On the Linear Complexity and Lattice Test of Nonlinear Pseudorandom Number Generators

Domingo Gómez-Pérez
University of Cantabria
Avd. Los Castros, s/n, Santander, Spain
domingo.gomez@unican.es

Jaime Gutierrez
University of Cantabria
Avd. Los Castros, s/n, Santander, Spain
jaime.gutierrez@unican.es

Abstract

One of the main contributions which Harald Niederreiter made to mathematics is related to pseudorandom sequences theory. In this paper we study several measures for asserting the quality of pseudorandom sequences, involving generalizations of linear complexity and lattice tests and relations between them.

Keywords. Nonlinear pseudorandom number generators, linear complexity, lattice test, lattice points.

Mathematics subject classification number: 11K45.

1 Introduction

Let $\mathbb{F}_q$ be the finite field with $q$ elements, where $q$ is an arbitrary prime power. Through the paper, we only consider purely periodic sequences $S = (s_n) = (s_0, s_1, s_2, \ldots)$ of elements of $\mathbb{F}_q$ and we denote its period by $T$ with $T > 1$.

We recall that the Linear Complexity $L(S)$ of the sequence $S$ is the smallest positive integer $L$ for which there exist coefficients $a_0, a_1, \ldots, a_{L-1} \in \mathbb{F}_q$ such that

$$s_{n+L} = a_{L-1}s_{n+L-1} + \cdots + a_1s_{n-1} + a_0s_n, \quad \forall n \geq 0.$$
Note that $L(S)$ is the length of the shortest linear feedback shift register that can generate $S$, so the following inequalities $1 \leq L(S) \leq T$ hold. The linear complexity of sequences is an important security measure for stream cipher systems (see [2, 11, 16, 24]). A measure closely related to the linear complexity is the lattice test. For a given integer $L \geq 1$, $S$ passes the $L$-dimensional Lattice Test if the vectors

$$\{ \vec{s}_{n} - \vec{s}_0 : \vec{s}_n = (s_n, s_{n+1}, \ldots, s_{n+L-1}), \quad 0 \leq n < T \},$$

span $\mathbb{F}_q^L$. The reason of the existence of this test comes from the use of pseudorandom numbers generated by linear congruences, which were introduced by Lehmer in [9]. Although linear generators like the one mentioned are popular, they also comprise severe deficiencies that make them improper in many applications, such as cryptography. Even more general generators with low linear complexity turned out to be undesirable for more traditional applications in Monte Carlo methods as well, see [13, 14, 18].

One particularly undesirable feature of these pseudorandom number sequences is their coarse lattice structure. This is the reason, that Marsaglia in [10] proposed a test to measure this special structure. This test was investigated and enhanced by Harald Niederreiter and Arne Winterhof, see [19, 20].

However, this measure is closely related with the linear complexity. Relations between lattice test and linear complexity for parts of the period are given in [3, 4, 5, 6].

The importance of the relationship comes from the fact that linear complexity is a well understood concept. For example, it is known the exact value of the number of sequences of a given length and linear complexity on finite fields, see [22, Theorem 7.1.6] and [12]. Indeed, the lattice structure has been thoroughly studied in pseudorandom number sequences generated for Monte Carlo methods and stream ciphers (see [7, 14, 15, 17, 21]). The following is a natural generalization of the linear complexity. We define the Quasi-Linear complexity $QL(S)$ of the sequence $S$ as the smallest nonnegative integer $L$ for which there exist coefficients $a_0, a_1, \ldots, a_{L-1} \in \mathbb{F}_q$ and integers $0 < d_1 < \cdots < d_{L-1} < T$ such that

$$s_{n+d_L} = a_{L-1}s_{n+d_{L-1}} + \cdots + a_1s_{n+d_1} + a_0s_n, \quad \forall n \geq 0.$$ (1)

Obviously, we have that $1 \leq QL(S) \leq L(S)$ so, in particular $QL(S) \leq T$. We will see that this last concept coincides essentially with the Lattice Test introduced in [21]. For given integers $L \geq 1$, $0 < d_1 < \cdots < d_{L-1} < T$, $S$ passes the $L$-dimensional Lattice Test with lags $d_1, \ldots, d_{L-1}$ if the vectors

$$\{ \vec{s}_{n} - \vec{s}_0 : \vec{s}_n = (s_n, s_{n+d_1}, \ldots, s_{n+d_{L-1}}), \quad 0 \leq n < T \},$$

span $\mathbb{F}_q^L$. The greatest dimension $L$ such that $S$ satisfies the $L$-dimensional lattice test for all lags $d_1, \ldots, d_{L-1}$ is denoted by $T(S)$.

