Journal of Statistical Physics, Vol. 73, No. 1/2, 1993

Transient and Cyclic Behavior of Cellular Automata with Null Boundary Conditions

John G. Stevens,[1] Ronald E. Rosensweig,[2] and A. E. Cerkanowicz[3,4]

Received October 14, 1992; final June 9, 1993

One-dimensional cellular automata (CA) over finite fields are studied in which each interior cell is updated to contain the sum of the previous values of its two nearest neighbors. Boundary cells are updated according to null boundary conditions. For a given initial configuration, the CA evolves through transient configurations to an attracting cycle. The dependence of the maximal transient length and maximal cycle length on the number of cells is investigated. Both can be determined from the minimal polynomial of the update matrix, which in this case satisfies a useful recurrence relation. With cell values from a field of characteristic two, the explicit dependence of the maximal transient length on the number of cells is determined. Extensions and directions for future work are presented.

KEY WORDS: Cellular automata; discrete dynamical systems.

1. INTRODUCTION

Cellular automata (CA) are discrete dynamical systems consisting of a regular array of cells, each of which can be assigned a value.[(1)] At discrete times, these values are changed (updated) according to a prescribed rule. In this work, we deal exclusively with update rules which are based on a deterministic, Markovian, local transition rule. Namely, a neighborhood rule is given which specifies a finite, ordered set (neighborhood) of cells for

[1] Department of Mathematics and Computer Science, Montclair State College, Upper Montclair, New Jersey 07043. E-mail: STEVENS@APOLLO.MONTCLAIR.EDU.

[2] Corporate Research, Exxon Research and Engineering Company, Annandale, New Jersey 08801. E-mail: REROSEN@ERENJ.COM.

[3] Department of Mechanical Engineering, New Jersey Institute of Technology, Newark, New Jersey 07102.

[4] Deceased.

any given cell. After updating, the new value of a cell is a specified function (the *local transition rule*) of the values of its neighborhood cells before updating.

An example which has been much studied is as follows. N cells are arranged circularly. The values stored in them are 0 or 1, regarded as the elements of the finite field with two elements, $GF(2)$. (Notation is summarized in Appendix B.) These cells are assigned the indices $0, 1, \ldots, N-1$, which will be added as elements of the integers modulo N, Z_N. The update times t_i are a strictly increasing sequence of positive reals. Cell values remain constant during (t_i, t_{i+1}). Initial values for all cells are specified at time $t=0$, with the value of the ith cell denoted $x_i^{(0)}$. The value of the ith cell after the jth update (at time t_j) is denoted $x_i^{(j)}$ and is determined by

$$x_i^{(j)} = x_{i-1}^{(j-1)} + x_{i+1}^{(j-1)}$$

where addition of cell values is in $GF(2)$ and of indices is in Z_N. The neighborhood cells of the ith cell are thus its nearest (contiguous) neighbors. The local transition rule is simply to add cell values. Because of the cell arrangement, this case is said to satisfy periodic boundary conditions.[1] In the general setting, we use $X^{(j)}$ or X_j to denote the ordered list of cell values during the jth time interval (the jth generation); $X_j = (x_0^{(j)}, x_1^{(j)}, \ldots, x_{N-1}^{(j)})$, termed a *configuration*. The exact update times are generally not important; rather, we are interested in the sequence of configurations $\{X_j\}$. We denote the operator corresponding to application of the update procedure by T, i.e., $X_{j+1} = T(X_j)$. X_j is the *predecessor* of X_{j+1}, its *successor*. To emphasize that j enumerates "temporal" updates, we will often use t in place of j.

We generalize the above by allowing certain boundary cells to follow a different update rule. For example, consider the case in which the periodic CA for $N+1$ cells as above is implemented with two-state devices the zeroth of which fails in such a way that $x_0^{(j)} = 0$ for $j \geqslant j_0$. The transition rule for cells numbered 2 through $N-1$ would be as before, but

$$x_1^{(j)} = x_2^{(j-1)} \qquad \text{and} \qquad x_N^{(j)} = x_{N-1}^{(j-1)}$$

Equivalently, N cells could have been numbered 1 through N and arranged linearly. The neighborhood rule $i \to (i-1, i+1)$ is then applicable to the interior cells (2 through $N-1$), but not to cells 1 or N. Having a "neighborhood which extends outside the CA" distinguishes the latter as *boundary* cells, with their update covered by boundary rules. These rules are the CA analogs of boundary conditions for differential equations. We call such systems cellular automata with boundary, CA/B. The particular example just given will be designated Rule 90 with null boundary conditions or simply our *usual* rule. This rule has been discussed by Wolfram,[1] which see for an explanation of the "Rule 90" description. With null boundary conditions, CA/B behave similarly to the periodic case and, in general, are of interest in their own right.

