Science and Technology, 7 (2002) 147-156
© 2002 Sultan Qaboos University

Simulation of m-Sequence's Properties Through MATLAB-SIMULINK

A. Ahmad, M. J. Al-Mushrafi and S. Al-Busaidi

Department of Information Engineering, College of Engineering, Sultan Qaboos University, P.O.Box 33, Al Khod 123, Muscat, Sultanate of Oman.
Email: afaq@squ.edu.om, mufeed@omantel.net.om, albusaid@squ.edu.om.

نمذجة خصائص متتاليات – م بواسطة ماتلاب – سيميولنك

آفاق احمد ، مفيد المشرفي و سمير البوسعيدي

خلاصـة: اعـتمادا علـى دراسة تحليلية لنظرية متتاليات –م تم تطوير وسيلة لاستخدامها من قبل المهندسين لتوليد متتالية-م لتطبيقها في مجال ما ، او لفحص متتالية-م أو لكلا الغرضين.

ان مجمـوع البرامج المطورة تمثل كل خواص متتاليات-م من عدد الوحدات والأصفار ، طول المتتالية ، زمن دورة النبض ن وكذلك خواص الارتباط الذاتي لمتتالية. ان هذه المجوعة المبسطة للبرامج طورت باستخدام ماتلاب- سيميولنك للرموز المطلوبة وكذلك القوالب المطلوبة. ان النموذج المطور كفؤ من ناحية الكلفة ، كما ان الملفات الخارجة من نماذج سيميولنك فان استخدامها من قبل أي برنامج أو طريقة نمذجة.

ABSTRACT: Based on an analytical study of the theory of m-sequences, a tool is developed to facilitate practicing engineers to either generate m-sequence for its application, or to test applied m-sequence, or both. The developed test–kit simulates all the properties of m-sequences including number of 1s and 0s, the run lengths, pulse periods as well as autocorrelation properties. This simplified tool is developed using MATLAB–SIMULINK with required codes as well as desired blocks. The model is cost-effective and the output files of the generated SIMULINK models can be utilized in any software program or simulation procedures.

KEYWORDS: LFSR, M- Sequence, PN Codes, Stream Cipher, Run Length

1. Introduction

The m-sequences are of great importance in many fields of engineering and sciences. Cryptography is the most prominent of these applications. One important way of generating such sequences is via Linear Feedback Shift Registers (LFSRs).

In cryptography, m-sequences are used in two ways; one for designing simple forms of encrypting systems and, the other for selecting cryptographic keys (Pless, 1977; Konheim, 1981; Meyer and Matyas, 1982; Barker, 1984; Siegenthaler, 1985; Rueppel, 1986 and 1991; Davis *et. al*, 1994; Schneier, 1996; Diffie and Hellman,1996; Golic, 1998 and 2000; and Ahmad *et. al,* 2001). Apart from the use of m-sequences as stream ciphers, in crypto-security, they have also found a wide range of applications including error control, coding and spread spectrum communications (Shannon, 1963; Neumann, 1963; Massey, 1969; Newbridge Microsystems, 1992; Glaise, 1997). Table1 below shows a few of such practical applications of m-sequences. The table also depicts the used lengths of the m-sequences and the sizes of the corresponding LFSRs which generate them. Besides these, there are numerous other applications of m-sequences as briefly summarized below:

1. m-sequences have shown more effective probing signal than traditional Gaussian white noise for studying nonlinear biological systems (Chen *et. al*, 1996).
2. For computer and video games, the m-sequence generator is a central component (Meyer, and Matyas, 1982; Newbridge Microsystems, 1992 ; and Moore, 1994).

Table 1: Specifications of systems using m-sequences

System Application	Model	Length of m-sequence	Size of LFSR
Cyclic Redundancy Check	CRC-12	4095	12
	CRC-16, CCITT	65535	16*
	AUTODIN-II	4294967295	32
Radio Amateurs (Spread-spectrum)	SS-7	127	7
	SS-13	8191	13
	SS-19, A5 - I	524287	19*
Cellular Telephone (European)	A5 - II	4194303	22
	A5 - III	8388607	23
	A5 - IV	131071	17
ATM Networks	CRC-32	4294967295	32
GPS Satellite	GPSS – I, GPSS - II	1023	10*

* The LFSR sizes are same but each has different structures.

