_2 + \cdots + x_n \bmod_m$ is mapped to $X_1 \cdot X_2 \cdot \cdots \cdot X_n$. The sequence $\{Z_n\}_{n \geq 1}$ can be thought of now as a random rotation on the unit circle, or a random walk on the cyclic group of order m.

For a random variable X on the mth roots of unity with distribution $\mathbf{p}$, we now express the eigenvalues of the associated $m \times m$ circulant matrix P, equation (2.1) in terms of the moments of X (in analogy to the case $m = 2$).

Proposition 5. *Let X be a random variable on the mth roots of unity with distribution $\mathbf{p}$, and associated circulant matrix R.*

Then for $k = 0, 1, \ldots, m-1$, the eigenvalues λ_k and eigenvectors v_k of R are given by

$$\lambda_k = E[X^k] \; [= p(0) + p(1)\omega_m^k + p(2)\omega_m^{2k} + \cdots + p(m-1)\omega_m^{(m-1)k}],$$
$$v_k = \left(1, \omega_m^k, (\omega_m^k)^2, \ldots, (\omega_m^k)^{m-1}\right).$$

Proof. It is easy to check by direct calculation that if λ_k and v_k are defined as in the proposition then $Pv_k = \lambda_k v_k$, for $k = 0, 1, \ldots, m-1$. Now it is possible for the sequence $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$ to contain repeats. However, the

matrix whose rows are the vectors v_k is

$$V = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_m & \omega_m^2 & \cdots & \omega_m^{m-1} \\ 1 & \omega_m^2 & (\omega_m^2)^2 & \cdots & (\omega_m^2)^{m-1} \\ & & \vdots & & \\ 1 & \omega_m^{m-1} & (\omega_m^{m-1})^2 & \cdots & (\omega_m^{m-1})^{m-1} \end{pmatrix}$$

which we recognize as a Vandermonde matrix. The determinant of V is $\prod_{j<k}(\omega_m^k - \omega_m^j) \neq 0$. Consequently, $\{v_0, v_1, \ldots, v_{m-1}\}$ is a linearly independent set, and provides us with a basis of eigenvectors for P. □

Notice that the matrix R depends only on the outcome probabilities for the variable X and not on the values taken by X (although the equality of λ_k and $E[X^k]$ does depend on the values of X). In analogy with Definition 2, we write $\lambda_k(X)$, for $k = 1, 2, \ldots, m-1$, for the eigenvalues of P.

Theorem 6. *Let $X_1, X_2, \ldots$ be independent variables on the mth roots of unity, and let $Z_n = X_1 \cdot X_2 \cdot \cdots \cdot X_n$, for $n \geq 1$. Then*

$$Z_n \Rightarrow \mu \iff \lim_{n\to\infty} \prod_{i=1}^{n} |\lambda_k(X_i)| = 0, \qquad \text{for each } 1 \leq k \leq m-1.$$

Proof. Since $E[Z_n^k] = E[X_1^k] \cdot E[X_2^k] \cdot \cdots \cdot E[X_n^k] = \prod_{i=1}^n \lambda_k(X_i)$ by Proposition 5, it is enough to show that $Z_n \Rightarrow \mu$ if and only if the first $m-1$ moments of Z_n tend to zero.

One direction is trivial. The condition $Z_n \Rightarrow \mu$ is equivalent to the assertion $\lim_n P(Z_n = \omega_m^l) = 1/m$ for all $l = 0, 1, \ldots, m-1$. This gives, for $1 \leq k \leq m-1$, that

$$\begin{aligned} \lim_{n\to\infty} E[Z_n^k] &= \frac{1}{m}(1 + \omega_m^k + \cdots + \omega_m^{(m-1)k}) \\ &= \frac{1}{m} \cdot \frac{1 - \omega_m^{mk}}{1 - \omega_m^k} \end{aligned}$$

which vanishes as $\omega_m^m = 1$.

For the other direction, suppose that for every $1 \leq k < m$, we have that $\lim_{n\to\infty} E[Z_n^k] = 0$. Expanding, we have

$$\lim_{n\to\infty} \sum_{\ell=0}^{m-1} (\omega_m^k)^\ell P(Z_n = \omega_m^\ell) = 0, \qquad \text{for } k = 1, \ldots, m-1. \tag{2.2}$$

Consider the sequence

$$y_n = \langle P(Z_n = 1), P(Z_n = \omega_m), \ldots, P(Z_n = \omega_m^{m-1}) \rangle$$

of points in $\mathbb{R}^m$ for $n \geq 1$. The assertion that $Z_n \Rightarrow \mu$ is equivalent to the claim that $y_n \to \langle 1/m, 1/m, \ldots, 1/m \rangle$ as $n \to \infty$. Suppose this is not

the case. That is, y_n has a subsequence $\{y_{n_j} : j \geq 1\}$ that is bounded away from $\langle 1/m, \ldots, 1/m\rangle$. Since the sequence $\{y_{n_j} : j \geq 1\}$ is uniformly bounded, it must itself have a convergent subsequence converging, let us say, to $\langle a_0, a_1, \ldots, a_m\rangle \neq \langle 1/m, \ldots, 1/m\rangle$.