When T is a linear transformation, its action can be represented by a matrix A (or A_N to indicate its dependence on the number of cells) with the configuration written as a column vector,

$$X_j = AX_{j-1} = A^j X_0$$

The fact that T comes from a local transition rule endows A with special structure. For example, the periodic case has been studied by Guan and He.[2] They consider the general linear rule of the form

$$x_i^{t+1} = \sum_{j=0}^{N-1} a_j x_{i+j}^{(t)}$$

where the indices are summed modulo N and $a_j \in GF(q)$. Such rules give rise to a *circulant* matrix A with rows obtained by successive circular permutation of the first row. Using results about circulants, they obtain a great deal of information about the general properties of periodic CA from the eigenvalues and Jordan form for A. Previously, Martin *et al.*[3] had obtained many of these results for Rule 90 using dipolynomials.

2. PRELIMINARIES

We now specialize to cellular automata with boundary having cell values in an arbitrary finite field, updated according to the usual rule. It is evident that the transition matrix A for this case is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & 0 & 1 & \cdot & \cdot & \cdot & 0 \\ - & - & - & - & - & - & - & - \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & 0 \end{pmatrix}$$

Because, for a fixed number of cells, there are only finitely many possible configurations, it is clear that starting from any initial configuration, the sequence of successive iterates eventually reaches a cycle (periodic orbit, including those of period one, the fixed points). The initial segment of configurations obtained before entering the cycle represents the transient behavior. In this paper we study the maximal transient length and maximal period as functions of the number of cells. We proceed with formal definitions of these quantities.

Consider a given initial configuration $X_0 \neq 0$ and its successor configurations $X_t = A^t X_0$, $t \in \mathbb{N}$. Either there exists $t \in \mathbb{N}$ such that $A^t X_0 = 0$ or there does not. If not, then because the total number of configurations is finite, it must happen for some i, j that $A^i X_0 = A^j X_0$. More precisely, let $\mathbb{N}_0 \times \mathbb{N}$ be ordered lexicographically, i.e., $(i, j) < (k, l)$ if and only if $i < k$ or $i = k$ with $j < l$. Then there exists a smallest pair (t, c) for which $A^t X_0 = A^{t+c} X_0$. Note that $c > 0$. We call t the transient length and c the cycle length for the configuration X_0, denoting the dependence explicitly by $t(X_0)$ and $c(X_0)$. The maximal transient length τ and maximal cycle length γ are then defined by

$$\tau = \underset{\text{initial configurations}}{\text{max over all}} \; t(X)$$

$$\gamma = \underset{\text{initial configurations}}{\text{max over all}} \; c(X)$$

When we wish to emphasize the dependence on the number of cells N, we use τ_N and γ_N. It may happen that for all X, there exists t such that $A^t X = 0$. In this nilpotent case, we define τ to be the maximum such t over all configurations and γ to be zero.

3. TWO IMPORTANT POLYNOMIALS

For $\gamma = 0$, A is a nilpotent matrix with $A^\tau = 0$, i.e., the minimal polynomial of A is λ^τ. For $\gamma \neq 0$, if Y and Z are configurations belonging to cycles of length c_1 and c_2, respectively, then $Y + Z$ belongs to a cycle of length equal to the least common multiple of c_1 and c_2. Thus, the length of any cycle is a divisor of the maximal cycle length. Moreover, if X_0 is an initial configuration which has maximal transient length $(\tau > 0)$, then there exists a configuration Y such that $A^\tau Y = 0$ and $A^{\tau - 1} Y \neq 0$. This statement follows because $A^\tau X_0 = A^\tau Z$, where Z is on the cyclic orbit and therefore $Y = X_0 - Z$ is as claimed.

By superposition, every cycle has transient "tails" of length τ, i.e., starting from $Y + X$, where X is on the cycle. Because $A^\tau X = A^{\tau + \gamma} X$ for all X, A must satisfy the "transient and cycle" polynomial $\Phi(\lambda) = \lambda^\tau - \lambda^{\tau + \gamma} = \lambda^\tau (1 - \lambda^\gamma)$ over $GF(q)$, i.e., $\Phi(A) = 0$. If the minimal polynomial of A is $\mu(\lambda)$ $[= \mu_N(\lambda)]$, then $\mu(\lambda) \mid \Phi(\lambda)$. If we know $\mu(\lambda)$, we show below that it is easy to find $\Phi(\lambda)$ and consequently to find τ and γ.

The determination of $\mu_N(\lambda)$ is provided by the observation that the invariant factors of $\lambda I - A$ are of the form 1, 1,..., 1, $p(\lambda)$, i.e., the characteristic polynomial of A is the minimal polynomial of A. [Recall that if

$\delta_1(\lambda),\ldots,\delta_N(\lambda)$ are the invariant factors of $\lambda I-A$ with $\delta_i(\lambda)\mid\delta_j(\lambda)$ for $i<j$, then $\delta_N(\lambda)$ is the minimal polynomial of A.] This observation follows readily from permuting the rows of $-(\lambda I-A)$ to obtain

$$\begin{pmatrix} 1 & -\lambda & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & -\lambda & 1 & \cdot & \cdot & \cdot & 0 \\ - & - & - & - & - & - & - & - \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & -\lambda & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -\lambda \\ -\lambda & 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}$$

from which it is clear that the 1's on the diagonal can be used to annihilate the entries in the bottom row below them and then the entries in their row.

