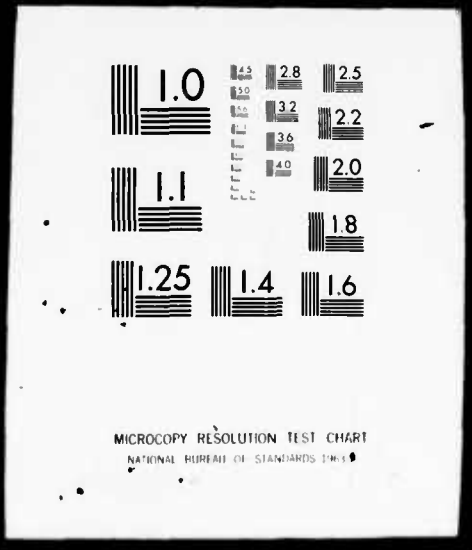
AD-A045 652 STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE F/G 9/2
ANALYSIS OF ADDITIVE RANDOM NUMBER GENERATORS.(U)
MAR 77 J F REISER N00014-76-C-0330
UNCLASSIFIED STAN-CS-77-601 NL

1 OF 1
AD A045652

END
DATE FILMED
11-77
DDC



(11)

(9) Technical rept.,

(6) ANALYSIS OF ADDITIVE RANDOM NUMBER GENERATORS.

by

(10) John E. Reiser
Fredrick

(14) STAN-CS-77-601
(11) MARCH 1977

(12) 40 p.

(15) N00014-76-C-0330

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

LELAND STANFORD JUNIOR UNIVERSITY
ORGANIZED 1891

094 120

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: STAN-CS-77-601	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE *(and Subtitle)*: ANALYSIS OF ADDITIVE RANDOM NUMBER GENERATORS	5. TYPE OF REPORT & PERIOD COVERED: technical, March 1977
6. PERFORMING ORG. REPORT NUMBER: STAN-CS-77-601	
7. AUTHOR(s): John F. Reiser	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0330
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Stanford University, Computer Science Department, Stanford, Ca. 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Department of the Navy, Arlington, Va. 22217	12. REPORT DATE: March 1977
	13. NUMBER OF PAGES: 34
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: ONR Representative: Philip Surra, Durand Aeronautics Bldg., Rm. 165, Stanford University, Stanford, Ca. 94305	15. SECURITY CLASS. *(of this report)*: unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

releasable without limitations on dissemination

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

analysis of algorithms, numerical analysis

20. ABSTRACT

This thesis presents an analysis of the distribution of residues generated by the k^{th}-order linear homogeneous recurrence $Y_{(n+k)} = a_{(k-1)}Y_{(n+k-1)} + \ldots + a_{(0)}Y_{(n)} \bmod p$ when $x^k - a_{k-1}x^{k-1} - \ldots - a_{(0)}$ is a primitive polynomial in $Z_{(p)}(x)$. It is shown that for $t \leq k$ the tuples of t consecutive residues are equidistributed in t dimensions in the limit as $\alpha \to \infty$, subject only to a much weaker condition on the distribution of the residues. When specialized to $|a_j| \leq 1$, the recurrence is the basis for a computer random number generator which can be efficiently implemented directly in floating-point arithmetic with no multiplication and little machine dependence. The results of empirical tests comparing generators of this type with standard linear congruential generators are also presented.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

ANALYSIS OF ADDITIVE RANDOM NUMBER GENERATORS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
John Fredrick Reiser
February 1977

The printing of this paper was supported in part by NSF grants MCS 72-03752 A03, by the Office of Naval Research contract N00014-76-C-0330, and by IBM Corporation. Reproduction in whole or in part is permitted for any purpose of the United States government.

Preface

This thesis presents an analysis of the distribution of residues generated by the k^{th}-order linear homogeneous recurrence $y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0 y_n \bmod p^\alpha$ when $x^k - a_{k-1}x^{k-1} - \cdots - a_0$ is a primitive polynomial in $Z_p[x]$. It is shown that for $t \le k$ the tuples of t consecutive residues are equidistributed in t dimensions in the limit as $\alpha \to \infty$, subject only to a much weaker condition on the distribution of the residues. When specialized to $|a_j| \le 1$, the recurrence is the basis for a computer random number generator which can be efficiently implemented directly in floating-point arithmetic with no multiplication and little machine dependence. The results of empirical tests comparing generators of this type with standard linear congruential generators are also presented.

I wish to thank Andrew Yao and Gene Golub for serving on my reading committee. Special thanks go to Donald Knuth for his advice and encouragement. I also thank Robert Maas for developing the typographical software, and my friends, colleagues, and relatives for their emotional support. The Graduate Fellowship Program of the National Science Foundation provided financial support.

My deepest appreciation is reserved for my teachers, especially Arnold Ross and Donald Knuth. And I am particularly fortunate to count among my teachers two of the best: my parents, Bil and Marleen Reiser.

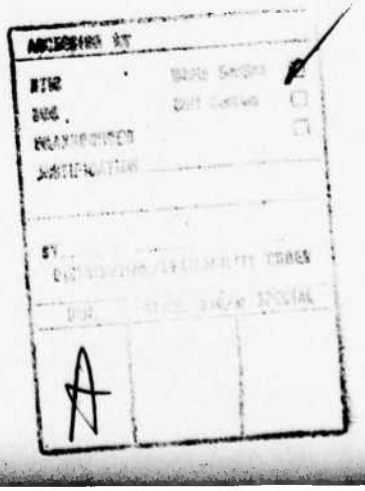


Table of Contents

Chapter 1. Introduction 1

Chapter 2. Discrepancy and Exponential Sums 8

Chapter 3. Reduction to a Weaker Criterion 13

Chapter 4. Analysis of the Number of Zeroes in a Cycle 19

Chapter 5. Practical Considerations 27

Bibliography 33

List of Tables

Table 1. Octal values of F_{12j+t} mod 64 14

Table 2. Octal values of F'_{12j+t} mod 64 15

Table 3. Results of tests on RANDU 31

Table 4. Results of tests on GOODLC 31

Table 5. Results of tests on ADDLC 32

Table 6. Results of tests on BESTX 32

CHAPTER 1

Introduction

Let α and k be positive integers and let p be a prime. Let $f(x) = x^k - a_{k-1}x^{k-1} - \cdots - a_0$ be a primitive polynomial in $Z_p[x]$. In other words, the residue classes modulo $f(x)$ of the polynomial ring $Z_p[x]$ form a finite field in which x is a generator of the multiplicative group. Consider the k^{th}-order linear homogeneous recurrence

$$y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0 y_n \bmod p^\alpha \tag{1.1}$$

for $n = 0, 1, 2, \ldots$ and initial values $(y_0, \ldots, y_{k-1}) \neq (0, \ldots, 0) \bmod p$. The sequence of fractions $\langle y_n/p^\alpha \rangle$ is a candidate for a pseudo-random sequence. If $\alpha = 1$ then the length of the period of $\langle y_n \rangle$ is p^k-1, and if $\alpha > 1$ the length of the period is $p^{\alpha-c}(p^k-1)$, where c is easy to compute and is often equal to 1. Recurrences of type (1.1) with $\alpha = k = 1$ and large p, or with $k = 1$, $p = 2$, and moderately large α, are the basis for some of the most acceptable and widely used computer random number generators [Knuth69]. Such generators have $|a_0| > 1$ and require that a multiplication modulo p^α be performed. Multiplication can be replaced by addition and subtraction if $|a_j| \leq 1$; of course then k must be greater than 1. Experimental evidence has been accumulating to the effect that recurrences of type (1.1) with $p = 2$ and moderately large α and k are quite successful. (See [Knuth69] p. 464 and [Brent73] pp. 163-164; also [Green59] and [Franklin64].) However, theoretical justification for such success has been lacking. We will show that recurrences of type (1.1) are indeed excellent random number generators by showing that for $t \leq k$ the t-tuples of consecutive residues become equidistributed in t dimensions in the limit as $\alpha \to \infty$, subject only to a much weaker condition on the distribution of the residues.

In the remainder of this chapter we shall discuss known results on the length of the period of sequences of type (1.1). Chapter 2 defines discrepancy, a means of measuring

equidistribution, and presents a formula of Harald Niederreiter which expresses the discrepancy in terms of an exponential sum. In Chapter 3, exponential sums are used to reduce the question of equidistribution of sequences of type (1.1) to a much weaker distribution criterion, and in Chapter 4 the sequences are analyzed with respect to this criterion. It is believed that the analysis of Chapters 3 and 4 is new. Chapter 5 considers implementation details and gives the results of empirical tests comparing higher order linear congruential generators of type (1.1) with standard linear congruential generators $y_{n+1} = ay_n + b \bmod p^\alpha$.

The simplest example of a sequence satisfying (1.1) with $k > 1$ is the Fibonacci sequence with $p = 2$ and initial conditions $(y_0, y_1) = (0, 1)$. The recurrence is $F_{n+2} = F_{n+1} + F_n$, corresponding to the primitive polynomial $x^2 - x - 1 = x^2 + x + 1$ in $Z_2[x]$. The period is $3 \cdot 2^{\alpha-1}$ and it is known that the sequence of ordered pairs $(2^{-\alpha}F_n, 2^{-\alpha}F_{n+1})$ becomes evenly distributed mod 1 as α increases. (See [Marsaglia72].) However, the Fibonacci sequence is not a suitable random number generator because successive triples are very poorly distributed in three dimensions. To achieve satisfactory performance we must consider recurrences of higher degree.

When considering such recurrences it is helpful to know some facts about the length of the period and some relationships between sequences satisfying the same linear congruence, but with different initial conditions. The papers [Ward31], [Ward33], and [Hall38a] present accounts of the theory for general linear recurrences. The length of the period of recurrence (1.1) for $\alpha = 1$ can be easily established using an idea from [Hall38a].

Lemma 1.1. If $f(x)$ is primitive in $Z_p[x]$ then the period of (1.1) for $\alpha = 1$ is $p^k - 1$.

