Using irreducible polynomials for random number generation

Tamás Herendi\textsuperscript{a}\thanks{The author has been partially supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund.}, Sándor Roland Major\textsuperscript{b}\thanks{The author has been partially supported by the SETIT Project (no. 2018-1.2.1-NKP-2018-00004), which has been implemented with the support provided by the National Research, Development and Innovation Fund of Hungary, financed under the 2018-1.2.1-NKP funding scheme.}

\textsuperscript{a}Department of Computer Science
University of Debrecen
Debrecen, Hungary
herendi.tamas@inf.unideb.hu

\textsuperscript{b}Department of Information Technology
University of Debrecen
Debrecen, Hungary
major.sandor@inf.unideb.hu

Abstract. A method is presented for generating random numbers with uniform distribution using linear recurrence sequences with very large period lengths. This method requires an irreducible polynomial modulo 2 to define the sequence. A suitable method for generating an infinite number of such polynomials is presented. The polynomials generated in this way can have an arbitrarily large degree, and a large enough order to make them suitable for practical applications.

Keywords: irreducible polynomials, finite fields, random number generation

AMS Subject Classification: 12-08 Computational methods for problems pertaining to field theory

1. Introduction

Pseudorandom number generation (PRNG) is an important component of many practical applications. Generators with different properties are used in a wide
range of fields, such as simulations [12], Monte Carlo methods [14]. Moreover, PRNGs recently play an essential role in many areas of cryptography, for example, key generation, stream ciphers, asymmetric cryptosystems, and authentication protocols [11].

The results presented in this paper relate to an algorithm detailed in [8], showing the construction of uniformly distributed linear recurrence sequences (LRS) modulo powers of 2, with theoretically arbitrarily large period lengths. A modified version of this algorithm is given in Section 3, optimizing it to be less computationally expensive.

2. Theory

The algorithm presented takes an irreducible polynomial over $\mathbb{F}_2$ as input.

Irreducible polynomials over finite fields are used in a wide variety of contexts, not just in pure mathematics and many areas of computer science, but practical applications as well.

In [8], we see a method of constructing linear recurring sequences with extremely long periods used for pseudorandom number generation. The sequence requires an irreducible polynomial to create, the degree of which is directly related to the resulting period length.

In coding theory, creating error correcting codes that can be used to reliably transmit information over noisy channels is a key practical application that can be found in many everyday electronic systems. These codes are almost always connected to the use of polynomials over finite fields. An in-depth discussion can be found in [9].

In cryptography, many encryption protocols use finite fields as their domain. Irreducible polynomials have been used in public key cryptosystems for decades, such as in [4].

As the previous examples show, irreducible polynomials over the finite field $\mathbb{F}_2$ are of special interest.

2.1. Irreducible polynomials

A univariate polynomial over the finite field $\mathbb{F}_2$ is

$$p(x) = \sum_{k=0}^{n} a_k x^k, \quad a_0, \ldots, a_n \in \mathbb{F}_2.$$

$\mathbb{F}_2[x]$ is the set of all polynomials over $\mathbb{F}_2$. A polynomial $p \in \mathbb{F}_2[x]$ of degree $k$ is irreducible if it has no nontrivial factors over $\mathbb{F}_2$. That is, $p(x) = p_1(x)p_2(x)$ can not hold if $\deg(p_1), \deg(p_2) > 0$. The natural way to prove a polynomial’s irreducibility is, therefore, to factor it and show that no such factors can be found.

The first algorithm for factoring a polynomial over a finite field was published by Berlekamp [2]. It is a deterministic algorithm that requires a square-free polynomial, and is well suited for cases where the cardinality of the finite field is small.
Later, the Cantor-Zassenhaus algorithm [3] provided a practical solution even for polynomials over large finite fields. This algorithm is probabilistic in nature. A detailed description of both methods can be found in [13].

Rabin’s test [16] provides a very simple algorithm. A polynomial over $\mathbb{F}_2$ is irreducible if and only if:

1. $p(x) \mid x^{2^k} - x$
2. $\forall t_i \ GCD(x^{2^k/t_i}, p(x)) = 1$,

where $t_i$ are the prime divisors of $k$. The test simply computes all $x^{2^k/t_i} \mod p(x)$ polynomials using repeated squaring, and polynomial modulo operations, then uses polynomial GCD to check condition 2.

