Algorithm for normal random numbers

Julio F. Fernández
Instituto de Ciencia de Materiales de Aragón, Consejo Superior de Investigaciones Científicas, and Universidad de Zaragoza, 50009-Zaragoza, Spain

Carlos Criado
Departamento de Física Aplicada I, Universidad de Málaga, 29071-Málaga, Spain
(Received 27 April 1999)

We propose a simple algorithm for generating normally distributed pseudorandom numbers. The algorithm simulates \( N \) molecules that exchange energy among themselves following a simple stochastic rule. We prove that the system is ergodic, and that a Maxwell-like distribution that may be used as a source of normally distributed random deviates follows in the \( N \to \infty \) limit. The algorithm passes various performance tests, including Monte Carlo simulation of a finite two-dimensional Ising model using Wolff’s algorithm. It only requires four simple lines of computer code, and is approximately ten times faster than the Box-Muller algorithm. [S1063-651X(99)04409-8]

PACS number(s): 02.70.Lm, 02.90.+p

I. INTRODUCTION

Pseudorandom number (PRN) generation is a subject of considerable current interest [1]. Deterministic algorithms lead to undesirable correlations, and some of them have been shown to give rise to erroneous results for random-walk simulations [2], Monte Carlo (MC) calculations [3,4], and growth models [5]. Most of the interest has been focused on PRN’s with uniform distributions. Less attention has been paid to nonuniform PRN generation.

Sequences of random numbers with Gaussian probability distribution functions (PDF’s) are needed to simulate computers Gaussian noise that is inherent to a wide variety of natural phenomena [6]. Their usefulness transcends physics. For instance, numerical simulations of economic systems that make use of so-called geometric Brownian models (in which noise is multiplicative) also need a source of normally distributed PRN’s [7]. There are several algorithms available for PRN’s with Gaussian PDF’s [8]. Some, such as the Box-Muller’s algorithm, require an input of uniform PRN’s, and their output often suffers from the pitfalls of the latter [9]. Robustness is therefore a relevant issue. In addition, Box-Muller’s algorithm is slow and can consequently consume significant fractions of computer-simulation times [10].

The comparison method demands several uniform PRN’s per normal PRN, and is therefore also slow [11]. Use of tables [12] is not a very accurate method. Algorithms that are related, but not equivalent, to the one we propose here have been published [10,13], but they are somewhat cumbersome to use. In addition, no proof of their validity has been given.

We propose here an algorithm for the generation of normally distributed PRN’s that is quite simple and fast. It is a stochastic caricature of a closed classical system of \( N \) particles. Their velocities provide a source of PRN’s. We prove that, for any initial state, their PDF becomes Maxwellian in the \( N \to \infty \) limit, after an infinite number of two-particle “collisions” takes place. To this end, we first prove that our system is ergodic [14,15]. The proof is not exceedingly difficult because our system is not deterministic. We also study its output as a function of \( N \), and establish useful criteria for its implementation. Correlation test results are also reported.

II. ALGORITHM

A. Motivation

Consider real numbers \( v_1, v_2, \ldots, v_N \), placed in \( N \) computer registers, in analogy to the velocities of \( N \) particles that make up a closed classical system in one dimension. Pairs of registers \( i \) and \( j \), say, selected at random without bias, are to “interact” somehow, conserving the kinetic energy, that is, quantity \( v_i^2 + v_j^2 \). The motivation for imposing this conservation rule comes from the foundations of statistical physics. The statistical distribution function of a system (say, any one of \( N \) particles) in equilibrium with a macroscopic system is an exponential function of the additive constants of the motion (the kinetic energy in our case) if the two systems are statistically uncorrelated [16]. By analogy with the approach to equilibrium that is believed to take place in statistical physics, we therefore expect that a sufficient number of iterations will lead to a Maxwellian (i.e., Gaussian) PDF of register values, from which the desired PRN’s may be drawn. (See, also Ref. [10].) We define below the simplest interaction we can think of in order that (i) implementation on a computer be very fast, and (ii) that we may be able to prove that a Gaussian PDF does indeed ensue.