Proof. Corresponding to the k-tuple $(y_0, \ldots, y_{k-1})$, associate the polynomial

$$Y(x) = y_0x^{k-1} + (y_1 - a_{k-1}y_0)x^{k-2} + \cdots + (y_{k-1} - a_{k-1}y_{k-2} - \cdots - a_1y_0) . \qquad (1.2)$$

Then $xY(x) = y_0x^k + (y_1 - a_{k-1}y_0)x^{k-1} + \cdots + (y_{k-1} - \cdots - a_1y_0)x = y_1x^{k-1} + (y_2 - a_{k-1}y_1)x^{k-2} + \cdots + (y_k - \cdots - a_1y_1)$ (modulo $f(x)$), and this is the polynomial associated with $(y_1, \ldots, y_k)$. Thus iterating the recurrence corresponds to multiplying the associated polynomial by x. Since $f(x)$ is primitive, x is a generator of the multiplicative group of the finite field $Z_p[x]/(f(x))$ and has period $p^k - 1$. Thus recurrence (1.1) has period $p^k - 1$. ∎

We will use generating functions and congruences to a double modulus to analyze the period and certain other properties of integer sequences satisfying (1.1) when $\alpha > 1$. The notion of congruence to a double modulus is an extension of the usual notion of congruence. If $f(x)$, $a(x)$, and $b(x)$ are polynomials with integer coefficients, and if there exist polynomials $u(x)$ and $v(x)$ with integer coefficients such that $a(x) = b(x) + f(x)u(x) + mv(x)$, then we will write $a(x) \equiv b(x)$ (modulo $f(x)$ and m), or also $a(x) \equiv b(x)$ (modd $m, f(x)$). The following four lemmas are essentially exercise 3.2.2-11 of [Knuth69].

Lemma 1.2. Assume that $f(0)$ is relatively prime to p and that $p^\alpha > 2$. If

$$x^T \equiv 1 \text{ (modd } p^\alpha, f(x)) \quad \text{and} \quad x^T \not\equiv 1 \text{ (modd } p^{\alpha+1}, f(x)),$$

then

$$x^{pT} \equiv 1 \text{ (modd } p^{\alpha+1}, f(x)) \quad \text{and} \quad x^{pT} \not\equiv 1 \text{ (modd } p^{\alpha+2}, f(x)).$$

Proof. By definition of congruence to a double modulus there exist polynomials $u(x)$, $v(x)$ such that $x^T = 1 + f(x)u(x) + p^\alpha v(x)$. In addition $v(x) \not\equiv 0$ (modd p, $f(x)$), or else $x^T = 1 + f(x)u(x) + p^\alpha(f(x)u_2(x) + pv_2(x)) = 1 + f(x)(u(x) + p^\alpha u_2(x)) + p^{\alpha+1}v_2(x) \equiv 1$ (modd $p^{\alpha+1}$, $f(x)$), contradicting the assumption about x^T (modd $p^{\alpha+1}$, $f(x)$). If we raise x^T to the p^{th} power, the binomial theorem makes it clear that $x^{pT} = 1 + p^{\alpha+1}v(x) + p^{2\alpha+1}v^2(x)(p-1)/2$ plus other terms which are congruent to zero modulo $f(x)$ and $p^{\alpha+2}$. Since $p^\alpha > 2$ we have $p^{2\alpha+1}(p-1)/2 \equiv 0$ modulo $p^{\alpha+2}$, and thus $x^{pT} \equiv 1 + p^{\alpha+1}v(x)$ (modd $p^{\alpha+2}$, $f(x)$). Suppose that $p^{\alpha+1}v(x) \equiv 0$ (modd $p^{\alpha+2}$, $f(x)$).

Then there exist polynomials $a(x)$, $b(x)$ such that $p^{\alpha+1}v(x) = a(x)f(x) + p^{\alpha+2}b(x)$, which implies that $p^{\alpha+1}(v(x) - pb(x)) = a(x)f(x)$. Since $f(0)$ is relatively prime to p we can apply Gauss's lemma about the g.c.d. of the coefficients of the product $a(x)f(x)$ and deduce that $p^{\alpha+1}$ divides $a(x)$. This means that $v(x) = p^{-(\alpha+1)}a(x)f(x) + pb(x) \equiv 0$ (modd p, $f(x)$), which is a contradiction. Thus $p^{\alpha+1}v(x) \not\equiv 0$ (modd $p^{\alpha+2}$, $f(x)$) and therefore $x^{pT} \not\equiv 1$ (modd $p^{\alpha+2}$, $f(x)$). ∎

Using generating functions we will next derive a relationship between period lengths and the powers τ for which $x^{\tau} \equiv 1$ (modd m, $f(x)$). Let $f(x) = 1 - a_1x - \cdots - a_kx^k$, and let $G(x) = 1/f(x) = A_0 + A_1x + A_2x + \cdots$. Denote by $\tau(m)$ the length of the period of $\langle A_n \bmod m \rangle$.

Lemma 1.3. $\tau(m)$ is the least positive integer τ such that $x^{\tau} \equiv 1$ (modd m, $f(x)$).

Proof. Since $\tau(m)$ is the period of $\langle A_n \bmod m \rangle$ we have $G(x) - x^{\tau(m)}G(x) \equiv A_0 + A_1x + A_2x^2 + \cdots A_{\tau(m)-1}x^{\tau(m)-1}$ (modulo m), which implies $1 - x^{\tau(m)} \equiv f(x)(A_0 + A_1x + \cdots + A_{\tau(m)-1}x^{\tau(m)-1})$ (modulo m), so $1 - x^{\tau(m)} \equiv 0$ (modd m, $f(x)$). This shows that $\tau(m) \geq \tau$, since τ was the smallest positive integer such that $1 - x^{\tau} \equiv 0$ (modd m, $f(x)$). Conversely, taking $x^{\tau} - 1 \equiv 0$ (modd m, $f(x)$) and multiplying by $G(x) = 1/f(x)$ gives $x^{\tau}G(x) - G(x) \equiv 0$ (modulo m). Equating coefficients of x gives $A_n - A_{n+\tau} \equiv 0$ (modulo m) for all $n \geq 0$. Thus $\tau(m) \leq \tau$. ∎

We now restrict our attention to prime power moduli $m = p^{\alpha}$ and show that when α is large enough, increasing α by 1 multiplies not only the modulus but also the period by p.

Lemma 1.4. If $p^{\alpha} > 2$ and $\tau(p^{\alpha}) \neq \tau(p^{\alpha+1})$ then $\tau(p^{\alpha+k}) = p^k\tau(p^{\alpha})$.

Proof. It suffices to show that $\tau(p^{\alpha}) \neq \tau(p^{\alpha+1})$ implies that $\tau(p^{\alpha+1}) = p\tau(p^{\alpha}) \neq \tau(p^{\alpha+2})$. From what we have already shown about the period, we know that $\tau(p^{\alpha+2}) \neq p\tau(p^{\alpha})$ and

that $\tau(p^{\alpha+1})$ divides $p\tau(p^{\alpha})$ but does not divide $\tau(p^{\alpha})$. Let $\tau(p^{\alpha}) = p^{e}q$ with p and q relatively prime. Then since $\tau(p^{\alpha+1})$ divides $p^{e+1}q$ but does not divide $p^{e}q$, it must be that $\tau(p^{\alpha+1}) = p^{e+1}d$ where d divides q. Since $p^{e+1}d$ is a period modulo $p^{\alpha+1}$, it is certainly also a period modulo p^{α}. The smallest period divides all other periods, so $p^{e}q$ divides $p^{e+1}d$. Thus q divides d, $q = d$ and $\tau(p^{\alpha+1}) = p\tau(p^{\alpha})$. ∎

We have determined the period of $\langle A_n \bmod p^{\alpha}\rangle$ where A_n is the coefficient of x^n in the generating function $G(x) = 1/f(x)$ and $f(x) = 1 - a_1x - \cdots - a_kx^k$. This sequence $\langle A_n \bmod p^{\alpha}\rangle$ satisfies (1.1), and we would like to know something about the set of all sequences satisfying (1.1). If $G_1(x) = y_0 + y_1x + \cdots$ where $y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0y_n$ for all $n \geq 0$, it is easy to see that $f(x)G(x)$ is the polynomial $g(x) = y_0 + (y_1 - a_1y_0)x + \cdots + (y_{k-1} - a_1y_{k-2} - a_2y_{k-3} - \cdots - a_{k-1}y_0)x^{k-1}$, hence $g(x)/f(x)$ is the generating function for the sequence with initial values $(y_0, y_1, \ldots, y_{k-1})$. The next lemma shows that the period of $\langle y_n \bmod m\rangle$ is the same as the period of $\langle A_n \bmod m\rangle$, in the cases of interest to us.

Lemma 1.5. Let $m = p^{\alpha}$. If $f(x)$ and $g(x)$ are relatively prime modulo p then the period of $\langle y_n \bmod m\rangle$ equals the period of $\langle A_n \bmod m\rangle$. In particular, this holds when $f(x)$ is irreducible and $(y_0, \ldots, y_{k-1}) \not\equiv (0, \ldots, 0)$ (modulo p).

Proof. Assume that τ is the period of $\langle y_n \bmod m\rangle$. Then $g(x)(1 - x^{\tau}) \equiv 0$ (modd m, $f(x)$). Because $f(x)$ and $g(x)$ are relatively prime modulo p, we can apply Hensel's lemma and find polynomials $a(x)$ and $b(x)$ such that $a(x)f(x) + b(x)g(x) \equiv 1 \bmod m$. Multiplying $g(x)(1 - x^{\tau}) \equiv 0$ (modd m, $f(x)$) by $b(x)$ gives $1 - x^{\tau} \equiv 0$ (modd m, $f(x)$), hence the period of $\langle y_n \bmod m\rangle$ is no shorter than the period of $\langle A_n \bmod m\rangle$. On the other hand it cannot be longer, since $\langle y_n \bmod m\rangle$ is a linear combination of sequences $\langle A_{n+j} \bmod m\rangle$ for various j. If $f(x)$ is irreducible in $Z_p[x]$ then $f(x)$ is relatively prime to every nonzero polynomial of lower degree, so the period of $\langle y_n \bmod m\rangle$ will be the same as the period of $\langle A_n \bmod m\rangle$ unless $g(x) \equiv 0 \bmod p$. This can only happen if all of the initial values $(y_0, \ldots y_{k-1})$ are

divisible by p (again because of Gauss's lemma). ∎

The results which have been derived so far about the period may also be derived by formulating the recurrence relation (1.1) in terms of a matrix-vector product. This alternative formulation is important because actual computation may be easier with the matrix than with the polynomials. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{k-1} \end{bmatrix} \tag{1.3}$$

where the a_j are the same coefficients as those in (1.1). Define $\mathfrak{y}_n^T = (y_n, y_{n+1}, \cdots, y_{n+k-1})$ to be the k-tuple of consecutive terms of (1.1) beginning with the n^{th} term. It is easy to see that $\mathfrak{y}_1 \equiv A\mathfrak{y}_0$ and hence by induction that $\mathfrak{y}_n \equiv A^n\mathfrak{y}_0$ mod p^α. By definition of the period $\tau = \tau(p^\alpha)$ we have $\mathfrak{y}_n \equiv A^\tau\mathfrak{y}_0$ mod p^α, so

$$A^\tau = I + p^\alpha B \tag{1.4}$$

for some matrix B. If p does not divide B then τ is not the period mod $p^{\alpha+1}$. However, $A^{p\tau} = (I + p^\alpha B)^p = I + p^{\alpha+1}B + p\cdot\frac{p-1}{2}\cdot p^{2\alpha}B^2$ plus other terms divisible by $p^{\alpha+2}$. If $p^\alpha > 2$ then we see that $A^{p\tau} \equiv I + p^{\alpha+1}B$ (modulo $p^{\alpha+2}$), and we have rederived the fact that $p\tau$ is the period modulo $p^{\alpha+1}$ but not modulo $p^{\alpha+2}$. We can also note that $\mathfrak{y}_{n+\tau/p} - \mathfrak{y}_n \equiv A^{\tau/p}\mathfrak{y}_n - \mathfrak{y}_n \equiv p^{\alpha-1}B\mathfrak{y}_n$ (modulo p^α). This says that as the modulus increases the difference between vectors a period apart is merely multiplied by p. In the binary case $p = 2$ this means that the exclusive-or of $\mathfrak{y}_{n+\tau/2}$ and $\mathfrak{y}_n$ gives the same pattern of bits, just shifted one place left as the modulus increases.