Ben-Or’s test [1] modifies this approach by computing $GCD(x^{2^i} \mod p(x), p(x))$ for every $i \in \{1, \ldots, \frac{k}{2}\}$. In practice, this improves average performance when testing random polynomials. A randomly selected polynomial is much more likely to have factors of small degrees than be the product of only large-degree factors. Since Ben-Or’s test checks for factors of small degrees first, these polynomials are very quickly eliminated. A comparison between the performance of Rabin’s test and Ben-Or’s test can be found in [7]. Victor Shoup also published a deterministic irreducibility test in [19], and a probabilistic algorithm is [18].

3. Algorithm for creating LRS

The following algorithm is for constructing uniformly distributed linear recurrence sequences modulo $2^s$, with very large period lengths. It is a modified version of the algorithm found in [8]. The version presented here is significantly less computationally expensive than the original, which enables the creation of sequences with larger period lengths.

The reduced time complexity speeds up the process of finding the desired coefficients for the LRS, while the reduced space complexity allows the algorithm to be carried out with significantly larger input parameters. Once the LRS is constructed, using it to generate the pseudorandom number sequence is unchanged compared to the original version.

1. Choose an integer $k$ and find a monic polynomial $q(x) \in \mathbb{Z}[x]$ of degree $k$, which reduction modulo 2 is irreducible in $\mathbb{F}_2[x]$.

2. Calculate the polynomials $p(x)$ of degree $k + 2$ and $p'(x)$ of degree $k + 1$ in the following way:

\[
p(x) \equiv (x^2 - 1)q(x) \mod 2 \quad \text{and} \quad p'(x) \equiv (x - 1)q(x) \mod 2,
\]

with the coefficients of $p(x)$ and $p'(x)$ in $\{0, -1\}$, except for the leading coefficients.
Calculate the four candidate polynomials:

\[ p_1(x) = p(x) \]
\[ p_2(x) = p(x) - 2 \]
\[ p_3(x) = p(x) - 2x \]
\[ p_4(x) = p(x) - 2x - 2 \]

We remark that the coefficients of the constant and linear terms of the candidate polynomials can be in \( \{0, -1, -2, -3\} \).

3. For \( i \in \{1, 2, 3, 4\} \), \( j \in \{0, 1, \ldots, k + 2\} \), let \( a_{ij} \) denote the coefficient of \( x^j \) in the polynomial \( p_i(x) \). Calculate \( S_i = \sum_{j=0}^{k+1} a_{ij} \) for each candidate polynomial. Keep the two candidates that satisfy \( S_i \equiv 1 \mod 4 \). Denote these two polynomials with \( c_1 \) and \( c_2 \).

4. Let \( \rho = \text{ord}(q) \) be the order of \( q(x) \), i.e., the smallest positive integer such that \( q(x) | x^\rho - 1 \).

We need to find the candidate that satisfies \( c_i(x) \nmid x^{2\rho} - 1 \mod 4 \). To do this, calculate

\[ r(x) \equiv x^\rho \mod (2, p(x)), \]

where \( \mod(2, p(x)) \) means calculating the polynomial remainder with \( p(x) \) over \( \mathbb{F}_2 \).

Then, find the candidate that satisfies

\[ 1 \not\equiv r(x)^2 \mod (4, c_i(x)), \]

where \( \mod(4, c_i(x)) \) means calculating the polynomial remainder with \( c_i(x) \) over \( \mathbb{F}_4 \).

Note that all of the computation in this step can be performed over \( \mathbb{F}_2 \), with the exception of the last step, which is performed over \( \mathbb{F}_4 \).

Denote the candidate that remains by \( c(x) \). This is the characteristic polynomial of the linear recurrence sequence we want to create. Let \( b_j, j \in \{0, 1, \ldots, k + 2\} \) be the coefficient of \( x^j \) in \( c(x) \). Then, our final recurrence relation is

\[ u_{n+k+2} = -b_{k+1} u_{n+k+1} - b_k u_{n+k} \ldots - b_0 u_n \]

5. Choose initial values for the sequence. Suppose we want \( s \)-bit long pseudorandom numbers. Choose random \( u_0, u_1, \ldots, u_k \in [0, 2^s - 1] \). Set these values as the initial values of the linear recurrence relation with characteristic polynomial \( p'(x) \). Compute the next element of the sequence, \( u'_{k+1} \). Find a random number \( u_{k+1} \in [0, 2^s - 1] \) such that \( u'_{k+1} \not\equiv u_{k+1} \mod 2 \).

