MATHEMATICS OF COMPUTATION
VOLUME 00, NUMBER 00
Xxxx 19xx, PAGES 000–000

DISTRIBUTION PROPERTIES OF MULTIPLY-WITH-CARRY RANDOM NUMBER GENERATORS

RAYMOND COUTURE AND PIERRE L'ECUYER

ABSTRACT. We study the multiply-with-carry family of generators proposed by Marsaglia as a generalization of the previous add-with-carry families of [7]. We define for them an 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-dimensional 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. Our results are stated in a general context and are applied to a related complementary multiply-with-carry family of generators.

1. INTRODUCTION

Marsaglia and Zaman introduced in [7] the add-with-carry (AWC) and subtract-with-borrow (SWB) families of uniform random number generators which combine both efficiency and very long period. They are all subsumed under the following scheme. We define a recursive *carry* generator of *order* r and *base* b (a positive integer) by means of a function

$$f : \Sigma \to \mathbf{Z}$$

where $\Sigma \subset \mathbf{Z}^{r+1}$ is the set of $\sigma = (x_{-1}, \ldots, x_{-r}, c)$ satisfying $0 \le x_i < b$. This set Σ is the state space of the generator. We refer to c as the *carry component* of the state σ. The state $\sigma \in \Sigma$ evolves according to the transformation $T : \Sigma \to \Sigma$ defined by $T(x_{-1}, \ldots, x_{-r}, c) = (x'_{-1}, \ldots, x'_{-r}, c')$, where

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

$$x'_{-1} + c'b = f(x_{-1}, \ldots, x_{-r}, c). \tag{2}$$

The integers x'_{-1} and c' are uniquely determined from (2) since we must have $0 \le x'_{-1} < b$, and therefore x'_{-1} is the least nonnegative residue of $f(x_{-1}, \ldots, x_{-r}, c)$ modulo b. From each state $(x_{-1}, \ldots, x_{-r}, c)$ a uniform pseudorandom number is obtained by using $x_{-1}/b \in [0, 1)$. As an example, if one takes

$$f(x_{-1}, \ldots, x_{-r}, c) = x_{-s} + x_{-r} + c$$

1991 *Mathematics Subject Classification.* Primary 65C10.
Key words and phrases. Random number generation, recurrences with carry, lattice structure .

This work has been supported by NSERC-Canada grant # OGP0110050 and FCAR-Québec grant # 93ER1654 to the second author. We wish to thank Brian Whitney who brought reference [6] to our attention.

where s is an integer with $0 < s < r$, one obtains an AWC generator.

The set Σ^{rec} of *recurrent* states—those for which $T^n(\sigma) = \sigma$ for some positive integer n—deserves special attention. In case of the AWC above, we know that any state $\sigma \in \Sigma$ will evolve into Σ^{rec}, in no more than $r + 1$ steps, so that we may as well assume $\sigma \in \Sigma^{\mathrm{rec}}$. Now the carry component of any recurrent state is either 0 or 1, and this allows to bypass the costly Euclidean division implied in (2), since we then have, if $x_{-s} + x_{-r} + c \geq b$,

$$x'_{-1} = x_{-s} + x_{-r} + c - b, \qquad c' = 1,$$

while, if $x_{-s} + x_{-r} + c < b$,

$$x'_{-1} = x_{-s} + x_{-r} + c, \qquad c' = 0.$$

There is another circumstance that will allow efficient calculation (on a binary computer) of x'_{-1} and c'. This is when the base b is equal to 2^{ω}, a power of 2. The binary representation of x'_{-1} and c' are then obtained, respectively, as the ω least significant bits and the remaining more significant bits of the right-hand side of (2), when this is positive. Using this device, and taking

$$f(x_{-1}, \dots, x_{-r}, c) = a_1 x_{-1} + \cdots + a_r x_{-r} + c, \tag{3}$$

for suitably chosen fixed non-negative integers a_l, Marsaglia [6], calls the carry generator thus defined, a *multiply-with-carry* (MWC) generator. In this case, the carry component of a recurrent state is non-negative. One may also allow the coefficients a_l to be negative. When none of them are positive, we call the defined generator a *complementary multiply-with-carry* generator. The carry component c of a recurrent state is now negative. The recurrence (2) can then be written, in terms of the related non-negative quantities $\tilde{c} = -c - 1$ and $\tilde{c}' = -c' - 1$,

$$(b-1) - x'_{-1} + \tilde{c}' b = (-a_1)x_{-1} + \cdots + (-a_r)x_{-r} + \tilde{c}, \tag{4}$$

and we may recover $(b-1) - x'_{-1}$ and $\tilde{c}'$ from the right-hand side, as in the case of a MWC. We shall see that, for the complementary MWC, each bit of the output value is fair, that is, the two binary digits will appear equally often in a full period, a property not shared by MWC generators.

In this paper, we study the d-dimensional uniformity of the output of MWC and complementary MWC generators. To be more specific, we note that the set Σ^{rec} of recurrent states will split, in general, into a certain number of T-invariant subsets, the *T-orbits*, over which the action of T is transitive. Each of these orbits defines a different random number generator, which we may refer to as a *minimal subgenerator*. Given one such orbit, and a positive integer d, we enquire about the number of states σ belonging to this orbit, and such that its *output d-tuple*, that is, the d-tuple of output values corresponding to $(\sigma, T(\sigma), \dots, T^{d-1}(\sigma))$, is equal to a given arbitrary d-tuple in the unit hypercube $[0,1)^d$. Marsaglia [6] obtained one such result in the special case of AWC/SWB generators. It states that almost every r-tuple of numbers of the form y/b, with y an integer satisfying $0 \leq y < b$, will appear exactly once as an output r-tuple in a full period. The method of proof is to show that some orbit in Σ^{rec} has its period close to the cardinality of Σ^{rec}, and the result follows from a characterization of recurrent states which implies that almost every r-tuple in $\{0, \dots, b-1\}^r$ figures as the first r components of a single recurrent state. This property of admitting a large-period orbit was, incidentally, the original motivation for the introduction of a carry component. The lagged-Fibonacci and more generally the multiple recursive generators have, in case of a power of two modulus, a maximal period much smaller than the cardinality of the set of recurrent states.

In Section 2, we give a characterization of the recurrent states, and show that any state in Σ will quickly evolve into Σ^{rec}. A close connection is also established between the recurrent states and a certain linear congruential generator (LCG). This connection was investigated in [11] and [1], again in the AWC/SWB case. We then distinguish two aspects of the question of d-dimensional uniformity, requiring different methods. These are discussed in Sections 3 and 4 respectively. We see in Section 3 that the problem leads to some arithmetical questions. In Section 4, we make use of the well-known spectral test. We examine in this respect some specific instances proposed in [6]. We also indicate a method of search for parameters which are good according to this test. The spectral test had been used in [11] and [1], to obtain distribution properties of the AWC/SWB generators for dimensions $d > r$. A preliminary version of this paper (without the proofs) was presented at the 1995 Winter Simulation Conference. For general references on random number generation, the reader can consult, e.g., [3, 4, 9].

2. Orbit structure

In this section, we deal with recursive carry generators defined by functions f of the form (3). We do not assume the coefficients a_l to be positive, but only that $m = -1 + a_1 b + \cdots + a_r b^r \neq 0$. This m may thus also be negative. In order to avoid trivial special cases we assume that at least one coefficient a_l is not 0. It is convenient to introduce a coefficient a_0 equal to -1. The base b can be an arbitrary positive integer. We will examine, for such generators, the orbit structure in the state space Σ, under the action of the transformation T. This is done by embedding a certain LCG into the carry generator.

