Space-Time Phases

Statistical Properties of Dynamics on Large Networks

Lecture notes by

Prof. Robert S. MacKay

prepared with the assistance of Dayal Strub

UNIVERSITY OF WARWICK

WARWICK MATHEMATICS INSTITUTE

April, 2011

Contents

Preface

Complexity science is the study of systems with many interdependent components. One of the main concepts is "emergence": the whole may be greater than the sum of the parts. The objective of this short course is to put emergence on a firm mathematical foundation in the context of dynamics of large networks. both stochastic and deterministic dynamics are treated. To minimise technicalities, attention is restricted to dynamics in discrete time, in particular to probabilistic cellular automata and coupled map lattices. The key notion is *space-time phases*: probability distributions for state as a function of space and time that can arise in systems that have been running for a long time. I say what emerges from a complex dynamic system is one or more space-time phases. The *amount of emergence* in a space-time phase is its distance from the set of product distributions over space, using an appropriate metric. A system exhibits *strong emergence* if it has more than one space-time phase. Strong emergence is the really interesting case.

These lecture notes are based on MSc or PhD courses I have given in Warwick in 2006/7, Paris April 2007, Warwick Spring 2009 and Autumn 2009, and Brussels Autumn 2010, and were written up during study leave in 2010/11 at the Université Libre de Bruxelles, to whom we are grateful for hospitality.

They provide an introduction to the theory of space-time phases, via some key examples of complex dynamic system.

Chapter 1

Stochastic Dynamics - Probabilistic Cellular Automata (PCA)

Although there is a wide range of types of stochastic dynamic system, we will restrict attention here to probabilistic cellular automata (PCA). These are discrete-time Markov processes on large products of finite-state units, for which the transitions at different sites given the whole current configuration are independent. It is usual, but not essential, to assume that the transition probability for the state of site s is independent of state off a finite set $N(s)$ called its neighbourhood.

In this chapter we will study two main examples: Stavskaya's PCA and the NEC majority voter PCA, but we will derive more general results in the process of treating them.

Some suggested background reading is [7] for a short introduction and [26] for an extended treatment.

1.1 Stavskaya's PCA

Stavskaya's PCA [24] is also known as directed percolation, oriented percolation, asymmetric contact process, discrete Regge field, Domany-Kinzel model, and possibly more names. It is a PCA in 1D, in which each unit at each time is in one of two states: 1 representing healthy and 0 representing infected. Each healthy unit can catch infection from its left hand neighbour only.

The state is $x = (x_s)_{s\in\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}}$. The parameter $\lambda \in (0,1)$, is both the probability of recovery in one step and the probability of avoiding infection from one's left neighbour (the probabilities can be taken different if desired, and also site-dependent, and recovery rate could be taken to depend on whether the neighbour is infected, but to reduce notational complexity we restrict to this case). At each time $t \in \mathbb{Z}_+$, for each site $s \in \mathbb{Z}$, if $\left(x^t_{s-1}, x^t_s\right) = (1,1)$ then $x^{t+1}_s = 1$, else if $\left(x^t_{s-1}, x^t_s\right) \neq (1,1)$ then $x^{t+1}_s = 1$ with probability λ, and 0 with probability $1-\lambda$, independently of all other $x^{t'}_{s'}$. Let $\underset{\sim}{1}$ denote the state of all 1, which is an absorbing state, and $\delta_{\underset{\sim}{1}}$ the probability distribution supported on $\underset{\sim}{1}$.

For the finite version of this system on $\mathbb{Z}_N = \mathbb{Z} \mod N$, we have the probability

$$P\{x^{t+1} = \underset{\sim}{1} | x^t\} \geq \lambda^N,$$

for all x^t (x^t means the state of the whole line at time t; we sometimes denote it by $\underset{\sim}{x}^t$ to emphasise that it is the state of the whole system). Therefore $P\{\text{not absorbed by time } t\} \leq \left(1-\lambda^N\right)^t \to 0$ as $t \to \infty$, and so $P\{\text{eventually } \underset{\sim}{1}\} = 1$.

This discussion of the finite case misses an important feature, however, namely that there exists a critical parameter value $\lambda_c \in (0,1)$, such that for $\lambda > \lambda_c$, the time to absorption is of order $\log_\gamma N$ for some $\gamma(\lambda) > 0$, and for $\lambda < \lambda_c$, the time to absorption is instead of of order $e^{\gamma N}$. This is best understood by considering the infinite system, for which the following hold:
for $\lambda \geq \lambda_c$, stationary probability measure $\delta_{\underset{\sim}{1}}$ attracts all initial probability measures,
for $\lambda < \lambda_c$, there exists an additional stationary probability ν_λ with $\nu_\lambda(x_s = 0) = c(\lambda) > 0$ ("endemic infection" or "eternal transient") and for all initial probabilities ν, we have

$$\nu P^t \to \sigma \delta_{\underset{\sim}{1}} + (1-\sigma)\,\nu_\lambda,$$

where $\sigma = \nu\{\text{eventual absorption}\}$, and P is the transition operator for the process (definition to be recalled in Section 1.1.1).

We will first treat λ near 1 and then λ near 0.

1.1.1 λ near 1

This is the regime of low infectivity and high recovery rate, so we expect any initial infection to die out. We will introduce a nice metric on a space of multivariate probabilities to prove (exponential) attraction of νP^t to $\delta_{\underset{\sim}{1}}$ for any initial probability ν, based on ideas of Dobrushin [8]. See [17].

We introduce the metric in the following general setting. Let S be a countable set, the "network" (on which we can have a metric, but it is not required at this stage). An element $s \in S$ is called a "site", and for all $s \in S$, (X_s, d_s) is a Polish (= complete separable metric) space of diameter $\leq$ some Ω. For example, let $S = \mathbb{Z}$, $X_s = \{0, 1\}$ and $d_s(0, 1) = 1$, $X = \prod_{s\in S} X_s$ with product topology[1]. Let $\mathcal{P}$ be the set of Borel probabilities[2] on X, and Z be the set of zero-charge measures[3] on X. The transition probability for updating the state of site s is denoted $p_s(x'_s|\underset{\sim}{x})$ (or $p_s(\mathrm{d}x'_s|\underset{\sim}{x})$ if X_s is not discrete), and their product is denoted $p(\mathrm{d}\underset{\sim}{x}'|\underset{\sim}{x})$.

Let BC denote the set of bounded continuous functions $f : X \to \mathbb{R}$. The transition operator P on BC is defined by

$$(Pf)(x) = \int f(x')p(\mathrm{d}x'|x).$$

Note that $P1 = 1$, where $1(x) = 1, \forall x \in X$.

The transition operator P can be extended to act on measures ρ on X by

$$(\rho P)(f) = \rho(Pf),$$

where $\rho(f)$ is the integral of f with respect to ρ. We write P as acting on the left on measures to fit with standard matrix notation, considering measures as row-vectors.

A *stationary probability* is $\rho \in \mathcal{P}$ for which $\rho P = \rho$.

For $f \in BC$, $s \in S$, let

$$\Delta_s(f) = \sup \frac{f(x) - f(x')}{d_s(x_s, x'_s)},$$

over $x_s \neq x'_s$, $x_r \neq x'_r$ $\forall r \neq s$, i.e. the best Lipschitz constant of f with respect to the state x_s on s (if finite). Let $F = \{f \in BC : \|f\|_F = \sum_{s\in S} \Delta_s(f) < \infty\}$ modulo addition of constants, which is a normed linear space. For $\mu \in Z$, let

$$\|\mu\|_Z = \sup_{f\in F\setminus C} \frac{\mu(f)}{\|f\|_F},$$

where C is the set of constant functions. Note that $(Z, \|\cdot\|_Z)$ is a Banach space (complete normed linear space), because it is the dual of a normed linear space.

Finally we define

$$D(\rho, \rho') = \|\rho - \rho'\|_Z,$$

for $\rho, \rho' \in \mathcal{P}$, which makes $\mathcal{P}$ a complete metric space. We call D *Dobrushin metric.*

For $L : Z \to Z$ linear, e.g. the restriction of P to Z, let

$$\|L\|_Z = \sup_{\mu\in Z\setminus 0} \frac{\|\mu L\|_Z}{\|\mu\|_Z}.$$

The point of $\|\cdot\|_Z$ is to make transition operators P, P' which intuitively look close be close, uniformly in the size of S.

