Chaotic Evolution via Generalized Probabilistic Automata (Probabilistic Arrays)

Azaria Paz[1]
Computer Science Department
Technion—IIT
Haifa Israel

Jack W. Carlyle
Computer Science Department
UCLA
Los Angeles CA 90095 USA

1 Summary

The finite-state probabilistic array was proposed some time ago [Paz68] as a theoretical model of polynomially nonlinear statistical interactions among the members of populations—multiway, multi-machine, or multi-parent interactions—producing state-distribution variabilities from generation to generation in the population, analogous to the variabilities studied in genetics or population biology. These machine-like stochastic systems are revisited here, with attention directed toward new details about asymptotic behavior. It is now possible to obtain more definitive results about the potential of such systems to exhibit profound disequilibrium, asymptotically over time, as the state-distributions evolve. The concepts and techniques used are suggested by chaos theory, aided by exploratory computational visualization of examples. The 2-state case is explored in detail, with brief remarks on systems with larger state sets. The principal results for 2-state systems are as follows: (i) there exist infinite-periodicity (i.e., chaotic) two-state systems of polynomial degree d (modeling d-way interactions), in particular for d as small as 5; (ii) two-state systems with quadratic nonlinearity (two-way or two-parent interactions, d=2) can evolve only non-chaotically, exhibiting periods of at most 2 in the time-sequence of state distributions.

2 Probabilistic Machines and Linear Iterations

We review very briefly the conventional probabilistic model in automata theory; the main objective is to observe that the corresponding distributional transformations are linear and have simple asymptotics under iteration.

For a time-sequential finite-state probabilistic machine which is autonomous (free of inputs, or with only a constant clocked-time input), the only significant quantities are the states, the state-to-state (conditional) transition probabilities, and the (unconditional) state-occupancy probabilities. In this setting, a machine with n states is defined by specifying its one-step transition probabilities $a_{ij}, i, j = 1, 2, \ldots, n$; the interpretation is that a_{ij} is the numerical label on the directed edge from state $\#i$ to state $\#j$ in the state graph, or the ij entry

[1]The contribution of the first author was supported by the Fund for the Promotion of Research at the Technion

in an $n \times n$ Markov matrix A:

$$a_{ij} \geq 0 \text{ for all } i, j, \text{ and } \sum_j a_{ij} = 1 \text{ for each } i \ . \tag{1}$$

The "next state" distribution is obtained from the "present state" distribution and the one-step transition probabilities:

$$P[\text{next state} = j] = \sum_{i=1}^{n} P[\text{present state} = i]\, a_{ij}, \quad j = 1, 2, \ldots, n, \tag{2}$$

This defines a mapping F of distribution-vectors into distribution-vectors, i.e., a *finite-state stochastic transformation* (f.s.s.t.), which in this case is specified by linear equations and can therefore be called a *linear* f.s.s.t.:

$$\begin{aligned} x' = F(x) \ , \text{ or } (x'_1, x'_2, \ldots, x'_n) \quad &= \quad F(x_1, x_2, \ldots, x_n) \ , \\ x'_j \quad &= \quad \sum_{i=1}^{n} x_i a_{ij}, \quad j = 1, 2, \ldots, n \ . \end{aligned} \tag{3}$$

This is an obvious stochastic extension of the concept of a mapping from the present state to the next state in conventional deterministic automata; here, "state" is generalized to "distribution over the states." The time-sequential action of a deterministic automaton generates a sequence of states starting from the initial state. In the same way, a probabilistic automaton generates a time-sequence of vectors, each of which is a distribution over the states, i.e., a list of the n state-occupancy probabilities at one instant in clocked discrete time. As time advances, the distribution vector evolves, following a trajectory determined by the iteration of the f.s.s.t. F as defined in (3), starting from some initial distribution vector π; the calculations are also expressible through repeated matrix multiplications by A:

$$\begin{aligned} \pi_0 = F^0(\pi) \quad &= \quad \pi = \pi A^0 \ (A^0 \equiv I) \\ \pi_1 = F(\pi) \quad &= \quad \pi A \\ \pi_2 = F(\pi_1) = F(F(\pi)) = F^2(\pi) \quad &= \quad \pi_1 A = \pi A A = \pi A^2 \\ &\ \ \vdots \\ \pi_t = F(\pi_{t-1}) = F(F(\cdots F(\pi)\cdots)) = F^t(\pi) \quad &= \quad \pi_{t-1} A = \pi A \cdots A A = \pi A^t \\ &\ \ \vdots \end{aligned} \tag{4}$$

Evidently it is appropriate to call this situation linear, referring either to the scalar linear relationships in (3) or to the vector/matrix linear operations in the right-hand sides of (4). Summarizing: {iterated linear finite-state stochastic transformations} $\equiv$ {repeated multiplication of probability vectors by a finite Markov matrix} $\equiv$ {state occupancy probability evolution in finite-state probabilistic machines}.

There is a unique correspondence between a linear map F and its realization as a probabilistic machine, expressible through

$$a_{ij} = j^{\text{th}} \text{ component of } F(\textit{state } i) = F_j(\textit{state } i) \tag{5}$$

where "*state i*" means the corresponding degenerate probability vector:

$$\textit{state } i \equiv \underset{(1 \text{ in } \mathrm{i}^{\mathrm{th}} \text{ position})}{(0, \ldots, 0, 1, 0, \ldots, 0)} \quad . \tag{6}$$

Deterministic machines, the conventional building blocks for models in automata theory, are included as a subcase of (3) in which all a_{ij} values are 0 or 1. In this case the transformations F can likewise be called deterministic. In the deterministic case, if the initial distribution π is degenerate—concentrated on a single state—then a sequence of iterations of F applied to π corresponds (through (6)) to the actual sequence of states of the machine as it operates in clocked time. However, if the initial distribution π is not degenerate, it is clear that even a deterministic F will in general produce a sequence of nondegenerate probability distributions upon iteration.

For linear finite-state stochastic mappings, the asymptotics of the time-sequence of iterations shown in (4) can be described using known facts about Markov chains. For these simple systems, questions related to convergence of the state-distribution to equilibrium (fixed point of the mapping), or existence of periodicities (as well as transients, absorbing states, etc.), can all be settled in a well-established manner via calculations using only the connectivities and numerical labels of the state graph, i.e., the information contained in the matrix A of state-transition probabilities. In particular, the time-sequence (4) can be written explicitly as a formula consisting of a linear combination of at most n terms, constant and real or complex exponentials, i.e., the solution to the linear difference equation $\pi_{t+1} = \pi_t A$, expressible through the eigenvalues/eigenvectors of A (see, e.g., [Kle75]). Thus the linear finite-state case is completely characterizable and has quite restricted asymptotic behavioral complexity. Consider the following elementary 2-state examples:

$$\#1\colon A = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix} \; ; \; \#2\colon A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \; ; \; \#3\colon A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; . \tag{7}$$

System #1 has a single fixed point $(0.5, 0.5)$ and all starting distributions $\pi = (x, 1-x)$ converge to this fixed point under the iterations (4). System #2 is the "identity" for which every point $\pi = (x, 1-x)$ is a fixed point trivially. System #3 has the single fixed point $(0.5, 0.5)$, but in addition it is immediately clear by inspection that the points $(0, 1)$ and $(1, 0)$ have period 2; more generally, for $\pi = (x, 1-x)$, $\pi A = (1-x, x)$ and $((\pi A)A) = (x, 1-x)$ yielding an indefinitely repeated 2-cycle, so for each x, $0 \leq x \leq 1$, the distribution $(x, 1-x)$ has period 2. Thus it can be said that System #3 has infinitely many periodic points, but again in a trivial manner, since all have the same period 2 (the fixed point and the period 2 arise from the two eigenvalues $+1$ and -1 of A). These examples should suffice to recall that, for a linear system, there can be at most finitely many distinct features (such as periodicities), bounded above by the number n of states or dimensions.

3 The Nonlinear Generalization

We now introduce certain *nonlinear* finite-state stochastic transformations. There are, of course, unlimited possibilities for defining nonlinear mappings. Here we consider a form of *polynomial* nonlinearity, which will be seen to be (i) a natural generalization of the linear case described above, (ii) interpretable as a stochastic model of interactions that are more intricate than those governed by single state-to-state transitions in finite-state machines, and, most significantly, (iii) with the potential for exhibiting an infinity of periods, leading to asymptotic disequilibrium or chaos.

