JOURNAL OF COMPUTATIONAL PHYSICS 32, 168–211 (1979)

The Simulation of Random Processes on Digital Computers with Čebyšev Mixing Transformations*

T. ERBER,[†]

Department of Physics, University of California, Los Angeles, California 90024

P. EVERETT,

Department of Electrical Engineering, Illinois Institute of Technology, Chicago, Illinois 60616

AND

P. W. JOHNSON

Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616

Received June 15, 1978; revised October 17, 1978

The iteration of the Čebyšev polynomial $x^2 - 2$ generates a mixing transformation on the interval $x \in [-2, 2]$. Extensive computer experiments have demonstrated that this is a convenient method for generating sequences of pseudo-random numbers. Despite the eventual domination of cumulative roundoff errors the asymptotic statistical features of the mixing are preserved. Multiple sequences of stochastically independent variables may be generated by these techniques. In practical computations the Čebyšev mixing eventually terminates in long pseudo-ergodic cycles. These results are linked with the general problem of simulating the stochastic behavior of physical systems by means of functional iteration.

1. INTRODUCTION

It is well known that the standard multiplicative congruence schemes for generating sequences of pseudo-random numbers have correlational defects [1]. The appearance of these correlations is consistent with Hamming's astringent criterion "... if we can

* Supported by the Research Corporation.

[†] On leave from Illinois Institute of Technology, Chicago, Illinois 60616.

0021-9991/79/080168–44$02.00/0

see no pattern with reference to the particular application, then the sequence of numbers are random for that application, and we must be content with little more" [2]. However, in large scale computer applications involving the parallel operation of many random number generators it is vital to try for a "little more." In particular in reactor and plasma simulations or the modeling of phase transitions in complex systems it is essential to have confidence from the outset that the results will reflect the inherent characteristics of the physical processes and not the latent stochastic dependences among the random number generators. One possible means for improving the performance of random number algorithms is to draw on some of the results obtained during the last few decades in statistical physics and ergodic theory. For example, if strings of digits are generated by means of mixing transformations, there is a wealth of information which guarantees—at least in a theoretical sense—that the resulting sequences will have all the "right" properties associated with random numbers, i.e., equidistribution, normality, auto- and cross-correlations equivalent to white noise, ergodicity, statistical stability, etc. By contrast, the multiplicative and additive congruence schemes which are recommended for most computer software have a narrower conceptual footing in the theory of normal numbers [3–8]. Since normality is in general not even preserved under a change of base [9], and practical simulations are liable to fall short of theoretical expectations, cf. Table III, it is not surprising that the congruence schemes are less versatile than mixing transformations in modeling the behavior of random systems.

The practical implementation of any random number algorithm—excepting hybrid devices [10]—is basically an experimental problem since digital networks are restricted to processing finite sets whereas all the theoretical schemes promise random behavior essentially only on sets of positive (non-atomic) measure. It is therefore a fortunate circumstance that the pseudo-random mixing sequences generated by the iteration of the second order Čebyšev polynomial $x^2 - 2$, on the interval $[-2, 2]$, can be approximated with excellent statistical fidelity on a wide variety of digital devices. In fact a central result of this work is that the numbers $(4/\pi)\cos^{-1}(\frac{1}{2}Z_{n+1}) - 2$, are to a very good approximation uniformly and "randomly" distributed between -2 and $+2$ when Z_{n+1} is calculated in double precision from the simple recurrence $Z_{n+1} = Z_n^2 - 2$ [see eq. (3.27)]. Extensive trials carried out with such diverse machines as the HP-25, SR-52, SR-56, and HP-9100 (programmable calculators), and computers such as IBM 360/195, and UNIVAC 1108 have demonstrated that the computer generated sequences continue to exhibit all the "right" pseudo-random features even after the cumulative roundoff and truncation errors have been amplified by the iterations to the point where they dominate the numerical aspects of the calculations. Of course all deterministic simulations of random processes on finite devices must eventually terminate in fixed points or fall into loops. In the Čebyšev case the computer experiments indicate that there are long intervals of "free running" preceding the appearance of terminal cycles: Specifically for N-state machines both the observed free running periods as well as the cycle lengths correspond to the optimum random value $N^{1/2}$; moreover, all the statistical properties of the Čebyšev mixing are preserved even on the cycles.

For purposes of exposition it is convenient to consider the practical computing aspects of the Čebyšev iterations separately from their deeper ergodic properties. For instance the dependence of the rate of mixing or randomization on the machine capacity is of direct interest to the numerical analyst, whereas the correspondence of the mixing rate with say the Kolmogorov entropy has a subtler connection with practical computations. Accordingly we begin with a brief recapitulation of general criteria for pseudo-random number generators (Section 2), and then present a summary of our practical experience in simulating the $x^2 - 2$ mixing on a variety of slow and fast machines (Section 3). The diagnostics which enable us to monitor how well the machine generated sequences maintain their random character are of course all statistical in nature: These tests can be used to verify that the computer generated sequences indeed conform to the general randomness criteria given in Section 2, and also furnish the link to the underlying mathematical theory of mixing transformations. In Section 4 we give a summary of this theory including results on product measures which are useful for the parallel generation of stochastically independent sequences [11, 12].

Since the $x^2 - 2$ mixing effectively acts as an amplifier for computing noise, e.g., roundoff or truncation errors, it is obvious that corresponding to every machine and programming mode there will be a characteristic threshold in the number of iterations beyond which all numerical accuracy is obliterated. However, in virtue of the non-vanishing topological entropy of the mixing the randomization "renews" itself with every iteration and therefore the progressive numerical distortions introduced by computing noise are countered by a statistical stabilization of the mixing sequences. Where the balance is struck in practise between these opposing tendencies is of cours an experimental question which can be studied with variable precision simulations, e.g., 1–60 bit routines (Section 5). The results confirm the intuitive expectation that larger capacity machines are superior in sustaining the pseudo-random character of the mixing.

A curious feature which can be interpreted as a kind of stability property of the terminal cycles became apparent during the computer trials. Specifically, if the mixing simulations are started at arbitrary machine numbers the iterations naturally merge into terminal cycles or fixed points. However, extensive checks have shown that there appear to be very few distinct terminal cycles corresponding to a given machine and programming mode; for instance on a 12 digit device (HP-9100) we have so far discovered only three terminal cycles. By mapping out the de Bruijn diagrams [13, 14] corresponding to the complete set of iterations or orbits which can be generated on a particular machine it appears that the terminal cycles are isolated in a topological sense (Section 6).

The scattered results available in the literature indicate that the empirical construction of functions whose iterates simulate chaotic behavior is not likely to lead to practical algorithms for generating pseudo-random sequences [15]. However, the experience gained with the Čebyšev mixing suggests that iterative processes for which there is an *a priori* theoretical assurance of pseudo-random behavior may be promising candidates for computer simulation. An extensive inventory of such processes ranging from Anosov flows [16], to broken linear (bakers') transformations [17], and billiard collisions in convex domains [18] has been studied in connection with general ergodic problems [19]. However, it remains to be seen whether any of these processes can be approximated on computers with a statistical fidelity matching the Čebyšev iterations. In Section 7 we consider some general mathematical constraints on the prospects for the success of this program based on the theorems of Šarkovskii [20] and Brolin [21]. For general functional iterates, such as the Čebyšev mixing, which cannot be embedded in flows, we show that the corresponding physical processes, e.g., product detection, cannot be described in terms of Hamiltonians.

2. Randomness Criteria

It is convenient to begin with a brief recapitulation of various randomness criteria. A complementary summary oriented towards practical computations may be found in Knuth [22]; the deeper logical and philosophical aspects of the subject have been discussed by Carnap [23].

(i) *Equidistribution.* Strings of random digits can in principle be generated by repeated independent drawings from the population 0, 1, 2,..., 9; where the probability of selecting any particular digit is one-tenth [8]. A deterministic computer simulation of this process corresponds to generating sequences $\{x_n\}$ which in first approximation will mimic the equidistribution of the digits. Formally this means that if the elements x_n are drawn from the interval $[0, Q)$, then

$$\lim_{N\to\infty} \frac{Q}{N} \sum_{\substack{a\leqslant x_n<b\\ 1\leqslant n\leqslant N}} 1 = b - a, \tag{2.1}$$

where $0 \leqslant a < b \leqslant Q$ [4]. The simplest illustration of equidistribution is the approximate equality of the number of heads and tails obtained in consecutive tosses of an "ideal" coin. Analogous behavior is expected for other random sequences of dichotomic variables [14].

(ii) *Normality.* Runs of consecutive heads and tails generated by tossing an ideal coin are not only equidistributed but presumably occur in definite proportions, i.e., about half the runs are expected to have length 1, one-fourth have length 2, one-eight have length 3, and so forth. A mathematical realization of this sort of apportionment is a *normal number*, that is a number in whose decimal expansion all blocks of digits of the same length occur with equal frequency. In particular if $x = 0.x_1x_2x_3 \cdots$, represents an infinite decimal expressed to base r, X_n denotes the block of digits $x_1x_2 \cdots x_n$, B_k a given block $b_1b_2 \cdots b_k$, and $N(B_k, X_n)$ designates the number of occurrences of the block B_k in X_n, then the condition for the normality of x to the base r is [5]

$$\lim_{n\to\infty} n^{-1}N(B_k, X_n) = r^{-k}, \qquad \text{for all } B_k, \quad k = 1, 2,\ldots. \tag{2.2}$$

This constraint of "proportional representation" can also be expressed in terms of equipartition criteria [4]. A theorem of Borel [5] asserts that almost all real numbers (in the sense of Lebesgue measure) are normal to every base, but in practise it is exceedingly difficult to determine whether or not any particular irrational number is normal. For example it is not known whether e or π are normal to *any* base. Furthermore a number which is known to be normal may not have its digits arranged randomly. Champernowne's number [5]

$$0.12345678910111213.... \tag{2.3}$$

is normal, transcendental, and obviously non-random.

(iii) *Auto-correlation.* Patterns of order such as those exhibited by (2.3) can often be detected by comparing the sequence $\{x_n\}$ and its shifts $\{x_{n+\tau}\}$. A formal measure of the degree of order in the sequence then is given by the autocorrelation function

$$C(\tau) = \lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} x_n x_{n+\tau}, \qquad \tau = 0, 1, 2,...; \tag{2.4a}$$

provided that the limit exists. In the case of random or "white noise" sequences it is intuitively obvious that the correlation function ought to approximate the behavior of a Kronecker delta, e.g.,

$$C(\tau) = \text{const.} \times \delta_{0\tau}, \qquad \text{where} \quad \delta_{0\tau} = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0. \end{cases} \tag{2.4b}$$

Indeed for some mathematically "white" sequences it is possible to derive the Kronecker form (2.4b) by rigorous means [4].

(iv) *Ergodicity.* The essence of the ergodic hypothesis regarding the "random" behavior of general systems is that there is a direct proportionality between the time that a system spends in a certain region of phase space and the volume of this region. Clearly the ergodic hypothesis is a physical counterpart of the normality condition (ii), cf. [24]. If we identify "time spent" with relative frequency of occurrence and "volume" with a measure of content such as probability measure then ergodicity can be linked to the computer simulation of random processes by means of the ergodic theorem. It is convenient to recall first that ergodicity and metric transitivity are connected by the following definitions:

(iv-a) *metric transitivity and ergodicity.* Let P be a probability measure defined on a σ-field of subsets of a set Ω. Let f be a (not necessarily invertible) function, which maps Ω into itself. Then f is ergodic or metrically transitive on Ω with respect to P if for any P-measurable subset S of Ω the condition $f^{-1}S = S$ implies that either $P(S) = 0$ or $P(S) = 1$ [24, 25]. Here f^{-1} is the inverse image of S under f, i.e., the set of all points x such that $f(x)$ is in S, and f is measure preserving $P(f^{-1}S) = P(S)$.

The practical embodiment of these abstractions is furnished by the

(iv-b) *ergodic theorem* [19, 24, 25, 26]. If f is ergodic on Ω with respect to P, then there is a subset Ω' of Ω with $P(\Omega') = 1$, such that

$$\lim_{n\to\infty} n^{-1} \sum_{m=0}^{n-1} \chi_S(f^m[x]) = P(S) \tag{2.5a}$$

for all $S \in \Omega$ and all $x \in \Omega'$. As usual χ_S denotes the characteristic function

$$\begin{aligned} \chi_S(\zeta) &= 0, \qquad \zeta \in S; \\ &= 1, \qquad \zeta \in S. \end{aligned} \tag{2.5b}$$

Obviously the left hand side of (2.5a) represents the average number of images of x, i.e., the iterates $f^m[x]$, which fall in the set S, and the theorem asserts that the relative frequency of this occurrence is in fact proportional to the measure of S. Finally the law of large numbers combined with (2.5a) leads to the interpretation that $P(S)$ is the probability that a point *selected at random* lies in S [23, 26, 27].

(v) *Mixing.* Metric transitivity implies that the successive iterations of f applied to $x \in \Omega$ correspond to a stirring or randomization of the elements of Ω. This interpretation can be made explicit by noting that the object $f^{-m}S_1 \cap S_2$—constructed for any two P-measurable subsets S_1 and S_2 of Ω—represents the miscibility of the two subsets after m stirrings, i.e., iterations. Indeed if the mapping f is ergodic then (2.5a) can be generalized in the form

$$\lim_{n\to\infty} n^{-1} \sum_{m=0}^{n-1} P(f^{-m}S_1 \cap S_2) = P(S_1)\, P(S_2), \tag{2.6}$$

and this guarantees that *on the average* the sets S_1 and S_2 are completely mingled by the iteration of the mapping f [24]. Of course theorem (2.6) does not preclude the sporadic recurrence of patterns of order: For instance in communications codes it is possible to conceal bursts of "signal" in noise, and there are physical systems such as electron spins which can show a recrudescence of order in virtue of spin "echo." See also [65].

It is possible to suppress these kinds of departure from random behavior by replacing the average mingling implied by ergodicity (2.6) by the much stronger asymptotic mixing condition [28]

$$\lim_{n\to\infty} P(f^{-n}S_1 \cap S_2) = P(S_1)\, P(S_2). \tag{2.7}$$

A function f satisfying (2.7) is said to be mixing on the set Ω with respect to the measure P. Clearly if f is mixing it is also ergodic, cf. (4.4) *et seq.* One rigorous consequence of (2.7) is that mixing will spread initially small sets throughout Ω [11]. This randomization, which is also amenable to computer simulation, is irreversible in the sense that with probability 1 the diffused sets will never contract.