Furthermore, many fields of research (e.g. physics and even finance) are increasingly relying on large computer simulations to study phenomenon that cannot be observed directly for obvious reasons. In these circumstances, the use of m-sequences is an alternate validation methodology to avoid an unforeseen subtle statistical interaction between the m-sequence generator and the properties of simulated phenomenon (Brillinger, 1981; Vattulainen *et al.*, 1994; and Brotherton-Ratcliff, 1995).

As it is evident from the above-mentioned facts that the study of the art and the theory of the generating and applying of m-sequences are becoming essential for engineers and scientists in this era of information technology. Moreover, the testing of m-sequences is rather becoming an updated task of system engineers and scientists rather than acquiring the knowledge of simply using the m-sequences. A lot of research papers are available in the literature which deals with the state of art of generating and testing of m-sequences by adapting different approaches (Moore, 1994; Brillinger, 1981; Vattulainen *et. al.*, 1994; Tausworthe, 1965; Geffe, 1973; Blum *et. al.*, 1986; Micali and Schnorr, 1991; Krawczyk, 1992; L'Ecuyer, 1992; Entacher and Leeb, 1995; Golomb, 1982; Knuth, 1982).

This paper presents a new, simple and systematic procedural study of properties of m-sequences. The state of generating as well as testing of such sequences is also considered in this paper. Further, based on the study of the properties of m-sequences a simplified tool is developed using MATLAB-SIMULINK for generating and testing of such sequences. The model is cost-effective and the output files of the generated SIMULINK models can be utilized in any software program or simulation procedures. The developed test-kit has an attribute of either generating and testing a generated sequence and declaring it as m-sequence or failing to be m-sequence and suggests the change of the LFSR structure. It can also load / read / register the so-called m-sequence and then tests it for its pass / fail of being m-sequence.

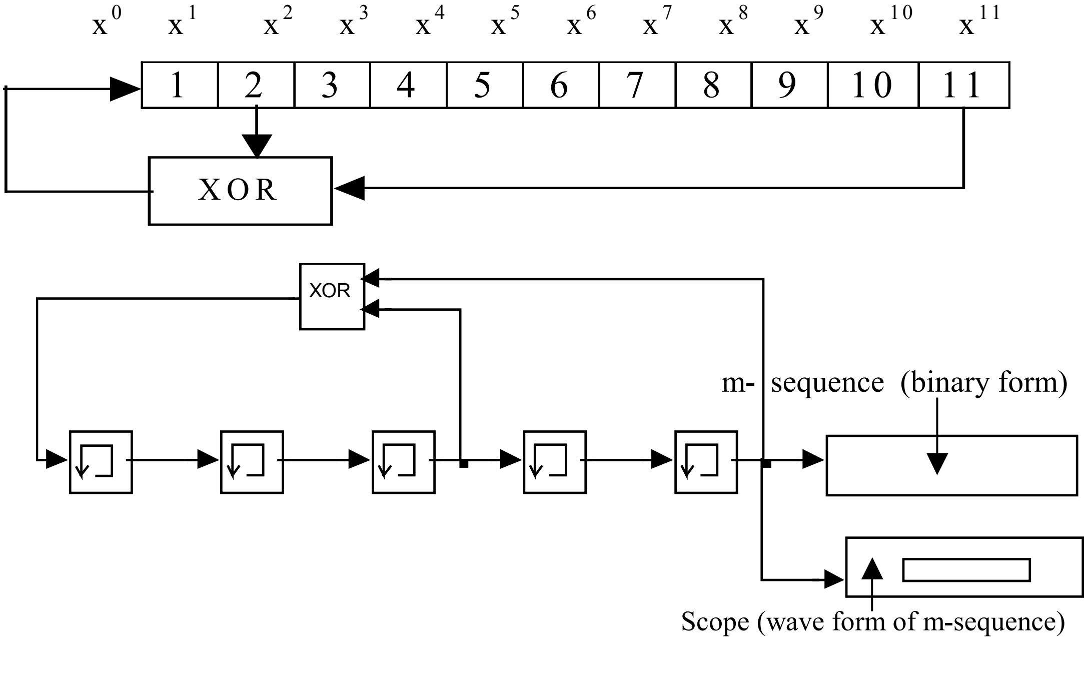
2. Theory of LFSR - The m-Sequence Generator

The idea of randomness reflects the impossibility of predicting the next bit of the sequence. If a generated sequence, $((s_{i\in N}) = s_1,s_2,...s_i.....)$ of symbols, s_i ; $i \in N$ from the finite field GF $(2^n\text{-}1)$, is not matching all the previous ones, it is called a pseudo-random or pseudo-noise (PN) sequence, the other name of the m-sequence. Pseudorandom sequences are the only known sequences that satisfy these properties and are generated by LFSRs, (Golomb, 1982; Knuth, 1982).

