Efficient multiply-with-carry random number generators with maximal period.

MARK GORESKY[1]
Institute for Advanced Study
and
ANDREW KLAPPER[2]
University of Kentucky

In this (largely expository) paper we propose a simple modification of the multiply-with-carry random number generators of [Marsaglia 1994], [Couture and L'Écuyer 1997]. The resulting generators are both efficient (since they may be configured with a base b which is a power of 2) and exhibit maximal period. These generators are analyzed using a simple but powerful algebraic technique involving b-adic numbers.

Categories and Subject Descriptors: G.3 [**Probability and Statistics**]: Random Number Generation; I.6.0 [**Simulation and Modeling**]: General; G.2.1 [**Discrete Mathematics**]: Combinatorics—*recurrences and difference equations*

General Terms: Algorithms, Experimentation, Theory

Additional Key Words and Phrases: fcsr, feedback shift register, k-distribution, lattice structure, m-sequences, multiply-with-carry, p-adic number, primitive element, random number generation

1. INTRODUCTION

1.1

A pseudo random number generator (RNG) for high speed simulation and Monte Carlo integration should have several properties: (1) it should have enormous period, (2) it should exhibit uniform distribution of d-tuples (for a large range of d), (3) it should exhibit a good structure (usually a lattice structure) in high dimensions, and (4) it should be efficiently computable (preferably with a base b which is a power of 2). Typically the RNG is a member of a family of similar generators with different parameters and one hopes that parameters and seeds may be easily chosen so as to guarantee properties (1), (2), (3) and (4). Generators with these properties are surprisingly rare. Perhaps the best candidates known at present are

1. The Institute for Advanced Study, School of Mathematics, (www.math.ias.edu/∼goresky/). Research partially supported by N.S.F. grant CCR 0002693.
2. Dept. of Computer Science, 779A Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046 and the Institute for Advanced Study, School of Mathematics. E-mail: klapper@cs.uky.edu.
Project sponsored by the National Science Foundation under grant CCR 9980429.

[L'Écuyer 1996] and [Matsumoto and Nishimura 1998].

1.2

In [Marsaglia and Zaman 1991], Marsaglia and Zaman showed that their add-with-carry (AWC) generators satisfy condition (1). By giving up on (4) and using an appropriate base b, they achieve good distribution properties of d-tuples for values d which are less than the "lag." It has been shown [Tezuka et al 1993] that these generators fail the spectral test [Coveyou and MacPherson 1967] for large d. A generalization, the multiply-with-carry (MWC) generator, was described in [Marsaglia 1994], [Couture and L'Écuyer 1997] and independently, (motivated by some questions in cryptography) in [Klapper and Goresky 1994], [Klapper and Goresky 1993], [Klapper and Goresky 1997], where it was called a feedback-with-carry shift register, or FCSR. (This paper, which is largely expository, combines both points of view.)

The MWC generator was proposed as a modification of the AWC generator which satisfies both conditions (1) and (4). That is, all computations are performed modulo a base b which is a power of 2. However the distributional properties (2) of MWC sequences are not optimal, and in fact they are rather difficult to determine. See [Couture and L'Écuyer 1997], where estimates on the distribution of d-tuples are derived (using some sophisticated techniques from number theory).

In this paper we show that a slight (almost trivial) modification of the MWC generator results in sequences with maximum period (from which it follows that the distribution of d-tuples is uniform, for all d less than the lag, d_0) and which continue to satisfy properties (1) and (4). It is relatively easy to find generators of this type with base b a power of 2 (say, $b = 2^{21}$), with d_0 around 100, and with periods around 10^{750}. As in [Couture and L'Écuyer 1997], one could use the spectral test to search for parameters which might satisfy (3) however it is expected that for large $d > d_0$ the lattice structure will suffer from the same shortcomings as those described in [Couture and L'Écuyer 1997].

In Theorems 2.1, 2.2, 2.3, 2.4, 3.1 and 4.1 we describe the main properties of these generators. The proofs of Theorems 3.1 and 4.1 are "elementary". Proofs of the other results may be distilled from the literature on AWC and MWC generators ([Couture and L'Écuyer 1994], [Couture and L'Écuyer 1997], [Marsaglia and Zaman 1991], [Tezuka et al 1993]). However there is a very illuminating algebraic technique which may be used to give short and efficient proofs of these results. It is a simple but not entirely obvious modification of the technique of [Klapper and Goresky 1997] (and, as such, it is a special case of the general technique of [Klapper and Xu 1996]). We have included these short proofs at the end of this paper for the benefit of the reader who may not be familiar with the language of discrete valuations and algebraic completions.