(vi) *Cross-correlation.* When several pseudo-random number generators are run in parallel in a computer program it is possible that latent stochastic dependences among the generators may introduce spurious features. The simplest means for detecting such dependences is the covariance or cross-correlation: Specifically if X_1 and X_2 are random variables with expectation values

$$\langle X_i \rangle = \iint dx_1\, dx_2\, x_i g(x_1, x_2), \qquad i = 1, 2; \tag{2.8a}$$

where g is the normalized joint probability density, then the cross-correlation between X_1 and X_2 is given by [27]

$$\mathscr{C}(X_1, X_2) = \iint dx_1\, dx_2\, (x_1 - \langle X_1 \rangle)(x_2 - \langle X_2 \rangle)\, g(x_1, x_2). \tag{2.8b}$$

In case X_1 and X_2 are independent the probability density may be factored into one dimensional distributions, e.g.,

$$g(x_1, x_2) = h_1(x_1)\, h_2(x_2),$$

and the cross-correlation vanishes. However, the converse may fail and this is the essential hazard in identifying stochastic dependences. One counter-example is furnished by the pseudo-random vector (X_1, X_2) distributed uniformly on the circumference of the unit circle: in this case $\mathscr{C}(X_1, X_2) = 0$ despite the fact that X_1 and X_2 are not independent [27]. When regression criteria such as (2.8b) fail, more powerful tests utilizing copulas are available for detecting the presence of stochastic dependences [29–31]; cf. Section 3.4.

(vii) *Statistical Stability.* If $\{x_n\}$ is a random sequence it is possible to introduce patterns of order by modifying the individual terms by additions or deletions. However, if these perturbations are sufficiently "weak" then it is intuitively plausible that the random character of the sequence should be unaffected. Under suitable restrictions, stability results of this kind can be established rigorously [32]. In the particular case of mixing transformations the meaning of "weak perturbation" can be made more precise by recalling the analogy with continuously stirred reaction vessels of the type used in chemical engineering applications: In these devices the rate at which the constituents are mixed is adjusted to exceed the rate at which at which new constituents are added; the homogenization of the effluent is thereby continuously maintained. In computer simulations of mixing, the homogenization or rate of randomization is scaled by the Kolmogorov entropy, and the inflow of new constituents corresponds to the perturbations which are introduced by computer noise. Statistical stability in this instance has the practical consequence that the random character of the Čebyšev mixing simulations can actually be sustained on sufficiently large machines, cf. Section 5.

3. Computer Simulation of Čebyšev Mixing

3.1 *Single Sequences of Pseudo-Random Numbers*

The second degree Čebyšev polynomial

$$f(x) = x^2 - 2 \tag{3.1}$$

maps the interval $[-2, 2]$ onto itself. Simple iteration leads to the quartic mapping

$$f[f(x)] = f^2(x) = x^4 - 4x + 2. \tag{3.2}$$

The nth order iterate of f can also be expressed in the closed form

$$f^n(x) = 2\cos[2^n \cos^{-1}(x/2)], \qquad n \geqslant 1. \tag{3.3}$$

According to the basic theorem of Adler and Rivlin, cf. (4.1a), the sequence of iterates $\{f(x)\}$ is mixing for almost all choices of the initial "seed" $x \in [-2, 2]$. Since mixing implies ergodicity both of the randomness criteria (iv) and (v) of Section 2 are satisfied. Furthermore it is easy to verify that the orthogonality of the Čebyšev polynomials implies that the autocorrelation function for $\{f^n(x)\}$ has the Kronecker form (2.4b) associated with white noise. It will shortly appear that the sequences $\{f^n(x)\}$ also satisfy the other randomness criteria discussed in Section 2.

For practical purposes it is sometimes convenient to shift the mixing region from $[-2, 2]$ to the general interval $[a, b]$ where $a < b$. In this case (3.1) is replaced by

$$\tilde{f}(\tilde{x}) = a + (2\tilde{x} - a - b)^2/(b - a), \qquad \tilde{x} \in [a, b]. \tag{3.4}$$

However, the mixing behavior does not occur for arbitrary quadratics. The possibilities are severely constrained by the following result announced in [34].

Lemma. *No quadratic function that maps the unit interval onto itself and has a minimum can be measure preserving,* a fortiori *mixing, with respect to any absolutely continuous measure if its minimum point falls outside the closed interval* [0.5, 0.647798...]. *A similar statement holds for quadratic functions that have maxima. The function whose minimum occurs at* 0.5 *is the Čebyšev polynomial*

$$\tilde{f}(\tilde{x}) = (2\tilde{x} - 1)^2, \qquad \tilde{x} \in [0, 1]. \tag{3.5}$$

The behavior of the Čebyšev mixing sequence $\{f^n(x)\}$ can be illustrated by considering the first few terms which arise from the initial "seed" $\pi - 3 \simeq 0.1416$, viz.,

$$0.1416,\ -1.9799,\ 1.9202,\ 1.6872,\ 0.8466,\ -1.2832,\ -0.3534. \tag{3.6}$$

These numbers seem to imitate an erratic wandering. Computer experiments show that such wandering sequences are obtained for practically all choices of the initial seed $x \in [-2, 2]$; compare Table II(a). The theoretical safeguard preventing collapse into fixed points or cyclic orbits is provided by the following

THEOREM [12]. *Let n be an integer $n \geqslant 1$. Then the set of cyclic points*

$$f^n(x) = x, \qquad x \in [-2, 2] \tag{3.7a}$$

of the Čebyšev polynomials is dense but countably infinite; consequently of measure zero. Specifically if

$$f^n(x_m) = x_m\,, \qquad m = 0, \ldots, 2^n - 1; \qquad n \geqslant 2; \tag{3.7b}$$

then the fixed points are given by

$$x_m = \begin{cases} 2 \cos\left[\dfrac{2m\pi}{2^n - 1}\right], & m \text{ even}; \\ 2 \cos\left[\dfrac{2m\pi}{2^n + 1}\right], & m \text{ odd}. \end{cases} \tag{3.7c}$$

Practical experience indicates that the computer iterations will collapse into the repulsive fixed point $+2$ at step n if at step $n - 2$ the sequence wanders sufficiently close to zero; this will occur if $|f^{n-2}(x)| \leqslant \delta$, and δ^2 vanishes within the accuracy of the machine arithmetic. The probability for such a collapse can be inferred from the distribution of the values of $\{f^n(x)\}$. Specifically since the Čebyšev mixing is invariant with respect to the measure, cf. (4.1a),

$$P_c(S) = \frac{1}{\pi} \int_S \frac{dx}{(4 - x^2)^{1/2}}\,, \qquad S \subset [-2, 2]; \tag{3.8}$$

it is clear that after $\mathscr{N}$ iterations the inequality

$$-2 \leqslant x_L \leqslant f^n(x) \leqslant x_U \leqslant 2, \qquad n \leqslant \mathscr{N}, \tag{3.9a}$$

will on the average be satisfied I times where

$$I(x_L\,, x_U\,; \mathscr{N}) = \frac{\mathscr{N}}{\pi} \{\sin^{-1}(x_U/2) - \sin^{-1}(x_L/2)\}. \tag{3.9b}$$

In other words the sequence $\{f^n(x)\}$ is likely to fall into the interval $[-\delta, \delta]$ at least once when

$$I(-\delta, \delta; \mathscr{N}) \simeq \mathscr{N}\delta/\pi \geqslant 1; \qquad \delta \ll 1. \tag{3.10}$$

If the simulations are performed on a machine that effectively operates with q digits to the base 10 then the accessible universe of numbers comprised in [-2, 2] consists of N_U distinct elements where

$$N_U = 4 \times 10^{q-1} + 1. \tag{3.11a}$$

The condition that δ^2 vanishes within the accuracy of the machine arithmetic then is equivalent to the bound

$$\delta \lesssim 10^{-(q-1)/2}. \tag{3.11b}$$

Finally if we anticipate the estimate that the average number of iterations is scaled by $\mathcal{N} \gtrsim N_U^{1/2}$, cf. (3.21c), it becomes apparent that the "+2 collapse" criterion (3.10) is marginally violated, i.e.,

$$I(-\delta, \delta; \mathcal{N}) \gtrsim N_U^{1/2}\delta/\pi \sim 2/\pi \not\gg 1. \tag{3.11c}$$

Practical experience confirms that when $q \gtrsim 10$ the mixing simulations hardly ever terminate on the $+2$ fixed point, cf. Tables II(a), II(b) and Table VI.

The distribution function (3.9b) furnishes a good statistical check on whether the computer-generated values of $\{f^n(x)\}$ conform to the theoretical Čebyšev measure (3.8). One drawback for practical applications is that this distribution is not flat. In principle it is of course easy to adjust the shape: for instance the sequence $(4/\pi)$ $\{\cos^{-1}[f^n(z)/2]\} - 2$ where $z = 2\cos[(\pi/4)(x + 2)]$, is uniformly distributed over $[-2, 2]$ and satisfies the equidistribution criterion (i) of Section 2; cf. [35]. A cruder flattening that is more economical in programming time can be achieved by eliding the first few digits of each term in the sequence $\{f^n(x)\}$: this corresponds to using a linear approximation for $\cos^{-1}(z + \zeta)$ when $|\zeta| \ll |z \neq 1$.

Another statistical index which can easily be derived from the underlying Čebyšev measure (3.8) is the *average agitation* associated with the mixing. It is plausible to identify this as the average fluctuation of the successive terms of $\{f^n)\}$, i.e.,

$$\mathscr{A} \equiv \lim_{n\to\infty} \langle | f^{n+1}(x) - f^n(x)| \rangle_{x\in[-2,2]}, \tag{3.12a}$$

where the notation indicates that the average is taken over $x \in [-2, 2]$. By specializing (4.5) this limit can be derived from a simple quadrature,

$$\mathscr{A} = \int_{-2}^{+2} dz \frac{| z^2 - z - 2 |}{\pi(4 - z^2)^{1/2}} = \frac{3^{3/2}}{\pi} \simeq 1.653\,986 \ldots . \tag{3.12b}$$

The agitation for the 7 terms displayed in (3.6) happens to be 1.69; longer simulation runs of $\{f^n(x)\}$ lead to the exact result (3.12b) within the latitude of statistical error; cf. Table VIII.

3.2 Computer Simulations: Growth of Round-Off Errors

The individual terms of the Čebyšev mixing sequence $\{f^n(x)\}$ can in principle be computed by starting with an initial "seed" x_0 and iterating the simple quadratic

$$f^{n+1}(x_0) \equiv x_{n+1} = x_n^2 - 2. \tag{3.13}$$

If we use a binary digital machine which carries out arithmetical operations exactly [illegible]

$$\bar{x}_0 = \operatorname{sgn}(x_0)\,\frac{1}{2^b}\,[2^b \mid x_0 \mid], \tag{3.14a}$$

and the iteration (3.13) can be approximated by the operations

$$\bar{x}_{n+1} = \frac{1}{2^b}\,[2^b \bar{x}_n^2] - 2. \tag{3.14b}$$

As usual in these expressions $[z]$ denotes the largest integer contained in the (non-negative) number z, and *sgn*(z) is the sign of z. Of course the precise realization of (3.14a, b) depends upon the nature of the microprograms of the computer as well as the way in which the Čebyšev iteration is programmed. In any event the distinction between (3.13) and (3.14b) will lead to a progressive divergence between the exact and computed values of $\{f^n(x)\}$. Let us first introduce the increments $\Delta x_n \sim \mathcal{O}(2^{-b})$ which correspond to the round-off when (3.14b) is rewritten in the form

$$\bar{x}_{n+1} = \bar{x}_n^2 - 2 + \Delta x_{n+1}\,. \tag{3.15}$$

Then the divergence

$$D_n \equiv x_n - \bar{x}_n \tag{3.16a}$$

will evolve according to the difference relation

$$D_{n+1} = (x_n^2 - \bar{x}_n^2) - \Delta x_{n+1} = 2x_n D_n - (D_n^2 + \Delta x_{n+1}). \tag{3.16b}$$

In view of the initial condition $D_0 = x_0 - \bar{x}_0 \sim 2^{-b} \ll 1$, it is easy to check that (3.16b) yields numbers D_n that tend to increase exponentially with n. After a few iterations the stepwise truncation error can be neglected relative to the cumulative truncations. The magnitude of the average deviation can then be estimated from

$$\begin{aligned} \langle \mid D_{n+1} \mid \rangle &\cong \langle \mid 2x_n D_n - D_n^2 \mid \rangle, \\ &\sim 2\langle \mid x_n \mid \rangle \langle \mid D_n \mid \rangle \qquad \text{for} \quad \langle \mid D_n \mid \rangle \ll \langle \mid x_n \mid \rangle = 4/\pi; \end{aligned} \tag{3.16c}$$

and this leads to the simple proportionality

$$\langle | D_{n+1} | \rangle \sim \frac{8}{\pi} \langle | D_n | \rangle, \tag{3.16d}$$

where the averages are carried out over $x \in [-2, 2]$. Eventually the growth of the deviation is limited by the bound $| D_n | \leqslant 4$. The number of machine iterations required to amplify the deviations from the "noise" level $D_0 \sim 2^{-b}$ to $D_{n_d} \sim 1$ is then given by

$$n_d \sim b(3 - \log_2 \pi)^{-1} \sim 0.74\, b. \tag{3.16e}$$

Clearly this index corresponds to the threshold beyond which the computer generated iterates $\{\bar{x}_n\}$ and the exact Čebyšev mixing sequences $\{x_n\}$ lose all numerical resemblance. These features have been confirmed with a variety of machine experiments.

The practical relevance of the computer iterations then hinges on the basic experimental observation that all the statistical properties of the exact Čebyšev mixing sequences $\{x_n\}$ are simulated faithfully by the machine generated sequences $\{\bar{x}_n\}$ even for values of n greatly in excess of the divergence threshold n_d (3.16*e*)! A concise illustration of the agreement between the observed and expected values of the frequency distribution I (3.9b) for a typical run with $\mathcal{N} = 10^5$ and $n_d \sim 26$ is given on Table I. If there is a good correlation between the experimental frequency distribution of the $\{\bar{x}_n\}$ and the underlying Čebyšev measure (3.8) it is reasonable to anticipate that all the other statistical indices, such as the agitation (3.12a) and the auto-correlation (2.4a), will simulate the behavior of the exact mixing sequences $\{f^n(x)\}$. These points have also been verified with computer experiments.