The main goal of this paper is comparing the three integers $L(S), QL(S)$ and $T(S)$. It is divided into five sections. In Section 2 we obtain the relation between $QL(S)$ and $T(S)$. The main result is presented in Section 3, where we
obtain a non-trivial inequality relating the linear complexity and the quasi-linear complexity. Section 4 is devoted to apply the presented results to some pseudorandom number generators and we show that our result rediscovers a special case of a previous result. Finally, we present an open problem in Section 5.

2 Lattice test and quasi-linear complexity

In this section we compare the two integers $QL(S)$ and $T(S)$ and obtain the main result of the section:

**Theorem 1.** With the above notation, we have

$$T(S) \leq QL(S) \quad \text{and} \quad QL(S) \leq 2T(S) + 2.$$

**Proof.** We fix the following notation, $QL(S) = M$ and $T(S) = L$, so by the definition of $QL(S)$, there exist $d_1, \ldots, d_M$ and $a_0, \ldots, a_{M-1} \in \mathbb{F}_q$ such that,

$$s_{n+1} = a_{M-1}s_{n+d_{M-1}} + \cdots + a_1s_{n+d_1} + a_0s_n, \quad \forall n \geq 0.$$  

Since $(a_0, a_1, \ldots, a_{M-1}, -1)$ is a non-zero vector, then we denote $H$ the following hyperplane

$$H = \{(x_0, \ldots, x_{M-1}, x_M) \in \mathbb{F}_q^{M+1} : a_0x_0 + \cdots + a_{M-1}x_{M-1} - x_M = 0\}.$$  

Moreover, the vectors $\tilde{s}_n = (s_n, s_{n+d_1}, \ldots, s_{n+d_{M-1}}, s_{n+d_M}) \in H$, for $0 \leq n < T$. So, $\{\tilde{s}_n - \tilde{s}_0 : 0 \leq n < T\} \subset H$ which implies $L < M + 1$ thus $T(S) \leq QL(S)$.

On the other hand, by definition of $T(S)$, there exist integers $0 < d_1 < \cdots < d_{L-1} < d_L < T$ such that the vector space $V$ generated by $\{\tilde{s}_n - \tilde{s}_0 : 0 \leq n < T\}$ is strictly contained in $\mathbb{F}_q^{L+1}$, where $\tilde{s}_n = (s_n, s_{n+d_1}, \ldots, s_{n+d_{L-1}}, s_{n+d_L}) \in \mathbb{F}_q^{L+1}$, for $0 \leq n < T$. So, there exits a non-zero vector $\vec{w} = (w_0, \ldots, w_L)$ satisfying $\langle \vec{w}, \tilde{s}_n - \tilde{s}_0 \rangle = 0$, where $\langle \cdot, \cdot \rangle$ denotes the usual inner product. We denote by $\delta$ the inner product $\langle \vec{w}, \tilde{s}_0 \rangle$, then for $0 \leq n$ the following equations holds:

$$w_0s_n + w_1s_{n+d_1} + \cdots + w_Ls_{n+d_L} = \delta, \quad w_0s_{n+1} + w_1s_{n+1+d_1} + \cdots + w_Ls_{n+1+d_L} = \delta, \quad (2)$$

Notice that $w_L \neq 0$ because $T(S) = L$, from the Equations (2) we have

$$w_Ls_{n+d_L+1} = w_0s_n - w_0s_{n+1} + \ldots + w_Ls_{n+d_{L-1}} - w_{L-1}s_{n+1+d_{L-1}} + w_Ls_{n+d_L}.$$  

Now, we distinguish two cases: the previous equation is trivial or not. Notice that if the previous equation is not trivial, then it is of the form of (1). In other case, the lags satisfy the following relation,

$$d_i = i \mod T, \quad i = 1, \ldots, L.$$  

This implies that $L = T - 1$ so the inequality is satisfied trivially. 