Set \( u_0, u_1, \ldots, u_{k+1} \) as the initial values of the sequence.
The original version of the above algorithm differs in steps 3 and 4. That algorithm requires computation using the companion matrices of the candidates. The companion matrix of $p_i(x)$ is

$$M_{(i)} = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_{i0} & -a_{i1} & \cdots & -a_{ik} & -a_{i,k+1} \end{pmatrix}$$

Step 3 calls for finding the two candidates that satisfy $M_{(i)} \vec{1} \equiv \vec{1} \mod 4$, where $\vec{1}$ is a vector of size $k + 2$ with all coordinates equal to 1. It is simple to see that this is equivalent to the step 3 of the above algorithm.

In step 4 of the old algorithm, the final candidate is the one for which $M^2_{(i)} \vec{\varphi} \not\equiv \vec{E} \mod 4$ holds, where $E$ is the identity matrix of size $(k+2) \times (k+2)$. This requires matrix exponentiation modulo 4, using matrices sized $(k+2) \times (k+2)$. Since we want $k$ to be as large as possible, these matrices quickly become inconvenient, both to store, and to perform multiplications on. The step 4, presented here, instead requires mainly polynomial modulo and squaring over $F_2$, and once over $F_4$. This makes storage more efficient, since the size of a polynomial is a linear function of its degree. Moreover, even a naive implementation of polynomial squaring over $F_2$ has time complexity $O(n)$. For polynomial modulo, a naive approach has time complexity $O(n^2)$, but faster algorithms are known, such as the result by Schönhage in [17], which enables polynomial division with remainder in $O(n \log n \log \log n)$ time. This is better than even the fastest current matrix multiplication algorithms, such as [6], which has a complexity over $O(n^{2.3})$.

4. Q-transform

A promising concept for constructing uniformly distributed high order linear recurring sequences is the application of the theory of $Q$-transform. In this section, we introduce some definitions and results that allow us to formulate infinite series of irreducible polynomials. Based on the idea described in Section 3, we can use such polynomials for creating uniformly distributed pseudorandom sequences with large period lengths.

During the section, $q$ is a prime power, $F$ denotes a field, $F_q$ is a finite field of $q$ elements, and $K$ is an algebraic extension field of $F$ or $F_q$, depending on the context.

**Definition 4.1.** Let $p \in F[x]$ be a polynomial of degree $d$. We say that the reciprocal polynomial of $p$ is $p^*(x) = x^d p(x^{-1})$. We call a polynomial $p$ self-reciprocal, if $p = p^*$.

**Remark 4.2.**

a) If $p \in F[x]$, then $p^* \in F[x]$ and $(p^*)^* = p$. 


b) Let \( p, r, s \in \mathbb{F}[x] \) be such that \( s = p \cdot r \). Then \( s^* = p^* \cdot r^* \).

c) For any \( p \in \mathbb{F}[x] \), \( p \) is irreducible if and only if \( p^* \) is irreducible.

d) Let \( p, r, s \in \mathbb{F}[x] \) be such that \( s = p \cdot r \). If \( p \) and \( r \) are self-reciprocal, then \( s \) is self-reciprocal, as well.

**Proposition 1.** Let \( p \in \mathbb{F}[x] \), and \( \mathbb{K} \) be the splitting field of \( p \). Then the following statements are equivalent.

a) \( p \) is self-reciprocal;

b) \( \forall \alpha \in \mathbb{K} \setminus \{0\} : p(\alpha) = 0 \implies p(\alpha^{-1}) = 0 \).