TABLE I

Experimental and Theoretical Frequency Distributions for Čebyšev Mixing Transformations[a]

Interval	−2.0 −1.5	−1.5 −1.0	−1.0 −0.5	−0.5 0	0 0.5	0.5 1.0	1.0 1.5	1.5 2.0
I^b (theory)	23 005	10 328	8623	8043	8043	8623	10 328	23 005
I (exp.)	22 912	10 249	8510	8007	8128	8515	10 288	23 391

[a] Starting point: $x_0 = \pi$-3; arithmetic: 15 decimal digits; computer: IBM 360/195; number of iterations: 10^5.

[b] See (3.9b).

We have already indicated in Section 2(vii) that the statistical correspondence between the exact sequences $\{x_n\}$ and the machine generated iterates $\{\bar{x}_n\}$ for $n > n_d$ is a manifestation of the stability of the mixing randomization. This stability can be destroyed by inserting large and structured increments y_n between successive iterations of (3.14b): obviously the modified sequences

$$\{\bar{x}_n\} \rightarrow \{\bar{x}_n + y_n\}$$

where $|y| \sim |\tilde{x}_n|$ need not be pseudo-random. In practice it is easy to perturb the mixing by varying the precision of the computer arithmetic. This leads to the practical constraint that machines operating with less than 25 bit or 8 decimal place arithmetic tend to throttle the randomization. In Section 5 we will discuss this problem in further detail and show that larger capacity devices can successfully simulate the Čebyšev mixing at levels close to the theoretical optimum.

3.3 *Computer Simulations: Free Running and Terminal Cycles*

The quantitative divergence between the computer generated iterates $\{\bar{x}_n\}$ and the exact Čebyšev mixing sequences $\{x_n\}$ not only reflects the cumulative truncation errors of the machine arithmetic but on a deeper level is connected with the asymptotic cycling of iterative processes on finite state devices. Specifically for any mapping $\mathscr{M}(x_i) \to x_j$ of a finite set $x_1, x_2, \ldots, x_N$ onto itself the sequences of iterated mappings

$$\mathscr{M}(x_i), \mathscr{M}^2(x_i), \ldots, \mathscr{M}^{n_f}(x_i), \ldots, \mathscr{M}^{n_L}(x_i), \ldots, \mathscr{M}^m(x_i) \tag{3.17}$$

must contain at least two identical elements, say $\mathscr{M}^{n_f}(x_i)$ and $\mathscr{M}^{n_L}(x_i)$, whenever $m \geqslant N$, irrespective of the choice of the initial "seed" x_i (pigeonhole principle [5]). This implies that (3.17) is actually comprised of a "free-running" subsequence $\mathscr{M}(x_i), \ldots, \mathscr{M}^{n_f-1}(x_i)$ containing no repetitions, and a contiguous subsequence $\mathscr{M}^{n_f}(x_i), \ldots, \mathscr{M}^{n_L}(x_i)$ which constitutes a terminal loop with $n_L - n_f$ distinct elements. In cases where the sequence of iterates $\{\mathscr{M}^m(x_i)\}$ has a random character it is possible to obtain quantitative estimates for the magnitudes of the free running index n_f and the loop index n_L by means of simple combinatorial arguments. The essential idea is to consider that the sequence (3.17) is generated by random selections with replacement from a sample of N_U distinct elements, cf. (3.11a): We must then determine how many choices have to be made on the average before at least one element is encountered twice. This is a simple variant of the "birthday" problem [27].

If we index the selections by i, and indicate the probability for a repetition by $P(i, N_U)$, then obviously

$$P(1, N_U) = 0; \tag{3.18a}$$

and the non-void cases can be built up according to the pattern

$$P(2, N_U) = N_U^{-1}, \tag{3.18b}$$

and

$$P(3, N_U) = P(2, N_U) + \frac{2}{N_U}[1 - P(2, N_U)]. \tag{3.18c}$$

One can easily check that the general recursion is given by

$$P(i + 1, N_U) = P(i, N_U) + \frac{i}{N_U}[1 - P(i, N_U)]; \tag{3.19a}$$

and consequently

$$1 - P(i+1, N_U) = \prod_{j=1}^{i} [1 - (j/N_U)] = \frac{\Gamma(N_U)}{N_U{}^i \Gamma(N_U - i)}. \tag{3.19b}$$

The object on the left hand side of (3.19b) is the probability that after $i + 1$ selections one has not encountered a repetition. As usual it is convenient to parameterize this expression by an exponential, i.e.,

$$e^{-\lambda} \equiv 1 - P(i+1, N_U), \tag{3.20}$$

and then to evaluate the probabilities with the help of the auxiliary simplifications $N_U \gg i$ and $N_U \gg 1$, which are satisfied in practise. Applying Stirling's approximation to the gamma functions in (3.19b) we then obtain

$$\lambda \simeq i + (i - N_U)\ln[N_U/(N_U - i)], \tag{3.21a}$$

and to leading order

$$\lambda \simeq \frac{i^2}{N_U} + \mathcal{O}(i^3/N_U{}^2). \tag{3.21b}$$

So for example at the 50% level, i.e., $e^{-\lambda} \sim 1/2$, a pseudo-random (albeit *deterministic*) sequence of the type indicated in (3.17) will merge into a terminal loop whenever the index n_L reaches the range

$$n_L \sim 0.83 N_U^{1/2}. \tag{3.21c}$$

Since the Čebyšev iterates are not uniformly distributed over the accessible numbers in the interval $[-2, 2]$, cf. (3.9) and (3.11a), the coefficient in (3.21c) is actually somewhat too large: these features are in agreement with the trends of the experimental results, cf. Tables II(a) and II(b).

A slight extension of these arguments also leads to an order of magnitude estimate for the free running index n_f. In this case we consider a universe of N_U distinct elements which contains a distinguishable subset of $n_L - n_f$ objects—free running then obviously corresponds to sampling N_U at random until one encounters one of the $n_L - n_f$ objects lying on the terminal loop. Since the probability of selecting a member of the loop is approximately $(n_L - n_f)/N_U$, the number of iterations required to reach the loop is of the order of the reciprocal $N_U/(n_L - n_f)$. If we consider the option that the loop length does not exceed the interval of free running, i.e., $n_L - n_f \leqslant n_f$, then evidently

$$n_f \sim \frac{N_U}{n_L - n_f} \sim N_U^{1/2}; \tag{3.22}$$

and this estimate also turns out to be in good correspondence with all the experimental results, cf. Table II(b). One can easily check that the assumption that the loop length greatly exceeds the interval of free running, i.e., $n_L - n_f \gg n_f$, is inconsistent with the prior estimate (3.21c). Similarly one can show that the existence of terminal loops with very few elements, $n_L - n_f \sim o(N_U^{1/2})$, is highly improbable: indeed no trace of such loops has been found in many trials involving machines with capacities satisfying the bound $N_U \lesssim 10^8$, cf. Section 6.

It is interesting to follow these trends on a large capacity machine, e.g., a Univac 1108 operating in double precision: In this instance 60 bits are available and therefore the exact Čebyšev mixing sequences $\{x_n\}$, cf. (3.13), and the machine generated

TABLE II(a)

Computer Simulations of Čebyšev Mixing on an *HP*-9100: 12 Significant Decimal Numbers[a]

	Initial number	Approximate interval of free running n_f (3.22)	Length of terminal loop $n_L - n_f$
1.	$\pi - 3$	343×10^3	95,447
2.	$\pi/2$	361×10^3	95,447
3.	$e - 1$	318×10^3	95,447
4.	$\pi^5 - 306$	277×10^3	95,447
5.	0.842 268 953 196	204×10^3	95,447
6.	1.280 799 970 80	206×10^3	95,447
7.	1.180 505 431 39	296×10^3	95,447
8.	1.255 073 742 11	209×10^3	95,447
9.	0.449 352 494 752	276×10^3	95,447
10.	0.054 997 654 6405	155×10^3	95,447
11.	1.956 541 430 01	232×10^3	95,447
12.	1.863 332 537 98	180×10^3	95,447
13.	1.166 449 883 52	334×10^3	95,447
14.	0.918 820 097 328	132×10^3	95,447
15.	0.835 869 910 785	309×10^3	95,447
16.	1.950 237 174 69	178×10^3	95,447
17.	0.763 877 929 030	387×10^3	95,447
18.	0.691 345 197 426	300×10^3	95,447
19.	1.644 424 334 36	186×10^3	95,447
20.	0.399 110 461 937	189×10^3	95,447
21.	1.754 673 704 92	333×10^3	95, 447
22.	$\pi^2 - 9$	299×10^3	104,694
23.	1.009 732 533 76	82×10^3	104,694
24.	0.989 320 505 142	94×10^3	104,694
25.	1.547 445 266 95	121×10^3	104,694
26.	1.085 062 746 99	147×10^3	104,694
27.	1.288 397 343 65	120×10^3	104,694
28.	1.952 359 515 65	77×10^3	104,694
29.	0.378 897 599 758	210×10^3	104,694
30.	$(\sqrt{2} - 1)^2$	72×10^3	39,965
31.	1.045 816 019 14	94×10^3	39,965
32.	1.555 555 555 55	13,597	+2 (fixed point)

[a] Approximate decimal equivalent including guard digit and discounting roundoff errors: $N_U \simeq 4 \times 10^{11}$, cf. (3.11a).

TABLE II(b)

Statistical Summary of Data in Table II(a): Comparison With Combinatorial Estimates

Terminal loop $n_L - n_f$	Percentage of initial numbers merging into terminal loop	Average interval of free running $\langle n_f \rangle$	$\langle n_f \rangle + n_L - n_f$ $\approx \langle n_L \rangle$	Theoretical estimate of $\langle n_L \rangle$ (3.21c)	$\frac{n_L - n_f}{\langle n_f \rangle}$
95,447	66 %	257,000	352,000	520,000	0.37
104,694	25 %	144,000	249,000	520,000	0.73
39,965	6 %	83,000	123,000	520,000	0.48
+2 (fixed point)	3 %	13,597	—	cf. (3.11c)	—

sequences $\{\bar{x}_n\}$, cf. (3.14b), should agree quantitatively until the number of iterations approaches the divergence index $n_d \sim 50$, cf. (3.16e); in practical trials it was in fact observed that the discrepancy D_n (3.16a) was amplified to the order of unity after 45–50 iterations. Beyond this point the machine generated sequences $\{\bar{x}_n\}$ continue to simulate all the statistical properties of the exact Čebyšev mixing—for instance the frequency distribution (3.9b), the Kronecker auto-correlation (2.4b), and the agitation (3.12b). In the interval $0 \leqslant n \leqslant n_f \sim 10^9$, cf. (3.11a) and (3.22), all elements of the sequence $\{\bar{x}_n\}$ should be distinct: for example with $\pi - 3$ as the initial seed, the interval of free running was found to be approximately 2.8×10^8; the corresponding terminal loop $\{\bar{x}_n\}$, $n_f \leqslant n \leqslant n_L$, had a length $n_L - n_f = 48, 424, 947$. This result is in fair accord with the estimate $n_L - n_f \gtrsim \mathcal{O}(N_U^{1/2})$ implied by (3.22). Tracing the behavior of the $\pi - 3$ iterations on the 1108 consumed nearly 1 hour of CPU time; most of the other numerical experiments were therefore carried out under less expensive conditions. Some representative results are summarized in Tables II(a) and II(b).

As indicated in the first column of Table II(a) we found it convenient to use a number of standard "seeds" such as $\pi - 3$, $e - 1$, and 1.555 555 ... for running the Čebyšev mixing simulations on a variety of computers; the other 25 twelve digit seeds were compiled from a table of random numbers. All runs were carried out at least twice to eliminate power line or internal "glitches," and precautions against non-commutative operations with guard digits were incorporated in the programs [36]. The terminal loops were identified with the help of the *a priori* estimate $n_L \sim N_U^{1/2}$ (3.21c): We simply allowed the computer iterations to run past a larger index $n_L^* \gtrsim 2N_U^{1/2}$, and then programmed the machines to look for a repeating sequence

$$x_{n_L^*+i} = x_{n_L^*}, \qquad x_{n_L^*+i+1} = x_{n_L^*+1}, \ldots, x_{n_L^*+i+5} = x_{n_L^*+5}. \tag{3.23}$$

The minimum value of i obviously correspond to the loop length $n_L - n_f$, and the five-fold redundancy implied by (3.23) insures that the recognition of the loop is "glitch" proof. Of course this simple recipe is not adequate to pinpoint precisely where the iterations originating from a particular seed x_0 first enter the terminal loops. However, once the loop has been identified the first point of entry $x_{n_f(x_0)}$ and the corresponding index $n_f(x_0)$ can be located by a variety of iterative methods. The imprecision in the values of n_f listed in column 2 of Table II(a) reflects the limits to which we pushed these methods. In any event it is clear that the observed and estimated values of the loop index n_L are in good agreement. As indicated previously the estimate (3.21c) is actually somewhat too large because of the non-uniform distribution of the Čebyšev iterates. The tabulations also confirm that the loop lengths tend to be smaller than the intervals of free running, i.e., $(n_L - n_f)/n_f < 1$. Finally it is apparent that the interval of free running is roughly scaled by $N_U^{1/2}$ as indicated by (3.22). Further results from machine trials with Čebyšev simulations are summarized in Tables V–VII of Section 5 (Variable Precision Routines), and Table VIII of Section 6 (Terminal Cycles).

Although detailed comparisons of the relative merits of standard multiplicative congruence methods and Čebyšev iterations as pseudo-random number generators are beyond the scope of the present work it is interesting to spot check the trends. For instance one procedure recommended for the SR-52 programmable calculator (10 digits + 3 guard digits) is to multiply an arbitrary initial seed by 7^9, elide the first eight digits of the product, and use the remaining five digits as the seed for the next multiplication by 7^9: the sequence of five digit numbers generated by these means is supposed to be pseudo-random. However as indicated on Table III this process terminates on three short loops ($n_L - n_f = 421$; 156; 77) after very brief intervals of free running. Specifically with $\pi - 3$ as the initial seed the "7^9" congruence method leads to a loop of length 77 after approximately 76 iterations, whereas on the *same* machine the Čebyšev mixing simulations merge into a loop of length 564,609 ($N_U^{1/2} \sim 2 \times 10^6$) after $\lesssim 10^6$ iterations. This striking disparity indicates that the practical performance of some standard random number generators may fall far short of the theoretical estimates [37].