Put $\mathbf{Z}_m = \{k \in \mathbf{Z} \,|\, 0 \leq k/m < 1\}$, and define the transformation $S : \mathbf{Z}_m \to \mathbf{Z}_m$ by $S(k) = k'$ where $k' \in \mathbf{Z}_m$ is subject to $bk' \equiv k \pmod{m}$. This transformation S is well defined and invertible, since b is prime to m. We first construct a one-to-one mapping $\iota : \mathbf{Z}_m \to \Sigma$ such that, identifying corresponding elements, S is identified with T (see Theorem 1 for a precise statement).

For $k \in \mathbf{Z}_m$, we define

$$\gamma(k) = \textstyle\sum_{i=-\infty}^{-1} \sum_{l=0}^{r} a_l y_{i-l} b^i, \tag{5}$$

$$\iota(k) = (y_{-1}, \ldots, y_{-r}, \gamma(k)), \tag{6}$$

where $y_{-1}, y_{-2}, \ldots$ are the digits in the b-adic expansion of k/m (note that these digits are uniquely determined by k/m since b is prime to m), so that $k/m = \sum_{i=-\infty}^{-1} y_i b^i$, and therefore

$$k = \sum_{i=0}^{r-1} \sum_{l=i+1}^{r} a_l y_{i-l} b^i + \gamma(k). \tag{7}$$

It follows from (7) that $\gamma(k) \in \mathbf{Z}$, and we have thus obtained a mapping $\iota : \mathbf{Z}_m \to \Sigma$.

Theorem 1. *The mapping $\iota : \mathbf{Z}_m \to \Sigma$, given by (5) and (6), is uniquely determined by its following two properties.*

(i) *For $k \in \mathbf{Z}_m$, we have $\iota(S(k)) = T(\iota(k))$.*

(ii) *If $k \in \mathbf{Z}_m$, then $x_{-1}/b \leq k/m < x_{-1}/b + 1/b$, where x_{-1} is the first component of $\iota(k)$.*

Proof. It is a simple verification that ι, given by (5) and (6), satisfies (i) and (ii). Consider now any mapping $\iota : \mathbf{Z}_m \to \Sigma$ satisfying these two properties. Take any $k \in \mathbf{Z}_m$, and let $k/m = \sum_{i=-\infty}^{-1} y_i b^i$ be its b-adic expansion. We will show that $\iota(k)$ is given by (5) and (6). For any non-negative integer n, we have the b-adic expansion $S^{-n}(k)/m = \sum_{i=-\infty}^{-1} y_{i-n} b^i$. By (1) and property (i), the jth component of $\iota(k)$, for $1 \le j \le r$, is equal to the first component of $\iota(S^{-j+1}(k))$, and is therefore equal to y_{-j} by property (ii). Thus the first r components of $\iota(k)$ are given by the first r digits in the b-adic expansion of k/m. Apply this to $S^{-n}(k)$ for n equal to $-i-1$ and $-i$, with i a negative integer. We then find, denoting by c_n the carry component of $\iota(S^{-n}(k))$, and using (2) with property (i), that $c_{-i-1} = (\sum_{l=0}^{r} a_l y_{i-l} + c_{-i}) b^{-1}$ and, by a recursive substitution, that $c_0 = \sum_{i=-\infty}^{-1} \sum_{l=0}^{r} a_l y_{i-l} b^i$. □

Property (ii) of the Theorem is extended as follows.

Corollary 1. *For $k \in \mathbf{Z}_m$ and any positive integer n, the nth digit in the b-adic expansion of k/m is equal to the first component of $\iota(S^{-n+1}(k))$.*

Proof. Let $k/m = \sum_{i=-\infty}^{-1} y_i b^i$ be the b-adic expansion of k/m. We then have $S^{-n+1}(k)/m = \sum_{i=-\infty}^{-1} y_{i-n+1} b^i$, and y_{-n} is the first component of $\iota(S^{-n+1}(k))$ by property (ii) of Theorem 1. □

As with Σ^{rec}, the set $\mathbf{Z}_m$ is decomposed, by means of the transformation S, into a set of orbits, which we may call S-orbits. By Theorem 1 (i), each S-orbit is mapped by ι onto a T-orbit. For d, a positive integer, and for any integer y satisfying $0 \le y < b^d$, we denote by $I_y^{(d)}$ the interval $\{x \in \mathbf{R} \,|\, y/b^d \le x/m < (y+1)/b^d\}$.

Corollary 2. *Let $K \subset \mathbf{Z}_m$ be a S-orbit, and $\iota(K)$ its corresponding T-orbit. Let d be any positive integer, and let $y_{-1}, \ldots, y_{-d}$ be given integers in $\{0, \ldots, b-1\}$. Put $y = \sum_{i=-d}^{-1} y_i b^{d+i}$. Then the number of states $\sigma \in \iota(K)$ with output d-tuple $(y_{-d}/b, \ldots, y_{-1}/b)$, is equal to the cardinality of $K \cap I_y^{(d)}$.*

Proof. By Corollary 1, for $k \in K$, and any positive integer n, the nth digit in the b-adic expansion of $S^{d-1}(k)/m$ is equal to the first component of $\iota(S^{d-n}(k)) = T^{d-n}(\iota(k))$. Thus the set of $k \in K$ such that $S^{d-1}(k) \in I_y^{(d)}$ is in a one-to-one correspondence, by ι, with the set of states $\iota(k)$, $k \in K$, with given output d-tuple $(y_{-d}/b, \ldots, y_{-1}/b)$. On the other hand, the former set is mapped one-to-one onto $K \cap I_y^{(d)}$ by S^{d-1}. □

The question of the distribution of the output d-tuples of the minimal subgenerator associated with the T-orbit $\iota(K)$ is thus reduced to the question of the distribution of the S-orbits K in $\mathbf{Z}_m$, into intervals of length $|m|/b^d$. It now arises whether every T-orbit in Σ^{rec} is of the form $\iota(K)$ for some S-orbit K. This turns out to be true with one exception, namely for the trivial orbit $\{\varsigma_1\}$ where $\varsigma_1 = (b-1, \ldots, b-1, a_0 + \cdots + a_r)$ is one of the only two states fixed by T, the other being $\varsigma_0 = (0, \ldots, 0) = \iota(0)$. It is a consequence of the fact that the set of recurrent states Σ^{rec} is equal to $\iota(\mathbf{Z}_m) \cup \{\varsigma_1\}$, which we now proceed to demonstrate.

First, introducing the mapping $\rho : \Sigma \to \mathbf{Z}$, defined by

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

we can rewrite (7) as

$$k = \rho(\iota(k)), \qquad k \in \mathbf{Z}_m. \tag{8}$$

We also note that property (ii) of Theorem 1 can be generalized to

$$b\,\rho(T(\sigma)) = \rho(\sigma) + x'_{-1}m, \tag{9}$$

where x'_{-1} is the first component of $T(\sigma)$.