B. Procedure

Before the algorithm is implemented, all \( N \) registers must be initialized to, say, \( v_i = 1 \) for all \( i \) satisfying \( 1 \leq i \leq N \), or all \( v_i \) may be read from a set of \( N \) register values saved from a previous computer run, which we assume to fulfill \( \sum v_i^2 = N \). Let \( U(1,N) \) and \( U_i(1,N) \) be unbiased integer random variables, both in the interval \([1,N]\), except that \( U_i \) cannot equal \( i \). The algorithm follows:

\[
\begin{align*}
  i &= U(1,N); \quad j = U_i(1,N), \\
  v_i &= (v_i + v_j)/\sqrt{2}, \quad \text{(2.1)}
\end{align*}
\]
The updated value of $v_i$, from Eq. (2.2), is used in Eq. (2.3). After an initial warmup phase (see below), $v_i$ and $v_j$ may be drawn each time transformations (2.1)–(2.3) are applied. These are the two desired PRN’s, with $\langle v_i \rangle = 0$ and $\langle v_i^2 \rangle = 1$ for all $i$. Their PDF becomes Gaussian (see below) in the $N \to \infty$ limit.

Transformation (2.1)–(2.3) may be thought of as a rotation of $\pm\pi/4$ with respect to a randomly chosen $ij$ plane ($\pm\pi/4$ and $-\pi/4$ are for the two possible index orderings, $ij$ and $ji$, respectively).

Numerical evidence that the system becomes nonergodic if $j=U[1,N]$ is replaced by $j=1+1$ mod $N$ in Eq. (2.1) is easily obtained. This remark may help to place the significance of the proof that follows into proper perspective.

III. PROOF

Let $P_n(v)$ be the probability density at $v = (v_1, v_2, \ldots, v_N)$, after transformation (2.1)–(2.3) has been applied $n$ times. We prove below, in three stages, that $P_n(v) \to \text{constant}$ over spherical surface $S_{N-1}$, if $N \to \infty$, in the $n \to \infty$ limit. We first prove $P_n(v) \to P_n(u)$ as $n \to \infty$ if $v$ and $u$ are related. From here on, we say that points $v$ and $u$ are related if successive transformations (2.1)–(2.3) of $v$ can lead to $u$. We then prove that the system’s “orbit” covers $S_{N-1}$ densely [that is, that any point $v \in S_{N-1}$ can be brought arbitrarily close to any other point $u \in S_{N-1}$ by applying transformations (2.1)–(2.3) to $v$ a sufficient number of times]. Then, the desired result follows easily.

To start the proof, let kernel $K(v,v')$ be defined by $P_{n+1}(v) = \int K(v,v') P_n(v') \, dv'$, and let

$$F_n = \int \{P_{n+1}(v) - P_n(v)\} \, dv.$$  \hspace{1cm} (3.1)

Note first that $F_n < 0$ implies that $P_{n+1}(v)$ is more uniform than $P_n(v)$, in the sense that $\int dv \, [P_{n+1}(v) - P] \leq 0$. This follows from the definition of $K(v,v')$

$$F_n = \int dv \left[ \int dv_1 K(v,v_1) P_n(v_1) - P_n(v) \right].$$  \hspace{1cm} (3.2)

Making use of the detailed balance condition, $K(v,v') = K(v',v)$, that our system satisfies, and the relation $\int dv K(v,v') = 1$, Eq. (3.2) can be cast into

$$F_n = -\frac{1}{2} \int dv \int dv_1 \int dv_2 \, Q(v,v_1,v_2),$$  \hspace{1cm} (3.3)

where $Q = K(v,v_1)K(v,v_2)P_n(v_1)P_n(v_2)$. Therefore, in the $n \to \infty$ limit, $P_n(v)$ becomes constant over each set in $S_{N-1}$ within which any two points $v$, $u$ are related.

