Complex Systems **1** (1987) 107–130

Power Spectra of Regular Languages and Cellular Automata

Wentian Li
Center for Complex Systems Research, University of Illinois at Urbana-Champaign, 508 South Sixth Street, Champaign, IL 61820
and
Physics Department, Columbia University, New York, NY 10027, USA

Abstract. The spatial structure of attractors produced by many one-dimensional cellular automata can be described by regular languages. This paper gives simulations and analytical results for the power spectra of such attractors. The power spectra are Fourier transforms of autocorrelation functions which are exponentially damped (sometimes with oscillations). The characteristic length scale is related to non-trivial eigenvalues of the arc-to-arc transition matrix in the regular language graph.

1. Introduction

Spectra or Fourier transforms are some of the most frequently used tools in physics to reveal regularity in seemingly irregular behaviour. Recent examples in experimental physics include such events as the discovery of the so-called quasicrystals in $MnAl_6$ by electron diffraction [1] which in effect is a spatial Fourier transform, and the observation of the period-doubling route to chaos in Rayleigh-Bernard cells found by looking at the temporal spectra of the temperature oscillation [2]. It is also an indispensable method in numerical studies of dynamical systems, where nonlinearities generally prevent one from getting exact solutions.

A cellular automaton (CA) is a spatially-extended dynamical system with discrete states on each site of a lattice, and dynamics in discrete time steps. For elementary CA (with 2 states, and nearest neighbour updating rule in 1-dimension) [6] a typical random initial configuration will be attracted to an "equilibrium" set which is called an *attractor*. [1]

[1] One uses the term *limiting set* for limiting configurations obtained from all *any* initial configurations, and *attractor* for those from *typical random* initial configurations. The attractor is thus a stable limiting set.

Generally it does not take very long for a cellular automaton to settle down onto a attractor. An attractor can either be a time-invariant configuration (possibly after a spatial shift), a temporally periodic cycle, or it may exhibit chaotic dynamics. For most CA rules, the properties of the attractor, such as the density or the spatial spectrum, are not sensitive to the initial configuration [12]. In addition most CA rules (except the rules with infinitely long transient times) have time-invariant statistical quantities which characterize their attractors. These facts ensure that the study of an attractor at any instant after it has reached the "equilibrium" state beginning in a typical random initial configuration (even if one has a poor-quality random number generator) will be a study of properties intrinsic to the rule itself. [2]

I have carried out a numerical discrete Fourier transformation (DFT) of the spatial configurations obtained after many large time steps for each of the elementary CA rules. Out of $2^8 = 256$ elementary rules 88 are independent. Of these 8 yield the all 0's configuration from every initial condition. Figure 1 shows power spectra for the remaining 80 rules. The number of sites used in the computations was $2^{12} = 4096$. The number of time steps before a measurement was made varied from 15 to more than 100. Each spectrum is an average of 18 runs. The fluctuations seen in the figure are due to statistics. Only half of the Fourier components are shown since the other half are redundant (N real data points but N complex Fourier components). The spatial-temporal patterns generated by the rules shown in figure 1 are given for example in [7].

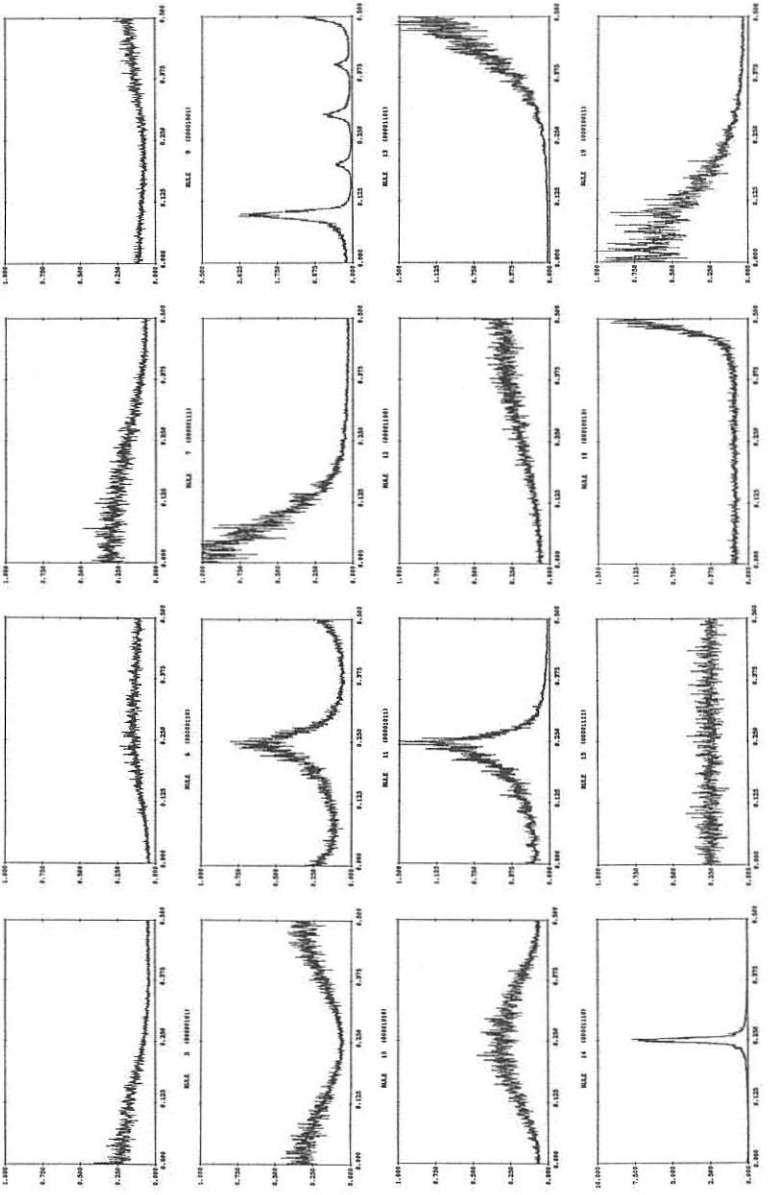
The spectra in figure show quite a variety of shapes. Some are white noise (constant spectrum), others Brownian noise (with $1/\omega^2$ tails). Many of the spectra show discrete peaks. Some spectra follow continuous distributions which are not fit by any $1/\omega^n$ laws. In this paper, I will show that many of these spectra, particularly those with peaks, can be calculated exactly.