The equations in (2.2) imply that

$$a_0 + \omega_m^k a_1 + (\omega_m^k)^2 a_2 + \cdots + (\omega_m^k)^{m-1} a_{m-1} = 0$$

for $k = 1, \ldots, m-1$. In addition we have the equation

$$a_0 + a_1 + \cdots + a_{m-1} = 1$$

as the limit probabilities $\langle a_\ell : 0 \leq \ell \leq m-1\rangle$ satisfy $\sum_\ell a_\ell = 1$. Thus we have a system of m linear equations whose matrix (if the last equation is moved to the top) is none other than V, which we have already determined to be nonsingular. Consequently, the system has a unique solution. Since

$$\langle a_0, a_1, \ldots, a_{m-1}\rangle = \langle 1/m, 1/m, \ldots, 1/m\rangle$$

is a solution, it is the only one contradicting our assumption. □

The convergence in measure holds in particular for 'non-degenerate' identically distributed independent variables. Motivated by the above result, let us say that the distribution $\mathbf{p}$ of a random variable X taking values on $\mathbb{Z}_m$ is *non-degenerate* when $|\lambda_k(X)| < 1$ for all $1 \leq k \leq m-1$. [Note that by the triangle inequality, $|\lambda_k(X)| \leq p(0)+\cdots+p(m-1) = 1$.] Or, in other words, the distribution is non-degenerate exactly when there is a 'positive spectral gap,' that is, a positive gap between the eigenvalue $\lambda_0(X) = 1$ and $\max_{1\leq k\leq m-1} |\lambda_k(X)|$. We also call a variable non-degenerate if its distribution has that property. We now develop concrete conditions for non-degeneracy of a random variable X.

Let us again think of $\mathbb{Z}_m$ as the additive group of integers modulo m. For $i, j \in \mathbb{Z}_m$, let $\langle j\rangle + i = \{jx + i : x = 0, 1, \ldots, m-1\}$ be the coset of the cyclic subgroup generated by j displaced by i.

Proposition 7. *Let X be a random variable on $\mathbb{Z}_m$ with distribution $\mathbf{p}$. Define $H = \{i \in \mathbb{Z}_m : p(i) > 0\}$. Then for $1 \leq k \leq m-1$, we have that*

$$|\lambda_k(X)| = 1 \iff H \subset \langle j\rangle + i$$

where $i = \min(H)$ and $j = m/\gcd(m, k)$.

Proof. The '⇒' proof follows from two observations.

(1) Let $z, w \in \mathbb{C} - \{0\}$. Then $|z + w| = |z| + |w|$ implies that $z = rw$ for some real number $r > 0$. This is the law of cosines.
(2) Let $p(1), \ldots, p(k)$ be positive real numbers, and $\alpha_1, \ldots, \alpha_k$ be distinct complex numbers of magnitude 1. If $|p(1)\alpha_1 + \cdots + p(k)\alpha_k| = p(1) + \cdots + p(k)$, then $k = 1$. We will show this momentarily.

Given (1) and (2), we prove the proposition. Assume $|\lambda_k| = 1$ for some $1 \leq k \leq m-1$. Then $|\sum_{i\in H} p(i)\omega_m^{ik}| = 1$. By (2) we have $i, l \in H$ implies $\omega_m^{ik} = \omega_m^{lk}$. Let $i = \min(H)$, $l \in H$, $d = \gcd(m,k)$, and $j = m/d$. Then $\omega_m^{ik} = \omega_m^{lk}$ implies m divides $(l-i)k$, which further implies m/d divides $(l-i)k/d$, and so m/d divides $(l-i)$. This is equivalent to $l \in \langle j\rangle + i$.

To finish the proof, we argue (2). Let $k > 1$ be the least counterexample of the statement. Then, by the triangle inequality,

$$\begin{aligned}|p(1)\alpha_1 + (p(2)\alpha_2 + \cdots + p(k)\alpha_k)| &= p(1) + p(2) + \cdots + p(k)\\ &\geq |p(1)\alpha_1| + |p(2)\alpha_2 + \cdots + p(k)\alpha_k|.\end{aligned}$$

Using the triangle inequality again, $|p(1)\alpha_1| + |p(2)\alpha_2 + \cdots + p(k)\alpha_k| \geq |p(1)\alpha_1 + \cdots + p(k)\alpha_k|$. This implies $p(2)+\cdots+p(k) = |p(2)\alpha_2+\cdots+p(k)\alpha_k|$. Therefore, by assumption, $k-1 = 1$, that is $k = 2$. By (1), we must have $p(2)\alpha_2 = rp(1)\alpha_1$ for some $r > 0$. This gives $\alpha_2 = r(p(1)/p(2))\alpha_1$. But, $|\alpha_1| = |\alpha_2| = 1$, and so $rp(1)/p(2) = 1$ and $\alpha_2 = \alpha_1$. This is a contradiction.