An n-state f.s.s.t. F is said to be *polynomial of degree d* if it can be expressed in the form

$$x'_j = \sum_{1 \le i_1, i_2, \ldots, i_d \le n} x_{i_1} x_{i_2} \cdots x_{i_d} a_{i_1 i_2 \cdots i_d j} \tag{8}$$

where the coefficients $a_{i_1 i_2 \cdots i_d j}$ satisfy

$$(1 \ge) \quad a_{i_1 i_2 \cdots i_d j} \quad \ge 0 \quad \text{for all } i_1, i_2, \cdots, i_d, j\ , \tag{9}$$

$$\sum_j a_{i_1 i_2 \cdots i_d j} \quad = 1 \quad \text{for all } i_1, i_2, \cdots, i_d\ . \tag{10}$$

Using the terminology introduced in an earlier study [Paz68], we call any array $A = \| a_{i_1 i_2 \cdots i_d j} \|$ satisfying (9,10) a *probabilistic array*; A is said to be an n-state $(d+1)$-array (referring to the number of indexing subscripts). For each fixed j the a-coefficients constitute a *sheet* of A; the sheets are substochastic (condition (9)) and sum to an all-ones sheet (condition (10)). The conditions (9,10) generalize the Markov-matrix conditions (1) to arrays, and the relationships (8) generalize the linear case (3) to a polynomial nonlinearity.

This view of the polynomial transformation—as obtained from a multi-sheeted array—is in agreement with the concept of array multiplication [Paz68]: (8) describes the n components of a nonlinear vector/array product $\pi' = \pi A$, a special case of an array/array product $C = AB$, which is not only noncommutative but also nonassociative. A simple illustration is given in the Appendix. The time-sequential display (4) of notations for repeated iterations, introduced for the linear case, can also be applied in the nonlinear case, but because of nonassociativity, the products on the right-hand sides of (4) must be interpreted to mean that proper parenthesization is maintained:

$$\pi AAA \text{ means } (((\pi A)A)A), \text{ etc., and } A^t \text{ is not used by itself.}$$

Polynomial transformations and probabilistic arrays can be referred to interchangeably; a transformation is "realized by" an array. However, this is not quite the simple one-to-one correspondence of the linear case (noted above in connection with (5)). Some of the array coefficients are adjustable, within limited ranges, while producing the same values x'_j in (8), as in the simple example (18) of the next Section; the Appendix has further details on coefficient relationships and ranges.

Nonlinear transformations like (8) are seen in theoretical modeling of, e.g., biological processes (genetics, population growth), genetic algorithms, belief networks (in artificial intelligence), and multiway machine interactions [Paz68, May76, Pea91, Rab95]. The following model interpretation of (8) is particularly easy to describe, and can serve as a guide to the notation. A "current generation" of individuals of n "types," represented by the n states and their probabilities (or "proportions"), become parents of the children who are the members of the "next generation." For d=2, this takes place in accordance with specified conditional probability values

$$a_{i_1 i_2\, j} = P[\text{child's state} = j \mid \text{state of parent \#1} = i_1 \ \& \ \text{state of parent \#2} = i_2].$$

The state distribution $(x_1, x_2, \ldots, x_n)$ for parents is then transformed into the children's distribution $(x'_1, \ldots, x'_n)$ using (8) with d=2:

$$x'_j = \sum_{1 \leq i_1, i_2 \leq n} x_{i_1} x_{i_2} a_{i_1 i_2\, j} \quad . \tag{11}$$

Here, the probabilities x_{i_1} and x_{i_2} are multiplied because the two parents are selected independently from the population and the parents are considered to be statistically identical, i.e., drawn as a random sample of size 2 with the common distribution $(x_1, x_2, \ldots, x_n)$. For $d > 2$, the interpretation is the same, with d parents per child. Iteration of (8) or (11) produces a trajectory (time sequence of state distribution vectors) which might be called an *evolution*, i.e., a history of generation-to-generation changes in the population characteristics.

4 The Two-State (One-Dimensional) Case

First, a remark on a familiar geometrical fact: the set of all n-state probability vectors forms a space of dimension $n-1$, namely the convex polytope formed by the intersection of the unit n-cube with the hyperplane $x_1+x_2+\cdots+x_n = 1$. For example, the set of all 3-component probability vectors consists of the boundary and interior of a triangle, whose vertices are identified with the states through the correspondence (6). Thus, in geometric terms, an n-state f.s.s.t. maps an $(n-1)$-dimensional simplex into itself. Therefore, two-state transformations are, in effect, one-dimensional mappings, and can accordingly be simplified, in notations and calculations, as follows.

With $n = 2$ and $\pi = (x_1, x_2) = (x_1, 1 - x_1), \pi' = (x'_1, x'_2) = (x'_1, 1 - x'_1)$, the two components of a stochastic mapping $\pi' = F(\pi)$ are

$$x'_1 \quad = \quad F_1(x_1, x_2) \geq 0, \quad x_1 \geq 0, \ x_2 \geq 0, \quad x_1 + x_2 = 1,$$

$$\begin{aligned} x_2' &= F_2(x_1,x_2) \geq 0, \quad x_1 \geq 0,\ x_2 \geq 0, \quad x_1+x_2=1, \\ x_1'+x_2' &= 1. \end{aligned}$$

Substituting $x_1 = x$, $x_2 = 1-x$, $f(x) = F_1(x, 1-x)$, this reduces to

$$\begin{aligned} x' &= f(x), \quad 0 \leq x \leq 1 \\ 1-x' &= 1-f(x), \quad 0 \leq x' \leq 1 \end{aligned} \tag{12}$$

and the second equality is redundant, so only a single mapping f of one scalar variable needs to be considered; f is a one-dimensional stochastic transformation, i.e., maps $[0,1]$ into $[0,1]$. Likewise, for the iterates of F,

$$F^t(x_1,x_2) = F(F(\ldots F(x_1,x_2)\ldots)) \ ,$$

it suffices to consider the one-dimensional iterates

$$f^t(x) = f(f(\ldots f(x)\ldots)).$$

For $n=2$, in the polynomial mapping definition (8), the coefficients $a_{i_1 i_2 \cdots i_d 1}$ and $a_{i_1 i_2 \cdots i_d 2} = 1 - a_{i_1 i_2 \cdots i_d 1}$ form the two $2 \times \cdots 2 \times 2$ [d-fold] sheets of a $2 \times \cdots 2 \times 2$ [$(d+1)$-fold] array. The condition that all $a_{i_1 i_2 \cdots i_d j}$ "sum to 1 over sheets," i.e., that the sum over the $(d+1)$st coordinate is 1 for each element, reduces for $n=2$ to the simple observation that the two sheets sum to the all-ones sheet. Therefore in the two-state case we can largely ignore the second sheet, which is obtainable from the first by subtractions from 1; this corresponds to the above-mentioned redundancy of $F_2(x, 1-x) = 1 - f(x)$. Thus to specify any particular 2-state example, it is actually only necessary to give one $2 \times \ldots \times 2$ [d-fold] sheet of coefficients $a_{i_1 i_2 \cdots i_d}$ which lie between 0 and 1; i.e., the multivariate polynomial definition (8) can be reduced to the one-variable specification

$$f(x) \ = \ x' \ = \sum_{\substack{i_1,i_2,\ldots,i_d=1,2 \\ (x_{i_1},\ldots,x_{i_d} \ = \ x \text{ or } (1-x))}} x_{i_1} x_{i_2} \cdots x_{i_d} a_{i_1 i_2 \cdots i_d} \tag{13}$$

$$= \sum_{k=0}^{d} c_k \ x^k (1-x)^{(d-k)} \ = \ \sum_{k=0}^{d} C_k \ x^k \ . \tag{14}$$

The alternate forms (14) simply take note of the fact that the terms of the polynomial can be collected in any convenient manner. Some coefficient values are not uniquely fixed for a given mapping f, as is easily seen, e.g., from the simple example (18) below.

For a two-state model with $d=2$, i.e., a two-state *quadratic* system, the defining relationship $x' = f(x)$, in terms of the $a-$coefficients, becomes simply

$$x' = a_{11}x^2 + (a_{12}+a_{21})x(1-x) + a_{22}(1-x)^2 \tag{15}$$

where the four $a-$coefficients form a substochastic 2×2 sheet

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \ .$$

For example,
$$x' = x^2 + (1-x)^2 \tag{16}$$
is a stochastic 2-state quadratic transformation realizable by the array (first sheet)
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
which is a deterministic realization. Likewise, the example
$$x' = 2x(1-x) \tag{17}$$
is produced by the sheet
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

To see that there may be some freedom of choice for "off-diagonal" a-values, it suffices to consider another simple example; e.g., the two distinct choices
$$\begin{pmatrix} 1 & 0.5 \\ 0.7 & 0.4 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0.2 \\ 1 & 0.4 \end{pmatrix}$$
produce the same mapping
$$x' = x^2 + 1.2x(1-x) + 0.4(1-x)^2 \ . \tag{18}$$
Such variations in the array realizing a given nonlinear function are not significant for our purposes; nonuniqueness is simply a consequence of the stochastic constraints in the alternate formulations in (13) and (14), in which the a-coefficients are "uncollected," while the c- or C-coefficients are "collected." (For the corresponding formulations in the general multistate case, see the Appendix.)