**Proof.** Since \( \deg(p) \) is odd, Proposition 1 implies that \( p \) has a root \( \alpha \) such that \( \alpha = \alpha^{-1} \). This is possible if and only if \( \alpha \in \{-1, 1\} \). Then either \( x - 1 \) or \( x + 1 \) is a divisor of \( p \). However, \( p \) is irreducible, thus either \( p = ax - a \) or \( p = ax + a \), but \( ax - a \) is self-reciprocal if and only if \( a = -a \). \( \square \)

**Corollary 4.3.** If \( p \in \mathbb{F}[x] \) is self-reciprocal and irreducible of odd degree, then \( p(x) = ax + a \), with some \( a \in \mathbb{F} \).

**Proof.** Since \( \deg(p) \) is odd, Proposition 1 implies that \( p \) has a root \( \alpha \) such that \( \alpha = \alpha^{-1} \). This is possible if and only if \( \alpha \in \{-1, 1\} \). Then either \( x - 1 \) or \( x + 1 \) is a divisor of \( p \). However, \( p \) is irreducible, thus either \( p = ax - a \) or \( p = ax + a \), but \( ax - a = \alpha \) is self-reciprocal if and only if \( a = -a \). \( \square \)

**Corollary 4.4.** Let \( p \in \mathbb{F}[x] \) be a self reciprocal polynomial, and \( p_1, \ldots, p_k \in \mathbb{F}[x] \) be distinct irreducible polynomials such that \( p = p_1^{n_1} \cdots p_k^{n_k} \). Then for each \( 1 \leq i \leq k \) there exists \( 1 \leq j \leq k \) such that \( p_i = p_j^* \) and \( n_i = n_j \).

**Remark 4.5.** In the previous corollary, \( i = j \) if and only if \( p_i \) is self-reciprocal.

**Definition 4.6.** Let \( p \in \mathbb{F}[x] \) be a polynomial of degree \( d \). The \( Q \)-transform of \( p \) is \( \hat{p}(x) = x^d p(x + x^{-1}) \).

**Remark 4.7.** If \( p \in \mathbb{F}[x] \), then \( \hat{p} \in \mathbb{F}[x] \), and \( \deg(\hat{p}) = 2 \deg(p) \).

**Proposition 2.** Let \( p, r, s \in \mathbb{F}[x] \) be such that \( s = p \cdot r \). Then \( \hat{s} = \hat{p} \cdot \hat{r} \).

Let \( p \in \mathbb{F}[x] \), and \( \alpha \in K \setminus \{0\} \). Then \( \hat{p}(\alpha) = 0 \) if and only if \( \hat{p}(\alpha^{-1}) = 0 \). By Proposition 1, we may state the following.

**Proposition 3.** If \( p \in \mathbb{F}[x] \), then \( \hat{p} \) is self-reciprocal.

**Proposition 4.** The \( Q \)-transform is an injection.

**Proof.** Let \( p \in \mathbb{F}[x] \), \( d = \deg(p) \), \( K \) be the splitting field of \( \hat{p} \), and \( \alpha_i, \beta_i \in K \) \((i = 1, \ldots, d)\) with the following properties:

\[
p(x) = a_d \prod_{i=1}^{d} (x - \alpha_i), \quad \text{and} \quad \beta_i = -\frac{1}{2} \alpha_i + \frac{1}{2} \sqrt{\alpha_i^2 - 4}.
\]

Then

\[
\hat{p}(x) = a_dx^d \prod_{i=1}^{d} (x + x^{-1} - \alpha_i) = a_d \prod_{i=1}^{d} (x^2 + 1 - \alpha_i x)
\]

41
\[ = a_d \prod_{i=1}^{d} (x - \beta_i) \prod_{i=1}^{d} (x - \beta_i^{-1}). \]

This means that there is a one-to-one correspondence between the roots of \( p \) and the pairs of roots of \( \tilde{p} \).

Let

\[ \mathcal{P}_q(d) = \{ p \mid p \in \mathbb{F}_q[x], \deg(p) = d \}, \]
\[ \mathcal{Q}_q(d) = \{ p \mid p \in \mathbb{F}_q[x], \deg(p) = 2d, p = p^* \}. \]

Since \( |\mathcal{P}_q(d)| = |\mathcal{Q}_q(d)| \), Proposition 4 implies the following.

**Corollary 4.8.** Let \( p \in \mathbb{F}_q[x] \) be a self-reciprocal polynomial. Then there exists a unique \( r \in \mathbb{F}_q[x] \) such that \( p = \tilde{r} \).