For the '$\Leftarrow$' argument, suppose $H \subset \langle j\rangle + i$, where $j = m/\gcd(m,k)$ and $i = \min(H)$. Then m divides kj, and so $\omega_m^{k(jx+i)} = \omega_m^{ki}$ for all $x \in \mathbb{Z}_m$. Hence, $|\lambda_k| = |\omega_m^{ki}\sum_{l\in H} p(l)| = 1$. □

We now state a few consequences of Proposition 7.

Corollary 8. *Let* $\mathbf{p}$ *be a distribution on* $\mathbb{Z}_m$*, and let* H *be its support. Define* H' *as the translate* $H' = H - i$ *where* $i = \min(H)$*. Then* $\mathbf{p}$ *is non-degenerate if* H' *is not contained in a proper subgroup of* $\mathbb{Z}_m$*.*

Proof. From Proposition 7, non-degeneracy of $\mathbf{p}$ is equivalent to $H' \not\subset \langle m/\gcd(m,k)\rangle$ for $k = 1, \ldots, m$. However, we observe that $\langle m/\gcd(m,k)\rangle$ for $1 \leq k \leq m-1$ are precisely the proper subgroups of $\mathbb{Z}_m$ to finish the argument. □

Corollary 9. *Let* $X_1, X_2, \ldots$ *be a sequence of identically distributed, independent non-degenerate random variables taking values in* $\mathbb{Z}_m$ *with distribution* $\mathbf{p}$*. Let* $Z_n = X_1 + X_2 + \cdots + X_n \ mod_m$ *for* $n \geq 1$*. Then*

$$Z_n \Rightarrow \mu \iff \mathbf{p} \text{ is non-degenerate.}$$

Proof. For $0 \leq k \leq m-1$, denote by λ_k the common value of $\lambda_k(X_i)$ for $i \geq 1$. By Theorem 6, $Z_n \Rightarrow \mu$ if and only if $|\lambda_k|^n$ converges to 0 for all $1 \leq k \leq m-1$. This can happen if and only if $|\lambda_k| < 1$ for all $1 \leq k \leq m-1$ which is equivalent to non-degeneracy of $\mathbf{p}$. □

We note if m is prime then the only proper subgroup of $\mathbb{Z}_m$ is $\{0\}$. Hence, in this case, Corollary 8 tells us that the only degenerate random variables are the constant ones.

On the other hand, consider the following examples when $m = 6$. Suppose $p_0 = p_2 = p_4 = 0$ and $p_1 = p_3 = p_5 = 1/3$; then, by Proposition 7, $\mathbf{p}$ is degenerate, and Z_n does not converge to μ. Indeed, Z_n takes on only odd values for n odd, and only even values for n even. However, when $p_0 = p_1 = p_2 = 0$ and $p_3 = p_4 = p_5 = 1/3$, one can verify that $\mathbf{p}$ is non-degenerate, and so $Z_n \Rightarrow \mu$.

3. Some Generalized Bias Functions

Although Theorem 6 gives equivalent conditions for convergence of $Z_n = X_1 + \cdots + X_n \bmod_{\mathrm{m}}$ to the uniform distribution on $\mathbb{Z}_m$ in terms of the eigenvalues $\left\{ \langle \lambda_k(X_i) : 1 \le k \le m \rangle \right\}_{i \ge 1}$, computing these spectral quantities is not immediate. Also, the equivalent conditions, except for the case $m = 2$, do not give a rate of convergence. Therefore, it is natural to ask if there are concrete, useful generalizations of the bias of a Bernoulli random variable to variables when $m > 2$. In particular, we would like to retain as many of the properties outlined after Definition 2 as possible. It turns out there are several reasonable candidates.

Let X be a random variable taking values in $\mathbb{Z}_m$ with distribution $\mathbf{p} = \langle p(0), p(1), \ldots, p(m-1) \rangle$ for $m \ge 2$. Order the probabilities in terms of their size as

$$\hat{p}_0 \le \hat{p}_1 \le \cdots \le \hat{p}_{m-1}. \tag{3.1}$$

We write $\lfloor x \rfloor$ and $\lceil x \rceil$ for the floor and ceiling of the real number x. Let $U(\mathbf{p})$ be the sum of the $\lfloor m/2 \rfloor$ largest probabilities among $\mathbf{p}$, and let $L(\mathbf{p})$ be the sum of the $\lfloor m/2 \rfloor$ smallest numbers among them. In terms of $\langle \hat{p}_i : 0 \le i \le m-1 \rangle$, $U(\mathbf{p}) = \sum_{i=\lceil m/2 \rceil + 1}^{m-1} \hat{p}_i$, and $L(\mathbf{p}) = \sum_{i=0}^{\lfloor m/2 \rfloor} \hat{p}_i$.