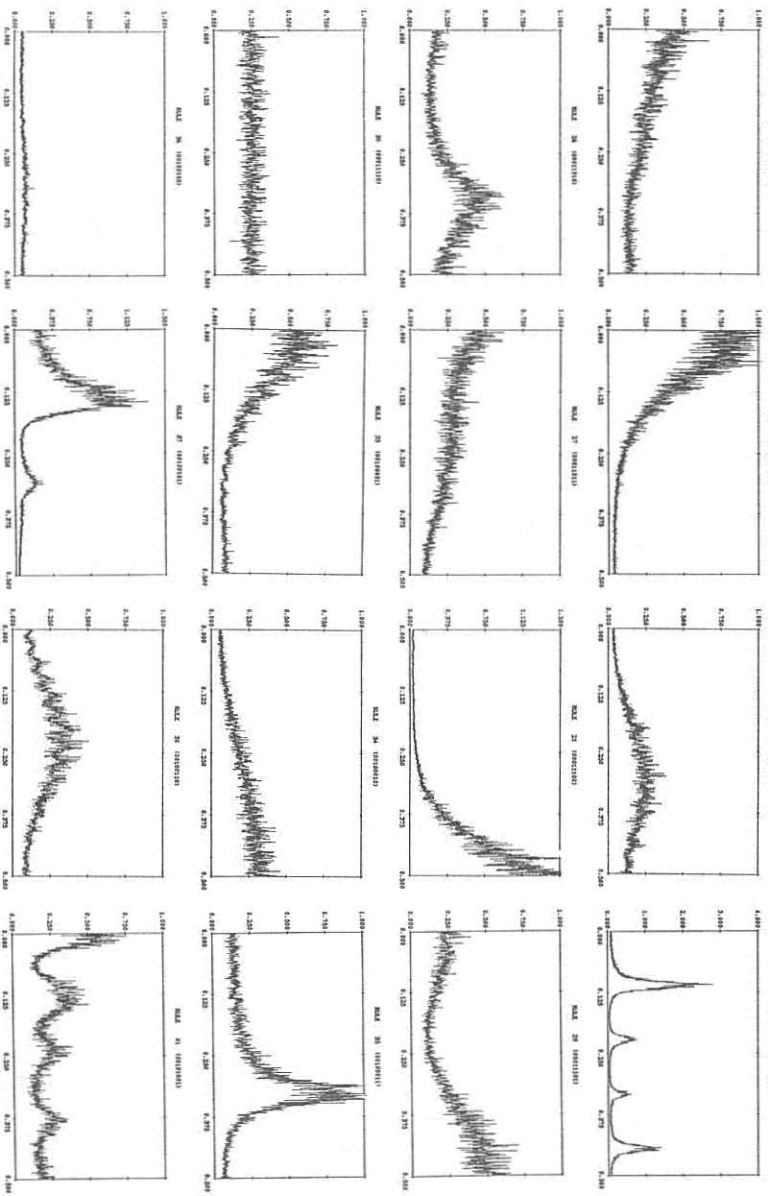
A key concept used in the paper is that of a class of formal languages known as regular languages. Following the standard terminology, a formal language is a set of strings satisfying certain properties. A string or a word is a finite sequence of symbols chosen from a finite alphabet of letters, numbers, etc. There are several ways to describe a regular language: *(i)* by regular expressions, *(ii)* as the language accepted by a finite automaton, and *(iii)* by a regular grammar [3]. Here we shall represent regular languages by directed graphs, called a regular language graphs (RLG) [3], having a finite number of nodes and arcs (edges with arrows). When a letter is attached to each arc, any path in this graph will generate a word of the regular language.

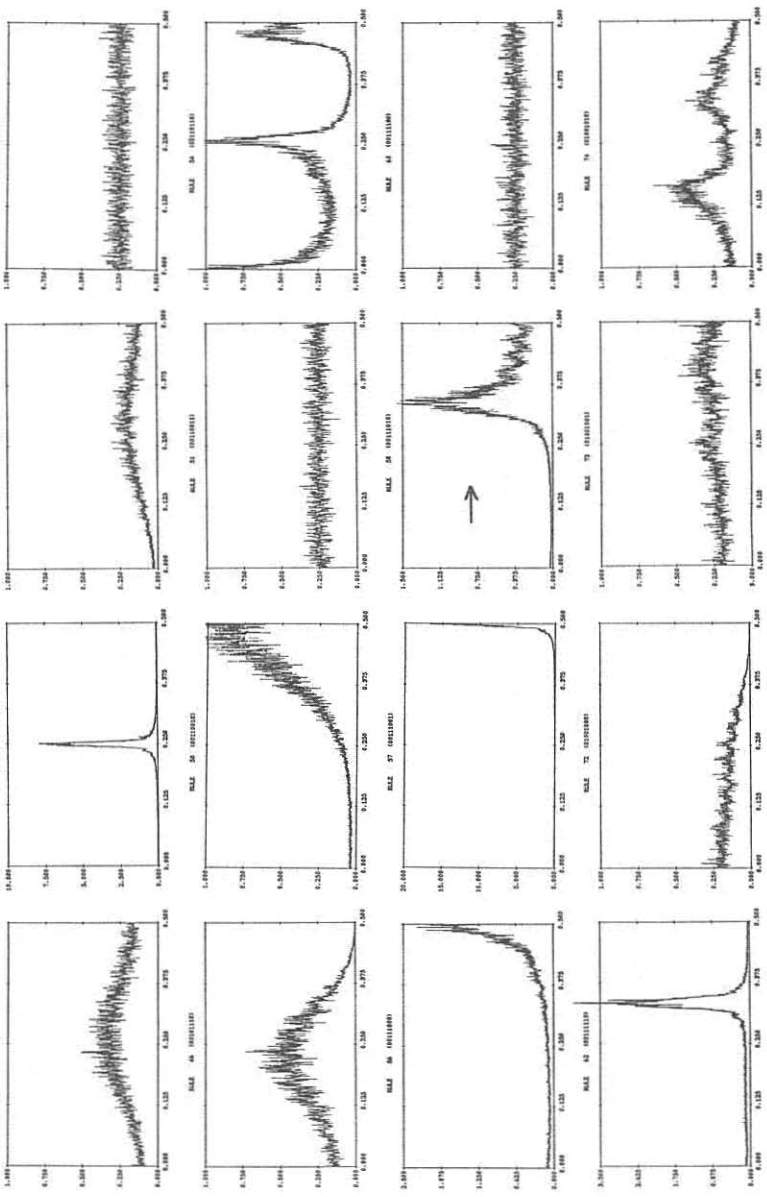
Before defining the "spectrum of a regular language", several comments should be made:

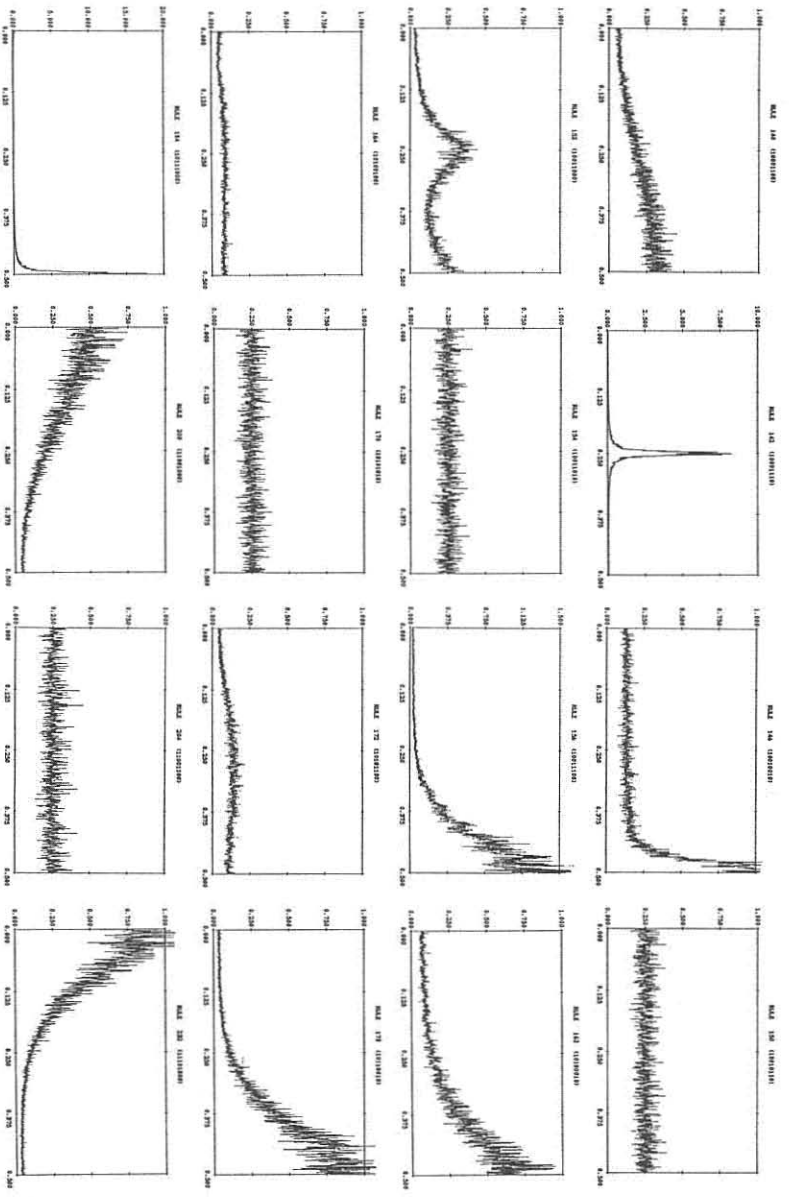
[2] The spatial spectrum of the two-cycle configuration at even time steps may differ from that at odd steps (e.g. rule 1). But it is not a problem for most of the rules.