Because $\mu_N(\lambda)$ is equal to the characteristic polynomial of A, we have

$$\mu_N(\lambda)=\det\begin{pmatrix} \lambda & -1 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & \lambda & -1 & \cdot & \cdot & \cdot & 0 \\ - & - & - & - & - & - & - \\ 0 & \cdot & \cdot & \cdot & -1 & \lambda & -1 \\ 0 & \cdot & \cdot & \cdot & 0 & -1 & \lambda \end{pmatrix}_{N\times N}$$

$$=\lambda\cdot\det\begin{pmatrix} \lambda & -1 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & \lambda & -1 & \cdot & \cdot & \cdot & 0 \\ - & - & - & - & - & - & - \\ 0 & \cdot & \cdot & \cdot & -1 & \lambda & -1 \\ 0 & \cdot & \cdot & \cdot & 0 & -1 & \lambda \end{pmatrix}_{N-1\times N-1}$$

$$-(-1)\det\begin{pmatrix} -1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda & -1 & 0 & \cdot & \cdot & 0 \\ 0 & -1 & \lambda & -1 & \cdot & \cdot & 0 \\ - & - & - & - & - & - & - \\ 0 & \cdot & \cdot & \cdot & 0 & -1 & \lambda \end{pmatrix}_{N-1\times N-1}$$

or $\mu_N(\lambda)=\lambda\mu_{N-1}(\lambda)-\mu_{N-2}(\lambda)$. The initial conditions are $\mu_1(\lambda)=\lambda$ and $\mu_2(\lambda)=\lambda^2-1$ or, equivalently, $\mu_0(\lambda)=1$ and $\mu_1(\lambda)=\lambda$. This recurrence relation can be used to generate the minimal polynomials quite simply.

It is of interest to develop an explicit representation for $\mu_N(\lambda)$ by the method of generating functions. Let

$$F(x)=\sum_{i=0}^{\infty}\mu_i x^i$$

Then

$$xF(x)=\sum_{i=1}^{\infty}\mu_{i-1}x^i \qquad \text{and} \qquad x^2F(x)=\sum_{i=2}^{\infty}\mu_{i-2}x^i$$

Therefore

$$F(x)-\lambda xF(x)+x^2F(x)=\mu_0+\mu_1x-\lambda\mu_0x+\sum_{i=2}^{\infty}(\mu_i-\lambda\mu_{i-1}+\mu_{i-2})\,x^i$$
$$=1+\lambda x-\lambda x+\sum 0\,x^i=1$$

Thus

$$\begin{aligned}F(x)&=\frac{1}{1-\lambda x+x^2}=\frac{1}{1-\{x(\lambda-x)\}}\\&=\sum_{i=0}^{\infty}x^i(\lambda-x)^i=\sum_{i=0}^{\infty}x^i\sum_{j=0}^{i}\binom{i}{j}(-1)^jx^j\lambda^{i-j}\\&=\sum_{i=0}^{\infty}\sum_{j=0}^{i}\binom{i}{j}(-1)^jx^{i+j}\lambda^{i-j}\\&=\sum_{i=0}^{\infty}\sum_{k=i}^{2i}\binom{i}{k-i}x^k\lambda^{2i-k}(-1)^{k-i}\\&=\sum_{k=0}^{\infty}\sum_{i=\lceil k/2\rceil}^{k}(-1)^{k-i}\binom{i}{k-i}\lambda^{2i-k}x^k\end{aligned}$$

where the binomial coefficients are to be interpreted in $GF(q)$ and $\lceil x\rceil$ is the least integer $\geqslant x$. Consequently,

$$\mu_N(\lambda)=\sum_{i=\lceil N/2\rceil}^{N}(-1)^{N-i}\binom{i}{N-i}\lambda^{2i-N}$$

We now examine the determination of $\Phi(\lambda)$ from $\mu_N(\lambda)$. The nilpotent case ($\gamma=0$) can be readily recognized by $\mu_N(\lambda)=\lambda^N$. Otherwise, $\mu(\lambda)\mid\Phi(\lambda)$, giving $\Phi(\lambda)=\lambda^\tau-\lambda^{\tau+\gamma}=\mu(\lambda)(a_0+a_1\lambda+\cdots+a_k\lambda^k)$ with $a_0\neq0$ by the minimality of τ. Consequently, the first (lowest degree) nonzero term of $\mu(\lambda)$ is of degree τ. Multiplication of $\mu(\lambda)$ by λ^i is equivalent to "shifting the components of the coefficient vector of $\mu(\lambda)$ i places to the right." Thus, i is chosen so that the lowest degree term of $\lambda^i\mu(\lambda)$ is of the same degree as the second lowest degree term in the product $\mu(\lambda)(a_0+\cdots+a_{i-1}\lambda^{i-1})$. The coefficient a_i is then chosen to annihilate that term and the process is repeated until only two terms of the form of $\Phi(\lambda)$ remain. With coefficients