\[\square\]
The following example shows that we cannot relax the first of the inequalities:

**Example 1.** The following sequence is defined in any field of odd characteristic, so these elements \( \{0, 1, -1\} \) are different. Take \( S = (s_0, s_1, \ldots) \) to be the sequence with even period defined by the following function, 

\[
s_i = \begin{cases} 
-1, & \text{if } i = T/2 - 1, \\
1, & \text{if } i = T - 1, \\
0, & \text{otherwise.}
\end{cases}
\]

It is clear that \( QL(S) = 1 = T(S) \), the reason is that the sequence satisfies the following recurrence:

\[
s_{n + T/2} = -s_n, \quad \forall n \geq 0.
\]

It is even easy to see that \( L(S) = T/2 \).

### 3 Quasi-linear and linear complexity

In this section we give a relationship between two measures, the linear complexity and quasi-linear complexity, under the extra condition that the period is a power of a prime number, without any extra conditions on the field \( \mathbb{F}_q \).

**Theorem 2.** If \( T \) is a power of prime number \( P \), \( T = P^t \), then the following inequality holds

\[
QL(S) + 1 \geq \frac{\log T}{l(\log T - \log L(S)) + 2},
\]

where \( \log \) denotes the binary logarithm.

We need the following result, which is proved in paper [25] for the linear complexity.

**Lemma 1.** Let \( a \) be a positive integer, we denote by \( S_a \) the sequence \( (s_{na}) = (s_0, s_a, s_{2a}, \ldots) \). If \( \gcd(a, T) = 1 \) then \( S_a \) has period \( T \), \( L(S) = L(S_a) \), \( QL(S) = QL(S_a) \) and \( T(S) = T(S_a) \).

**Proof.** We prove only \( QL(S) = QL(S_a) \) and the proof of the other properties is done similarly. We write \( L = QL(S) \), there exist coefficients \( a_0, a_1, \ldots, a_{L-1} \in \mathbb{F}_q \) and integers \( 0 < d_1 < \ldots < d_L < T \) satisfying Equation (1), for all \( 0 \leq a \leq n \).

Evaluating Equation (1) in the integers of the form \( na \), we have

\[
s_{an + d_L} = a_{L-1}s_{an + d_{L-1}} + \cdots + a_1s_{an + d_1} + a_0s_{an}, \quad \forall n \geq 0.
\]

We take the positive integer \( 1 \leq r < T \) such that \( ar \equiv 1 \mod T \), then there exist a positive integer \( \lambda_i \) such that \( d_i + \lambda_iT = ard_i \), for \( i = 0, \ldots, L \). Since \( T \) is the period, we obtain \( s_{a(n + rd_i)} = s_{an + ard_i} = s_{an + d_i + \lambda_iT} = s_{an + d_i} \). So, we obtain:

\[
s_{a(n + rd_L)} = a_{L-1}s_{a(n + rd_{L-1})} + \cdots + a_1s_{a(n + rd_1)} + a_0s_{an}.
\]
This implies $QL(S_n) \leq QL(S)$. To conclude the proof, we consider the sequence $S_{ar}$, i.e., $(s_{nar}) = (s_0, s_{ar}, s_{2ar}, \ldots)$. By the same argument used in the above part of the proof, we have $QL(S_{ar}) \leq QL(S_n)$. Clearly, $(s_{nar}) = (s_n)$, because $ar \equiv 1 \mod T$. \hfill \qed

Another trivial remark, which we will use in the proof, is the following connecting the values of the lags and the linear complexity.

**Remark 1.** Let $\Delta$ be a positive integer, $a_0, a_1, \ldots, a_{L-1}, a_L$ non zero elements of $\mathbb{F}_q$ and integers satisfying $(-\Delta) \leq d_1 < d_2 < \cdots < d_L \leq \Delta$ and suppose that the sequence $S$ satisfies the quasi-linear recurrence

$$a_L s_n + d_L = a_{L-1} s_{n+d_{L-1}} + \cdots + a_1 s_{n+d_1} + a_0 s_n, \quad \forall \Delta \leq n,$$

for all $\Delta \leq n$, then it is trivial that $L(S) \leq 2\Delta$.

