The Annals of Probability
1987, Vol. 15, No. 3, 1148–1165

RANDOM WALKS ARISING IN RANDOM NUMBER GENERATION

BY F. R. K. CHUNG, PERSI DIACONIS[1] AND R. L. GRAHAM

Bell Communications Research, Stanford University and AT & T Bell Laboratories

Random number generators often work by recursively computing $X_{n+1} \equiv aX_n + b \pmod{p}$. Various schemes exist for combining these random number generators. In one scheme, a and b are themselves chosen each time from another generator. Assuming that this second source is truly random, we investigate how long it takes for X_n to become random. For example, if $a = 1$ and $b = 0$, 1, or -1 each with probability $\frac{1}{3}$, then cp^2 steps are required to achieve randomness. On the other hand, if $a = 2$ and $b = 0$, 1, or -1, each with probability $\frac{1}{3}$, then $c \log p \log\log p$ steps always suffice to guarantee randomness, and for infinitely many p, are necessary as well, although, in fact, for almost all odd p, $1.02 \log_2 p$ steps are enough.

1. Introduction. Computers often generate pseudorandom sequences using recurrences such as

$$X_{n+1} \equiv aX_n + b \pmod{p},$$

where p is a fixed integer (with $2^{31} - 1$ and 2^{32} being popular choices), and the integers a and b are chosen so that the sequence $X_0 = 0, X_1, X_2, \ldots$ has some of the properties of a random sequence. An extensive discussion of these matters can be found in Knuth [5].

Of course, however, the sequence X_n is deterministic and exhibits many regular aspects. To increase "randomness," several different generators are often combined or "shuffled." We investigate properties of the process

$$X_{n+1} \equiv a_n X_n + b_n \pmod{p}, \tag{1}$$

where a_n and b_n are independent random variables. These might be the output of another generator, or they might be the result of a "truly random" source produced, for example, by electrical noise or radioactive decay. The earliest pseudorandom number generators, due to Lehmer, were of this type. He used the recurrence $X_{n+1} \equiv aX_n \pmod{p}$ run at a constant rate (say, 1000 times per second). Calls to the generator depended on the execution times of various steps in the program, resulting in a random multiplier (see Knuth [5] for further details).

Received July 1985.

[1]Research partially supported by National Science Foundation Grant DMS 86-00235.

AMS 1980 *subject classifications*. Primary 60B15; secondary 60J15.

Key words and phrases. Random walk, Fourier analysis, discrete Fourier analysis, random number generation.

Most distributions for a_n and b_n in (1) lead to a uniform distribution for X_n as n tends to infinity. We will use the following measure of "closeness to uniformity" for X_n. Let

$$P_n(j) := \text{Prob}\{X_n = j\}, \qquad 0 \le j \le p-1,$$

and let U denote the uniform distribution, defined by $U(j) = 1/p$ for all j. Define the *variation distance* between P_n and U by

$$\|P_n - U\| := \tfrac{1}{2}\sum_j |P_n(j) - 1/p|. \tag{2}$$

It is well known (and easy to show) that

$$\begin{aligned} \|P_n - U\| &= \max_{A \subseteq \mathbb{Z}_p} |P_n(A) - U(A)| \\ &= \tfrac{1}{2} \sup_{\|f\|=1} |P_n(f) - U(f)|, \end{aligned} \tag{3}$$

where $\mathbb{Z}_p$ denotes $\mathbb{Z}/p\mathbb{Z}$, the integers modulo p, for $A \subseteq \mathbb{Z}_p$,

$$P_n(A) := \sum_{j \in A} P_n(j),$$

and for a function $f\colon \mathbb{Z}_p \to \mathbb{C}$,

$$\|f\| := \max_j f(j)$$

and

$$P_n(f) := \sum_j P_n(j) f(j).$$

Our primary focus in this paper will be on sequences X_n generated by

$$X_{n+1} \equiv aX_n + b_n \pmod{p},$$

where p is odd and $b_1, b_2, \ldots, b_n, \ldots$ are independent with common distribution μ on $\mathbb{Z}_p$. Our basic goal will be to estimate as sharply as we can the number N of steps needed to guarantee that $\|P_N - U\|$ is close to 0. Typically, we show the existence of a "threshold" N_T so that for $N \ll N_T$, $\|P_N - U\|$ is close to 1 while for $N \gg N_T$, $\|P_N - U\|$ goes to 0 exponentially fast.

The remainder of the paper is organized as follows. In Section 2 we review basic facts from Fourier analysis we will need. In Section 3 we discuss the (classical) case of $a = 1$. It is easy to show in this case that for general μ with bounded support, $N = c(p)p^2$ steps with $c(p) \to \infty$ as $p \to \infty$ are necessary and sufficient to drive $\|P_N - U\|$ to 0.

The situation is changed drastically if $a = 1$ is replaced by $a = 2$. In Section 4 we derive an upper bound of the form $c \log p \log\log p$ for N which, in Section 5, we show is best possible for infinitely many p. Finally, in Section 6 we prove that for almost all odd p, $N = c \log p$ steps suffice to drive $\|P_N - U\|$ to 0 for a fixed constant c (slightly larger than 1).