Table I. Transient and Cycle Lengths

Number of cells N	Transient length τ	Cycle length γ
2	0	2
3	3	0
4	0	6
5	1	4
6	0	14
7	7	0
8	0	14
9	1	12
10	0	62
11	3	8
12	0	126
13	1	28
14	0	30
15	15	0
16	0	30
17	1	28
18	0	1022
19	3	24
20	0	126
21	1	124
22	0	4094
23	7	16
24	0	2046
25	1	252
26	0	1022
27	3	56
28	0	32766
29	1	60
30	0	62
31	31	0
32	0	62
33	1	60
34	0	8190
35	3	56
36	0	174762
37	1	2044
38	0	8190
39	7	48
40	0	2046

in a Galois field, the argument for the existence of $\Phi(\lambda)$ shows that this process must terminate. (This fact is also evident from the theory of finite fields, as we shall see below.) This "shift and annihilate" algorithm is presented more formally in Appendix A. Output from a program based on this algorithm is given in Table I over $GF(2)$ for N up to 40. A straightforward modification allows the determination of cases for which $\mu(\lambda) \mid \lambda^{\tau} + \lambda^{\tau+h}$. In this situation $A^{\tau+h}X = -A^{\tau}X$. If the state values [components of X in $GF(q)$] have been graphically represented by colors, then this occurrence would be a "color reversal," i.e., the color of the ith cell at the $\tau+h$ generation is the "negative" of its color at generation τ, giving $\gamma = 2h$. Similarly, one can search for other color permutations, i.e., $A^{\tau+k}X = \alpha A^{\tau}X$, $\alpha \in GF(q)$.

It is desirable to be able to compute τ_N and γ_N easily. Certainly, the algorithm discussed above represents an improvement over the direct exercise of the CA rule, even if the generation to be repeated is known in advance. Below we show that τ_N is easy to obtain. Indeed, $\tau_{2m} = 0$ immediately by induction, using $\mu_0(\lambda) = 1$ and the recurrence relation. γ_N, on the other hand, appears to be quite difficult, even though interesting relationships can be developed. We proceed to examine an algebraic approach to determining γ which displays some of the inherent difficulties.

4. DETERMINATION OF THE MAXIMAL CYCLE LENGTH, γ_N

Instead of beginning with a general treatment, we look at several specific cases over $GF(2)$. Focusing on $\gamma_8 = 14$ from Table I, we see that $\mu_8(\lambda) = 1 + \lambda^4 + \lambda^6 + \lambda^8 = (1 + \lambda^2 + \lambda^3 + \lambda^4)^2$. [Over $GF(p^k)$, $(q(\lambda))^p = q(\lambda^p)$.] Because there are an even number of terms in the latter, it is clear that it is divisible by $1 + \lambda$. Thus $(1+\lambda)^2 (1 + \lambda + \lambda^3)^2$ is the factorization of $\mu_8(\lambda)$ into irreducible polynomials. We recall several useful facts about polynomials over finite fields. For a polynomial $f(\lambda)$ over a finite field such that $f(0) \neq 0$, the *order of* $f(\lambda)$, ord f, is the smallest natural number e such that $f(\lambda) \mid (\lambda^e - 1)$. If $f(0) = 0$ and $f(\lambda) = \lambda^h g(\lambda)$, with $h \in \mathbb{N}$ and $g(0) \neq 0$, then ord f is defined to be ord g. Consequently, $\gamma_N = \operatorname{ord} \mu_N(\lambda)$. It can be easily shown that if $f(\lambda)$ is irreducible of degree m over $GF(q)$, then ord f equals e if and only if every root of $f(\lambda) = 0$ has period e. Moreover, $e \mid q^m - 1$. Such a polynomial is called *primitive* if it has a root in $GF(q^m)$ of order $q^m - 1$. The multiplicative group of $GF(q^m)$ is a cyclic group generated by a primitive root, the existence of which is guaranteed. Evidently, the order of a primitive polynomial is maximal, i.e., $q^m - 1$.

Returning to the case $N = 8$, we see that the factor $1 + \lambda + \lambda^3$ has order 7, because $2^3 - 1 = 7$ is prime. Over a field of characteristic p, it is straightforward to show that $\operatorname{ord} f(X)^m = m \operatorname{ord} f(X)$ for m a power of p. Consequently, $\operatorname{ord}(1+\lambda+\lambda^3)^2 = 14$. Because $(1+\lambda)^2 \mid 1+\lambda^{14}$, $\mu(\lambda) \mid 1+\lambda^{14}$. Thus 14 is the smallest exponent for which this divisibility condition holds. It follows that $\Phi_8(\lambda) = 1+\lambda^{14}$ and $\gamma_8 = 14$, as previously observed.

We pursue this approach further by examining the case

$$\mu_{12}(\lambda) = 1+\lambda^2+\lambda^8+\lambda^{10}+\lambda^{12} = (1+\lambda+\lambda^4+\lambda^5+\lambda^6)^2 \quad \text{with } \gamma_{12} = 126$$

Factoring a polynomial into a product of irreducible polynomials can be accomplished via Berlekamp's algorithm.[4] Incorporating such algorithms, symbolic computing products, such as Mathematica, provide the facility for rapid factorization.[5] In addition, a table of irreducible polynomials over $GF(2)$ and their orders through degree 19 is available.[6] In this case, $1+\lambda+\lambda^4+\lambda^5+\lambda^6$ is irreducible. Its order must be a divisor of $2^6-1=63$. From the table, we find that it has order 63, giving $\gamma_{12} = \operatorname{ord} \mu_{12} = 2\cdot 63 = 126$.