This metric $D(\rho, \rho')$ is good because, e.g. for ρ_λ equal to the product of N independent probabilities $(1 - \lambda, \lambda)$ on $\{0, 1\}$, then the speed of change of ρ_λ with respect to λ is

$$v = \left\|\frac{\mathrm{d}\rho}{\mathrm{d}\lambda}\right\|_Z = 1.$$

[1] i.e. the open sets are those generated from the "cylinder sets" $\prod_{s\in S'} C_s \times \prod_{s\in S\setminus S'} X_s$ for $S' \subset S$ finite, C_s an open subset of X_s, by arbitrary union and finite intersection

[2] probabilities which are defined for all subsets of X generated from the open subsets by countable unions and intersection and complementation

[3] Borel measures μ for which $\mu(X) = 0$

In contrast, most of the standard metrics on probability spaces give speeds that go to infinity as $N \to \infty$: for total variation (TV) metric, Kullback-Leibler (KL) divergence, Hellinger metric, and Fisher information metric, one obtains $v \sim \sqrt{N}$, and for projective metric and l_∞-transportation metric, one obtains $v \sim N$. One could divide these metrics by $\sqrt{N}$ or N respectively, but then many localised changes would have small distances, and even all changes would have small distances for some of the above (e.g. TV metric has diameter 1, so $1/\sqrt{N}$ after division by $\sqrt{N}$)!

To bound a transition operator L (or their differences) we can use that $L : Z \to Z$ bounded and linear induces $L : F \to F$ by

$$\mu(Lf) = (\mu L)(f) \ \forall \mu \in Z,$$

and so

$$\|L\|_Z \le \|L\|_F.$$

Proof.

$$\frac{|\mu LF|}{\|f\|_F} \le \frac{\|L\|_F}{\|Lf\|_F} |\mu LF| \le \|L\|_F \|\mu\|_Z,$$

so

$$\|\mu L\| \le \|L\|_F \|\mu\|.$$

□

Actually $\|L\|_Z = \|L\|_F$, using the Hahn-Banach theorem, but we will not need this observation.

In particular, we obtain the useful bound

$$\|P\|_Z \le \delta = \sup_{r \in S} \sum_{s \in S} k_{rs} = \|k\|_\infty,$$

where the elements of the *dependency matrix* are defined by

$$k_{rs} = \sup_{x, \tilde{x}} \frac{D_T\left(p_r^x, p_r^{\tilde{x}}\right)}{d_s\left(x_s, \tilde{x}_s\right)},$$

with $x, \tilde{x}$ agreeing off s, and differing on s, where

$$p_r^x\left(x_r'\right) = p_r(x_r' | \underline{x}),$$

and the *transportation metric* D_T on probabilities on a metric space X is

$$D_T(\rho, \tilde{\rho}) = \sup_{f \in L \setminus C} \frac{\rho(f) - \tilde{\rho}(f)}{\|f\|_L},$$

with L the Lipschitz functions $f : X \to \mathbb{R}$, and $\|f\|_L$ the best Lipschitz constant of f. k_{rs} bounds the influence of the state at site s now on the state at site r next step, e.g. $k_{rs} = 0$ if $s \notin N(r)$.

Proof. Use an alternative way of writing the transportation distance

$$D_T\left(\rho, \tilde{\rho}\right) = \inf_{\tau \in \mathcal{P}(X \times X)} \int d(x, \tilde{x}) \tau(dx, d\tilde{x}),$$

such that marginals of τ on the two copies of X are $\rho, \tilde{\rho}$ respectively. See Fig. 1.1. Such τ are usually called "couplings", but "joinings" is a better name because for me "couplings" implies some non-trivial effect of at least one on the other. Kantorovich and Rubinstein [11] proved equality of the two definitions of the transportation distance.

Take $x, \tilde{x}$ agreeing off s and take optimal joinings $\tau_r^{x,\tilde{x}}$ of p_r^x to $p_r^{\tilde{x}}$ $\forall r$ (if there is no optimal joining, take τ close to optimal and restrict the functions f to be independent of the state on all but finitely many sites). By the marginals property of $\tau_r^{x,\tilde{x}}$, for any $f \in F$

$$\left(Pf\right)(x) - \left(Pf\right)(\tilde{x}) = \int \left(f(x') - f(\tilde{x}')\right) \prod_{r \in S} \tau_r^{x,\tilde{x}}\left(dx_r', d\tilde{x}_r'\right).$$

But

$$f(x') - f(\tilde{x}') \le \sum_r \Delta_r(f) d_r(x_r', \tilde{x}_r'),$$

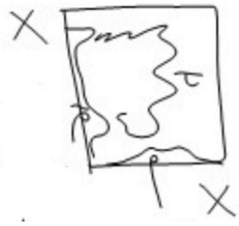

Fig. 1.1: Illustration of the marginals property of a joining of two probabilities.

Fig. 1.2: Example of Monge's earth movement problem in 1D.

and

$$\int d_r(x'_r, \tilde{x}'_r)\tau_r^{x,\tilde{x}}(dx'_r, d\tilde{x}'_r) \le k_{rs} d_s(x_s, \tilde{x}_s).$$

therefore

$$(Pf)(x) - (Pf)(\tilde{x}) \le \sum_r \Delta_r(f) k_{rs} d_s(x_s, \tilde{x}_s),$$

and

$$\Delta_s(Pf) \le \sum_r \Delta_r(f) k_{rs}.$$

Now sum this result over $s \in S$:

$$\|Pf\|_F = \sum_s \Delta_s(Pf) \le \|k\|_\infty \sum_r \Delta_r(f),$$

i.e. $\|Pf\|_F \le \|k\|_\infty \|f\|_F$, cf. Maes [18]. □

D_T is called transportation metric because of its origins in Monge's earth movement questions [11]. For example, see Fig. 1.2, where we want to minimise the integral of the mass moved times the distance moved. In $1D$ this is simple to find (Kolmogorov's formula), however in higher dimensions it is not easy (even $2D$), and the literature is frustrating as it strays from the original problem to other cost functions.

Example. An optimal joining. Let $\rho = (1-\lambda, \lambda)$ and $\tilde{\rho} = (1-\mu, \mu)$ on $\{0, 1\}$, with $d(0,1) = 1$, $\lambda > \mu$. Then we have a unique optimal joining τ and $D_T(\rho, \tilde{\rho}) = \lambda - \mu$. See Fig. 1.3.

Now, we will prove existence of exponentially attracting, unique stationary probabilities for "weakly dependent" PCAs, i.e. those for which $\|P\| < 1$. This is simply because P is then a contraction on $\mathcal{P}$, so it has a unique stationary probability ρ ($\mathcal{P}$ is complete with respect to our norm) and ρ attracts all $\nu \in \mathcal{P}$ exponentially:

$$D\left(\nu P^t, \rho\right) \le \|P^t\| D(\nu, \rho) \le \|P\|^t D(\nu, \rho).$$

Example. Show that $\|P\| \le 2(1-\lambda)$ for Stavskaya's PCA. This can be seen by using the dependency matrix k_{rs}. Recall that $D_T((1-\lambda, \lambda), (1-\mu, \mu)) = |\lambda - \mu|$. $\|P\| \le \sup_r \sum_s k_{rs}$, where the dependency

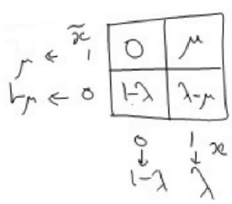

Fig. 1.3: Joining τ for example 1.1.1.

matrix is

$$k_{rs} = \sup_{x,\tilde{x}} \frac{D_T\left(p_r^x, p_r^{\tilde{x}}\right)}{d_s\left(x_s, \tilde{x}_s\right)}, \text{ where } x, \tilde{x} \text{ agree off } s, \text{ and differ on } s.$$

Now, for

(i) $k_{r,r}$:
if $x_{r-1} = 0$ then $p_r^x = (1-\lambda, \lambda)$ independently of x_r, so $D_T = 0$,
else if $x_{r-1} = 1$ then $p_r^x = (1-\lambda, \lambda)$ if $x_r = 0$, or $p_r^x = (0, 1)$ if $x_r = 1$, so $D_T = 1 - \lambda$,
so $k_{r,r} = 1 - \lambda$.

(ii) $k_{r,r-1}$:
if $x_r = 0$ then $D_T = 0$,
else if $x_r = 1$ then $D_T = 1 - \lambda$,
so $k_{r,r-1} = 1 - \lambda$.