[3] Also called transition diagram of the finite automaton.









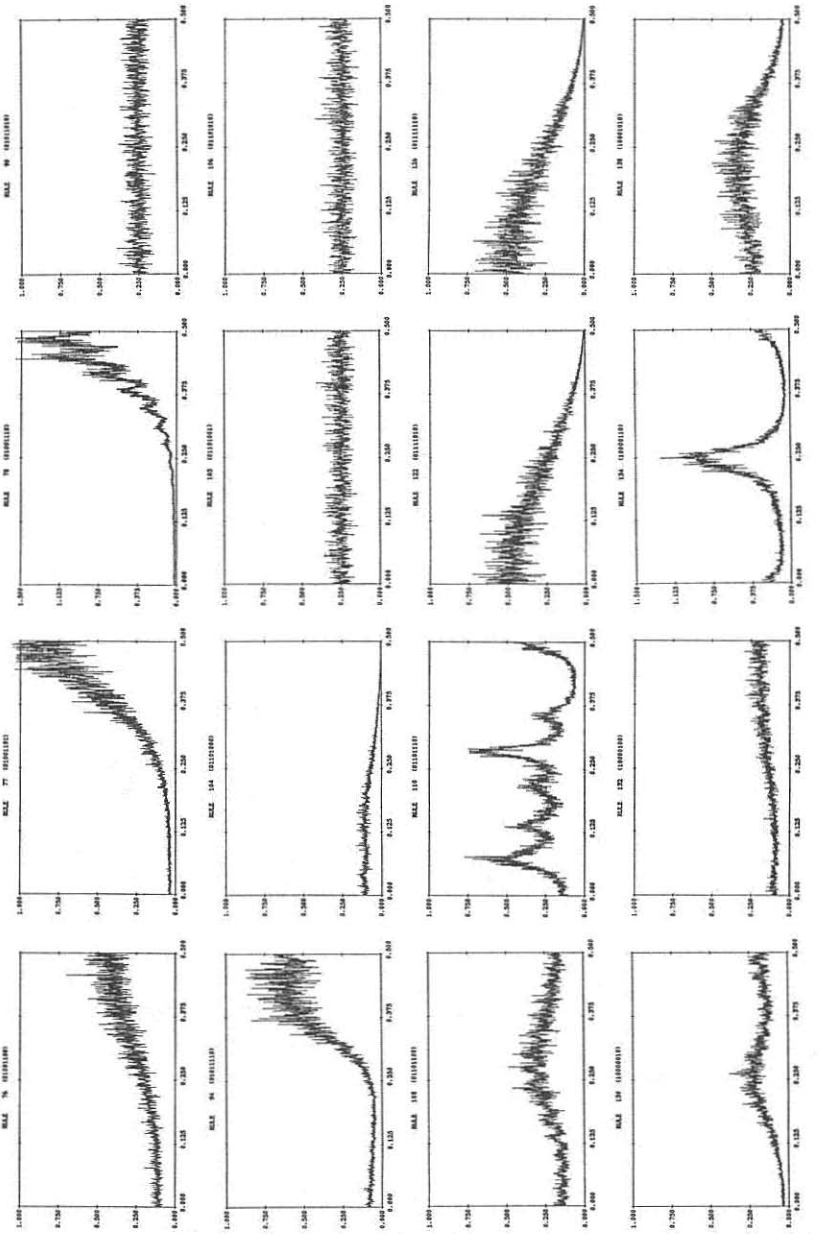


Figure 1: The spatial power spectra for all the independent elementary CA rules (except the trivial null rules). The x-axis is the frequency ν/N, y-axis is power $|A(\nu)|^2 N$ (see Appendix). Analytical results for the spectrum of rule 58 (indicated by an arrow) will be given below.

(1) Numbers will be used as the symbols. In the case of elementary CA rules they are the numbers {0,1}.

(2) Transient parts in the graph will not be considered here. (see Figure 2).

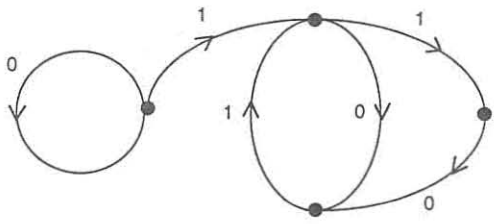


Figure 2: A regular language graph (RLG) with non-stationary (transient) part. See Figure 8 for the graph without a transient part.

(3) At each node the probabilities for making transitions to other nodes are given. Such "weighted" RLG were not considered in [3] and [8].

(4) Only words of infinite length will be considered.

The regular language considered above can be viewed as a stationary Markov chain. In this context, the limitations have the following consequences. *(2)* makes the process stationary. *(3)* allows a Markov transition matrix to be defined. *(4)* makes fluctuations negligible. We will use the notation $P_{n\times n}$, where n is the number of arcs, to denote the symbol-to-symbol (or arc-to-arc) transition probability matrix.

It is important to note that not every possible set of sequences can be represented as a regular language, or for that matter, as any kind of standard formal language. The finiteness of the description of the words in a regular language makes possible many computations, including the calculations of spectra given here.

From the definition of a regular language, we can immediately deduce some of the properties of its spectrum.

First of all, it is a "local language" which means any part of a word cannot be correlated with other parts far away from it. This is because the finite automaton which generates or recognizes the regular language has no internal memory. From the point of view of stochastic processes, this fact is reflected in the exponential decay of the autocorrelation function for a Markov process. (As discussed for example in [4] no eigenvalues of a Markov matrix can be larger than one, so the powers of the eigenvalues

decay exponentially). Now suppose the autocorrelation function [4]

$$R(\tau) \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)x(t+\tau)dt - \left(\lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)dt\right)^2 \tag{1.1}$$

is a normalized exponential function with correlation length τ_i:

$$R_0(\tau) = \left(\frac{1}{2\tau_i}\right) e^{-\frac{|\tau|}{\tau_i}}. \tag{1.2}$$

Then by the convolution theorem and the eveness of $R(\tau)$, the power spectrum [5]

$$\begin{aligned} S(\omega) &\equiv \lim_{T\to\infty} \left|\frac{1}{T}\int_{-\infty}^{\infty} e^{-i2\pi\omega t}x(t)dt\right|^2 \cdot T \\ &= \int_{-\infty}^{\infty} e^{-i2\pi\omega\tau}(R(\tau) + \text{const.})d\tau = 2\int_0^{\infty} cos(2\pi\omega\tau)R(\tau)d\tau \end{aligned} \tag{1.3}$$

takes on the "Lorentzian form":

$$S_0(\omega) = \frac{(1/\tau_i)^2}{(2\pi\omega)^2 + (1/\tau_i)^2} \tag{1.4}$$

with a maximum at $\omega = 0$ and half-width equal to $1/\tau_i$.

A second feature of regular languages is they can repeat substrings many times (as gauranteed by the "pumping lemma"). This leads to loop structures in the RLG, and gives peaks at nonzero frequencies in the power spectra. The simple exponential form used above does not show this phenomenon. To see it, instead, some of the eigenvalues of the Markov matrix must be negative or complex. When this happens, the autocorrelation function acquires an *oscillating* exponential form. This modulation splits the $\omega = 0$ peak into two peaks (the peak at negative frequency simply reflects the eveness of the power spectrum). (See e.g. [17]:page 108)

The plan of this paper is as follows. Section 2 treats the relation between CA attractors and regular languages. Section 3 gives the detailed procedures for calculating power spectra of regular languages. Section 4 shows several examples, and section 5 gives some further discussion and conclusion.

