LINEAR RECURRENCES WITH CARRY AS UNIFORM RANDOM NUMBER GENERATORS

Raymond Couture
Pierre L'Ecuyer

Département d'IRO
Université de Montréal, C.P. 6128, Succ. Centre-Ville
Montréal, H3C 3J7, CANADA

ABSTRACT

We study the multiply-with-carry family of generators proposed by Marsaglia as a generalisation of the previous add-with-carry and subtract-with-borrow families of Marsaglia and Zaman (1991). We define for them a general (infinite) state space and focus our attention on the (finite) subset of recurrent states. This subset will, in turn, split into possibly several subgenerators. We discuss the uniformity of the d-dimentional distribution of the output of these subgenerators over their full period. In order to improve this uniformity for higher dimensions, we propose a method for finding good parameters in terms of the spectral test.

1 THE MWC GENERATORS

Marsaglia (1994) proposed an extension of the add-with-carry (AWC) and subtract-with-borrow (SWB) families of uniform random number generators. The new generators are called multiply-with-carry (MWC). Let r and b be positive integers. Then, a MWC of *order* r, and *base* b, is defined as follows. Consider the set Σ of those $\sigma = (x_{-1}, \ldots, x_{-r}, c) \in \boldsymbol{Z}^{r+1}$ with $0 \leq x_i < b$ for $-r \leq i \leq -1$. This set is the state space of the generator. We refer to c as the *carry component* of the state σ. In state σ, the generator outputs the pseudo-random number $x_{-1}/b \in [0, 1)$ and then evolves into another state $\sigma' = (x'_{-1}, \ldots, x'_{-r}, c') \in \Sigma$ determined by the conditions

$$x'_i = x_{i+1} \quad \text{for} \quad -r \leq i < -1, \tag{1}$$

$$x'_{-1} + c'b = a_1 x_{-1} + \ldots + a_r x_{-r} + c, \tag{2}$$

where the coefficients a_l are suitably-chosen fixed integers. Note that the integers x'_{-1} and c' are uniquely determined from (2) since we must have $0 \leq x'_{-1} < b$, and therefore x'_{-1} is the least positive residue of $a_1 x_{-1} + \ldots + a_r x_{-r} + c$ modulo b. Marsaglia (1994) also proposes to take $b = 2^w$, a power of 2, so that x'_{-1} and c' are easily obtained (on a binary computer). We put $a_0 = -1$,

$$m = \sum_{l=0}^{r} a_l b^l,$$

and we assume throughout that $m \neq 0$.

In this paper, we study the d-dimensional uniformity of the output of MWC generators. Results are stated without proof; further developments and proofs will be given in a forthcoming extended version of the paper. We first show in section 2, the importance, in this respect, of the subset of recurrent states and we give a characterization for these states. Lemma 1 is a generalization of a statement of Marsaglia and Zaman (1991) for the case of the AWC/SWB generators. This lemma is further generalized to arbitrary dimension in Section 3, by connecting the MWC generators with linear congruential generators (LCG's). This connection was investigated by Tezuka, L'Ecuyer, and Couture (1994) and Couture and L'Ecuyer (1994), again in the AWC/SWB case. We also distinguish two aspects of the question of d-dimensional uniformity, requiring different methods. These are discussed in Secions 4 and 5 respectively. We see in Section 4 that the problem leads to some arithmetical questions; these will be discussed elsewhere. In Section 5, the method used is the well-known spectral test. We examine in this respect some specific instances proposed by Marsaglia (1994). We also indicate a method to search for parameters, which are good with respect to this test. For general references on random number generation, the reader can consult, e.g., Knuth (1981), L'Ecuyer (1994) and Niederreiter (1992).

2 RECURRENT STATES

Consider the transformation $T : \Sigma \to \Sigma$ defined by $T(x_{-1}, \ldots, x_{-r}, c) = (x'_{-1}, \ldots, x'_{-r}, c')$, subject to (1) and (2). A state $\sigma \in \Sigma$ is *recurrent* (with respect to T) if $T^n(\sigma) = \sigma$, for some positive integer n. According to the theorem below, any state will quickly evolve into the set Σ_{r} of recurrent states. We define the function $\delta : \Sigma \to \boldsymbol{R}$ by

$$\delta(x_{-1}, \ldots, x_{-r}, c) = c - \sum_{i=-r}^{-1} \sum_{l=0}^{r+i} a_l x_{i-l} b^i.$$