**Notation 1.** Let \( p \in \mathbb{F}[x] \) and \( k \in \mathbb{N} \). We denote by \( \tilde{p}^{(k)} \) the following iterated \( Q \)-transform:

\[
\text{if } k = 0, \text{ then } \tilde{p}^{(k)} = p; \\
\text{if } k > 0, \text{ then } \tilde{p}^{(k)} = \tilde{r}, \text{ where } r = \tilde{p}^{(k-1)}. 
\]

**Corollary 4.9.** Let \( p \in \mathbb{F}_q[x] \) be a self-reciprocal polynomial. Then there exists a unique \( r \in \mathbb{F}_q[x] \), not a self-reciprocal polynomial, and \( k \in \mathbb{N} \) such that \( p = \tilde{r}^{(k)} \).

**Corollary 4.10.** Let \( p \in \mathbb{F}_q[x] \) be irreducible. Then \( \tilde{p} \) is either irreducible or there exist \( p_1, p_2 \in \mathbb{F}_q[x] \) irreducible polynomials such that \( \tilde{p} = p_1 \cdot p_2 \), and \( p_1 = p_2^* \).

**Proof.** Assume contrary that there exists an \( r \in \mathbb{F}_q[x] \) self-reciprocal polynomial with \( 1 \leq \deg(r) < 2 \deg(p) \), such that \( r|\tilde{p} \). By Corollary 4.8, there exists \( s \in \mathbb{F}_q[x] \) satisfying \( s \mid p, \deg(s) < \deg(p) \), and \( r = \tilde{s} \), which is a contradiction.

**Proposition 5.** Let \( p \in \mathbb{F}_2[x] \) be an irreducible polynomial in the form \( p(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1x + 1 \). Then \( \tilde{p} \) is irreducible if and only if \( a_{d-1} = a_1 = 1 \). Furthermore, the coefficient of the linear term of \( \tilde{p} \) is 1.

**Proof.** The proposition is proven in a more general settings in [10].

**Corollary 4.11.** Let \( p \in \mathbb{F}_2[x] \) be an irreducible polynomial, and \( p(x) = x^d + x^{d-1} + a_{d-2}x^{d-2} + \cdots + a_2x^2 + x + 1 \). Then \( \tilde{p}^{(k)} \) is irreducible for all \( k \in \mathbb{N} \).

This result implies that any irreducible polynomial in the form as in Proposition 5 determines an infinite sequence of irreducible \( Q \)-iterated polynomials. Every self-reciprocal polynomial of even degree is contained in exactly one of such sequences.

**Proposition 6.** Let \( p \in \mathbb{F}_q \) be an irreducible polynomial, accomplishing \( \deg(p) = 2d \). Then \( p \) is self-reciprocal if and only if \( \text{ord}(p) \mid q^d + 1 \).
Proof. The proposition is stated e.g. in [5].

For the construction of pseudorandom number sequences with high period length, we need irreducible polynomials of high order. Actually, the period length is proportional to the order. Based on our experience, we have the following conjecture.

Conjecture 1. Let \( p \in \mathbb{F}_q[x] \) be an irreducible self-reciprocal polynomial of degree \( \deg(p) = 4d \). Then \( q^d + 1 < \text{ord}(p) \).

Furthermore, we have encountered \( Q \)-iterated polynomials having maximal order in many cases.

5. Statistical testing

In this section, we describe a test carried out to examine the statistical properties of the pseudorandom number sequences generated using the previously detailed method. Two irreducible polynomials of large degree were created, one using a brute force method and one using \( Q \)-transformations. The pseudorandom sequences generated using these polynomials were tested using the NIST statistical test suite.

The software and documentation of the NIST test suite are available at [15]. The suite includes 15 tests designed to examine the properties of pseudorandom bit sequences, such as:

- **Frequency test**: a simple check to determine the proportion of ones and zeroes in a binary sequence.
- **Runs test**: checking the number of runs (uninterrupted sequence of identical bits) of various lengths to see how closely matches the expected value in a truly random sequence.
- **DFT (Spectral) test**: determining the peak heights in the Discrete Fourier Transform of the sequence, with the purpose of finding periodic features.
- **Template matching test**: finding occurrences of predetermined target strings, to detect generators producing too many such patterns. Both overlapping and non-overlapping tests are included.
- **Maurer’s “Universal Statistical” test**: checking whether or not the sequence can be significantly compressed without loss of information.
- **Linear complexity test**: attempting to determine the length of the LRS that characterizes the sequence.