Now, we define five types of biases for the distribution of X. Certainly, others can also be envisioned.

$$\begin{aligned}
B_1(\mathbf{p}) &= \frac{m}{m-1} \max_i |p(i) - 1/m|, \\
B_2(\mathbf{p}) &= \frac{m}{m-1} \Big\lceil \frac{1}{2} \sum_i |p(i) - 1/m| \Big\rceil, \\
B_3(\mathbf{p}) &= \hat{p}_{m-1} - \hat{p}_0, \\
B_4(\mathbf{p}) &= \frac{1}{2} \max_{r \in \mathbb{Z}_m} \sum_{l \in Z_m} \big| p(l) - p(l-r) \big|, \\
B_5(\mathbf{p}) &= U(\mathbf{p}) - L(\mathbf{p})
\end{aligned}$$

where $l - r$ in the definition of B_4 is addition modulo m.

It is nice to observe that all of the biases are between 0 and 1. Also, the biases are not degenerate in that, for each $1 \le i \le 4$, $B_i(\mathbf{p}) = 0$ exactly when $\mathbf{p} = \mu$, the uniform distribution on $\mathbb{Z}_m$. In addition, with respect

to coin-flips, when $m = 2$, these biases all reduce to $|p(0) - p(1)|$, the bias introduced in Section 1. We note also, when $m = 3$, $B_3(\mathbf{p}) = B_4(\mathbf{p}) = B_5(\mathbf{p})$, and, when $m = 4$, $B_4(\mathbf{p}) = B_5(\mathbf{p})$. Reasonably then, all of these biases can be thought of as generalizations of the $m = 2$ case.

However, none of them are without flaws. While it is true that for $i \in \{1, 2, 3\}$ we have $B_i(\mathbf{p}) = 1$ precisely when X is constant, this desirable property fails for B_4 and B_5. We also note that B_4 is not invariant under a permutation of the outcomes of X. More important for our purposes, none of these bias functions exhibit an equality that generalizes Proposition 3 beyond the case $m = 2$. We are able to obtain inequalities that lead to sufficient conditions for convergence of Z_n to μ, but not necessary ones. In fact, in view of the $m-1 > 1$ (non-trivial in general) relations in Theorem 6, it would be too much to hope for to find a single bias function under which necessary and sufficient conditions for convergence of Z_n to μ would hold when $m > 2$.

There is a useful trick that can be helpful when manipulating sums such as the one in the definitions of B_2 and B_4. For a real number x, let $(x)_+$ denote $\max(x, 0)$ (the 'positive part of x') and $(x)_- = -\min(x, 0)$. Then $x = (x)_+ - (x)_-$ and $|x| = (x)_+ + (x)_-$. We can compute as follows.

$$\begin{aligned}
\sum_{i=0}^{m-1} |p(i) - 1/m| &= \sum_i \Big[\big(p(i) - 1/m\big)_+ + \big(p(i) - 1/m\big)_- \Big] \\
&= 2 \sum_i \big(p(i) - 1/m\big)_+ - \sum_i \Big[\big(p(i) - 1/m\big)_+ - \big(p(i) - 1/m\big)_- \Big] \\
&= 2 \sum_i \big(p(i) - 1/m\big)_+ - \sum_i (p(i) - 1/m) \\
&= 2 \sum_i \big(p(i) - 1/m\big)_+
\end{aligned}$$

since $\sum_i p(i) = \sum_i 1/m = 1$. From this computation, as well as its dual, we obtain

$$B_2(\mathbf{p}) = \frac{m}{m-1} \sum_i \big(p(i) - 1/m\big)_+ = \frac{m}{m-1} \sum_i \big(p(i) - 1/m\big)_- . \tag{3.2}$$

With this same trick, and also noting $(a - b)_+ = a - \min(a, b)$, $B_4(\mathbf{p})$ can rewritten as

$$\begin{aligned}
B_4(\mathbf{p}) &= \max_{r \in \mathbb{Z}_m} \sum_{l \in \mathbb{Z}_m} \big(p(l) - p(l - r)\big)_+ \\
&= 1 - \min_{r \in \mathbb{Z}_m} \sum_{l \in \mathbb{Z}_m} \min\big(p(l),\, p(l - r)\big).
\end{aligned} \tag{3.3}$$

Also, at first glance, B_5 may seem strange, however, the following proposition gives it another more natural interpretation which appears new.