If we consider $A^{j\tau(\alpha)}$ for $j = p^\alpha$ we see that

$$A^{p^\alpha\tau(\alpha)} = (I + p^\alpha B)^{p^\alpha} = I + p^\alpha p^\alpha B + p^\alpha\cdot\frac{p^\alpha-1}{2}p^{2\alpha}B^2 + \cdots \equiv I + p^{2\alpha}B \bmod p^{3\alpha-1} \tag{1.5}$$

When Eq. (1.4) is viewed as defining B as a function of α then Eq. (1.5) says that $B(\alpha) \equiv B(2\alpha)$ (modulo $p^{\alpha-1}$). In other words, as α increases B converges in the p-adic sense.

CHAPTER 2

Discrepancy and Exponential Sums

In order to make meaningful statements about how well a sequence is distributed it is necessary to have a measure of equidistribution. Discrepancy is such a measure. Let I be a unit interval, and for intervals $J \subset I$ let $A(J,N)$ be the number of n, $0 \le n < N$, with $y_n \in J$. Then the discrepancy of the points $y_0, \ldots, y_{N-1}$ is defined as

$$D_N(y_0, \ldots, y_{N-1}) = \sup_{J \subset I} \left| \frac{A(J,N)}{N} - \lambda(J) \right| \tag{2.1}$$

where $\lambda(J)$ is the measure of J. Thus the discrepancy measures the maximum difference between the actual fraction of hits in an interval and the expected fraction of hits. It is easy to see that $0 \le D_N \le 1$ and that the sequences produced by a good random number generator should have small discrepancy. Definition 2.1 can be extended naturally to define discrepancy for sequences of points y_n lying in a multidimensional unit interval I, and we shall often use the extended definition.

Harald Niederreiter has developed an inequality relating the discrepancy to certain exponential sums [Niederreiter78]. This inequality is important because it bounds the discrepancy in terms of functions having nice mathematical properties. The properties will be exploited when analyzing the discrepancy of linear congruential sequences (1.1).

Niederreiter's inequality is easier to state if some notation is introduced. For integers m and h define

$$r(h, m) = \begin{cases} 1 & \text{if } h \equiv 0 \pmod m \\ m \sin \pi \|h/m\| & \text{if } h \not\equiv 0 \pmod m \end{cases}$$

where $\|t\|$ is the distance from t to the nearest integer. For lattice points $\underline{h} = (h_1, \ldots, h_s) \in Z^s$ we write

$$r(\underline{h}, m) = \prod_{j=1}^{s} r(h_j, m) .$$

The summation symbol $\sum_{h \bmod m}$ will designate a sum over the complete system of numerically least residues modulo m, consisting of all integers h with $-m/2 < h \le m/2$. The summation symbol $\sum_{h \bmod m}^{*}$ refers to the same sum but with $h = 0$ deleted from the range of summation. The symbols $\sum_{\underline{h} \bmod m}$ and $\sum_{\underline{h} \bmod m}^{*}$ will refer to analogous sums over the complete system of representatives of $Z^s/(mZ)^s$ consisting of all $\underline{h} = (h_1, \ldots, h_s) \in Z^s$ with $-m/2 < h_j \le m/2$ for $1 \le j \le s$, possibly omitting $\underline{h} = (0, \ldots, 0)$. The notation $\underline{x} \cdot \underline{y}$ represents the standard inner product of two vectors. The function $e(t)$ is defined for real values t as $e(t) = e^{2\pi i t}$.

Lemma 2.1 (Niederreiter's lemma). Let $\underline{y}_0, \ldots, \underline{y}_{N-1}$ be N lattice points in Z^s. Then for any integer $m \ge 2$ the discrepancy D_N of the points $m^{-1}\underline{y}_0, \ldots, m^{-1}\underline{y}_{N-1}$ satisfies

$$D_N \le \frac{s}{m} + \sum_{\underline{h} \bmod m}^{*} \frac{1}{r(\underline{h}, m)} \left| \frac{1}{N} \sum_{n=0}^{N-1} e(\underline{h} \cdot \underline{y}_n / m) \right| .$$

Proof. For $\underline{k} = (k_1, \ldots, k_s) \in Z^s$, let $A(\underline{k}; N) = A(k_1, \ldots, k_s; N)$ be the number of n, $0 \le n \le N-1$, such that $\underline{y}_n \equiv \underline{k}$ (modulo m), and let $c_{\underline{k}}$ be the characteristic function of the coset $\underline{k} + (mZ)^s$ of $Z^s/(mZ)^s$. Then for $\underline{x} = (x_1, \ldots, x_s) \in Z^s$ we have

$$c_{\underline{k}}(\underline{x}) = \frac{1}{m^s} \prod_{j=1}^{s} \Big(\sum_{h_j \bmod m} e(h_j(x_j - k_j)/m) \Big) = \frac{1}{m^s} \sum_{\underline{h} \bmod m} e(\underline{h} \cdot (\underline{x} - \underline{k})/m) .$$

Therefore,

$$A(\underline{k}; N) = \sum_{n=0}^{N-1} c_{\underline{k}}(\underline{y}_n) = \frac{1}{m^s} \sum_{n=0}^{N-1} \sum_{\underline{h} \bmod m} e(\underline{h} \cdot (\underline{y}_n - \underline{k})/m) = \frac{1}{m^s} \sum_{\underline{h} \bmod m} e(-\underline{h} \cdot \underline{k}/m) \sum_{n=0}^{N-1} e(\underline{h} \cdot \underline{y}_n / m) ,$$

and so

$$A(\underline{k};N)-\frac{N}{m^s}=\frac{1}{m^s}\sum_{\underline{h}\bmod m}^{*}e(-\underline{h}\cdot\underline{k}/m)\sum_{n=0}^{N-1}e(\underline{h}\cdot\underline{z}_n/m)\,. \tag{2.2}$$

Now let $J=[\alpha_1,\beta_1)\times\ \dots\ \times[\alpha_s,\beta_s)$ be an arbitrary half-open subinterval of $[0,1)^s$. For each j, $1\le j\le s$, we choose the largest closed subinterval of $[\alpha_j,\beta_j)$ of the form $[u_j/m, v_j/m]$ with integers $u_j\le v_j$. The case where for some j no such subinterval of $[\alpha_j,\beta_j)$ exists can be dealt with easily, since we have then $A(J,N)=0$ and $\beta_j-\alpha_j<\frac{1}{m}$, hence

$$\left|\frac{A(J;N)}{N}-V(J)\right|=V(J)<\frac{1}{m}. \tag{2.3}$$

In the remaining case, the integers $u_1,\dots,u_s,\ v_1,\dots,v_s$ are well defined, and we obtain

$$A(J;N)-NV(J)=\sum_{\underline{k};\,u_j\le k_j\le v_j}\left(A(\underline{k};N)-\frac{N}{m^s}\right)+\frac{N}{m^s}(v_1-u_1+1)\cdots(v_s-u_s+1)-NV(J)$$

$$=\frac{1}{m^s}\sum_{\underline{h}\bmod m}^{*}\Big(\sum_{\underline{k};\,u_j\le k_j\le v_j}e(-\underline{h}\cdot\underline{k}/m)\Big)\Big(\sum_{n=0}^{N-1}e(\underline{h}\cdot\underline{z}_n/m)\Big)+N\Big(\frac{(v_1-u_1+1)\cdots(v_s-u_s+1)}{m^s}-V(J)\Big)$$

by using (2.2). It follows that

$$\left|\frac{A(J;N)}{N}-V(J)\right|\le\frac{1}{m^s}\sum_{\underline{h}\bmod m}^{*}\left|\sum_{\underline{k};\,u_j\le k_j\le v_j}e(\underline{h}\cdot\underline{k}/m)\right|\left|\frac{1}{N}\sum_{n=0}^{N-1}e(\underline{h}\cdot\underline{z}_n/m)\right|$$

$$+\left|\frac{(v_1-u_1+1)\cdots(v_s-u_s+1)}{m^s}-V(J)\right|. \tag{2.4}$$

For fixed $\underline{h}=(h_1,\dots,h_s)\in Z^s$ we have

$$\left|\sum_{\underline{k};\,u_j\le k_j\le v_j}e(\underline{h}\cdot\underline{k}/m)\right|=\left|\sum_{\underline{k};\,0\le k_j\le v_j-u_j}e(\underline{h}\cdot\underline{k}/m)\right|=\prod_{j=1}^{s}\left|\sum_{k_j=0}^{v_j-u_j}e(h_jk_j/m)\right|.$$

Now

$$\left|\sum_{k_j=0}^{v_j-u_j}e(h_jk_j/m)\right|=v_j-u_j+1\le m=\frac{m}{r(h_j,m)}$$

if $h_j\equiv 0 \bmod m$, and

$$\left|\sum_{k_j=0}^{v_j-u_j} e(h_j k_j/m)\right| = \frac{|e(h_j(v_j-u_j+1)/m)-1|}{|e(h_j/m)-1|} \le \frac{2}{|e(h_j/m)-1|} = \frac{1}{\sin \pi \|h_j/m\|} = \frac{m}{r(h_j,m)}$$

if $h_j \not\equiv 0 \bmod m$, and so

$$\left|\sum_{\underline{k};\, u_j \le k_j \le v_j} e(\underline{h}\cdot\underline{k}/m)\right| \le \prod_{j=1}^{s} \frac{m}{r(h_j,m)} = \frac{m^s}{r(\underline{h},m)}. \tag{2.5}$$