3.4 *Multiple Sequences of Pseudo-Random Numbers*

If the Čebyšev iterations are started at two initial seeds, x_0 and y_0, the ensuing mixing sequences $\{x_n\}$ and $\{y_n\}$ behave like independent random variables for almost all choices of x_0 and y_0, cf. (4.5). As a consequence both the cross-correlation (2.8a) as well as the copula measures of dependence [30, 31] vanish identically. Since the Cartesian squares of ergodic transformations need not be ergodic [12, 25] it is clear that the transition from the weaker mingling condition (2.6) to mixing (2.7) is crucial for generating multiple sequences of pseudo-random numbers. In practical trials it turns out to be convenient to check on the suppression of stochastic dependences by monitoring the dispersive properties of the machine generated iterates. Specifically this implies that for almost all choices of x_0 and y_0 the mixing process disperses the values of the iterates $\{x_n\}$ and $\{y_n\}$ so thoroughly that eventually their relative distributions become *independent* of the starting points. For instance the probability density that the difference between two iterates $|\{x_n\} - \{y_n\}|_{n\to\infty}$ falls between the values z and $z + dz$ where $0 \leqslant z \leqslant 4$ is given by [12]

$$G(z) = \frac{8}{\pi^2(4 + z)} K\left(\frac{4 - z}{4 + z}\right), \tag{3.24}$$

TABLE III

Free Running and Terminal Loops for the "7^9" Random Number Generator: SR-52 (13 digits)

Initial seed	Length of terminal loop	Upper bound for interval of free running[a]
0.01	421	474
0.02	421	478
0.03	421	895
0.04	421	474
0.05	421	53
0.06	156	192
0.07	421	474
0.08	156	192
0.09	421	474
0.10	421	895
0.11	156	348
0.12	421	474
0.13	421	474
0.14	156	192
0.15	421	474
0.16	421	474
0.17	156	348
0.18	421	474
0.19	421	53
0.20	421	474
0.21	421	474
0.22	421	53
0.23	421	53
0.24	421	474
0.25	421	474
0.26	421	474
0.27	421	474
⋮	⋮	⋮
0.52	421	579
0.14159	77	76
0.15159	421	579
0.16159	156	220
0.17159	421	579
0.18159	421	579
0.15574	421	578
0.16989	77	307
0.18404	421	579
0.19819	77	153

[a] Number of iterations required for the initial seed x_0 to encounter the number x_{3000} for the *first* time. Average value for upper bound of free running ~408.

where K denotes a complete elliptic integral of the first kind, cf. (4.6). The mean distance or asymptotic dispersion of the Čebyšev mixing on the interval $\Omega = [-2, 2]$ can then be obtained by quadrature [38]:

$$D_c = \lim_{n\to\infty} \langle |\{x_n\} - \{y_n\}| \rangle_{x_0, y_0 \in \Omega} \tag{3.25a}$$

$$= \int_0^4 dz\, zG(z) = \frac{16}{\pi^2} \cong 1.621\,138\ldots. \tag{3.25b}$$

The corresponding dispersive behavior of the machine generated iterates may be checked by straightforward methods. For example the mean distance can be approximated by selecting λ-pairs of initial seeds $\bar{x}_0{}^i$, $\bar{y}_0{}^i$, $1 \leqslant i \leqslant \lambda$, cf. (3.14a), all equally spaced; $| \bar{x}_0{}^i - \bar{y}_0{}^i | = \Delta \leqslant 4$, and then computing the average distance

$$z(n) = \lambda^{-1} \sum_{i=1}^{\lambda} | \bar{x}_n{}^i - \bar{y}_n{}^i | \tag{3.26}$$

as a function of the number of iterations, cf. (3.14b). If the machine generated sequences properly simulate the behavior of the multiple Čebyšev iterates, then the statistical metric z should exhibit a transition from $z(0) = \Delta$ to $z(n) \to 16/\pi^2$ with increasing values of n. We have verified this behavior under a variety of conditions, e.g., $200 < \lambda < 2000$ and $10^{-4} < \Delta < 10^{-2}$, on a Univac 1108 operating in double precision [12, 38]. Some of the results are given on Table IV.

The ergodic nature of the mixing furnishes another simple test: Clearly after a sufficient number of iterations—roughly $n \gtrsim 20$ according to Table IV—the mixing should obliterate any initial pattern of order in the selection of the seed pairs $\bar{x}_0{}^i$, $\bar{y}_0{}^i$. In fact the frequency distribution of the iterates $\{\bar{x}_n{}^i\}$, $\{\bar{y}_n{}^i\}$ should approach the Čebyšev density (3.9b) for practically all choices of the initial distributions of $\bar{x}_0{}^i$ and $\bar{y}_0{}^i$. Numerical trials have also confirmed that this memory loss is correctly simulated by the machine sequences [38].

These statistical indices have a general utility in monitoring the behavior of multiple sequences of random numbers. The basic idea is that if Ω is any metric space endowed with a probability measure and $\mathscr{M}$ is a measure preserving transformation on Ω, then for any $z > 0$ and almost all pairs of points (x, y) in the Cartesian product $\Omega \times \Omega$, there exists a distribution function $F_{x,y}(z)$—the probabilistic metric—such that the fraction of times the distance between the points $\mathscr{M}^n(x)$ and $\mathscr{M}^n(y)$ is less than z convergences to $F_{x,y}(z)$ as $n \to \infty$ [39]. As a consequence arbitrary mixing transformations may be characterized by their asymptotic dispersion. An illustrative variant of our preceding results is the function,

$$\zeta_n = \frac{4}{\pi} \{\cos^{-1}[f^n(z)/2]\} - 2, \qquad z = 2\cos\left[\frac{\pi}{4}(x+2)\right] \tag{3.27}$$

TABLE IV

[illegible]on of the Statistical Distance $z(n)$ Towards its Asymptotic Value[a]

Number of iterations n	0	1	[illegible]	3	4	5	6	7	8	9						
$z(n)$ (3.26)	10^{-4}	2×10^{-4}	[illegible]10^{-4}	8.01×10^{-4}	1.61×10^{-3}	3.22×10^{-3}	6.50×10^{-3}	1.34×10^{-2}	2.73×10^{-2}	5.15×10^{-2}						
n	10	11	12	[illegible]	14	15	16	17	18	19	20	21	22	23	24	25
$z(n)$	0.103	0.201	0.393	[illegible]	1.34	2.04	2.13	1.32	1.51	1.71	1.80	1.62	1.67	1.58	1.68	1.58
n	26	27	28	29	[illegible]	31	32	33	34	35	36	37	38	39	40	100	$n \to \infty$
$z(n)$	1.62	1.60	1.61	1.6[illegible]	[illegible]	1.62	1.65	1.65	1.63	1.60	1.61	1.62	1.58	1.63	1.59	1.59	$\frac{16}{\pi^2} \sim 1.62$

[a] $\Delta = 10^{-4}$; $\lambda = 2000$; all entries [illegible]d from double precision calculations on an 1108 Univac.

which maps the interval $[-2, 2]$ onto itself, mixes with respect to Lebesgue measure; and leads to a uniform distribution of the iterates $\{\zeta_n\}$. In this instance the probabilistic metric and asymptotic dispersion corresponding to (3.24) and (3.25b) are given by [12]

$$\tilde{G}(z) = (4 - z)/8, \qquad 0 \leqslant z \leqslant 4; \tag{3.28a}$$

and

$$D_{CU} = \int_0^4 dz\, z\tilde{G}(z) = 4/3. \tag{3.28b}$$

The entries in Table IV describe the initial growth of the statistical distance $z(n)$ as well as its ultimate stabilization in the vicinity of $16/\pi^2$. Since the rate of growth is directly related to the dispersive nature of the mixing iterations and also reflects the underlying Kolmogorov entropy, cf. (4.12), it is useful to derive its evolution. We note first that the seed pairs x_0^i and y_0^i may be expressed in the symmetric form

$$x_0^i = \cos(\theta_i + \phi/2),\ y_0^i = \cos(\theta_i - \phi/2) \tag{3.29}$$

where the auxiliary constraint $| x_0^i - y_0^i | \ll 4$ requires that ϕ be a very small offset angle. The mean square of the statistical distance then is given by

$$\langle z^2(n)\rangle = \langle | x_n^i - y_n^i |^2\rangle_{\theta_i \in [0,\pi]}\,, \tag{3.30a}$$

or, substituting from (3.3) and (3.13),

$$\langle z^2(n)\rangle = \frac{4}{\pi}\int_0^\pi d\theta\, \{\cos[2^n(\theta + \phi/2)] - \cos[2^n(\theta - \phi/2)]\}^2. \tag{3.30b}$$

The dependence on the offset angle ϕ can be factored out of the integral, and consequently

$$\langle z^2(n)\rangle = 8\left(1 - \frac{1}{2^{n+1}\pi}\right)\sin^2[2^{n-1}\phi]. \tag{3.30c}$$

From this expression it is apparent that the initial rate of growth of the r.m.s. distance depends exponentially on n, viz.,

$$[\langle z^2(n)\rangle]^{1/2} \sim 2^{n+1/2}\phi, \qquad \text{when} \quad | 2^{n-1}\phi | \ll 1. \tag{3.30d}$$

A slight variant of this argument for the mth order Čebyšev polynomial ($m \geqslant 2$) leads to the estimate

$$[\langle z_m{}^2(n)\rangle]^{1/2} \sim m^{n+1/2}\phi,\ | m^{n+1/2}\phi | \ll 1. \tag{3.30e}$$

We will eventually link this result with the Kolmogorov entropy, cf. Section 4.

The essential implication of these estimates for the machine simulations of mixing is that the behavior of the computed values of the statistical distance $z(n)$ does in fact agree with the theoretical predictions (Table IV). However, if we recall that the growth of the computing errors is also scaled by an exponential dependence on n, cf. (3.16d)

$$\langle | D_n | \rangle \sim (8/\pi)^n \langle | D_0 | \rangle, \tag{3.31}$$

it becomes clear that there must be some practical constraints on the machine capacities so that the mixing simulations can be sustained. In particular the double precision routines employed for constructing Table IV require 50 iterations to obliterate the accuracy of the computations, cf. (3.16e), whereas the transition to a "steady-state" mixed regime requires only about 20 iterations. It is plausible that the randomization of the mixing simulations is statistically stabilized, by the mechanism indicated in Section 2(vii), precisely because of this disparity [70]. Of course with a finer mesh of initial seed pairs, i.e., $| \bar{x}_0{}^i - \bar{y}_0{}^i | = \Delta < 10^{-4}$, more iterations are required to reach the asymptotic mixed state $\langle z(\infty) > \sim 16/\pi^2$; and with a smaller computer the dominance of the roundoff errors sets in after fewer iterations. These opposing trends lead to a practical cross-over point where the dispersion of the mixing iterations tends to be throttled. A detailed discussion of mixing with variable precision routines is given in Section 5.

4. Čebyšev Mixing Theorems

All the numerical evidence we have obtained up to this point indicates that the $x^2 - 2$ computer iterations generate long "free running" strings of numbers that satisfy all the pseudo-randomness criteria listed in Section 2. We have also spot checked the iterative behavior of other surjective quadratic polynomials [34]: In all cases differing from (3.4) the computer iterations tend to collapse rapidly into cycles or fixed points. Clearly then there is some practical association between the mixing theory and the machine simulations despite the fact that the digital networks operate exclusively on sets of measure zero whereas it is precisely such sets that are omitted or ignored in the mathematical treatment of mixing. Similar results have been inferred from other numerical simulations of irregular physical systems—particularly non-linear coupled oscillator models in the vicinity of the stochastic transition, e.g., [40, 41]. In these cases the nature of the demarcation between ordered and disordered behavior has not yet been clarified theoretically [42] and extensive efforts have gone into computer experiments in order to gain an empirical feeling for the conditions leading to a threshold of randomization. The Čebyšev mixing problems at least have the conceptual advantage that most of the mathematical features concerning randomness have been worked out in detail.

All of the essential results are a consequence of the

Theorem (Adler and Rivlin [43]). *Let Ω be the interval* $[-2, 2]$, *and* $P_c(S)$ *the probability measure*

$$P_c(S) = \int_S \frac{dx}{(4 - x^2)^{1/2}}, \tag{4.1a}$$

for each measurable subset S of Ω. Let C_n be the modified Čebyšev polynomial defined on Ω by

$$C_n(x) = 2\cos[n \cos^{-1}(x/2)], \qquad (\textit{principal} \cos^{-1}). \tag{4.1b}$$

Then the sequence of iterates $\{C_n(x)\}$ is mixing on Ω with respect to the measure P_c, cf. Section 2(v).

The C_n's are closed under composition, with $C_m \circ C_n = C_{mn}$. Therefore if we specialize to the second order Čebyšev polynomial

$$f(x) = C_2(x) = x^2 - 2, \tag{4.2a}$$

then the nth iterate of f is given by $f^n = C_{2^n}$, or cf. (3.3)

$$f^n(x) = 2\cos[2^n \cos^{-1}(x/2)]; \tag{4.2b}$$

and the sequence $\{f^n(x)\}$ is strongly mixing. The iterative behavior of the general quadratic polynomial (on a suitably chosen interval)

$$ax^2 + (b+1)x + c, \qquad a \neq 0 \tag{4.3}$$

is completely equivalent to the "standard" Čebyšev form (4.2a) for all values of the coefficients satisfying the condition $b^2 - 4ac = 9$; cf. (3.4), (3.5), and reference [44].

Suppose that the inverse image $f^{-1} S$ of the set S is invariant under the mapping f; then

$$P_c(f^{-n}S \cap S) = P_c(S). \tag{4.4}$$

On the other hand the basic mixing condition (2.7) implies $P_c(S) = P_c(S)\,P_c(S)$, and this can only happen if $P_c(S)$ is 0 or 1. According to Section 2(iv-a) this coincides with metric transitivity and shows that the sequence $\{f^n(x)\}$ is also ergodic [38].

The essential link between mixing and the generation of multiple series of uncorrelated random numbers derives from the following

Lemma (Product Transformations [11, 12]). *If f is mixing with respect to the measure P on the set Ω, then $f^{(n)}$ is mixing with respect to $P^{(n)}$ on $\Omega^n = \Omega \times \Omega \times \cdots \times \Omega$ for each positive integer n, where $f^{(n)}$ denotes the n-fold Cartesian product $f \times f \times \cdots \times f$, and similarly $P^{(n)}$ is the n-fold Cartesian product of the measure. It follows that $f^{(n)}$ is also measure preserving and ergodic on Ω^n with respect to $P^{(n)}$.*

For $n = 2$ we therefore have

$$\lim_{n\to\infty} n^{-1} \sum_{m=0}^{n-1} \chi_{S_1}(f^m[x])\, \chi_{S_2}(f^m[y]) = P_c(S_1)\, P_c(S_2) \tag{4.5}$$

for almost all pairs (x, y) and all P_c-measurable subsets $S_1, S_2 \subseteq \Omega$; where as usual χ_S denotes the characteristic function (2.5b). This relation is a generalization of the ergodic theorem (2.5a) and permits us to assert that for almost all (x, y) the mixing sequences $\{f^m(x)\}$ and $\{f^m(y)\}$ behave like independent random variables—each distributed with the density $\pi^{-1}(4 - t^2)^{-1/2}$; cf. (4.1a) and (3.9b).