Proposition 10. *Let* $\mathbf{p}$ *be a distribution on* $\mathbb{Z}_m$*. Then, with respect to the set of permutations* S_m *on* $\mathbb{Z}_m$*, we have*

$$B_5(\mathbf{p}) = \max_{\sigma \in S_m} \frac{1}{2} \sum_{\ell \in \mathbb{Z}_m} |p(\ell) - p(\sigma_\ell)| = \max_{\sigma \in S_m} \sum_{\ell \in \mathbb{Z}_m} \big(p(\ell) - p(\sigma_\ell)\big)_+.$$

Proof. It is enough to show, for any permutation σ, that $T := \sum_{\ell \in \mathbb{Z}_m} \big(p(\ell) - p(\sigma_\ell)\big)_+ \leq B_5(\mathbf{p})$. Some of the terms in the expression for T are 0 and others are positive. If we throw away the 0 terms and rearrange, there are elements $\{k_1, k_2, \ldots, k_t, j_1, \ldots, j_t\} \subseteq \mathbb{Z}_m$ such that

$$T = \big(p(k_1) + p(k_2) + \cdots + p(k_t)\big) - \big(p(j_1) + p(j_2) + \cdots + p(j_t)\big); \quad (3.4)$$

$$p(k_1) \geq p(k_2) \geq \cdots \geq p(k_t) \text{ and } p(j_1) \leq p(j_2) \leq \cdots \leq p(j_t). \quad (3.5)$$

Now the indices $k_1, \ldots, k_t$ come from distinct values of σ, hence they are all distinct. Similarly, the j_i's are all distinct. If any $k_i = j_\ell$, then we could cancel them in equation (3.4). Therefore, we can assume that there are no repeats in the list $k_1, k_2, \ldots, k_t, j_1, \ldots, j_t$. Thus we can conclude that $t \leq m/2$.

It follows from equation (3.5) that for every $1 \leq i \leq t$, the value of $p(k_i)$ is at most the ith largest in $\mathbf{p}$, i.e., $p(k_i) \leq \hat{p}_{m-i+1}$. Similarly, $p(j_i) \geq \hat{p}_i$. From these two inequalities we obtain $\big(p(k_i) - p(j_i)\big) \leq (\hat{p}_{m-i+1} - \hat{p}_i)$. Once again rearranging the terms in the sum for T:

$$\begin{aligned} T &= \big(p(k_1) - p(j_1)\big) + \big(p(k_2) - p(j_2)\big) + \cdots + \big(p(k_t) - p(j_t)\big) \\ &\leq (\hat{p}_m - \hat{p}_1) + (\hat{p}_{m-1} - \hat{p}_2) + \cdots + (\hat{p}_{m-t+1} - \hat{p}_{t-1}) \leq B_5(\mathbf{p}) \end{aligned}$$

as desired. □

A natural question now is how do these biases relate to each other.

Theorem 11. *Let* $\mathbf{p}$ *be a distribution on* $\mathbb{Z}_m$*. Then*

$$B_1(\mathbf{p}) \leq B_j(\mathbf{p}) \leq B_4(\mathbf{p}) \leq B_5(\mathbf{p})$$

where $j \in \{2, 3\}$*. However,* $B_2(\mathbf{p})$ *and* $B_3(\mathbf{p})$ *are not comparable for* $m \geq 4$*, although* $B_2(\mathbf{p}) \leq B_3(\mathbf{p})$ *for* $m \leq 3$*.*

We remark all inequalities above can be strict for $m \geq 5$. For example, when $m \geq 5$ is even, consider $\mathbf{p} = \frac{2}{3m}\langle 1, \ldots, 1, 2, 1, 2, \ldots, 2\rangle$ with equal numbers of 1's and 2's. We leave it to the interested reader to verify that $B_1(\mathbf{p}) < B_3(\mathbf{p}) < B_2(\mathbf{p}) < B_4(\mathbf{p}) < B_5(\mathbf{p})$, and to construct a similar example for odd $m \geq 5$.

Proof of Theorem 11. With respect to $\mathbf{p}$, let $\hat{p}_i$, $i = 0, 1, \ldots, m-1$ be the ordered probabilities as in (3.1). As $\frac{m-1}{m} B_1(\mathbf{p})$ is equal to either $(\hat{p}_{m-1} - 1/m)_+$ or $(\hat{p}_0 - 1/m)_-$, the inequality $B_1(\mathbf{p}) \leq B_2(\mathbf{p})$ follows from the equations in (3.2).

We now prove the inequality $B_1(\mathbf{p}) \leq B_3(\mathbf{p})$. As the minimum and maximum probabilities must satisfy $\hat{p}_0 \leq 1/m \leq \hat{p}_{m-1}$, we can write

$$\begin{aligned} B_1(\mathbf{p}) &= \frac{m}{m-1} \max\left\{ \frac{1}{m} - \hat{p}_0,\ \hat{p}_{m-1} - \frac{1}{m} \right\} \\ &= \max\left\{ \frac{1-\hat{p}_0}{m-1} - \hat{p}_0,\ \hat{p}_{m-1} + \frac{\hat{p}_{m-1} - 1}{m-1} \right\}. \end{aligned}$$

Now, $1 - \hat{p}_0 = \hat{p}_1 + \cdots + \hat{p}_{m-1} \leq (m-1)\hat{p}_{m-1}$, and so $(1-\hat{p}_0)/(m-1) - \hat{p}_0 \leq \hat{p}_{m-1} - \hat{p}_0 = B_3(\mathbf{p})$. Similarly, $\hat{p}_{m-1} + (\hat{p}_{m-1} - 1)/(m-1) \leq B_3(\mathbf{p})$ to finish the argument.