The example
$$x' = 3.6x(1-x) \ , \tag{19}$$
which is in appearance similar to the realizable (17), cannot be a stochastic two-state quadratic transformation since it has no probabilistic array realization: from (15), with substochastic values a_{12} and a_{21}, we see that the coefficient of $x(1-x)$ cannot be 3.6; in fact it cannot exceed 2.

The transformations (17) and (19) are instances of the *logistic mapping*
$$x' = Kx(1-x) \ , \tag{20}$$
which has been used, e.g., in modeling population growth [May76], and is the canonical example leading to the study of chaos in iterated mappings [Fei80]. As a growth model, (20) describes the number of individuals in the current population as a function of the number in the immediately preceding generation (where "number" may mean a scaled or normalized fraction of individuals of a certain type), and the progress of the iterations shows how the population varies from generation to generation. We shall review definitions of the term "chaos" below, but accepting the intuitive concept of "complex and unsettled

asymptotic behavior with an infinity of periods," it is well-known [e.g., Fei80] that (20) begins to exhibit chaotic behavior when the "growth factor" $K \geq 3.57$ approximately; (17) is not chaotic.

Since (19) cannot be realized as a 2-state probabilistic array, although the mapping is stochastic, the following question immediately arises: are there any two-state quadratic stochastic mappings which are chaotic? We show here that there are none (Section 8). This leads to the question: allowing the degree d to be sufficiently large (more than 2), are there any chaotic two-state stochastic mappings? This is answered in the affirmative (Section 7), with examples and details showing values of d sufficient to ensure the possibility of chaos.

We make frequent use of the family of examples of the form

$$x' = x^d + (1-x)^d \tag{21}$$

which defines, for each d, a one-dimensional (2-state) stochastic mapping of degree d. This can be realized by the choice in (14) of $c_0 = a_{11\cdots 1} = 1$, $c_d = a_{22\cdots 2} = 1$, and all other $c_k = 0$; note that this choice gives a deterministic realization. For odd degree d, collecting powers of x in (21) yields a polynomial of degree $d-1$ in x, but the system as a whole is still properly said to be of degree d, since it cannot be realized with a lower-degree probabilistic array; this is a reflection of the fact that the ignored "redundant" polynomial $F_2(x, 1-x) = 1 - f(x)$ contains a degree-d term. The same remarks apply to d for multistate mappings in general; see the Appendix.

5 Chaos Defined

We use standard terminologies such as the following, for a general transformation $F\colon y = f(x)$ which maps some set E, e.g., a metric space, into itself.

An orbit:	a set $\{x, F(x), F(F(x)), \ldots\}$ (unordered set of points of "trajectory" or sequence of iterations from initial x)
a fixed point x of F:	a point x such that $F(x) = x$
periodic point x, period k:	$F^k(x) = x$ and k is least $(F^j(x) \neq x, j < k)$
periodic orbit or k-cycle:	$F(p_1) = p_2, \ldots, F(p_{k-1}) = p_k\,, F(p_k) = p_1$ (each p_i has period k; fixed points of F^k)
the periods of F:	set of integers, the periods of all periodic points

and, for one-dimensional maps f which are differentiable functions of one real variable,

critical point:	x such that $f'(x) = 0$
attracting periodic point:	x of period k and $\lvert (f^k)'(x)\rvert < 1$ (an attractor or sink)
repelling periodic point:	x of period k and $\lvert (f^k)'(x)\rvert > 1$.

Reducing the concept of "chaos" in iterated maps to a formal definition is subject to some variability in different application contexts; however, one widely quoted mathematical characterization [Dev89] consists of the following three conditions, for a mapping $F\colon E \to E$.

(c1) F has sensitive dependence on initial conditions
(c2) F is topologically transitive
(c3) periodic points are dense in E .

It has been shown that the three conditions are not independent and can be reduced to two by eliminating (c1) [Ban92], or all three can be replaced with a single combined condition [Tou97]

(c0) each pair of nonempty open sets shares a periodic orbit.

Nevertheless, the original three conditions are of interest because they have reasonable intuitive meanings. First, "sensitive dependence..." suggests future unpredictability from small deviations from a starting point, and is defined simply by requiring that, given any x, neighborhood N of x, and $\delta > 0$, there is $y \in N$ and k such that $|F^k(x) - F^k(y)| > \delta$. Topological transitivity embodies the idea of mixing or indecomposability, i.e., rules out a breakup into two or more noninteracting subsystems: for any pair of nonempty open subsets U, V of E, $F^k(U) \cap V$ is nonempty for some sufficiently large k. The density of periodic points suggests a residue of regularity embedded within the unpredictability. Clearly all three concepts are in accord with informal ideas of chaotic behavior. The definition itself (or one of its combined replacements) does not yield generally usable procedures for deciding whether a particular example is chaotic; typically, some ad-hoc analysis is needed, although it is occasionally possible to make use of topology-preserving conversions of previously established chaotic maps.

We direct our attention mainly to a related condition, which also suggests chaos intuitively, namely

(c4) the existence of an infinity of integers each of which is a period.

Condition (c4) is adequate for our purposes, and is of particular interest because it directly contrasts with the very restricted situation in the linear case (not more than n periods) illustrated by the linear examples (7). Superficially (c4) appears to be possibly both wider and narrower than (c3); e.g., the infinity of periodic points guaranteed by (c3) might not all have different periods, and the infinity of periods from (c4) might not yield density of periodic points, and (c1,c2) also need to be considered. However, for examples of one-dimensional polynomial maps such as those discussed in Section 7 below, after establishing (c4) it is feasible (on a case-by-case basis) to verify the other conditions as well; we do not pursue the details here.[1]

[1] Condition (c4) also would suffice in some situations where the set E is restricted ($F\colon E \to E$), e.g., $f(x) = Kx(1-x)$ for $K > 2 + \sqrt{5}$ which is known to be chaotic on a Cantor set [Dev89] (this is not a stochastic example).

6 Graphical Tools

One-dimensional examples can be illustrated and explored graphically with several types of charts. First, straightforward graphical overlays of f and some of its iterates f^k (functional compositions), together with the 45-degree line $I(x)=x$, will assist in identifying fixed points and some periodic points (new crossings of I by graphs f^k). Such points may then be accurately evaluated numerically, whether or not f is chaotic, since no temporal asymptotics are involved; i.e., the graphs are used in a "purely analytic" manner. See, e.g., Figs. 2,8,12.[1] In contrast, there are charts depicting what might be called "experimental data," even though computed reproducibly with apparently high precision; these charts show points, typically large numbers of points, generated by computing repeated iterations of f starting from some chosen initial x-value; i.e., charts of $x, f(x), f(f(x)), \ldots$. The appearance of such "experimental" charts may vary with the initial condition, but often may be strikingly similar for a range of initial conditions, therefore suggesting trends to study in seeking formal verification of chaos/nonchaos for f. In the accompanying Figures, three types of experimental charts appear, as follows.

An "orbit chart" is a one-dimensional view in which iterated values are plotted on the x axis, perhaps with elongated vertical markers for better visual display; the appearance is reminiscent of a "bar-code." For examples, see Figs. 1 and 6.

A "sorted chart" is a two-dimensional figure, a nondecreasing plot of iterated values as ordinates, with their positions in the sorted list of values (minimum to maximum) as abcissas. If suitably normalized, a sorted chart can be thought of as an experimental estimate of a cumulative probability distribution of points visited in the long run. In a sorted chart, periodicities may be revealed by the appearance of multiple horizontally flat segments or "stairtreads." The possibility of chaos is suggested by monotone increasing segments; e.g., a sorted chart having the appearance of a diagonal line would suggest that, in the long sequence of iterations, "all" (i.e., arbitrarily many) values have been visited with roughly equal frequency. Vertical segments suggest "gaps," i.e., intervals which may not be (re)visited. For examples of sorted charts, see Figs. 4,7,11. In Fig. 4, the dominance of two levels suggests an attracting 2-cycle (there is one, as can be deduced from the corresponding function plots in Fig.2).

A "diagonal phase portrait" (using the terminology of [Dev89]) is an (x, y) chart in which, starting with x_0 and generating successive iterated values x_t, a vertical line segment is drawn from the point (x_t, x_t) on the diagonal line to the point $(x_t, f(x_t) = x_{t+1})$, then a horizontal segment to the point (x_{t+1}, x_{t+1}), etc., thus showing the relationship of the trajectory to the f graph and the 45-degree line; both may optionally appear in the same graph as overlays. In a diagonal phase portrait, progress toward an attracting fixed point may be revealed as a sequence of "stairsteps" of decreasing size, or by a shrinking square-sided spiral;

[1] All Figures are collected at the end of the paper. Many of the Figures are plots in the unit square $0 \leq x \leq 1, 0 \leq y \leq 1$, but the axes may be unequally sized for convenience in display; the diagonal $y \equiv x$ is indeed a 45-degree line when these scalings are taken into account.

an attracting period 2 is suggested by a box as limit-cycle, etc., while chaos is suggested by a "largely black" plotted result. Figs. 3,5,10 show diagonal phase portraits. In the simple example of Fig. 3, from the starting point near the upper right corner, a downward stairstep approaches the fixed point, and then spirals outward (since the fixed point is repelling) toward a period 2 limit box.