The asymptotic dispersion of the pairs (x, y) under successive iterations can also be described in terms of a probabilistic metric. Specifically if we let $D(z)$ be the set of points (x, y) in the square $[-2, 2] \times [-2, 2]$ satisfying the condition $|x - y| \leqslant z$, then it is possible to establish the existence of the limit [39]

$$\lim_{n\to\infty} n^{-1} \sum_{m=0}^{n-1} \chi_{D(z)}(f^m[x], f^m[y]) = F_c(z), \tag{4.6}$$

for almost all pairs (x, y). $F_c(z)$ is a distribution function whose value can be interpreted as the probability that the distance between two points chosen independently in $[-2, 2]$, but with each choice weighted by the measure P_c (4.1a), is $\leqslant z$. It is easy to check that the derivative of this probabilistic metric in the Čebyšev case coincides with the previous $G(z)$ (3.24).

Another simple consequence of the dispersion (4.6) is that the Čebyšev mixing is totally unstable. Let x and y be two initial points separated by an arbitrarily small distance, i.e., $|x - y| < \delta$. Then a stable evolution of the mixing would require that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f^n(x) - f^n(y)| < \epsilon, \qquad \text{for } \textit{all} \quad n \geqslant 1. \tag{4.7}$$

However, we have already seen that the asymptotic dispersion is $16/\pi^2$, cf. (3.25a) and (3.25b), and therefore the stability criterion (4.7) will be violated for $\epsilon < 16/\pi^2$.

A much stronger result is implied by the following

THEOREM [11, 45]. *Let (Ω, d) be a metric space and (Ω, P) a probability space. Suppose that the domain of the probability measure P includes all Borel sets of Ω, and $P(Q) > 0$ for every open ball Q in Ω. Then if T is a mixing transformation on (Ω, P) and S is a subset of Ω with positive measure*

$$\lim_{n\to\infty} d(T^n S) = d(\Omega), \tag{4.8}$$

where as usual the diameter of a set $\mathfrak{I}$ is defined by

$$d(\mathfrak{I}) \equiv \sup\{d(\omega_1, \omega_2) | \omega_1, \omega_2 \in \mathfrak{I}\}.$$

Loosely speaking this means that arbitrarily small sets of positive measure are dispersed throughout Ω by mixing processes: In this respect it is clear that Hopf's mathematical formulation of mixing [28] accomplishes precisely what Gibbs had originally intended [46].

The dispersive nature of the mixing is manifested both in the asymptotic "steady state" distribution described by the probabilistic metric $F_c(z)$ (4.6), as well as in the divergence of the individual orbits which appears at every iteration. These global and local properties are connected by the mean square statistical distance $\langle z^2(n)\rangle$ defined in eq. (3.30a). For instance averaging (3.30c) over all n yields

$$\lim_{n\to\infty} n^{-1} \sum_n \langle z^2(n)\rangle = 8\langle \sin^2(2^{n-1}\phi)\rangle_n = 4; \tag{4.9a}$$

and this value agrees with the mean square distance inferred directly from the probabilistic metric (3.24), i.e.,

$$\int_0^4 dz\, z^2 G(z) = 4. \tag{4.9b}$$

On the other hand the magnitude of the local dispersion (3.30d) is scaled by the initial mesh size because the offset angle ϕ determines the spacing of all the seed pairs $x_0{}^i$ and $y_0{}^i$ in (3.29). Nevertheless there are intrinsic ways of gauging the rate of dispersal [67]: One first eliminates the scale dependence by taking derivatives, cf. (3.30d)

$$\frac{\partial}{\partial\phi} [\langle z^2(n)\rangle]^{1/2} \Big|_{\phi\to 0+} = 2^{n+1/2}; \tag{4.10}$$

and then averages with respect to n in order to estimate the dispersion per iteration. In view of the exponential dependence on n it is obviously most convenient to average the logarithms. This finally leads to a simple measure of the intrinsic dispersion

$$\mathscr{D} = \lim_{n\to\infty} n^{-1} \ln \left\{ \frac{\partial}{\partial\phi} [\langle z^2(n)\rangle]^{1/2} \Big|_{\phi\to 0+} \right\} = \ln 2. \tag{4.11a}$$

It is easy to check that the corresponding result for the mth order Čebyšev polynomials is, cf. (3.30e)

$$\mathscr{D}(C_m) = \ln m, \qquad m \geqslant 2. \tag{4.11b}$$

A far more abstract way of characterizing the dispersion rate associated with mixing is by means of the metric and topological entropies of Kolmogorov and Adler [47]: Let $\mathscr{A}$ be an open cover of a compact space Ω, and $N(\mathscr{A})$ denote the minimum cardinality of all sub-covers of $\mathscr{A}$. Then $H(\mathscr{A}) = \log N(\mathscr{A})$ is defined to be the entropy of $\mathscr{A}$. The join of the two covers $\mathscr{A}$, $\mathscr{B}$ is the cover $\mathscr{A} \vee \mathscr{B} \equiv \{A \cap B;\ A \in \mathscr{A},\ B \in \mathscr{B}\}$. Now suppose that T is a continuous map of Ω into itself; then the connection with mixing enters by considering the joins associated with the iterates of T. Specifically the entropy of T with respect to the covering $\mathscr{A}$ is defined to be the limit

$$h(T, \mathscr{A}) = \lim_{n\to\infty} n^{-1} H(\mathscr{A} \vee T^{-1}\mathscr{A} \vee T^{-2}\mathscr{A} \vee \cdots \vee T^{-n+1}\mathscr{A}). \tag{4.12}$$

The intrinsic dispersion of T can then be obtained by averaging over the auxiliary coverings $\mathscr{A}$. Topologically this can be done with refining sequences: in particular we say that a cover $\mathscr{B}$ is a refinement of the cover $\mathscr{A}$ if every set of $\mathscr{B}$ is a subset of some set of $\mathscr{A}$. If we denote this relation by $\mathscr{A} \prec \mathscr{B}$, then a refining sequence is a set of open covers satisfying $\mathscr{A}_m < \mathscr{A}_{m+1}$, which is complete in the sense that for every open cover $\mathscr{B}$ of Ω there is some $\mathscr{A}_m$ such that $\mathscr{B} < \mathscr{A}_m$. Finally the topological entropy of the mapping T is defined in terms of the limit

$$h(T) = \lim_{m \to \infty} h(T, \mathscr{A}_m). \tag{4.13}$$

In the particular case of the Čebyšev polynomials it has been shown by Adler and McAndrew that the topological entropy is given by [35]

$$h(C_m) = \ln m, \qquad m \geqslant 1. \tag{4.14}$$

The numerical coincidence of (4.11b) and (4.14) may be significant insofar as it suggests a "microscopic" interpretation for the topological entropy. From a physical point of view it is of course clear that "entropy" is a misnomer in this context because the topological entropy has no connection with the properties of the "equilibrium" state, i.e., the asymptotic dispersion described by the probabilistic metric $F_c(z)$; cf. (4.6) and footnote 22 of reference [48]. Rather the simple criterion (4.11a) indicates that the underlying concept is the *rate* of entropy production, or numerically the rate of randomization.

Clearly the topological entropy rate (4.14) discriminates between the non-mixing Čebyšev polynomial $C_1 = x$ and all the mixing polynomials. There is a general surmise that the demarcation between the ordered and the disordered evolution of arbitrary complex systems corresponds to a transition to non-vanishing values of the topological entropy rate. Complex systems tend to have numerous instability zones and these generate patches with positive entropy rates [68]. Unfortunately applications relevant to physics (turbulence?) are still far too speculative to warrant discussion in print. However, some promising headway has been made with computer simulations of the iterated reflections of ideal "billiards" confined to convex domains. The key result is that billiard flows inside polygons and circles have vanishing entropy [19], whereas billiards in a stadion—a region whose boundary consists of two equal parallel segments joined by two semi-circles—are a K-flow and therefore mixing [18]. The point of the computer simulations then is to follow the collapse of the stadion billiard trajectories into ordered patterns as the stadion is smoothly deformed into a circular colosseum. It has been shown by Benettin *et al.* [41, 49] that this transition can be described by a numerical version of the topological entropy; in fact their arguments provide a direct link between the intrinsic dispersion rate (4.11b) and the topological entropy (4.14). The quantum version of this problem is discussed in [69].

We note first that the intrinsic dispersions (4.11a) and (4.11b) correspond precisely to the quantities denoted by $k(\tau, x, d)$ in Section IIb of reference [41]. These objects in turn approximate the so-called maximum Lyapunov characteristic of the flow. This relation is quite plausible since the highly technical construction of the Lyapunov characteristic is in fact a sophisticated (albeit differentiable) version of the definition (4.11a). In essence one considers an n-dimensional differentiable manifold $\mathscr{M}$, a vector field X defined on $\mathscr{M}$, and a flow $\{T^n\}$ on $\mathscr{M}$ induced by X. For $x \in \mathscr{M}$, the tangent space to $\mathscr{M}$ at x and the norm induced on it by the metric of $\mathscr{M}$ are denoted by E_x and $\| \cdots \|$ respectively. The tangent mapping of E_x onto $E_{T^n(x)}$ induced by the diffeomorphism T^n then is dT_x^n. It is also assumed that the flow $\{T^n\}$ preserves a normalized measure μ. All of this machinery can then be combined to show that there exists a measurable set $\mathscr{M}_1 \subset \mathscr{M}$, such that for every $x \in \mathscr{M}_1$ and for almost every vector $e \in E_x$, $e \neq 0$ the limit

$$\lim_{n\to\infty} n^{-1} \ln \| dT_x^n(e) \| = \lambda(x, e) \tag{4.15}$$

exists, is finite, and non-negative. For the Čebyšev mixing the Lyapunov characteristic is given by

$$\mathscr{D}(C_m) = \ln m \to \lambda(x, e), \qquad m \geqslant 2. \tag{4.16}$$

This identification substantiates the interpretation of $k(\tau, x, d)$ given in Section 2B of reference [49].

Next we use Piesin's Theorem [50] to connect the *metric* entropy of Kolmogorov [51] with the Lyapunov characteristic, viz.

$$h_\mu(C_m) = \int_{\mathscr{M}} \lambda(x, e)\, d\mu(x). \tag{4.17}$$

In the present instance the integration is trivial, and accordingly the metric entropy of the Čebyšev mixing, with respect to the measure (4.1a), is simply

$$h_{P_c}(C_m) = \ln m, \qquad m \geqslant 2. \tag{4.18}$$

Finally, the technical step of verifying the agreement of the metric and topological entropies for the Čebyšev mixing has been carried out by Ranade [52]. Q.E.D.

5. Variable Precision Simulations

Small computers are unsuitable for simulating the behavior of Čebyšev mixing. For instance in the extreme case of a device which can only compute with the digits 0, ±1, ±2, the $x^2 - 2$ iterations degenerate into the following pattern:

input	first iteration	second iteration		nth iteration
$+2$	$+2$	$+2$	→	$+2$
$+1$	-1	-1	→	-1
0	-2	$+2$	→	$+2$
-1	-1	-1	→	-1
-2	-2	$+2$	→	$+2$

(5.1)

Obviously there are no cycles, and after only two iterations everything has collapsed into the two fixed points -1 and $+2$. However, as the machine capacity is increased more numerical and combinatorial complexity can appear in the iterations, and gradually the statistical regularities associated with the Čebyšev mixing begin to be superposed on the computer generated sequences. The transition from the combinatorial to the statistical regime can be modeled by programming a large computer, such as a Univac 1108, to carry out the Čebyšev iterations with adjustable N-bit modular arithmetic. In this way one can empirically establish that beyond the range $N \gtrsim 24$ the computer simulations furnish usable strings of pseudo-random numbers. In addition to deliberately "expanding" the computer it is also interesting to model the effects of perturbing the mixing flow. This can be done with a slight programming change in which variable truncations play the role of the perturbations. As indicated in Section 2(vii) we expect a greater fidelity in the simulations when the intrinsic mixing rate (4.11b) dominates the rate of immigration of the round-off errors (3.31).

In practise the variable truncation studies were carried out on an 1108 in double precision arithmetic. We chose as the initial "seeds" the numbers $x_0{}^i = i \times 10^{-2}$, $i = 1, 2, 3, \ldots, 200$; and used a slight modification of (3.14b) to generate iterates with variable truncations, viz.

$$\bar{x}^i_{n+1} = \frac{1}{10^{N_0}} [10^{N_0} \bar{x}_n{}^2] - 2, \qquad N_0 = 1, 2, \ldots, 7. \tag{5.2}$$

The iterations were continued until either (a) we reached the index $n = 10^4$, or (b) located two positive integers k_i and l_i satisfying the condition cf. (3.7a)–(3.7c)

$$\bar{x}^i_{k_i} = \bar{x}^i_{k_i+l_i}\,. \tag{5.3}$$

Obviously k_i corresponds to the index where the iterations stemming from $\bar{x}_0{}^i$ enter a terminal loop, and l_i is the length of the loop, cf. (3.17) and (3.23). It is convenient to summarize all of this information by introducing two auxiliary indices:

$$K(N_0) = \text{Max}\{k_i\}, \tag{5.4}$$

and

$$L(N_0) = \prod_{(\text{distinct } l_i)} l_i\,. \tag{5.5}$$

Clearly $K(N_0)$ represents the effective onset of cyclic behavior because for $n > K(N_0)$ *all* the sequences $\{\bar{x}_n{}^i\}$ have entered terminal cycles. Figure 1 is a semi-logarithmic plot showing the variation of K with N_0: the essential feature of this graph is the appearance of a sharp "knee" in the vicinity of $N_0 \gtrsim 7$. The object $L(N_0)$ represents the common periodicity of all the terminal loops; the growth of this index is displayed on Fig. 2. It is also evident from this graph that the mixing simulations "take off" when $\geqslant 7$ decimal place arithmetic is available on the computer.