The LFSR most often implemented through hardware designs, is the basis of stream ciphers and other applications. In an LFSR a string of bits is stored in a chain of memory cells, where the clock pulses advance the bits towards its next succeeding memory cells. The XOR of certain positions of the cells is used to produce the new bit in the string for each clock pulse with the condition that the last cell position is always used in the XOR process. If each of the memory cells are initially not loaded (initial condition) with 0s, the produced sequence will cycle through a period of more than one. The produced sequence could be cycled through its maximum periodicity of $T = 2^n - 1$, where n is the number of the memory cells used in the LFSR. This maximum periodicity of the sequence can only be achieved through XOR-ing only some combinations of a few particular positions of memory cells of the LFSRs. A diagram shown in Figure 1, illustrates an LFSR consisting of 11 memory cells, where a combination of the memory cells 2nd and 11th is XOR-ed to produce the new bits in the string of the sequence at each clock pulse. The sequence produced by this particular structure of the LFSR of Figure 1 has periodicity of 2047. To make it more readable the following definitions and illustrations are given below:

Definition 1

The structure of an LFSR described by its XOR-ed positions in polynomial form is termed as characteristic polynomial of the LFSR. For example the structure of LFSR of Figure 1 is described by characteristic polynomial $C_{11}(x) = 1 + x^2 + x^{11}$.



x^0 x^1 x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11}
1 2 3 4 5 6 7 8 9 10 11
XOR
XOR
m- sequence (binary form)
Scope (wave form of m-sequence)

Figure 1. An 11-bit LFSR.

Definition 2

The characteristic polynomial of a structure based on an n - memory cells LFSR, which generates a sequence of maximum periodicity ($T = 2^n - 1$), is termed primitive polynomial.

Definition 3

If the produced sequence generated by an LFSR has maximum periodicity ($T = 2^n - 1$), then that sequence is known as m-sequence.

Example 1

Let us consider a 3-bit LFSR, described by a characteristic polynomial, $C_3(x) = 1 + x^2 + x^3$. Have the initial loadings as all 1s. Then the generated sequence by this LFSR will be 0100111, which is of length $2^3 - 1$; thus, the generated sequence is an m-sequence, whereas the characteristic polynomial $C_3(x) = 1 + x^2 + x^3$ is primitive.

3. Properties of m-sequences

Statistical tests on m-sequence can be performed to provide a quantitative measure of randomness. They measure the relative frequencies of certain patterns of 0s and 1s in the sequence s_i (Golomb, 1982; Knuth, 1982). We give below, the systematic study results of the properties of m-sequences.

Property 1

In every period of m-sequence generated by an n-bit LFSR, the total number of 1s will be equal to 2^{n-1}.

Property 2

In every period of m-sequence generated by an n-bit LFSR, the total number of 0s will be one less than number of 1s i.e. number of 0s will be equal to $2^{n-1} - 1$.

Property 3

A period of m-sequence generated by an n-bit LFSR, has an occurrence of n 1s in succession.

Property 4

A period of m-sequence generated by an n-bit LFSR, does not have any occurrences of $(n-1)$ 1s in succession.

Property 5

A period of m-sequence generated by an n-bit LFSR, has an occurrence of $(n-1)$, 0s in succession.