Tracing the first few iterations as a diagonal phase portrait, for various trial starting x-values, can assist in demonstrating that an example f has, e.g., an attracting or repelling fixed point or a periodic cycle. This is called *graphical analysis* in [Dev89] and has been widely used in the literature on iterated maps; we make use of the idea in several arguments. Of course this is only intended to be a shorthand indication that a formal verification is easy to obtain and therefore need not be written out in detail, at least in visually simple cases.

As noted above, chaotic behavior is suggested by an extensive "blackened" appearance for orbit charts and phase portraits, and by the presence of intervals of steady increase in sorted charts. Of course, these computational/visual aids must be used with caution and are exploratory only, suggesting an avenue for formal investigation. The infinity of periodic points associated with chaotic maps is ultimately inaccessible via finite-precision computer-generated data and plots, but plots do make "close visits," and are therefore useful in the sense that they may display "silhouettes" yielding clues about underlying behavior. (Various "close-visiting" properties are known to hold (e.g., the "shadowing lemma" [Pei92]).)

Our illustrative calculations and chart generations are done to 80 bit IEEE extended precision or better. In the Figures, examples of experimental-data plots are shown for moderate numbers of iterations (100 to 1000), but we have examined much larger iteration numbers (up to 20,000 in some cases), and have varied the initial x choices, to increase confidence that the plots are "typical." As the number of iterations increases, aside from the ultimate limitations of finite precision computer arithmetic, the effects of even coarser granularities of graphic/printed displays will eventually begin to mask subtle details; again this emphasizes that such charts are to be taken as guides to further study rather than as solutions. On the other hand, as already noted, graphs of functions used in a purely analytic manner (e.g., for rootfinding) do lead to solutions with whatever accuracy is desired through use of standard numerical routines, and as suggested in [Dev89], nests of small containing intervals may likewise be accessible to accurate calculation, in tracing the progress of iterations.

7 Chaos in Two-State Systems

We have noted that the chaotic logistic map (20) with $K = 3.6$ is not realizable with a two-state array of degree d=2; see the remark after (19). However, the question of its array-realizibility with two states but higher degree can be pursued: is there a one-variable polynomial $f(x)$ which can be written in the form (14) for some $d > 2$ but is actually equal to $3.6x(1-x)$? A brute-force approach can be taken, choosing a trial degree d and equating to zero all coefficients of powers of x except for $3.6x$ and $-3.6x^2$; the coefficients must also be constrained to lie within ranges derived from the stochastic conditions on the array a-values, or equivalently by the requirement that each c_k-coefficient in (14) must lie between 0 and $\binom{d}{k}$. A solution is obtainable, e.g., for $d = 9$; we omit these calculations (they are summarized in the Appendix).

This establishes that there do indeed exist higher-degree chaotic 2-state array-realizable systems. However, instead of relying on the known behavior of a particular logistic map, we choose to discuss in detail a different example which is of considerably lower degree, $d = 5$, and is independently analyzable.

The mapping to be studied is (21) with $d = 5$:

$$f(x) = x^5 + (1-x)^5 \quad [= 5x^4 - 10x^3 + 10x^2 - 5x + 1\,]\,. \tag{22}$$

This mapping is realizable using a $2 \times 2 \times 2 \times 2 \times 2$ substochastic sheet which can be selected to be *deterministic*: the elements of the sheet are

$$a_{11111} = 1,\ a_{22222} = 1,\ a_{-----} = 0 \text{ otherwise},$$

and of course the complementary values $(0 \to 1, 1 \to 0)$ appear in the ignored second sheet, to complete a 2-state 6-array realization. In Fig. 8 the function (22) and its iterates up to the sixth are plotted. It is clear that the sixth iterate has new crossings of the diagonal—crossings not shared by the function itself or the first five iterates—so 6 is one of the periods of f.

We shall now make use of the Sarkovskii ordering; the following is a very brief summary (see, e.g., [Dev89]). Let $\succ$ denote the following ordering of the positive integers:

$$\begin{aligned} 3 \succ 5 \succ 7 \succ 9 \succ \cdots \succ 2{\cdot}3 \succ 2{\cdot}5 \succ \cdots 2^2 3 \succ 2^2 5 \succ \cdots \succ 2^3 3 \succ \cdots \succ \cdots \\ \cdots \succ \cdots \succ 2^4 \succ 2^3 \succ 2^2 \succ 2 \succ 1 \end{aligned} \tag{23}$$

(the first line begins with all odd integers except 1, and eventually lists all integers which are not powers of 2; the second line lists all powers of 2 in decreasing order ending with $2^0 = 1$). The principal result is the

> *Theorem*: if f is a continuous real function, k is a period of f, and $k \succ \ell$, then ℓ is a period of f.

Consequently, an infinity of periods is guaranteed whenever it can be shown that f has period 3, or any other period which is not a pure power of 2.

Returning to our example, the period $6{=}2^1 3$ appears early in the Sarkovskii ordering (23), and therefore we can immediately conclude that f exhibits infinitely many periods. Note that for this example, 6 is the smallest number available for application of Sarkovskii's result (graphs of the third and fifth iterates do not yield new diagonal crossings).

Some of the "experimental evidence," as provided by sorted charts of iteration values, is shown in Fig. 7, supporting the idea that f is chaotic. These charts do not exhibit a visible collection of flat "stair-treads." Instead there is a generally monotone increasing trend, consistent with the known presence of an infinity of periodicities, and further suggesting that there may be no attracting periodic cycles; this will be verified below. Further support for the hypothesis of chaoticity is given by diagonal phase portraits such as Figs. 5 and 10, which do not seem to exhibit consistently traversed limit-cycle "boxes," although this is more difficult to decide in a phase portrait if there are overlying "black" (supposedly chaotic) regions.

Another obvious feature in the sorted charts is the "gap" (vertical segment) in the range of values, indicating an interval not visited (except possibly for the first few iterations when a starting x-value is chosen in the gap interval). As can be seen in Fig. 8, the fixed point of f ($x \approx 0.245$) lies within this interval and is a repelling fixed point (i.e., at this point the the function has a slope which is larger than 1 in absolute value). The interval can be quantified, aided by graphical analysis with diagonal phase portraits; the calculations are not pursued here, but the phenomenon can be seen by inspecting the phase portraits in Figs. 5 and 10, which show an inner "white box" region that is not entered (or, if initially entered, is left and not reentered).

Returning to the hypothesis that there exist no attracting periodic cycles (which in a sense suggests the "most chaotic" of behaviors), this is verifiable for the example, through a combination of further graphical/numerical inspections and use of known theorems; we give only a brief sketch of the ideas.

It is instructive, although not necessary, to evaluate numerically the periodic points and slopes up to period 6; see the Appendix. It suffices here simply to observe that, as can be seen by close inspection of Figs. 8, there are 12 new crossings of the diagonal by the 6th iterate, all of which can be calculated easily to any desired accuracy and will be found to fall into two 6-cycles; the slopes for the 6th iterates, evaluated on these 6-cycles, are greater than one in absolute value, and therefore these cycles are repelling. The same can be said for smaller periods (there is a 2-cycle and a 4-cycle). Note that for evaluation of slopes, we can make use of the identity

$$(f^k)'(p_i) = f'(p_1)f'(p_2)\cdots f'(p_k) \tag{24}$$

which applies to the points of a k-cycle $\{p_1, p_2, \ldots, p_k\}$. This identity follows from differentiating $(f^k)(x)$ via the chain rule and substituting the cyclic relationships among the p_i and $f(p_i)$. The Appendix lists values obtained using (24) with $f'(x) = 5(x^4 - (1-x)^4)$. Again, accurate slope values are not important here; we only observe that they are larger than one in absolute value.

To investigate the existence/nonexistence of attracting cycles in general, we appeal to the negativity of the Schwartzian derivative, and its consequences. Reviewing the facts briefly [Dev89]: The Schwartzian derivative Sf is defined by

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}^2\right) . \tag{25}$$

If this is negative (values of $-\infty$ are allowed), and if f has m critical points (points of zero slope), then f can have at most m attracting periodic orbits (plus possibly two more associated with excursions to $\pm\infty$, which need not concern us here because of the stochastic constraint on the mappings studied); i.e., to each of these orbits at least one critical point of f must be attracted.