The "expanding computer" experiments were carried out in b-bit modular arithmetic for b varying between $1 \leqslant b \leqslant 24$. In this case it was convenient to rescale the itera-

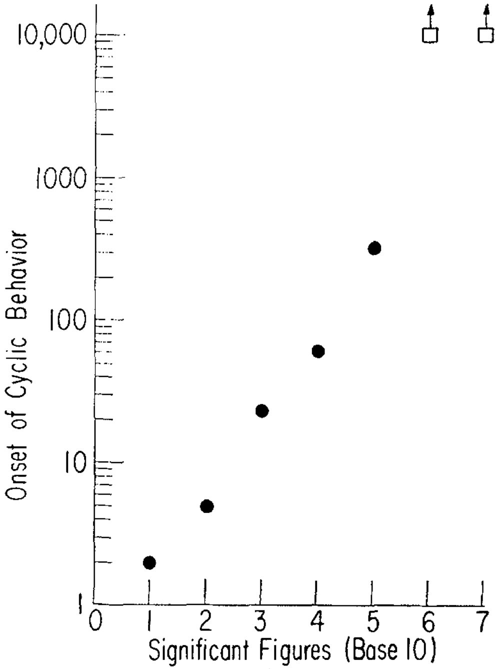

FIG. 1. Onset of cyclic behavior in Čebyšev mixing simulations. The dots indicate the variation of the function $K(N_0)$, cf. (5.4). In all figures the boxes indicate lower limits ($\geqslant 10^4$).

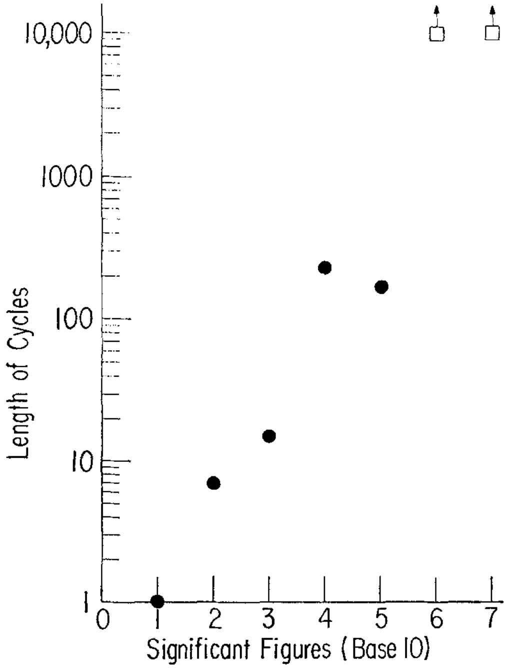

FIG. 2. Common periodicity of cycles in Čebyšev mixing simulations. The figure displays the variation of the func tion $L(N_0)$, cf. (5.5).

tions to the interval $\tilde{x} \in [-1, 1]$ and arrange the program to compute, cf. (3.4) and (3.14b),

$$\tilde{x}_{n+1} = \frac{1}{2^b} [2^{b+1} \tilde{x}_n{}^2] - 1. \tag{5.6}$$

For $b \leqslant 9$ we traced the iterative behavior of *all* $2^{b+1} + 1$ binaries in the interval $[-1, 1]$. Figure 3 shows that the trivial fixed point collapse (5.1) is augmented by the appearance of a single terminal cycle of length 3 when $b = 4$. Beyond this threshold there is an increase both in the number of cycles as well as their length. However, due to practical limitations of computer time these results could only be spot checked in the range $9 < b \leqslant 24$: this demarcation is indicated by the "exact" and "statistical" portions on Figs. 3 and 4. Presumably the interplay between the combinatorial and statistical elements tends to produce a scatter in the cycle lengths. Nevertheless there appears to be a regularity in the increase of the maximum cycle lengths in the range $5 \leqslant b \leqslant 23$. At $b \gtrsim 24$ the maximum cycle length jumps to a value exceeding 10^4; this behavior parallels the "take off" in the mixing exhibited on Figs. 1 and 2.

These features are also apparent on Fig. 4 which summarizes information on the free running intervals, cf. (3.17). In the range $1 \leqslant b \leqslant 9$ we show the onset of cyclic

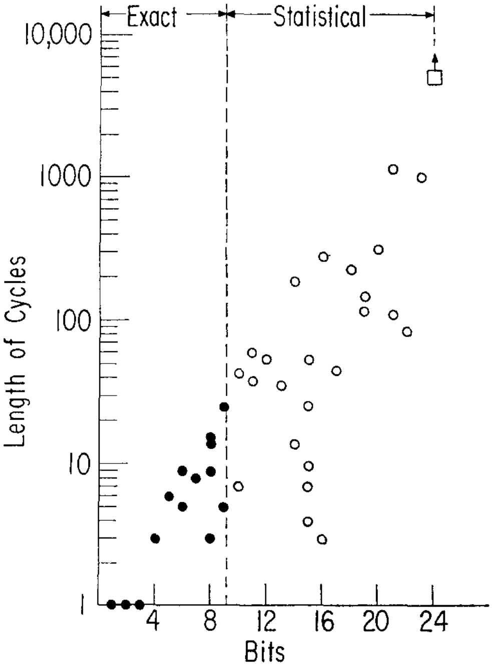

FIG. 3. Cyclic behavior in Čebyšev mixing simulations. The lengths of the individual terminal cycles are shown as functions of the machine accuracy. For $b < 9$ all accessible binary numbers were followed to their terminal cycles; the results for $b \geqslant 9$ are statistical spot checks.

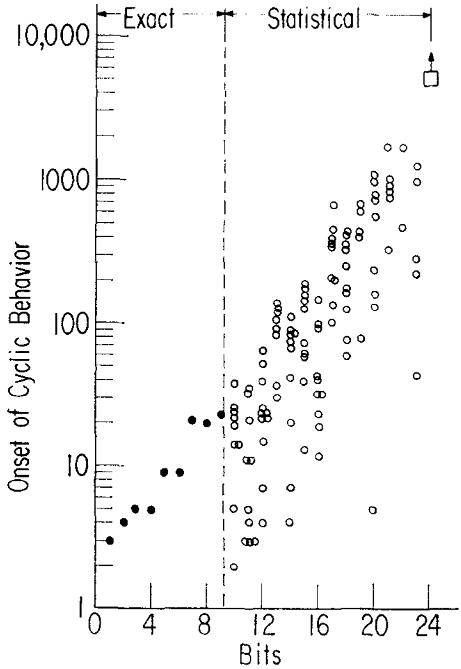

FIG. 4. Onset of cyclic behavior in Čebyšev mixing simulations. The points indicate the lengths of the intervals of free running. See the text for a complete explanation.

behavior by plotting $K(b)$ which is the binary analogue of $K(N_0)$, cf. (5.4). For larger values of b, $9 < b \leqslant 24$, we display the overall trends including the statistical scatter of the *individual* free running intervals. Once again there is an indication of a sharp increase in the vicinity of $b \gtrsim 24$. In view of the effective conversion between binaries and decimals, i.e., $b \sim (\log_2 10)q \sim 3.32\, q$, it is obvious that all of the graphs shown on Figs. 1–4 are consistent, exhibiting symptoms of a "stochastic transition" at a computer capacity of the order of 7 decimal figures.

It is also interesting to check on the growth of computer noise by comparing the influence of simple truncation ("T") and roundoff ("R"), i.e., increasing the last available guard digit by 1/2 and then truncating. In the extreme case (5.1) these distinctions are of course irrelevant. However, with increasing computing capacity variations in the truncation routines and inherent differences in the computers give rise to inequivalent perturbations of the mixing.

Let us first consider an *HP*-25 (10 digits + 1 guard) and an *SR*-52 (10 digits + 3 guard) programmed to carry out the $x^2 - 2$ iterations according to (5.2), but restricting the arithmetic to 3 place accuracy, i.e., $N_U = 4001$ cf. (3.11a), by using roundoff. In this case we spot checked the iterative behavior of some of the same seeds that were used in Table II(a): The results are given in Table V—the essential point being that the *same* results were obtained on both computers. When the arithmetic accuracy is increased the mixing simulations become more sensitive to computing

TABLE V

Computer Simulations of Čebyšev Mixing: $N_U = 4001$, cf. (3.11a)

Initial number	Interval of free running	Length of terminal loop
$\pi - 3$	16	23
$\pi/2$	40	23
$(\sqrt{2} - 1)^2$	71	23
$e - 1$	42	23
1.555	78	23

TABLE VI

Computer Simulations of Čebyšev Mixing:
Comparison of Different Machines and Truncation Methods.
$N_U = 4 \times 10^8 + 1$, (3.11a)

Initial number	Approximate interval of free running				Length of terminal loop			
	HP-25		HP-9100		HP-25		HP-9100	
	T[a]	R[b]	T	R	T	R	T	R
$\pi - 3$	1,556	—	5,537	6,606	2,498	+2[c]	35	3,109
$(\sqrt{2} - 1)^2$	541	16,720	2,398	5,677	4,612	10,525	+2[c]	+2[c]
$\pi/2$	4,956	14,091	7,415	17,743	2,498	10,525	35	3,109
$\pi^2 - 9$	13,054	5,992	3,045	11,127	2,498	10,525	+2[c]	3,109
$\pi^5 - 306$	10,271	6,703	8,703	2,194	2,498	10,525	35	3,109

[a] $T \equiv$ simple truncation.

[b] $R \equiv$ roundoff (see text).

[c] Iterations terminate on fixed point +2.

TABLE VII

Variable Precision Simulations of Čebyšev Mixing on an HP-25 with Truncation

	Approximate interval of free running				Length of terminal loop			
$N_U - 1$ (3.11a)	4×10^3	4×10^6	4×10^8	4×10^9	4×10^3	4×10^6	4×10^8	4×10^9
Initial number								
$\pi - 3$	3	2,140	1,556	16,064	18	402	2498	4158
$\pi/2$	46	539	4,956	21,491	18	402	2498	4158
$(\sqrt{2} - 1)^2$	11	1,038	541	21,348	18	402	4612	4158
$e - 1$	9	—	3,534	14,211	12	—	4612	4158
$\pi^2 - 9$	—	830	13,054	12,773	—	402	2498	4158
$\pi^5 - 306$	—	496	10,271	26,654	—	402	2498	4158
1.555 ...	—	—	2,676	30,878	—	—	2498	4158

noise and the results depend both on the machine as well as the truncation conventions —all of these features are illustrated in Table VI. Finally, in Table VII, we present some information supplementing the data given on Figs. 1–4.

6. Terminal Cycles

It is interesting to see how the computer iterations represent a synthesis of the pseudo-random character of the Čebyšev mixing and the asymptotic regularity imposed by the terminal cycles. Extensive numerical trials have shown that on large machines ($N_U \lesssim 10^{10}$) with long terminal cycles, i.e., $n_L - n_f \sim N_U^{1/2}$ (3.22), most of the statistical properties of the Čebyšev mixing are preserved on the cycles. An exception is the auto-correlation (2.4a) which becomes a Kronecker function with period $n_L - n_f$. Some of the statistical data concerning the behavior of the agitation (3.12a) and the frequency distribution (3.9b) on terminal cycles is summarized on Table VIII. On the basis of all the evidence from the computer experiments it may not be too fanciful to describe this as a situation in which order masquerades as chaos. Naturally these features are relevant to communications codes. It was originally

TABLE VIII

Statistical Properties of Terminal Cycles: Agitation (3.12a) and Normalized Frequency Distribution (3.9b). Data for the Three Terminal Loops of Table II(a).

	Length of loop	39,965	95,447	104,694	Theoretical values
x_U	x_L	$\mathscr{A}$ 1.653	1.654	1.656	1.654
+2.0	1.8	0.143	0.144	0.143	0.144
1.8	1.6	0.0624	0.0612	0.0606	0.0613
1.6	1.4	0.0476	0.0483	0.0488	0.0484
1.4	1.2	0.0404	0.0426	0.0422	0.0420
1.2	1.0	0.0388	0.0376	0.0375	0.0382
1.0	0.8	0.0357	0.0357	0.0366	0.0357
0.8	0.6	0.0351	0.0342	0.0335	0.0340
0.6	0.4	0.0326	0.0322	0.0324	0.0329
0.4	0.2	0.0323	0.0331	0.0330	0.0322
0.2	0.0	0.0316	0.0309	0.0320	0.0319
0.0	−0.2	0.0325	0.0330	0.0317	0.0319
−0.2	−0.4	0.0318	0.0317	0.0317	0.0322
−0.4	−0.6	0.0320	0.0330	0.0341	0.0329
−0.6	−0.8	0.0346	0.0337	0.0340	0.0340
−0.8	−1.0	0.0360	0.0355	0.0352	0.0357
−1.0	−1.2	0.0388	0.0384	0.0382	0.0382
−1.2	−1.4	0.0423	0.0418	0.0423	0.0420
−1.4	−1.6	0.0492	0.0478	0.0482	0.0484
−1.6	−1.8	0.0611	0.0611	0.0615	0.0613
−1.8	−2.0	0.143	0.144	0.144	0.144

pointed out by Shannon [53] that mixing transformations could be adapted to scrambling messages. Since decrypting relies heavily on statistical analyses the camouflage of order by mixing is enhanced in cases where the probabilistic metric information is equilateral [39]; specifically, as in (4.6), the distribution function $F_c(z)$ is independent of the underlying (x, y) "message" configuration when the product transformation is ergodic. These concealment possibilities can be extended still further in virtue of the sensitivity of the terminal cycles to perturbations (Table VI), as well as the statistical indistinguishability of the terminal cycles from the intervals of free running (Tables I and VIII), see also reference [54].

Crude numerical arguments based on (3.21c) and (3.22) indicate that the total number of terminal cycles for a given computer and programming mode is rather small. To some extent this can be confirmed experimentally. Let us consider a b-bit machine with a total "universe" of $N_U \sim 2^b$ numbers; (5.6) *et seq.* Furthermore let $C(b)$ denote the corresponding number of iterates which are members of terminal cycles. Then the data accumulated in connection with Figs. 1–4 and Tables V–VIII show that N_U and $C(b)$ are empirically related by the ratio

$$\frac{C(b)}{N_U} \sim 2^{-0.57b} \tag{6.1}$$

which diminishes rapidly with increasing computer size. More precisely for $b \gtrsim 4$ the combinatorics of the iterations lead to an *increase* in the number of terminal cycles; however, for still larger values of b, say $b \gtrsim 25$, the number of terminal cycles which can be located by *statistical* methods remains rather small.