We end this section by commenting on two further problems which motivate the careful study of the special cases reported here. One basic technique we

employ is the "upper bound" lemma (Lemma 1) of Section 2. This has been used in numerous other random walk problems as detailed in Chapter 3 of [2]. In most problems, it tends to give sharp estimates, such as in the case $a = 1$. The case $a = 2$ is an example of a problem in which the upper and lower bound for specific values of p are of different orders of magnitude. Our investigations indicate that this is not an artifact arising from the use of the upper bound lemma but rather it is a consequence of the true state of affairs here.

We also point out that Markov chains very similar to (1) arise in computer graphics algorithms which draw realistic pictures of leaves, trees, and clouds by using iterated maps. Diaconis and Shahshahani [4] investigate such problems in a continuous setting.

2. Preliminaries. Many of our arguments involve the Fourier transform on $\mathbb{Z}_p$. To define this, let $q = q(p) = e^{2\pi i/p}$. For any complex-valued function f on $\mathbb{Z}_p$, define the Fourier transform $\hat{f}\colon \mathbb{Z}_p \to \mathbb{C}$ by

$$\hat{f}(j) = \sum_{k \in \mathbb{Z}_p} q^{kj} f(k). \tag{4}$$

The Fourier inversion theorem and the Plancherel theorem give, respectively,

$$f(k) = \frac{1}{p} \sum_{j \in \mathbb{Z}_p} q^{-kj} \hat{f}(j), \tag{5}$$

$$\sum_{k \in \mathbb{Z}_p} |f(k)|^2 = \frac{1}{p} \sum_{j \in \mathbb{Z}_p} \hat{f}(j)^2, \tag{6}$$

where $|x|^2 = x\bar{x}$. These follow easily from (4) (see also Chapter 6 of Serre [6]). A further useful fact concerning the uniform distribution U is

$$\hat{U}(j) = \begin{cases} 1, & \text{if } j = 0, \\ 0, & \text{otherwise.} \end{cases}$$

By combining these facts we obtain the following upper bound.

LEMMA 1.

$$\begin{aligned} \|P - U\|^2 &= \tfrac{1}{4}\Big(\sum_j |P(j) - U(j)|\Big)^2 \\ &\le \tfrac{1}{4} p \sum_j |P(j) - U(j)|^2 \qquad (7) \\ &= \tfrac{1}{4} \sum_{j \neq 0} |\hat{P}(j)|^2. \end{aligned}$$

In Lemma 1 we have used the Cauchy–Schwarz inequality and the fact that $\hat{P}(0) = 1$ (for any probability measure P).

This bound will usually be used when P is close to uniform, and then the Cauchy–Schwarz inequality is fairly accurate. The bound in (7) was used in the noncommutative case in [3].

Let P and Q be probability distributions on $\mathbb{Z}_p$ and let $P*Q$ denote the *convolution* of P and Q, defined by

$$P*Q(k) = \sum_{j\in\mathbb{Z}_p} P(j)Q(k-j).$$

The following facts (with proofs omitted) will be used in what follows:

(i) $\widehat{P*Q} = \hat{P}\cdot\hat{Q}$.

(ii) $\|P*Q - U\| \le \|Q - U\|$.

(iii) Let T: $\mathbb{Z}_p \to \mathbb{Z}_p$ be 1–1 and let PT^{-1} and QT^{-1} denote the corresponding image measures. Then

$$\|PT^{-1} - QT^{-1}\| = \|P - Q\|.$$

3. The case $a = 1$. We first discuss for illustrative purposes the process

$$X_{n+1} \equiv X_n + \varepsilon_n \pmod p, \tag{8}$$

where $\varepsilon_n = 0, 1$, or -1, each with probability $\frac{1}{3}$, and $X_0 = 0$.

FACT. There exist positive constants α and β such that

$$e^{-\alpha N/p^2} < \|P_N - U\| < e^{-\beta N/p^2}. \tag{9}$$

PROOF. The ε_n's in (8) have the common distribution μ where $\mu(0) = \mu(-1) = \mu(1) = \frac{1}{3}$. Thus,

$$\hat{\mu}(j) = \sum_k q^{kj}\mu(j) = \frac{1}{3} + \frac{2}{3}\cos\frac{2\pi j}{p}.$$

After N steps, the corresponding measure P_N is the convolution of μ with itself N times. By Lemma 1,

$$\|P_N - U\|^2 \le \frac{1}{4}\sum_{j\neq 0}|\hat{P}_N(j)|^2 = \frac{1}{4}\sum_{j\neq 0}\hat{\mu}^{2N}(j)$$

$$= \frac{1}{4}\sum_{j\neq 0}\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi j}{p}\right)^{2N}.$$

The right-hand side of (9) now follows by using elementary bounds on the cosine function, such as

$$\tfrac{1}{3} + \tfrac{2}{3}\cos x \le e^{-2x^2/9} \quad \text{for } 0 \le x \le \pi/2.$$

To prove the left-hand side of (9) we use (3) with the choice

$$f(j) = \cos\frac{2\pi j}{p}.$$

In this case,

$$U(f) = 0$$

and since

$$\begin{aligned} P_N(f) &= \sum_j P_N(j) f(j) \\ &= \sum_j P_N(j) \cos\frac{2\pi j}{p} \\ &= \operatorname{Re} \sum_j P_N(j) e^{\frac{2\pi j}{p}} \\ &= \operatorname{Re} \hat{P}_N(1) = \left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi}{p}\right)^N, \end{aligned}$$

then

$$\|P_N - U\| \geq \frac{1}{2}\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi}{p}\right)^N \geq e^{-\beta N/p^2}$$

for a suitable β. □

A more careful analysis shows that the correct rate is $e^{-\alpha N/p^2}$ with $\alpha = \frac{4}{3}\pi^2$, but we will not prove this here. By using the Berry–Esseen theorem to bound the random walk, it can be shown that a "cap" $\{k \in \mathbb{Z}_p: |k| \leq c\}$ for a fixed c, has probability close to 1 under P_N, but probability close to 0 under U, if N/p^2 is small (which in turn also implies that $\|P_N - U\|$ is close to 1).