In order to estimate the second term on the right-hand side of (2.4), one shows first by induction on s that

$$|\gamma_1 \cdots \gamma_s - \delta_1 \cdots \delta_s| \le \sum_{j=1}^{s} |\gamma_j - \delta_j|$$

whenever $0 \le \gamma_j, \delta_j \le 1$ for $1 \le j \le s$. Consequently,

$$\left|\frac{(v_1-u_1+1)\cdots(v_s-u_s+1)}{m^s} - V(J)\right| = \left|\prod_{j=1}^{s}\frac{v_j-u_j+1}{m} - \prod_{j=1}^{s}(\beta_j-\alpha_j)\right|$$

$$\le \sum_{j=1}^{s}\left|\frac{v_j-u_j+1}{m} - (\beta_j-\alpha_j)\right|.$$

From the definition of u_j and v_j it follows that

$$\alpha_j \le \frac{u_j}{m} < \alpha_j + \frac{1}{m} \quad \text{and} \quad \beta_j - \frac{1}{m} \le \frac{v_j}{m} < \beta_j$$

so that

$$\left|\frac{v_j-u_j+1}{m} - (\beta_j-\alpha_j)\right| < \frac{1}{m} \quad \text{for } 1 \le j \le s.$$

Therefore,

$$\left|\frac{(v_1-u_1+1)\cdots(v_s-u_s+1)}{m^s} - V(J)\right| < \frac{s}{m},$$

and by combining this with (2.4) and (2.5), we arrive at

$$\left|\frac{A(J;N)}{N} - V(J)\right| \le \frac{s}{m} + \sum_{\underline{h} \bmod m}^{*} \frac{1}{r(\underline{h},m)} \left|\frac{1}{N}\sum_{n=0}^{N-1} e(\underline{h}\cdot\underline{y}_n/m)\right|.$$

In view of (2.3), this inequality holds for all J, and by forming the supremum over J on the left-hand side, we obtain the desired inequality for D_N. ∎

The latter part of this proof indicates that it is somewhat unfair to use arbitrary intervals J when computing the discrepancy. The points $\langle y_n \rangle$ lie on the coordinate lattice, and we may as well assume that J is an interval of the type $[a_1/m, b_1/m) \times \cdots \times [a_s/m, b_s/m)$. This assumption disposes of the term s/m in the statement of Lemma 2.1.

Lemma 2.2. For any integer $m \geq 2$ we have

$$\sum_{\underline{h} \bmod m} \frac{1}{r(\underline{h}, m)} < \left(\frac{2}{\pi} \log m + \frac{7}{5}\right)^s .$$

Proof. [Niederreiter78] Since

$$\sum_{\underline{h} \bmod m} \frac{1}{r(\underline{h}, m)} = \left(\sum_{h \bmod m} \frac{1}{r(h, m)} \right)^s ,$$

it suffices to estimate the sum on the right-hand side. We have

$$\sum_{h \bmod m} \frac{1}{r(h, m)} = 1 + \frac{1}{m} \sum_{h \bmod m}^{*} \csc \pi \left\| \frac{h}{m} \right\| \leq 1 + \frac{2}{m} \sum_{h=1}^{[m/2]} \csc \frac{\pi h}{m} ,$$

and by comparing sums with integrals we get

$$\sum_{h=1}^{[m/2]} \csc \frac{\pi h}{m} = \csc \frac{\pi}{m} + \sum_{h=2}^{[m/2]} \csc \frac{\pi h}{m} \leq \csc \frac{\pi}{m} + \int_1^{[m/2]} \csc \frac{\pi x}{m} \, dx$$

$$\leq \csc \frac{\pi}{m} + \frac{m}{\pi} \int_{\pi/m}^{\pi/2} \csc t \, dt = \csc \frac{\pi}{m} + \frac{m}{\pi} \log \cot \frac{\pi}{2m} < \csc \frac{\pi}{m} + \frac{m}{\pi} \log \frac{2m}{\pi} .$$

For $m \geq 6$ we have $(m/\pi) \sin(\pi/m) \geq (6/\pi) \sin(\pi/6)$, hence $\sin(\pi/m) \geq 3/m$. This implies

$$\sum_{h=1}^{[m/2]} \csc \frac{\pi h}{m} < \frac{m}{\pi} \log m + \left(\frac{1}{3} - \frac{1}{\pi} \log \frac{\pi}{2}\right) m \quad \text{for } m \geq 6,$$

and so

$$\sum_{h=1}^{[m/2]} \csc \frac{\pi h}{m} < \frac{m}{\pi} \log m + \frac{m}{5} \quad \text{for } m \geq 6.$$

This last inequality is easily checked for $m = 3$, 4, and 5, so that

$$\sum_{h \bmod m} \frac{1}{r(h, m)} < \frac{2}{\pi} \log m + \frac{7}{5} \quad \text{for } m \geq 3.$$

For $m = 2$, Lemma 2.2 is shown by inspection. ∎

The papers [Niederreiter72], [Niederreiter74], and [Niederreiter76] contain further theory of discrepancy and exponential sums.

CHAPTER 3

Reduction to a Weaker Criterion

Chapter 1 introduced the linear recurrences we are investigating and gave some of the known results on the length of the period. Chapter 2 defined discrepancy as a measure of the distribution of a set of points and presented Niederreiter's lemma, which bounds the discrepancy by an exponential sum. Computing the exponential sum is straightforward but costly; even for small cases the amount of computation becomes prohibitive. Since we are interested in the exponential sum mainly as a bound on the discrepancy and hence as an indicator of the goodness of the distribution of the sequence, in this Chapter 3 we bound the exponential sum by a function involving the number of zeroes occurring in a related sequence. The number of zeroes can be much larger than expected and the discrepancy of the original sequence will still approach zero. The question of equidistribution of the original sequence is thus reduced to a weaker distribution criterion.

In the case of higher order congruences we cannot expect the discrepancy to be less than m^{-s}, since some points are never generated by the recurrence; for example, no integer congruent to 6 mod 8 will ever appear in the Fibonacci sequence. The problem in proving small discrepancy lies in showing that the values which occur more often than expected do not occur *too* often. Intuition for this problem comes from considering recurrences based on primitive trinomials with unit coefficients and looking at the carries that occur when the addition in the recurrence is performed in radix-2 positional notation. The distribution of carries should say something about the distribution of the sequence.

In order to study the relationship between carries and distribution of digits, let us consider the top 3 bits and the bottom 3 bits of $\langle F_n \bmod 64 \rangle$. The period of $\langle F_n \bmod 8 \rangle$ is 12 and

the period of $\langle F_n \bmod 64\rangle$ is 96. In the matrix formulation of Eqs. (1.3) and (1.4) we have $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ and $A^{12} = \begin{pmatrix} 89 & 144 \\ 144 & 233 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 8\begin{pmatrix} 11 & 18 \\ 18 & 29 \end{pmatrix} = I + 8B$. Taking the difference of two vectors 12 apart, we see that $\underline{y}_{t+12} - \underline{y}_t = A^{12}\underline{y}_t - \underline{y}_t = (A^{12} - I)\underline{y}_t = 8B\underline{y}_t$, and in general

$$\underline{y}_{12j+t} - \underline{y}_t = (A^{12j} - I)\underline{y}_t = ((I + 8B)^j - I)\underline{y}_t = (8jB + j\cdot\frac{j-1}{2}\cdot 64B^2 + \cdots)\underline{y}_t \equiv 8jB\underline{y}_t \bmod 64 . \quad (3.1)$$

Breaking the period of $\langle F_n \bmod 64\rangle$ into 8 blocks of 12 and considering each block as a point in 12-space, the points are $\underline{a} + 8t\underline{b} \bmod 64$, where $\underline{a}, \underline{b} \in Z^{12}$ and $0 \le t \le 7$. The coordinates of $\underline{a}$ are the first 12 terms of $\langle F_n \bmod 64\rangle$. The coordinates of $\underline{b}$ satisfy the Fibonacci recurrence $b_n = b_{n-1} + b_{n-2} \bmod 8$, since they are the difference of two Fibonacci sequences. Here are two tables illustrating the relationship implied by Eq. (3.1). The first gives $\langle F_n \bmod 64\rangle$ and the second gives $\langle F'_n \bmod 64\rangle$ where $\langle F'_n\rangle$ satisfies the Fibonacci recurrence, but with initial conditions $F'_0 = 2$, $F'_1 = 5$. The entries in each table are octal integers.

Table 1. Octal values of $F_{12j+t} \bmod 64$

j \ t	0	1	2	3	4	5	6	7	8	9	10	11	
0	00	01	01	02	03	05	10	15	25	42	67	31	
1	20	51	71	42	33	75	30	25	55	02	57	61	
2	40	21	61	02	63	65	50	35	05	42	47	11	
3	60	71	51	42	13	55	70	45	35	02	37	41	
4	00	41	41	02	43	45	10	55	65	42	27	71	
5	20	11	31	42	73	35	30	65	15	02	17	21	
6	40	61	21	02	23	25	50	75	45	42	07	51	
7	60	31	11	42	53	15	70	05	75	02	77	01	
	20	50	70	40	30	70	20	10	30	40	70	30	row difference

Table 2. Octal values of F'_{12j+t} mod 64

j \ t	0	1	2	3	4	5	6	7	8	9	10	11	
0	02	05	07	14	23	37	62	21	03	24	27	53	
1	02	55	57	34	13	47	62	31	13	44	57	23	
2	02	25	27	54	03	57	62	41	23	64	07	73	
3	02	75	77	74	73	67	62	51	33	04	37	43	
4	02	45	47	14	63	77	62	61	43	24	67	13	
5	02	15	17	34	53	07	62	71	53	44	17	63	
6	02	65	67	54	43	17	62	01	63	64	47	33	
7	02	35	37	74	33	27	62	11	73	04	77	03	
	00	50	50	20	70	10	00	10	10	20	30	50	row difference

This observation can be applied directly to the evaluation of the sum in Niederreiter's Lemma. Let $m = p^{2\alpha}$, $\langle \underline{y}_n \rangle$ = successive tuples of $\langle y_n \bmod m \rangle$, and N = length of one period mod m. Then

$$\sum_{n=0}^{N-1} e(\underline{h}\cdot\underline{y}_n/p^{2\alpha}) = \sum_{j=0}^{N/p^{\alpha}-1} \sum_{t=0}^{p^{\alpha}-1} e(\underline{h}\cdot(\underline{a}_j + p^{\alpha}t\underline{b}_j)/p^{2\alpha}) = \sum_{j=0}^{N/p^{\alpha}-1} e(\underline{h}\cdot\underline{a}_j/p^{2\alpha}) \sum_{t=0}^{p^{\alpha}-1} e(\underline{h}\cdot t\underline{b}_j/p^{\alpha}) \quad (3.2)$$

The inner sum is a geometric series with ratio $e(\underline{h}\cdot\underline{b}_j/p^{\alpha})$. In fact, the sum is zero if $\underline{h}\cdot\underline{b}_j \neq 0$ mod p^{α}, and is p^{α} if $\underline{h}\cdot\underline{b}_j = 0$ mod p^{α}. Also, the sum depends only on $\underline{h}$ mod p^{α}, which is $\underline{h}$ mod $n^{1/2}$.