We now prove $B_2(\mathbf{p}) \leq B_4(\mathbf{p})$. Note that $\mu = \langle \frac{1}{m}, \ldots, \frac{1}{m} \rangle R$ where R is the matrix in (2.1). Write

$$\begin{aligned} \frac{m-1}{m} B_2(\mathbf{p}) &= \frac{1}{2} \sum_j \left| p(j) - \frac{1}{m} \right| = \frac{1}{2} \sum_j \left| p(j) - \frac{1}{m} \sum_i R(i,j) \right| \\ &= \frac{1}{2} \sum_j \left| \frac{1}{m} \sum_i p(j) - p(j-i) \right| \leq \frac{1}{2m} \sum_{i,j} |p(j) - p(j-i))|. \end{aligned}$$

By pulling out a maximum over $i \neq 0$, we have the further bound

$$\frac{m-1}{m} \cdot \frac{1}{2} \max_{i \neq 0} \sum_j |p(j) - p(j-i)| = \frac{m-1}{m} B_4(\mathbf{p}).$$

We now argue $B_3(\mathbf{p}) \leq B_4(\mathbf{p})$. Observe that there are indices $j, k \in \mathbb{Z}_m$ such that $\hat{p}_{m-1} = p(j)$ and $\hat{p}_0 = p(k)$. Taking $r = j - k \bmod_m$, we have $B_3(\mathbf{p}) = \hat{p}_{m-1} - \hat{p}_0 = p(j) - p(j-r) \leq \sum_l \big(p(l) - p(l-r)\big)_+ \leq B_4(\mathbf{p})$.

Also, the inequality $B_4(\mathbf{p}) \leq B_5(\mathbf{p})$ is a consequence of Proposition 10.

Last, we give an example when $m = 4$ to show $B_2(\mathbf{p})$ and $B_3(\mathbf{p})$ are not comparable. Let $\mathbf{p}_1 = \langle 1/8, 1/8, 3/8, 3/8 \rangle$ and $\mathbf{p}_2 = \langle 1/8, 1/4, 1/4, 3/8 \rangle$. Then $B_2(\mathbf{p}_1) = 1/3$ and $B_2(\mathbf{p}_2) = 1/6$, but $B_3(\mathbf{p}_1) = B_3(\mathbf{p}_2) = 1/4$. We leave it to the interested reader to generalize the example to $m \geq 4$, and also the verification that always $B_2(\mathbf{p}) \leq B_3(\mathbf{p})$ when $m \leq 3$. □

We now turn to some inequalities which bound the bias of Z_n in terms of a product of biases. The first and second parts of the following proposition are due to Dvoretzky and Wolfowitz [6], and Horton and Smith [13], respectively. We omit its proof as we develop a sharper bound in Proposition 13.

Proposition 12. *Let X and Y be independent random variables on $\mathbb{Z}_m$ with distributions $\mathbf{p}_X$, and $\mathbf{p}_Y$. Also, let $Z = X + Y \bmod_m$ and let $\mathbf{p}_Z$ be its distribution. Then we have*

(i) $B_1(\mathbf{p}_Z) \le B_5(\mathbf{p}_X)B_1(\mathbf{p}_Y)$, *and*
(ii) $B_3(\mathbf{p}_Z) \le B_5(\mathbf{p}_X)B_3(\mathbf{p}_Y)$.

By iterating the inequalities in parts (i) and (ii) above, along with noting, from Theorem 11, that $B_1(\mathbf{p}), B_3(\mathbf{p}) \le B_5(\mathbf{p})$, we have $B_1(\mathbf{p}_{Z_n}) \le \prod_{i=1}^n B_5(\mathbf{p}_i)$ and $B_3(\mathbf{p}_{Z_n}) \le \prod_{i=1}^n B_5(\mathbf{p}_i)$.

However, the next bound is a little better than the inequalities in Proposition 12 in that on both sides of the relation the same type of bias is used.

Proposition 13. *Consider independent random variables X and Y on $\mathbb{Z}_m$ with respective distributions $\mathbf{p}_X$ and $\mathbf{p}_Y$. Let $\mathbf{p}_{X+Y}$ be the distribution of $X + Y \bmod_m$. Then $B_4(\mathbf{p}_{X+Y}) \le B_4(\mathbf{p}_X)B_4(\mathbf{p}_Y)$.*

Proof. Let $r(j) = \mathbf{p}_{X+Y}(j) = \sum_k \mathbf{p}_X(k)\mathbf{p}_Y(j-k)$. Then $B_4(\mathbf{p}_{X+Y}) = \max_s \sum_l (r(l) - r(l-s))_+$. Write

$$\begin{aligned}\sum_l (r(l) - r(l-s))_+ &= \sum_l \left(\sum_k \mathbf{p}_X(k)(\mathbf{p}_Y(l-k) - \mathbf{p}_Y(l-s-k)) \right)_+ \\ &= \sum_l \left(\sum_u \mathbf{p}_X(l-u)(\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s)) \right)_+ \\ &= \sum_{l \in A} \left(\sum_u \mathbf{p}_X(l-u)(\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s)) \right)\end{aligned} \tag{3.6}$$

where A consists of those indices l for which

$$\sum_u \mathbf{p}_X(l-u) \cdot \big(\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s)\big) > 0.$$