We are very grateful to the editor, and to an anonymous referee, for their thoughtful comments and suggestions.

2. MULTIPLY WITH CARRY GENERATORS

2.1

Fix an integer "base", $b \geq 2$ and fix integer coefficients $a_0, a_1, \ldots, a_r$ with a_0 chosen to be relatively prime to b. (If b is a power of 2, this simply means that a_0 is odd.) A MWC generator of order r and base b consists of a state

$$\sigma = (x_{-1}, \ldots, x_{-r}; c)$$

where $0 \leq x_i < b$ and $c \in \mathbf{Z}$; and a transformation rule $T : \sigma \mapsto \sigma' = \left(x'_{-1}, \ldots, x'_{-r}, c'\right)$ which is defined as follows. If $i < -1$ then $x'_i = x_{i+1}$. The numbers x'_{-1} and c' are the unique solutions to

$$a_0 x'_{-1} + c'b = \sum_{i=1}^{r} a_i x_{-i} + c \tag{1}$$

with $0 \leq x'_i < b;$. The values of x'_{-1} and c' may be computed as follows. Calculate once and for all

$$A = a_0^{-1} \pmod{b} \tag{2}$$

and realize this as an integer between 0 and $b-1$. Set $\tau = \sum_{i=1}^{r} a_i x_{-i} + c$. Then:

$$x'_{-1} = (A\tau) \pmod{b} \tag{3}$$

$$c' = (\tau - a_0 x'_{-1})/b. \tag{4}$$

The integer c is called the "carry" or the "memory" of the state. The *output* of the state $\sigma = (x_{-1}, \ldots, x_{-r}; c)$ is the integer $\texttt{OUT}(\sigma) = x_{-r}$ and the *normalized output* is the real number x_{-r}/b.

Since $c \in \mathbf{Z}$ is arbitrary, there are infinitely many different states and infinitely many different output sequences. However there are only finitely many *periodic* states, in which case the carry c remains within a certain finite interval $w^- \leq c \leq w^+$ according to Theorem 3.1 below. Moreover, from any initial state, the generator will eventually enter a periodic state. Consequently, for any initial state, the output sequence from the generator is *eventually periodic*; it has an initial transient segment whose size depends roughly on how far c is from this interval.

The analysis of the MWC generator relies heavily on the number theoretic properties of the *connection integer*

$$m = -a_0 + \sum_{i=1}^{r} a_i b^i, \tag{5}$$

(so named because it plays the same role as the connection polynomial of a linear feedback shift register). It follows that m is relatively prime to b. Moreover, every $m > 0$ which is relatively prime to b has a unique representation of the form (5) with $0 \leq a_i < b$ $(0 \leq i \leq r)$ and with a_0 relatively prime to b (and $a_0 \neq 0$). In this paper, however, we allow the a_i to be arbitrary integers, so for a given connection integer m the representation (5) is not necessarily unique: one could even take $a_0 = -m$. It would be interesting to study to what extent the computations (3) and (4) might be optimized by appropriate choice of coefficients a_i.

As originally defined in [Marsaglia 1994; Couture and L'Écuyer 1994], the coefficient a_0 was equal to 1. If the base b is chosen to be a power of 2, then these

generators admit efficient implementations; however the connection integer will be constrained to be of the form $m = Nb - 1$ for some integer N. In this case, b is never a primitive root modulo m which implies (see Corollary 2.3) that the generator will never have maximal period. A similar criticism applies to the (original) subtract-with-borrow (SWB) generator. The introduction of a nontrivial value for a_0 (as first described in [Klapper and Xu 1996]) comes with the cost of two more multiplications per round, but it has the benefit that the connection integer m may be chosen so that b is primitive modulo m and this leads to properties (1), (2) and (4) listed above: (1) the period of the generator is $m-1$, which is maximal; (2) the d dimensional distribution properties of this generator are optimal, for each $d < r$ and (4) the modulus b may be taken to be a power of 2.