(iii) $k_{r,s} = 0$ for $s \notin \{r, r-1\}$.

So the dependency matrix has diagonal and sub-diagonal entries $1 - \lambda$ and all the rest are zero. So $\|k\|_\infty = 2(1-\lambda)$, and $\|P\| \leq 2(1-\lambda)$. Hence for all $\lambda > \frac{1}{2}$ there exists a unique stationary probability and it attracts everything exponentially. However, we already know one stationary probability, $\delta_{\underset{\sim}{1}}$, so $\delta_{\underset{\sim}{1}}$ attracts all of $\mathcal{P}$ exponentially. See Fig. 1.4.

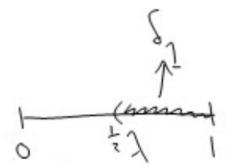

Fig. 1.4: The part of parameter space for the Stavskaya PCA for which absorption is established from weak dependence.

Rapid Absorption for Finite System when $\lambda > \frac{1}{2}$

On $\mathbb{Z}_N$, we still have $\|P\| \leq \delta = 2(1-\lambda) < 1$. Take $f(\underset{\sim}{x})$ to be the number of zeros in $\underset{\sim}{x}$. So $\|f\|_F = N$, from the definition, and $\|P^t f\|_F \leq \delta^t N$. Thus $P\{\text{not absorbed at time } t\} \leq \delta^t N$. See Fig. 1.5. So, in any meaningful sense, the time to absorption is of order $\frac{\log N}{\log 1/\delta}$, which is relatively short (compared with N).

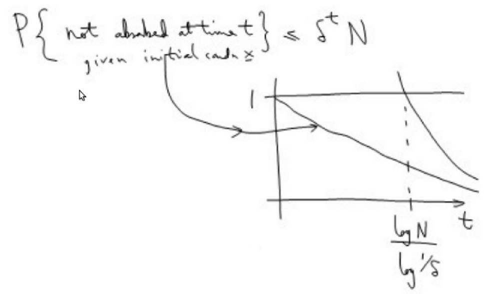

Fig. 1.5: Absorption time is logarithmic in N for $\lambda > \frac{1}{2}$.

Robustness of Exponentially Attracting Stationary Probability

So far we have shown existence of unique and exponentially attracting stationary probability for all weakly dependent PCA, but the same result holds for some PCA that are not weakly dependent. In particular, it is an open property, with the following explicit estimates.

Say a (stationary) probability $\rho_0 \in \mathcal{P}$ *attracts exponentially* for a transition operator P_0 if there exist C and $r < 1$ such that

$$D(\sigma P_0^t, \rho_0) \leq C r^t D(\sigma, \rho_0).$$

Theorem 1.1.1. *If P_0 has stationary probability ρ_0 which attracts exponentially, then all P near P_0, i.e. $\|P - P_0\|_Z = \delta < \frac{1-r}{C}$, also have exponentially attracting stationary probability ρ.*

Proof. Introduce the adapted norm on Z,

$$\|\mu\|_r = \sup_{n\geq 0} \|\mu P_0^n\| r^{-n}.$$

Then

$$\|\mu\| \leq \|\mu\|_r \leq C\|\mu\|,$$

so the nroms are equivalent, and

$$\|\mu P_0\|_r = \sup_{n\geq 0} \|\mu P_0^{n+1}\| r^{-n} \leq r\|\mu\|_r,$$

so $\|P_0\|_r \leq r$. Next

$$\|P - P_0\| = \delta \Rightarrow \|P - P_0\|_r \leq C\delta,$$

so

$$\begin{aligned}\|P\|_r &\leq \|P_0\|_r + \|P - P_0\|_r \\ &\leq r + C\delta < 1,\end{aligned}$$

since $\delta < \frac{1-r}{C}$. So $\|P^n\|_r \leq (r + C\delta)^n$ and hence $\|P^n\| \leq C(r + C\delta)^n$. So P has a unique stationary probability ρ and it attracts all exponentially. □

Also, we show that

$$\|\rho - \rho_0\| \leq \frac{C\|\rho_0(P - P_0)\|}{1 - r - C\delta}.$$

Proof. $\rho_0 P_0 = \rho_0$ and $\rho P = \rho$, so

$$(\rho - \rho_0)(Id - P) = \rho_0(P - P_0).$$

Hence, using the r-norm,

$$\|\rho - \rho_0\|_r \leq \frac{1}{1 - \|P\|_r} \|\rho_0(P - P_0)\|_r,$$

where $\|P\|_r \leq r + C\delta < 1$. Use results in the previous proof to obtain the result. □

It would be nice to give an explicit example of use of these results, e.g. to extend the parameter range for unique stationary probability of the Stavskaya PCA to some larger interval than $(\frac{1}{2}, 1]$, but I didn't see an easy way to achieve this (though it can be done by methods of Maes and Shlosman, see [18], which consist roughly speaking in showing that $\|P^n\| < 1$ for some $n > 1$). Nevertheless, the results of this subsection have conceptual importance.

1.1.2 λ near 0

This is the low recovery rate, high infectivity case for the Stavskaya PCA. We will show, following Toom et al. [26], that

$$\delta_{\underset{\sim}{0}} P^t \nrightarrow \delta_{\underset{\sim}{1}} \text{ as } t \to \infty \text{ for } \lambda < \frac{1}{54},$$

where $\delta_{\underset{\sim}{0}}$ is the probability distribution supported on all infected and $\delta_{\underset{\sim}{1}}$ is that on all healthy. Note that 1/54 is not an intrinsically important number, it is just to show that one can do concrete calculations.

Start from $\underset{\sim}{0}$ (all infected), the "worst case", and we ask for $P\{x_s^t = 1\}$ for $t > 0$. Without loss of generality, $s = 0$. The trick is to consider an equivalent way of generating the probability distribution on space-time configurations $\underset{\sim}{x} \in \{0,1\}^{\mathbb{Z}\times\mathbb{Z}^+}$, where $\mathbb{Z} \times \mathbb{Z}^+$ represents space $\times$ time, starting from $\underset{\sim}{0}$. We generate probability distributions over space-time configurations from $\underset{\sim}{0}$ at $t = 0$ by putting "stoppers" in $\mathbb{R}^2$ (embedding the discrete space-time $\mathbb{Z} \times \mathbb{Z}^+$ in $\mathbb{R}^2$) from $\left(s - \frac{1}{2}, t - \frac{1}{4}\right)$ to $\left(s + \frac{1}{2}, t - \frac{1}{4}\right)$ with independent probabilities λ for each (s,t) and using the deterministic rules given in Fig. 1.6 (the numbers $1/2, 1/4$ are chosen such that the stoppers "block" straight lines joining the space-time points). For example, see Fig. 1.7. This generates the same probability distribution as the PCA with initial

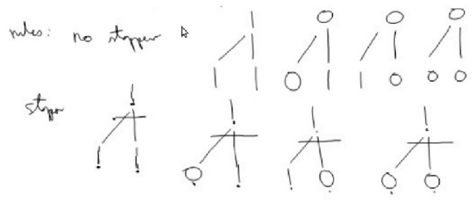

Fig. 1.6: Stopper generating rules for construction of space-time configurations.

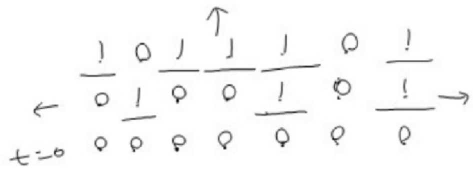

Fig. 1.7: Example of space-time configuration constructed using stoppers.

condition $\underset{\sim}{0}$.

Therefore $x_0^t = 1$ (recall that without loss of generality $s = 0$) if and only if every path along ups and up-rights from $t = 0$ encounters a stopper, that is if and only if there exists a "fence" around $(0, T)$ starting at $\left(\frac{1}{2}, T + \frac{3}{4}\right)$ formed from stoppers (going from left to right) and down-left and ups, e.g. see Fig. 1.8.

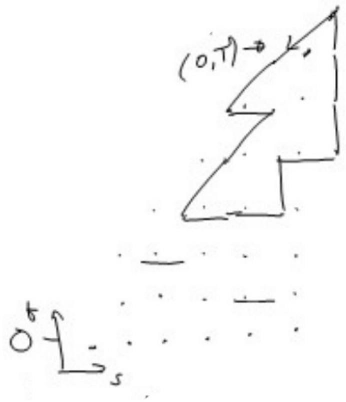

Fig. 1.8: Example of a "fence" around $(0, T)$.