To apply this result to our example, note that a function of the form $x^d+(1-x)^d$ has exactly one critical point in [0,1], namely $x = 0.5$, where f is minimum; thus if the Schwartzian derivative is negative, it follows immediately that there is at most one attracting cycle. In addition, further graphical analysis of (22) starting from the single critical point 0.5 can be used to aid in verifiying that there are no attractions of this point (details omitted, but the phenomenon is clearly visible in Fig. 11(a)). The overall conclusion, then, is that this f has no attracting periodic points, if Sf is negative for this example f. The Schwartzian derivative of (22) is

$$f(x) = \frac{(120x-60)/(5x^4-5(1-x)^4)-(3/2)((20x^3+20(1-x)^3)}{(5x^4-5(1-x)^4))^2}\,, \qquad (26)$$

which is easily verified (numerically or analytically) to be negative, including the expected $-\infty$ value at the critical point 0.5; the Schwartzian is an even function about this point, and is in fact negative for all real x. A plot is shown in Fig. 9.

To summarize, the $d = 5$ example (22) has an infinity of periods, so is chaotic in this sense, and moreover has no attracting cycles. To establish conformance to a full formal definition (Section 5), one would need to verify that periodic points are dense in $[0, 1]$ or a subset and transitivity holds (or that an equivalent combined condition holds); although this is not needed for our purposes, we remark that it is feasible to adapt techniques [Dev89] involving the "first return map."

The family of examples $x' = x^d + (1-x)^d$ for other values of d can also be examined, aided by graphical inspection, suggesting chaotic behavior for, e.g., d=6,7,8,9. However for d=3,4 apparently regular behavior is observed, with simple limit cycles; for d=4 see Figs. 1,2,3,4. We return to this issue in Section 10. The chaotic case d=9 is of some interest because there is an attracting periodic cycle, evidence of which appears in, e.g., sorted charts such as those in Fig. 11. This example is indeed chaotic because it can easily be found numerically to have period 3, the first entry in the Sarkovskii ordering. We omit the calculations of periodic points and slopes; the results show that there is a period-8 orbit with slope less than one in absolute value, verifying that the orbit is attracting. This must be the only attracting orbit: the Schwartzian is again negative, and calculations will verify that the single critical point x=0.5 is immediately attracted to the 8-orbit as in Fig. 11(a), and the points of a small containing interval, say (0.499, 0.501), are likewise attracted. For other starting values, the attracting cycle is visibly embedded in chaotic behavior as in Fig. 11(b), and both the starting point and the number of iterations plotted will affect the apparent "percentage" of embedding (this is somewhat illusory, of course, since the example possesses an infinity of periods no matter what seems to dominate in a particular computed picture).

8 Regularity of Two-State Quadratic Systems

Proposition. For $n = 2$ states and degree $d = 2$, chaos cannot occur; i.e., a one-dimensional stochastic quadratic map f of the form (15) can exhibit only finitely many periods, in fact at most period 2, for any permitted values of the four a_{ij}-coefficients.

This is established through a case analysis showing that, in addition to the obvious fixed point(s), there may be period-2 points for some coefficient values, but that no greater periodicities will be encountered, because the relevant points are attracting; i.e., in all cases either |slope of f| < 1 at a fixed point or if not then a subsequently-appearing 2-cycle is such that |slope of f^2| < 1 at the period-2 points.

For notational convenience in this analysis, the coefficients of the quadratic (15) are renamed as follows:

$$\begin{aligned} x' = f(x) &= ax^2 + bx(1-x) + c(1-x)^2 \quad [\, 0 \le a, c \le 1, 0 \le b \le 2 \,] \\ &= f(1)x^2 + (4f(0.5) - f(1) - f(0))x(1-x) + f(0)(1-x)^2 \\ &= (a + c - b)x^2 + (b - 2c)x + c \\ &= Ax^2 + Bx + c \ . \end{aligned} \tag{27}$$

This one-dimensional mapping f arises from the first substochastic sheet of a 2×2×2 array (i.e., a 2-state 3-array) of the form

$$\left(\begin{pmatrix} a & b_1 \\ b_2 & c \end{pmatrix} \begin{pmatrix} 1-a & 1-b_1 \\ 1-b_2 & 1-c \end{pmatrix} \right)$$

with off-diagonal values satisfying $b_1 + b_2 = b$ ($b \le 2$ since $0 \le b_1, b_2 \le 1$).

The slope and curvature of the quadratic f are given by

$$\begin{aligned} \text{Slope :} \quad & f'(x) = 2Ax + B \\ \text{Curvature :} \quad & f''(x) = 2A = -2(b - a - c) \end{aligned} \tag{28}$$

Without loss of generality f can be assumed to be convex-cap; i.e., the constant curvature f'' can be taken to be negative: $A < 0, b > a + c$. Positive curvature can be handled in an identical manner, or through consideration of the mapping arising from the second sheet of the probabilistic array. Cases of zero curvature, i.e., reduction of f to a linear or constant function (i.e., realizable by a 2-state Markov chain), can be dismissed immediately, aided by simple graphical inspection showing one of: a single attracting fixed point in the interval $[0, 1]$; or a trivial "all-points-are-fixed" situation when $f(x) = x$; or a trivial "all-2-cycles" situation when $f(x) = 1 - x$ (corresponding to the three Markovian examples (7)). The case $c = 0$ ($f(0) = 0$) can also be dismissed: $f(x) = (Ax + B)x$ has a fixed point at $x = 0$ with slope $B = b$, attracting all points if $b \le 1$; if $b > 1$ the additional fixed point at $x_0 = (b - 1)/(b - a)$ ($b - a = -f''/2 > 0$) appears with $|f'(x_0)| = 2 - b \le 1$ and all points iterate toward x_0. Furthermore, $a = f(1) < 1$ can be assumed (otherwise, with f a convex-cap quadratic, there is only the trivial attracting fixed point $x = 1$.)

To summarize, it suffices to restrict attention to cases in which

$$0 \leq a < 1, \quad 0 < c \leq 1, \quad a + c < b \leq 2 \tag{29}$$

and further restrictions on the coefficients to be studied will be introduced in stages below. First, some typical functions f (and f^2) are graphed in Fig. 12; graphs (i) through (v) show functions satisfying the constraints (29), while for contrast (vi) is the logistic map (19) which does not satisfy the constraints.

In cases where the slope of f at its fixed point p_0 is less than one in absolute value, e.g., Fig. 12(i), it is immediately clear (with the aid of simple graphical analysis) that all starting x-values in the unit interval will be attracted to p_0; in these cases there can be no periodic cycles and further study is not needed, because there can be no new crossings of the diagonal $I(x) = x$ by any of the functions $f^2, f^3, \ldots$.

In Figs. 12(iii) and (iv), the slope of f at p_0 is > 1 in absolute value, and the presence of two new crossings of I by the second iterate f^2 shows that there is a 2-cycle; in Fig. 13 the cycle is outlined as a square, as it would appear as a limit in a diagonal phase portrait. The slopes of f^2 at the periodic points are small in absolute value, so the cycle is attracting.

It must be demonstrated that these are the only behaviors possible. Specifically, it remains to show that if there is a 2-cycle, the slopes at the period-2 points cannot be larger than one in absolute value. We approach this through a series of further restrictions to those cases in which a 2-cycle can actually occur, with verification that several significant x and y values are then constrained in such a way that the main conclusion follows.

Let p_0 be the fixed point of f in the interval $[0, 1]$, and let s_0 be the slope at p_0. Let e be the point at which f assumes its maximum value m as a convex-cap quadratic on the real line; i.e., in some cases the maximizing point e may lie outside $[0, 1]$. Then

$$0 = f'(e) = 2Ae + B, \quad e = \frac{B}{-2A} = \frac{b - 2c}{2[b - (a + c)]} \tag{30}$$

$$m = f(e) = Ae^2 + Be + c = -Ae^2 + c = \frac{B^2}{-4A} + c \tag{31}$$

$$s_0 = f'(p_0) = 2Ap_0 + B = 2A(p_0 - e) \tag{32}$$

$$p_0 = \frac{B - s_0}{-2A} = e + \frac{-s_0}{-2A}\ . \tag{33}$$

The fixed point p_0 and the slope s_0 may be expressed in terms of the coefficients of f by solving $Ax^2 + Bx + c = x$:

$$p_0 = \left[-(B - 1) - \sqrt{(B - 1)^2 - 4Ac}\right] / 2A \tag{34}$$

$$s_0 = 1 - \sqrt{(B - 1)^2 - 4Ac} = 1 - \sqrt{(b - 1)^2 + 4c(1 - a)} \tag{35}$$

$$p_0 = \frac{b - 2c}{2[b - (a + c)]} + \frac{\sqrt{(b - 1)^2 + 4c(1 - a)} - 1}{2[b - (a + c)]} \tag{36}$$