The first irreducible polynomial tested, denoted by \( t_1 \), was generated using irreducibility testing methods described in previous sections. The implementation uses the NTL (Number Theory Library) available at [20].
The degree of $t_1$ was chosen to be 216091. The reason for this choice is that $2^{216091} - 1$ is a Mersenne prime. Choosing a value this way simplifies Step 4 of the algorithm described in Section 2. Note that this step requires the computation of the order of the irreducible polynomial, which is a divisor of $2^d - 1$, where $d$ is the degree of the polynomial. If $d$ is large, this step becomes computationally impractical, but choosing $2^d - 1$ to be a prime gives a simple solution to the problem.

The second irreducible polynomial tested, denoted $t_2$, was created using iterated $Q$-transform, using the following method:

1. Let $q$ be a self-reciprocal irreducible monic polynomial, with $\text{deg}(q) = d$.
2. Run the algorithm described in Section 3, using $q$ as input. Let $p$ be the candidate polynomial that remains after Step 4. Determine $s, r \in \mathbb{Z}[x]$ such that $p = sq + r$, and $\text{deg}(r) < \text{deg}(q)$.
3. Compute $t = s\tilde{q}^{(n)} + r$, where $\tilde{q}^{(n)}$ is the iterated $Q$-transform, described in Notation 1 of Section 4. Use $t$ to construct the linear recurrence sequence.

Based on practical observation, if the sequence produced by $p$ has uniform distribution, then the sequence produced by $t$ will also have uniform distribution. However, the proof of this conjecture is currently an open question.

To create $t_2$, the following polynomial was used as a starting point:

$$q_2 = x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^7 + x^5 + x^4 + x^3 + x^2 + x + 1.$$ 

As it is stated in Proposition 6, the order of a self-reciprocal irreducible monic polynomial of degree $d$ is at most $2^d + 1$. The above polynomial was chosen because $\text{ord}(q), \text{ord}(\tilde{q})$, and $\text{ord}(\tilde{q}^{(2)})$ all reach this maximum value. Theoretically, it does not guarantee that this maximality property will hold after further $Q$-transformations, but practical observations suggest that the order of $\tilde{q}^{(n)}$ will grow at a rate that is sufficient for use in the applications described in this paper. We stated our related experience in Conjecture 1.

Using the above method, $p_2 = s_2q_2 + r_2$ was determined, and the polynomial to be used was set as $t_2 = s_2\tilde{q}_2^{(14)} + r_2$. Note that $\text{deg}(\tilde{q}_2^{(14)}) = 229376$, and $\text{deg}(t_2) = 229378$.

Using $t_1$ and $t_2$, two LRSs were created to generate the pseudorandom sequences to be tested, denoted $L_1$ and $L_2$ respectively. Both LRSs generate 64-bit words. Following the recommendations in the documentation of the NIST test suite, 16MB ($2^{21}$ words) of test data were generated using $L_1$ and $L_2$ each.

For each of these two streams, the NIST suite split the data into 100 bitstreams. The testing software provides a detailed output of the tests, as well as a summary showing the number of bitstreams that passed each test. The minimum pass rate for a test is considered to be 96 out of a sample size of 100.

Tables 1 and 2 show some of the result obtained from the tests. The full report can be found at https://arato.inf.unideb.hu/major.sandor/statistical_results/.
Using irreducible polynomials for random number generation

The results show the two generators to have very similar statistical properties, with $L_2$ being only slightly weaker in some tests. Since the order of $t_2$ is significantly lower than the order of $t_1$, this result is to be expected. Also of note is that both generators produced very high quality pseudorandom sequences, passing all relevant benchmarks set by the test suite.

This shows that using the $Q$-transformation described above to generate irreducible polynomials of very large degree is completely suitable for use in generating uniformly distributed pseudorandom linear recurrence sequences.

References