If a fence has k stoppers, then it has precisely k ups and k down-lefts. Let N_k be the number of fences with k stoppers, then

$$N_k \leq \tbinom{3k}{k\,k\,k} \leq 27^k$$

(a large over-estimate because we could exclude those which self-intersect and those which don't surround $(0, T)$, and without loss of generality one could always start with a down-left and end with an up). So

$$\begin{aligned} P^0\{x_0^T = 1\} \leq \sum_{k \geq 1} N_k \lambda^k \leq \frac{27\lambda}{1 - 27\lambda}, &\quad \text{for } \lambda < \frac{1}{27} \\ < 1, &\quad \text{for } \lambda < \frac{1}{54}, \end{aligned}$$

independently of T.

Thus $P^0\{x_0^T = 1\} \not\to 1$ as $T \to \infty$, and $P^0\{x_0^T = 0\} \geq c(\lambda) = \frac{1-54\lambda}{1-27\lambda} > 0\ \forall T$. So $\delta_{\underset{\sim}{0}} P^t \not\to \delta_{\underset{\sim}{1}}$ as $t \to \infty$.

Where does $\delta_{\underset{\sim}{0}} P^t$ go? More generally what about νP^t for arbitrary initial ν?

To analyse this question we will use "monotonicity" of Stavskaya's PCA (also called "attractivity"). We define a partial order $\underset{\sim}{x} \leq \underset{\sim}{x}' \in \{0,1\}^{\mathbb{Z}}$ if for all $s \in \mathbb{Z}$ we have that $x_s \leq x_s'$ (with $0 < 1$). Also, say $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ is *non-decreasing* if $\underset{\sim}{x} \leq \underset{\sim}{x}'$ implies that $f(\underset{\sim}{x}) \leq f(\underset{\sim}{x}')$. Say that probabilities $\rho \leq \rho'$ on

$\{0,1\}^{\mathbb{Z}}$ if for all non-decreasing f, we have that $\rho(f) \le \rho'(f)$. Say that transition operator P is *monotone* if $\rho \le \rho'$ implies that $\rho P \le \rho' P$, equivalently if f non-decreasing implies Pf non-decreasing.

We prove Stavskaya's PCA is monotone by joining ("coupling") processes from two initial conditions $\underset{\sim}{x} \le \underset{\sim}{x}'$ in such a way that for all $t > 0$ we have that $\underset{\sim}{x}^t \le \underset{\sim}{x}'^t$, so it follows that

$$\left(P^t f\right)(\underset{\sim}{x}) \le \left(P^t f\right)(\underset{\sim}{x}') \;\forall \text{ non-decr. } f.$$

We denote the state $x_s^t x_s'^t$ of the joint process by $00, 01, 11$. The state 10 does not occur as initial condition because initially $\underset{\sim}{x} \le \underset{\sim}{x}'$ and our choice of joining never generates it. We choose the transition rates for the joint process to be:

$$00 \to \begin{cases} 11 \; P\{0 \to 1 | \underset{\sim}{x}\} \\ 01 \; P\{0 \to 1 | \underset{\sim}{x}'\} - P\{0 \to 1 | \underset{\sim}{x}\} \\ 00 \; P\{0 \to 0 | \underset{\sim}{x}'\}, \end{cases}$$

$$01 \to \begin{cases} 11 \; P\{0 \to 1 | \underset{\sim}{x}\} \\ 01 \; P\{0 \to 0 | \underset{\sim}{x}\} - P\{1 \to 0 | \underset{\sim}{x}'\} \\ 00 \; P\{1 \to 0 | \underset{\sim}{x}'\}, \end{cases}$$

$$11 \to \begin{cases} 11 \; P\{1 \to 0 | \underset{\sim}{x}\} \\ 01 \; P\{1 \to 0 | \underset{\sim}{x}\} - P\{1 \to 0 | \underset{\sim}{x}'\} \\ 00 \; P\{1 \to 0 | \underset{\sim}{x}'\}. \end{cases}$$

It can be checked that the marginals on the first and second components are just copies of the Stavskaya process. This proves monotonicity of the Stavskaya process.

Then $\delta_{\underset{\sim}{0}} P \ge \delta_{\underset{\sim}{0}}$, so $\delta_{\underset{\sim}{0}} P^t$ is a non-decreasing sequence. It is bounded above by $\delta_{\underset{\sim}{1}}$, and so converges to its supremum in the partial order $\le$, see Fig. 1.9 (consider $\delta_{\underset{\sim}{0}} P^t$ acting on arbitrary non-decreasing f). Call the limit ν_λ, so $\delta_{\underset{\sim}{0}} P^t \to \nu_\lambda$ as $t \to \infty$.

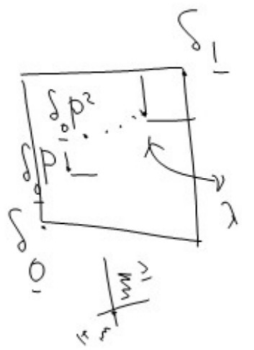

Fig. 1.9: Monotonicity of the dynamics of P on $\mathcal{P}$.

In addition,

- ν_λ is non-decreasing with respect to λ (prove this by joining processes for $\lambda < \lambda'$ to make $\underset{\sim}{x}^t(\lambda) \le \underset{\sim}{x}^t(\lambda')$).
- there exists a value $\lambda_c \in \left(\frac{1}{54}, \frac{1}{2}\right)$ such that $\nu_\lambda = \delta_{\underset{\sim}{1}}$ for all $\lambda \ge \lambda_c$, and $\nu_\lambda < \delta_{\underset{\sim}{1}}$ for all $\lambda < \lambda_c$ (this follows from ν_λ non-decreasing with λ).
- If $\lambda < \lambda_c$ then $\nu_\lambda(\underset{\sim}{1}) = 0$.

Proof of the latter. Else if $\nu_\lambda(\underset{\sim}{1}) = \rho > 0$, then the conditional probability $\mu = \nu_\lambda(\cdot|\text{ not } \underset{\sim}{1})$ is stationary and $\nu_\lambda = p\delta_{\underset{\sim}{1}} + (1+p)\,\mu$. So $\mu < \nu_\lambda$. However, $\mu \ge \delta_{\underset{\sim}{0}}$, so

$$\mu = \mu P^t \ge \delta_{\underset{\sim}{0}} P^t \to \nu_\lambda \text{ as } t \to \infty,$$

and finally, we reach the contradiction $\mu \ge \nu_\lambda$. □

- For $\lambda < \lambda_c$ and any subset $A \subset \mathbb{Z}$, let $\alpha_A = P\{\text{infection never dies} \mid \text{initial infected set is } A\}$, then $\alpha_A = 1$ if A is infinite (if A is finite $\alpha_A \le 1 - \lambda^{2|A|} < 1$).

- For all $\nu \in \mathcal{P}$,
$$\nu P^t \to \gamma \delta_{\underset{\sim}{1}} + (1-\gamma)\,\nu_\lambda \text{ as } t \to \infty,$$
with $\gamma = P\{\text{eventual absorption}\}$ $(= 1-\alpha_A$ if $\nu = \delta_A)$, so in particular there are no other stationary probabilities than the convex combinations of ν_λ and $\delta_{\underset{\sim}{1}}$.

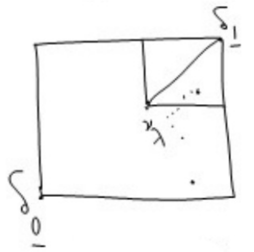

Fig. 1.10: Schematic of $\mathcal{P}$ indicating the line segment of stationary probabilities from ν_λ to δ_1.

For finite versions (on $\mathbb{Z}_N$) the estimate of $P^0\{x_0^T = 1\}$ does not apply for $T > N$ because fences can wrap round the cylinder $\mathbb{Z}_N \times \mathbb{Z}_N$ and the counting has to be changed, but one can obtain estimates depending on T, N and prove exponentially long time to absorption (with respect to N) for λ small.