The term run may in general be defined as a succession of items of the same class. In a period of m-sequence the distribution of sequential occurrences of groups of 1s, and 0s (runs property for$1 \le x \le n-2$), is governed by a rule presented in the form a following theorem:

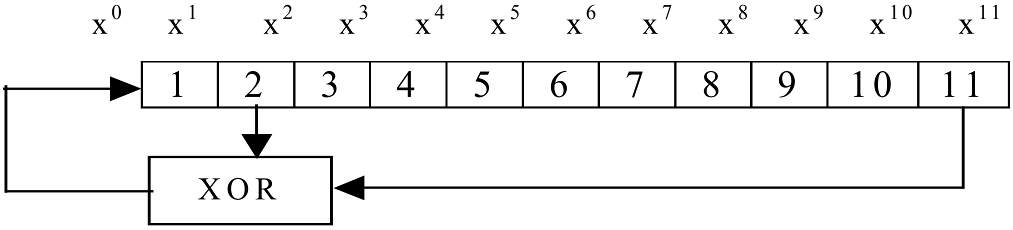
Theorem 1

In a period T of m-sequence generated by an n-bit LFSR, there will be 2^{x-1}runs of ($n - x - 1$) 1s, as well as 0s, for$1 \le x \le n-2$.

Example 2

Figure 2 shows a LFSR simulated using SIMULINK – MATLAB, whose degree of polynomial is n = 5, has characteristic polynomial $(1 + x^3 + x^5)$, initial condition (11111), and generates the m-sequence [s] that has period 31 (i.e. 2^5- 1). The output file of m-sequence in binary form is given as below (Equation 1) whereas the oscilloscope waveform is shown in Figure 3.

$$1111100011011101010000100101100 \qquad (1)$$



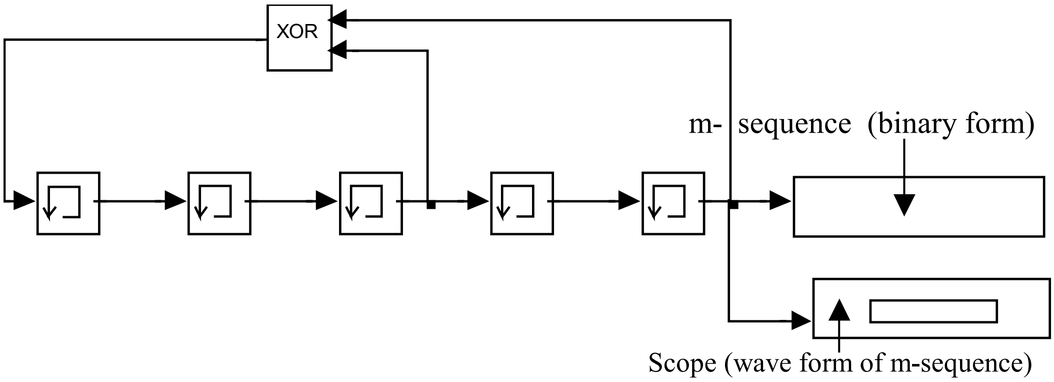


Figure 2. An implementation of a 5-bit LFSR with characteristic polynomial ($1 + x^3 + x^5$) using MATLAB – SIMULINK.

The Properties 1 – 5, and Theorem 1 can be tested for the above sequence of Equation (1). For better explanation, the test result is presented below in the form of the Table 2.

Table 2: Run counts / total number of 1s and 0s.

Frequency of Runs	1	0	0	1	1	1	2	2	4	4	Total of 1s	Total of 0s
Succession of Groups of 1s and 0s	5 1s	5 0s	4 1s	4 0s	3 1s	3 0s	2 1s	2 0s	1 1s	1 0s	16	15

Property 6

The Property of m-Sequences – To Function As Pulse Generator:

It is also interesting to note that the m-sequence generates pulses of different frequencies. The study also, reveals that the pulse width and frequency of different pulses have definite relation with the others (see Figure 3). Table 3, describes this property for an m-sequence of periodicity 2^n-1 with assumption that the clock pulse of LFSR has time period T.

Table 3: Pulses generated in m-sequence.

Number of Pulses	1	1	2^{x-1}; for $1 \le x \le n-2$	2^{x-1}; for $1 \le x \le n-2$
Pulse Width	nT	(n-1)T	(n-x-1)T	(n-x-1)T
Nature of Pulse	Active high	Active low	Active high	Active low

Property 7

The Property of Auto-correlation of *m*-sequences:

To study the statistics and behaviors of m-sequences, it is important to analyze them through their correlation functions. Correlation function of two sequences can be described as the comparison of two sequences to see how much they correspond with one another. Various parameters effect the correlation of two sequences including the length of sequence, phase between the sequences, and clock rate of LFSR. The act of correlating a signal through all variations of itself is known as autocorrelation. The autocorrelation function, AC (k) of an m-sequence $((s_{i\in N}) = s_1, s_2, \ldots s_i \ldots\ldots)$ where, N = 1 to 2^n-1 can be given for its kth shift as:

$$AC(k) = \frac{1}{N}\sum_{i=1}^{N} s_i s_{i+k} \; ; \; 0 \le k \le N-1 \tag{2}$$

Where, s_i is the value of the ith - position of the m-sequence.