4. The case $a = 2$—an upper bound. We consider here the process

$$X_{n+1} \equiv 2X_n + \varepsilon_n \pmod p, \qquad X_0 = 0, \tag{10}$$

where, as usual, p is odd and the ε_n all share a common distribution μ with $\mu(0) = \mu(-1) = \mu(1) = \frac{1}{3}$.

THEOREM 1. *Suppose X_n satisfies* (10) *and $c > 1/\log 9$. Then for $N \geq c \log p \log\log p$, we have $\|P_N - U\| \to 0$ as $p \to \infty$.*

PROOF. After N steps of the process we have

$$X_N \equiv \sum_{a=0}^{N-1} 2^{N-1-a}\varepsilon_a \pmod p. \tag{11}$$

Thus, the corresponding measure P_N induced by X_N is just the *convolution* of the N measures $\mu^{(a)}$ given by

$$\mu^{(a)}(0) = \mu^{(a)}(-2^a) = \mu^{(a)}(2^a) = \tfrac{1}{3}, \qquad 0 \leq a \leq N-1,$$

where, of course, the arguments of $\mu^{(a)}$ are reduced modulo p. Therefore,

$$\hat{P}_N = \prod_{a=0}^{N-1} \hat{\mu}^{(a)}$$

and, by Lemma 1,

$$\begin{aligned} \|P_N - U\|^2 &\le \frac{1}{4} \sum_{k \neq 0} \left|\hat{P}(k)\right|^2 \\ &= \frac{1}{4} \sum_{k \neq 0} \prod_{a=0}^{N-1} \left(\frac{1}{3} + \frac{2}{3} \cos \frac{2\pi 2^a k}{p} \right)^2. \end{aligned} \tag{12}$$

Let

$$g(x) := \left(\frac{1}{3} + \frac{2}{3} \cos 2\pi x \right)^2$$

and define the function h: $[0, 1] \to \mathbf{R}$ by

$$h(x) = \begin{cases} \frac{1}{9}, & \text{if } x \in \left[\frac{1}{4}, \frac{3}{4}\right), \\ 1, & \text{otherwise.} \end{cases}$$

Note that $g(x) \le h(x)$ for $0 \le x \le 1$. Thus, with $\{x\}$ denoting the fractional part of x, we have

$$\|P_N - U\|^2 \le \frac{1}{4} \sum_{k \neq 0} \prod_{a=0}^{N-1} h\left(\left\{ \frac{2^a k}{p} \right\} \right). \tag{13}$$

If we write $x \in [0, 1)$ in its binary expansion

$$x = \alpha_1 \alpha_2 \alpha_3, \ldots, \qquad \alpha_i = 0 \text{ or } 1 \quad (\text{where } \alpha_i = 0 \text{ infinitely often}),$$

then

$$h(x) = \tfrac{1}{9} \quad \text{if and only if } \alpha_1 \neq \alpha_2.$$

Thus, if $A_x(N)$ denotes the number of "alternations" in the first N binary digits of x, i.e.,

$$A_x(N) := \left| \{ 1 \le i < N \colon \alpha_i \neq \alpha_{i+1} \} \right|,$$

then

$$\prod_{a=0}^{N-1} h\left(\left\{ \frac{2^a k}{p} \right\} \right) = 9^{-A_{k/p}(N+1)}. \tag{14}$$

Define the integer t to satisfy

$$2^{t-1} < p < 2^t.$$

We shall choose N to be of the form rt for a large integer $r = r(t)$ to be specified later.

Consider the first $N = rt$ binary digits of the binary expansion of

$$k/p = \alpha_1\alpha_2 \cdots \alpha_t\alpha_{t+1} \cdots \alpha_{2t} \cdots \alpha_{rt} \cdots .$$

Partition this string into r disjoint blocks B_{ki}, $1 \le i \le r$, each of length t, by defining

$$B_{ki} = \alpha_{(i-1)t+1}\alpha_{(i-1)t+2} \cdots \alpha_{it}.$$

Let $A(B_{ki})$ denote the number of alternations in the block B_{ki}. Thus,

$$\prod_{a=0}^{N-1} h\left(\left\{\frac{2^a k}{p}\right\}\right) \le \prod_{i=1}^{r} 9^{-A(B_{ki})}, \tag{15}$$

where the inequality sign allows for the fact that we have ignored the possible interblock alternations, and that we are only using the first N digits of k/p (and not $N + 1$ as specified by (14)).

Next observe that as k ranges over $\mathbb{Z}_p \setminus \{0\}$, the blocks B_{k1} are all distinct and all have at least one alternation (by the choice of t). Furthermore, since $(2, p) = 1 = (2^t, p)$, then for each i, the set of blocks $\{B_{ki}: 1 \le k \le p - 1\}$ is identical with the set $\{B_{k1}: 1 \le k \le p - 1\}$.