Next we observe that $\langle \underline{h}\cdot\underline{y}_n \rangle$ satisfies the same recurrence as $\langle y_n \rangle$. To show this, let $y'_n = \underline{h}\cdot\underline{y}_n$. Then

$$\sum_{j=0}^{k-1} a_j y'_{n+j} = \sum_{j=0}^{k-1} a_j \underline{h}\cdot\underline{y}_{n+j} = \underline{h}\cdot\sum_{j=0}^{k-1} a_j \underline{y}_{n+j} = \underline{h}\cdot\underline{y}_{n+k} = y'_{n+k} \quad (3.3)$$

Furthermore, if $\underline{y}_n$ contains an element relatively prime to p, and if $s \leq k$, the g.c.d. of $\langle \underline{h}\cdot\underline{y}_n \rangle$ is equal to the g.c.d. of $\underline{h}$; that is, $\langle \underline{h}\cdot\underline{y}_n \rangle$ does not degenerate into a sequence in which all elements are divisible by a power of p higher than the highest power of p dividing $\underline{h}$. It suffices to prove the case for which the g.c.d. of $\underline{h}$ is 1. Extend $\underline{h}$ to a k-tuple $(h_0, \ldots, h_{k-1})$ by adding $k-s$ zero coordinates, and define the polynomial $H(x)$ by $H(x) = h_0x^0 + \cdots + h_{k-1}x^{k-1}$. Then the polynomial associated with $\langle \underline{h}\cdot\underline{y}_n \rangle$ by Eq. (1.2) is

$H(x)Y(x) \pmod{p, f(x)}$. In the field $Z_p[x]/(f(x))$ both $H(x)$ and $Y(x)$ are nonzero, hence their product is also nonzero. Therefore p does not divide all of the coefficients of $H(x)Y(x)$ and thus $\langle \underline{h}\cdot\underline{y}_n \rangle$ contains an element relatively prime to p. The foregoing observations about $\langle \underline{h}\cdot\underline{y}_n \rangle$ allow us to get rid of vector operations and return to plain integer sequences. Let $a'_j = \underline{h}\cdot\underline{a}_j$, $b'_j = \underline{h}\cdot\underline{b}_j$, $y'_n = \underline{h}\cdot\underline{y}_n$. Then

$$\sum_{n=0}^{N-1} e(\underline{h}\cdot\underline{y}_n/p^{2\alpha}) = \sum_{n=0}^{N-1} e(y'_n/p^{2\alpha}) = \sum_{j=0}^{N/p^{\alpha}-1} e(a'_j/p^{2\alpha}) \sum_{t=0}^{p^{\alpha}-1} e(tb'_j/p^{\alpha}) \tag{3.4}$$

The full sum, and hence an estimate of the discrepancy for sequences which are 2α bits wide, depends on the occurrences of 0 in $\langle b'_j \rangle$, which is only α bits wide. Again we note that the inner sum is a geometric series with a term ratio of $e(b'_j/p^{\alpha})$, and its sum is 0 if $b'_j \not\equiv 0 \bmod p^{\alpha}$ or p^{α} if $b'_j \equiv 0 \bmod p^{\alpha}$. Therefore

$$\sum_{j=0}^{N/p^{\alpha}-1} e(a'_j/p^{2\alpha}) \sum_{t=0}^{p^{\alpha}-1} e(tb'_j/p^{\alpha}) = p^{\alpha} \sum_{\substack{j=0 \\ b'_j \equiv 0 \bmod p^{\alpha}}}^{N/p^{\alpha}-1} e(a'_j/p^{2\alpha}) \tag{3.5}$$

Taking the absolute value,

$$\left| p^{\alpha} \sum_{\substack{j=0 \\ b'_j \equiv 0 \bmod p^{\alpha}}}^{N/p^{\alpha}-1} e(a'_j/p^{2\alpha}) \right| \le p^{\alpha} \sum_{\substack{j=0 \\ b'_j \equiv 0 \bmod p^{\alpha}}}^{N/p^{\alpha}-1} 1$$

$$= p^{\alpha} \cdot (\text{number of zeroes in } \langle b'_j \bmod p^{\alpha} \rangle). \tag{3.6}$$

The number of zeroes in $\langle b'_j \bmod p^{\alpha} \rangle$ will be large if a high power of p divides $\underline{h}$, since $b'_j = \underline{h}\cdot\underline{b}_j$. To analyze this situation, let us go back to the original exponential sum of Niederreiter's lemma and choose $m = p^{\alpha}$. Set

$$S(\alpha) = \sum_{\underline{h} \bmod p^{\alpha}}^{*} \frac{1}{r(\underline{h}, p^{\alpha})} \left| \frac{1}{\tau(p^{\alpha})} \sum_{n=0}^{\tau(p^{\alpha})-1} e(\underline{h}\cdot\underline{y}_n/p^{\alpha}) \right| \tag{3.7}$$

where $\tau(p^{\alpha})$ is the period modulo p^{α}, and let α_0 be an exponent for which

Reduction to a Weaker Criterion

$\tau(p^{\alpha_0+j}) = p^j \tau(p^{\alpha_0})$ for all $j \geq 0$. Then for $\alpha \geq \alpha_0$

$$S(\alpha) = \sum_{d=1}^{\alpha} \sum_{\substack{\underline{h} \bmod p^\alpha \\ \gcd(\underline{h}) = p^{\alpha-d}}} \frac{1}{r(\underline{h}, p^\alpha)} \left| \frac{1}{\tau(p^\alpha)} \sum_{n=0}^{\tau(p^\alpha)-1} e(\underline{h}\cdot\underline{x}_n / p^\alpha) \right|$$

$$= \sum_{d=1}^{\alpha} \sum_{\substack{\underline{h} \bmod p^\alpha \\ \gcd(\underline{h}) = p^{\alpha-d}}} \frac{1}{p^{s(\alpha-d)} r(\underline{h}/p^{\alpha-d}, p^\alpha/p^{\alpha-d})} \left| \frac{1}{\tau(p^\alpha)} \sum_{n=0}^{\tau(p^\alpha)-1} e(\underline{h}/p^{\alpha-d}\cdot\underline{x}_n / (p^\alpha/p^{\alpha-d})) \right|$$

$$= S(\alpha_0 - 1) + \sum_{d=\alpha_0}^{\alpha} p^{-s(\alpha-d)} \sum_{\substack{\underline{h} \bmod p^d \\ \gcd(\underline{h}) = 1}} \frac{1}{r(\underline{h}, p^d)} \left| \frac{1}{\tau(p^d)} \sum_{n=0}^{\tau(p^d)} e(\underline{h}\cdot\underline{x}_n / p^d) \right| .$$

Thus for $\alpha \geq \alpha_0$

$$S(\alpha) = p^{-s} S(\alpha - 1) + \sum_{\substack{\underline{h} \bmod p^\alpha \\ \gcd(\underline{h}) = 1}} \frac{1}{r(\underline{h}, p^\alpha)} \left| \frac{1}{\tau(p^\alpha)} \sum_{n=0}^{\tau(p^\alpha)-1} e(\underline{h}\cdot\underline{x}_n / p^\alpha) \right| . \tag{3.8}$$

Collapsing the chain of inequalities back to the statement of Niederreiter's lemma, for $\alpha \geq \alpha_0$

$$D_{\tau(p^{2\alpha})} \leq \frac{s}{p^{2\alpha}} + \frac{S(2\alpha - 1)}{p^s} + \sum_{\substack{\underline{h} \bmod p^{2\alpha} \\ \gcd(\underline{h}) = 1}} \frac{1}{r(\underline{h}, p^{2\alpha})} \left| \frac{1}{\tau(p^{2\alpha})} \sum_{n=0}^{\tau(p^{2\alpha})} e(\underline{h}\cdot\underline{x}_n / p^{2\alpha}) \right|$$

$$\leq \frac{s}{p^{2\alpha}} + \frac{S(2\alpha - 1)}{p^s} + \sum_{\substack{\underline{h} \bmod p^{2\alpha} \\ \gcd(\underline{h}) = 1}} \frac{1}{r(\underline{h}, p^{2\alpha})} \cdot \frac{p^\alpha}{\tau(p^{2\alpha})} \cdot (\text{number of zeroes in } \langle b_j' \bmod p^\alpha \rangle) . \tag{3.9}$$

For $\alpha \geq \alpha_0$ the g.c.d. of the components of $\underline{b}_j$ is 1, by Lemma 1.4. Thus by an earlier remark the g.c.d. of the $\langle b_j' \rangle$ is 1, because $b_j' = \underline{h}\cdot\underline{b}_j$ and the sum (3.9) is restricted to those $\underline{h}$ for which g.c.d.$(\underline{h}) = 1$. Replacing the number of zeroes in $\langle b_j' \bmod p^\alpha \rangle$ by the maximum number of zeroes in a nondegenerate cycle mod p^α and applying the bound of Lemma 2.2

gives

$$D_{T(p^{2\alpha})} \le \frac{s}{p^{2\alpha}} + \frac{S(2\alpha-1)}{p^{s}} + (c_1 2\alpha)^{s} \cdot \frac{p^{\alpha}}{\tau(p^{2\alpha})} \cdot (\text{maximum number of zeroes in nondegenerate cycle mod } p^{\alpha}) . \quad (3.10)$$

for some c_1 which depends on p but not on α. Remembering that $\tau(p^{2\alpha})$ is proportional to $p^{2\alpha}$, we see that the discrepancy in $s \le k$ dimensions will approach zero if the maximum number of zeroes in a cycle modulo p^{α} is $o(\alpha^{-s} p^{\alpha})$. This is a very weak condition. If the elements of the cycle are evenly distributed then the expected number of zeroes is a constant. Equation 3.10 shows that the discrepancy will tend to zero even if the number of zeroes in a cycle is exponentially increasing, as long as the rate of increase is $p - \epsilon$ for some $\epsilon > 0$.

CHAPTER 4

Analysis of the Number of Zeroes in a Cycle

The result of Chapter 3 expresses a bound on the discrepancy in terms of the number of zeroes occurring in sequences satisfying the recurrence. (See [Hall38b] for an early application of a similar bound.) In this chapter we will investigate the number of zeroes appearing in any cycle, and try to bound it from above by a function which is small enough to force the discrepancy to zero as the modulus increases.

For the Fibonacci sequence modulo 2^α the bound is particularly small.

Theorem 4.1. At most two zeroes appear in a cycle satisfying the Fibonacci recurrence $y_{n+2} \equiv y_{n+1} + y_n$ modulo 2^α when y_0 and y_1 are relatively prime.