We next prove that the orbit of any point $p$ in $S_{N-1}$ under the action of the transformation group defined by Eqs. (2.1)–(2.3) covers the sphere densely. Let $H_N$ denote the subgroup of $SO(N)$ corresponding to transformations (2.1)–(2.3). To show that the orbit $H_N(p)$ of any $p$ in $S_{N-1}$ is dense in $S_{N-1}$, we have to prove that for any $q$ in $S_{N-1}$ and any real number $e > 0$ there exists an element $h$ in $H_N$ such that $h(p)$ is within distance $e$ of $q$, that is, that $d(h(p),q) < e$. For this purpose, it is sufficient to prove that $H_N$ is dense in $SO(N)$, by which we mean that for any $g$ in $SO(N)$ and any real number $e > 0$ there exists an element $h$ in $H_N$ such that $d_{SO(N)}(g,h) < e$, where $d_{SO(N)}$ is the distance between elements $g$ and $h$ and is defined to be the supremum of the distances on $S_{N-1}$, $d(g,p,h(p))$, for $p$ varying in $S_{N-1}$. To show that this is indeed sufficient, let $g$ in $SO(N)$ and $h$ in $H_N$ be such that $q = g(p)$ and $d_{SO(N)}(g,h) < e$. The desired relation, $d(h(p),q) < e$, then follows from the definition of $d_{SO(N)}(g,h)$.

We now prove for $N = 3$ that $H_N$ is dense in $SO(3)$. The proof is extended to higher dimensions by induction. Note first that $H_3$ does not belong to the set of finite rotation groups in three dimension [17], and is therefore an infinite group. Let group $SO(3)$ be covered by disks of radius $\epsilon/2$ each. A finite number of them is sufficient, since $SO(3)$ is compact [18]. It follows that there must be at least one disk with two elements of $H_3$ in it, since $H_3$ has an infinite number of elements. Let these two elements be $r$ and $s$, and let $g(u,e)$ be element $rs^{-1}$ of $H_3$, which is a rotation by an angle smaller than $e$ about some undetermined $u$ axis. We will build elements of $H_3$ that are as arbitrarily close to any given rotation. To this end, it is sufficient to show that it can be done for a set of infinitesimal generators of rotations [19]. One such set is made up of infinitesimal rotations about three linearly independent axes. Consider axes $u_1$, $u_2$, and $u_3$ that are obtained from $u$ by rotations $g(1,\pi/2)$, $g(2,\pi/2)$, and $g(3,\pi/2)$ about each one of the coordinate axes by angle $\pi/2$. The corresponding infinitesimal rotations are given by [18], $g(u,e) = g(i,\pi/2)g(u,e)g^{-1}(i,\pi/2)$. This concludes the proof for three dimensions.

We now prove by induction that for any element $g(i,\alpha)$ of the rotation about plane $ij$ by angle $\alpha$ of the rotation group $SO(N)$, there exists an element $g$ of group $H_N$ that lies arbitrarily close to it, for $N \geq 3$. By hypothesis, any $g(i,\alpha)$, for $i,j=1,2,\ldots,N$, can be approximated by an element $g$ of $H_N$. We show now that $g(i,N+1,\alpha)$, for $i=1,2,\ldots,N$, can also be approximated by elements of $H_{N+1}$. We take $g \in H_N$ within distance $e$ of $g(i,\alpha)$. Now, since rotations preserve distances, it follows that $g(i,N+1,\alpha) = g(i,N+1,\pi/2)g(i,\alpha)g^{-1}(i,N+1,\pi/2)$ is within distance $e$ of $g' \in H_{N+1}$, given by $g' = g(i,N+1,\pi/2)g^{-1}(i,N+1,\pi/2)$. This proves dense coverage in $N \geq 3$ dimensions. This is a kind of stochastic generalization of Jacobi’s theorem [15] to more than two dimensions.

To conclude the proof that $P_n(v) \to \text{constant}$ in the $n \to \infty$ limit, consider any two points $V$ and $U'$ as centers of disks $D(V)$ and $D(U')$, both of radius $r$, in $S_{N-1}$. Since the system’s orbit covers $S_{N-1}$ densely for $N \geq 3$, it follows that a point $U$ that is related to $V$ exists arbitrarily close to $U'$. Consider now two equal size disks $D(V)$ and $D(U)$. The fact that there exists at least one sequence of rotations in $H_N$ that take $V$ to $U$ implies that there exists a rotation $g$ in $H_N$ that transforms $V$ into $U$. Since $g$ is a rotation, it transforms $D(V)$ rigidly into $D(U)$. It follows that $\int dV dV' P_n(v) \to \int dV dU P_n(u)$, as $n \to \infty$. Since $r$ is arbitrary, and $V$ and $U'$ are any two points in $S_{N-1}$, it follows that $P_n(v)$
constant over $S_{N-1}$ (except, perhaps, on a set of measure zero). Ergodicity follows [15].