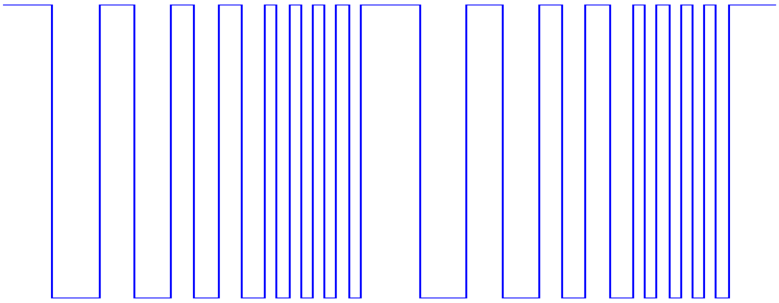


Figure 3 (a). Continuous waveform of generated m-sequence of Figure 2.

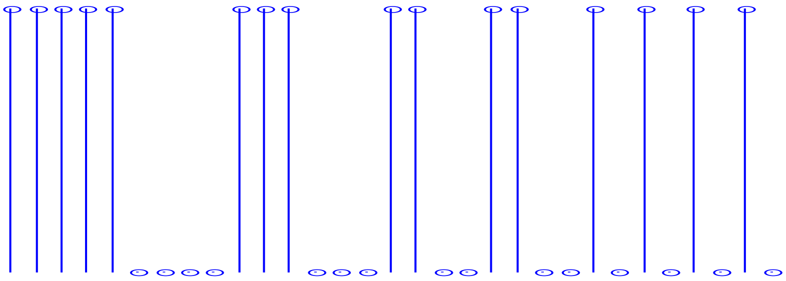


Figure 3 (b). Discrete waveform of generated m-sequence of Figure 2.

Theorem 2

The autocorrelation function of an m-sequence reaches a maximum of 2^n-1 at zero shifts. For other shifts ($1 \le k \le N-2$) its value will be equal to –1.

4. A Test–kit for *m*-Sequence Generator

Using the SIMULINK - MATLAB (Weizheng wang, 1997), a test – kit, shown in Figure 4 is developed to test the Properties 1- 7, and Theorems 1 – 2. The developed kit consists of two separate counters one which monitors the counting of total numbers of 1s and 0s. The second counter monitors the run length properties of the m-sequence. There are two scopes provided in the model of the test–kit. They are dedicated to provide the waveforms of the autocorrelation function and of m-sequence itself. A provision is also made to load the binary form of the m-sequence. All the counters are simulated based on the algorithms to check the desired counts of runs. The developed kit reads the outputs of the second counter as in Table 4.

Table 4: Outputs of second counter

Output number (Function)	Output number (Function)
OUT1 - runs of 1 – 1s;	OUT11 - runs of 1 – 0s;
OUT3 - runs of 2 – 1s;	OUT13 - runs of 2 – 0s;
OUT5 - runs of 3 – 1s;	OUT15 - runs of 3 – 0s;
OUT7 - runs of 4 – 1s;	OUT17 - runs of 4 – 0s;
OUT9 - runs of 5 – 1s;	OUT19 - runs of 5 – 0s;