Next, we define $\delta : \Sigma \to \mathbf{R}$ by

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

where $\sigma = (x_{-1}, \dots, x_{-r}, c) \in \Sigma$, so that

$$m \sum_{i=-r}^{-1} x_i b^i = \rho(\sigma) - \delta(\sigma). \tag{10}$$

We further have, writing $\sigma' = (x'_{-1}, \dots, x'_{-r}, c') = T(\sigma)$,

$$b\,\delta(\sigma') = \delta(\sigma) + m x_{-r} b^{-r}. \tag{11}$$

Note that, in general, for $\sigma = (x_{-1}, \dots, x_{-r}, c) \in \Sigma$, the integer $k = \rho(\sigma)$ may well not be contained in $\mathbf{Z}_m$. However, it results from (10) that, taking $\delta(\sigma)/m$ non-negative and sufficiently small, one can arrange that this be the case, that the x_i's be the first r-digits in the b-adic expansion of k/m, and that $\gamma(k)$ be close to c. The function δ is thus indicative as to the extent to which a given state falls short of being recurrent (see the next theorem for a precise statement).

Theorem 2. (i) *A state $\sigma \in \Sigma$ belongs to $\iota(\mathbf{Z}_m)$ if and only if*

$$0 \le \delta(\sigma)/m < 1/b^r. \tag{12}$$

(ii) *A state $\sigma \in \Sigma$ is equal to ς_0 if and only if $\delta(\sigma) = 0$.*
(iii) *For any state $\sigma \in \Sigma$, we have $T^n(\sigma) \in \iota(\mathbf{Z}_m) \cup \{\varsigma_1\}$ if the non-negative integer n satisfies*

$$n \ge \max(0, \log_b|\delta(\sigma)| - \log_b|m| + r) + \max(r, \log_b|m|) + 1. \tag{13}$$

(iv) *A state $\sigma \in \Sigma$ satisfies $T^n(\sigma) \in \{\varsigma_0, \varsigma_1\}$ for some non-negative integer n, if and only if $\rho(\sigma) \equiv 0 \pmod m$.*

For $h \in \mathbf{R}$, we denote by Σ_h the set of states $\sigma \in \Sigma$ for which $\delta(\sigma)/m < h/b^r$. We put $\Sigma' = \Sigma_1 \backslash \Sigma_0$. This is precisely the set of states σ for which (12) holds. We will say that the two states $\sigma = (x_{-1}, \dots, x_{-r}, c)$ and $\bar\sigma = (\bar x_{-1}, \dots, \bar x_{-r}, \bar c)$ are equivalent, in symbols $\sigma \sim \bar\sigma$, when $x_i = \bar x_i$, $i = -1, \dots, -r$. In the proof of Theorem 2, we make use of the following facts.

Lemma 1. (i) *If σ and $\bar\sigma \in \Sigma$ satisfy $\sigma \sim \bar\sigma$ and $T(\sigma) \sim T(\bar\sigma)$, we then have $(c' - \bar c')b = c - \bar c$, where $c, \bar c, c'$ and $\bar c'$ are the respective carry components of σ, $\bar\sigma$, $T(\sigma)$ and $T(\bar\sigma)$.*
(ii) *The set Σ_h is T-invariant if $h \ge 1$. The complementary set Σ_h^c is T-invariant if $h \le 0$. In particular, Σ' is T-invariant.*
(iii) $\rho(\Sigma') \subset \mathbf{Z}_m$.
(iv) *If $\sigma \in \Sigma'$ then $\sigma \sim \iota(\rho(\sigma))$.*
(v) *If $\sigma \in \Sigma'$ then $T(\iota(\rho(\sigma))) = \iota(\rho(T(\sigma))$.*

(vi) *If $\sigma \in \Sigma$ has carry component c and satisfies $\sigma \sim \varsigma_0$, then $\delta(\sigma) = c$, while if $\sigma \sim \varsigma_1$, then $\delta(\sigma) = c - (a_0 + \cdots + a_r) + mb^{-r}$.*

(vii) *For $\sigma \in \Sigma$ and n, a positive integer with $n \geq \log_b|\delta(\sigma)| - \log_b|m| + r$, we have $T^n(\sigma) \in \Sigma_2 \backslash \Sigma_{-1}$.*

(viii) *Let $\sigma \in \Sigma_2 \backslash \Sigma_1$ (resp. $\Sigma^c_{-1} \backslash \Sigma^c_0$). If $\sigma \not\sim \varsigma_1$ (resp. ς_0), then $T^r(\sigma) \in \Sigma'$, while if $\sigma \sim \varsigma_1$ (resp. ς_0), then either $\sigma = \varsigma_1$ (resp. ς_0), or there exists a positive integer n such that $T^n(\sigma) \not\sim \varsigma_1$ (resp. ς_0), and $n \leq \log_b|m| - r + 1$.*

Proof. Statement (i) follows from the recurrence formulas for T, (1) and (2), and from (3). Using (11) we obtain

$$\frac{1}{b}\frac{\delta(\sigma)}{m} \leq \frac{\delta(T(\sigma))}{m} \leq \frac{1}{b}\frac{\delta(\sigma)}{m} + \frac{1}{b^r} - \frac{1}{b^{r+1}}, \tag{14}$$

so that, if $h \geq 1$ and $\sigma \in \Sigma_h$, then $\delta(T(\sigma))/m < (1 + (h-1)/b)/b^r \leq h/b^r$, and $T(\sigma) \in \Sigma_h$ while, if $h \leq 0$ and $\sigma \in \Sigma^c_h$, then $\delta(T(\sigma))/m \geq h/b^{r+1} \geq h/b^r$ and $T(\sigma) \in \Sigma^c_h$. This proves statement (ii). Let $\sigma = (x_{-1}, \ldots, x_{-r}, c) \in \Sigma'$. By (10), $0 \leq \rho(\sigma)/m < 1$, so that $\rho(\sigma) \in \mathbf{Z}_m$, and $x_{-1}, \ldots, x_{-r}$ are the first r digits of the b-adic expansion of $\rho(\sigma)/m$. This gives statement (iii) and, using the definition of ι, statement (iv). If $\sigma \in \Sigma'$ then, by (ii), we also have $T(\sigma) \in \Sigma'$ and therefore, by (iii), $\rho(\sigma)$ and $\rho(T(\sigma)) \in \mathbf{Z}_m$. Statement (v) then follows from (9) and Theorem 1 (i). Statement (vi) is a straightforward calculation from the definition of δ. If $\sigma \in \Sigma$ and n is a positive integer then, by repeated application of (14), we have

$$\frac{1}{b^n}\frac{\delta(\sigma)}{m} \leq \frac{\delta(T^n(\sigma))}{m} \leq \frac{1}{b^n}\frac{\delta(\sigma)}{m} + \frac{1}{b^r} - \frac{1}{b^{r+n}},$$

so that, if $|\delta(\sigma)/m| \leq b^{n-r}$, then $-1/b^r \leq \delta(T^n(\sigma))/m < 2/b^r$, and statement (vii) follows. Finally, we prove statement (viii). Assume first that $\sigma = (x_{-1}, \ldots, x_{-r}, c) \in \Sigma_2 \backslash \Sigma_1$. If $\sigma \not\sim \varsigma_1$, then $x_i \leq b - 2$ for some value i_0 of the index i so that by (11), $T^{r+i_0+1}(\sigma) \in \Sigma'$ and therefore, $T^r(\sigma) \in \Sigma'$. If $\sigma \sim \varsigma_1$ then, by our hypothesis, we have $1/b^r \leq \delta(\sigma)/m < 2/b^r$ or, using (vi),