2.2

Throughout the rest of this section we fix a modulus b and consider the MWC generator corresponding to a connection integer m as in (5), where b is relatively prime to m.

Suppose $\sigma = (x_{-1}, x_{-2}, \ldots, x_{-r}; c)$ is a state of the generator. This state determines an integer

$$h = b^r c + a_0 \sum_{j=0}^{r-1} x_{-r+j} b^j - \sum_{k=1}^{r-1} b^k \sum_{i=1}^{k} a_i x_{-r+k-i}. \tag{6}$$

Conversely, the number h determines the state σ (an observation for which we thank an anonymous referee). For, reading equation (6) modulo b allows us to recover x_{-r} from h; then reading modulo b^2 and knowing x_{-r} allows us to recover x_{-r+1}. Continuing this way by induction we recover x_i for $-r \le i \le -1$. Finally, knowledge of these x_i and of h allows us to recover c. Several important properties of the state can best be expressed in terms of h (cf. Theorems 2.1 and 2.2).

Let us say that a state of the generator is degenerate if the output remains constant. The "bottom" state, in which all $x_i = 0$ and $c = 0$ is degenerate with output 0 (and $h = 0$). The "top" state, in which all $x_i = b - 1$ and

$$c = -a_0 + \sum_{i=1}^{r} a_i$$

is degenerate with output $b - 1$ (and $h = m$). (There may be more degenerate states.)

THEOREM 2.1. *The output sequence is strictly periodic if and only if* $0 \le h \le m$.

Define $B = b^{-1} \pmod{m}$ and represent it as a non-negative integer $0 < B < m$. Define A as in (2).

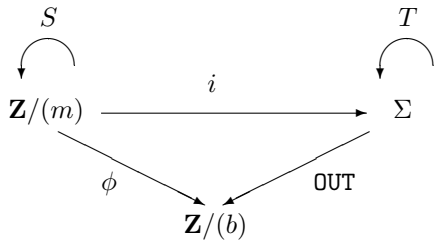
THEOREM 2.2. *Suppose the register is in a strictly periodic state. Then for all* $i \ge 0$ *we have*

$$x_{-r+i} = -A \left(hB^i \pmod{m} \right) \pmod{b}. \tag{7}$$

Equation (7) means that first the number $hB^i = hb^{-i}$ is computed modulo m, and is represented as a number between 0 and $m - 1$. Then this number is multiplied by $-A$ and reduced modulo b to give an integer between 0 and $b - 1$.

2.3

Let Σ be the set of all possible states $(x_{-1}, \ldots, x_{-r}; c)$ where $0 \le x_i < b$ and where $c \in \mathbf{Z}$. Let $i : \mathbf{Z}/(m) \to \Sigma$ be defined by (6), associating $h \in \mathbf{Z}/(m)$ with the state σ. Let $S : \mathbf{Z}/(m) \to \mathbf{Z}/(m)$ be the mapping $S(h) = Bh \pmod m$. Let $\phi : \mathbf{Z}/(m) \to \mathbf{Z}/(b)$ be the mapping $\phi(h) = -Ah \pmod b$. Theorem 2.2 says the following diagram "commutes", that is, $T(i(h)) = i(S(h))$ and $\phi(h) = \texttt{OUT}(i(h))$ for all $h \in \mathbf{Z}/(m)$.



COROLLARY 2.3. *If m is prime and if b is a primitive root modulo m then the period of the MWC generator is $m - 1$.*

In this case we say the resulting periodic sequence is a (generalized) ℓ-sequence (or long sequence), because of the many properties it shares with m-sequences (or maximal length sequences) from the theory of linear feedback shift registers and finite fields.

THEOREM 2.4. *Suppose m is prime and b is a primitive root* $\pmod m$. *Fix $d \ge 1$ and let $\mathbf{z} = (z_1, z_2, \ldots, z_d)$ with $0 \le z_i < b$. Then the number $N(\mathbf{z})$ of occurrences of the d-tuple $\mathbf{z}$ which begin in any fixed period of the sequence* (7) *can vary at most by 1. That is, $N(\mathbf{z})$ is either*

$$\left\lfloor \frac{m-1}{b^d} \right\rfloor \quad \textit{or} \quad \left\lfloor \frac{m-1}{b^d} \right\rfloor + 1.$$

In particular, if $b^d < m - 1$ then every d-tuple occurs at least once in any fixed period.