The desired result follows easily. Let the single register PDF $p(v)$ be the $n \to \infty$ limit of $p_n(v)$, where $p_{n}(v_i) = \int p_n(v) d\nu_2 d\nu_3 \ldots d\nu_N$. Since $p_n(v) \to constant$ as $n \to \infty$, it follows that

$$p(v) \propto \left(1 - \frac{v^2}{N}\right)^{(N-3)/2}.$$  (3.4)

Clearly, $p(v) \to C \exp(-\sqrt{2}v/2)$ in the $N \to \infty$ limit, which is the desired result.

IV. NUMERICAL TESTS

We now address the following practical issues: (i) how good an approximation to a Gaussian PDF of PRN’s is achieved with a necessarily finite set of $N$ registers, (ii) how long must the initial warm-up phase be, (iii) period length, and (iv) what correlations, if any, are found numerically.

A. Distribution

Frequencies of events from sequences of $10^6$, $10^8$, and $10^{10}$ PRN’s generated with transformation (2.1)–(2.3), with $N = 1024$, are exhibited in Fig. 1. In order to determine how many PDF’s obtained for finite $N$ deviate from the desired Gaussian distribution, it is convenient to rewrite Eq. (3.4) as follows,

$$p(v) \propto e^{-v^2/2} e^{g_N(v)/N},$$  (4.1)

where $g_N(v) = v^2(3 - v^2/2)/2 + O(1/N)$. $N^{-1}g_N(v)$ is approximately the fractional deviation, $\delta p(v)/p(v)$, from Gaussian form if $\delta p(v)/p(v) \ll 1$. We have checked this behavior numerically. The results obtained are shown in Fig. 2. Clearly, the number of registers $N$ that must be used increases with the number $M$ of PRN’s one intends to generate. This is because the value of the largest PRN generated increases, on the average, with $M$. More precisely, the value of $v$ beyond which PRN’s are only generated with probability $q$ is approximately given by $v^2 = 2 \ln(M/q\delta)$. Now, it follows from Eq. (4.1) that the fractional error $\delta p/P$ in the probability density at $v$ is approximately $N^{-1}v^2(3 - v^2/2)/2$ for very large $N$. (It is pointless to require this error to be too small since a PRN is expected to be generated beyond $x$ with a small probability $q$.) It then follows that $[\ln(M/q\delta)]^2 \leq n \delta I/P$ must be satisfied by $N$. Thus, approximately $10^4$ registers are sufficient in order to generate as many as $10^{15}$ PRN’s, with roughly a 10% error in the probability for the largest PRN in the sequence.

B. Warmup

Our algorithm must be applied a number $n_p N$ of times before it is ready for use unless all $v_i$ are initialized with ‘equilibrium’ values (stored from some previous computer run). The distribution of all register values then evolves toward equilibrium, as illustrated in Fig. 3. Deviations from equilibrium are statistically insignificant for $n_p \geq 2$ and $N = 1024$, and for $n_p \geq 4$ and $N = 1048 576$. Since $n_p$ is expected to increase as $\ln N$, $n_p = 8$ should provide ample warmup for any foreseeable applications.

C. Recurrence

The number of PRN’s that must be generated before each PRN in sequence $v_1, v_2, \ldots, v_N$ returns within distance $r$ from its initial value is exponential in $N$. More specifically, we estimate it to be $(\tau/\sqrt{N}) (1/r)^N$ for $N \gg 1$, where $\tau$ is the period of the algorithm used to select $i$ and $j$ in Eq. (2.1). The estimation is based on $P_n(v) \to constant$ over $S_{N-1}$ as $n \to \infty$. We have numerically checked this for small values of $N$. The data points shown as $\bullet$, $\bigcirc$, and $\bigtriangleup$ are for $n_p = 2, 4$, and 10, respectively, for $N = 1048 576$. The two straight lines stand for $C \exp(-\sqrt{2}v/2)$ for two values of $C$. Figure 3.
Fig. 4. Average energy per spin, obtained from MC simulations using Wolff’s algorithm, versus the inverse of the number of registers used for the generation of PRN’s with Gaussian PDF’s. The ●, ■, and △ stand for data points that follow from feeding our algorithm with the following uniform PRN generators: GGL, r(250,103,XOR), and RAN3, respectively. Unacceptable energy values that have been obtained in Refs. [3] using r(250,103,XOR), and RAN3 are also shown as bars next to the y axis.