Let Σ' be the set of states σ for which

$$0 \le \delta(\sigma)/m < 1/b^r. \qquad (3)$$

and put $\varsigma_1 = (b-1, \ldots, b-1, a_0 + \ldots + a_r)$. Note that $T(\varsigma_1) = \varsigma_1$, and that $\delta(\varsigma_1) = m/b^r$. We thus have $\varsigma_1 \in \Sigma_{\mathrm{r}}$ but ς_1 is not contained in Σ'. There is only one other state in Σ fixed by T, namely $\varsigma_0 = (0, \ldots, 0)$, and it is contained in Σ'.

Theorem 1 *For any state $\sigma \in \Sigma$, we have $T^n(\sigma) \in \Sigma_{\mathrm{r}}$ if the non-negative integer n exceeds*

$$\max(0, \log_b |\delta(\sigma)| - \log_b |m| + r) + \max(r, \log_b |m|) + 1.$$

Moreover, $\Sigma_{\mathrm{r}} = \Sigma' \cup \{\varsigma_1\}$.

It follows from this characterisation of Σ_{r} that the real interest centers on the set Σ' rather than on Σ_{r} itself. Clearly, T is an invertible transformation of Σ'. In the next section, we will give another description of the set Σ' which will allow a further study of the action of T on it. The following two lemmas are easy consequences of (3).

Lemma 1 *Given any r-tuple $(x_{-1}, \ldots, x_{-r}) \in \boldsymbol{Z}^r$, with $0 \le x_i < b$ for $-r \le i \le -1$, the number of values of c such that $(x_{-1}, \ldots, x_{-r}, c) \in \Sigma'$ differs from $|m|/b^r$ by less than 1.*

This first Lemma can be interpreted as a result on the uniformity of the output r-tuples of successive values over the whole Σ'. It will be further generalized in the next section. The next lemma will be used in Section 5.

Lemma 2 *If $a_l \ge 0$ for $1 \le l \le r$, then the carry component c of a state in Σ' satisfies $0 \le c < \sum_{l=1}^{r} a_l$. These inequalities are best possible if $m > b^r$.*

3 ORBIT STRUCTURE

Since T is invertible on the set Σ', the latter will split, in general, into a number of disjoint subsets—the *orbits*—on each of which the action of T is transitive. Each of these orbits defines a different random number generator, which we may refer to as a *subgenerator*. Note that ς_0 constitutes an orbit by itself. The subgenerator defined by it is trivial.

Put $\boldsymbol{Z}_m = \{k \in \boldsymbol{Z} \,|\, 0 \le k/m < 1\}$. Define the mapping $S : \boldsymbol{Z}_m \to \boldsymbol{Z}_m$ by $S(k) = k'$ where $k' \in \boldsymbol{Z}_m$ is subject to $bk' \equiv k \pmod m$. The following theorem allows to reduce the study of the action of T on Σ' to that of S on $\boldsymbol{Z}_m$, which is well known.

Theorem 2 *There exists a mapping $\iota : \boldsymbol{Z}_m \to \Sigma$, uniquely determined by the following two properties.*

1) *The transformation S is mapped by ι into T, that is, we have $\iota(S(k)) = T(\iota(k))$ for $k \in \boldsymbol{Z}_m$.*

2) *If $k \in \boldsymbol{Z}_m$, we have $y_{-1}/b \le k/m < y_{-1}/b + 1/b$, where y_{-1} is the first component of $\iota(k)$.*

The mapping ι, subject to the above conditions, is then one to one, and its image is $\iota(\boldsymbol{Z}_m) = \Sigma'$. Also, $\iota(0) = \varsigma_0$.

We have a similar decomposition of $\boldsymbol{Z}_m$, with respect to the tranformation S into a set of orbits, which we may call S-orbits. The set of S-orbits in $\boldsymbol{Z}_m$ correspond by ι to the set of T-orbits of Σ'. Consider now, one such T-orbit, and the subgenerator it defines. Let d be any positive integer, and let $y_{-1}, \ldots, y_{-d}$ be given integers in $\{0, \ldots, b-1\}$. It then follows from Theorem 2, that the number of times the output d-tuples of successive values of this subgenerator assume the value $(y_{-d}/b, \ldots, y_{-1}/b)$ over its full period is equal to the number of integers k in the corresponding S-orbit satifying $\sum_{i=-d}^{-1} y_i b^i \le k/m < \sum_{i=-d}^{-1} y_i b^i + b^{-d}$.