Lattice theory will play an important role in the proof of Theorem 2, specially the well known Minkowski theorem.

Let $\{\vec{b}_1, \ldots, \vec{b}_s\} = B$ be a set of linearly independent vectors in $\mathbb{R}^r$. The set

$$\Lambda = \{\vec{z} : \vec{z} = c_1 \vec{b}_1 + \cdots + c_s \vec{b}_s, \quad c_1, \ldots, c_s \in \mathbb{Z}\}$$

is called an $s$-dimensional lattice with basis $\{\vec{b}_1, \ldots, \vec{b}_s\}$.

To each lattice $\Lambda$ one can naturally associate its volume

$$\text{vol} (\Lambda) = \left( \det \left( (\vec{b}_i, \vec{b}_j)_{i,j=1}^s \right) \right)^{1/2},$$

which does not depend on the choice of the basis $\{\vec{b}_1, \ldots, \vec{b}_s\}$.

For a vector $\vec{u}$, let $\|\vec{u}\|$ denote its infinity norm. The famous Minkowski theorem, see Theorem 5.3.6 in Section 5.3 of [8], gives the upper bound

$$\min \left\{ \|\vec{z}\| : \vec{z} \in \Lambda \setminus \{\vec{0}\} \right\} \leq \text{vol} (\Lambda)^{1/s}$$

on the shortest nonzero vector in any $s$-dimensional lattice $\Lambda$ in terms of its volume. Now, we have all ingredients to proof the Theorem 2.

**Proof of Theorem 2.** We write $M = QL(S)$, then there exits an equation of the form (1). By definition of $QL(S)$, there exist coefficients $b_0, b_1, \ldots, b_{M-1} \in \mathbb{F}_q$ and integers $0 < d_1 < \cdots < d_M < T$ such that

$$s_{n+d_M} = b_{M-1} s_{n+d_{M-1}} + \cdots + b_1 s_{n+d_1} + b_0 s_n, \quad \forall n \geq 0.$$  \hfill (4)

We introduce the following notation,

$$d_i = \overline{d_i} P^{i-1} + r_i, \quad 0 \leq r_i < P^{i-1}, \quad 0 \leq \overline{d_i} < P, \quad i = 1, \ldots, M,$$  \hfill (5)

and consider the lattice $\Lambda$ generated by the columns of the following matrix $B \in \mathbb{Z}^{M+1\times M+1}$

$$
\begin{pmatrix}
\overline{d_M} & 0 & \cdots & 0 & P \\
0 & \cdots & P & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\overline{d_1} & P & \cdots & 0 & 0 \\
1 & 0 & \cdots & 0 & 0
\end{pmatrix}
$$

5
Clearly, the volume of the lattice is \( \text{vol}(\Lambda) = P^M \), so, by Minkowski theorem (3), if \( \vec{v} \) is the shortest vector in the lattice with infinity norm, then

\[
\|\vec{v}\| \leq P^{1-\frac{1}{M+1}} \tag{6}
\]

Since \( \vec{v} \in \Lambda \) there exist integers \( \lambda_i, i = 1, \ldots, M \) and a positive integer \( R \) such that

\[
\vec{v} = (R\vec{d} + \lambda_M P, \ldots, R\vec{d} + \lambda_1 P, R).
\]

Notice that \( \gcd(R, P) = \gcd(R, T) = 1 \), because \( R < P \) and \( P \) is a prime. Now, we consider \( a \) satisfying \( aR \equiv 1 \mod T \) and the sequence \( S_a = (s_0, s_a, s_2, \ldots) = (s'_0, s'_1, s'_2, \ldots) \). From Equation (4), \( S_a \) satisfies the following quasi-linear recurrence

\[
s'_{n+Rd_M} = b_{M-1} s'_{n+Rd_{M-1}} + \cdots + b_1 s_{n+Rd_1} + b_0 s'_n, \quad \forall n \geq 0. \tag{7}
\]