---

[4] Some authors distinguish the autocorrelation function from the autocovariance function by whether it is normalized; others distinguish them by whether the "DC" component is subtracted. To avoid confusion, the "autocorrelation function" will be used exclusively in this paper.

Although notations such as T, t and τ are used, the spectra to be discussed are mainly spatial Fourier transforms at fixed time.

[5] The discreteness of CA makes the discrete Fourier transformation seem more appropriate. On the other hand, the discreteness introduces periodicity into the spectra which makes some features difficult to discern. I will use the continuous Fourier transform as an *approximation* of the DFT to illustrate the features of the spectra and mention DFT only when these two differ greatly. The equivalent DFT formulas are in the Appendix. Remember that they are the exact formulas for discrete processes like CA.

2. Cellular automata attractors and regular languages

Several class of behaviour in cellular automata can be identified from computer experiments.

1. The evolution leads from a typical random initial configuration to a time-invariant configuration (perhaps after a simple spatial shift). The elementary CA rules f of this kind are (where C denotes the CA configuration, and s a shift to the right):

> Rule 4, 12, 13, 36, 44, 72, 76, 77, 78, 104, 132, 140, 164, 172, 200, 204, 232 with $f(C) = C$.
>
> Rule 2, 10, 34, 42, 46, 58, 130, 138, 162, 170 with $f(C) = s^{-1}(C)$.
>
> Rule 24, 56, 152 with $f(C) = s(C)$.
>
> Rule 57 and rule 184 can satisfy either $f(C) = s(C)$ or $f(C) = s^{-1}(C)$, depending on initial conditions. The reason for this is that there exist some particle-like substrings (also called gliders) which move either to the right or left. Pairs of such gliders annihilate when they collide. But due to the finite size of simulations, which direction of glider survives depends on which predominates in the initial condition.

2. Some rules lead to temporally periodic cycles with period equal to two. They are:

> Rule 1, 5, 19, 23, 28, 29, 33, 37, 50, 51, 108, 156, 178 with $f^2(C) = C$.
>
> Rule 3, 7, 25, 27, 35 with $f^2(C) = s(C)$.
>
> Rule 6, 38, 74, 134 with $f^2(C) = s^{-2}(C)$.
>
> Rule 9, 11, 14, 15 with $f^2(C) = s^2(C)$.
>
> Rule 142 with either $f^2(C) = s^2(C)$ or $f(C) = s^-1(C)$ depending on initial conditions.

3. Some rules evolve to cycles with longer periods.

4. Some rules do not ever enter cycles in the infinite lattice size limit.

The attractors of the first two kinds of CA rules can be characterized by simple regular languages; those of the third kind by more complicated regular languages. The chaotic attractors of the fourth class may or may not be described by regular languages.

The way to write down the RGL corresponding to the attractor for a particular CA is very simple (see [8]). Take the fixed point case, for example. Imagine a "window" with length equal to two sites, which is moved from the left side of the configuration to the right. Each such two-site window will be a node in the RLG. Now suppose this window is moved by one site. It then covers a three-site region which will become an arc in the graph. For the elementary CA rules, each block of three sites at time t will yield the value of a single site at time $t+1$, and we attach this value to

the corresponding arc in the RLG. Only some of the arcs will not violate the invariance condition (e.g. for $f(C) = C$ rules, the middle site value of the three-site block at t must be equal to the single site value at $t+1$); other arcs must be deleted. The remaining graph is then the RLG for the fixed point attractor (see Figure 3 (a) for a RLG of a fixed point attractor which characterizes one row in the spatial-temporal pattern in Figure 3 (b)).

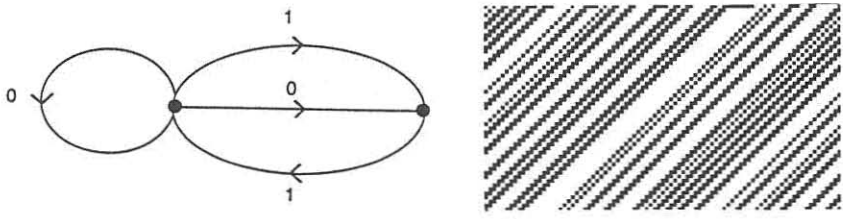


Figure 3: (a) The RLG for a fixed-point attractor of elementary CA rule 42. (b) Spatial-temporal pattern of rule 42.

For two-cycles or k-cycles ($k > 2$), only trivial extensions are needed. Instead of using a two-site window, one uses four-site window, or in general a $2k$-site windows. The maximum number of nodes and arcs are, respectively, 16 and 32 for two-cycles, 2^{2k} and 2^{2k+1} for k-cycles (see Figure 4 (a) for a RLG of a 2-cycle attractor which charaterizes one row in the spatial-temporal pattern shown in Figure 4 (b) [10]). In fact, the actual number of nodes and arcs is strictly smaller than this upper bound; otherwise the graph is equivalent to the simplest "white noise" graph (see Figure 5), corresponding to the trivial regular language in which all possible sequences can occur.

Generally simplifications may be carried out by combining equivalent nodes. The distinction between deterministic finite automata (DFA) (at every node all the symbols attached to the outgoing arcs are different) and non-deterministic finite automata (NDFA) is not important for our purposes. The conversion from NDFA to DFA is desirable only if the resulting graph has fewer arcs, but in practice even the minimum DFA can have more arcs than the equivalent NDFA. [6]

The spatial spectra of CA attractors are related to the autocorrelation functions for the configurations they contain. Having shown above that at least some such attractors can be characterized by regular languages, or, equivalently, by Markov chains, we can borrow techniques used in the theory of stochastic processes to calculate the necessary autocorrelation functions and thus the power spectra.

[6] Unlike DFA, there is no theorem about the uniqueness of minimal NDFA, and no standard algorithm to simplify a NDFA.

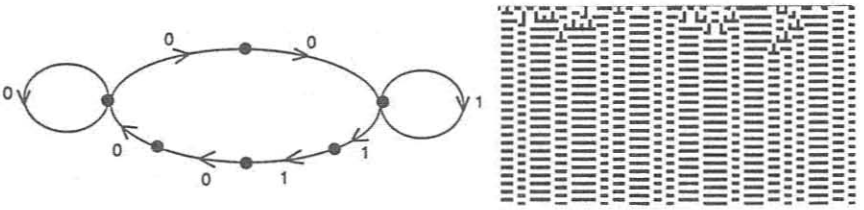


Figure 4: (a) The RLG for a 2-cycle attractor of elementary CA rule 37. (b) The spatial-temporal pattern of rule 37.

3. Calculating the power spectrum of a regular language

The procedure for calculating power spectra of regular languages is as follows.

1. Write down the arc-to-arc transition matrix $P_{n\times n}$. The sum of elements in each row is equal to one since it is the total probability for any letter to follow the letter corresponding to that row. No such meaning can be given to the sum of elements in the same column. The usual node-to-node transition matrix, or "weighted adjacency matrix", is not the one we are using here, although the two share common elements, and some common eigenvalues.

2. Calculate the τ*-th power of the matrix* P. The elements in P^τ are the τ-step transition probabilities from one arc to another. Since it is the main part of the computation, I will write the steps in more detail (see [5], [4]):

Let $\lambda_1, \lambda_2, \cdots \lambda_n$ be the eigenvalues of the matrix $P_{n\times n}$, $\mathbf{r}_j$ be any nonzero right eigenvector ($j = 1, \cdots n$) and $\tilde{\mathbf{l}}_i$ be any nonzero left eigenvector ($i = 1, \cdots n$) ($\mathbf{r}_j$ is a column vector and $\tilde{\mathbf{l}}_i$ is a row vector).