3. BOUNDS ON THE CARRY

3.1

Throughout this section we consider a MWC generator of order r with base b, coefficients $a_0, a_1, \ldots, a_r$ and state $\sigma = (x_{-1}, x_{-2}, \ldots, x_{-r}; c)$ as described in §2.1. Recall that c and a_i are integers and that $0 \le x_i < b - 1$. We show that the carry rapidly converges to a narrow range.

There are two generators which we refer to a the *extremal* generators. The first extremal generator has $a_0 > 0$ and all the remaining coefficients $a_i \le 0$ (for $1 \le i \le r$). The second has $a_0 < 0$ and all the remaining coefficients $a_i \ge 0$.

3.2

If $a_0 > 0$ define

$$w^+ = \sum_{\substack{a_i > 0 \\ 1 \le i \le r}} a_i \quad \text{and} \quad w^- = -a_0 + \sum_{\substack{a_i < 0 \\ 1 \le i \le r}} a_i.$$

If $a_0 < 0$ define

$$w^+ = -a_0 + \sum_{\substack{a_i>0 \\ 1\le i\le r}} a_i \quad \text{and} \quad w^- = \sum_{\substack{a_i<0 \\ 1\le i\le r}} a_i.$$

THEOREM 3.1. *Suppose the generator is not extremal. If the generator is in a strictly periodic state then the carry c lies in the range*

$$w^- < c < w^+. \tag{8}$$

If $c > w^+$ then it will drop monotonically and exponentially until it lies within this range and it will remain within this range thereafter. If $c < w^-$ then it will rise monotonically and exponentially until it lies within this range, and it will remain within this range thereafter. If the generator is extremal then c will move monotonically until it lies within the range

$$w^- \le c \le w^+$$

and it will remain within this range thereafter.

3.3 Proof

Let us assume $a_0 > 0$. (The proof for $a_0 < 0$ is completely parallel.) From (1), since $0 \le x_i \le b-1$ we have

$$c' = \frac{1}{b}\left[\sum_{i=1}^{r} a_i x_{-i} + c - a_0 x'_{-1}\right] \le \left(\frac{b-1}{b}\right) w^+ + \frac{c}{b}. \tag{9}$$

If $c < w^+$ this gives $c' < w^+$. If $c = w^+$ this gives $c' \le w^+$. If $c > w^+$ this gives

$$c' - c \le (w^+ - c)\left(\frac{b-1}{b}\right) < 0,$$

hence the carry decreases monotonically. Moreover, if $c > 0$ then $c' - w^+ \le (c - w^+)/b$, which is to say that $c - w^+$ decreases exponentially. It is easy to see that there are no strictly periodic states with $c = w^+$ unless the generator is extremal. For if $c = c' = w^+$ then (1) gives

$$(b-1)w^+ = \sum_{i=1}^{r} a_i x_{-i} - a_0 x'_{-1}.$$

The right side of this equation achieves its maximum value, $(b-1)w^+$, when $x'_{-1} = 0$ and

$$x_{-i} = \begin{cases} 0 & \text{whenever } a_i < 0 \\ b-1 & \text{whenever } a_i > 0. \end{cases}$$

Eventually this $0 = x'_{-1}$ will get shifted into one of the positions where $a_i > 0$ and then the value of c will drop below w^+. (If $a_i \le 0$ for all i, then the generator is extremal, $w^+ = 0$, and the degenerate "bottom" (all-zero) state satisfies $c = w^+$.) In summary, if the generator is not extremal and if the carry starts out at any positive value, it will drop until $c < w^+$ and will remain there forever.