On the other hand, since \( T \) is the period of the sequence \( S_a \) and recalling Equation (5), we have,

\[
s'_{n+Rd_1} = s'_{n+R\vec{d}P^{-1} + Rr} = s'_{n+R\vec{d}P^{t-1} + R\lambda_1 P t} = s'_{n+(R\vec{d} + \lambda_1 P)P^{t-1} + Rr_i},
\]

for all \( i = 1, \ldots, M \). Then from (7), we get

\[
s'_{n+Rd_M + \lambda_M T} = b_{M-1} s'_{n+Rd_{M-1}} + \cdots + b_1 s_{n+Rd_1} + b_0 s'_n, \quad \forall n \geq 0.
\]

By Bound (6), we have for \( i = 1, \ldots, M \),

\[
|Rd_i + \lambda_i T| = |(R\vec{d}_i + \lambda_i P)P^{t-1} + Rr_i| \leq 2P^{t-1}P^{1-\frac{1}{M+1}} = 2T^{1-\frac{1}{M+1}}.
\]

By Remark 1, \( L(S_a) \leq 4T^{1-\frac{1}{M+1}} \). Now, using Theorem 1

\[
L(S) = L(S_a) \leq 4T^{1-\frac{1}{M+1}},
\]

and operating we obtain the result.

This result generalize the one presented in [1]. Indeed, in the particular case that \( T \) is a prime number, we can give this improved result:

**Corollary 1.** If \( T \) is a prime number, then the following inequality

\[
QL(S) \geq \frac{\log T}{\log T - \log L(S) + 1},
\]

holds, where \( \log \) denotes the binary logarithm.
4 Applications of our results

This bound is very general and applicable to several pseudorandom number generators. Apart from the fact that the period must be a power of a prime number, the other condition to obtain a nontrivial result for the lattice test or the quasi-linear complexity is that the linear complexity of the pseudorandom number generator is known to be large.

There are several sequences which have large linear complexity, nearly as big as the period, see surveys [26, 27]. For example, if $S$ is the inverse recursive generator and has prime period, then $S$ has linear complexity greater than $(T - 1)/2$. This implies that,

$$QL(S) \geq \frac{\log T}{3}.$$  

Similar bounds can be used to find similar bounds for the Legendre sequence and the Sidelnikov sequence. Here, we want to comment another application of our result where the best results known are of similar strength.

The authors in [23] studied lattice test for digital explicit inversive generators and they obtained bounds, even in parts of the sequence. We cite their result only in the case of full period.

**Theorem 3.** Let $S$ be a sequence arising from a digital explicit inverse generator defined over $\mathbb{F}_q$ with $q = p^t$, then we have that,

$$T(S) \geq \frac{\log T - \log \log T - 1}{t - 1} - 1,$$

if $t > 1$. For $t = 1$ the inequality

$$T(S) \geq \frac{T}{2} - 1,$$

holds.

To apply Theorem 2, we need the following bound from [28]. We cite it restricted to the special case of a sequence arising from a digital explicit inverse generator.

**Lemma 2.** Let $S$ be a sequence arising from a digital explicit inversive generator defined over $\mathbb{F}_q$. Then we have

$$L(S) \geq \frac{q(p - 1)}{p} \geq \frac{q}{2}.$$  

This result and a direct application of Theorem 2 gives

$$QL(S) \geq \frac{\log T}{3t} - 1.$$  

Using that $QL(S) \leq T(S)$, we obtain a lower bound which is of the same order as the result obtained in Theorem 3.
Although our bound seems to be weak, it is also quite general. Indeed, for sequences defined by Fermat quotients, we know the exact value of the quasi-linear complexity, which is two, and our bound only gives that the quasi-linear complexity is greater than one.

5 An open problem

We think that Theorem 2 can be formulated for sequences of period $T$ under some restrictions, but not necessarily power of a prime number. However, Example 1 shows that it is not true for arbitrary $T$, but software computations show that our bounds hold in many cases and we think that, under some mild restrictions, it should be possible to prove a lower bound in the quasi-linear complexity depending only on the linear complexity and the period. Also, we would like to know a framework to study the real value of the quasi-linear complexity like in the linear complexity case.

Acknowledgements

The authors want to thank Arne Winterhof for very valuable discussions during our visit to Banff International Research Station and for the preparation of the paper. We also want to thank the anonymous referee for the comments which improved the paper. This work is supported in part by the Spanish Ministry of Science, project MTM2011-24678.

References