For N even, equal to $2m$, the recurrence relation or the explicit form for $\mu_N(\lambda)$ over $GF(2)$ shows that

$$\mu_{2m} = \sum_{i=0}^{m} a_i \lambda^{2i} = [\tilde{\mu}_{2m}(\lambda)]^2$$

where

$$\tilde{\mu}_{2m}(\lambda) = \sum_{i=0}^{m} a_i \lambda^i, \qquad a_i \in GF(2), \quad a_0 = 1$$

Thus $\gamma_{2m} \leqslant 2(2^m-1)$. This bound will occur when $\tilde{\mu}_{2m}(\lambda)$ is irreducible and primitive. In this sense, the period for the case $N=12$ is maximal. Through 40, the even values of N having maximal period are 2, 4, 6, 10, 12, 18, 22, and 28.

One final example for N even will suffice, namely, the factorization of

$$\begin{aligned}\mu_{38}(\lambda) = [&(1+\lambda)(1+\lambda+\lambda^4+\lambda^5+\lambda^6) \\ &\times(1+\lambda+\lambda^3+\lambda^4+\lambda^5+\lambda^6+\lambda^7+\lambda^8+\lambda^{11}+\lambda^{12})]^2\end{aligned}$$

The orders to which the factors in brackets belong are 1, 63, and 1365, respectively. By their irreducibility, these polynomials have no common roots. Because $\lambda^m - 1 \mid \lambda^n - 1$ if and only if $m \mid n$, the smallest value of k for which $1+\lambda^k$ is divisible by the product of these factors is 4095, the least common multiple of 1, 63, and 1365. Consequently, $\mu_{38}(\lambda)$ first divides $1+\lambda^{2\cdot 4095}$, i.e., $\gamma_{38} = \operatorname{ord} \mu_{38} = 8190$, which agrees with the tabulated value.

Cases with N odd can be treated similarly by first writing $\mu_N(\lambda)$ as $\lambda^\tau[\tilde{\mu}_N(\lambda)]^2$, where $\tilde{\mu}_N(\lambda) = \sum_{i=0}^{g} a_i\lambda^i$, $g = (N-\tau)/2$, $a_0 \neq 0$. For example, $\mu_9(\lambda) = \lambda(1 + \lambda^4 + \lambda^8) = \lambda(1 + \lambda + \lambda^2)^4$. Consequently, $\gamma_9 = 4 \cdot \mathrm{ord}(1 + \lambda + \lambda^2) = 4 \cdot 3 = 12$. Similarly, $\mu_{27}(\lambda) = \lambda^3(1 + \lambda^{16} + \lambda^{24}) = \lambda^3(1 + \lambda^2 + \lambda^3)^8$, giving $\gamma_{27} = 8 \cdot 7 = 56$.

For the case of N odd, we show below a stronger relation between γ_N and τ_N, but for now we state the following result, which summarizes the findings for N odd or even:

For given N over $GF(2)$ [or $GF(2^k)$], the minimal polynomial $\mu_N(\lambda)$ can be written as $\lambda^\tau[\tilde{\mu}_N(\lambda)]^2$, where $\tilde{\mu}_N(\lambda) = \sum_{i=0}^{g} a_i \lambda^i$, $g = (N - \tau_N)/2$, and $a_0 \neq 0$. If $\tilde{\mu}_N(\lambda) = f_1(\lambda) \cdots f_s(\lambda)$ with these factors pairwise relatively prime, then γ_N is twice the least common multiple of ord $f_1, \ldots,$ ord f_s. In particular, $\gamma_N \leqslant 2(2^g - 1)$.

5. DETERMINATION OF THE MAXIMAL TRANSIENT LENGTH τ_N

We can discover more about γ_N and τ_N by exploiting the fundamental recurrence relation more fully. Note that $\mu_N = \lambda\mu_{N-1} - \mu_{N-2}$ can be written

$$\mu_N = \mu_1 \mu_{N-1} - \mu_0 \mu_{N-2}$$

which prompts us to suspect the general form

$$\mu_N = \mu_i \mu_{N-i} - \mu_{i-1} \mu_{N-i-1}$$

This relationship is easily proved by induction:

$$\begin{aligned} \mu_N &= \lambda\mu_{N-1} - \mu_{N-2} \\ &= \lambda(\mu_i \mu_{N-1-i} - \mu_{i-1}\mu_{N-2-i}) - (\mu_i \mu_{N-2-i} - \mu_{i-1}\mu_{N-3-i}) \\ &= \mu_i(\lambda\mu_{N-1-i} - \mu_{N-2-i}) - \mu_{i-1}(\lambda\mu_{N-2-i} - \mu_{N-3-i}) \\ &= \mu_i \mu_{N-i} - \mu_{i-1}\mu_{N-i-1} \end{aligned}$$