To obtain the lower bound on c, equation (1) gives

$$c' = \frac{1}{b}\left[\sum_{i=1}^{r} a_i x_{-i} + c - a_0 x'_{-1}\right] \geq \frac{b-1}{b} w^- + \frac{c}{b}.$$

If $c > w^-$ then $c' > w^-$. If $c = w^-$ then $c' \geq w^-$. If $c < w^-$ then

$$c' - c \geq \left(\frac{b-1}{b}\right)(w^- - c) > 0$$

so the value of c will increase monotonically (and exponentially). Let us examine the possible periodic states with $c = c' = w^-$. For such a state, equation (4) gives

$$(b-1)w^- = \sum_{i=1}^{r} a_i x_{-i} - a_0 x'_{-1}.$$

The right side of this equation achieves its minimum value, $(b-1)w^-$ when $x'_{-1} = b-1$ and

$$x_i = \begin{cases} b-1 & \text{whenever } a_i < 0 \\ 0 & \text{whenever } a_i > 0. \end{cases}$$

If some coefficient a_i is positive (which is to say, if the generator is not extremal) then this $b-1 = x'_{-1}$ will eventually be shifted into the ith position, and the value of c will rise above w^-. However, if the generator is extremal (that is, if $a_i \leq 0$ for $1 \leq i \leq r$) then this argument fails and indeed, the degenerate "top" state satisfies $c = w^-$ and $x_i = b-1$ for all i. In summary, if the generator is not extremal and if the carry starts out at some negative value, then it will rise until $c > w^-$ and it will remain there forever. □

4. LATTICE STRUCTURE

4.1

Consecutive d-tuples $(x_k, x_{k+1}, \ldots, x_{k+d-1})$ of numbers generated by the MWC generator (1) do not form a d-dimensional lattice. However in [Couture and L'Écuyer 1994], R. Couture and L. Écuyer discovered the remarkable fact that these vectors very nearly lie on the lattice of vectors formed by the associated linear congruential generator with base b, multiplier B, and modulus m. To be precise, using the notation of §2.3, we have the following result.

THEOREM 4.1. *[Couture and L'Écuyer 1994] For every $z \in \mathbf{Z}/(m)$,*

$$\frac{\phi(z)}{b} \leq \frac{S(z)}{m} \leq \frac{\phi(z)+1}{b}.$$

The sequence of numbers $z, S(z), S^2(z), \ldots$ form the output of the LCG with base b, multiplier B, and modulus m (so consecutive d-tuples in this sequence form vectors which lie on a lattice in $\mathbf{R}^d$).

4.2 Proof

Consider $z \in \mathbf{Z}/(m)$ to be an integer $0 \leq z \leq m-1$. Since b is relatively prime to m there exists a unique $u \in \mathbf{Z}/(m)$ so that $bu \equiv z \pmod{m}$, or $u = S(z) =$

$b^{-1}z \pmod m$. Realizing u as an integer, $0 \le u \le m-1$ gives

$$bu = z + em \tag{10}$$

from which it also follows that $0 \le e \le b-1$. (On the one hand, $z = bu - em < (b-e)m$ implies $e \le b-1$. On the other hand, $bu = z + em < (1+e)m$ implies $e \ge 0$.) Dividing (10) by m gives $e \le bu/m \le e+1$ while reading (10) modulo b gives $z = -ea_0 \pmod b$ or $e = \phi(z)$. Hence, $\phi(z) \le bS(z)/m \le \phi(z)+1$. □

5 b-ADIC NUMBERS

5.1

As in §2.1 we fix a base b and consider the MWC generator which corresponds to a connection integer m of equation (5). In the literature, it is customary to analyze this generator by associating to each fraction h/m its fractional digital expansion in base b. Instead, we use the equivalent, but more abstract expansion of $-h/m$ as an element of the ring $\mathbf{Z}_b$ of b-adic numbers. (One expansion is just the reverse of the other.) The proofs become cleaner since various number-theoretic operations, such as $\pmod b$, may be applied to elements of $\mathbf{Z}_b$.

5.2

A b-adic number (or, more precisely, a b-adic integer) $\alpha \in \mathbf{Z}_b$ is a formal power series,

$$\alpha = x_0 + x_1 b + x_2 b^2 + \dots \tag{11}$$

with $0 \le x_i < b$. The sequence $x_0, x_1, \dots$ is referred to as the coefficient sequence of α. Addition and multiplication in $\mathbf{Z}_b$ are performed "with carry." That is, $xb^r + (b-x)b^r = b^{r+1}$. It is clear that $\mathbf{Z}_b$ contains the positive integers, but it also contains the negative integers since $-1 = (b-1) + (b-1)b + (b-1)b^2 + \dots$ as may be seen by adding 1 to both sides. It also contains all fractions of the form h/m where m is relatively prime to b. In fact, if a positive integer m is expanded in base b,

$$m = m_0 + m_1 b + \dots + m_r b^r \tag{12}$$

then m is relatively prime to b if and only if m_0 is invertible in $\mathbf{Z}/(b)$. Then

$$\frac{1}{m} = a_0 + a_1 b + \dots \tag{13}$$

where $m_0 a_0 \equiv 1 \pmod b$ and where the higher order coefficients a_i may be computed, one at a time, by substituting (12) and (13) in the equation $m\frac{1}{m} = 1$.