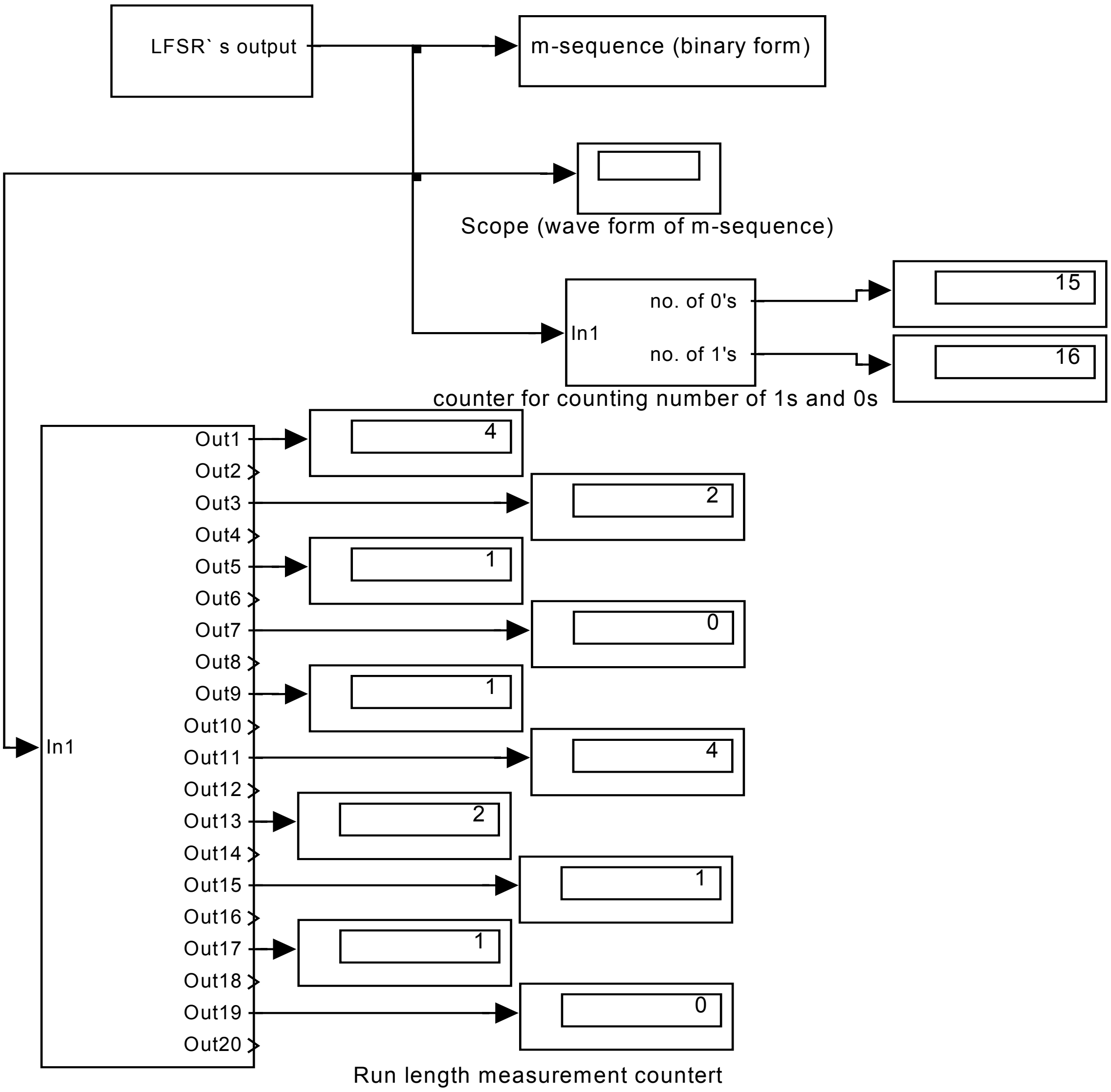


Figure 4. A implementation of test-kit using SIMULINK- MATLAB

5. Results and Discussions

We run the developed MATLAB-SIMULINK model for the generated sequence of Figure 2. The results for the counts of number of 1s and number of 0s as well as run lengths can be observed (according to Table 4); as it is monitored through Figure 4. The observed values are identical as demonstrated in Table 2, which verifies the properties 1-5 and Theorem 1. Figure 3 is the output of the scope, which is nothing but the wave–form of the generated m-sequence. The study of this wave–form reveals that the pulse properties are according to the rules stated in Table 3. The output of another scope is shown in Figure 5. It can be visualized through Figure 5 that the peaks of autocorrelation values (AC (k); Equation 2) are 31 at zero shifts, and for other shifts the values are –1, which verifies the Theorem 2. Further, it can be seen that it repeats in each cycle of the generated m-sequence. The result is satisfying the autocorrelation property of the m-sequence as given in Equation 2.