For the remainder of this section, let us suppose that the underlying field is of characteristic 2. Then

$$\begin{aligned} \mu_{2N+1} &= \mu_N \mu_{N+1} + \mu_{N-1}\mu_N \\ &= \mu_N(\lambda\mu_N + \mu_{N-1}) + \mu_{N-1}\mu_N \\ &= \mu_N(\lambda\mu_N + \mu_{N-1} + \mu_{N-1}) \end{aligned}$$

or

$$\mu_{2N+1} = \lambda\mu_N^2$$

Let the binary expansion of N be $\sum_{i=0}^{k} a_i 2^i$, where $k = \lfloor \log_2 N \rfloor$ and $a_i = 0$ or 1. Define $l = l_N$ as $\min\{i \mid a_i = 0\}$, where $a_{k+1} = 0$. Then we claim

$$\tau = \tau_N = 2^l - 1$$

If N is even, then l equals zero, giving $\tau = 0$ as required. If N is odd, we can repeatedly apply $\mu_N = \lambda \mu^2_{\lfloor N/2 \rfloor}$ to obtain

$$\begin{aligned} \mu_N &= \lambda(\lambda(\cdots(\lambda\mu_m)^2 \cdots)^2 \\ &= \lambda^t \mu_m^{2^l} \end{aligned}$$

where $t = 1 + 2 + \cdots + 2^{l-1}$ and $m = (N - t)/2^l$ is even. The constant term of μ_m is one, thus identifying $\tau_N = t = 2^l - 1$ and $\gamma_N = 2^l \gamma_m$.

We illustrate this result with several examples. First, $l_N = k + 1$ if and only if $N = 2^{k+1} - 1$, in which case $\tau_N = 2^l - 1 = 2^{k+1} - 1 = N$. In other words, $N = 2^{k+1} - 1$ if and only if $\mu_N = \lambda^N$. Equivalently, for any initial configuration, its Nth successor X_N equals the quiescent configuration $0 = (0, \ldots, 0)$ and, for some initial configuration, $X_{N-1} \neq 0$.

Next consider $N = 87 = (1010111)_2$, for which $l = 3$, $\tau = 2^3 - 1 = 7$, and $m = (87 - 7)/2^3 = 10$. Thus

$$\begin{aligned} 87 = (1010111)_2 &= (1010)_2 \times 2^3 + (111)_2 \\ &= 10 \times 2^3 + 7 \end{aligned}$$

Indeed, in general,

$$\begin{aligned} N &= (\overset{\substack{k\text{th} \\ \downarrow}}{1} \cdots \overset{\substack{l\text{th} \\ \downarrow}}{0} \underbrace{1 \cdots \overset{\substack{\text{0th-order bit} \\ \downarrow}}{1}}_{\text{all 1's}})_2 \\ &= m \cdot 2^l + \tau \end{aligned}$$

where $\tau = \tau_N$ and $\gamma_N = 2^l \gamma_m$.

6. SOME SPECIAL CASES

For N even, equal to $2m$, over a field of characteristic two,

$$\begin{aligned} \mu_N = \mu_{2m} &= \mu_m \mu_m + \mu_{m-1} \mu_{m-1} \\ &= (\mu_m + \mu_{m-1})^2 \end{aligned}$$

which identifies $\tilde{\mu}_N$ [i.e., $\mu_N = (\tilde{\mu}_N)^2$, as discussed earlier] as $\mu_m + \mu_{m-1}$. As we know, $\tilde{\mu}_N$ may or may not be further reducible (viz. $N = 12$) and consequently no general factorization can result from this approach. A fairly quick route to obtaining μ_N is provided, however, e.g.,

$$\begin{aligned}\mu_{20} &= (\mu_{10}+\mu_9)^2\\ &= [\mu_5^2+(1+\lambda)\,\mu_4^2]^2\\ &= [(1+\lambda+\lambda^2)\,\mu_2^4+(1+\lambda)\,\mu_1^4]^2\\ &= [(1+\lambda+\lambda^2)(1+\lambda^8)+(1+\lambda)\,\lambda^4]^2\\ &= (1+\lambda+\lambda^2+\lambda^4+\lambda^5+\lambda^8+\lambda^9+\lambda^{10})^2\end{aligned}$$

In certain instances, however, we can determine γ_N using this approach. If $N=2^k$, we claim $\mu_N(\mu_N+\lambda^N)=(1+\lambda^{N-1})^2=1+\lambda^{2(N-1)}$. We proceed by induction on k with

$$\mu_N(\mu_N+\lambda^N)=(\mu_{N/2}^2+\lambda^{N-2})(\mu_{N/2}^2+\lambda^{N-2}+\lambda^N)$$

having used

$$\mu_N=\mu_{N/2}^2+\mu_{N/2-1}^2=\mu_{N/2}^2+(\lambda^{N/2-1})^2=\mu_{N/2}^2+\lambda^{N-2}$$

Thus

$$\begin{aligned}\mu_N(\mu_N+\lambda^N)&=\mu_{N/2}^4+\mu_{N/2}^2\lambda^N+\lambda^{2N-4}+\lambda^{2N-2}\\ &=[\mu_{N/2}(\mu_{N/2}+\lambda^{N/2})]^2+\lambda^{2N-4}+\lambda^{2N-2}\end{aligned}$$

which, upon applying the induction hypothesis to the first term, is equal to

$$\begin{aligned}(1+\lambda^{N/2-1})^4+\lambda^{2N-4}+\lambda^{2N-2}&=1+\lambda^{2N-4}+\lambda^{2N-4}+\lambda^{2N-2}\\ &=(1+\lambda^{N-1})^2\end{aligned}$$