It is easy to see that the fractions $\alpha = h/m$ (with $h, m \in \mathbf{Z}$ and m relatively prime to b) are precisely the elements of $\mathbf{Z}_b$ whose coefficient sequence (11) is eventually periodic. We also refer to the coefficient sequence as the *b-adic expansion* of h/m.

The ring of b-adic numbers $\mathbf{Z}_b$ is isomorphic to the direct product $\prod_p \mathbf{Z}_p$ of the p-adic numbers $\mathbf{Z}_p$ over all prime factors b of p.

5.3

By summing the relevant geometric series, it is easy to see that the fractions h/m with $-m \le h \le 0$ are precisely the elements of $\mathbf{Z}_b$ whose coefficient sequence is *strictly* periodic (cf. [Klapper and Goresky 1997] Thm. 2.1). The case $h = 0$

corresponds to the coefficient sequence $0, 0, \ldots$ and the case $h = -m$ corresponds to the coefficient sequence $b-1, b-1, \ldots$

5.4

Now suppose we have a MWC generator with base b and with parameters $a_0, a_1, \ldots, a_r$, where a_0 is relatively prime to b. Let $m = -a_0 + \sum_{i=1}^{r} a_i b^i$ be the connection integer as in equation (5). Choose a seed state $\sigma = (x_{-1}, x_{-2}, \ldots, x_{-r}; c)$ as in §2.1. The output sequence x_{-r}, x_{-r+1}, $\ldots, x_{-1}$, x_0, $x_1, \ldots$ correspond to the following b-adic number

$$\alpha = x_{-r} + x_{-r+1}b + \ldots + x_0 b^r + x_1 b^{r+1} + \ldots \tag{14}$$

LEMMA 5.1. *Let $\sigma = (x_{-1}, x_{-2}, \ldots, x_{-r}; c)$ be the seed state of the generator and define the integer $h \in \mathbf{Z}$ by equation* (6)*. Then the resulting b-adic number α is the b-adic expansion of the fraction $-h/m$.*

5.5 Proof

This is a special case of [Klapper and Xu 1996] Theorem 3. Alternatively, one may easily adapt the proof of [Klapper and Goresky 1997] Theorem 4.1, replacing 2 by b. (In both cases, the proof is parallel to the original method of [Golomb 1982] Sect. 2.5.) □

5.6 Proof of Theorem 2.1

This follows immediately from Lemma 5.1 and §5.3. □

5.7 Proof of Theorem 2.2

To a given state $\sigma = (x_{-1}, x_{-2}, \ldots, x_{-r}; c)$ we associate the b-adic number $f(\sigma) = \alpha$ of (14). By Lemma 5.1, $\alpha = -h/m$ for some integer h. (The precise value of h is given by (6) however this fact is not needed for the argument.) If $\sigma' = (x'_{-1}, \ldots, x'_{-r}; c')$ represents the next state then

$$f(\sigma') = x_{-r+1} + x_{-r+2}b + \ldots = -h'/m$$

for some integer h'. So the following equation holds in $\mathbf{Z}_b$:

$$bf(\sigma') + x_{-r} = f(\sigma)$$

or

$$h = bh' - mx_{-r}. \tag{15}$$

Although this is an equation in $\mathbf{Z}_b$, all the terms are integers, so it is an equality among integers. Reading this equation modulo b gives $x_{-r} \equiv -m^{-1}h \equiv -Ah \pmod{b}$ (since $m \equiv a_0 \pmod{b}$). In other words, the output is $\mathtt{OUT}(\sigma) = -Ah \pmod{b}$.