By (13) and (15) we have

$$\|P_N - U\|^2 \le \tfrac{1}{4} \sum_{k \ne 0} \prod_{i=1}^{r} 9^{-A(B_{ki})}. \tag{16}$$

Since for $a \le a'$, $b \le b'$, and $0 < \gamma < 1$ we have

$$\gamma^{a+b'} + \gamma^{a'+b} \le \gamma^{a+b} + \gamma^{a'+b'}, \tag{17}$$

then by successively interchanging pairs of exponents $A(B_{ki})$, $A(B_{k'i})$ (using (17) with $\gamma = \frac{1}{9}$ and the fact that the sets $\{B_{ki}: 1 \le k \le p - 1\}$ are identical for $1 \le i \le r$), we obtain

$$\sum_{k \ne 0} \prod_{i=1}^{r} 9^{-A(B_{ki})} \le \sum_{k \ne 0} 9^{-rA(B_{k1})}. \tag{18}$$

Since the B_{k1}, $1 \le k \le p - 1$, are all distinct and have at least one alternation, the right-hand side of (18) is upper-bounded by summing over all blocks B of length t having at least one alternation:

$$\sum_{k \ne 0} 9^{-rA(B_{k1})} \le \sum_{\substack{\text{length } B = t \\ A(B) > 0}} 9^{-rA(B)}. \tag{19}$$

If $M(j)$ denotes the number of blocks of length t with exactly j alternations then

$$M(j) \le 2\binom{t-1}{j} \le 2\binom{t}{j}.$$

Therefore, combining the preceding estimates, we have

$$\begin{aligned}
\|P_N - U\|^2 &\le \tfrac{1}{4} \sum_{k \neq 0} 9^{-rA(B_{k1})} \\
&\le \tfrac{1}{4} \sum_{\substack{\text{length } B = t \\ A(B) > 0}} 9^{-rA(B)} \\
&\le \tfrac{1}{4} \sum_{j=1}^{t} M(j) 9^{-rj} \qquad (20) \\
&\le \tfrac{1}{2} \sum_{j=1}^{t} \binom{t}{j} 9^{-rj} \\
&= \tfrac{1}{2}\left\{(1 + 9^{-r})^t - 1\right\} \\
&\le \tfrac{1}{2}(e^{t9^{-r}} - 1).
\end{aligned}$$

Thus, for

$$r \ge \frac{\log t}{\log 9} + d,$$

$$\|P_N - U\|^2 \le \tfrac{1}{2}(e^{9^{-d}} - 1),$$

which goes to 0 as $d \to \infty$. This proves the theorem. □

REMARK. For the much easier special case that $p = 2^t - 1$, a probabilistic argument can be given for Theorem 1 (see [1]).

5. The case $a = 2$—a lower bound. The main result of this section deals with moduli p of a special form, namely, $p = 2^t - 1$. One reason for suspecting that these p might have particularly slow convergence of P_N to uniform is that for each k, all the blocks B_{ki}, $1 \le i \le r$, introduced in the preceding section, are identical, so that equality holds in (18). It turns out that for these p (and many others as well) the estimate in Theorem 1 is essentially best possible.

THEOREM 2. *Suppose X_n satisfies* (10). *Then for a suitable $c' > 0$, if $p = 2^t - 1$ and $N \le c' \log p \log\log p$ then $\|P_N - U\|$ is bounded away from 0 as $t \to \infty$.*

PROOF. Our strategy will be to introduce what might be called a "separating" function $f\colon \mathbb{Z}_p \to \mathbb{C}$, defined by

$$f(k) = \sum_{j=0}^{t-1} q^{k2^j}.$$

We will take $N = rt$ where r is an integer of the form $r = \alpha \log t - d$ for a fixed constant α to be specified later.

Define

$$\Pi_j := \prod_{a=0}^{t-1}\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi 2^a(2^j-1)}{p}\right)$$

for $1 \le j \le t-1$. We will compute the mean and variance of f with respect to the two measures U and P_N.

To begin with, it is easy to see that for U the mean is

$$E_U(f) = \sum_k U(k)f(k) = 0 \tag{21}$$

and the second moment is

$$\begin{aligned} E_U(f\bar{f}) &= \sum_k U(k)f(k)\overline{f(k)} \\ &= \frac{1}{p}\sum_k \sum_{j,j'} q^{k(2^j-2^{j'})} = t. \end{aligned} \tag{22}$$

For the measure P_N we have

$$\begin{aligned} E_{P_N}(f) &= \sum_k P_N(k)f(k) \\ &= \sum_k \sum_{j=0}^{t-1} P_N(k)q^{k2^j/p} \\ &= \sum_{j=0}^{t-1} \hat{P}_N(2^j) \\ &= \sum_{j=0}^{t-1}\prod_{a=0}^{t-1}\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi 2^a 2^j}{p}\right)^r = t\Pi_1^r. \end{aligned} \tag{23}$$

In the same way we compute

$$\begin{aligned} E_{P_N}(f\bar{f}) &= \sum_k P_N(k)f(k)\bar{f}(k) \\ &= \sum_k \sum_{j,j'} P_N(k)q^{k(2^j-2^{j'})}/p \\ &= \sum_{j,j'} \hat{P}_N(2^j - 2^{j'}) \\ &= t\sum_{j=0}^{t-1} \Pi_j^r. \end{aligned} \tag{24}$$

Thus, the variances of f under U and P_N are

$$\begin{aligned} \mathrm{Var}_U(f) &= t, \\ \mathrm{Var}_{P_N}(f) &= t\sum_{j=0}^{t-1} \Pi_j^r - t^2\Pi_1^{2r}. \end{aligned} \tag{25}$$

We will need the following complex form of Chebyshev's inequality:

$$\Pr\left\{x: |f(x) - E_Q(f)| \geq \alpha\sqrt{\mathrm{Var}_Q(f)}\right\} \leq 1/\alpha^2, \tag{26}$$

for an arbitrary probability distribution Q (on $\mathbb{Z}_p$) and an arbitrary constant $\alpha > 0$.

