Notes on Arnold cat map symbolic dynamics

Li Han

School of Physics, Georgia Institute of Technology, Atlanta, GA 30332

In this notes we describe the computer experiment results of symbolic dynamics of Arnold cat map using some linear encoding.

I. SYMBOLIC DYNAMICS BACKGROUND

Usually we study the evolution of a dynamical system by specifying the dynamics rule and initial conditions. When the dynamics is chaotic, rather than working with initial conditions of real-valued vectors in phase space one can instead partition the phase space into several regions and study the sequence of regions the system visits during time evolution.

The following calculator experiment is a good example to illustrate the idea of symbolic dynamics above: Start with an arbitrary positive generic number, say 20000. Press the "Ln" (natural log) button again and again until the number becomes negative, then press the "+/-" button to make it positive again. This essentially realizes the mapping $x \mapsto \ln(|x|)$, and the initial value 20000 corresponds to the iteration sequence

$$\{20000., 9.90349, 2.29289, 0.829812, -0.186556, -1.67902, 0.518211, -0.657373, \cdots\}. \tag{1}$$

Equivalently, we can only write down the positions at which we insert a minus sign. That is, the real-valued initial condition can be rewritten as an (infinite) sequence from an alphabet of binary symbols $\{m_1 m_2 \cdots\}$, $m_i \in \{+, -\}$. The dynamics extracts 1 more bit of information contained in the initial condition at each time step. This mapping is chaotic, and at sufficiently large length the symbol sequences will differ for arbitrarily close initial conditions. For example,

$$\begin{aligned} 20000 &= \{++++--+-----+----+--++-+--++-++\cdots\} \\ 20001 &= \{++++--+-----+----+--++-++-+--+-\cdots\}. \end{aligned}$$

In this simple example we have used two symbols $\{``+" \equiv \{x|x>0\}, ``-" \equiv \{x|x<0\}\}$ to construct the symbolic dynamics of $f : x \mapsto \ln(|x|)$. Furthermore, all symbol sequences $\{m_1 m_2 \cdots\}$ are admissible (at fixed length n there are a total of 2^n of them and each occurs with a non-zero probability). This is not the case for general dynamics as we will see for the Arnold cat map discussed next.

II. SYMBOLIC DYNAMICS OF ARNOLD CAT MAP

The Arnold cat map is an area-preserving and piecewise linear map of the unit torus onto itself. It is the action by a 2×2 matrix $A \in SL(2, \mathbb{Z})$ modulo the integer lattice. Denote $s = \mathrm{Tr}A$. For the mapping to be chaotic we need $s > 2$ and the original Arnold's cat map refers to $s = 3$. For our purpose it is convenient to choose the phase space variable as (q_{t-1}, q_t), where t is the discrete time index. The single cat map then corresponds to $A = \begin{pmatrix} s & -1 \\ 1 & 0 \end{pmatrix}$ and is equivalent to the 2-step recurrence relation

$$q_{t+1} = sq_t - q_{t-1} \bmod 1, \tag{2}$$

where $q_t \in [0, 1)$ for all t. Such definition naturally leads to the symbolic encoding

$$m_t \equiv sq_t - (q_{t-1} + q_{t+1}) \tag{3}$$

as m_t is always an integer taking values in $\{-1, 0, \cdots, s-1\}$. The term linear coding follows from the formal linear mapping between phase space trajectories $\{q_t\}_t$ and symbol sequence $\{m_t\}_t$.

In Fig. 1 we show how the above linear encoding m_t partitions the phase space into 4 regions $m = \{-1, 0, 1, 2\}$ for the original Arnold cat map ($s = 3$) (the top-left panel), as well as their time evolution under the map (the other panels). In Fig. 2 we show the phase space regions corresponding to a sequence of symbols of length 2, 3, 4 respectively, which are obtained from taking the superpositions of consecutive frames $\{m_0, m_1\}$, $\{m_0, m_1, m_2\}$, $\{m_{-1}, m_0, m_1, m_2\}$ from Fig. 1 respectively.

The symbolic sequences give rise to an interesting counting problem. The bi-infinite symbol sequence $\{m_t\}_{t=-\infty}^{\infty}$ encodes all the information about the dynamics and can be used to recover the phase phase trajectory. A symbol sequence $\{m_1 m_2 \cdots m_\ell\}$ of finite length ℓ provides an approximation to the trajectory and the error is exponentially small in ℓ. However, not all of the $(s+1)^\ell$ sequences, i.e., length-ℓ words constructed from the alphabet of $(s+1)$ symbols are admissible. Certain words are pruned and the admissible ones are known as the gramar rules.