Reading equation (15) modulo m gives $h' \equiv Bh \pmod{m}$. Now suppose the state σ is a nonzero, strictly periodic state. Then the same is true for σ', hence by Theorem 2.1, $0 < h, h' < m$. So we have the following equality,

$$h' = Bh \pmod{m},$$

provided we interpret the instructions $\pmod{m}$ to mean: reduce modulo m then represent this quantity as an integer between 0 and $m-1$.

It follows that the ith state $\sigma^{(i)}$ will correspond to the fraction $f(\sigma^{(i)}) = h^{(i)}/m$ where $h^{(i)} = B^i h \pmod{m}$, and the output will therefore be

$$\mathtt{OUT}(\sigma^{(i)}) = -Ah^{(i)} \pmod{b} = -A(B^i h \pmod{m}) \pmod{b}. \quad \square$$

5.8 Proof of Theorem 2.4

(This proof is parallel to [Klapper and Goresky 1997] Sect. 13.4.) A purely periodic nonzero sequence $\mathbf{x} = (x_0, x_1, \ldots)$ with connection integer m is the b-adic expansion of a rational number $-h/m$ with $0 < h < m$. Since b is chosen to be primitive, the different nonzero choices of h correspond to cyclic shifts of $\mathbf{x}$. Thus, a d-digit subsequence $\mathbf{z} = (z_1, z_2, \ldots, z_d)$ occurs in $\mathbf{x}$ if and only if it occurs as the first d digits in the b-adic expansion of some rational number $-h/m$. Moreover, two such rational numbers $-h_1/m$ and $-h_2/m$ have the same first d digits if and only if

$$-\frac{h_1}{m} \equiv -\frac{h_2}{m} \pmod{b^d},$$

that is, if and only if $h_1 \equiv h_2 \pmod{b^d}$. So we only need to count the number of h with a given first d digits and with $0 < h < m$.

Suppose that $b^r < m < b^{r+1}$. If $d > r$ then there is at most one such h and the result follows. Thus we may assume that $d \le r$. Now we count the number of possible h with $0 < h < m$ whose first d digits are fixed. Write

$$h = (h_0 + h_1 b + \ldots + h_{d-1} b^{d-1}) + b^d (h_d + \ldots + h_r b^{r-d}) = h' + b^d h'' \tag{16}$$

with $0 \le h_i < b$. Similarly set $m = m' + b^d m''$. Then

$$m'' = \left\lfloor \frac{m}{b^d} \right\rfloor = \left\lfloor \frac{m-1}{b^d} \right\rfloor$$

and

$$0 \le h', m' < b^d. \tag{17}$$

First note that $h'' \le m''$. For if $h'' \ge m'' + 1$ then

$$h \ge b^d h'' \ge b^d m'' + b^d > b^d m'' + m' = m$$

which contradicts $h < m$. We now consider two cases.

Case 1: $h' \ge m'$ Every choice of $h'' \le m'' - 1$ will give $0 < h < m$ since, by (17),

$$h = h' + b^d h'' < b^d + b^d h'' \le b^d + b^d (m'' - 1) \le b^d m'' + m' = m.$$

There are m'' such choices.

Case 2: $h' < m'$ Any choice of $h'' \le m''$ will give $h < m$. If $h' \ne 0$ then then all such choices give $0 < h < m$ and there are $m'' + 1$ possible such choices. If $h' = 0$ then all nonzero choices of $h'' \le m''$ give $0 < h < m$ and there are m'' such choices. $\square$

We remark that if $b = 2$ and $d = 1$ (that is, when counting the number of occurrences of a single bit in a binary ℓ-sequence), then $m' = 1$ so the two cases are: $h' = 1$ and $h' = 0$. In particular, Case 2 with $h' \ne 0$ never occurs. In other words, the sequence (7) is balanced: the number of 0's equals the number of 1's, and this number is $(m-1)/2$.