For the cases $Q = U$ and $Q = P_N$, (26) implies

$$\begin{aligned} &\Pr_U\{x: |f(x)| \geq \alpha t^{1/2}\} \leq 1/\alpha^2, \\ &\Pr_{P_N}\left\{x: |f(x) - t\Pi_1^r| \geq \beta\left(t\sum_{j=0}^{t-1}\Pi_j^r - t^2\Pi_1^{2r}\right)^{1/2}\right\} \leq 1/\beta^2. \end{aligned} \tag{27}$$

Thus, if A and B denote the *complements* of these two sets, respectively, then

$$\Pr_U(A) \geq 1 - 1/\alpha^2, \qquad \Pr_{P_N}(B) \geq 1 - 1/\beta^2. \tag{28}$$

If in fact the sets A and B are disjoint, then from (28) we immediately obtain the lower bound

$$\|P_N - U\| \geq 1 - 1/\alpha^2 - 1/\beta^2. \tag{29}$$

We next specify r to be an *even* integer of the form

$$r = \frac{\log t}{2\log(1/|\Pi_1|)} - \lambda \tag{30}$$

for a fixed number λ. In this case,

$$t\Pi_1^r = t^{1/2}|\Pi_1|^{-\lambda} \gg ct^{1/2} \tag{31}$$

for any fixed c, provided λ is sufficiently large since $|\Pi_1|$ is bounded away from both 0 and 1 as $t \to \infty$.

CLAIM.

$$\frac{1}{t}\sum_{j=0}^{t-1}\left(\frac{\Pi_j}{\Pi_1^2}\right)^r \to 1 \quad \text{as } t \to \infty. \tag{32}$$

REMARK. It is easy to see that (32) implies that

$$\left(\mathrm{Var}_{P_N}(f)\right)^{1/2} = o\left(E_{P_N}(f)\right). \tag{33}$$

Hence, by choosing $\alpha = \beta = 2$ and $\lambda \geq 1$, for example, we see that A and B are disjoint for t sufficiently large and $\|P_N - U\| \geq \frac{1}{2}$. Since in this case

$$N = rt = \left(\frac{\log t}{2\log(1/|\Pi_1|)^{-1}}\right)t > c'\log p \log\log p,$$

the theorem then follows.

It remains to prove the claim. To begin with, define

$$G(x) = \left|\tfrac{1}{3} + \tfrac{2}{3}\cos 2\pi x\right|. \tag{34}$$

Thus,

$$|\Pi_j| = \prod_{a=0}^{t-1} G\left(\frac{2^a(2^j-1)}{p}\right). \tag{35}$$

Note that

$$0 \le x < y \le \tfrac{1}{3} \Rightarrow G(x) > G(y),$$

$$\tfrac{1}{3} \le y < x \le \tfrac{1}{2} \Rightarrow G(x) > G(y).$$

FACT 1. $|\Pi_j| \le |\Pi_1|$ for all $j \ge 1$.

PROOF OF FACT 1. The strategy will be to pair off each factor x of Π_j with a corresponding factor $\pi(x)$ of Π_1 so that $|x| \le |\pi(x)|$. (Strictly speaking, we do this for all but two factors of Π_j and Π_1, as we shall see.) Since $G(x) = G(1-x)$, we can assume without loss of generality that $2 \le j \le t/2$. Thus,

$$\Pi_j = \prod_{i=0}^{t-j-1} G\left(\frac{2^{i+j}-2^i}{p}\right) \prod_{i=0}^{j-1} G\left(\frac{2^{i+t-j}-2^i}{p}\right). \tag{36}$$

We make the following association:

$$x \leftrightarrow \pi(x),$$

$$G\left(\frac{2^{i+j}-2^i}{p}\right) \leftrightarrow G\left(\frac{2^{i+j-1}}{p}\right), \quad \text{for } 0 \le i \le t-j-2,$$

$$G\left(\frac{2^{t-1}-2^{t-j-1}}{p}\right) \leftrightarrow G\left(\frac{2^{t-1}}{p}\right),$$

$$G\left(\frac{2^{i+t-j}-2^i}{p}\right) \leftrightarrow G\left(\frac{2^i}{p}\right), \quad \text{for } 0 \le i \le j-3,$$

$$G\left(\frac{2^{t-1}-2^{j-1}}{p}\right) G\left(\frac{2^{t-2}-2^{j-2}}{p}\right) \leftrightarrow G\left(\frac{2^{j-2}}{p}\right) G\left(\frac{2^{t-2}}{p}\right).$$

It is straightforward to check that the product of the factors under x is Π_j, and the product of the factors under $\pi(x)$ is Π_1. Furthermore, $|x| \le |\pi(x)|$ for each pair of associated terms (as well as for the bottom pair of associated products). This proves Fact 1. □

By a similar (but slightly more complex) argument, the following result can be proved.