Put all the $\mathbf{r}_j$ together to define a "right" matrix

$$R = \{r_{ij}\} \equiv (\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_n) \tag{3.1}$$

and do the same for $\tilde{\mathbf{l}}_i$ to define a "left" matrix

$$L = \{l_{ij}\} \equiv \begin{pmatrix} \tilde{\mathbf{l}}_1 \\ \tilde{\mathbf{l}}_2 \\ \vdots \\ \tilde{\mathbf{l}}_i \end{pmatrix}. \tag{3.2}$$

If all the λ_i's are distinct [7], then the matrix

$$C \equiv LR \tag{3.3}$$

[7] If there are degeneracies among the λ_i, the matrix may not be diagonizable. But if it can be diagonized, then with a little care in choosing the $\mathbf{r}_j$'s and $\tilde{\mathbf{l}}_i$'s, the C will still be diagonal with nonzero diagonal elements.

is diagonal for any non-zero $\mathbf{r}_j$'s and $\tilde{\mathbf{l}}_i$'s with all $c_i \neq 0$. We may write

$$C = diag(c_1, c_2 \ldots c_n). \tag{3.4}$$

We also define

$$\Lambda \equiv diag(\lambda_1, \lambda_2 \cdots \lambda_n). \tag{3.5}$$

Then

$$\begin{aligned} P^\tau &= R\Lambda^\tau R^{-1} = R\Lambda^\tau C^{-1} L \\ &= R \begin{pmatrix} \lambda_1^\tau c_1^{-1} & & & \\ & \lambda_2^\tau c_2^{-1} & & \\ & & \ddots & \\ & & & \lambda_n^\tau c_n^{-1} \end{pmatrix} L. \end{aligned} \tag{3.6}$$

Note that although we could simply calculate R^{-1} instead of left eigenvectors, the latter is somewhat easier to handle, especially for a large matrix.

3. Calculate the autocorrelation function $R(\tau)$. If the alphabet set is {0,1}, then by eqn. (*A*.3)

$$\begin{aligned} R(\tau) &= (\sum_{j=1}^{n}\sum_{i=1}^{n} b_i (P^\tau)_{ij} x_i x_j) - (\sum_{i=1}^{n} b_i x_i)^2 \\ &= \sum_{i'j'(x_{i'}=1, x_{j'}=1)} b_{i'} (P^\tau)_{i'j'} - (\sum_{i'(x_{i'}=1)} b_{i'})^2 \end{aligned} \tag{3.7}$$

where $\{b_i\}$ is the equilibrium probability distribution for each arc; the $\{x_i\}$ are the values attached to the arcs i. Obviously only nonzero arc values enter the expression.

In the present calculations, the transitions probabilities $\{p_{ij}\}$ from which the $\{b_i\}$ are determined are taken as given. Although to my knowledge there is no way to predict these parameters from the first principles, one can use a mean-field theoretic scheme (e.g. [11]) to calculate the limiting density and possibly also the $\{p_{ij}\}$.

In the case where all the eigenvalues are real (with λ_{pi} positive and λ_{ni} negative),

$$R(\tau) = \sum_i \left[a_i \lambda_{pi}^\tau + b_i (-|\lambda_{ni}|)^\tau \right]. \tag{3.8}$$

The coefficients a_i and b_i have a simple relation with the elements of the eigenvectors. Since all powers of 1 or 0 are still 1 or 0, the trivial eigenvalues ($\lambda = 0, 1$) can contribute only a constant term to $R(\tau)$.

4. The power spectrum $S(\omega)$. For the simplest case with all λ_i real, the spectrum is:

$$S(\omega) = \sum_i \left[a_i \frac{2/\tau_{pi}}{(2\pi\omega)^2 + (1/\tau_{pi})^2} + b_i \left(\frac{1/\tau_{ni}}{(2\omega+1)^2\pi^2 + (1/\tau_{ni})^2} + (\omega \to -\omega) \right) \right] \tag{3.9}$$

where $\tau_{pi} = \log(1/\lambda_{pi})$ and $\tau_{ni} = \log(1/|\lambda_{ni}|)$

In general when the eigenvalues are complex, the position of the peak is no longer at $\omega = 1/2$ as given by the above formula, but at the positions determined by the angles of the eigenvalues in the complex plane. In a case where the RLG is very complicated, the number of eigenvalues of the matrix is very large, and we expect the shape of the spectrum will also be complicated.

4. Examples

In this section several examples of the spectra of regular languages are given to illustrate the method presented in the last section. Example 1 is a simple random process which generates "white noise". Example 2 contains a two-loop in the graph which produces a peak at $\omega = 1/2$. Example 3 illustrates the relationship between the cycle structure of the RLG and the eigenvalues of the corresponding transition matrix. Example 4 has both a two-loop and a three-loop in the graph, and with continuous changes in the transition probability the peak can move between $\omega = 1/3$ and $\omega = 1/2$. Example 5 gives a simple model for "Brownian noise" in terms of a regular language.

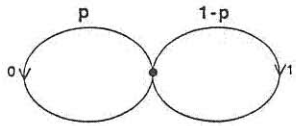


Figure 5: RLG for a biased random sequence. This regular languages characterizes the initial CA configuration used (with $p = 0.5$) and the attractors for some chaotic rules.

Example 1 (see Figure 5): The graph is a simple representation of a coin-tossing sequence with bias p. It becomes purely random when $p = 1/2$. The arc-to-arc transition matrix is ($i = 1$ is arc 0; $i = 2$ is arc 1):

$$P_{2\times 2} = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix} \tag{4.1}$$

$$\det(P - \lambda I) = (\lambda - 1)\lambda \tag{4.2}$$

The eigenvalues are

$$\lambda_1 = 1, \quad \lambda_2 = 0$$

and

$$R = \begin{pmatrix} 1 & 1-p \\ 1 & -p \end{pmatrix}, \quad L = \begin{pmatrix} p & 1-p \\ 1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \tag{4.3}$$

$$P^\tau = R \begin{pmatrix} 1^\tau & 0 \\ 0 & 0^\tau \end{pmatrix}, \quad L = \begin{cases} P & \tau \neq 0 \\ I & \tau = 0 \end{cases} \tag{4.4}$$

Defining b_1 as the stationary probability for arc 1, and $\overline{x}$ the average value for the sequence, one may write $b_1 = 1 - p = \overline{x}$, and the autocorrelation function becomes:

$$R(\tau) = b_1(P^\tau)_{22} - \overline{x}^2 = \begin{cases} 0 & \tau \neq 0 \\ p(1-p) & \tau = 0 \end{cases} \tag{4.5}$$

The power spectrum defined by $(A.2)$ then assumes a constant value

$$S(\nu) = \frac{1}{N} R(0) = \frac{p(1-p)}{N} = const. \tag{4.6}$$

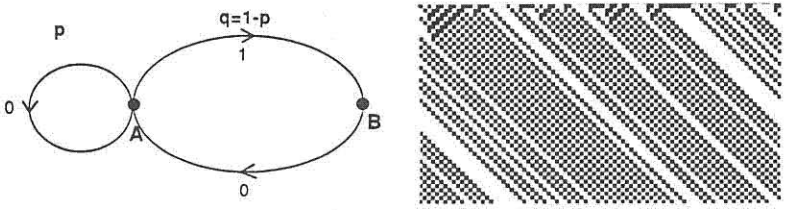


Figure 6: (a) RLG for the attractor of rule 56. (b) Spatial-temporal pattern for rule 56.