At this point, it is clear that $\gamma_N \mid 2(N-1)$; it is in fact equal to $2(N-1)$. We have not found an "algebraic" proof, but a "geometric" one can be obtained by showing that for $e=(1,0,\ldots,0)_N$, $A^ie\neq e$ for $0<i<2(N-1)$. We omit the details here. Further, note that for $N=2^k$

$$\begin{aligned}\mu_N&=\lambda\mu_{N-1}+\mu_{N-2}=\lambda\cdot\lambda^{N-1}+\mu_{N-2}\\ &=\lambda^N+\mu_{N-2}\end{aligned}$$

Thus $\mu_N+\lambda^N=\mu_{N-2}$, giving $\mu_N\mu_{N-2}=1+\lambda^{2(N-1)}$. Again a geometric argument can be used to convert $\gamma_{N-2}\mid 2(N-1)$ to $\gamma_{N-2}=2(N-1)$.

We summarize these results in Table II. The γ values for $N=2^k-3$ and 2^K+1 are obtained using the reduction for odd degree which results in known even cases (2^{k-1} and $2^{k-1}-2$, respectively). From the table, it is clear that there exists a cluster of values for the number of cells around

Table II. Cycle Lengths Around 2^k

Number of cells N	Corresponding value of γ	γ as function of N
2^k-3	$2(2^k-2)$	$2(N+1)$
2^k-2	$2(2^k-1)$	$2(N+1)$
2^k-1	0	0
2^k	$2(2^k-1)$	$2(N-1)$
2^k+1	$2(2^k-2)$	$2(N-3)$

the powers of two for which the cycle lengths are small compared to the maximal value of $2(2^g-1)$, $g=(N-\tau_N)/2$. Additional relations can be obtained, e.g., for 2^k+3, but γ cannot be obtained in an obvious fashion for other values, e.g., 2^k+2. Indeed, while $\gamma_{16}=30$, $\gamma_{18}=1022$, the maximal value. For 2^k+2, γ is generally "large," though not necessarily maximal.

7. SUMMARY AND DISCUSSION

Although we have emphasized the determination of τ_N and γ_N, many other facets of CA behavior can be elicited by the present approach. For the rule under consideration, A is nonsingular if and only if N is even. In this case, the state diagram (the directed graph with vertices corresponding to configurations and directed edges, to the successor relationship) consists of pure cycles. The structure of the state diagram can be obtained from the knowledge of the elementary divisors of A (in our case the irreducible factors of μ) and their orders. For example, for $N=8$, $\mu(\lambda)=(\lambda+1)^2(\lambda^3+\lambda+1)^2$ implies the existence of seventeen 14-cycles, two 7-cycles, one 2-cycle, and two fixed points (one being the zero vector). The matrix which brings A into its classical canonical form[7] can be used to obtain an explicit configuration labeling of the state diagram.

For N odd, the nullity of A is one. Thus a configuration has either no predecessor or two predecessors. The state diagram then consists of cycles with binary trees of height τ_N rooted at each vertex of a cycle. Again, the elementary divisors of A determine the state diagram.

We emphasize that for any linear rule (local transition plus boundary) which updates according to $X_{t+1}=AX_t$ the minimal polynomial of the update matrix plays an important role. The determination of γ_N as the order of μ_N (and by the algorithm of Appendix A) and τ_N as the degree of the first nonzero term of μ_N are general results.

We now summarize our findings for the CA/B updated by Rule 90 with null boundary conditions. Two facts are responsible for the results particular to this rule. First, the minimal polynomial of A is its characteristic polynomial, independent of the characteristic of the underlying field. This felicitous situation may not be the case for other rules, e.g., for periodic boundary conditions, $\mu_4(\lambda)$ is λ^2 if the characteristic is 2 and is $\lambda^3 - 4\lambda$ in all other cases. Second, the structure of A provides a valuable recurrence relation for μ_N which is valid over any finite field. Thus our results are derived from the special structure of A for the rule being studied, as can results for the periodic case be obtained from the properties of circulant matrices. However, quite generally A will possess banded or nearly banded structure because it is based on a local transition rule, with boundary rules affecting at most a few of the first and last rows. Moreover, because a "near copy" of A_{N-1} will be a submatrix of A_N, the search for recurrence relations between the characteristic and minimal polynomials for N, $N-1$,..., may provide a fruitful line of attack.

Further results were then obtained by exploiting the recurrence relation in a particular finite field or one of given characteristic. Such an approach allowed our determination of γ_N for special values of N over fields of characteristic two. τ_N was found to be $2^j - 1$ where 2^j divides $N+1$ but 2^{j+1} does not. We speculate that over GF(p^k), τ_N for our rule is in general related to the divisibility of $N+1$ by p, e.g., we conjecture that over $GF(3)$, $\tau_N = 3^j$ (N odd) where 3^j divides $N+1$ but 3^{j+1} does not.