FACT 2. There is an absolute constant c_0 such that for $t^{1/3} \le j \le t/2$ we have

$$\frac{\Pi_j}{\Pi_1^2} \le 1 + \frac{c_0}{2^j}. \tag{37}$$

A straightforward calculation now shows that

$$\sum_{t^{1/3} \le j \le t/2} \left| \left(\frac{\Pi_j}{\Pi_1^2} \right)^r - 1 \right| \le \frac{c_1 tr}{2^{t^{1/3}}} < \frac{c_2}{2^{t^{1/4}}} \tag{38}$$

for absolute constants c_1 and c_2. Therefore, since $\Pi_j = \Pi_{t-j}$, then

$$\begin{aligned} \frac{1}{t} \sum_{j=0}^{t-1} \left(\frac{\Pi_j}{\Pi_1^2} \right)^r &\le \frac{2}{t} \left(\sum_{0 \le j < t^{1/3}} \left(\frac{\Pi_j}{\Pi_1^2} \right)^r + \sum_{t^{1/3} \le j \le t/2} \left(\frac{\Pi_j}{\Pi_1^2} \right)^r \right) \\ &= 1 + o(1) \end{aligned}$$

as $t \to \infty$. This completes the proof of the Claim and Theorem 2. □

Essentially the same argument can be applied to other p, such as those of the form $2^t + 1, 2^{2t} \pm 2^t + 1$, etc., for which the value of $\operatorname{ord}_2(p)$ is small, where $\operatorname{ord}_2(p)$ is defined to be the least positive integer satisfying

$$2^{\operatorname{ord}_2(p)} \equiv 1 \pmod{p}.$$

However, these p are rather exceptional, as the result in the next section shows.

6. A bound for almost all odd p. In this section, we show that in fact almost all odd p require far fewer steps to bring P_N close to uniform than the upper bound given in Theorem 1.

THEOREM 3. *Suppose X_n satisfies* (10). *Then for almost all odd p, if*

$$N \ge \frac{\log p}{\log(9/5)} + c,$$

then

$$\|P_N - U\| = O\left(\left(\tfrac{5}{9} \right)^{c/2} \right).$$

PROOF. We will proceed as in the proof of Theorem 1, defining $g(x)$, $h(x)$, and $A_x(N)$ in the same way. Further, define $f_k(p)$ and $f(p)$ by

$$\begin{aligned} f_k(p) &:= \prod_{a=0}^{N-1} h(\{2^a k/p\}), \\ f(p) &:= \sum_{k=1}^{p-1} f_k(p). \end{aligned} \tag{39}$$

By applying the argument used to establish (15), it follows that

$$f_k(p) \le 9^{-A_{k/p}(N)}. \tag{40}$$

Note that by (13),

$$\|P_N - U\|^2 \le \frac{1}{4} f(p). \tag{41}$$

Choose a large fixed integer t and consider the set of all $f_k(p)$ where

$$2^{t-1} < p < 2^t, \qquad 1 \le k \le p-1$$

(and where we always assume p is odd).

Define

$$S := \sum_{2^{t-1}<p<2^t} \sum_{k=1}^{p-1} f_k(p).$$

Of course, if $k/p = k'/p'$ then $f_k(p) = f_{k'}(p')$. Let $M(r/s)$ denote the number of pairs (k, p) with $2^{t-1} < p < 2^t$, $1 \le k \le p-1$, which satisfy $k/p = r/s$ where $(r, s) = 1$. Therefore,

$$\begin{aligned} S &= \sum_{\substack{1\le s<2^t \\ 1\le r\le s-1 \\ (r,s)=1 \\ s \text{ odd}}} M\left(\frac{r}{s}\right) f_r(s) \\ &= \sum_{\omega=1}^{t} {\sum}' M\left(\frac{r}{s}\right) f_r(s) \\ &\le 2^t \sum_{\omega=1}^{t} {\sum}' \frac{f_r(s)}{s} \qquad \left(\text{since } M\left(\frac{r}{s}\right) < \frac{2^t}{s}\right) \\ &\le 2^t \sum_{\omega=1}^{t} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)} \quad \text{(by (40))}, \end{aligned} \tag{42}$$

where Σ' denotes the sum taken over all r and s satisfying

$$2^{\omega-1} < s < 2^\omega, \qquad 1 \le r \le s-1, \qquad (r, s) = 1 \qquad \text{and } s \text{ odd}.$$

We will partition the sum

$$S^* := \sum_{\omega=1}^{t} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)}$$

into three sums as follows:

$$S_1 := \sum_{\omega \le t/\log t} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)},$$

$$S_2 := \sum_{t/\log t < \omega \le N/2} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)},$$

$$S_3 := \sum_{N/2 < \omega \le t} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)}.$$

We should keep in mind that we are assuming $N = \log p/\log \frac{9}{5} + c$ for a fixed positive c.