The question of the distribution of the output d-tuples of subgenerators is thus equivalent to the question of the distribution of the S-orbits in $\boldsymbol{Z}_m$, into intervals of length $|m|/b^d$.

We may now distinguish two cases, according to whether $|m|/b^d$ is larger or smaller than 1. If it is smaller than 1, each output d-tuple can appear only once, and we are then interested in the set formed by these d-tuples. When it is larger than 1, it more a question of the frequencies of given output values. We may refer to these two cases as the *large* and *small interval* cases.

4 LARGE INTERVALS

We assume, from now on, that the coefficients a_l have been chosen in such a way that m is prime.

The simplest case arises if b is a primitive root modulo m. Then the action of T on $\Sigma' \backslash \{\varsigma_0\}$ is transitive, and we have, besides $\{\varsigma_0\}$, a single orbit. It then follows from the previous section that, when $|m|/b^d$ is large, the d-tuples of succesive output values of the subgenerator defined by $\Sigma' \backslash \{\varsigma_0\}$, will be equal essentially equally often, over the full period, to any given d-tuple of numbers of the form y/b, where $y \in \{0, \ldots, b-1\}$.

If $b = 2^w$ with w an integer greater than 2, then the Legendre symbol

$$\left(\frac{2}{m}\right) = (-1)^{(m^2-1)/8} = 1,$$

and therefore b is a quadratic residue modulo m. Thus b cannot be a primitive root modulo m, though it may still generate the group of all quadratic residues. Assuming this is so, the set $\Sigma' \backslash \{\varsigma_0\}$ splits into two orbits of the same cardinality $(m-1)/2$, corresponding respectively to the subsets in $\boldsymbol{Z}_m \backslash \{0\}$ of quadratic residues and non-residues modulo m, and defining two non-trivial subgenerators. For both of these subgenerators, the question whether the r-tuples of successive outputs are well distributed amounts to the following question. Are there, generally, in an interval of length $|m|/b^r$, (roughly) as many quadratic residues as non-residues modulo m. Since m will be taken usually very large, one cannot answer by merely listing all quadratic residues. It is generally believed that the quadratic residues modulo a prime are fairly uniformly distributed. However results in this respect are scarce, and not easy to obtain. The question is all the more difficult as the length of the interval is smaller. The simplest case is for the interval $(0, |m|/2)$, and is equivalent to asking if the most significant binary digit of the output of the subgenerators, is as often 0 as 1 over the full period.

5 SMALL INTERVALS

We now examine the distribution of d-tuples of successive outputs of a (sub)generator in a case where $|m|/b^d < 1$. In this circumstance, any such d-tuple will appear only once in the full period, and we are led to study them as a set, rather than a set with multiplicities.

We assume, in this section, that

(i) the basis is a power of two: $b = 2^w$, with $w \in \boldsymbol{Z}$ greater than 2,

(ii) all coefficients a_l, $l = 1, \ldots, r$ are non-negative, the greater weight being given to the leading coefficient a_r, so that $a_r > a_l \geq 0$ for $l = 1, \ldots, r-1$, and

(iii) the carry component c of all recurrent states satisfies $0 \leq c < b$, which, according to Lemma 2, amounts to the inequality $\sum_{l=1}^r a_l \leq b$. We will also exclude the trivial case where $a_r = b$ (and therefore $a_l = 0$, $l = 1, \ldots, r-1$).

The exponent w in assumption (i) will normally be the computer's word length. Thus, all components of a state, save the carry, can each be stored in one word. Assumption (iii) guarantees that the carry can also be stored in one word and, therefore, that the sum on the right-hand-side of (2) can be accumulated in a double word register.

Let d be any positive integer. We denote by $e_1, \ldots, e_d$, the canonical basis in $\boldsymbol{R}^d$. Let Λ_d be the lattice in $\boldsymbol{R}^d$ generated by $v^* = 1/m \sum_{j=1}^d b^{d-j} e_j$ and $\boldsymbol{Z}^d$. The intersection of Λ_d with $[0,1)^d$ is then precisely the set of d-tuples $(k/m, \ldots, S^{d-1}(k)/m)$, with $k \in \boldsymbol{Z}_m$. By Theorem 2, the study of the distribution of the set of d-tuples of successive outputs over Σ' is by large reduced to the study of the lattice Λ_d (we use here Σ' instead of the orbits contained in it only for simplicity's sake, since in our main applications below, consideration of these orbits would lead to the same lattice Λ_d). Let $\Lambda^{(d)}$ denote the lattice dual to Λ_d.