N. Thus, an effectively infinite recurrence time follows for any reasonably large value of N.

D. Correlations

Correlations between a finite number of PRN’s clearly vanish as $N \to \infty$, since $i$ and $j$ in Eq. (2.1) are supposedly independent PRN’s. We have searched for correlations in $m$ successively generated PRN’s $v_1, v_2, \ldots, v_m$, for $m = 3, 4, \ldots, 6$, performing a chi-square isotropy test over the corresponding m-dimensional space. An $m$-tuple $v = v_1, v_2, \ldots, v_m$ was said to belong to the $i$th cone of 1024 randomly oriented cones with axes $w_1, w_2, \ldots, w_{1024}$, if $0.99 \leq v \cdot w \leq 1$. No significant deviations from isotropy were observed for $10^6$ generated $m$ tuples.

Implementation of Wolff’s algorithm [20] in MC calculations of the Ising model’s critical behavior is a demanding test that some well-known uniform PRN generators have failed [3]. Large clusters are then flipped as a whole, and this tests correlations in very long sequences. We have used normal PRN’s generated by our algorithm as input into a MC simulation of an Ising system of $16 \times 16$ spins at the critical temperature. [For that, we note that $v_i^2 + v_j^2 > 2x$ as often as $u > \exp(-x)$ if $v_i$ and $v_j$ (u) are PRN’s with Gaussian (uniform) PDF’s, respectively.] The energy obtained is shown in Fig. 4 as a function of the number of registers $N$. The following uniform PRN algorithms were used to select $i$ and $j$ in Eq. (2.1): GGL [3], r(250,103,XOR) [2,3], and RAN3 [21].

We tried the latter two algorithms, which have been shown to lead by themselves to unacceptable results for the Ising model [3], in order to test our algorithm’s robustness. The results shown in Fig. 4 are gratifying.

Similarly, the specific heat $c$ and magnetization $m$ fluctuations data points obtained follow approximately the relations $c = c_0 + 8.4/N$, and $(\delta m)^2 = \chi_0 + 33/N$, respectively, where $c_0 = 1.497(1)$ and $\chi_0 = 0.5454(2)$, in agreement with the known exact values [3,22].

V. CONCLUSIONS

Double precision is recommended. It prevents excessive drift of the sum $\Sigma v_i^2$ away from its assigned value. Even then, single precision accuracy is to be expected at the end of a sequence of some $10^{16}$ PRN’s, unless the sum is normalized several times during the run.

In summary, we have shown that implementation of Eqs. (2.1)–(2.3) provides a source of PRN’s with an approximately Gaussian PDF. Some $10^4$ registers (molecules) are sufficient for some purposes, but up to $10^5$ or more may be necessary for more demanding tasks. (Having to make a decision about the number of registers to be used may sometimes be an unwelcome task. On the other hand, it is a virtue of the algorithm that one can control, through the value of $N$, how close the output is to be from sequences of truly independent random numbers with Gaussian PDF’s.) Initial warmups for arbitrary initial conditions are necessary; it is sufficient to let each register initially interact an average number of, say, eight times. The system’s recurrence time was shown to be exponential in $N$, and therefore effectively infinite. Its behavior appears to be robust. The proposed algorithm runs an order of magnitude faster on computers than the most often used Box-Muller method [8,9]. For a fortran code of our algorithm or other questions, please write JFF@Pipe.Unizar.Es.

ACKNOWLEDGMENTS

Continuous help from Dr. Pedro Martínez with computer systems is deeply appreciated by J.F.F. We are indebted to Professor P. Grassberger for a leading comment. One of us (J.F.F.) is indebted to Dr. Juan Rivero for valuable help received while visiting the Instituto Venezolano de Investigaciones Científicas (IVIC), on a travel grant from a CSIC-IVIC agreement. J.F.F. and C.C. are grateful for financial support from DGESIC of Spain, through Grant Nos. PB95-0797 and PB97-1080, respectively.

[6] For old work, see N. Wax, Noise and Stochastic Processes
(Dover, New York, 1954); for a more recent general account, see C.W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1990).