The last expression in (3.6), after interchange of summation and rewriting $(\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s))$ in terms of positive and negative parts, yields

$$\sum_u \left\{ \left[\sum_{l \in A} \mathbf{p}_X(l-u) \right] (\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s))_+ - \left[\sum_{l \in A} \mathbf{p}_X(l-u) \right] (\mathbf{p}_Y(u) - \mathbf{p}_Y(u-s))_- \right\}$$

and is less than

$$\max_u \left[\sum_{l\in A} \mathbf{p}_X(l-u)\right]\left[\sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(u-s))_+\right] - \min_u \left[\sum_{l\in A} \mathbf{p}_X(l-u)\right]\left[\sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(l-u))_-\right].$$

As $\sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(l-u))_+ = \sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(l-u))_-$, the last expression equals

$$\begin{aligned}
&\left[\sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(l-u))_+\right]\left[\max_{u,u'} \sum_{l\in A} (\mathbf{p}_X(l-u)-\mathbf{p}_X(l-u'))\right] \\
&\le \left[\sum_u (\mathbf{p}_Y(u)-\mathbf{p}_Y(l-u))_+\right]\left[\max_{u,u'} \sum_{l} (\mathbf{p}_X(l-u)-\mathbf{p}_X(l-u'))_+\right] \\
&\le B_4(\mathbf{p}_X)B_4(\mathbf{p}_Y).
\end{aligned}$$

As s is arbitrary, this finishes the proof. □

We can now write a sufficient condition for convergence of $\mathbf{p}_{Z_n}$ to μ in terms of the bias B_4.

Theorem 14. *Let $\{X_i\}_{i\ge 1}$ be independent random variables on $\mathbb{Z}_m$ with respective distributions $\{\mathbf{p}_i\}_{i\ge 1}$. Let $Z_n = X_1 + \cdots + X_n \bmod_m$ and $\mathbf{p}_{Z_n}$ be its distribution for $n \ge 1$. Then $B_4(\mathbf{p}_{Z_n}) \le \prod_{i=1}^n B_4(\mathbf{p}_i)$ and*

$$Z_n \Rightarrow \mu \text{ if } \lim_{n\to\infty} \prod_{i=1}^n B_4(\mathbf{p}_i) = 0.$$

Proof. The product inequality is a consequence of Proposition 13, and the convergence follows since the bias B_4 vanishes exactly at the uniform distribution μ. □

However, we remark condition '$\prod_{i\ge 1} B_4(\mathbf{p}_i) \to 0$' is only sufficient as there are distributions $\mathbf{p}$ where $B_4(\mathbf{p}) = 1$ but $\mathbf{p}$ is non-degenerate in the sense of the last section. In particular, distributions $\mathbf{p} = \langle p, 1-p, 0, 0\rangle$ for $0 < p < 1$ are such examples when $m = 4$.

Finally, we furnish a simple condition for exponential convergence of Z_n to the uniform distribution. This condition in a different form for identically distributed random variables $\{X_i\}_{i\ge 1}$ on more general groups was given by Kloss [15]. See also in this context Aldous and Diaconis [1].

Corollary 15. *Under the setting of Theorem 14, suppose in addition there is a constant $0 < c < 1$ such that $p_i(j) \ge c/m$ for all $i \ge 1$ and $j \in \mathbb{Z}_m$. Let $\mathbf{p}_{Z_n}$ be the distribution of Z_n Then, for $n \ge 1$, we have*

$$B_4(\mathbf{p}_{Z_n}) \le (1-c)^n.$$

Proof. By Proposition 13, we have that $B_4(\mathbf{p}_{Z_n}) \leq \prod_{i=1}^{n} B_4(\mathbf{p}_i)$. To finish the proof, we now observe from (3.3) that, for each $i \geq 1$,

$$\begin{aligned} B_4(\mathbf{p}_i) &= 1 - \min_{r \in \mathbb{Z}_m} \sum_{l \in \mathbb{Z}_m} \min(p_i(l), p_i(l-r)) \\ &\leq 1 - m(c/m) \;=\; 1 - c. \end{aligned}$$

□

4. Concluding Remarks

Corollary 9 is a case of the well-known Ito-Kawade Theorem which gives equivalent conditions on a distribution ν in a compact group for the n-fold convolutions ν^{*n} to converge to the Haar measure of the group; see Grenander [9], Heyer [11], and Högnäs and Mukherjea [12].