The numbers $N_\ell(s)$ of total pruning rules are listed in Table I. They are calculated through a lengthy computation on the computer. The missing entries are those taking significantly long time (order of week(s)) on a personal computer. For each given ℓ the number $N_\ell(s)$ as a function of $s = \mathrm{Tr}A$ satisfies an empirical polynomial law of order $\ell - 1$, while there is no simple rule for $N_\ell(s)$ as a function of the sequence length ℓ.

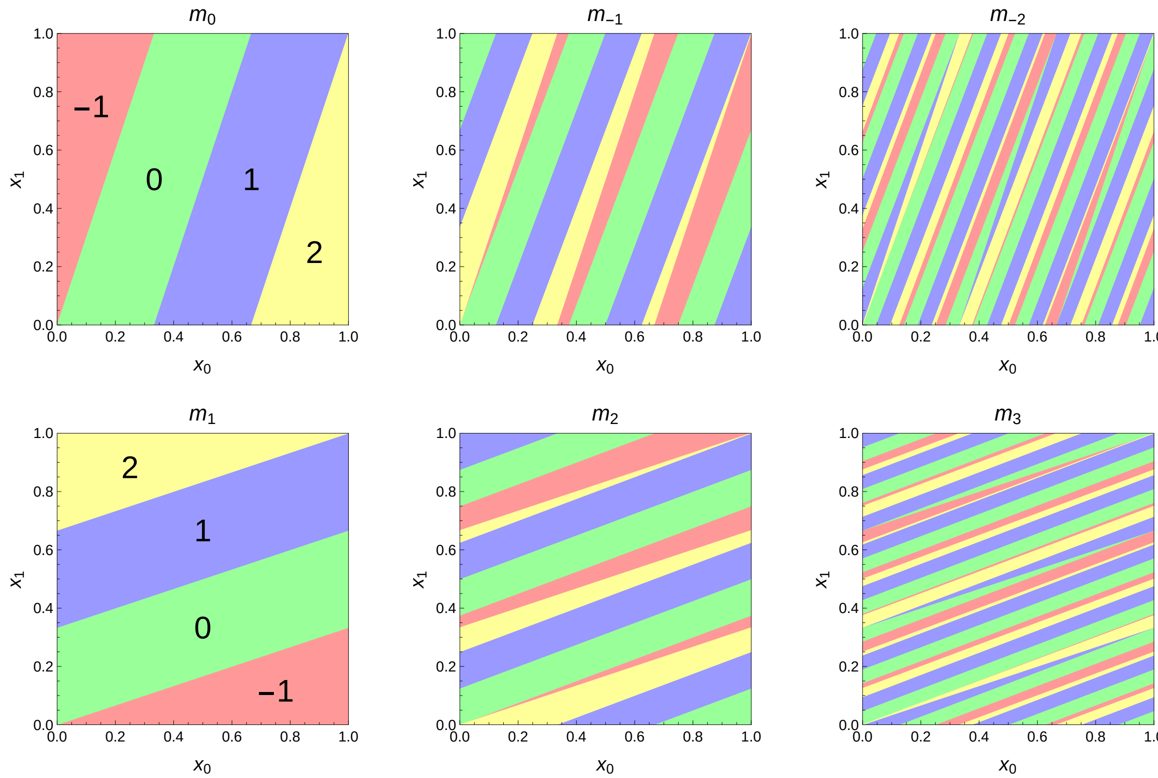

FIG. 1: Symbol regions of Arnold cat map using linear encoding m_t: the different panels show the time evolution (forward and backward) of the symbol regions under the map.