Proof. If a zero appears at all then shift the cycle so that the zero is the first element. Then $y_0 \equiv 0$, $y_1 \equiv a$, and a must be odd since y_0 and y_1 are relatively prime. Thus the cycle is merely a multiple of the Fibonacci sequence, where $y_0 = 0$, $y_1 = 1$. Modulo 8 the Fibonacci sequence is 0 1 1 2 3 5 0 5 5 2 7 1. By inspection there is one zero mod 2, one zero mod 4, and there are two zeroes mod 8. We will show by induction that $F_{3\cdot 2^{\alpha-2}} = 2^\alpha q$ and $F_{3\cdot 2^{\alpha-2}-1} = 1 + 2^{\alpha-1} r$ where $\alpha \geq 3$ and q, r are odd integers. This is true by inspection for $\alpha = 3$. Assume that it is true for $\alpha = j$. Then using the relations $F_{2n+1} = F_n^2 + F_{n+1}^2$ and $F_{2n} = F_n(2F_{n+1} - F_n)$ we have $F_{3\cdot 2^{j-1}} = 2^{j+1} q(1 + 2^{j-1} r - 2^{j-1} q)$ and $F_{3\cdot 2^{j-1}+1} = 1 + 2^j r + 2^{2j-2} r^2 + 2^{2j} q^2$, which is the property for $\alpha = j+1$. Therefore the period doubles and the number of zeroes remains constant as the modulus doubles. ∎

We can use this fact to estimate the discrepancy in two dimensions of a complete cycle of the Fibonacci sequence modulo 2^α. Assume that $S(\alpha) < c\alpha^2 2^{-\alpha/2}$ in the notation of Eq.

(3.8). For $\alpha \geq 6$ we have $7/5 + (2/\pi)\log 2^{\alpha} < \log 2^{\alpha}$. By Eq. (3.9), $S(\alpha+1) \leq 2^{-2}S(\alpha) + 2\cdot 2^{1/2}(\alpha+1)^2 2^{-(\alpha+1)/2} = c\alpha^2 2^{-\alpha/2-2} + (\alpha+1)^2 2^{3/2-(\alpha+1)/2} = c(\alpha+1)^2 2^{-(\alpha+1)/2}\left(2^{-3/2}\left(\frac{\alpha}{\alpha+1}\right)^2 + c^{-1}2^{3/2}\right) < c(\alpha+1)^2 2^{-(\alpha+1)/2}$ if $c > 8/(2^{3/2}-1)$. If we also select c large enough so that $S(6)$ satisfies the inequality then a simple induction will establish the bound for $\alpha \geq 6$. Therefore $D_N \leq 2^{1-\alpha} + c\alpha^2 2^{-\alpha/2}$.

For recurrences of higher degree than the Fibonacci sequence, the number of zeroes in a cycle can increase as the modulus increases.

Theorem 4.2. If $k > 2$ and $\alpha \geq \alpha_0$ then the maximum number of zeroes in a nondegenerate cycle modulo $p^{2\alpha}$ is at least p^{α}.

Proof. Construct k initial elements $y_0, \ldots, y_{k-1}$ as follows. Set $y_0 \equiv 0 \bmod p^{2\alpha}$. Choose $y_1, \ldots, y_{k-1} \bmod p^{\alpha}$ so that $\underline{e}_1 B \underline{y}_0 \equiv 0 \bmod p^{\alpha}$, where B is the matrix defined in Eq. (1.4) and $\underline{e}_1 = (1, 0, 0, \ldots, 0)$ is a unit vector with a 1 in the first coordinate. Since y_0 has been specified, the constraint on $y_1, \ldots, y_{k-1}$ is one equation in $k-1$ unknowns. For $k = 2$ (as in the Fibonacci sequence), y_1 would have to be 0 mod p^{α} and g.c.d.(y_0, y_1) would be p^{α}. For $k > 2$, well-known methods guarantee the existence of a nontrivial solution for $y_1, \ldots, y_{k-1}$. It is easy to see that $\underline{y}_{jT} \equiv \underline{e}_1 A^{jT} \underline{y}_0 \equiv \underline{e}_1 (I + jp^{\alpha}B)\underline{y}_0 \equiv \underline{e}_1 \cdot \underline{y}_0 + \underline{e}_1 jp^{\alpha} B \underline{y}_0 \equiv 0 \bmod p^{2\alpha}$ for $0 \leq j < p^{\alpha}$ and that $\underline{y}_{jT}$ are distinct elements of a cycle modulo $p^{2\alpha}$. ∎

In order to obtain a good bound on the discrepancy, it is necessary to show that, as the modulus goes from p^{α} to $p^{\alpha+1}$, the number of zeroes increases by a factor which is eventually less than p. We will try to show that the number of zeroes eventually increases at a rate no faster than $p^{(k-2)/(k-1)}$, where k is the degree of the recurrence.

The fundamental idea will be counting in two different ways the total number of zeroes in the cycles corresponding to all of the possible initial conditions. Consider a large tabular

array in which each cycle appears as a row. The first column of the array contains the first element of each cycle; the second column is composed of the second element from each cycle, and so on. Let b_j be the number of zeroes in the j^{th} row of the array. Then $\sum_j \binom{b_j}{t}$ counts the number of t-tuples of zeroes in the array, with all elements of a tuple required to appear in the same row. The number of t-tuples of zeroes can be counted another way. For each t-tuple of column indices, count the number of rows which have zeroes in all of those t columns. Since the two methods of counting must agree, a bound achieved by considering one counting method can be applied to the other counting method.

To count the zeroes appearing in all of the cycles, we must first know how many cycles there are. Each cycle is determined by its first k elements and each element can take on p^α values, so there are $p^{\alpha k}$ possible cycles. Some of these cycles are isomorphic under cyclic shift, and of course the number of zeroes in cycles which are cyclic shifts of one another is the same. To reduce our counting effort by applying knowledge of the isomorphism, we should divide by the period. A cycle in which all the elements are divisible by p has a shorter period than a cycle containing an element relatively prime to p, so it is necessary to count the cycles according to the highest power of p dividing all elements. Let a_j be the number of cycles for which p^j is this highest power. Then by counting all the cycles we have $a_0 + a_1 + \cdots + a_\alpha = p^{\alpha k}$. If p^α is large enough so that the period is multiplied by p as the modulus increases from p^α to $p^{\alpha+1}$, then by considering what happens when the zero unit's digit is removed from a cycle in which all the elements are divisible by p, we see that $a_1 + \cdots + a_\alpha = p^{(\alpha-1)k}$. Thus $a_0 = p^{\alpha k} - p^{(\alpha-1)k} = (p^k - 1)p^{(\alpha-1)k}$. Considering only those cycles containing an element relatively prime to p and allowing for cyclic shifts, there are $(p^k - 1)p^{(\alpha-1)k}/(p^k - 1)p^{\alpha-1-(c-1)} = c_1 p^{(\alpha-1)(k-1)}$ cycles.

Some of these cycles are isomorphic under multiplication by a number prime to p. Cycles which are multiples of each other have the same number of zeroes. In fact, a cycle may be non-trivially isomorphic to itself under such a multiplication. For instance, in the

Analysis of the Number of Zeroes in a Cycle

Fibonacci sequence modulo 8 the second half cycle 0 5 5 2 7 1 is 5 times the first half cycle 0 1 1 2 3 5. In this situation the number 5 is called a multiplier. If a cycle has a multiplier then obviously the number of zeroes in the cycle is a multiple of the number of zeroes appearing in the first partial cycle, which is called a block [Ward33]. Multipliers complicate some of the later analyses. Any constant number of multipliers can be tolerated, but it is preferable not to deal with multipliers at all. Therefore we must be able to detect when a cycle has a multiplier.

Suppose that a is a multiplier after n iterations of the recurrence, i.e., $y_n \equiv ay_0$, $y_{n+1} \equiv ay_1$, ... Let $g(x)$ be the polynomial corresponding to the initial conditions $(y_0, y_1, \ldots, y_{k-1})$. Then

$$x^n g(x) \equiv a g(x) \text{ (modd } p^\alpha, f(x))$$
$$(x^n - a)g(x) \equiv 0 \text{ (modd } p^\alpha, f(x))$$
$$(x^n - a)g(x) = f(x)u(x) + p^\alpha v(x) \quad \text{for some } u(x), v(x).$$

Since $f(x)$ is primitive in $Z_p[x]$, either $f(x)$ divides $x^n - a$ or $f(x)$ and $x^n - a$ are relatively prime. Suppose they are relatively prime. Then by Hensel's Lemma we can find $c(x)$, $d(x)$ such that $(x^n - a)c(x) + f(x)d(x) \equiv 1 \bmod p^\alpha$. Multiplying by $g(x)$ we find

$$(x^n - a)g(x)c(x) + f(x)g(x)d(x) \equiv g(x) \bmod p^\alpha$$
$$f(x)u(x)c(x) + p^\alpha v(x)c(x) + f(x)g(x)d(x) \equiv g(x) \bmod p^\alpha$$
$$0 \equiv g(x) \text{ (modd } p^\alpha, f(x)).$$

Thus a would be a multiplier for the trivial cycle (0) only. Consider the other possibility, that $f(x)$ divides $x^n - a$ in $Z_p[x]$ so that $x^n \equiv a$ (modd p, $f(x)$). The multiplier a must be relatively prime to p or else the recurrence could be run backwards to deduce that all of the initial values were non-prime to p. Let j be the order of a mod p. Then $x^{jn} \equiv a^j \equiv 1$ (modd p, $f(x)$). Thus $\tau(p)$ divides jn. Noting that $p - 1$ divides $p^k - 1$ and that j divides $p - 1$, we have $\tau(p)/(p-1)$ divides n. Specializing to $p = 2$, $\tau(2)$ divides n. If we consider the case in which n is the smallest number of iterations producing a multiplier then n

divides $\tau(p^\alpha)$. For $p = 2$ these two divisibility conditions restrict n to the form $2^{\alpha-c}(2^k - 1)$, and thus a is a 2^j root of unity for some j. The powers a^2, a^4, a^8, ... must also be multipliers, and in particular one of the square roots of 1 must be a multiplier. The square roots of 1 mod 2^α are +1, -1, $2^{\alpha-1}+1$, and $2^{\alpha-1}-1$. The case of +1 is trivial. If -1 or $2^{\alpha-1}-1$ were a multiplier, it would correspond to $2^{-1}\tau(2^\alpha) = \tau(2^{\alpha-1})$ iterations, contradicting the known period *mod* $2^{\alpha-1}$. Therefore $2^{\alpha-1}+1$ must be a multiplier corresponding to one-half the period. In matrix formulation of Eq. (1.3) and (1.4),

$$A^{\tau/2}\underline{x} \equiv (2^{\alpha-1}+1)\underline{x} \bmod 2^\alpha$$
$$(2^{\alpha-1}B+I)\underline{x} \equiv (2^{\alpha-1}+1)\underline{x} \bmod 2^\alpha$$
$$2^{\alpha-1}(B-I)\underline{x} \equiv 0 \bmod 2^\alpha$$

Thus B-I is singular mod 2. For recurrences of moderate degree this is not hard to determine. In the nonsingular cases there can be no multipliers.