$$0 \leq \frac{c - (a_0 + \cdots + a_r)}{m} < \frac{1}{b^r}. \tag{15}$$

For any positive integer n, denote by c_n the carry component of $T^n(\sigma)$ and assume that $T^n(\sigma) \sim \varsigma_1$ for $0 < n \leq \log_b|m| - r + 1$. Let n' be the integral part of $\log_b|m| - r + 1$. By repeated application of (i), we obtain from (15)

$$|c_{n'} - (a_0 + \cdots + a_r)| < \frac{|m|}{b^{r+n'}}.$$

The left-hand side, being a non-negative integer, must be equal to 0 since the right-hand side does not exceed 1. Another appeal to (i) leads to $c = a_0 + \cdots + a_r$, and therefore $\sigma = \varsigma_1$. Assume now that $\sigma = (x_{-1}, \ldots, x_{-r}, c) \in \Sigma^c_{-1} \backslash \Sigma^c_0$. This case is symmetrical to the previous one, and we use the same notations. If $\sigma \not\sim \varsigma_0$, then $x_i \geq 1$ for some value i_0 of the index i so that by (11), $T^{r+i_0+1}(\sigma) \in \Sigma'$ and therefore, $T^r(\sigma) \in \Sigma'$. If $\sigma \sim \varsigma_0$ then, by our hypothesis, we have

$$-\frac{1}{b^r} \leq \frac{\delta(\sigma)}{m} < 0$$

or, using (vi),

$$-\frac{1}{b^r} \leq \frac{c}{m} < 0. \tag{16}$$

Assume that $T^n(\sigma) \sim \varsigma_0$ for $0 < n \le \log_b|m| - r + 1$. By repeated application of (i), we obtain from (16)

$$|c_{n'}| < \frac{|m|}{b^{r+n'}}.$$

We conclude as in the preceding case that $c = 0$. □

Proof of Theorem 2. Assume that $\sigma = (x_{-1}, \dots, x_{-r}, c) = \iota(k)$ for $k \in \mathbf{Z}_m$. Let $y_{-1}, y_{-2}, \dots$ be the digits in the b-adic expansion of k/m. It then follows from the definition of ι that $x_i = y_i$ for $i = -1, \dots, -r$. But, from (10) and (8), we have $\sum_{i=-r}^{-1} x_i b^i = k/m - \delta(\sigma)/m$ so that $\delta(\sigma)/m = \sum_{i=-\infty}^{-r-1} y_i b^i$, and (12) follows. We thus have shown that $\iota(\mathbf{Z}_m) \subset \Sigma'$ and we now prove the converse inclusion. For this, we show that, if $\sigma \in \Sigma'$, then the conclusion of Lemma 1 (iv) can be strenghtened to $\sigma = \iota(\rho(\sigma))$. For any non-negative integer n, put $\sigma_n = T^n(\sigma)$, $\bar\sigma_n = T^n(\iota(\rho(\sigma))$ and denote by c_n and $\bar c_n$ their respective carry components. By Lemma 1 (ii) and (v), we have $\sigma_n \in \Sigma'$ and $\bar\sigma_n = \iota(\rho(\sigma_n))$ so that, by Lemma 1 (iv), $\sigma_n \sim \bar\sigma_n$. This implies, using Lemma 1 (i), that

$$(c_{n+1} - \bar c_{n+1})b = c_n - \bar c_n, \quad n \ge 0. \tag{17}$$

Since $c_n - \bar c_n$ is an integer, it must therefore be equal to zero if n is sufficiently large. But then (17) implies that it is zero for all $n \ge 0$, and we obtain $\sigma = \iota(\rho(\sigma))$. We have thus shown that $\iota(\mathbf{Z}_m) = \Sigma'$. This is statement (i).

Clearly $\delta(\varsigma_0) = 0$. Conversely, if a state $\sigma = (x_{-1}, \dots, x_{-r}, c) \in \Sigma$ satisfies $\delta(\sigma) = 0$ then, by (10), we must have $x_i = 0$, $i = -r, \dots, -1$, since b is prime to m. Therefore $c = \delta(\sigma) = 0$, and $\sigma = \varsigma_0$. This proves statement (ii).

Consider any state $\sigma = (x_{-1}, \dots, x_{-r}, c) \in \Sigma$, and a non-negative integer n satifying (13). There then exist two non-negative integers n_1 and n_2 satisfying $n = n_1 + n_2$, $n_1 \ge \log_b|\delta(\sigma)| - \log_b|m| + r$, and $n_2 \ge \max(r, \log_b|m|)$. By Lemma 1 (vii), $T^{n_1}(\sigma) \in \Sigma_2 \backslash \Sigma_{-1}$, and by Lemma 1 (viii), either $T^{n_1}(\sigma) = \varsigma_1$, or $T^n(\sigma) \in \Sigma'$. Combined with statement (i), this proves statement (iii).

It follows from (9) that, for any state $\sigma \in \Sigma$, $\rho(\sigma) \equiv 0 \pmod m$ if and only if $\rho(T(\sigma)) \equiv 0 \pmod m$. We obtain statement (iv) from this and statement (i), since ς_0 and ς_1 are the only states in $\iota(\mathbf{Z}_m) \cup \{\varsigma_1\}$ mapped by ρ on an integer multiple of m. □

Theorem 1 (i) implies that $\iota(\mathbf{Z}_m) \subset \Sigma^{\mathrm{rec}}$, and we now obtain from Theorem 2 (iii) that $\Sigma^{\mathrm{rec}} = \iota(\mathbf{Z}_m) \cup \{\varsigma_1\}$. It follows that any non-trivial T-orbit in Σ^{rec} is of the form $\iota(K)$ for some S-orbit K in $\mathbf{Z}_m$. We are now in a position to apply Corollary 2. The simplest case arises when m is prime, and b is a primitive root modulo m, so that $K = \mathbf{Z}_m \backslash \{0\}$. Let d be a positive integer and let $y_{-1}, \dots, y_{-d}$ be given integers in $\{0, \dots, b-1\}$. Let ν be the largest integer smaller than $|m|/b^d$. It then follows from Corollary 2, that the number of those states in $\iota(K)$ for which the output d-tuple is equal to $(y_{-d}/b, \dots, y_{-1}/b)$, is either ν or $\nu + 1$. Let N_0 and N_1 be the number of such d-tuples for which this is ν, and $\nu + 1$ respectively. Then we have $N_0 + N_1 = b^d$, and $\nu N_0 + (\nu + 1)N_1 = m - 1$, from which we obtain

$$N_0 = (\nu + 1)b^d - m + 1 \tag{18}$$

$$N_1 = m - \nu b^d - 1. \tag{19}$$

For instance, in case of an AWC meeting the above conditions on m and b, we have $m = -1 + b^s + b^r$ where the integer s satisfies $0 < s < r$. We obtain, for $s < d \le r$,

$\nu = b^{r-d}$, $N_0 = b^d - b^s + 2$, $N_1 = b^s - 2$, and for $0 < d \le s$, $\nu = b^{r-d} + b^{s-d}$, $N_0 = 2$, $N_1 = b^d - 2$.