1.1.3 Variations on Stavskaya's PCA

Some possible variations on Stavskaya's PCA are:

- one can make the recovery probability λ_1, and avoiding infection probability λ_2 differ (and make recovery depend on the state of left-hand neighbour) and obtain similar results.
- one can add infectivity from the other side: by monotonicity one obtains similar results.
- I don't know about the effects of small changes breaking monotonicity (one would need to go back to drawing fences), but the large λ regime is still easy.
- one can go to higher dimensions and inhomogeneous networks: see [13] for continuous time (presumably there is a literature in discrete time too). To go to higher dimensions, a useful method is to compare the "radial process" of the extent of infection with Stavskaya's PCA on the half-line $\mathbb{Z}_+$, using monotonicity.
- there is an alternative treatment of Stavskaya's PCA (under the name "oriented percolation") by Durrett [9], which involves looking at the velocity of propagation of a front between infected and healthy zones.

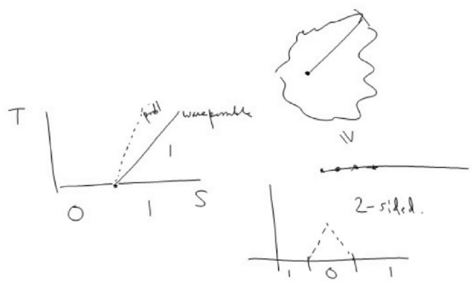

Fig. 1.11: Alternative treatment of Stavskaya's PCA by Durrett.

1.2 The Majority Voter PCA

Stavskaya's PCA has the feature of an absorbing state, so is not communicating (a Markov process is *communicating* if it is possible to get from any state to any other and back). The majority voter PCA has non-unique phase even though communicating. It is important to distinguish it from many other voter models (see e.g. [13]) which have absorbing states and I prefer to call "opinion-copying models", the randomness coming solely from the choice of which neighbour's opinion to copy.

The majority voter PCA was proposed and studied numerically by [27] and analysed by Toom. We concentrate on the NEC (north-east-centre) model: $S = \mathbb{Z}^2$, $x_s \in \{0,1\} \approx \{-,+\}$ and

$$x_s^{t+1} = \begin{cases} \text{majority of } x_s^t, x_{s+E}^t, x_{s+N}^t, \text{ i.e. NEC voters at time } t, & \text{with probability } 1-\lambda, \\ \text{the opposite with probability } \lambda, \end{cases}$$

where $E = (1,0)$ and $N = (0,1)$.

Our treatment will be much more sketchy than for Stavskaya's PCA.

Key properties of the model are:

- there is an exponentially attracting, unique stationary probability for λ near $1/2$ (because weakly dependent).
  Exercise: prove that $|\lambda - \frac{1}{2}| < \frac{1}{6}$ implies weakly dependent.
- P is monotone for $\lambda \leq \frac{1}{2}$, so $\delta_{\underset{\sim}{0}} P^t$ approaches some $\underline{\mu}$, and $\delta_{\underset{\sim}{1}} P^t$ approaches some $\bar{\mu}$, where $\underline{\mu} \leq \bar{\mu}$ (reflections of each other via $+ \leftrightarrow -$). See Fig. 1.12.
- for small enough λ, we have $\underline{\mu} < \bar{\mu}$.

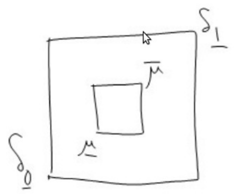

Fig. 1.12: The space $\mathcal{P}$ of probabilities for the NEC majority voter PCA.

Idea of Proof. For $\lambda = 0$, no islands of 1 in a sea of 0 will survive. Toom calls the deterministic cellular automaton an "eroder". Given an island of 1 in a sea of 0, draw its south-west envelope: the island can't grow beyond it. However, every north-east corner is eroded, so after finitely many steps the island will disappear. See Fig. 1.13.

Now turn on $\lambda > 0$ and show that

$$\underline{\mu}\left(x_{00}^t = 1\right) \leq \text{ some } \psi(\lambda) \to 0 \text{ as } \lambda \to 0,$$

by showing that

$$\begin{aligned} P^0\left(x_{00}^t = 1\right) &= \sum_{m=0}^{\infty} \sum_{\Gamma} P\left(G(\underline{x}) = \Gamma\right), \text{ where } \Gamma \text{ are certain graphs on S} \times \text{T with } m \text{ edges} \\ &\leq \sum_m N_m \, P(\text{error at each vertex}) \\ &\leq \sum_m 48^{2m} \lambda^{m/4+1} \\ &= \frac{\lambda}{1 - 48^2 \lambda^{1/4}} \text{ for } \lambda < \frac{1}{48^8} \\ &< \frac{1}{2} \text{ for } \lambda \text{ small enough } (1 - 48^2\lambda^{1/4} > 2\lambda, \text{ e.g. } \lambda < \frac{1}{2} 48^{-8}). \end{aligned}$$

□

For details, see Toom's original proof [25] or lecture notes [26] or [12] from which the above estimates are taken.

Thus, we have at least two stationary probabilities (plus their convex combinations). I call these "ferromagnetic" phases (cf. statistical mechanics). There might be yet more stationary probabilities.

Possible variations:

Fig. 1.13: Island of 1s in a sea of 0s for the NEC majority voter model.

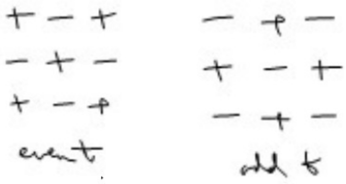

Fig. 1.14: Antiferromagnetic NEC voter PCA taking λ near 1 to obtain "antiferromagnetic period-2" phases.

- For λ near 1 the model can be called an "antimajority" voter. The model is equivalent to parameter value $1-\lambda$ by recoding x^t for odd t as $-x^t$. So if $\delta_{\underset{\sim}{0}} P_\varepsilon^t \to \underline{\mu}$, then $\delta_{\underset{\sim}{0}} P_{1-\varepsilon}^t \to$ a 2-cycle, with $\underline{\mu}$ at even t and $\bar{\mu}$ at odd t. So on $\{0,1\}^{S\times T}$ we have two (or more) probabilities, which we call "period-2" phases. This is an example of "non-trivial collective behaviour", or "asymptotic periodicity".
- We can also make PCA which have "antiferromagnetic" phases, by considering $x_s^{t+1} =$ majority of $(x_s^t, -x_{s+E}^t, -x_{s+N}^t)$. One can prove this by recoding the usual NEC PCA by multiplication by $(-)^{\text{parity of } s}$ where the parity of s is $s_1 + s_2 \mod 2$.
- By taking λ near 1 in the antiferromagnetic model we obtain "antiferromagnetic period-2" phases. See Fig. 1.14.
- One can break the $+ \leftrightarrow -$ symmetry and still prove non-unique stationary probabilities for λ small.
- One can make the interaction anisotropic (see [4]) and still obtain non-unique stationary probability.

Contrast the 2D Ising model where non-unique phase for low temperature is lost as soon as one adds a magnetic field.

Many other variants of the NE majority voter PCA are possible, e.g. [7].

Open questions:

(i) What about adding south and west neighbour influence? One can allow the error rate to depend weakly on the state of the S and W neighbours and still keep non-unique stationary probability, but numerically one obtains the same with the strong perturbation of enlarging the majority neighbourhood to NECSW, yet we are not aware of any proof.

(ii) Majority voter models in continuous time? Note there exist continuous-time voter models (see e.g. [13]), but they are very different because they choose to duplicate a random neighbour's state precisely, so $\underset{\sim}{+}, \underset{\sim}{-}$ are absorbing.

1.3 General Properties for Phases of PCA

1.3.1 Phases and Emergence

Define a *phase* of a PCA as a limit point of probability on space-time configurations started from an initial space probability in the distant past. Any convex combination of phases is a phase, so it suffices to consider "extremal" phases, i.e. those which are not combinations of any others. Thus, Stavskaya's PCA has two extremal phases for λ small enough: those generated by time evolution of $\delta_{\underset{\sim}{1}}$ and ν_λ. Phases do not have to be time-translation invariant, e.g. the period-2 phases of the NEC voter PCA for $\lambda = 1 - \varepsilon$.

Emergence maps the dynamical model to the set of phases. The amount of emergence expressed by a phase μ is

$$D(\mu, \{\text{product of independent dynamics}\}) \approx \text{ distance from mean-field models,}$$

using Dobrushin metric D (now on S-T phases, rather than on probabilities on configuration space). Weak emergence means we get $D > 0$.

For strong emergence the set of phases is not a singleton, i.e. non-unique phase. The amount of strong emergence is given by the diameter of the set of phases, e.g. for Stavskaya's PCA,

$$\operatorname{diam}(\{\text{phases}\}) = D(\nu_\lambda, \delta_{\underset{\sim}{1}}),$$

which by analogy to a computation in [7] is $c(\lambda)$ (the probability of a given site being infected in ν_λ). See Fig. 1.15.