Example 2 (Figure 6): This regular language characterizes many one-dimensional CA attractors (e.g. rule 56)

Assigning arc $\overline{AB} = 1$ to $i = 1$; $\overline{BA} = 0$ to $i = 2$; $\overline{AA} = 0$ to $i = 3$, the transition matrix becomes (define $q = 1 - p$ for simplicity):

$$P_{3\times 3} = \begin{pmatrix} 0 & 1 & 0 \\ q & 0 & p \\ q & 0 & p \end{pmatrix} \tag{4.7}$$

$$\det(P - \lambda I) = -(\lambda - 1)\lambda(\lambda + q) \tag{4.8}$$

The eigenvalues are:

$$\lambda_1 = 1, \qquad \lambda_2 = 0, \qquad \lambda_3 = -q$$

$$R = \begin{pmatrix} 1 & p & 1 \\ 1 & 0 & -q \\ 1 & -q & -q \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 1 & p/q \\ 0 & 1 & -1 \\ 1 & -1/q & -1/q \end{pmatrix} \tag{4.9}$$

$$C = \begin{pmatrix} (1+q)/q & & \\ & q & \\ & & 1+q \end{pmatrix} \tag{4.10}$$

The average value $\overline{x}$ is given by $\overline{x} = q/(1+q)$, and

$$P^\tau = R \begin{pmatrix} \overline{x} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (-q)^\tau/(1+q) \end{pmatrix} L \tag{4.11}$$

$$= \begin{pmatrix} \overline{x}+(-q)^\tau/(1+q) & \overline{x}-(-q)^\tau/q(1+q) & p/q(\overline{x}+(-q)^\tau/(1+q)) \\ \overline{x}+(-q)^{\tau+1}/(1+q) & \overline{x}-(-q)^{\tau+1}/q(1+q) & p/q(\overline{x}+(-q)^\tau/(1+q)) \\ \overline{x}+(-q)^{\tau+1}/(1+q) & \overline{x}-(-q)^{\tau+1}/q(1+q) & p/q(\overline{x}+(-q)^\tau/(1+q)) \end{pmatrix} \tag{4.12}$$

It is easy to show that $b_{\overline{AB}}$, the equilibrium probability for arc $\overline{AB}$, is equal to $\overline{x}$ since it is the only arc with symbol 1, so that

$$R(\tau) = b_{\overline{AB}}(P^\tau)_{11} - \overline{x}^2 = \frac{q}{(1+q)^2}(-q)^\tau \tag{4.13}$$

The spectrum according to (1.3) is then given by

$$S(\omega) = 2\int_0^\infty cos(2\pi\omega\tau)\frac{q}{(1+q)^2}q^\tau \cos(\pi\tau)d\tau$$

$$= \log(\frac{1}{q})\frac{q}{(1+q)^2}\left[\frac{1}{(\omega-\frac{1}{2})^2 4\pi^2 + \log^2(q)} + (\omega \to -\omega)\right] \tag{4.14}$$

($cos(\pi\tau)$ is used here as the continuous analog of $(-1)^\tau$. It will not cause any problem in this example. Also see (*A*.5).)

The spectrum (4.14) has a maximum at $\omega = 1/2$. The existence of this peak can be understood directly in terms of the structure of the RLG in figure 6(a). The loop on the right-hand side of the graph has length two, yielding a period two component in the spatial configurations. and thus a peak at $\omega = 1/2$ in the power spectrum. The existence of this peak can also be understood from the eigenvalues (4.9). It is the presence of a negative real eigenvalue which yields a peak at exactly $\omega = 1/2$. (The existence of negative eigenvalue can correspondingly thus be seen as a consequence of the RLG structure.) Eqn. (4.14) also shows that the half width $\log(1/q)$ of the peak decreases as q increases, as expected from the fact that the period two part of figure 6(a) is then entered more often.

Example 3 (Figure 7): The graph shows a pure loop structure with length equal to n (see [13]:chapter 2). The transition matrix [8] is:

$$P_{n\times n} = \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 0 & 1 \\ 1 & & & & & 0 \end{pmatrix}_{n\times n} \tag{4.15}$$

[8] In fact it is a adjacency matrix with all the elements equal to either 1 or 0. The relation between a graph and the characteristic polynomial of its adjacency matrix is well studied. For more details see [13].

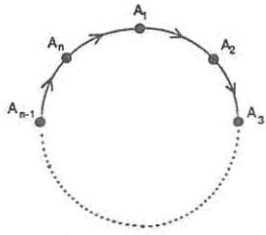


Figure 7: A simple loop structure with length n. It is not a generic RLG since all the transitions are deterministic.

$$\det(P - \lambda I) = (-1)^n(\lambda^n - 1) \tag{4.16}$$

All the n eigenvalues equally divide the unit circle, and the autocorrelation function involves $\cos(2\pi\tau/n)$ terms which will generate peaks at $\omega_0 = 1/n$, and possibly the multiples of ω_0.

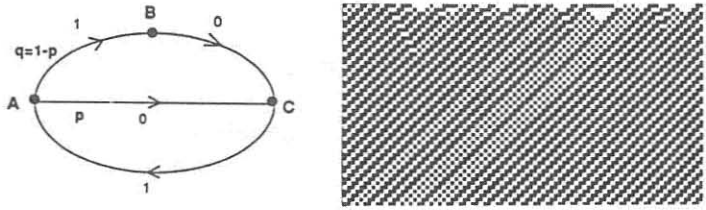


Figure 8: (a) RLG for the attractor of rule 58. (b) Spatial-temporal pattern of rule 58.

Example 4 (Figure 8): This is a more complicated regular language with 4 arcs and 3 nodes, representing the attractor for elementary CA rule 58. After assigning $\overline{AB} = 1$ to $i = 1$; $\overline{BC} = 0$ to $i = 2$; $\overline{CA} = 1$ to $i = 3$; $\overline{AC} = 0$ to $i = 4$, the transition matrix is ($q = 1 - p$):

$$P_{4\times 4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ q & 0 & 0 & p \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{4.17}$$

$$det(P - \lambda I) = (\lambda - 1)\lambda(\lambda^2 + \lambda + q) \tag{4.18}$$

The eigenvalues are

$$\lambda_1 = 1, \quad \lambda_2 = 0, \quad \lambda_3 = \frac{-1 + \sqrt{1 - 4q}}{2}, \quad \lambda_4 = \frac{-1 - \sqrt{1 - 4q}}{2}$$

$$R = \begin{pmatrix} 1 & p & 1 & 1 \\ 1 & 0 & \lambda_3 & \lambda_4 \\ 1 & 0 & \lambda_3^2 & \lambda_4^2 \\ 1 & -q & \lambda_3 & \lambda_4 \end{pmatrix} \quad L = \begin{pmatrix} 1 & 1 & 1/q & p/q \\ 0 & 1 & 0 & -1 \\ 1 & 1/\lambda_3 & \lambda_3/q & p/q \\ 1 & 1/\lambda_4 & \lambda_4/q & p/q \end{pmatrix} \tag{4.19}$$

$$C = \begin{pmatrix} (2+q)/q & & & \\ & q & & \\ & & (3q+2p\lambda_3)/q & \\ & & & (3q+2p\lambda_4)/q \end{pmatrix} \tag{4.20}$$

$$P^\tau = R \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3^\tau & 0 \\ 0 & 0 & 0 & \lambda_4^\tau \end{pmatrix} \begin{pmatrix} q/(2+q) & 0 & 0 & 0 \\ 0 & 1/q & 0 & 0 \\ 0 & 0 & q/(3q+2p\lambda_3) & 0 \\ 0 & 0 & 0 & q/(3q+2p\lambda_4) \end{pmatrix} L \tag{4.21}$$

We define the probabilities of 1's as $b_{\overline{AB}} = q/(2+q)$, $b_{\overline{CA}} = 1/(2+q)$, while the average $\overline{x} = b_{\overline{AB}} + b_{\overline{CA}} = (1+q)/(2+q)$.