Changing $-\surd$ to $+\surd$ in (34) produces an irrelevant fixed point, outside of [0,1], to the left of the origin: $f(x)$ is a convex-cap quadratic behaving asymptotically like $-x^2$ as $x \to \infty$, so a negative-x crossing of the 45-degree line is necessarily present, whenever there is a positive crossing point p_0. According to (35),we can say the following about the slope at the relevant fixed point p_0 in $[0,1]$:

$$s_0 < 0 \quad \text{if} \quad (b-1)^2 + 4c(1-a) > 1 \tag{37}$$

$$s_0 < -1 \quad \text{if} \quad (b-1)^2 + 4c(1-a) > 4 \tag{38}$$

This is clearly possible within the coefficient ranges (29). The choices $a = 0, b = 2, c = 1$ yield the largest attainable magnitude:

$$\max[(b-1)^2 + 4c(1-a)] = 5 \quad (\text{then } s_0 = 1 - \sqrt{5} = -1.23607). \tag{39}$$

The slope s_0 when positive must be < 1, since the largest attainable positive value is $s_0 = 1 - \sqrt{\epsilon}$ for small positive ϵ, e.g., $\epsilon = 4c(1-a)$ with c close to 0 and a close to 1 (satisfying (29)). Therefore, in the cases of interest here, at least (37) is added to the list of restrictions on coefficients, or the stronger (38) is assumed as appropriate.

With s_0 negative, a graph showing f and the 45-degree line I must have the feature that f crosses I from above/left to below/right; i.e., the pair (p_0, p_0) will be seen below/right of the pair (e, m), since the latter is the "peak of the cap" of the quadratic curve and the region of negative slopes is $[x > e]$: $m = f(e) > f(p_0)$ with $e < p_0$. (It can also be seen from (32) that $p_0 > e$, for negative s_0 and negative curvature $2A$). The maximizing point e may sometimes be to the left of $[0, 1]$ (an exact evaluation is not needed here—both cases are allowed in the argument).

Now consider the second iterate $f^2 = f(f)$; this function potentially has four fixed points, the roots of the quartic $f^2(x) - x$. Two real fixed points are always present, identical to those of f: p_0 and an irrelevant outside point (in Fig. 12(v), this point would be below the lower left corner of the displayed area). If there are two more real roots of $f^2(x) - x$ in $[0, 1]$, they are the source of a 2-cycle. We wish to verify that, as anticipated, the two new points appear when $s_0 < -1$. Call these period-2 points p_1 and p_2 ($p_1 < p_2$); their locations and f^2-slopes are to be studied.

The slope and curvature of f^2 are

$$(f^2)'(x) \quad = \quad f'(f(x))f'(x) \tag{40}$$

$$(f^2)''(x) \quad = \quad f''(f(x))[f'(x)]^2 + f'(f(x))f''(x) \tag{41}$$

There are potentially three critical points—zero-slope points—for f^2, namely the maximum point e for f ($f'(e) = 0$), and possibly two solutions (if real-valued) of $f(x) = e$. These solutions correspond to the intersections, if they exist, of the convex-cap quadratic f with the horizontal line $y \equiv e$; two intersections if $e < m$, one if $e = m$, and none if $e > m$. The latter two cases are inapplicable here, with the restriction to $e < p_0 < m$ already in effect. So $f(x) = e = -B/2A$

has two real solutions, which can be written in the form

$$\frac{B \pm \sqrt{B^2 - 2B - 4Ac}}{-2A} = e \pm \sqrt{e^2 - (c-e)/A} \quad . \tag{42}$$

The two points are symmetrically located about e, which is an obvious property of $f^2 = f(f)$: considered as a function defined on the entire real line, $f(f(x))$ inherits from $f(x)$ its even symmetry about $x = e$. Let r denote the significant solution, the one using $+\sqrt{\ }$ in (42), so $r > e$. At e, f^2 has a local minimum, while there are local maxima at both r and its companion point ((42) with $-\sqrt{\ }$), as can be seen from the curvature values:

$$\begin{aligned} (f^2)''(e) &= (2A)(0) + f'(m)(2A) = (-)(-) \ > 0 & (43) \\ (f^2)''(r) &= (2A)[f'(r)]^2 + (0)(2A) = (-)(+) \ < 0 & (44) \end{aligned}$$

(the second holds at the companion point as well). For the signs we note that $A < 0$ is assumed, and $m > e =$ point of zero slope for f, so f has negative slope at m.

The two local maxima of f^2 have the same value as the maximum of f: $f(f(r)) = f(e) = m$. For large $|x|$ the quartic $f^2(x)$ is asymptotic to A^3x^4 with $A < 0$, so r and its companion are in fact points of global maxima for f^2 when its domain is extended to the real line. Thus the horizontal line $x \equiv m$ touches all three maxima, one for f and two for f^2, the curves of f and f^2 otherwise lie below this line, and the local minimum of f^2 occurs midway between its maxima, at the same point e of even symmetry where f is maximum. The extended graph of f^2 is therefore verified to have a smooth symmetric "M-shape" in the cases we are now considering; see Fig. 12(v). The portions of the graphs of f and f^2 seen within the unit square may encompass one or more of the three extrema; this is evident in (i) through (iv) of Fig. 12. The $-\sqrt{\ }$ choice in (42) lies to the left of e and, for the cases being considered, does not play a role in determining crossings of the 45-degree line, so this companion point need not be mentioned further.

The local minimum of f^2 at e has value $f(f(e)) = f(m)$. In (i), (ii), and (v) of Fig. 12, this value appears to be positioned above e on the vertical line $y = e$ and therefore also above the 45-degree line, and in the "borderline" case (iii) the location is at the origin; formal verification is needed that other locations can be ruled out. This will be seen to be of importance in establishing non-chaotic behavior; in contrast, for the classical chaotic logistic map (19), the minimum of f^2 dips below the 45-degree line, as illustrated in Fig. 12(vi). To verify $f(m) > e$, we may equivalently show that $m < r$: both r and m lie to the right of e and thus are in the region of the x-axis where f has negative slope, so $m < r \iff f(m) > f(r)$. We calculate the difference $r - m$, the left side of (42) (+ sign) minus the right side of (31), and show that it is positive for the cases being considered:

$$\begin{aligned} r - m &= \frac{2B + 2\sqrt{B^2 - 2B - 4Ac} - B^2 + 4Ac}{-4A} = \frac{\sqrt{\beta} - (\beta/2)}{-2A} & (45) \\ \text{where } \beta &= B^2 - 2B - 4Ac = [(b-1)^2 + 4c(1-a)] - 1 = (1 - s_0)^2 - 1 & (46) \end{aligned}$$

Since the function $\sqrt{x} - x/2$ is positive for $0 < x < 4$, β should be checked to see if it lies in this range. Using (46) with (39)(37)(38), we see that $\beta \leq 5-1 = 4$ and $\beta > 1-1 = 0$, so β is indeed restricted as desired. Fig. 12(iii) shows a boundary case of $\beta = 4$ ($s_0 = 1 - \sqrt{5}$); (ii) illustrates $\beta = 3$ ($s_0 = -1$), (iv) depicts a typical intermediate case with $s_0 < -1$, and Fig. 13 repeats this case with labels and lines drawn to emphasize the relationships that have been established.

With Fig. 13 as a guide, the argument can now be completed. Let us trace the progress of the graph of f^2 from left to right. Beginning at its local minimum, which is left/above (e, e), f^2 passes through (p_0, p_0) and reaches its maximum below/right of the 45-degree line I. According to (40) (or (24)), the slope of f^2 at p_0 is

$$(f^2)'(p_0) = (f'(p_0))^2 \tag{47}$$

which is positive and greater than one precisely when $s_0 < -1$. In this case, the trace just described must progress from left to right by crossing I at (p_0, p_0) from below to above; therefore, there must be two intermediate crossings of I, at p_1 and p_2, positioned below and above p_0, and these two crossings of I must be left-to-right with positive slopes smaller than one (slope values greater than one at these points would require new intermediate critical points and/or crossings of f^2, but none can appear since all possibilities have previously been accounted for). This completes the argument as desired, verifying a small (positive and $<$ 1) slope value at the 2-cycle points whenever they appear. From (24), these two slopes have the common numerical value given by

$$(f^2)'(p_1) = (f^2)'(p_2) = f'(p_1)f'(p_2) \quad . \tag{48}$$

We observe that for strongly negative s_0-slope values, the points e and r are typically near the left and right ends inside the interval [0,1] as in Fig. 13, but these points could also fall at the ends as in Fig. 12(iii), or even slightly outside the interval to the left and right, within the coefficient constraints. In these cases,the conclusions reached regarding 2-cycle slopes would obviously continue to be valid (e.g., the x axis of the template Fig. 13 could be extended to accommodate the points and the same constructions would hold).