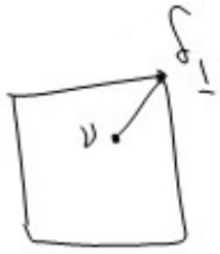

Fig. 1.15: Amount of strong emergence for Stavskaya's PCA.

One word of caution: a system may exhibit strong emergence and yet none of its space-time phases exhibit weak emergence. For example, take a weakly coupled network of bistable units. The good conjecture would be that for indecomposable systems exhibiting strong emergence, at least one of the phases exhibits weak emergence.

See my paper in Nonlinearity for some more open problems [16].

1.3.2 Gibbsian View of Space-Time Phases

The space-time phases for PCA can be re-defined as the probabilities μ on the set of space-time configurations $\underset{\sim}{x}$ such that for all bounded $\Lambda \subset S \times T$ and configurations $\underset{\sim}{x}_{\Lambda^c}$ on the complement $\Lambda^c = (S \times T) \backslash \Lambda$, the conditional probability of μ given $\underset{\sim}{x}_{\Lambda^c}$ is proportional to

$$\prod_{R : R \cap \Lambda \neq \emptyset} P_R(\underset{\sim}{x}),$$

where R ranges over subsets of $S \times T$ of the form $\{(s, t+1)\} \cup (N(s) \times \{t\})$ and $P_R(\underset{\sim}{x}) = p_s(s_s^{t+1} | \underset{\sim}{x}^t)$.

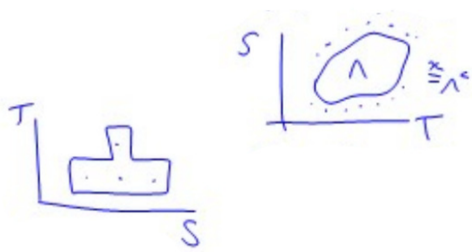

Fig. 1.16: Gibbsian view of space-time phases.

One can rewrite the product as $\exp\left(-\sum_{R:R\cup\Lambda\neq\emptyset}\phi_R(\underset{\sim}{x})\right)$, with $\phi_R(\underset{\sim}{x}) = -\log P_R(\underset{\sim}{x})$. This is the defining (Dobrushin-Lanford-Ruelle) condition for Gibbs phases in equilibrium statistical mechanics (where we would have configurations on only S and not $S \times T$), with "energy" function $\sum_R \phi_R(\underset{\sim}{x})$. Recall from equilibrium statistical mechanics that for energy function $H = \sum_R H_R$ and "coolness" $\beta \in \mathbb{R}^+$ (inverse temperature), the conditional probability for configuration $\underset{\sim}{x}$ with respect to counting measure is given by

$$
\begin{aligned}
P(\underset{\sim}{x}_\Lambda | \underset{\sim}{x}_{\Lambda^c}) &= \frac{1}{Z_\Lambda(x_{\Lambda^c})} e^{-\beta\sum_{R:R\cap\Lambda\neq\emptyset} H_R(\underset{\sim}{x})} \\
&= e^{-\beta(\sum_{R:R\cap\Lambda\neq\emptyset} H_R(\underset{\sim}{x}) - F_\Lambda(\underset{\sim}{x}_{\Lambda^c})|\Lambda|)},
\end{aligned}
$$

where Z_Λ is a normalisation constant called the partition function and $F = \lim_{|\Lambda|\to\infty} F_\Lambda$ is called the free energy per site, see e.g. Lebowitz, Maes and Speer [12]. One can absorb β into H and F, so in our interpretations $\beta = 1$.

Note, however, that the "energy functions" for PCA the have special feature that $F = 0$, because

$$
\sum_{\underset{\sim}{x} \text{ on } \Lambda} \prod_{R:R\cap\Lambda\neq\emptyset} P_R(\underset{\sim}{x}) = P\,(\text{upper boundary values} \mid \text{lower ones}),
$$

which scales like the surface area of Λ, not the volume of Λ.

Fig. 1.17: Surface to volume effect.

So generic statements for equilibrium statistical mechanics do not apply, e.g. Gibbs phase rule that says in equilibrium statistical mechanics generically co-existence of N extremal phases is of codimension $(N-1)$, such as in the $2D$ Ising model.

One can ask how the phase or set of phases varies with parameters. In the exponentially attracting regime the (unique) phase varies smoothly with parameters (interpreted as in [17]). In general, the set of phases depends upper semi-continuously on parameters, but it need not be lower semi-continuous. For example, consider discrete-time Glauber dynamics of the 2D Ising model below the critical temperature: as the magnetic field h crosses zero the set of phases jumps. Non-unique phase for the NEC majority voter PCA is robust to small changes in the model, because the perturbations can be compared with the NEC majority voter with slightly larger error rate. This gives an explicit example where Gibbs phase rule fails for PCA, but the argument depends on monotonicity and the question of generic robustness of non-unique phase for PCA remains open.

1.3.3 Indecomposability

If the state-space decomposes into two or more components which do not communicate then we obtain non-unique phases trivially, because each communicating component supports at least one phase, but I don't want to count this as strong emergence, because it suffices to check in which component you started to understand which phase you will see.

Say a PCA is "indecomposable" [15] if there exists a $D_0 \in \mathbb{R}_+$ such that for all finite subsets $A, B \subset S \times T$ with separation greater than or equal to D_0 (with respect to a metric on $S \times T$, e.g. the sum of the displacements in S and T), and two realisations x, x' then there exists a realisation z agreeing with

x on A and x' on B, cf. "specification property" in dynamical systems theory, e.g. for the NEC voter, one can take $D_0 = 1$.

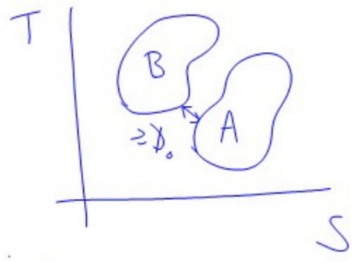

Fig. 1.18: Indecomposability of PCA.

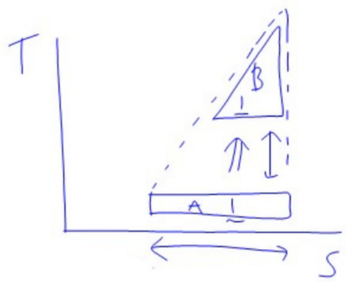

Fig. 1.19: Failure of Stavskaya's PCA to be indecomposable.

Stavskaya's PCA does not satisfy this indecomposability property because of the absorbing state $\underset{\sim}{1}$, but deserves to be called strongly emergent because the phase corresponding to ν_λ is not associated with another communicating component but with the transient set. The finite versions of Stavskaya's PCA have a single communicating component (the absorbing state) even though it is not the whole state space. Therefore, we say a PCA is "pre-indecomposable" if it has a unique indecomposable subset and it can be attained in bounded time and we count non-unique phase as strong emergence for pre-indecomposable systems.

Chapter 2

Deterministic Case - Coupled Map Lattices (CML)

2.1 Examples with Non-unique Space-Time Phase

A CML (coupled map lattice) is a map $f : M \to M$, where $M = \prod_{s\in S} M_s$, M_s is a finite dimensional manifold, e.g. $[-1, +1]$, and S is a countable metric space. We want to turn our PCA examples into CML examples.

Given a PCA on $\Sigma = \prod_{s\in S} \Sigma_s$, with transition probabilities p_s^σ for $s \in S, \sigma \in \Sigma_s$, partition $[-1, +1]$ into $|\Sigma_s|$ equal length intervals labelled by the possible values of $\sigma_s \in \Sigma_s$ (Fig. 2.1), and construct a piecewise affine map $f : M \to M$ with

$$\frac{\partial f_s}{\partial x_s}(x) = \frac{1}{p_s^{\sigma(x)}(\sigma'_s)}, \quad \frac{\partial f_s}{\partial x_r} = 0 \text{ for } r \neq s, \text{ for } x_s \in \sigma_s, x_r \in \sigma_r,$$

where σ'_s is the symbolic state of $f_s(x)$, e.g. for the NEC voter, see Fig. 2.2.

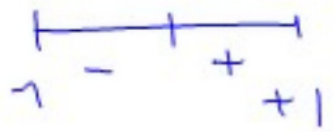

Fig. 2.1: Partition of $[-1, +1]$ into two intervals for $\Sigma_s = \{-, +\}$.