If $w \in \Lambda^{(d)} \backslash \{0\}$ and $n \in \boldsymbol{Z}$, then the region $\{v \in \boldsymbol{R}^d \mid n < v \cdot w < n+1\}$ is the set of points between two parallel hyperplanes, apart by a distance of $1/\|w\|$, and it contains no point of Λ_d. We are thus concerned with the presence of small vectors in $\Lambda^{(d)}$ as they produce wide gaps in the distribution of points of Λ_d.

Assume now that $d > r$. This is the small interval case in this situation. Put

$$w_1 = \sum_{j=d-r}^{d} a_{d-j} e_j.$$

We will also write $w_1^{(d)}$ when it is convenient to indicate the dimension. We have, in the case $d = r+1$,

Theorem 3 *The non-zero vectors of minimal length in $\Lambda^{(r+1)}$ are the vectors $\pm w_1$.*

(Note that here one can replace assumption (i) 1 above, by the condition that $b \geq 6$.) The squared length of this vector w_1 is equal to $1 + \sum_{l=1}^r a_l^2$. Thus, a better $(r+1)$-dimensional uniformity is obtained by choosing the coefficients a_l, $l = 1, \ldots, r$, so as to maximize $\sum_{l=1}^r a_l^2$, subject to the conditions in (ii–iii), namely that $0 \leq a_l < a_r < b$ for $l = 1, \ldots, r-1$, and $\sum_{l=1}^r a_l \leq b$. Clearly these conditions imply that $\sum_{l=1}^r a_l^2 < b^2$. Now, requiring good uniformity in still higher dimension imposes further constraints on the choice of the coefficients a_l. In fact, for $d > r+1$, small vectors in $\Lambda^{(d)}$ may arise as follows.

Define $w_j = -e_{j-1} + b e_j$ for $j = 2, \ldots, d$. If $d > r$, then set of vectors $w_1, \ldots, w_d$ is a lattice basis for $\Lambda^{(d)}$, and an arbitrary vector $w \in \Lambda^{(d)}$ can thus be written as $w = \sum_{j=1}^d z_j w_j$ with integer coefficients z_j. Associate with such a w the vector $w' = z_1 w_1 + b \sum_{j=2}^d z_j e_j$. We have then the inequality

$$\|w\| \leq (1 + b^{-1}) \|w'\| + \|w_1\| b^{-1} |z_1|. \qquad (4)$$

Thus, if there exists integers $z_1, \ldots, z_d$, not all zero, such that $|z_1|$ and $\|w'\|$ are small, we obtain a small non-zero vector $w \in \Lambda^{(d)}$. This condition does not depend on the dimension d for $d > r$, and amounts to the existence of a small non-zero integer multiple $z_1 w_1^{(r+1)}$ of $w_1^{(r+1)}$ sufficiently close to a vector of the lattice $b\boldsymbol{Z}^{r+1}$. We illustrate this using two sets of parameters proposed by Marsaglia (1994). Both have $b = 2^{16}$, and $r = 8$. His choice of coefficients makes m and $(m-1)/2$ prime, also in both cases, so that b generates the group of quadratic residues modulo m, and we thus have two non-trivial orbits.

In the first case he chooses $a_1 = 1941$, $a_2 = 1860$, $a_3 = 1812$, $a_4 = 1776$, $a_5 = 1492$, $a_6 = 1215$, $a_7 = 1066$, and $a_8 = 12013$. His second choice is $a_1 = 1111$, $a_2 = 2222$, $a_3 = 3333$, $a_4 = 4444$, $a_5 = 5555$, $a_6 = 6666$, $a_7 = 7777$, and $a_8 = 9272$.

In Table 1, we give the minimum squared length for a non-zero vector $w \in \Lambda^{(d)}$, for dimensions $8 < d < 15$.