If $4p - 3 > 0$ or $q < \frac{1}{4}$, the autocorrelation function and the power spectrum by (1.3) are:

$$\begin{aligned} R(\tau) &= b_{\overline{AB}}(P_{11}^\tau + P_{13}^\tau) + b_{\overline{CA}}(P_{31}^\tau + P_{33}^\tau) - \overline{x}^2 \\ &= c_1(-r_3)^\tau + c_2(-r_4)^\tau \end{aligned} \tag{4.22}$$

$$S(\omega) = \frac{c_1 \log(1/r_3)}{(2\omega+1)^2\pi^2 + \log^2 r_3} + \frac{c_2 \log(1/r_4)}{(2\omega+1)^2\pi^2 + \log^2 r_4} + (\omega \to -\omega) \tag{4.23}$$

with $r_{3,4} = -\lambda_{3,4}$ and $c_{1,2} = (pr_3 - q(2q+1))/((2+q)^2(1-4q))$.

If $4p - 3 < 0$ or $q > \frac{1}{4}$, the results are:

$$R(\tau) = d_1 r^\tau \cos(\tau\theta) - d_2 r^\tau sin(\tau\theta) \tag{4.24}$$

$$S(\omega) = \frac{d_1 \log(1/r) - d_2(2\pi\omega + \theta)}{(2\pi\omega+\theta)^2 + \log^2 r} + (\omega \to -\omega) \tag{4.25}$$

with $re^{i\theta} = \lambda_3$, where $r = \sqrt{q}$, $\theta = \tan^{-1}(-\sqrt{4q-1}) + \pi$ and $d_1 = (1+q)/(2+q)^2$, $d_2 = p/((2+q)^2\sqrt{4q-1})$.

Notice there is a "critical" probability $p_c = 3/4$ such that when $p > p_c$ the two-loop dominates and the peak stays at $\omega = 1/2$, but when $p < p_c$ the position of the peak is determined by $\theta/2\pi$, as a compromise between a two-loop and a three-loop. At $p = p_c$ the matrix is not diagonalizable, the above procedure may not give a correct result; in this case, one may use the Jordan canonical form to calculate P^τ. Eqn. (4.25) is not always a good approximation to the DFT and gives negative results at some parameter values. Figure 9 shows the spectrum obtained in this case using the exact

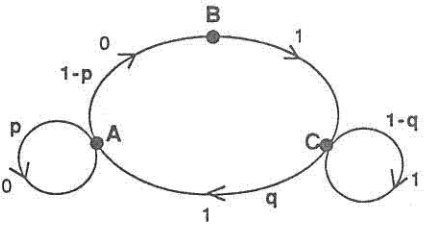


Figure 9: The spatial power spectrum obtained from eqn. (*A*.7) for rule 58 with $p = 0.3$. This is to be compared with the result of simulations shown in figure 1.

DFT formula (A.7). The similarity with the simulation (rule 58 in figure 1) is clear.

Example 5 (Figure 10): This regular language specifies the attractor for rule 200, which is a typical one with a "Brownian spectrum", having a $1/\omega^2$ tail. The feature distinguishing these regular languages from others is that the "one-loop" labeled with symbol 1 has a large probability. This is similar to the situation of a random walk (or Brownian motion) where the increment is random. The difference is that in Brownian motion the increment is in state space while in rule 200 the state is limited to $\{0, 1\}$, so the increment is in the width of domains in which all sites have value 1.

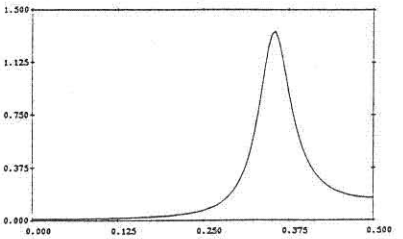


Figure 10: RLG for the attractor of rule 200, which has a Brownian spectrum.

Notice that the simple spectrum (1.4) has a $1/\omega^2$ tail when the first term in the denominator $(2\pi\omega)^2$ overwhelms the second term $(1/\tau_i)^2$; in order to keep the peak at $\omega = 0$ all the eigenvalues must in addition be positive real numbers. In our example, the transition matrix is

$$P_{5\times 5} = \begin{pmatrix} p & 1-p & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1-q & q \\ 0 & 0 & 0 & 1-q & q \\ p & 1-p & 0 & 0 & 0 \end{pmatrix} \tag{4.26}$$

with $(\overline{AA}, \overline{AB}, \overline{BC}, \overline{CC}, \overline{CA})$ as the $i = 1, 2, 3, 4, 5$. The characteristic poly-

nomial is

$$\det(P - \lambda I) = -(\lambda - 1)\lambda^2(\lambda^2 + (q - p)\lambda + (1 - p)q). \tag{4.27}$$

The resulting eigenvalues

$$\lambda_1 = 1 \; \lambda_{2,3} = 0 \; \lambda_{4,5} = \frac{(p - q) \pm \sqrt{(p + q + 2\sqrt{q})(p + q - 2\sqrt{q})}}{2} \tag{4.28}$$

are positive and real when

$$p > q \qquad \text{and} \qquad p + q > 2\sqrt{q}$$

This condition is satisfied when either q is small enough or p is big enough, or both. The configuration so generated will have distinct regions of 1's and 0's where the length of the "1-region" is controlled by parameter $1 - q$, while the length of the "0-region" is controlled by p.

5. Extensions and Conclusions

The method for calculating the spectra of regular languages described in this paper is applicable mainly to simple cellular automaton rules. The spectra of some more complicated chaotic rules can also potentially be related to regular languages. This is the case particularly when the spectrum is flat (white noise), or has sharp peaks.

Rules 30 and 90 are two examples of chaotic rules which evolve to maximum entropy, so that all possible configurations occur with equal probability, and white noise power spectra are obtained. Rules 54 and 110 show complicated behaviour, but yield comparatively simple power spectra with definite peaks. The configurations produced by these rules can be viewed as consisting of "gliders" showing complicated motion, on a rather simple periodic background. At large times, the gliders tend to annihilate, so that the background "crystalline" form dominates. The transient time required for this to occur is very long; these rules have Lyapunov exponents (defined by the rate of spatial expansion of perturbations) which are smaller than for most chaotic rules. As a result, power spectra obtained after just a few time steps may not reflect the true "attractor" results. The peaks for these rules shown in Figure 1 become progressively sharper at later times. Although it is hard to predict the complete language which characterizes the attractor, it is conceivable it is heavily weighted on some subsets on which it is a simple regular language.

The method used here to calculate spatial spectra in cellular automata can also be used to calculate spectra in other dynamical systems. One-dimensional mappings of the interval provide an example. These are deterministic processes, but when they have positive Lyapunov exponents, and when initial conditions are coarse grained, their coarse-grained dynamics may be indistinguishable from a genuinely stochastic or random process.

The probabilities for one coarse-grained subinterval to be mapped to others can then be used to form the elements of a Markov transition matrix, and the formalism discussed above may then be applied.

When the dynamics is the same with or without such "Markov partition" coarse graining, the temporal spectra computed by these are exact. This is the case for piecewise linear mappings (even though the mapping as a whole is extremely non-linear in this case) [15]. In other cases, the calculations required may be much than for simple CA, since the number of intervals (and thus the number of arcs in the RLG) needed to obtain a good approximation may be very large. Nevertheless, we expect that the bifurcation in the simple logistic map, (i.e. the accumulation of peaks with higher order sub-harmonics and the onset of broad-band noise) can be related to some qualitative changes in the eigenvalues of the Markov transition matrix. (For a different way of approximating such spectra see [16]).

This paper has concentrated on systems which can be described by regular languages. One may also consider systems described by more complicated formal languages. In particular, one may imagine using context-free languages (the next level in the Chomsky hierarchy) to explain for example the spatial spectra of some CA (e.g. rule 126).

Context free languages can be obtained from derivation trees [3], at each step of which some branches grow leaves, which then in turn grow further branches. One famous example of a context-free language is the Fibonacci sequence. This can be generated by any fixed number of steps of the generative grammar: $A \to B$, $B \to BA$, $A \to a$, $B \to b$ where A and B are "non-terminals" (intermediate states) and a, b are terminals (symbols appearing in the words at the end of the derivation).