It seems, however, that the convergence of convolutions of non-identical distributions in general groups are not as well understood (cf. [12, Section 2.4]). But see there, in particular, [12, Theorem 2.49] which gives a necessary and sufficient condition for convergence of convolutions of non-identical measures on discrete groups. In this context, Theorem 6 is a particular case of this result with respect to the finite group $\mathbb{Z}_m$ first proved by Dvoretzky and Wolfowitz [6] with Fourier methods. See also Aldous and Diaconis [1], and Goel and Gulati [8] for applications to card shuffling and statistics among other things.

The bias $B_2(\mathbf{p})$ is $m/(m-1)$ times the variational distance between $\mathbf{p}$ and the uniform measure $\mu = \langle 1/m, \ldots, 1/m \rangle$. Also, $B_4(\mathbf{p})$ is the maximum variational distance between $\mathbf{p}$ and its translates, whereas $B_5(\mathbf{p})$ is the maximum variational distance between permutations of $\mathbf{p}$. Moreover, the bias B_4 is the 'contraction coefficient' of the circulant matrix R in (2.1). The contraction coefficient of a general stochastic matrix is a measure of its distance from a stochastic matrix with equal rows, and is useful in the analysis of Markov chains. Proposition 13 is a case of a more general contraction coefficient inequality. See Levin, Peres and Wilmer [17, chapter 4] for more on variational distances in the context of Markov chains, and Isaacson and Madsen [14, chapter 5] and Griffeath [10] for more discussion on the contraction coefficient.

Circulant matrices appear naturally in many applications. A proof of Corollary 9 purely in terms of circulant properties is found in Krafft and Schaefer [16]. The book Davis [3] is a comprehensive reference. See also Diaconis [5] for a discussion of interesting generalizations of these matrices and their properties.

Finally, returning to the opening example, it seems that, no matter the number N of extra spins demanded of the friends, the skilled spinner will still have an advantage, although the bias vanishes exponentially in N.

However, if N is allowed to be 'random', then it seems the friends may have equal odds of winning. Namely, N should be a 'strong stationary time', which can be constructed in this case along the lines of Aldous and Diaconis [2, Proposition 3.2]. See also Levin, Peres and Wilmer [17, Section 6.4] for more discussion and extensions to related questions.

References

[1] D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly, **93** (1986), 333–348.

[2] D. Aldous and P. Diaconis, *Strong uniform times and finite random walks*, Adv. Appl. Math., **8** (1987), 69–97.

[3] P. Davis, *Circulant Matrices*, Wiley, New York, 1979.

[4] P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes-Monograph Series, **11**, Hayward, CA, 1988.

[5] P. Diaconis, *Patterned matrices*, Proceedings of Symposia in Applied Mathematics, Ed. C. Johnson, **40** (1990), 37–58.

[6] A. Dvoretzky and J. Wolfowitz, *Sums of random integers reduced modulo m*, Duke Math. Journal, **18** (1951), 501–507.

[7] W. Feller, *An Introduction to Probability Theory and its Applications I*, Wiley, New York, 1968.

[8] P. Goel and C. Gulati, *Monotone decreasing distance between distributions of sums of unfair coins and a fair coin*, Math. Sci., **26** (2001), 34–40.

[9] U. Grenander, *Probabilities on Algebraic Structures*, Wiley, New York, 1963.

[10] D. Griffeath, *Uniform coupling of non-homogeneous Markov chains*, J. Appl. Probability, **12** (1975), 753–762.

[11] H. Heyer, *Probability Measures on Locally Compact Groups*, Springer, Berlin-Heidelberg-New York, 1975.

[12] G. Högnäs and A. Mukherjea, *Probability Measures on Semigroups*, Plenum Press, New York, 1995.

[13] H. B. Horton and R. T. Smith, *A direct method for producing random digits in any number system*, Ann. Math. Stat., **20** (1949), 82–90.

[14] D. Isaacson and R. Madsen, *Markov Chains: Theory and Applications*, Wiley, New York, 1976.

[15] B. M. Kloss, *Probability distributions on bicompact topological groups*, Theory. Probab. Appl, **4** (1959), 237–270.

[16] O. Krafft and M. Schaefer, *Convergence of the powers of a circulant stochastic matrix*, Linear Alg. and Appl., **127** (1990), 59–69.

[17] D. Levin, Y. Peres, and E. Wilmer, *Markov Chains and Mixing Times*, AMS, Providence, Rhode Island, 2009.

[18] M. Matsui, *Linear cryptanalysis method for DES cipher*, Advances in Cryptology—Eurocrypt '93, Lecture Notes in Computer Science Vol. 765, pp. 386–397, Springer-Verlag, New York, 1993.

MSC2010: 60B10, 60B15

Key words and phrases: random walk, circulant matrix, contraction coefficient, cyclic group

Department of Mathematics, Iowa State University, Ames, IA 50011
E-mail address: cbergman@iastate.edu

Department of Mathematics, The University of Arizona, 617 N. Santa Rita Ave., PO Box 210089, Tucson, AZ 85721-0089
E-mail address: sethuram@math.arizona.edu