9 Remarks on Multidimensional Maps

In a lengthy series of investigations beginning in the late 1950s, Ulam and Stein [Ste93] considered many examples of polynomially nonlinear multidimensional mappings, using exploratory computational visualization. In our terminology, these examples can be seen to be finite-state stochastic transformations realizable by 3-state or 4-state deterministic arrays with degrees $d = 2$ or $d = 3$. It was observed, within the limits of the experimental computational tools available at the time of these studies, that quadratic mappings all seemed to lead to non-chaotic behavior, but that there were interesting degree-3 examples which seemed to be chaotic (instead of the more recent term "chaos," a richly complex trajectory or "mess" was precisely the same idea referred to in that work). One of these examples is

$$x' = 6xyz + 3xz^2 + 3yz^2 + 3y^2z + z^3 \tag{49}$$
$$y' = x^3 + y^3 + 3x^2z \tag{50}$$
$$z' = 3xz^2 + 3x^2z \tag{51}$$
$$[\ z = 1 - x - y,\ \ z' = 1 - x' - y'\ ]$$

(the state-probability vector $(x_1, x_2, x_3) = (x, y, z)$ here). This is easily seen to be realizable by a 3-state 4-array which is deterministic (three $3 \times 3 \times 3$ sheets, one each for the x, y, and z dimensions, with all 0 and 1 element values, obtained from inspection of the defining equations above; e.g., $a_{---1} = 1$ for $---=$ one of the 6 permutations of $\{1, 2, 3\}$, etc.). There are three coordinate dimensions but the actual geometric dimension is two because of the stochastic constraint, so there are various choices of axes for plotting purposes; two of the three xyz axes can be shown, in a planar plot of, e.g., the x and y iteration values, or a barycentric coordinate transformation of the problem to a triangular region can be used. A typical (x, y) orbit chart is shown in Fig. 14. The orbit is plotted as a collection of points (small crosses); a complex "attracting basin" is revealed. The data are reused in Fig. 15, showing a "trajectory" version of the chart, in which successive (x, y) points are connected with straight lines; this is redundant, but serves visually to emphasize the "motion" of the trajectory—the fact that successive points can be widely separated—another view of the complexity of the behavior. The attracting basin shown is typical for some ranges of initial (x, y, z) values; other starting points can yield patterns appearing to be concentrated on straight lines but still richly complex.

This is generally regarded as a chaotic example, although it remains to prove exact conformance with some choice of formal definition. An acceptable working concept of multidimensional chaos needs to be specified more precisely before these and other examples can be classified fully. It should be kept in mind that some of the tools available in one dimension are specific to that situation and do not have exact counterparts in more dimensions; e.g., there is no analog of the Sarkovskii ordering, so one cannot depend on an appeal to such a general theorem about periodicities. An approach based on studying projections onto the axes may have some merit; however, relationships among projections, one-dimensional iterations, and multidimensional iterations then need to be clarified.

10 Conclusions and Conjectures

In summary, polynomially nonlinear interaction, as modeled by a finite-state probabilistic array, has strong potential for chaos, except for small numbers of states and low degree. In particular, 2-state or one-dimensional quadratic maps are guaranteed to be nonchaotic, while there exist 2-state degree-5 chaotic examples and many of higher degree. There remains a small gap to investigate in one dimension, since no chaotic 2-state arrays of degree 3 or 4 have been found;

based on investigation of examples, we conjecture that, at least for degree 3, nonchaoticity prevails. The degree-4 example $x^4 + (1-x)^4$ appears regular, with a period 2 (see Figs. 1,2,3,4 with comments in Section 5); however, degree 4 remains a speculative borderline situation (the chaotic d=5 example (22) is actually a degree-4 polynomial, although as already noted the system as a whole is still properly of degree 5). For more dimensions or states, there are apparently chaotic examples of degree as low as 3 (as in the Ulam/Stein map discussed above); i.e., there may be a dimension(states)/degree tradeoff in achieving a level of complexity necessary for prevalence of chaos. In general, as the array size (combining states and degree) increases, chaotic behavior is not only possible but may be typical in a variety of situations. Details of the chaotic/nonchaotic dichotomy remain to be explored; e.g., are there simply-expressed constraint conditions (or classes of arrays) that ensure non-chaotic or chaotic behavior?

References

[Ban92] J. Banks *et al*, "On Devaney's definition of chaos," *Amer. Math. Monthly*, v99#4, 1992.

[Bus90] S. Buss, C. Papadimitriou, J. Tsitsiklis, "On the predictability of coupled automata: an allegory about chaos," *IEEE Proc. 31st Symp. FOCS*, 1990.

[Dev89] R. Devaney, *An Introduction to Chaotic Dynamical Systems, 2nd ed.*, Addison-Wesley, 1989.

[Far92] O. Farhangi, "Stochastic arrays and chaos," M.S. Thesis, Computer Science Department, UCLA, 1992.

[Fei80] M. Feigenbaum "Universal behavior in nonlinear systems," *Los Alamos Science*, v1#1 pp4–27, 1980.

[Kle75] L. Kleinrock, *Queueing Systems, Vol. 1: Theory*, J. Wiley, 1975. (§ 2.3.)

[May76] R. May, "Simple mathematical models with very complicated dynamics," *Nature*, v261 pp459–467, 1976.

[Paz68] A. Paz, "Probabilistic arrays," *Proc. Princeton Conf. on Inf. Sci & Sys.*, IEEE-68-CT-1 pp194–199, 1968.

[Pea91] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, revised prtg., Morgan Kaufmann, 1991. (See § 4.3.1.)

[Pei92] H-O Peitgen, H. Jürgens, D. Supe, *Chaos and Fractals*, Springer, 1992.

[Rab95] Y. Rabani, Y. Rabinovich, A. Sinclair, "A computational view of population genetics (prel. vers.)," *ACM Proc. 27th STOC*, pp83–92, 1995.

[Ste89] P. Stein, "Iteration of maps, strange attractors, and number theory—an Ulamian potpourri," in *From Cardinals to Chaos*, N. Cooper, ed., Cambridge, 1989.

[Tou97] P. Touhey, "Yet another definition of chaos," *Amer. Math. Monthly*, v104#5 pp 411–414, 1997.

Appendix (Details, Calculations)

Array products; nonassociativity. A general n-state k-array $A = \| a_{i_1 i_2 \cdots i_k} \|$ has indexed elements (which may be from an arbitrary commutative algebra, i.e., not necessarily real numbers satisfying probabilistic constraints), with each of the k indexes ranging from 1 to n. If A is an n-state k-array and B is an n-state ℓ-array, then the *product* $C = AB$ is defined to be the n-state k-array with elements

$$c_{i_1 i_2 \cdots i_k} = \sum_{1 \le j_1, \ldots, j_{\ell-1} \le n} \left\{ \prod_{p=1}^{\ell-1} a_{i_1 i_2 \cdots i_{k-1} j_p} \right\} b_{j_1 j_2 \cdots j_{\ell-1} i_k} \ .$$

To see that this is a nonassociative product, consider the 2-state 3-array

$$A = \left(\begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix} \right)$$

and the 2-state 1-array (i.e., 2-vector) $v = (x, y)$. Briefly summarizing results obtainable from direct applications of the array product definition:

$$\begin{aligned} vA &= (x^2 + xy + y^2, xy) = (\text{quadratic (in } x \ \&\ y), \ \text{quadratic}) \\ \text{and } ((vA)A) &= (\text{quartic}, \text{quartic}); \\ AA &= \text{another } 2-\text{state } 3-\text{array}, \\ \text{so } \ (v(AA)) &= \text{another } (\text{quadratic}, \text{quadratic}) \neq ((vA)A) \ . \end{aligned}$$

Polynomials; collected powers; coefficients. We may write (8) in an alternative form which displays the powers of the x_i explicitly:

$$x'_j = \sum_{\substack{0 \le d_1, \ldots, d_n \le d \\ d_1 + \cdots + d_n = d}} x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n} c_{d_1 d_2 \cdots d_n j} \ . \tag{52}$$

The a- and c- coefficients must satisfy

$$c_{d_1 d_2 \cdots d_n j} = \sum_{\substack{d_k \text{ of } i_1, \ldots, i_d = k \\ k = 1, 2, \ldots, n}} a_{i_1 i_2 \cdots i_d j} \ , \tag{53}$$

$$\sum_{j=1}^{n} c_{d_1 d_2 \cdots d_n j} = \frac{d!}{d_1! d_2! \cdots d_n!} \ . \tag{54}$$

The ranges that are possible for a-values realizing a given mapping can be deduced using (53) and (54). Formula (53) follows immediately from comparison of the polynomials (8) and (52). To obtain (54) consider the identity

$$x'_1 + x'_2 + \cdots + x'_n = 1 = (x_1 + x_2 + \cdots + x_n)^d \tag{55}$$

(i.e., components of probability vectors must sum to 1). Substituting (52) in the left side of (55), using the usual multinomial expansion on the right side, and equating coefficients of like powers, yields (54). In other words, the left side of (55) simply represents a partitioning of the multinomial expansion of the right side; this suggests an easy method for constructing examples. In one dimension (two states), (14) amounts to a specialization of (54), and the relationships reduce to the simple observation that in the term $c_k x^k (1-x)^k$, the value of c_k (which is itself a sum of a-values) must lie between zero and the binomial coefficient $\binom{d}{k}$.