As a consequence of the characterization (12) in Theorem 2, if $|m| > b^r$, then there exists a state $(x_{-1}, \ldots, x_{-r}, c) \in \iota(\mathbf{Z}_m)$, for any choice of $(x_{-1}, \ldots, x_{-r}) \in \{0, \ldots, b-1\}^r$. It is possible, in this case, given the coefficients $a_1, \ldots, a_r$, to determine the smallest interval containing the carry component of all states in $\iota(\mathbf{Z}_m)$. For $j = -1, \ldots, -r$, put $m_j = \sum_{0 \le l < -j} a_l b^l$ and, for $x \in \mathbf{R}$, write $x^+ = \max(x, 0)$, $x^- = -\min(x, 0)$.

Corollary 3. *The carry component c of any state in $\iota(\mathbf{Z}_m)$ satisfies*

$$-(b-1) \sum_{j=-r}^{-1} b^j \left(\frac{m_j}{m}\right)^- \le \frac{c}{m} < (b-1) \sum_{j=-r}^{-1} b^j \left(\frac{m_j}{m}\right)^+ + \frac{1}{b^r}. \tag{20}$$

These inequalities are best possible when $|m| > b^r$. If $a_l \ge 0$, $l = 1, \ldots, r$, they amount to

$$0 \le c < \sum_{l=1}^{r} a_l, \tag{21}$$

while if $a_l \le 0$, $l = 1, \ldots, r$, to

$$\sum_{l=1}^{r} a_l \le c \le 0. \tag{22}$$

In the latter case, the carry c is equal to 0 only when $x_i = 0$ for $i = -r, \ldots, -1$.

Proof. The inequalities (12) can be rewritten as

$$\sum_{j=-r}^{-1} x_j b^j m_j / m \le c/m < \sum_{j=-r}^{-1} x_j b^j m_j / m + 1/b^r, \tag{23}$$

and the inequalities (20) follow by taking the minimum and the maximum, over all $(x_{-1}, \ldots, x_{-r}) \in \{0, \ldots, b-1\}^r$, of the left bound and the right bound in (23) respectively. When $|m| > b^r$, there is always an integer c satisfying (23), and therefore the inequalities (20) are best possible. Assume now that $a_l \ge 0$, $l = 1, \ldots, r$. Let k be the smallest integer l such that $a_l > 0$. Then $(m_j/m)^-$ (resp. $(m_j/m)^+$) is equal to $1/m$ (resp. 0) if $-k \le j \le -1$, and to 0 (resp. m_j/m) if $-r \le j < -k$. The lower bound in (20) is thus equal to $(-1 + b^k)/m$, and the upper bound to $(b-1)/m \sum_{-r \le j \le -k} b^j m_j + 1/b^r = 1/m \sum_{l=1}^{r} a_l - 1/(mb^r)$. Since c is an integer, these bounds are equivalent to the inequalities $0 \le c < \sum_{l=1}^{r} a_l$. The case of non-positive coefficients a_l is similar. ☐

3. Large intervals

We will consider a case where application of Corollary 2 leads to the study of the distribution of the S-orbits in $\mathbf{Z}_m$, into large intervals. We assume that $b = 2^\omega$ for some positive integer ω greater than 2. We also assume that m is prime, so that we may consider $\mathbf{Z}_m \backslash \{0\}$ as a group with respect to multiplication modulo m. Let K_0 be the subgroup of $\mathbf{Z}_m \backslash \{0\}$ generated by b. A non-trivial S-orbit $K \subset \mathbf{Z}_m$ is then given by any coset of K_0 in $\mathbf{Z}_m \backslash \{0\}$. Since the Legendre symbol

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

2 is always a quadratic residue, and we will assume that 2 generates the subgroup of quadratic residues. It follows that the number of non-trivial S-orbit K is equal to $2\omega_0$ with ω_0 equal to the greatest common divisor of ω and $(|m|-1)/2$.

We first consider the simplest case of $m>0$, namely when $\omega_0=1$. A non-trivial S-orbit $K\subset \mathbf{Z}_m$ can now be either the set of all quadradic residues or of all non-quadratic residues in $\mathbf{Z}_m\backslash\{0\}$. It suffices to consider the former case. Let d be a positive integer. Corollary 2 then leads us to the study of the distribution of the set of quadratic residues in the intervals $I_y^{(d)}$, $0\le y<b^d$ and, in particular, to the question of how many residues there are in the interval

$$I=\cup_{0\le y<b^d/2}I_y^{(d)}=\{x\in\mathbf{R}\,|\,0\le x/m<1/2\}.$$

Let D_m denote the difference between the number of residues and non-residues in I. The number of non-residues in I is equal to the cardinality of $K\backslash I$. Thus, D_m is equal to the difference between the number of times the most significant binary digit of the output value of the minimal subgenerator associated with K is equal to 0 and the number of times it is equal to 1, over the full period. This statement remains valid if we use the nth most significant digit, $n\le\omega$, instead of the first. Indeed, the correspondence $k\mapsto k'$, where $k,\ k'\in\mathbf{Z}_m$ satisfies $2^nk'\equiv k \pmod m$, maps K one-to-one onto K, and the nth digit of the binary expansion of k/m, $k\in\mathbf{Z}_m$, is nothing but the first digit of that of k'/m. If we expect this minimal subgenerator to be a uniform random number generator, the parameters a_l, and therefore m, should thus be chosen so as to make $|D_m|$ as small as possible and, at any rate, not significantly larger than $\sqrt{(|m|-1)/2}$, the standard deviation of a sum of $(|m|-1)/2$ independent Bernouilli trials, each equal to 1 or -1 with the same probability $1/2$.

As m will normally be very large, direct computation of D_m is not to be considered. It may be of interest to see what can be obtained by means of Weyl's method. For any integer n, define $\chi(n)$ to be the Legendre symbol (n/m) if n is prime to m, and 0 otherwise. This function χ is a Dirichlet character for the modulus m. This means it is multiplicative, that is $\chi(n_1n_2)=\chi(n_1)\chi(n_2)$ for any pair of integers n_1, n_2, it is equal to 0 precisely for integers not prime to m, and it is periodic with period m. We consider sums of the type

$$\sum_{k\in\mathbf{Z}_m}\chi(k)p\left(\frac{2\pi k}{m}\right) \tag{24}$$

where p is a 2π-periodic function. This expression is equal to D_m if we take $p=p_0$ defined by

$$p_0(t)=\begin{cases}1/2 & 0<t<\pi\\ 0 & t=0,\ \pi\\ -1/2 & -\pi<t<0.\end{cases}$$

Weyl's method is based on the fact that we know the values of (24) for the functions $p(t)=\exp(int)$, $n\in\mathbf{Z}$. They are the well-known Gauss sums, and are equal, in this case, to $i|\sqrt{m}|\chi(n)$. Expanding p_0 into a Fourier series,

$$p_0(t)=\frac{1}{\pi i}\sum_{n\equiv 1\,(2)}\frac{1}{n}e^{int},$$

we find that (24), with $p = p_0$, is equal to

$$\frac{|\sqrt{m}|}{\pi} \sum_{n\equiv 1\ (2)} \frac{\chi(n)}{n} = \frac{|\sqrt{m}|}{\pi}\left(1 - \frac{\chi(2)}{2}\right)\sum_{n\neq 0}\frac{\chi(n)}{n} = \frac{|\sqrt{m}|}{\pi}\sum_{n=1}^{\infty}\frac{\chi(n)}{n} \tag{25}$$

since $\chi(2) = 1$, and $\chi(-1) = -1$. By the analytic class number formula of Dedekind (see §51 of [2]), the right-hand side of (25), and therefore D_m, is nothing but the number h_Δ of ideal classes of the imaginary quadratic field $\mathbf{Q}(\sqrt{\Delta})$ of discriminant $\Delta = -m$. In particular $D_m > 0$, that is, there are always more residues than non-residues in I. In fact, it has been proved by Siegel [10] that, for any $\epsilon > 0$, there exists a constant C_ϵ such that $h_\Delta \geq C_\epsilon|\Delta|^{1/2-\epsilon}$ for every discriminant Δ of an imaginary quadratic field. On the other hand, we have the inequality (see p. 389 of [8]) $h_\Delta \leq (1/\pi)\sqrt{|\Delta|}\ \log|\Delta| + (2/\pi)(1+\log(2/\pi))\sqrt{|\Delta|}$. We therefore have

$$C_\epsilon m^{1/2-\epsilon} \leq D_m \leq \frac{1}{\pi}\sqrt{m}\log m + \frac{2}{\pi}\left(1 + \log\left(\frac{2}{\pi}\right)\right)\sqrt{m}. \tag{26}$$