(i) *Bounding* S_1. Let $N' = [N/2\omega]$ and for each $r/s = \alpha_1\alpha_2 \cdots \alpha_N \cdots$, partition the first $2\omega N'$ binary digits of r/s into N' blocks $B_i(r/s)$, each of length 2ω, as follows:

$$\frac{B_1(r/s)}{\alpha_1\alpha_2 \cdots \alpha_{2\omega}} \frac{B_2(r/s)}{\alpha_{2\omega+1} \cdots \alpha_{4\omega}} \frac{B_N(r/s)}{\cdots \alpha_{2\omega N'}}.$$

Observe that since $(2, s) = 1$, then for

$$R(s) := \{1 \le x < s: (x, s) = 1\}$$

we have

$$\{2r \pmod s: r \in R(s)\} = R(s).$$

It follows that all of the sets

$$\{B_i(r/s): r \in R(s)\}, \qquad 1 \le i \le N',$$

are *equal*, i.e., independent of i. Furthermore, since for

$$2^{\omega-1} < s, s' < 2^\omega, \qquad (r, s) = (r', s') = 1,$$

if $r/s \ne r'/s'$, then

$$|r/s - r'/s'| \ge 1/ss' > 2^{-2\omega}.$$

Hence, all the blocks

$$B_1(r/s), \qquad 2^{\omega-1} < s < 2^\omega, \qquad 1 \le r \le s-1, \qquad (r, s) = 1,$$

are distinct. Also, since $s < 2^\omega$, each $B_1(r/s)$ has at least one alternation.

Next, for each s, we can apply the "interchange" technique used in deriving (18) and obtain

$$\sum_{r \in R(s)} 9^{-\sum_{i=1}^{N'} A(B_i(r/s))} \le \sum_{r \in R(s)} 9^{-N'A(B_1(r/s))}.$$

Therefore,

$$\begin{aligned}
\sum{}' \frac{1}{s} 9^{-A_{r/s}(N)} &\ll 2^{-\omega} \sum_{\substack{\text{length}(B)=2\omega \\ A(B)\ge 1}} 9^{-N'A(B)} \\
&\le 2^{-\omega} \sum_{j=1}^{2\omega} \binom{2\omega}{j} 9^{-jN'} \\
&= 2^{-\omega}\left[(1 + 9^{-N'})^{2\omega} - 1\right] \\
&\le 2^{-\omega}(e^{9^{-2\omega N'}} - 1) \ll \omega 2^{-\omega} 9^{-N'}.
\end{aligned}$$

Continuing the inequality for S_1 we have

$$S_1 \ll \sum_{\omega \le t/\log t} \omega 2^{-\omega} 9^{-N'}$$

$$\ll 9^{-(N \log t)/2t},$$

since $\sum_{\omega \ge 1} \omega 2^{-\omega}$ is bounded.

Thus, for $N = t/\log_2(\frac{9}{5}) + c$,

$$S_1 \ll t^{-(1/2)\log 9(1/\log(9/5)+c/t)}. \tag{43}$$

(ii) *Bounding* S_2. Starting with the definition

$$S_2 = \sum_{t/\log t < \omega \le N/2} {\sum}' \tfrac{1}{2} 9^{-A_{r/s}(N)},$$

we can proceed as in the case of S_1 until we come to the point where we have

$$S_2 \ll \sum_{t/\log t < \omega \le N/2} 2^{-\omega}\left((1 + 9^{-N'})^{2\omega} - 1\right).$$

We now continue as follows:

$$\begin{aligned} S_2 &\ll N \max_{t/\log t < \omega \le N/2} 2^{-\omega}\left((1 + 9^{-N'})^{2\omega} - 1\right) \\ &\ll N \max_{t/\log t < \omega \le N/2} (50/81)^{\omega} \qquad (44) \\ &< N(50/81)^{t/\log t}, \end{aligned}$$

since $N' \ge 1$.

(iii) *Bounding* S_3. We begin by partitioning

$$[0,1) = \bigcup_{\alpha} I(\alpha),$$

where α ranges over all binary rationals $u/2^N$, $0 \le u < 2^N$, and

$$I(\alpha) = [a, a + 2^{-N}).$$

Therefore,

$$S_3 = \sum_{N/2 < \omega \le t} {\sum}' \frac{1}{s} 9^{-A_{r/s}(N)}$$

$$= \sum_{\alpha} \sum_{N/2 < \omega \le t} \left(\sum_{\substack{2^{\omega-1} < s < 2^{\omega} \\ 1 \le r \le s-1 \\ (r,s)=1 \\ s \text{ odd} \\ r/s \in I(\alpha)}} \frac{1}{s} \right) 9^{-A(\alpha)},$$

where, of course, $A(\alpha)$ denotes the number of alternations in the binary N-tuple α.

Our next task is to estimate the inner sum. This we do as follows. First, we write

$$\sum_{r/s\in I(\alpha)}{}' \frac{1}{s} := \sum_{\substack{2^{\omega-1}<s<2^{\omega}\\ 1\le r\le s-1\\ (r,s)=1\\ s \text{ odd}\\ r/s\in I(\alpha)}} \frac{1}{s} \le 2^{-(\omega-1)}\left|\left\{\frac{r}{s}: \frac{r}{s}\in I(\alpha)\right\}\right|.$$

However, since $r/s \neq r'/s'$ implies

$$\left|\left\{\frac{r}{s}: \frac{r}{s}\in I(\alpha)\right\}\right| \ll 2^{2\omega-N},$$

consequently,

$$\sum_{r/s\in I(\alpha)}{}' \frac{1}{s} \ll 2^{\omega-N}.$$

Thus,

$$\begin{aligned} S_3 &\ll \sum_{\alpha}\sum_{N/2<\omega\le t} 2^{\omega-N}9^{-A(\alpha)} \\ &\ll 2^{t-N}\sum_{\alpha} 9^{-A(\alpha)} \\ &\ll 2^{t-N}\sum_{j=0}^{N}\binom{N}{j}9^{-j} = 2^t\left(\tfrac{5}{9}\right)^N. \end{aligned}$$

For $N = t/\log_2(\frac{9}{5}) + c$, this gives

$$S_3 \ll \left(\tfrac{5}{9}\right)^c. \tag{45}$$

Now we can combine the estimates for S_1, S_2, and S_3. Since for fixed $c \ge 0$, $S_1 = o(1)$ and $S_2 = o(1)$ as $t \to \infty$, we have

$$S \ll 2^t(S_1 + S_2 + S_3) < c'2^t\left(\tfrac{5}{9}\right)^c \tag{46}$$

for an absolute constant c'.