The number of terminal cycles is significant because the entire iterative scheme can be envisaged as evolving in reverse order—this corresponds to the set $C(b)$ expanding throughout the universe N_U. Mathematically this reversion presents no difficulties because the pre-images are well defined even though the Čebyšev mixing is not invertible. For instance the first and second pre-images of x are

$$\begin{tikzcd}
 & +(x+2)^{1/2} \arrow[r] \arrow[dr] & +((x+2)^{1/2}+2)^{1/2}, \\
x \arrow[ur] \arrow[dr] & & -((x+2)^{1/2}+2)^{1/2}, \\
 & -(x+2)^{1/2} \arrow[r] \arrow[dr] & +(-(x+2)^{1/2}+2)^{1/2}, \\
 & & -(-(x+2)^{1/2}+2)^{1/2};
\end{tikzcd} \tag{6.2}$$

and obviously at the nth stage the pre-images comprise 2^n numbers (of course there is no contradiction with the preservation of measure). At first sight one might then suppose that the terminal cycles are surrounded by sunbursts of cascading pre-images somewhat resembling the pattern indicated on Fig. 5. However, it is easy to show by direct enumeration that this is actually an impossible configuration on a finite-bit machine. Consider for instance the terminal loop of length 4158 which occurs on an *HP*-25 calculator, cf. the last column of Table VII. If in fact every element of the

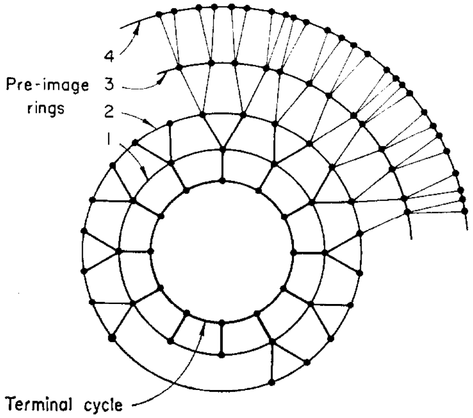

FIG. 5. An "impossible" configuration of pre-images, cf. (6.2).

loop gave rise to a pre-image chain, as indicated in eq. (6.2), then by the nth stage the total set of numbers included in the "sunburst" would be $\sim 4158 \times 2^n$. Since the total universe of accessible numbers is $4 \times 10^9 + 1$, it is apparent that the reverse iterations will saturate the computer capacity in the vicinity of $n \sim 30$. However, this estimate differs by a factor of 10^3 from the theoretical (3.22) and experimental (column 5, Table VII) values for the average length of the interval of free running!

The flaw of this argument lies in the presumption that the pre-image chain (6.2) can be simulated on a computer. A little experimentation will show that for any digital device the limited accuracy of the machine arithmetic entails the existence of gaps so

on points. The genealogical analogy implied by (6.2) suggests that we call these "orphan" or "originating" points. In any event it is clear that the Čebyšev simulations can be visualized as flow networks originating from a discrete set of orphan points; aggregating into larger streams by confluence; and finally merging into a small set of terminal cycles. The inherent computational asymmetry of this process arises from the simple but deep distinction between the pre-image chain (6.2) and the forward iterations (3.17). In coding terminology this corresponds to the effective construction of so-called one way functions [54].

Figure 6 indicates schematically the structure of a flow network for an SR-56 programmed to iterate in 3 place accuracy. In this case the Čebyšev mixing simulations lead to a single terminal loop with 76 numbers. Starting from each of these numbers in turn, it is possible to trace out the complete pattern of pre-images. In this way the graphical linkages—induced by the Čebyšev iterations—of all the $4 \times 10^3 + 1$ numbers in the computer universe can be determined. Unfortunately at this low level of precision the Čebyšev simulations are perturbed by combinatorial factors (cf. Section 5) and therefore the structural peculiarities exhibited in Fig. 6 may not be typical.

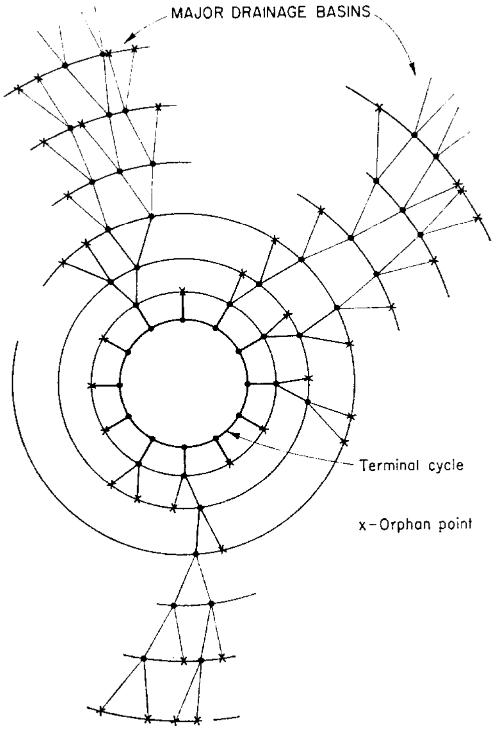

FIG. 6. Flow network or de Bruijn diagram [13] for a small scale Čebyšev mixing simulation (SR-56 programmed in 3-place accuracy). The terminal loop is insulated from the rest of the computer "universe" by a cluster of "orphan" points. Note the contrast with the pre-image pattern shown on Fig. 5.

The essential point of these tedious studies was to gain some insight into the nature of the terminal loops. In this respect the two most striking features of Fig. 6 are: (i) the almost total separation of the terminal loop from the rest of the universe by sets of orphan points; and (ii) the tendency for most of the rest of the points in the universe to cluster about a few major "drainage basins" in the flow network.

The insulation of the terminal loops can easily be described in terms of the pre-image chain (6.2). Suppose in particular that x is a point on a terminal loop linked to descendants and ancestors in the following sequence:

$$\cdots \leftarrow x^4 - 4x^2 + 2 \leftarrow x^2 - 2 \leftarrow x \leftarrow -(x + 2)^{1/2} \leftarrow +(-(x + 2)^{1/2} + 2)^{1/2} \leftarrow \cdots \tag{6.3}$$

It is then evident that the terminal loop must be surrounded by a *first pre-image ring* each of whose elements is the negative of a number on the terminal loop. In particular the ancestors corresponding to (6.3) on the first pre-image ring are the numbers

$$-(x^4 - 4x^2 + 2),\ -(x^2 - 2),\ -x,\ +(x + 2)^{1/2},\ -(-(x + 2)^{1/2} + 2)^{1/2}. \tag{6.4}$$

Carrying the reversion back one more step leads to the *second pre-image ring*: the ancestors corresponding to (6.4) then are the elements

$$\pm(x^4 - 4x^2)^{1/2},\ \pm(-x^2 + 4)^{1/2},\ \pm(-x + 2)^{1/2},\ \pm((x + 2)^{1/2} + 2)^{1/2},$$
$$\pm[-(-(x + 2)^{1/2} + 2)^{1/2} + 2]^{1/2}; \tag{6.5}$$

and as indicated previously *these numbers may not be accessible on the computer.* Figure 6 and Table IX display the curious feature that more than 80 % of the elements of the second and third pre-image rings are orphan points; in other words egress from the terminal cycle is possible only at a few junction points. Similar results have been obtained for the 95,447 terminal cycle on an *HP*-9100 (Table VIII), and we consider it likely that this may be a general property of the flow networks.

There may be some correlation between the frequency distribution of the Čebyšev mixing and the tendency exhibited in Table IX for points near ± 2 to be preferred ports of entry onto the terminal loop. However, it does not seem plausible that metric concepts are adequate to describe the characteristics of the iteration patterns. After all, the essential import of the information on Table VIII is to confirm that the asymptotic dispersion (4.6) is equilateral, i.e., independent of (x, y), and therefore in a metric sense the terminal loops are indistinguishable from the rest of the flow network. Moreover, it is impossible to characterize the numbers on the terminal loop by metric stability criteria: these require some type of norm such as (4.7), but in virtue of the fact that $x^2 - 2$ has no fractional iterates whatsoever—a basic point to which we shall return in Section 7—it is futile to look for an association between metric resemblance and genealogical kinship (6.2)! The conspicuous clustering of the orphan points around the terminal cycles may, however, be interpreted as a clue pointing in the direction of a topological stability concept.

Finally it should be noted that the entire situation is altered if the mixing simulations are not restricted to simple iterations without memory. It is then easy in practise to extend the intervals of free running by interleafing or cross-feeding the elements of several mixing simulations run in parallel. A more interesting possibility is to perturb the simple mapping sequence (3.17) with a memory-dependent feedback. Specifically if we have arrived at the mth term of the sequence $\{x_m\}$, it is possible to associate a "memory" with the values of the preceding j terms by forming the product

$$\prod_{m-j}^{m-1} x_i = p_j(x_m).$$

The standard Čebyšev iteration (3.17) can then be perturbed with the increments $\delta p_j(x_m)$, i.e.,

$$\cdots \to \mathscr{M}^m[x_m + \delta p_j(x_m)] \to \mathscr{M}^{m+1}[x_{m+1} + \delta p_j(x_{m+1})] \to \cdots, \tag{6.6}$$

where δ is adjusted to be sufficiently small so that the perturbation does not spoil the statistical stability of the flow. In this case both of the sequences (3.17) and (6.6) will

TABLE IX

Pre-Image Chains for a 76 Member Terminal Loop: SR-56 with 3-place Accuracy

Flux[a] (%)	23.2	23.0	11.5	2.40	2.35	2.20	1.70	1.25	1.20	1.15	0.75	$\leqslant 0.4$	0
Element of Loop	-1.996	0.842	-1.041	1.960	1.928	-0.070	0.056	-1.790	1.980	-0.131	1.984	19 points	40 points
Label[b]	33	10	43	2	70	5	32	16	7	68	34		

[a] Flux ≡ 100 × (total number of elements of universe whose iteration chains (3.17) enter the terminal loop at this element)/(total number of elements in universe).

[b] Enumeration of the 76 elements of the terminal loop: "-1.990" is labeled no. 1, "$+1.960$" is no. 2, etc.

tend to encounter the same element again after about $N_U^{1/2}$ iterations, but owing to the memory dependence (6.6) will not degenerate into a clock.

7. Other Simulations of Random Processes

The procedures developed for the Čebyšev mixing simulations can in principle be applied to other mathematical models of random processes. A familiar group of examples is the bakers' transformations [28, 17] which are mappings of the unit square with periodic boundary conditions. In one version the mixing is generated by iterating the simple two-dimensional mapping

$$g(x, y) \rightarrow (x + y, x + 2y) \qquad (\text{mod } 1) \tag{7.1}$$

in which g plays a role analogous to f in (3.1). One can then easily check that the nth iterate is given by [55]

$$g^n(x, y) \rightarrow (F_{2n-1}x + F_{2n}y, F_{2n}x + F_{2n+1}y) \qquad (\text{mod } 1) \tag{7.2}$$

where F_N denotes the Nth Fibonacci number ($F_N = F_{N-1} + F_{N-2}$; $F_1 = F_2 = 1$). This result parallels the nth iterate Čebyšev expression (3.3). The fidelity of the mixing simulations can also be monitored by statistical tests similar to those developed for the Čebyšev iterations. The probabilistic metric corresponding to (3.24) for the bakers' transformation is [55]

$$B(z) = \begin{cases} 2\pi z, & 0 \leqslant z < 1/2; \\ z[2\pi - 8 \cos^{-1}(1/2z)], & 1/2 \leqslant z < 2^{-1/2}; \end{cases} \tag{7.3}$$

where z denotes the Euclidean distance between x and y, and the upper bound $z < 2^{-1/2}$ stems from the periodic boundary conditions (torus mapping). The asymptotic dispersion or mean distance between the points x and y can be obtained by elementary quadratures, cf. (3.25a) and (3.25b). In the present instance this is given by

$$\begin{aligned} D_B &= 2\pi \int_0^{2^{-1/2}} dz\, z^2 - 8 \int_{1/2}^{2^{-1/2}} dz\, z^2 \cos^{-1}(1/2z), \\ &= \tfrac{1}{6}[\sqrt{2} + \ln(1 + \sqrt{2})] \cong 0.382\,597\,8. \end{aligned} \tag{7.4}$$

The frequency distribution (3.9b), agitation (3.12a), and other statistical properties of the mixing sequences (7.2) may then be derived by straightforward means. The remaining basic statistical index, the Kolmogorov rate of entropy production (4.12), can also be determined explicitly since the bakers' transformation is isomorphic to a (1/2, 1/2) Bernoulli shift [56, 52]. It is interesting that the result for the bakers' transformation

$$h(g) = \ln 2 \tag{7.5}$$

coincides numerically with the entropy production for the second degree Čebyšev polynomial (4.14)—auguring well for the computer simulations—although the Čebyšev mixing and two-sided Bernoulli shifts are *not* isomorphic [55]. This difference is crucial for physical interpretations (*vide infra*).

The Čebyšev and bakers' transformations are only two isolated examples of a large class of mixing processes. It was shown by Brolin [21] that *any* non-linear polynomial has a restriction which is mixing with respect to an appropriate measure. Informally this means that any polynomial with degree $\geqslant 2$ mixes somewhere in the complex plane. In general, the set over which the mixing occurs is perfect but nowhere dense. In the special case that the mixing set is a Jordan curve, i.e., an arc or a closed curve without self-intersections, it was shown by Ranade [52] that the corresponding nth degree polynomials are isomorphic to one-sided Bernoulli n-shifts and therefore have topological and metric entropy production rates equal to $\ln n$, cf. (4.12) and (4.14). From this perspective the entire Čebyšev mixing theory is merely a special case of Brolin's theorem in which the Jordan curve is a real interval, cf. (3.4), and the polynomial degree assumes its minimum value. The implication of these results for practical computer experiments is twofold: not only is the choice of the iterating polynomial crucial, but it is also important to find the region in the complex plane where the mixing actually occurs.