Since the right-hand side becomes large compared with $\sqrt{(m-1)/2}$, when m is large, a more precise determination of D_m is still wanting.

We now assume that $m < 0$, and that ω is divisible by 4. We then have $-1 \in K_0$, and it follows that for any non-trivial S-orbit K, $K \cap I$ and $K\backslash I$ have the same cardinality. Thus, in this circumstance, all ω output bits are fair. Again using Weyl's method, it is further possible to study the independence of contiguous output bits. We consider for instance the two most significant bits. Since these two bits are fair, the pairs 01 and 10, as well as the pairs 00 and 11, will appear equally often as the two most significant bits of the output values over the full period of each minimal subgenerator. Therefore, all pairs of binary digits will appear equally often if the pairs 00 and 01 do so. We may thus measure the independence of the two most significant bits by the difference between the number of times these two pairs, 00 and 01, appear in the full period. But the sum $\tilde{D}_m$ of these differences corresponding to all S-orbits contained in the group of quadratic residues is equal to (24) with $p = p_1$, where p_1 is given by $p_1(t) = p_0(t+\pi/2)$. In this case the Gauss sums are equal to $|\sqrt{m}|\chi(n)$ and, expanding p_1 into a Fourier series,

$$p_1(t) = \frac{1}{\pi}\sum_{n\equiv 1\ (2)} \frac{(-1)^{(n-1)/2}}{n} e^{int},$$

we find that $\tilde{D}_m$ is equal to

$$\frac{|\sqrt{m}|}{\pi}\sum_{n\equiv 1\ (2)}\frac{\chi(n)(-1)^{(n-1)/2}}{n} = \frac{|\sqrt{4m}|}{\pi}\sum_{n=1}^{\infty}\frac{\chi'(n)}{n},$$

where χ' is that Dirichlet character for the modulus $4m$, which is the product of χ and the only nontrivial Dirichlet character for the modulus 4. Written in this form we recognize that $\tilde{D}_m$ is the class number of the imaginary quadratic field $\mathbf{Q}(\sqrt{\Delta})$ of discriminant $\Delta = 4m$, and we obtain for it, bounds similar to (26).

The above results and estimates are however too weak for the quantities D_m or $\tilde{D}_m$ to be useful as effective uniformity criteria. For this, sufficiently precise approximations to them must be developed.

4. Small intervals

When the dimension d is large enough so that the length $|m|/b^d$ of the intervals $I_y^{(d)}$ is smaller than 1, any d-tuple in the unit hypercube $[0,1)^d$ can appear at most once, in the full period of a minimal subgenerator, as an output d-tuple. It is sufficient in this case to locate in the unit cube those d-tuples that do appear. We will construct a lattice in $\mathbf{R}^d$ such that all output d-tuples are approximated by a lattice point. This lattice is then studied via the spectral test [3, 5].

We denote by $e_1, \ldots, e_d$, the canonical basis in $\mathbf{R}^d$. Put $v^* = 1/m \sum_{j=1}^d b^{d-j} e_j$ and let $\Lambda_d = \mathbf{Z}v^* + \mathbf{Z}^d$ be the lattice in $\mathbf{R}^d$ generated by v^* and $\mathbf{Z}^d$. The intersection $\Lambda_d \cap [0,1)^d$ is then precisely the set of d-tuples $(k/m, S(k)/m, \ldots, S^{d-1}(k)/m)$, $k \in \mathbf{Z}_m$.

By Theorem 1, the study of the distribution of the set of d-tuples of successive outputs of the carry generator, restricted to $\iota(\mathbf{Z}_m)$, is by large reduced to the study of the lattice Λ_d. Let $\Lambda^{(d)} = \{w \in \mathbf{R}^d \,|\, w \cdot \Lambda_d \subset \mathbf{Z}\}$ denote the lattice dual to Λ_d. If $w \in \Lambda^{(d)} \backslash \{0\}$ and $n \in \mathbf{Z}$, then the region $\{v \in \mathbf{R}^d \,|\, n < v \cdot w < n+1\}$ 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.

Define $a_l = 0$ if $l > r$, and put $w_1 = (\sum_{l \geq d-1} a_l b^{l-d+1}) e_1 + \sum_{j=2}^d a_{d-j} e_j$, and $w_j = -e_{j-1} + b e_j$ for $j = 2, \ldots, d$. These vectors belong to $\Lambda^{(d)}$ since they have integer coefficients, and since

$$v^* \cdot w_1 = 1, \qquad v^* \cdot w_j = 0, \quad j = 2, \ldots, d. \tag{27}$$

Let $H^{(d)} \subset \mathbf{R}^d$ be the subspace containing all vectors orthogonal to v^*, and let $\Lambda_H^{(d)}$ be the lattice generated in $H^{(d)}$ by the vectors w_j, $j = 2, \ldots, d$.

Proposition 1. *A lattice basis for $\Lambda^{(d)}$ is given by the set of vectors w_j, $j = 1, \ldots, d$. We have $\Lambda^{(d)} \cap H^{(d)} = \Lambda_H^{(d)}$, and the vectors of minimal length in $\Lambda_H^{(d)} \backslash \{0\}$ are the vectors $\pm w_j$, $j = 2, \ldots, d$.*

Proof. See [1], Propositions 1 and 4. ☐

The next theorem describes the set of shortest vectors of $\Lambda^{(d)} \backslash \{0\}$, for $d = 1, \ldots, r+1$, in an important special case, namely when

1) all coefficients a_l, $l = 1, \ldots, r$ are either non-negative or non-positive, the greater weight being given to the leading coefficient a_r,
2) the carry component c of any recurrent state satisfies $0 \leq c < b$, if all coefficients $a_l, l = 1, \ldots, r$ are non-negative, or $-b \leq c \leq 0$, if they are non-positive.

The first condition is to ensure that the density of points in Λ_d is large, as this density is equal to the group theoretical index $[\Lambda_d : \mathbf{Z}^d] = |m|$. The second condition, which is equivalent, by Corollary 3, to the condition $|\sum_{l=1}^r a_l| \leq b$, derives from implementation considerations. In case $b = 2^\omega$, with ω equal to the computer's word length, the first r components of a state can each be stored in one word. In case of non-negative coefficients, the condition guarantees that the carry component of any recurrent state can also be stored in one word, and that the corresponding sum (3) can be accumulated in a double-word register. A similar statement holds when the coefficients are non-positive. If c is the carry component of a recurrent state, then $\tilde{c}$ can be stored in one word and the right-hand side of (4) can be stored in a double-word.