Back to considering cycles isomorphic under multiplication by a number relatively prime to p, we now assume that the recurrence itself has no multipliers. We can divide the number of cycles by $\phi(p^\alpha)$, which in the case $p = 2$ leaves $c_1 2^{(\alpha-1)(k-1)}/2^{\alpha-1} = c_1 2^{(\alpha-1)(k-2)}$ cycles containing an odd element and non-isomorphic under cyclic shift and multiplication by a constant.

Knowing the total number of cycles, we begin counting the zeroes by counting the rows which have zeroes in specified columns.

Theorem 4.3 [Hall38a]. There exists a sequence $\langle y_n \rangle$ satisfying the recurrence (1.1) and not identically zero modulo p^α for which $y_n \equiv 0 \bmod p^\alpha$ for $k-1$ arbitrary values of n.

Proof. Let the arbitrary values of n be $n_1, \ldots, n_{k-1}$. Write $\langle y_n \rangle = \langle c_0 w_n + c_1 w_{n+1} + \cdots + c_{k-1} w_{n+k-1} \rangle$ where $\langle w_n \rangle$ is the unit sequence with initial conditions (0, 0, ..., 0, 1) and the c's are to be determined by the $k-1$ congruences

$y_{n_1} \equiv y_{n_2} \equiv \cdots \equiv y_{n_{k-1}} \bmod p^\alpha$. These are $k-1$ homogeneous linear congruences in the k variables $c_0, \ldots, c_k$, so there must exist a solution in which not all the c's vanish and y_n does not vanish identically. ∎

Thus at least one cycle has zeroes for each set of $k-1$ arbitrary positions. To calculate the exact number of cycles having zeroes in the specified positions, we need to know the rank of the coefficient matrix. Suppose that the rank was $k-1$. Then there would be one free parameter in the solution set and the number of solutions would be proportional to p^α as α varied. Varying the parameter would merely generate a solution which was a multiple of the previous solution, so that the corresponding cycles would also be multiples of each other. Thus if the rank of the coefficient matrix was $k-1$ then there would be a constant number of cycles having zeroes in the designated columns. To take advantage of the cyclic shift isomorphism, we can demand that first column be one of the columns containing a zero. This leaves $k-2$ other columns which can be specified, so altogether the number of systems of simultaneous homogeneous linear equations we are considering is proportional to $p^{\alpha(k-2)}$. We have seen that if the coefficient matrices are of rank $k-1$ then each system corresponds to a constant number of cycles. Letting b_j be the number of zeroes in the j^{th} row, this argument shows that $\sum \binom{b_j}{k-1} \le c_1 p^{\alpha(k-2)}$, where the sum is over nondegenerate cycles with a zero in the first position. This implies that $b_j^{k-1} \le c_2 p^{\alpha(k-2)}$ for each j, and hence the maximum number of zeroes in a cycle mod p^α would be bounded by $c_3 p^{\alpha(k-2)/(k-1)}$ for some constants c_2, c_3. This bound is good enough to force the discrepancy to zero as α increases.

For recurrences of degree 3 that have no multipliers, the coefficient matrix is of rank $2 = k-1$. The rank cannot be zero because no k consecutive terms vanish, and if the rank were 1 then the second row would be a multiple of the first. Hence the preceding argument proves the following theorem.

Analysis of the Number of Zeroes in a Cycle

Theorem 4.4. If $f(x)$ is a primitive polynomial of degree 3 in $Z_p[x]$ then the discrepancy over an entire period of t-tuples generated by (1.1) tends to zero as α tends to infinity, for $t \le 3$.

The cases in which the coefficient matrix determined by the $k-1$ column positions is of rank $k-1-t$ for $t>0$ generate a number of essentially distinct cycles proportional to $p^{\alpha t}$. Since the matrix is not of full rank there is some non-trivial linear combination of the rows which is the zero vector. Using the isomorphism (1.2) between sequences and polynomials, the same non-trivial linear combination of the powers of x corresponding to the designated column positions must be the zero polynomial modd $p^\alpha, f(x)$. This means that we are now interested in the number of polynomials with $k-t$ terms (one of which is the constant term c_0x^0) which are congruent to zero modd $p^\alpha, f(x)$. We can choose $k-t-1$ exponents of x from a number of possible exponents proportional to p^α. The coefficient for each of the $k-t$ terms can be chosen in p^α ways, but this gives p^α times too many choices because of constant multiples. Thus the number of polynomials to consider is proportional to $p^{\alpha(k-t-1)}p^{\alpha(k-t)}p^{-\alpha} = p^{2\alpha(k-t-1)}$. If these polynomials were evenly distributed among the $p^{\alpha k}$ residue classes modd $p^\alpha, f(x)$ then the probable number of polynomials congruent to zero would be $p^{\alpha(k-2t-2)}$, and the number of $(k-1)$ tuples of zeroes would be proportional to $p^{\alpha(k-t-2)}$. Summing for t from 0 to $k-2$ gives a number proportional to $p^{\alpha(k-2)}$ and therefore $b_j < c_4 p^{\alpha(k-2)/(k-1)}$.

Another way to attack the problem of bounding the maximum number of zeroes in a cycle modulo p^α uses the matrix formulation of the recurrence. Let A be the matrix (1.3). Assume that α is large enough so that the period is multiplied by p as α increases by 1, and that $y_n \equiv 0 \bmod p^\alpha$. Let $\tau = \tau(p^\alpha)$ be the period mod p^α, let $\underline{y} = (y_n, \ldots, y_{n+k-1})^T$ be the k-tuple of residues at y_n, let B be the matrix (1.4) so that $A^\tau = I + p^\alpha B$, and let $\underline{z} = (z_n, \ldots, z_{n+k-1})^T = B\underline{y}$. Let $p^{\alpha+i} \| y_n$ and $p^j \| z_n$. If $i<j$ then modulo $p^{\alpha+j+1}$ there will be no zeroes with indices congruent to n modulo $\tau(p^\alpha)$. If $i \ge j$ then as α increases

there will eventually be $(p-1)p^j$ zeroes in each period with indices congruent to n modulo $\tau(p^\alpha)$. This shows that, given any initial conditions over the integers, the number of zeroes in a cycle generated from those initial conditions will eventually be a constant. The problem is that as the modulus increases the sets of initial conditions which are allowed also varies; the number of zeroes in a cycle converges pointwise (for each set of initial values over the integers), but the question of uniform behavior is as yet unanswered.

Establishing a nontrivial bound on the number of zeroes in a period of a recurrence satisfying (1.1) is a worthwhile future project. Any bound which is $o(\alpha^{-k}p^\alpha)$ can be used in Eq. (3.10) to show that the discrepancy of a full period of k-tuples tends to zero as α tends to infinity. The conjecture below might possibly be proved by establishing suitable bounds on the distribution among the residue classes modd p^α, $f(x)$ of polynomials with k or. fewer terms and degree less than $\tau(p^\alpha)$, or by other means.

Conjecture 4.1. The maximum number of zeroes in a nondegenerate period of a recurrence generated by Eq. (1.1) is less than $cp^{\alpha(k-2)/(k-1)}$ for some c depending on p and k but not on α.

CHAPTER 5

Practical Considerations

In Chapter 4 we saw indications that, when considered over an entire cycle, the tuples of residues probably become equidistributed as the recurrence is computed with more and more bits. In any practical situation only a moderately large, fixed number of bits are used; most current machines use from 24 to 48 bits to store the fraction of a single-precision floating-point number. Even so, it is highly unlikely that an entire period will ever be used. Choosing $p = 2$, $\alpha = 24$, and $k > 50$ implies a period of about 2^{70}, which is in the range of the total number of states that all the computers in the world have ever been in (2^{20} machines $\cdot$ 2^{5} years $\cdot$ 2^{25} seconds/year $\cdot$ 2^{20} states/machine-second). A truly random sequence is locally non-random in some places; such places should not occur frequently in the portions of a computer-generated sequence which are likely to be used. To be recommended for general use, a random number generator must be efficient and easy to implement on a variety of machines. This chapter addresses these practical considerations and presents the results of comparisons with a standard linear congruential generator $y_{n+1} = ay_n + b \bmod 2^{\alpha}$.

A random number generator based on recurrence (1.1) can be efficiently implemented. Only the k most recent values of y need to be remembered, and these can be stored in an array which is accessed cyclically. The number of arithmetic operations involved can be minimized by choosing $f(x)$ to be a primitive trinomial modulo 2. (Some primitive polynomials mod 2 have been tabulated in [Watson62] and [Zierler68].) If $f(x)$ is chosen in this way then the recurrence is $y_n = \pm y_{n-j} \pm y_{n-k} \bmod 2^{\alpha}$, which requires no multiplication or division, only one addition or subtraction, and one reduction modulo 2^{α}. Many

applications require only the fractions $y_n/2^\alpha$. These fractions can be computed exactly in floating-point arithmetic without intermediate integer computation and conversion. This is done by dividing each term of the recurrence by 2^α, so that the computation is done in floating-point numbers modulo 1. It is necessary to prevent any floating-point operation from generating a result greater than 1.0, which would involve a right shift of the fraction part and the possibility of losing one bit, making the computation no longer exact. This can be insured by choosing the signs so that $f(x) = x^k - x^{k-j} + 1$ and the recurrence is $y_n = y_{n-j} - y_{n-k} \bmod 2^\alpha$. This is allowed because $+1 \equiv -1 \bmod 2$ and therefore $x^k - x^{k-j} + 1$ is the same polynomial in Z_2 as $x^k + x^{k-j} + 1$. If the subtraction $y_{n-j}/2^\alpha - y_{n-k}/2^\alpha$ produces a negative result then we add 1.0 to bring the value back into the range $0 \le y_n/2^\alpha < 1$. When performing this addition we must remember to actually perform two additions of 0.5 in order to guarantee that no bits will be lost. (Computers such as the Control Data Corporation 6000 series effectively do not use any guard digits and hence one bit can be lost when preshifting $y_n/2^\alpha$ to align the radix point before the addition.)

To illustrate the fact that a random number generator based on (1.1) can be implemented with little machine dependence, here is a coding in FORTRAN of a generator based on $x^{55} - x^{31} + 1 \bmod 2$. Each time XRAND is called it returns the next random number.

```
      FUNCTION XRAND
      COMMON /XRANDX/ I, J, X(55)
      DATA I /55/
      I = I -1
      IF (I .LE. 0) I = 55
      J = I - 31
      IF (J .LE. 0) J = J + 55
      X(I) = X(I) - X(J)
      IF (X(I) .LT. 0.0) X(I) = (X(I) + 0.5) + 0.5
      XRAND = X(I)
      RETURN
      END
```

The machine dependence lies in the initialization of the first 55 values of the array X. The initial values must be chosen so that $0 \le X(I) < 1$, $2^\alpha X(I)$ is an integer, and $2^\alpha X(I)$ must

be odd for at least one value of I. One way to do this is to set $X(1) = 1/2^{\alpha}$ and $X(I) = 0$ for $2 \le I \le 55$. However, this puts one of the local non-random areas at the start of the generated sequence. It would be better to run the recurrence a million times (say), write out the values of X(I), and have some other routine place these values into X before using XRAND. Alternatively, an auxiliary linear congruential generator can supply the initial values; see [Brent73].