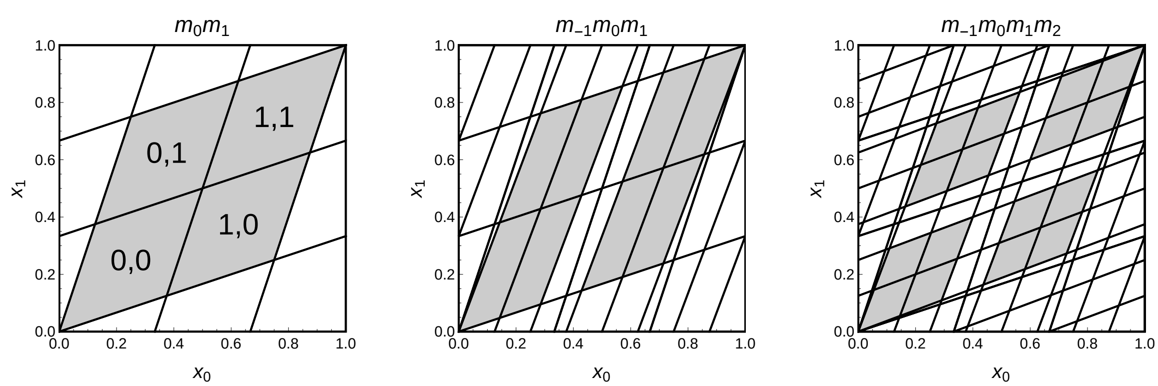

FIG. 2: From left to right: phase space regions corresponding to a sequence of 2 (left), 3 (middle), and 4 (right) symbols.

Note that the number $N_\ell(s)$ counts those rules that are already pruned at length $\ell' < \ell$. Therefore, it is reasonable to look at the number of new pruning rules $\tilde{N}_\ell(s)$, which is the number of length ℓ forbidden words exclusively for ℓ: all of their length $\ell' < \ell$ substrings are admissbile. In Table II we list the numbers $\tilde{N}_\ell(s)$ from available data.

s	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N
3	0	2	22	132	684	3164	13894	58912	244678	1002558	4073
4	0	2	32	254	1714	10264	58006	315218	1671326	8712716	4489
5	0	2	44	436	3616	26548	183390	1216280	7856850	49821822	3117
6	0	2	58	690	6780	58916	480892	3766162	28702960	214561416	15814
7	0	2	74	1028	11668	117076	1102342	9955928	87460838	753246422	
8	0	2	92	1462	18814	214024	2283234	23361722	232435698		
9	0	2	112	2004	28824	366524	4369006	49947328	555149746		
10	0	2	134	2666	42376	595588	7844920	99101450			
11	0	2	158	3460	60220	926956	13369542	184939952			
12	0	2	184	4398	83178	1391576	21811822	327903298			
13	0	2	212	5492	112144	2026084	34291774				
14	0	2	242	6754	148084	2873284	52224756				
15	0	2	274	8196	192036	3982628	77369350				
16	0	2	308	9830	245110	5410696	111878842				
$N_\ell(s)$	0	2	s^2+3s $+4$	$2s^3+6s^2$ $+6s+6$	$3s^4+11s^3$ $+13s^2+7s$ $+6$	$4s^5+17s^4$ $+24s^3+15s^2$ $+8s+8$	$5s^6+24s^5$ $+41s^4+33s^3$ $+18s^2+11s$ $+10$	$6s^7+32s^6$ $64s^5+65s^4$ $48s^3+33s^2$ $14s+10$			

TABLE I: Number of total pruning rules $N_\ell(s)$ for different $s = \mathrm{Tr}A$ and sequence length ℓ. The last row shows the empirical fit of N_ℓ as a function of s.

Take the $s = 3$ case (first row in each table) for example: there are four symbols $\{-1, 0, 1, 2\}$. At length 1 all symbols can occur with probabilites $1/6, 1/3, 1/3, 1/6$ respectively (the top-left panel of Fig. 1). That gives the number 0 and 0 in both Table I and Table II. At length 2 there can be a total of $4^2 = 16$ symbol sequences, while two of them $\{-1, -1\}$, $\{3, 3\}$ are not admissible (left panel of Fig. 2). That gives the number $N_2 = 2$ in Table I and $\tilde{N}_2 = 2$ in Table II. At length 3 there are more non-adimissible sequences, which include those already pruned at length 2, e.g., any length-3 sequence that contains $\{-1, -1\}$ as a subsquence. There are a total of 22 such non-adimissible sequences at length 3 and that's the number $N_3 = 22$ in Table I. Among them 8 are new: $\{-1, 0, 1\}$, $\{0, -1, 0\}$, $\{0, -1, 1\}$, $\{0, 2, 1\}$ and $2 - \{\cdots\}$ (those obtained by reflection symmetry). This is the number of new pruning rules $\tilde{N}_3 = 8$ in Table II.