Theorem 3. *Assume that $b \geq 6$, that $a_r \neq 0$, $0 \leq a_l/a_r < 1$, $l = 1, \ldots, r-1$, and that $|\sum_{l=1}^{r} a_l| \leq b$. The vectors of minimal length in $\Lambda^{(d)} \backslash \{0\}$ are then given by*

(i) $\pm w_j$, $j = 2, \ldots, d$ *if* $d < r$, *or if* $d = r$ *and* $|a_r| > 1$,
(ii) $\pm w_j$, $j = 1, \ldots, d$ *if* $d = r$ *and* $|a_r| = 1$, *or if* $d = r+1$ *and* $|a_r| = b$,
(iii) $\pm w_1$ *if* $d = r+1$, *and if* $|a_r| < b$.

Proof. Consider an arbitrary linear combination $w = \sum_{j=1}^{d} z_j w_j$ with real coefficients z_j. We have $\|v^*\|^2 = m^{-2} \sum_{j=1}^{d} b^{2(j-1)} = m^{-2}(b^{2d} - 1)/(b^2 - 1)$ and, by (27), $z_1 = v^* \cdot w$. Using the Cauchy-Schwarz inequality, we obtain

$$|z_1| \leq \frac{1}{m} \left(\frac{b^{2d} - 1}{b^2 - 1} \right)^{1/2} \|w\|. \tag{28}$$

Assume that $w \in \Lambda^{(d)} \backslash \{0\}$ so that the coefficients z_j are now integers, not all 0. Since all coefficients a_l, $l = 1, \ldots, r$, are of the same sign (or 0), we have

$$\frac{b^d}{m} \leq \frac{1}{|a_r| b^{r-d} - b^{-d}}, \tag{29}$$

and since $b \geq 4$, we have

$$\frac{1}{2 - b^{-d}} < \left(\frac{b^2 - 1}{b^2 + 1} \right)^{1/2}. \tag{30}$$

Under either assumptions in (i) we have $2 \leq |a_r| b^{r-d}$ and, combining this with (28), (29) and (30), we obtain

$$|z_1| < \frac{\|w\|}{(b^2 + 1)^{1/2}}. \tag{31}$$

We then obtain (i) from the last statement of Proposition 1 since, using (31), if $\|w\| \leq (b^2 + 1)^{1/2}$, then $z_1 = 0$, and therefore, $w \in \Lambda_H^{(d)}$.

Put $\varepsilon = a_r/|a_r|$. Under either assumptions in (ii), we have $w_1 = \varepsilon b e_1 - e_d$, and therefore, $w = w' + w''$ with $w' = \varepsilon b z_1 e_1 + b \sum_{j=2}^{d} z_j e_j$, and $w'' = -\sum_{j=1}^{d} z_{j+1} e_j$, where we take z_{d+1} to be equal to z_1. We have $\|w'\| = b(\sum_{j=1}^{r} z_j^2)^{1/2}$, and $\|w''\| = (\sum_{j=1}^{d} z_j^2)^{1/2}$. It follows that $\|w\| \geq (b-1)(\sum_{j=1}^{r} z_j^2)^{1/2}$, and this exceeds $(b^2+1)^{1/2}$ for integer coefficients z_i, unless at most one is not 0 and equal to ± 1. This implies (ii).

Assume now that $d = r+1$, that $|a_r| < b$, and that $\|w\| \leq \|w_1\|$, with $z_1 \geq 0$. We will see that this implies that $w = w_1$. Our hypothesis implies that $|a_l| < b$, $l = 1, \ldots, r$, and $\sum_{l=1}^{r} |a_l| \leq b$ so that $\sum_{l=1}^{r} a_l^2 < b^2$. We thus have $\|w_1\|^2 \leq b^2$, and therefore

$$\|w\| \leq b. \tag{32}$$

It follows that we cannot have $z_1 = 0$ since we would then have $w \in \Lambda_H^{(d)} \backslash \{0\}$, contradicting the last statement of Proposition 1.

We first consider the case $r > 1$. Since $b \geq 4$, we have

$$b^{2(d+1)} < (b+1)^2 (b^r - 1)^2 (b^2 - 1). \tag{33}$$

Also, using (28), we have

$$z_1^2 a_r^2 (b^r - 1)^2 \leq z_1^2 m^2 \leq \frac{b^{2d} - 1}{b^2 - 1} \|w\|^2.$$

Combining this with (33) we obtain

$$z_1^2 a_r^2 < (b+1)^2 \frac{\|w\|^2}{b^2}, \tag{34}$$

and therefore, using (32),

$$0 < z_1|a_r| \le b. \tag{35}$$

This implies that

$$0 \le z_1|a_j| \le b-2, \qquad 1 \le j < r, \tag{36}$$

since $|a_j| < |a_r|$ and $|a_j| + |a_r| \le b$, for $1 \le j < r$. We next show that

$$-1 \ \le \ \varepsilon z_j \ \le 0, \qquad 2 \le j \le r, \tag{37}$$

$$0 \ \le \ z_{r+1} \ \le 1. \tag{38}$$

We have

$$\|w\|^2 = (z_1 a_r - z_2)^2 + \sum_{j=2}^{r} (z_1 a_{r-j+1} + z_j b - z_{j+1})^2 + (z_{r+1} b - z_1)^2, \tag{39}$$

so that by (32), $(z_{r+1}b - z_1)^2 \le b^2$, and since $0 < z_1 \le b$, we must have $0 \le z_{r+1} \le 2$. Now z_{r+1} cannot be equal to 2, unless $z_1 = b$, $|a_r| = 1$ by (35), and therefore $a_1 = 0$, which would imply $\|w\|^2 \ge (z_r b - 2)^2 + b^2 > b^2$. This proves (38). For $2 \le j \le r$, we have, by (32) and (39), $(z_1 a_{r-j+1} + z_j b - z_{j+1})^2 \le b^2$ so that, by (36), $-2 \le \varepsilon z_{j+1} \le 1$ implies that $-2 \le \varepsilon z_j \le 1$ and $\varepsilon z_j = -2$ implies that $\varepsilon z_{j+1} = -2$. From this and (38) we obtain that $-1 \le z_j \le 1$ for $2 \le j \le r$. Using this and (39), we see that if $\varepsilon z_{j_0} = 1$ for some index j_0 with $2 \le j_0 \le r$, then

$$(z_1 - 1)^2 + (b-1)^2 + (b - z_1)^2 \le \|w\|^2$$

if $z_{r+1} = 1$, while

$$b^2 + z_1^2 \le \|w\|^2$$

if $z_{r+1} = 0$. But this is excluded by (32) since, in both cases, the left-hand side exceeds b^2 as $b \ge 6$. We have thereby proved (37). For $1 \le j \le d$, define q_j as the square of the jth coordinate of w minus a_{r+1-j}^2, the square of the jth coordinate of w_1. Let J be the set of indices j, for which $q_j < 0$ and let J' be the set of those $j \in J$ for which $\varepsilon z_j = -1$. Since $q_1 \ge (z_1^2 - 1)a_r^2 \ge 0$, we have $1 \notin J$. Also, $q_{r+1} \ge 0$, and $r+1 \notin J$. Indeed, assuming otherwise $q_{r+1} < 0$, this would imply $z_1 = b$, $z_{r+1} = 1$, and by (35), $|a_r| = 1$. But clearly then (32) cannot be satisfied. Thus, if we put