There are a wide variety of problems in fluid mechanics [15, 57] and epigenetic networks [58] which involve processes that are apparently non-periodic and irregular. In computer simulations it is therefore essential to utilize mathematical models that will stave off collapse into terminal cycles as long as possible. Mixing processes are of course plausible candidates because they simulate pseudo-random behavior and have long periods of free running before merging into terminal cycles. However, in this respect the iteration of rational functions is only a special case. Other likely prospects for modeling irregular behavior with very long cycles can be identified with the help of more general, albeit weaker, results of Šarkovskii [20]. The key point is that complicated iterative structure can, so to speak, be "imposed from below"; this notion is also forcefully expressed by the title of reference [59], "Period Three Implies Chaos." Let us recall that a point x_m is said to be a fixed point of order k of a function h if $h^k(x_m) = x_m$ and $h^j(x_m) \neq x_m$ for all $j = 1, 2,\ldots, k - 1$; cf. (3.7b). Suppose that h belongs to the class C of all continuous, real-valued functions defined on $(-\infty, \infty)$, i.e., it is a mapping of the real line into itself. Now define an ordering relation $\prec$ for positive integers as follows: $k \prec n$ if for any function $h \in C$ the existence of a fixed point of order k of h implies the existence of a fixed point of order n of h, but not conversely. Šarkovskii's basic result then is that the set of positive integers is ordered by the relation $\prec$ in the following way:

$$\begin{aligned} 3 \prec 5 \prec 7 \prec 9 \prec 11 \prec \cdots \prec 3 \times 2 \prec 5 \times 2 \prec \cdots \prec 3 \times 2^2 \prec 5 \times 2^2 \\ \prec \cdots \prec 2^3 \prec 2^2 \prec 2 \prec 1. \end{aligned} \tag{7.5}$$

So for example if a function has a fixed point of order 8, it necessarily has fixed points of order 4, 2, and 1, but nothing more can be asserted; on the other hand if it has a fixed point of order 3, then there must exist fixed points with *arbitrarily* long periods: Furthermore this entails the existence of an uncountable set of points, not necessarily with positive measure, which wander perpetually under iteration since they do *not* asymptotically approach *any* periodic point [59].

All of these results are consistent with the experimental finding that functional iteration can indeed be adapted to the computer simulation of random processes. Moreover, the simple functional form of some of the mixing transformations suggests direct physical realizations. For instance the quadratic Čebyšev polynomial describes the operation of a biased product detector or, equivalently, an amplitude limited nonlinear amplifier. In either case (3.13) corresponds to a voltage transducer with the property that [38]

$$V(\text{output}) = [V(\text{input})]^2 - 2, \tag{7.6}$$

where

$$-2 \leqslant V(\text{input}) \leqslant +2 \text{ volts} \rightarrow -2 \leqslant V(\text{output}) \leqslant +2 \text{ volts}.$$

A cascade of such devices linked together on an analogue computer will exhibit all the erratic characteristics of mixing. However, it is important to realize that the Čebyšev mixing *cannot* be associated with the evolution of any Hamiltonian system. Sinai's wry definition "...ergodic theory consists of the study of the statistical properties of the groups of motions of non-random objects" [19] also points up the narrow scope of the physical models inherited from the historical association of Hamiltonian mechanics and ergodic theory [60, 61].

The gap between Hamiltonian mechanics and iteration arises from an embedding problem: Specifically let us suppose that the state of a physical system can be described by a function $F(t, X)$, where X corresponds to the initial conditions and t denotes some index of the evolution—usually the time. If the evolution has no memory dependence, i.e., no hysteresis, then the state of the system at "time" $t_1 + t_2$ will be identical to the state reached in the two-step transition $0 \rightarrow t_1$, $t_1 \rightarrow t_2$; consequently F must satisfy the basic translation or flow equation [62]

$$F(t_1 + t_2, X) = F[t_1, F(t_2, X)]. \tag{7.7}$$

The elementary transcription $F(t, X) \equiv g_t(X)$ then permits us to rewrite (7.7) with a somewhat different emphasis:

$$g_{t_1+t_2}(X) = g_{t_1}[g_{t_2}(X)] = g_{t_2}[g_{t_1}(X)]. \tag{7.8}$$

This form exhibits the connection between evolution and functional iteration. Any family of functions $\{g_t\}$ satisfying the basic relation (7.8) for all t_1, $t_2 \geqslant 0$ is called a *one-sided flow*. A *two-sided flow* is a similar family of functions in which the index t ranges over all real numbers [63]. Flows are special cases of Abelian semi-groups. A function f is *embeddable in a flow* if there exists a flow $\{g_t\}$ such that $f = g_t$ for some positive t (the positivity is not an essential restriction). As a consequence of the preceding we have the

LEMMA [63]. *If a function f is embeddable in a flow, then f has iterative roots of* all *orders, i.e., there exist functions h such that*

$$h^n(X) = f(X) \qquad \text{for all integers} \quad n \geqslant 2. \tag{7.9}$$

In the Hamiltonian case $F(t, X)$ corresponds to the principal function, and the associated (differentiable) flow $g_t(X)$ describes the continuous sequence of infinitesimal canonical transformations that governs the evolution of the characteristic surfaces [64]. On the other hand the Čebyšev polynomials satisfy the following

THEOREM (Sklar [63]). *A sufficient condition for the nth order Čebyšev polynomial C_n to have* no fractional iterates whatever *is that*

$$n^p - n \qquad \text{is not divisible by} \quad p^2 \tag{7.10}$$

for any prime $p \leqslant n$. Direct calculation shows that this condition is satisfied for $n =$ 2, 3, 6, 11, 14, *etc.* (*C_5 has iterative square roots but these are not even measure preserving* [55, 66].)

This theorem and the preceding lemma imply that the Čebyšev iterations cannot be embedded in any one sided or two sided flow; therefore it is impossible to construct a Hamiltonian system whose evolution describes the Čebyšev iterations. Clearly there is a profound difference between systems in which the t-dependence in (7.8) is so smooth as to be continuous or even differentiable, and non-embeddable systems in which the t variation has an inherent graininess or quantization. The physical implications of non-embeddability can be quite drastic: If we recall the electrical realization of C_2 as a biased product detector, then (7.9) and (7.10) show that it is impossible to construct n ($\geqslant 2$) identical "black boxes" which, when connected in cascade, reproduce the voltage transformation (7.6). Of course mathematical theorems cannot prohibit us from looking "inside" the product detectors or the computers which are programmed to carry out the Čebešev iterations: At this microscopic level one might presume to find "hidden variables" which interpolate in a smooth and well defined way "between" the quantum steps of the non-embeddable flow. However, these distinctions between the evolution of structures and Hamiltonian systems involve epistemological questions which would take us too far afield.

ACKNOWLEDGMENTS

We are extremely grateful to our colleague A. Sklar for innumerable helpful conversations on all aspects of this work. Many of the computer experiments were carried out by J. Nuti and T. M. Rynne. It is also a pleasure for T. E. to acknowledge the hospitality of the Department of Physics at UCLA.

References

1. R. L. Coldwell, *J. Computational Physics* **14** (1974), 223.
2. R. W. Hamming, "Numerical Methods for Scientists and Engineers," McGraw–Hill, New York, 1962.
3. A. Rotenberg, *J. Assoc. Comput. Mach.* 7 (1960), 75.
4. J. N. Franklin, *Math. Comput.* **17** (1963), 28.
5. I. Niven, "Irrational Numbers," Carus Math. Monograph No. 11, Math. Assoc. Amer., 1956.
6. H. Weyl, *Math. Ann.* **77** (1916), 313.
7. J. G. van der Corput, *Acta Math.* **56** (1931), 373.
8. M. Zelen and N. C. Severo, *in* "Handbook of Mathematical Functions," (M. Abramowitz and I. A. Stegun, Eds.), N.B.S. Applied Math. Series, No. 55, U. S. Govt. Printing Office, Washington, D.C., 1964.
9. W. Schmidt, *Pacific J. Math.* **10** (1960), 661.
10. R. L. T. Hampton, *in* "Proceedings of the Spring Joint Computer Conference, 1964," p. 287.
11. T. Erber, B. Schweizer, and A. Sklar, *Comm. Math. Phys.* **29** (1973), 311.
12. P. Johnson and A. Sklar, *J. Math. Anal. Appl.* **54** (1976), 752.
13. N. G. de Bruijn, *Koninklijke Nederlandse Akad. Wetenschappen, Proc. 49*, Part 2 (1946), 758.
14. S. W. Golomb, "Shift Register Sequences," Holden–Day, San Francisco, 1967.
15. E. N. Lorenz, *J. Atmos. Sci.* **20** (1963), 130.
16. D. V. Anosov, *Dokl. Akad. Nauk. SSSR* **151** (1963), 1250; *Soviet Math. Dokl.* 4 (1963), 1153.
17. A. Gervois and M. L. Mehta, *J. Math. and Phys.* **18** (1977), 1476.
18. L. A. Bunimovich, *Funkcional Anal. i Priložen* **8** (1974), 73; *Functional Anal. Appl.* **8** (1974), 254.
19. Ya. G. Sinai, "Introduction to Ergodic Theory," Mathematical Notes, Princeton Univ. Press, Princeton, N. J., 1976.
20. A. N. Šarkovskii, *Ukrain. Mat. Ž.* **16** (1964), 61.
21. H. Brolin, *Arkiv. Mat.* **6** (1965), 103.
22. D. E. Knuth, "The Art of Computer Programming. Vol. 2. Seminumerical Algorithms," Addison–Wesley, Reading, Mass., 1969.
23. R. Carnap, "Logical Foundations of Probability," Univ. of Chicago Press, Chicago, 1962.
24. P. Billingsley, "Ergodic Theory and Information," Wiley, New York, 1965.
25. P. R. Halmos, "Lectures on Ergodic Theory," Chelsea, New York, 1956.
26. J. C. Oxtoby and S. M. Ulam, *Ann. of Math.* **42** (1941), 874.
27. W. Feller, "Introduction to Probability Theory and Its Applications," Vol. 1, Wiley, New York, 1950.
28. E. Hopf, *J. Math. Phys.* **13** (1934), 51.
29. A. Rényi, "Probability Theory," North–Holland, Amsterdam, 1970.
30. A. Sklar, *Kybernetica* **9** (1973), 449.
31. B. Schweizer and E. F. Wolff, *C. R. Acad. Sci. Paris Sér A* **183** (1976), 659.
32. G. Rauzy "Propriétés statistiques de suites arithmétiques," Le Mathématicien, No. 15, Collection SUP, Presses Universitaires de France, Paris, 1976.
33. A. Sklar, *Aequationes Math.* **4** (1970), 244.
34. T. Erber, P. Everett, P. Johnson, and A. Sklar, *SIAM Rev.* **19** (1977), 380.
35. R. L. Adler and M. H. McAndrew, *Trans. Amer. Math. Soc.* **121** (1966), 236.
36. B. L. Osofsky, *Notices Amer. Math. Soc.* **23** (1976), 422.
37. H. Niederreiter, *Math. Comp.* **26** (1972), 793.
38. T. Erber and A. Sklar, *in* "Modern Developments in Thermodynamics" (B. Gal-Or, Ed.), Israel Univ. Press and Wiley, Jerusalem and New York, 1974; pp. 281–301.

[illegible]

41. G. Benettin, L. Galgani, and J. M. Strelcyn, *Phys. Rev. A* **14** (1976), 2338.
42. J. Moser, "Stable and Random Motions in Dynamical Systems," Annals of Math. Studies, No. 77, Princeton, N. J., 1973.

43. R. L. Adler and T. J. Rivlin, *Proc. Amer. Math. Soc.* **15** (1964), 794.
44. R. E. Rice and A. Sklar, "Fractional Iteration of Polynomial Functions," to be published.
45. R. E. Rice, *Aequationes Math.* **17** (1978), 104.
46. J. W. Gibbs, "Elementary Principles in Statistical Mechanics," Yale Univ. Press, New Haven, Conn., 1902; reprinted by Dover, New York, 1960.
47. R. L. Adler, A. G. Konheim, and M. H. McAndrew, *Trans. Amer. Math. Soc.* **114** (1965), 309.
48. A. S. Wightman, Statistical mechanics and ergodic theory: An expository lecture, *in* "Statistical Mechanics at the Turn of the Decade" (E. G. D. Cohen, Ed.), pp. 1–29, Dekker, New York, 1971.
49. G. Benettin and J. M. Strelcyn, *Phys. Rev. A* **17** (1978), 773.
50. Ya. B. Piesin, *Dokl. Akad. Nauk. SSSR* **226** (1976), 774; *Soviet Math. Dokl.* **17** (1976), 196.
51. A. N. Kolmogorov, *Dokl. Akad. Nauk. SSSR* **119** (1958), 861.
52. M. S. Ranade, "Ergodic Properties of Polynomials," Doctoral dissertation, I.I.T., 1974.
53. C. E. Shannon, *Bell Syst. Tech. J.* **28** (1949), 656.
54. W. Diffie and M. E. Hellman, *IEEE Trans. Information Theory* **IT-22**, No. 6 (1976), 644; M. E. Hellman, *IEEE Trans. Information Theory* **IT-23**, No. 3 (1977), 289.
55. A. Sklar, private communication.
56. D. S. Ornstein, *Bull. Amer. Math. Soc.* **77** (1971), 878.
57. J. McLaughlin. *J. Statist. Phys.* **15** (1976), 307.
58. S. A. Kauffman, *J. Theoret. Biol.* **22** (1969), 437.
59. T. Y. Li and J. A. Yorke, *Amer. Math. Monthly* **82** (1975), 985.
60. A. N. Kolmogorov, *in* "Proceedings of the International Congress on Mathematics, Amsterdam," Vol. 1, North–Holland, Amsterdam, 1957; Appendix D of R. Abraham, "Foundations of Mechanics," Benjamin, New York, 1967.
61. L. Markus and K. R. Meyer, "Generic Hamiltonian Dynamical Systems Are neither Integrable nor Ergodic," *Mem. Amer. Math. Soc.* No. 114 (1974).
62. A. Beck, "Continuous Flows in the Plane," Springer–Verlag, New York, 1974.
63. A. Sklar, "Fractional Iteration and the Embedding of Functions in Flows," to be published.
64. C. Lanczos, "The Variational Principles of Mechanics," Univ of Toronto Press, Toronto, 1962.
65. J. P. Heller, *Amer. J. Phys.* **28** (1960), 348.
66. P. Walters, *J. London Math. Soc.* **44**, Part 1 (1969), 7.
67. G. M. Zaslavsky and B. V. Chirikov, *Usp. Fiz. Nauk* **14** (1972), 195; *Sov. Phys. Usp.* **14** (1972), 549.
68. T. Erber, S. A. Guralnick, and H. G. Latal. *Ann. Phys. (N.Y.)* **69** (1972), 161.
69. S. W. McDonald and A. N. Kaufman, *Phys. Rev. Letters* **42** (1979), 1189.
70. R. Shaw, "Strange Attractors, Chaotic Behavior, and Information Flow" (preprint, 1978; Santa Cruz).