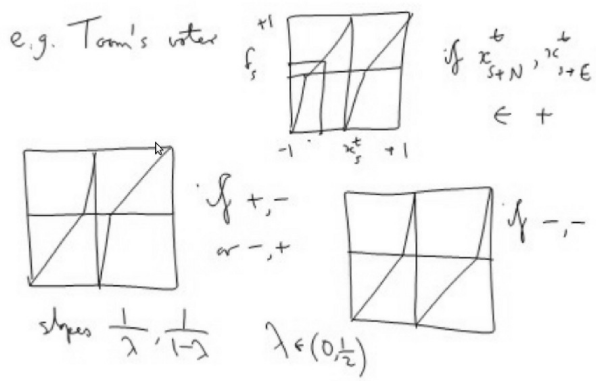

Fig. 2.2: CML for the NEC voter model.

Then for any initial probability on M with absolutely continuous marginals on finite subsets Λ of S

and density h_Λ satisfying a Hölder condition

$$|\log h_\Lambda(x) - \log h_\Lambda(y)| \le C \sum_{s\in\Lambda} |x_s - y_s|^\delta$$

in the distant past, $\sigma(x)$ is distributed according to a phase of the PCA [10]. So, for example, one can make CML with ferromagnetic phases.

The idea is that the initial probability is repeatedly stretched and cut, and so becomes more uniform in each cylinder set (set with prescribed symbols), and the action on probabilities that are uniform on cylinder sets is precisely the PCA (see Fig. 2.3) (cf. random number generators).

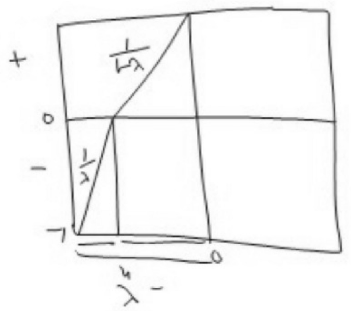

Fig. 2.3: Action of f on probabilities.

The Hölder condition on the densities is certainly stronger than necessary, but we did not determine a better condition.

This construction provided a response to a challenge by Sinai and Bunimovich [6] to provide a CML with non-unique phase. A similar construction was made by [22] to simulate discrete-time Glauber dynamics of the 2D Ising model, with variable partition according to the symbolic state of the neighbours and constant slope within partition element, but it is not clear that paper contained a proof.

Similarly, we can make CML with period-2 phase (simulate the NEC voter with $\lambda \in (1/2, 1)$), anti-ferromagnetic, persistent infection phases, etc.

One can make invertible examples by replacing $[-1, +1]$ by a solid torus $[-1, +1] \times D^2$ modulo identification of the ends, and use distorted versions of the solenoid map, depending on the symbolic state of the neighbours. Recall the standard (Smale-Williams) solenoid map [20]: $(x, w) \in S^1 \times D^2$, where $S^1 = \mathbb{R}/2\mathbb{Z}$ (note I'm taking S^1 of length 2), $D^2 = \{w \in \mathbb{C} : |w| \le 1\}$, and

$$\begin{cases} x' = 2x, \\ w' = \lambda w + \mu e^{i\pi x}, \end{cases}$$

with $\lambda < \mu,\ \mu + \lambda \le 1$, to make it map $S^1 \times D^2$ into itself and one-to-one. The intersection of the forward images of the solid torus under the solenoid map is a strange attractor - a Cantor set of lines locally (see Fig. 2.4).

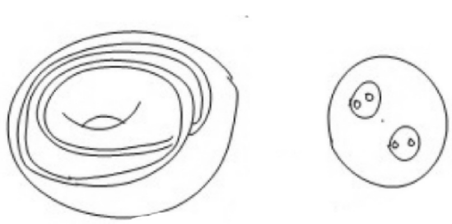

Fig. 2.4: Solenoid map.

Our distorted solenoid maps are maps on $[-1, +1] \times D^2$ depending on the symbolic state of neighbours, where we set $\sigma_s = \operatorname{sign}(x_s)$, e.g. for $++$ neighbours:

$$\begin{cases} x' = g_{++}(x), \\ w' = \lambda w + \mu e^{i\pi x}, \end{cases}$$

where g_{++} denotes the first map of Fig. 2.2. The addition of the D^2 dynamics makes the map one-to-one. To make it onto, we need to extend the solenoid maps outside $S^1 \times D^2$ to say $S^1 \times \sqrt{2}D^2$ to make backward orbits exist for all initial points, e.g. by $w' = (2 - |w|^2)(\lambda w + \mu e^{i\pi x}) + (|w|^2 - 1)w$ for $1 \leq |w| \leq sqrt2$; see Fig. 2.5.

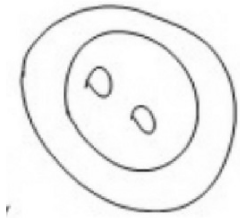

Fig. 2.5: Extending solenoid map outside $S^1 \times D^2$.

Our CML is not continuous, however, because of the jumps at changes of symbolic state of neighbours. Bardet and Keller [3] overcame this defect, they even made C^∞ examples. Their main idea was to replace symbolic coupling by discrete diffusion

$$x_s^{t+1} = (1-\varepsilon)e(x_s^t) + \frac{\varepsilon}{2}(e(x_{s+N}^t) + e(x_{s+E}^t)),$$

with e a carefully chosen map, but it is not clear how to make it invertible.

2.2 Statistical Phases for Uniformly Hyperbolic Attractors of Finite-Dimensional Deterministic Dynamical Systems

We shall now review the statistical theory of "chaos". Thus for this section we drop the study of spatially extended systems and consider finite-dimensional dynamical systems only. Instead of space-time phases, we will obtain just time-phases. This is only "complexity science" if we think of the states at successive times as the interdependent components. It is a warm-up for the theory of space-time phases for spatially extended deterministic differentiable dynamical systems. The key idea is to introduce a symbolic description of the orbits, so the approach is limited to systems for which this is possible. We shall see, however, that this is a large class, in particular C^1-open.

The idea (due to Sinai [23]) is that under suitable conditions, the deterministic dynamics on an attractor is equivalent to a stochastic dynamics on a finite symbol set, of a form called "Gibbsian".

The setting is iteration of a C^1-diffeomorphism $x^{t+1} = f(x^t)$, $x^t \in M$, a C^1 manifold with norm $|\cdot|$ on tangent vectors and associated Finsler metric. A key trick is to note that its orbits $\underline{x} = (x^t)_{t\in\mathbb{Z}}$ are the fixed points of a "super-map" $F : M^{\mathbb{Z}} \to M^{\mathbb{Z}}$ defined by $F(\underline{x})^t = f(x^{t-1})$. F is a C^1 map when $M^{\mathbb{Z}}$ is endowed with the supremum metric over its components. The tangent space $T(M^{\mathbb{Z}})$ at a sequence x is the set of sequences ξ of tangent vectors ξ^t to M at x^t with

$$\|\underset{\sim}{\xi}\|_\infty = \sup_{t\in\mathbb{Z}} |\xi^t| < \infty.$$

We sometimes denote it $TM_\infty^{\mathbb{Z}}$.

Definition. An orbit $\underline{x}$ is *uniformly hyperbolic* if it is a non-degenerate fixed point of F, i.e. $F(\underline{x}) = \underline{x}$ and

$$I - DF_{\underline{x}} = \begin{pmatrix} \ddots & & & & & \\ \ddots & I & & & & \\ & -f'_{-1} & I & & & \\ & & -f'_0 & I & & \\ & & & -f'_1 & I & \\ & & & & \ddots & \ddots \end{pmatrix}$$

has bounded inverse, where f'_t denotes $f'(x^t)$. A set Λ of orbits is uniformly hyperbolic if $\|(I - DF_{\underline{x}})^{-1}\|$ is bounded for $\underline{x} \in \Lambda$.

Example. Arnol'd's cat map. Consider

$$f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ on } M = \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2.$$

The whole of $\mathbb{T}^2$ is uniformly hyperbolic, which is seen by solving $(I - DF)\underline{\xi} = \underline{\eta}$ for $\underline{\xi} \in TM_\infty^{\mathbb{Z}}$, given $\underline{\eta} \in TM_\infty^{\mathbb{Z}}$. This can be done by splitting ξ^t, η^t into components $\xi^t_\pm, \eta^t_\pm$ along the eigenvectors of f.