Table 1: Spectral test for Marsaglia's examples

d	First example	Second example
9	162 815 416	258 774 925
10	162 815 416	7 917 146
11	57 479 774	4 922 735
12	13 628 741	1 248 822
13	3 545 576	627 603
14	1 311 482	591 467
15	589 430	441 038

We notice, in dimension $d = r+2 = 10$, a minimal length vector smaller by a factor near 5 for the second case relative to the first case. This vector is given by $w_{\min} = 177w_1 - 25w_2 - 21w_3 - 18w_4 - 15w_5 - 12w_6 - 9w_7 - 6w_8$. Its length is approximately equal to 2813.74. The vector $177w_1$ happens to be of least distance to the lattice $b\boldsymbol{Z}^d$ among all vectors $z_1 w_1$ with $z_1 \in \boldsymbol{Z}$, and $0 < |z_1| < 2000$. This distance is approximately 2788.15, and this accounts, in view of the inequality (4), for the presence in the lattice $\Lambda^{(d)}$ of the small vector $w_{\min}$.

It is easy to find coefficients a_l, which will satify the conditions in (ii–iii), and which will make the distance of $z_1 w_1$ to $b\boldsymbol{Z}^d$ much larger than 2788.15 for a wider range of values of z_1. For instance, we found that the choice $a_1 = 16$, $a_2 = 20$, $a_3 = 147$, $a_4 = 1500$, $a_5 = 2083$, $a_6 = 5276$, $a_7 = 10551$, and $a_8 = 45539$, gives a minimal distance to $b\boldsymbol{Z}^d$ approximately equal to 18163.47, for the set of vectors $z_1 w_1$, $0 < |z_1| < 3000$. We then found nearby coefficients which further satisfy the conditions that m is prime, and that b generates the group of quadratic residues. They are $a_1 = 14$, $a_2 = 18$, $a_3 = 144$, $a_4 = 1499$, $a_5 = 2083$, $a_6 = 5273$, $a_7 = 10550$, and $a_8 = 45539$. We give in Table 2 the minimum squared length for a non-zero vector $w \in \Lambda^{(d)}$, for dimensions $8 < d < 15$, for those coefficients.

Table 2: Spectral test for another example

d	The other example
9	2 219 514 697
10	305 990 559
11	92 513 087
12	18 472 574
13	4 862 652
14	1 910 260
15	705 271

ACKNOWLEDGMENTS

This work has been supported by NSERC-Canada grant # OGP0110050 and FCAR-Québec grant # 93-ER-1654.

REFERENCES

Couture, R., and P. L'Ecuyer. 1994. On the lattice structure of certain linear congruential sequences related to AWC/SWB generators. *Mathematics of Computation*, 62(206):798–808.

Knuth, D. E. 1981. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms.* second ed., volume 2. Addison-Wesley.

L'Ecuyer, P. 1994. Uniform random number generation. *Annals of Operations Research*, 53:77–120.

Marsaglia, G. 1994. Yet another rng. Posted to the electronic billboard `sci.stat.ma`, August 1.

Marsaglia, G., and A. Zaman. 1991. A new class of random number generators. *The Annals of Applied Probability*, 1:462–480.

Niederreiter, H. 1992. *Random Number Generation and Quasi-Monte Carlo Methods.* volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics.* Philadelphia: SIAM.

Tezuka, S., P. L'Ecuyer., and R. Couture. 1994. On the add-with-carry and subtract-with-borrow random number generators. *ACM Transactions of Modeling and Computer Simulation*, 3(4):315–331.

AUTHOR BIOGRAPHIES

RAYMOND COUTURE works as a research assistant at the University of Montreal. He received a Ph.D. in mathematics in 1981, from Laval University, then was at McGill University for three years, under an NSERC post-doctoral scholarship. He has several publications in approximation theory, harmonic analysis, number theory, and pseudorandom number generation.

PIERRE L'ECUYER is a professor in the department of "Informatique et Recherche Opérationnelle" (IRO), at the University of Montreal. He received a Ph.D. in operations research in 1983, from the University of Montreal. From 1983 to 1990, he was with the computer science department, at Laval University, Québec. His research interests are in Markov renewal decision processes, sensitivity analysis and optimization of discrete-event stochastic systems, random number generation, and discrete-event simulation in general. He is the Departmental Editor for the Simulation Department of *Management Science* and an Area Editor for the *ACM Transactions on Modeling and Computer Simulation*.