Notice that (*i*) Context-free languages contain regular languages but the impression that a context-free language is *always* more complicated than a regular language is not correct. The Heaviside step function 0^n1^n, for example, is very simple, yet it is a context-free language which is not regular. (*ii*) Sometimes a context-free language does generate interesting non-trivial sequences such as self-similar sequences. In fact, this self-similarity is inherited in the context-free language itself since the same grammar governs branches at all stages and thus all "sizes" in the derivation tree.

The self-similarity aspect of context-free languages may make them relevant to the study of $1/f$ noise [9]. One school in $1/f$ theory uses a large number of Lorentzian spectra (1.4) with correlation lengths $\{\tau_i\}$ following a log-normal distribution [19]. This approach ensures that there will be no intrinsic time scale (or length scale if we consider spatial spectra), which is one of the most important features of $1/f$ noise. Since Lorentzian spectra are typical for regular languages, the above approach is equivalent to having a regular language with very large number of arcs and a certain

[9] $1/f$ noise appears in many different natural phemomena and there is still no satisfactory theory to explain all of them. For $1/f$ in condensed matter physics see the recent review article [18]

distribution for its eigenvalues. In order to have a $1/f$ spectrum we may either construct a complicated regular language with this constraint, or perhaps use a simple context-free language as an approximation to it. The latter stores the long range correlation in the operation of the grammar, while the former build up this correlation by a large assembly of "local" components.

In conclusion, the study of the spectra of regular languages is not only useful in explaining numerical results from cellular automata but it also provides a framework unifying many different types of spectra. Many features of the spectra can be found by examining the eigenvalues of the weighted arc-to-arc transition matrix and the structure of the RLG, without detail calculations of coefficients. This is similar to the situation in chemistry where the energy levels of a molecule can be related to its topological configuration of atoms. (One difference here is that our graphs are weighted and directed.)

Acknowledgements

I am grateful to S. Wolfram for suggesting the study of cellular automata spectra and many fruitful discussions. I also thank Z. Yin and H. Gutowitz for recommending references and for discussions, J. Lu for help in getting familiar with the C programming language. I would also like to thank B. Mel, R. Shaw and especially O. Martin for reading an earlier draft of this paper. Financial support from Columbia University when this project was started is also very much appreciated. Work at the Center for Complex Systems Research was supported in part by the National Center for Supercomputing Applications.

Appendix A. Discrete Fourier Transform formulas

The definition of the autocorrelation function from a finite number of data points $x(t)(t = 0, 1, \cdots N-1)$ is meaningful only if an appropriate boundary condition is specified. Generally the periodic boundary condition $x(t) = x(t-N)$ $(t \geq N)$ is used, so that the DFT spectrum of these data is both discrete (since the function is periodic) and periodic (since the function is discrete).

If the DFT is defined as

$$A(\nu) = \frac{1}{N} \sum_{t=0}^{N-1} e^{-i2\pi(\nu/N)t} x(t) \qquad (\nu = 0, 1, \cdots N-1) \qquad (A.1)$$

its power spectrum $S(\nu) \equiv |A(\nu)|^2$ will be related to the autocorrelation function as follows (remember that periodic boundary conditions are imposed):

$$S(\nu) = \frac{1}{N} \sum_{\tau=0}^{N-1} e^{-i2\pi(\nu/N)\tau} R(\tau) \approx \frac{R(0)}{N} + \frac{2}{N} \sum_{\tau=1}^{N/2} cos(2\pi \frac{\nu}{N} \tau) R(\tau) \qquad (A.2)$$

where the autocorrelation function is defined as:

$$R(\tau) = \left[\frac{1}{N}\sum_{t=0}^{N-1} x(t)x(t+\tau)\right] - \left[\frac{1}{N}\sum_{t=0}^{N-1} x(t)\right]^2 \qquad (A.3)$$

The three relevant formulas after neglecting higher order terms are $(R(\tau) \leftrightarrow NS(\nu))$:

$$e^{-\lambda\tau} \leftrightarrow 1 + 2\left(\frac{e^{\lambda}cos(2\pi\nu/N) - 1}{e^{2\lambda} - 2e^{\lambda}cos(2\pi\nu/N) + 1}\right) \qquad (A.4)$$

$$(-1)^{\tau}e^{-\lambda\tau} \leftrightarrow 1 - 2\left(\frac{e^{\lambda}cos(2\pi\nu/N) + 1}{e^{2\lambda} + 2e^{\lambda}cos(2\pi\nu/N) + 1}\right) \qquad (A.5)$$

$$cos(\tau\theta)e^{-\lambda\tau} \leftrightarrow 1 + \left(\frac{e^{\lambda}cos(2\pi\nu/N + \theta) - 1}{e^{2\lambda} - 2e^{\lambda}cos(2\pi\nu/N + \theta) + 1} + (\nu \rightarrow -\nu)\right) \qquad (A.6)$$

The spectrum for Example 4 when $4p - 3 > 0$ by $(A.2)$ is:

$$d_1e^{-\lambda\tau}cos(\tau\theta) - d_2e^{-\lambda\tau}sin(\tau\theta) \leftrightarrow$$

$$d_1 + \frac{d_1(e^{\lambda}cos(2\pi\nu/N + \theta) - 1) - d_2(e^{\lambda}sin(2\pi\nu/N + \theta))}{e^{2\lambda} - e^{\lambda}cos(2\pi\nu/N - \theta) + 1} + (\nu \rightarrow -\nu) \qquad (A.7)$$

References

[1] D. Shechtman, I. Blech, D. Gratias, J. W. Cahn,"Metallic phase with long-range orientational order and no translational symmetry", *Physical Review Letters* **53** (1984) 1951.

[2] A. Libchaber, J. Maurer, *Journal de Physique Collq.* **41** (1980) 3-51.

[3] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Language and Computation* (Addison-Wesley 1979).

[4] S. Karlin, H. M. Taylor, *A Second Course in Stochastic Processes* (Academic Press, 1981), and S. Karlin, *A First Course in Stochastic Process* (Academic Press, 1968).

[5] J. J. Hunter, *Mathematical Techniques of Applied Probability*, Volume 2, (Academic Press, 1983).

[6] S. Wolfram, "Statitical mechanics of cellular automata", *Reviews of Modern Physics* **55** (1983) 601-644.

[7] Appendix in *Theory and Applications of Cellular Automata*, ed. S. Wolfram, (World Scientific, 1986).

[8] S. Wolfram, "Computation Theory of Cellular Automata", *Communications in Mathematical Physics* **96** (1984) 15-57.

[9] L. Hurd, "Formal Language Characterizations of Cellular Automaton Limit Sets", *Complex System* (1987) (this issue).

[10] W. Li and S. Wolfram, "Cellular Automata Spectra", *CA-86 Conference* (Boston, May 1986).

[11] H. Gutowitz, J. Vicor, B. Knight, "Local structure theory for cellular automata", *Physica D* (1986) (submitted).

[12] W. Li (unpublished) and H. Gutowitz (private communication).

[13] D. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs* (VED Berlin, 1982).

[14] P. Grassberger, "Toward a Quantitative Theory of Self-Generated Complexity", *International Journal of Theoretical Physics* **25** (1986) 907-938.

[15] S. Grossmann and S. Thomae, "Invariant distributions and stationary correlation functions of 1-d discrete process" *Z. Naturforsch* **32a** (1977) 1353-1363.

[16] S. Thomae and S. Grossmann, "Correlation and spectra of periodic chaos generated by the logistic parabola", *Journal of Statistical Physics* (1981) **26** 485-504.

[17] R. N. Bracewell, *The Fourier Transform and Its Applications*, (McGraw-Hill, 1986).

[18] M. B. Weissman, $1/f$ *noise and Slow Non-exponential Kinetics in Condensed Matter* (University of Illinois preprint 1986).

[19] E. Montroll and M. Shlesinger, "Maximum entropy, Fractals, Scaling phenomena, and $1/f$ noise: A tale of tails" *Journal of Statistical Physics* **32** (1983) 209-230.