6. EXAMPLES

6.1

Let $m > 2$ be a prime number and let $b = 2^\omega$ with $\omega \geq 1$. Then b is a primitive root modulo m if and only if 2 is a primitive root modulo m and ω is relatively prime to $m-1$. Moreover 2 is *not* primitive modulo m if and only if

$$2^{(m-1)/p} \equiv 1 \pmod{m} \tag{18}$$

for some prime factor p of $m-1$. If 2 is a primitive root modulo m then $m \equiv 3$ or $5 \pmod 8$. These facts make it fairly easy to find large primes for which 2 is a primitive root. The following examples were found in a few hours using MAPLE. They use auxiliary primes p and q. According to Theorem 2.4, in each of these cases, the resulting MWC generator will have period $m-1$ and the resulting d-tuples will be uniformly distributed, with every d-tuple occurring whenever $d \leq d_0$. The last column, $T = m-1$, gives the approximate period of the generator.

b	p	q	m	d_0	T
2^{21}	$b^{14} - b^2 + 1$	$b^{58} - b^{36} + 1$	$4pq + 1$	71	10^{455}
2^{21}	$b^{52} - b^7 - 1$		$4p^2 + 1$	103	10^{657}
2^{21}	$b^{60} - b^{13} - 1$	$b^{60} - b^{26} - 1$	$2pq + 1$	119	10^{758}
2^{23}	$b^{12} + b^7 + 1$	$b^{25} + b^{19} + 1$	$2pq + 1$	37	10^{256}
2^{23}	$b^{14} - b^7 - 1$	$b^{27} + b^{26} + 1$	$4pq + 1$	41	10^{284}
2^{24}	$b^{48} - b^{46} - b^{38} - b^{14} + 1$		$2p + 1$	47	10^{347}
2^{24}	$b^{41} - b^{38} - 2b^{14} + 1$		$2p + 1$	40	10^{296}
2^{25}	$b^6 - b^4 - 1$	$b^{16} - b^{11} - 1$	$2pq + 1$	21	10^{166}
2^{31}	$b^7 + b^4 + 1$	$b^{30} + b^{14} - 1$	$4pq + 1$	37	10^{345}
2^{32}	$b^{33} - b^{20} - b^{14} - b^{11} - b^4 + 1$		$4p + 1$	32	10^{318}
2^{33}	$b^3 + b^2 + 1$	$b^{27} + b^{14} + 1$	$4pq + 1$	30	10^{298}
2^{35}	$b^2 + b - 1$	$b^{41} - b^{28} + 1$	$4pq + 1$	43	10^{453}

Coveyou, R. and MacPherson, R. 1967. Fourier analysis of uniform random number generators. *J. ACM 14*, 100–119.

Couture, R. and L'Écuyer, P. 1994. On the lattice structure of certain linear congruential sequences related to AWC/SWB generators. *Math. Comp. 62*, 799–808.

Couture, R. and L'Écuyer, P. 1997. Distribution properties of multiply-with-carry random number generators. *Math. Comp. 66*, 591–607.

Golomb, S. 1982. *Shift Register Sequences*. Aegean Park Press, Laguna Hills, CA.

Klapper, A. and Goresky, M. 1993. Feedback shift registers, combiners with memory, and arithmetic codes. *Univ. of Kentucky Dept. of Comp. Sci. Tech. Rep. No. 239-93.*

Klapper, A. and Goresky, M. 1994. 2-adic shift registers. In *Fast Software Encryption, Cambridge Security Workshop, Cambridge UK, December, 1993*, R. Anderson, Ed. *Lecture Notes in Computer Science 809*, Springer Verlag, N.Y., 174-178.

Klapper, A. and Goresky, M. 1997. Feedback shift registers, 2-adic span, and combiners with memory. *J. Crypt. 10*, 111–147.

Klapper, A. and Xu, J. 1996. Algebraic feedback shift registers. *Theor. Comp. Sci. 226*, 61–92.

L'Écuyer, P. 1996. Maximally equidistributed combined Tausworthe generators. *Math. Comp. 65*, 8-30.

Marsaglia, G. 1992. The mathematics of random number generators. In *The Unreasonable Effectiveness of Number Theory, Proc. Symp. Pure Math. 46*, Amer. Math. Soc., Providence R.I., 73-90.

Marsaglia, G. 1994. yet another rng. Posted to electronic bulletin board *sci.stat.math*, Aug 1, 1994.

Marsaglia, G. and Zaman, A. 1991. A new class of random number generators. *Ann. Appl. Probabl. 1*, 462–480.

Matsumoto M. and Nishimura T. 1998. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Computer Simulation 8*, 3-30.

Tezuka, S., L'Ecuyer, P., and Couture, R. 1993. On the lattice structure of add-with-carry and subtract-with-borrow random number generators. *ACM Trans. Model. Comp. Sim. 3*, 315–331.

...