As suggested by the above example, there are many possible extensions and generalizations of the subject at hand. We delineate only a few. For our usual rule, is it possible to obtain a simple formula or rule for γ_N? How do τ_N and γ_N depend on the underlying field? These questions are, of course, of interest for other local transition and boundary rules as well. Our results depend on the underlying algebraic object being a *field*, which thus provides the tools of linear algebra. What results can be obtained in more general cases and by what means, e.g., when the whole is carried over Z_n?

The present approach can be extended to two- or higher-dimensional CA, in some cases rather easily. For the update rule

$$x_{i,j}^{(t+1)} = x_{i-1,j}^{(t)} + x_{i+1,j}^{(t)} + x_{i,j-1}^{(t)} + x_{i,j+1}^{(t)}$$

with null boundary conditions, if X_t represents the matrix of cell values, then it is updated by $X_{t+1} = X_t A + A X_t$, where A is our usual one-dimensional update matrix. Using this representation, one-dimensional results have immediate application in the two-dimensional case. We shall not pursue this connection further here. Finally, a two-dimensional array of cells can, by suitable specification of neighborhood and boundary rules, be given a variety of "boundary treatments," i.e., identified so as to have the cells on a cylinder, Möbius strip, torus, Klein bottle, or projective plane. It would be interesting to investigate the role of their topology in the behavior of such cellular automata.

APPENDIX A. ALGORITHM TO DETERMINE $\Phi(\lambda)$ FOR GIVEN NUMBER OF CELLS

1. For given number of cells N, compute $\mu(\lambda)=\mu_N(\lambda)$; $D_\mu=$ degree $\mu(\lambda)$. (Here $D_\mu=N$.)

2. Determine the degree τ of the term of $\mu(\lambda)$ of lowest degree, with its nonzero coefficient denoted α.

3. If $\tau=D_\mu$, print $(\tau, 0)$ and stop; otherwise proceed.

4. Replace $\mu(\lambda)$ by $\alpha^{-1}\mu(\lambda)$. (The lowest degree term will now have coefficient 1.)

5. $P(\lambda)\leftarrow\mu(\lambda)$; $D_P\leftarrow D_\mu$.

6. Determine the degree D of the second lowest degree term of $P(\lambda)$ having nonzero coefficient c.

7. If $D_P=D$ and $c=-1$, exit loop (step 10).

8. $P(\lambda)\leftarrow P(\lambda)-c\lambda^{D-\tau}\mu(\lambda)$; $D_P\leftarrow D_\mu+D-\tau$.

9. Repeat from step 6.

10. Print $(\tau, \gamma=D_P-\tau)$.

11. Stop.

APPENDIX B. NOTATION

A	$=A_N$, the update matrix such that $X_{t+1}=AX_t$ (Section 1)
$c(X_0)$	Length (period) of the cycle reached from the initial configuration X_0 (Section 2)
$GF(q)$	The finite field of q elements where $q=p^k$ for some prime p; $GF(p)$ is Z_p
$\mathbb{N}$	The natural numbers $\{1, 2, 3, \ldots\}$
$\mathbb{N}_0$	The nonnegative integers $\{0, 1, 2, \ldots\}$
$t(X_0)$	The transient length for the initial configuration X_0 (Section 2)
$x_i^{(j)}$	The value of the ith cell after j updates of the initial configuration; equivalently, the ith component of the configuration X_j (Section 1)
X_j	The N-vector of cell values after j updates of the initial configuration X_0 (Section 1)
Z_n	The integers modulo n

γ $=\gamma_N$, the maximal cycle length over all initial configurations (Section 2)

λ Indeterminate of the ring of polynomials over $GF(q)$

μ $=\mu_N(\lambda)$, the minimal polynomial of A_N (Section 3)

τ $=\tau_N$, the maximal transient length over all initial configurations (Section 2)

Φ $=\Phi(\lambda)$, the transient and cycle polynomial (Section 3)

$\lfloor x \rfloor$ The greatest integer less than or equal to x

$\lceil x \rceil$ The least integer greater than or equal to x

ACKNOWLEDGMENTS

We thank Prof. L. E. Scriven for useful discussions in the course of this work. During the revision of this paper, our colleague and friend, A. E. Cerkanowicz, died on May 28, 1993. He will be greatly missed by us and all who knew him.

REFERENCES

1. S. Wolfram, Statistical mechanics of cellular automata, *Rev. Mod. Phys.* **55**:601 (1983).
2. P. Guan and Y. He, Exact results for deterministic cellular automata with additive rules, *J. Stat. Phys.* **43**:463 (1986).
3. O. Martin, A. Odlyzko, and S. Wolfram, Algebraic properties of cellular automata, *Commun. Math. Phys.* **93**:219 (1984).
4. R. Lidl and G. Pilz, *Applied Abstract Algebra* (Springer-Verlag, New York, 1984), p. 172.
5. S. Wolfram, *Mathematica®, A System for Doing Mathematics by Computer*, 2nd ed. (Addison-Wesley, Redwood City, California, 1991), p. 602.
6. R. W. Marsh, Table of Irreducible Polynomials over GF(2) Through Degree 19, U. S. Department of Commerce, PB161693 (1957).
7. N. Jacobsen, *Lectures in Abstract Algebra, Vol. II—Linear Algebra* (Van Nostrand, New York, 1953), p. 73.