Since there are 2^{t-2} odd integers p in $(2^{t-1}, 2^t)$ and we have just shown that

$$\sum_{\substack{2^{t-1}<p<2^t\\ p \text{ odd}}} f(p) < c'2^t\left(\tfrac{5}{9}\right)^c$$

for $N = t/\log_2(\frac{9}{5}) + c$, then the number of odd p's in $(2^{t-1}, 2^t)$ which have $f(p) > c''(\frac{5}{9})^c$ is at most $(c'/c'')2^t$.

Since c'' is arbitrary then it follows that for each fixed $c \ge 0$, almost all odd integers p have, by (41),

$$\|P_N - U\|^2 = O\left(\tfrac{5}{9}\right)^c$$

for $N = \log p/\log \frac{9}{5} + c$. This proves Theorem 3. □

7. Concluding remarks. If a better upper bound to $g(x)$ is used, rather than the simple step function $h(x)$, the preceding estimates can all be

strengthened, but at the expense of considerably more complicated arguments. We plan to focus on these in a later paper—here, we will just summarize some of the results.

Define c^* by

$$c^* = \prod_{k=1}^{\infty}\left(\frac{1}{3} + \frac{2}{3}\cos\frac{2\pi}{2^k}\right)^{-2} = 143.659\cdots .$$

The "correct" values of the constants c and c' in Theorems 1 and 2, respectively, turn out to both be equal to $(\log c^*)^{-1}$.

In the case of Theorem 3, our bound is not as sharp. The best upper bound we can obtain is this:

For any $\varepsilon > 0$ and almost all odd p, if $N \geq (\hat{c} + \varepsilon)\log_2 p$ then $\|P_N - U\| < \varepsilon$, where

$$\hat{c} = \left(1 - \log_2\left(\frac{5 + \sqrt{17}}{9}\right)\right)^{-1} = 1.01999186\cdots .$$

It is conceivable that in fact $(1 + o(1))\log_2 p$ steps are enough for almost all p to force P_N to converge to uniform.

On the other hand, we are at present unable to exhibit for any fixed constant c, an explicit sequence of p's for which $c \log p$ steps suffice.

We should point out that similar techniques can be applied to more general sequences of the form

$$X_{n+1} \equiv \alpha X_n + b_n \pmod{p} \tag{47}$$

where the b_n share a common distribution μ of bounded support. These problems can usually be dealt with by applying the remarks (ii) and (iii) mentioned earlier together with the methods we have employed. For example, it is easy to show that if $\alpha = 2$ and $\mu(1) = \mu(-1) = \frac{1}{2}$ in (47) then $N = c \log p$ steps suffice, for some absolute constant c. If instead, we now consider (47) with $\alpha = 2$ and $\mu'(1) = \mu'(-1) = \frac{1}{4}$, $\mu'(0) = \frac{1}{2}$ then this measure μ' is (a scaled version of) the convolution of μ with itself. Thus, by (ii), the corresponding measure P'_N converges to uniform at least as rapidly as P_N, and so, $c \log N$ steps also suffice for this case, as well.

This, in fact, strikes the authors as somewhat curious since all three of the processes are of the form

$$X_{n+1} \equiv 2X_n + b_n \pmod{p},$$

where the b_n share the common distribution μ with $\mu(1) = \mu(-1) = \beta$, $\mu(0) = 1 - 2\beta$. Thus, if $\beta = \frac{1}{4}$ or $\beta = \frac{1}{2}$, $c \log p$ steps suffice. However, if $\beta = \frac{1}{3}$ then $c \log p \log\log p$ steps may be required (e.g., when $p = 2^t - 1$). It would be very interesting to know which value of β maximizes the value of N required for $\|P_N - U\| \to 0$.

Acknowledgments. The authors wish to thank Rob Calderbank, Peter Frankl, Narendra Karmarkar, Colin Mallows and Andrew Odlyzko for helpful suggestions at various points of this research.

REFERENCES

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

[2] DIACONIS, P. (1987). *Group Theory in Statistics*. IMS, Hayward, Calif. To appear.

[3] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions, *Z. Wahrsch. verw. Gebiete* **57** 159–179.

[4] DIACONIS, P. and SHAHSHAHANI, M. (1986). Products of random matrices and random walks on groups. *Contemp. Math.* **50** 183–195.

[5] KNUTH, D. E. (1973). *The Art of Computer Programming* **2**, 2nd ed. Addison-Wesley, Reading, Mass.

[6] SERRE, J. P. (1977). *Linear Representations of Finite Groups*. Springer, New York.

F. R. K. CHUNG
BELL COMMUNICATIONS RESEARCH
ROOM 2L-387
435 SOUTH STREET
MORRISTOWN, NEW JERSEY 07960

PERSI DIACONIS
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

R. L. GRAHAM
AT & T BELL LABORATORIES
600 MOUNTAIN AVENUE
MURRAY HILL, NEW JERSEY 07974