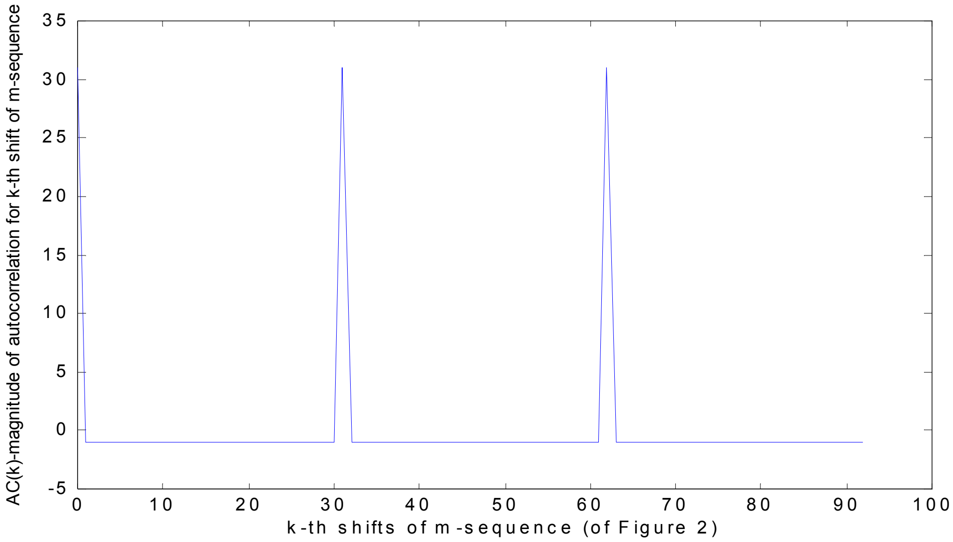


Figure 5. Autocorrelation property of m-sequence of Figure 2.

Thus, based on a systematic and procedural study of the theory of m-sequences we developed a tool to help the practicing engineers to either generate m-sequence for its application or to test applied m-sequence or both. The study presented in general in Section 1 and in particular in Table 1, reveals how important an m-sequence is. We tried to provide the knowledge of generating and testing of m-sequences with the least mathematical involvements to make the paper more suitable for general readers and especially for practicing engineers in the area of computer and communication. Since the security is a vital issue in this age of information technology, and finally, it seems that the security responsibility has to come in any form on the shoulders of practicing engineers of all fields to avoid the litigations.

References

AHMAD, A., AL-MUSHARAFI, M.J., AL-BUSAIDI, S., AL-NAAMANY, A., and JERVASE, J.A., 2001. 'An NLFSR Based Sequence Generator for Stream Ciphers' Proceedings of Seta01 (Sequences and Their Applications - an International Conference, held at Norway in May' 2001), pp. 11-12

BAKER, W.G., '*Cryptanalysis of Shift-Register Generated Stream Cipher Systems*,' Aegean Park Press, 1984.

BLUM, L., BLUM, M., and SHUB, M., 1986. 'A Simple Unpredictable Pseudo-Random Number Generator,' *SIAM Journal on Computing*, **15(2):** .

BLUM, L., BLUM, M., and SHUB, M., 1986. A Simple Unpredictable Pseudo-random Number Generator, SIAM Journal of Computing, **15(2):** 364-383.b

BRILLINGER, DAVID., 1981. '*Time Series: Data Analysis and Theory*, Holden-Day.

BROTHERTON-RATCLIFF, R., 1995. Using Quasi-Random Sequences in Monte-Carlo Valuation of Path–Dependent Options', Risk Magazine, December 1994, and also in *Canadian Treasure*, **11(2):** 36-38.

CHEN, H.W., AINE, C.J.E., BEST, D., RANKEN, HARRISON, R.R., FLYNN, E.R. and WOOD, C.C., 1996. 'Nonlinear Analysis of Biological Systems Using Short m-sequences and Sparse-Simulation Techniques', *Annals of Biomedical Engineering*, **24:** 513-536.

DAVIS, D., IHAKA, R., and FENSTERMACHER, P., 1994. 'Cryptographic Randomness from Air Turbulence in Disk Drives, Advances in Cryptology – Crypto-94, Springer-Verlag Lecture Notes in Computer Science No. 839.