$$\begin{aligned} Q_1 &= q_1 + \sum_{j \in J'} q_j, \\ Q_2 &= \sum_{j \in J \setminus J'} q_j + q_{r+1}, \end{aligned}$$

we have $\|w\|^2 - \|w_1\|^2 \ge Q_1 + Q_2$, and by our assumption on w, we obtain that

$$Q_1 + Q_2 \le 0. \tag{40}$$

Denoting by $\#J'$ the cardinality of J', we now show that

$$\#J' < z_1. \tag{41}$$

If $J' \neq \emptyset$, and $j \in J'$, then $|z_1 a_{r+1-j} - b - z_{j+1}| < |a_{r+1-j}|$ so that, by (37) and (38), we have $(z_1+1)|a_{r+1-j}| \geq b$, and therefore, since $|a_r| > |a_{r+1-j}|$,

$$\frac{(\#J+1)b}{z_1+1} < |a_r| + \sum_{j \in J'} |a_{r+1-j}|.$$

As the right-hand side does not exceed b, we obtain (41), and therefore, since $q_j \geq -a_{r-j+1}^2$,

$$Q_1 \geq (z_1^2 - z_1)a_r^2. \tag{42}$$

Now, $J \backslash J' \subset \{r\}$ and, if $r \in J \backslash J'$, then $q_r = (z_1 a_1 - z_{r+1})^2 - a_1^2 < 0$, and we must have $z_1 = z_{r+1} = 1$, so that $q_r + q_{r+1} = b^2 - 2b - 2a_1 + 1 = (b-1)^2 - 2a_1 > 0$. Therefore, in any case, $Q_2 \geq 0$ with equality holding only if $J' = J$. It follows from this, (40), and (42) that $Q_1 = Q_2 = 0$, and therefore that $z_1 = 1$, $J = \emptyset$, and $z_{r+1} = 0$. This proves that $w_1 = w$.

Finally, we deal with the remaining case $r = 1$, and $d = 2$. In this case, combining (28) with our assumption that $\|w\| \leq \|w_1\|$ gives

$$0 < z_1^2 \leq \frac{(b^2+1)(a_1^2+1)}{(a_1 b - 1)^2}.$$

Since $a_1 \geq 1$, and $b \geq 4$, the right-hand side is less than 4, and we must therefore have $z_1 = 1$. The inequality $\|w\| \leq \|w_1\|$ can thus be written as

$$(a_1 - z_2)^2 + (z_2 b - 1)^2 \leq a_1^2 + 1,$$

which clearly implies that $0 \leq z_2 \leq 1$. If $z_2 = 1$, then the left-hand side is equal to $a_1^2 + 1 + (b-1)^2 - 2a_1$, which clearly exceeds $a_1^2 + 1$ since $b - 1 \geq a_1$, and $b - 1 > 2$. Therefore we must have $z_2 = 0$, and $w = w_1$. □

Under the hypothesis of the previous theorem, we discuss short vectors of $\Lambda^{(d)}$ for $d \geq r+1$. For $d = r+1$, the squared length of 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 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 $\|w_1\|^2 = 1 + \sum_{l=1}^r a_l^2 \leq b^2$. 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.

With an arbitrary vector $w = \sum_{j=1}^d z_j w_j \in \mathbf{R}^d$ associate the vector

$$w^\star = z_1 w_1 + b \sum_{j=2}^d z_j e_j.$$

We then have the following inequality

Lemma 2.

$$\|w\| \leq (1 + b^{-1})\|w^\star\| + |z_1|. \tag{43}$$

Proof. We have $w = w^\star - \sum_{j=2}^d z_j e_j$, and $b(\sum_{j=2}^d z_j^2)^{1/2} \leq |z_1|\|w_1\| + \|w^\star\|$. Therefore

$$\|w\| \leq \|w^\star\| + (\sum_{j=2}^d z_j^2)^{1/2} \leq (1 + b^{-1})\|w^\star\| + \|w_1\| b^{-1} |z_1|,$$

and since $\|w_1\| \leq b$, this proves the lemma. □

Thus, if there exists integers $z_1, \ldots, z_d$, not all zero, such that $|z_1|$ and $\|w^\star\|$ 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 $b\mathbf{Z}^{r+1}$. We illustrate this using two sets of parameters proposed by Marsaglia [6]. Both have $b = 2^{16}$, and $r = 8$. In both cases, the choice of coefficients a_l makes m and $(m-1)/2$ prime, so that b generates the group of quadratic residues modulo m, and we thus have two non-trivial orbits.

The first set of parameters is $a_1 = 1941$, $a_2 = 1860$, $a_3 = 1812$, $a_4 = 1776$, $a_5 = 1492$, $a_6 = 1215$, $a_7 = 1066$, and $a_8 = 12013$. The second 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. Squared length of shortest dual vector 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 $b\mathbf{Z}^d$ among all vectors

$$z_1 w_1, \qquad 0 < |z_1| < 2000,\ z_1 \in \mathbf{Z}.$$

This distance is approximately 2788.15, and this accounts, in view of the inequality (43), for the presence in the lattice $\Lambda^{(d)}$ of the small vector $w_{\min}$.

It is easy to find coefficients a_l which satify the conditions in Theorem 3, and which make the distance of $z_1 w_1$ to $b\mathbf{Z}^d$ 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 $b\mathbf{Z}^d$ approximately equal to 18163.47, for the set of vectors

$$z_1 w_1, \qquad 0 < |z_1| < 3000,\ z_1 \in \mathbf{Z}.$$

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 these coefficients.

TABLE 2. Squared length of shortest dual vector 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

References

1. R. Couture and P. L'Ecuyer, *On the lattice structure of certain linear congruential sequences related to AWC/SWB generators*, Mathematics of Computation **62** (1994), no. 206, 798–808.
2. E. Hecke, *Lectures on the theory of algebraic numbers*, Springer-Verlag, New York, 1981.
3. D. E. Knuth, *The art of computer programming, volume 2: Seminumerical algorithms*, second ed., Addison-Wesley, Reading, Mass., 1981.
4. P. L'Ecuyer, *Uniform random number generation*, Annals of Operations Research **53** (1994), 77–120.
5. P. L'Ecuyer and R. Couture, *An implementation of the lattice and spectral tests for multiple recursive linear random number generators*, To appear, Circa 1997.
6. G. Marsaglia, *Yet another rng*, Posted to the electronic billboard `sci.stat.math`, August 1, 1994.
7. G. Marsaglia and A. Zaman, *A new class of random number generators*, The Annals of Applied Probability **1** (1991), 462–480.
8. W. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, PWN–Polish Scientific Publishers, Warsaw, 1974.
9. H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, SIAM CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63, SIAM, Philadelphia, 1992.
10. C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arithmetica **1** (1935), 83–86.
11. S. Tezuka, P. L'Ecuyer, and R. Couture, *On the add-with-carry and subtract-with-borrow random number generators*, ACM Transactions of Modeling and Computer Simulation **3** (1994), no. 4, 315–331.

DÉPARTEMENT D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE, UNIVERSITÉ DE MONTRÉAL, C.P. 6128, SUCC. CENTRE-VILLE, MONTRÉAL, H3C 3J7, CANADA

E-mail address: `couture@iro.umontreal.ca` and `lecuyer@iro.umontreal.ca`