The behavior of $\tilde{N}_\ell(s)$ depends on the parity of ℓ. For even ℓ, an interesting observation is that $\tilde{N}_\ell(s) = 2$ except for the anomaly at $\ell = 8$. The 2 pruning rules are $\{-1, (\ell-2) \text{ of } 0, -1\}$

s	$\tilde N_1$	$\tilde N_2$	$\tilde N_3$	$\tilde N_4$	$\tilde N_5$	$\tilde N_6$	$\tilde N_7$	$\tilde N_8$	$\tilde N_9$	$\tilde N_{10}$	$\tilde N_{11}$	$\tilde N_{12}$	$\tilde N_{13}$	$\tilde N_{14}$	$\tilde N_{15}$	$\tilde N_{16}$	$\tilde N_{17}$	$\tilde N_{18}$
3	0	2	8	2	30	2	70	16	198	2	528	2	1326	124	3410	2	9264	2
4	0	2	14	2	72	2	248	30	968	2	3620	2	13288	446				
5	0	2	22	2	140	2	630	48	3110	2	14876	2						
6	0	2	32	2	240	2	1328	70	7908	2	45976							
7	0	2	44	2	378	2	2478	96	17262	2								
8	0	2	58	2	560	2	4240	126	33808									
9	0	2	74	2	792	2	6798	160	61038									
10	0	2	92	2	1080	2	10360	198										
11	0	2	112	2	1430	2	15158	240										
12	0	2	134	2	1848	2	21448	286										
13	0	2	158	2	2340	2	29510											
14	0	2	184	2	2912	2	39648											
15	0	2	212	2	3570	2	52190											
16	0	2	242	2	4320	2	67488											
$\tilde N_\ell(s)$	0	2	s^2-s+2	2	s^3+s^2-2s	2	$s^4+\frac{2}{3}s^3-3s^2-\frac{2}{3}s$	$2s^2-2$	$s^5+\frac{2}{3}s^4-3s^3-\frac{8}{3}s^2+2s$									

TABLE II: Number of new pruning rules $\tilde N_\ell(s)$ for different $s = \mathrm{Tr}A$ and sequence length ℓ. The last row shows the empirical fit of $\tilde N_\ell$ as a function of s. Note the "anomaly" at $\tilde N_8$ and the fractional fit for $\tilde N_7$.

and $\{s-1, (\ell-2) \text{ of } s-2, s-1\}$, which are related by an inversion symmetry of the cat map. They are indeed present for all ℓ including the odd ones and 8. For odd ℓ the empirical fit of $\tilde N_\ell(s)$ is again a polynomial of s, with a leading order of $(\ell+1)/2$. The empirical polynomial of $\tilde N_8(s)$ versus s has order 2. Furthermore, the fitting polynomial $\tilde N_7(s)$ has fractional coefficients while being even. Similar to the case of $N_\ell(s)$, for a given s there is no simple rule for the dependence of $\tilde N_\ell(s)$ on ℓ in each parity class of ℓ.

The natural question to ask is what are the explicit forms of total pruning rules $N_\ell(s)$ and new pruning rules $\tilde N_\ell(s)$. Here we make two reasonable guesses on the two quantities:

Total pruning rules: the large ℓ asymptotics of $N_\ell(s)$ satisfies that

$$(s+1)^\ell - N_\ell(s) \sim \lambda^\ell, \quad \lambda = \frac{s+\sqrt{s^2-4}}{2},$$

where the LHS is the total number of admissible rules and λ is the bigger eigenvalue of the $SL(2,Z)$ matrix of the cat map (the Lyapunov exponent of the mapping is $h = \ln\lambda$). For

Arnold cat map ($s = 3$), $\lambda = \frac{3+\sqrt{5}}{2} \approx 2.618$. That is, the dynamics produces an amount of information corresponding to roughly 2.618 symbols at each time step.

New pruning rules: the lengths ℓ at which $\tilde{N}_\ell(s)$ displays the "anomalous" behavior happen at a periodicity of 6, i.e. at $\ell = 2, 8, 14, 20, \cdots$. Note at all even lengths $\ell = 2m$ there are always 2 new pruning rules $\{-1, 0, 0, \cdots, 0, -1\}$ and (reflection symmetry related sequence) $\{s-1, s-2, s-2, \cdots, s-2, s-1\}$). At $\ell = 2$ the anomalous new pruning rules are just vanishing. This periodicity 6 might be related to the fact that the Eisenstein series E_{2k} when expressed in terms of E_4 and E_6 also shows a periodicity of 6 in number of basis.

---

[1] Boris Gutkin, Predrag Cvitanović, Li Han, Rana Jafari, and Adrien K Saremi, *Spatiotemporal symbolic dynamics for coupled cat map lattices*, Work in progress.

[2] Ian Percival and Franco Vivaldi, *A linear code for the sawtooth and cat maps*, Physica D **27**, 373 (1987).

[3] Predrag Cvitanović, Chaos, Classical and Quantum, http://chaosbook.org.