The efficiency of generator XRAND compares favorably with that of standard first-order linear congruential generators. On the Digital Equipment Corporation PDP-10, XRAND can be encoded in 12 instructions, while a standard generator requires 10 instructions. (The number of instructions is for a routine which returns a floating-point value between 0 and 1, is guaranteed to not cause any interrupts, and includes the normal subroutine linkage.) Execution time per call on a KL-10 processor is approximately equal because XRAND uses no multiplication. On the IBM S/370, comparable encodings are 16 instructions (46 bytes) for XRAND and 10 instructions (27 bytes) for a first-order generator.

Four generators were compared using several tests. The generators were

RANDU: $y_n = (2^{16}+3)y_{n-1} \bmod 2^{31}$, $y_0 = 1$. This generator is notorious for its bad distribution in three dimensions.

GOODLC: $y_n = 3141592653y_{n-1} + 2718281829 \bmod 2^{35}$, $y_0 = 0$. This is a standard "linear congruential" generator.

ADDLC: $y_n = y_{n-55} - y_{n-24} \bmod 2^{27}$, $y_0 = \cdots = y_{53} = 0$, $y_{54} = 1$. The tests began with $y'_0 = y_{65536}$. This is an additive generator based on the primitive polynomial $x^{55} - x^{31} + 1 \bmod 2$.

BESTX: $y_n = x_n$ XOR z_n where $x_n = (3141592653x_{n-1} + 2718281829) \bmod 2^{35}$, $x_0 = 0$, $z_n = 314159270z_{n-1} \bmod (2^{35} - 31)$, and $z_0 = 1$. The number $2^{35} - 31$ is the

largest prime less than 2^{35}, and 314159270 is a primitive root. Here XOR is the exclusive-or bit operation.

These generators were compared under the following tests. (See [Knuth69] sec. 3.3.2.)

1D: A one-dimensional distribution test with the interval [0,1) divided into 4096 equal subintervals, and 8 hits expected in each subinterval.

2D: A two-dimensional distribution test with the interval [0,1) × [0,1) divided into 64 × 64 equal subintervals, and 8 hits expected in each subinterval. The ordered pairs used for the test were non-overlapping, i.e., $(y_0, y_1), (y_2, y_3), \ldots$.

3D: A three-dimensional distribution test with the interval [0,1) × [0,1) × [0,1) divided into 16 × 16 × 16 equal subintervals, and 5 hits expected in each subinterval. The ordered triples used for test were $(y_0, y_1, y_2), (y_3, y_4, y_5), \ldots$.

GAP: A test which considers the length of consecutive subsequences $y_j, y_{j+1}, \ldots y_{j+r}$ in which $0 \le \alpha \le y_{j+r} < \beta < 1$ for two fixed real numbers α and β, but the other y's do not. This test was performed with $\alpha = 0, \beta = 0.5$, and thus was a test of "runs above the mean". Gaps of length 0 through 5 and gaps of length greater than 5 were counted until 500 gaps had been tabulated.

MAX10: The maximum element of blocks of 10 consecutive values was selected until 1500 maxima had been chosen. The distribution of the maxima was tested against the theoretical distribution function z^{10} by the Kolmogorov-Smirnov test.

RUN: The length r of consecutive subsequences for which $y_j > y_{j+1} > \cdots > y_{j+r} \le y_{j+r+1}$ was tabulated until 500 runs had occurred. New runs were started at y_{j+r+2}, and runs of length 5 or more were grouped together for the analysis.

PERMUT: The order relations among consecutive blocks of 4 values were

tabluated, with each of the 24 possible orderings expected to occur 150 times.

Except for the test MAX10, where the Kolmogorov-Smirnov (KS) test was used, the chi-squared statistic was calculated using the appropriate probabilities. Exceptions to the expected distribution were counted when the calculated statistic lay in the 5% tail at either end of the theoretical distribution. Thus a perfectly random sequence would be expected to fail 10% of the tests in each category. The chi-squared values themselves were tested in groups of 16 by the KS test (both KS+ and KS- tests) against the hypothesis that they came from a chi-square distribution. Exceptions were noted for the 5% tails at both ends of the KS distribution.

Each test was repeated until a conveniently large percentage of the first $3 \cdot 2^{20}$ values of each generator had been tested. The following tables summarize the results.

Table 3. Results of tests on RANDU

test	repetitions	± 5% tails	KS tests	KS ± 5% tails
1D	96	12	12	3
2D	48	5	6	1
3D	48	48	6	6
MAX10	384	50		
GAP	192	20	24	4
RUN	192	26	24	6
PERMUT	192	11	24	0

Table 4. Results of tests on GOODLC

test	repetitions	± 5% tails	KS tests	KS ± 5% tails
1D	96	10	12	3
2D	48	5	6	1
3D	48	4	6	0
MAX10	384	39		
GAP	192	19	24	2
RUN	192	29	24	9
PERMUT	192	20	24	2

Table 5. Results of tests on ADDLC

test	repetitions	± 5% tails	KS tests	KS ± 5% tails
1D	96	13	12	0
2D	48	6	6	0
3D	48	5	6	0
MAX10	384	34		
GAP	192	19	24	3
RUN	192	29	24	8
PERMUT	192	17	24	3

Table 6. Results of tests on BESTX

test	repetitions	± 5% tails	KS tests	KS ± 5% tails
1D	96	15	12	0
2D	48	6	6	1
3D	48	3	6	0
MAX10	384	41		
GAP	192	16	24	3
RUN	192	21	24	12
PERMUT	192	13	24	0

The tests confirm the bad three-dimensional distribution of the values generated by RANDU. All generators had difficulty with the RUN test. The code for this test was carefully examined for systematic errors, but none were found. The exceptional chi-square values tended to be extremely small (less than 1.0) or just above the upper 5% tail cutoff.

Generator ADDLC compares favorably with GOODLC and BESTX in these tests. To the extent that this testing procedure is valid for a particular task requiring random numbers, ADDLC can be recommended as an acceptable generator.

Bibliography

[Brent73]
Richard P. Brent, *Algorithms for Minimization without Derivatives*, (Englewood Cliffs, New Jersey: Prentice-Hall) 1973, 163–164.

[Franklin64]
Joel N. Franklin, *Equidistribution of matrix-power residues modulo one*, Math. Comp. 18 (1964), 560–568.

[Green59]
Bert F. Green, Jr., J. E. Keith Smith, and Laura Klem, *Empirical tests of an additive random number generator*, JACM 6 (1959), 527–537.

[Hall38a]
Marshall Hall, *An isomorphism between linear recurring sequences and algebraic rings*, Trans. AMS 44 (1938), 196–218.

[Hall38b]
Marshall Hall, *Equidistribution of residues in sequences*, Duke Math. J. 4 (1938), 691–695.

[Knuth69]
D. E. Knuth, *Seminumerical Algorithms*, (Reading, Massachusetts: Addison-Wesley) 1969, Sec. 3.2.2, pp. 25–34.

[Marsaglia72]
George Marsaglia, *The structure of linear congruential sequences*, in S. K. Zaremba (ed.), *Applications of Number Theory to Numerical Analysis*, (New York: Academic Press) 1972, 266–267.

[Niederreiter72]
H. Niederreiter, *On the distribution of pseudo-random numbers generated by the linear congruential method*, Math. Comp., 26 (1972) 793–795; II, 28 (1974) 1117–1132; III, 30 (1976) 571–597.

[Niederreiter74]
H. Niederreiter, *Some new exponential sums with applications to pseudo-random numbers*, Colloquium on Number Theory (Debrecen, 1974).

[Niederreiter76]
H. Niederreiter, *On the cycle structure of linear recurring sequences*, Math. Scand. 38 (1976), 53–77.

[Niederreiter78]
H. Niederreiter, *Pseudo-random numbers and optimal coefficients*, Adv. Math., to

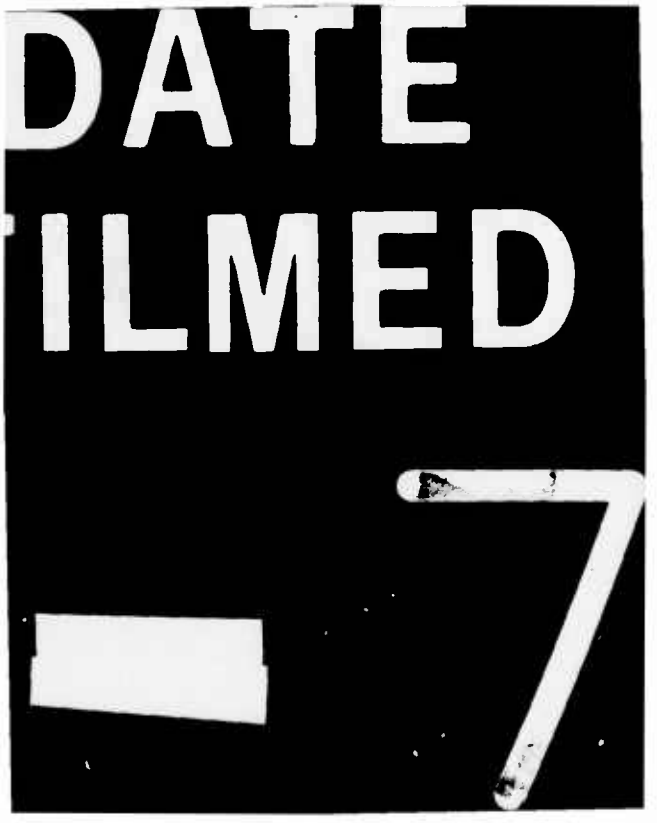
appear 1978; advance report in *Statistical independence of linear congruential pseudo-random numbers*, Bull. AMS 82 (1976), 927-929.

[Ward31]
Morgan Ward, *The distribution of residues in a sequence satisfying a linear recursion relation*, Trans. AMS 33 (1931), 166-190.

[Ward33]
Morgan Ward, *The arithmetical theory of linear recurring series*, Trans. AMS 35 (1933), 600-628.

[Watson62]
E. J. Watson, *Primitive polynomials (mod 2)*, Math. Comp. 16 (1962), 368-369.

[Zierler68]
Neal Zierler and John Brillhart, *On primitive trinomials (mod 2)*, Information and Control 13 (1968), 541-554.


DATE
ILMED