DIFFIE W., and HELLMAN, M.E., 1996. 'New Directions in Cryptography,' IEEE Transactions on Information Technology.

ENTACHER, K., and LEEB, H., 1995. 'Inversive pseudorandom number generators: empirical results. In Proceedings of the Conference Parallel Numerics 95, Sorrento, Italy, September 27-29, 1995, pp 15-27.

GEFFE, P.R, 1973. 'How to protect data with Ciphers that are really hard to break', Electronics.

GLAISE, R.J., 'A Two-Step Computation of Cyclic Redundancy Code CRC-32 For ATM Networks', *IBM Journal*, **41(6)** –Non-Topical Issue, 1997.

GOLIC, J.D., 1998.'Recent advances in stream cipher cryptanalysis,' *Publications de L'Institut Mathematique,* **64/78:** 183-204.

GOLIC, J.D., May 2000. Cryptanalysis of Three Mutually Clock-Controlled Stop/Go Shift Registers', IEEE Transactions on Information Technology, **46(3):**1081-1090.

GOLOMB, S.W. 1982.'*Shift Register Sequences*, Aegean Park Press, Revised Edition.

KNUTH, D.E., 1982. '*The Art of Computer Programming*, Volume 2: Semi numerical Algorithms, Chapter 3: Random Numbers. Addison Wesley Publishing Company, Second Edition.

KONHEIM, A.G., 1981. 'Cryptography: A Primer,' A Wiley-Inter-science Publication, John Wiley & Sons.

KRAWCZYK, H., 1992. 'How to Predict Congruential Generators, *Journal of Algorithms*,' **13(4)** December.

L'ECUYER, P., 1992. Testing Random Number Generators , Proceedings of the 1992 Winter simulation Conference, IEEE press, pp. 305-313.

MASSEY, J.L., 1969. Shift register synthesis and BCH Decoding', *IEEE Transactions On Information Technology*, Vol. IT-15.

MEYER, C.H., and MATYAS, S.M., 'Cryptography: *A New Dimension in Computer Data Security*,' A Wiley-Inter-science Publication, John Wiley & Sons, 1982.

MICALI, S., and SCHNORR, C.P., 1991. Efficient, Perfect Polynomial Random Number Generators,' Journal of Cryptology, **3:** 157-172.

MOORE, L., $600,000 put aside for keno winner, The Gazette, Montreal, April 23, 1994, page A6.

NEUMANN, V., 1963. 'Various techniques used in connection with random digits,' von Neumann's Collected Works, Vol. 5, Pergamon Press.

NEWBRIDGE MICROSYSTEMS, 1992. RBG1210 Random Bit Generator, data sheet published in Newbridge Microsystems' 3rd issue of CMOS Products data book, Newbridge Microsystems, Kanata, Ontario, Canada.

PLESS, V.S., 1977. 'Encrypting Schemes for Computer Confidentiality', IEEE Trans. On Computer, Vol. C-26, No. 11.

RUEPPEL, R.A., 1986. *Analysis and Design of Stream Ciphers*. New York, NY: Springer.

RUEPPEL, R.A., 1991. 'Stream Ciphers in Contemporary Cryptography: The Science of Information Integrity, Editor – G. Simmons, *IEEE Press*, 1991, pp. 65-134.

SCHNEIER, B., 1996. '*Applied Cryptography: Protocols, Algorithms, and Source Code in C*,' John Wiley & Sons.

SIEGENTHALER, T. 1985. 'Decrypting a class of Stream Ciphers Using Ciphertext Only', IEEE Trans. On Computer, Vol. C-34, No.'1.

SHANNON, C.E., 1963. '*The Mathematical Theory of Communication*,' University of Illinois Press. (Originally from: Bell System Technical Journal, July and October 1948).

TAUSWORTHE, R.C., 1965.Random numbers generated by linear recurrence modulo two, Mathematics of Computation, **19:** 201-209.

VATTULAINEN, I., ALA-NISSILA, T., and KANKAALA, K., 1994. Physical tests for Random Numbers in Simulations, Physical Review Letters, Vol. 73, Number 19, 7, pp. 2513-2516.

WEIZHENG WANG, 1997. 'SIMULINK (User's Guide – Version 2),' the MathWorks, Inc.

Received 27 June 2001
Accepted 22 September 2001