Degree-9 chaotic example derived from logistic map. The numbers $c_0, \ldots, c_9$ must be upperbounded by 1, 9, 36, 84, 126, 126, 84, 36, 9, 1. Expand the left side of (14) to express the 10 C_k as linear combinations of the c_k (i.e., $C_0 = c_9$, $C_1 = c_8 - 9c_9, C_2 = c_7 - 8c_8 + 36c_9$, etc.). Attempt to choose nonnegative c_k values so that their upperbounds are satisfied and the C_k vanish except for $C_1 = 3.6, C_2 = -3.6$. Solution: $c_0, \ldots, c_9 = 0, 3.6, 25.2, 75.6, 126, 126, 75.6, 25.2, 3.6, 0$. Degree 9 is the least for which this simple forced fit of $3.6x(1-x)$ is successful; lower degrees are easily checked, by the same direct-comparison approach, and found to be infeasible.

Formal vs. actual degree. The degrees of some of the n polynomials defined in (8) may be less than d, after simplification by collection of terms with like powers as in (52). This suggests that there may sometimes be a need to distinguish between the formal degree d of the polynomial collection (8) and the actual degree ($\leq d$) of a particular evaluated polynomial. However, the sum over j in (8) must be 1, and this fact, combined with (10), shows that not all of the n polynomials can have degree less than d; i.e., the formal degree d of a system (8) always agrees with the actual degree of one or more of its component polynomials, so in this sense it is unambiguous to say that the degree of a probabilistic system is d. This can also be seen by considering the consequences of the constraint (55).

Numerical values for the degree-5 example (22). Fixed points and periods are easily computed by applying a standard numerical rootfinding routine (e.g., false position) to the functions $f^k(x) - x$, after inspection of their graphs (see Fig. 8) to select appropriate subintervals for initial estimates. The results are as follows, rounded to adequate precision for the illustration (see also [Far92]). Fixed point of f: 0.2454621. New fixed points of f^2: 0.0939116 and 0.6107423, a single 2-cycle. New fixed points of f^3: none. New fixed points of f^4: 0.0730441, 0.1532676, 0.4353266, 0.6843791; numbering these points in order, and evaluating f at each point, we see that the 4-cycle is: $1 \to 4 \to 2 \to 3$. New fixed points of f^5: none. There are 12 new fixed points of f^6: 0.0632016(#1), 0.0650534, 0.1186134, 0.1353771, 0.1879693, 0.1971804, 0.3337936, 0.3533058, 0.4832531, 0.5319263, 0.7143889, 0.7214919(#12), which constitute two 6-cycles:

$$1 \to 12 \to 6 \to 7 \to 4 \to 9; \quad 2 \to 11 \to 5 \to 8 \to 3 \to 10.$$

Evaluating f' at the periodic points and using (24), we obtain $(f^6)'(p) = -2.28$ for p in the first 6-cycle, and -4.05 for the second; likewise, slopes for 2-cycles and 4-cycles are -1.96 and -3.26, and the slope of f at its fixed point is -1.6, so all of these points are repelling (slope absolute values exceed 1), as expected.

Figure 4: $f(x)=x^4+(1-x)^4$, sorted chart of 500 iterations, starting from x=0.98.

Figure 5: $f(x)=x^5+(1-x)^5$, diagonal phase portrait of 100 iterations starting from x=0.3.

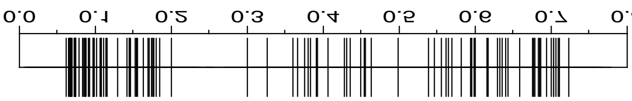

Figure 6: $f(x)=x^5+(1-x)^5$, orbit chart, 100 iterations starting from x=0.3.

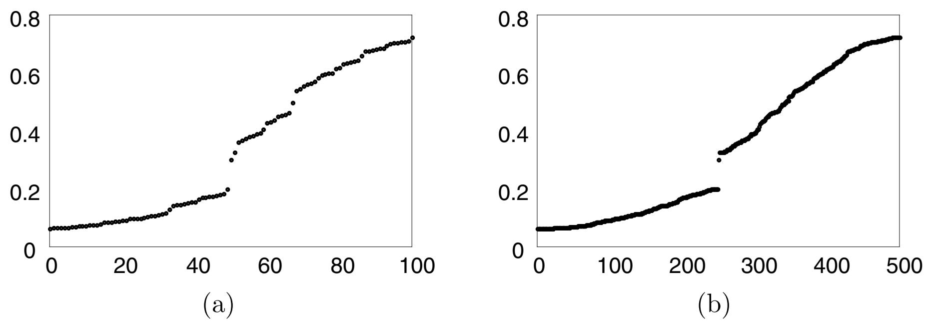

Figure 7: $f(x){=}x^5{+}(1{-}x)^5$, sorted charts, starting from $x{=}0.3$; 100 iterations shown in (a), 500 iterations in (b).

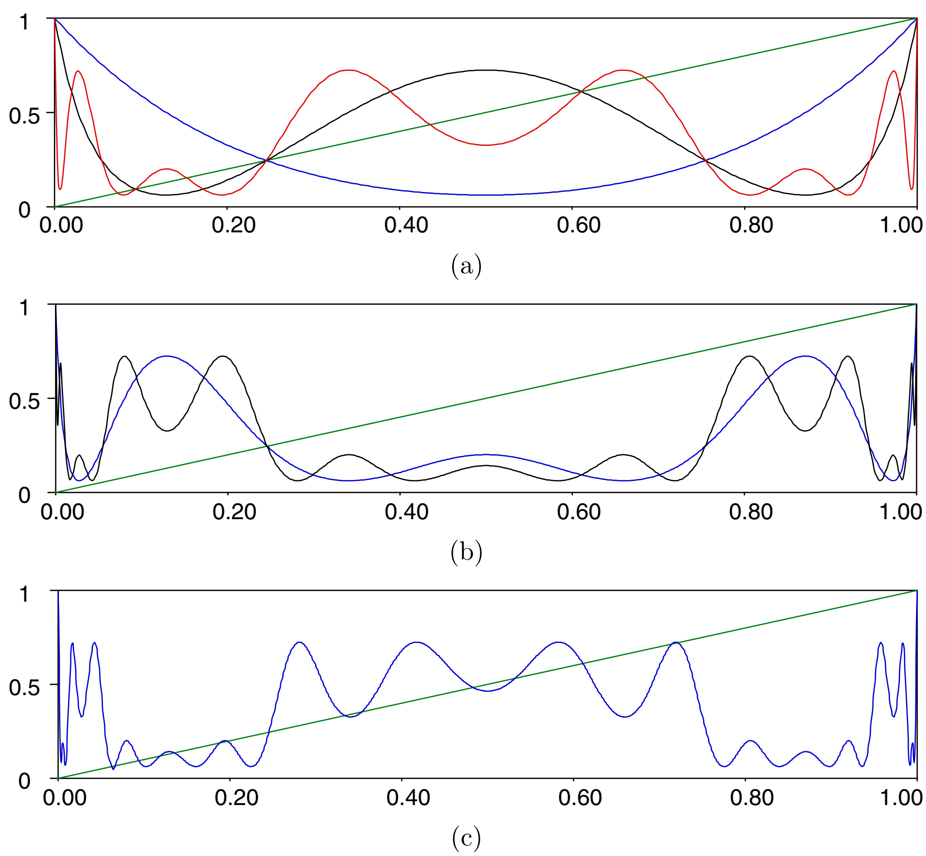

Figure 8: $f(x){=}x^5{+}(1{-}x)^5$, plots of f and its iterates up to the 6^{th}. (a): f, f^2, f^4. (b): f^3, f^5. (c): f^6.

(iv) (v) (vi)

Figure 12: plots of $f(x)$ and $f(f(x))$ for the quadratics $f(x) = ax^2 + bx(1-x) + cx^2$. Values of (a, b, c): $(13/40, 1.75, 5/8)$ in (i); $(1/8, 2, 6/7)$ in (ii); $(0, 2, 1)$ in (iii); $(0.02, 2, 0.88)$ in (iv) and (v); $(0, 3.6, 0)$ in (vi) ((vi) is the logistic map with K=3.6).

Figure 15: As above; successive (x, y) points joined with lines; 100 iterations shown.