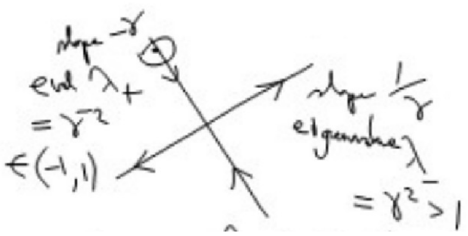

Fig. 2.6: Eigenvectors of f.

The notation is: $+$ means forward contracting, and $-$ means backward contracting. Many books use $+ = s$ and $- = u$, standing for stable and unstable. This is bad terminology, however, as can be seen by considering points near the eigen-directions, and noticing that actually the "stable" direction is unstable and vice-versa.

The splitting disconnects the equations into separate equations for the $+$ components and the $-$ components:
$+$ component: $\xi^t_+ - \lambda_+ \xi^{t-1}_+ = \eta^t_+$. Rewrite it as $\xi^t_+ = \eta^t_+ + \lambda_+ \xi^{t-1}_+$ to see it has unique bounded solution $\xi^t_+ = \sum_{n \geq 0} \lambda_+^n \eta_+^{t-n}$.
$-$ component: $\xi^t_- - \lambda_- \xi^{t-1}_- = \eta^t_-$. Rewrite it as $\xi^{t-1}_- = \frac{1}{\lambda_-}(\xi^t_- \eta^t_-)$ which has unique bounded solution $\xi^t_- = -\sum_{n \geq 1} \lambda_-^{-n} \eta_-^{t+n}$.

Thus we have unique bounded solution $\underline{\xi} = \underline{\xi}_+ + \underline{\xi}_-$ and with respect to Euclidean norm on the tangent space $T\mathbb{T}^2$ for example, we have that

$$\|\underline{\xi}_+\| \leq \frac{\|\underline{\eta}_+\|}{1 - \lambda_+}, \quad \|\underline{\xi}_-\| \leq \frac{\|\underline{\eta}_-\|}{\lambda_- - 1}, \quad \text{and so} \quad \|(I - DF)^{-1}\| \leq \frac{1}{1 - \lambda_+}.$$

The usual definition of uniform hyperbolicity is: there exists an invariant splitting of

$$TM_{x^t} = E^+_{x^t} \oplus E^-_{x^t}$$

along orbit $\underline{x}$, and $C \in \mathbb{R}^+, \lambda \in (0,1)$ such that $\xi^t \in E^\pm_{x^t}$ implies

$$|\xi^{t+s}| \leq C\lambda^{|s|}|\xi^t|,$$

for all $s > 0$ for E^+, all $s < 0$ for E^-.

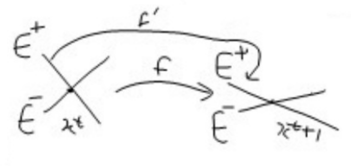

Fig. 2.7: Invariant splitting of uniform hyperbolicity.

This definition of uniform hyperbolicity is more or less equivalent to mine. In particular, it is implied by mine if $f'_t = f'(x^t)$, $t \in \mathbb{Z}$ is bounded (which is automatic if M is compact).

Proof. Suppose $\left(I - DF_{\underline{x}}\right)\underline{\xi} = \underline{\eta}$ has unique bounded solution $\underline{\xi}$ for any bounded $\underline{\eta}$. In particular take $\eta^t = 0$ for all $t \neq 0$. Let $\tilde{\xi}^t = z^{|t|}\xi^t$, for some $z > 1$ and try to prove $\underline{\tilde{\xi}}$ is bounded.

Now $\xi^t - f'_{t-1}\xi^{t-1} = \eta^t$, therefore

$$\begin{cases} \tilde{\xi}^t - z f'_{t-1}\tilde{\xi}^{t-1} = 0, \ t > 0, \\ \tilde{\xi}^t - \frac{1}{z} f'_{t-1}\tilde{\xi}^{t-1} = \eta^t, \ t \le 0 \ (= 0 \text{ for } t < 0). \end{cases}$$

So $(I - DF - E)\,\tilde{\xi} = \underline{\eta}$ with

$$E = \begin{pmatrix} \ddots & & & & \\ \ddots & 0 & & & \\ & \left(\frac{1}{z} - 1\right) f'_{-1} & 0 & & \\ & & (z-1)\, f'_0 & 0 & \\ & & & \ddots & \ddots \end{pmatrix}.$$

Then $\|E\| \le |z-1| l$ with $l = \sup_t |f'_t|$ and so $(I - DF - E)$ has bounded inverse if

$$\|E\| < K := \|(I - DF)^{-1}\|^{-1},$$

and then

$$\|(I - DF - E)^{-1}\| \le \frac{1}{K - \|E\|}.$$

So $\|\tilde{\xi}\| \le \frac{|\eta^0|}{K - |z-1| l}$, i.e. $|\xi^t| \le \frac{z^{-|t|}|\eta^0|}{K - |z-1| l}$ decays exponentially both ways in t. One obtains the same for any initial time s.

Then split $\eta^0 = \eta^0_+ + \eta^0_- = \xi^0 - f'_{-1}\xi^{-1}$ and check that this defines a complementary pair of projections $P^\pm$ with $P^+\eta^0 = \eta^0_+$ and $P^-\eta^0 = \eta^0_-$, $P^2 = P$, $P^+ + P^- = I$ and invariance of splitting and uniform exponential decay estimates. □

Similarly, the usual definition plus some extra hypotheses (notably that the angle of splitting is bounded away from 0) if M is non-compact implies my definition.

A nice feature of uniformly hyperbolic orbits is "robustness": for all $\tilde{f}$ C^1-close to f, there exists a unique orbit $\tilde{x}$ of $\tilde{f}$ uniformly near to $\underset{\sim}{x}$, i.e. $d(\tilde{x}^t, \underset{\sim}{x}^t) \le \delta$ (prove by applying implicit function theorem to $F(\underset{\sim}{x}) = \underset{\sim}{x}$). Furthermore, $\tilde{x}^0\left(x^0\right)$ is Hölder continuous and one-to-one, so given a uniformly hyperbolic set $\Lambda \subset M$ (which is the union of points of a uniformly hyperbolic set of orbits), for all $\tilde{f}$ C^1-close to f, there exists $h : \Lambda \to M$, a near-identity homeomorphism onto $h(\Lambda)$ such that

$$\tilde{f}h = hf \text{ on } \Lambda.$$

Similarly for t-dependent perturbations (but with h time-dependent).

Now we will show that the dynamics on a uniformly hyperbolic attractor is equivalent to a Gibbsian stochastic process on a symbol space (a generalisation of a Markov chain). We follow Adler's choice of definition of Markov partition [1].

Definition. A Markov partition of an invariant set Λ is a finite set of disjoint open subsets A_i whose closures cover Λ, such that, letting Γ be the graph with nodes A_i and edges $i \to j$ if

$$fA_i \cap A_j \neq \emptyset,$$

every doubly infinite path in Γ occurs as the symbolic trajectory of one and only one orbit in Λ.

Example: Solenoid attractor. Consider the attractor of $f : S^1 \times D^2 \to S^1 \times D^2$ ($S^1 = \mathbb{R}/2\mathbb{Z}$) given by

$$\begin{cases} x' = 2x \in S^1, \\ w' = \lambda w + \mu e^{i\pi x} \in D^2 \subset \mathbb{C}. \end{cases}$$

This has Markov partition: $0 = - = \{(x, w) : x \in (-1, 0)\}$, $1 = + = \{x \in (0, 1)\}$. See figures 2.8, 2.9.

Second Example: Arnol'd's cat map. Consider

$$f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ on } \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2.$$

Fig. 2.8: Markov partition for solenoid map.

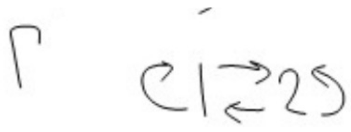

Fig. 2.9: The graph of allowed transitions for the solenoid attractor (relabel!).

The partition $\{1, 2\}$ of Fig. 2.10 is almost a Markov partition, but to achieve unique orbit for each symbol sequence, one has to subdivide region 1 into e.g. $\{1a, 1b, 1c\}$ (not shown). Rather than draw the graph of allowed transitions on $\{1a, 1b, 1c, 2\}$ we prefer to use multiple edges in the graph on $\{1, 2\}$ to indicate the two transitions $1a \to 1, 1c \to 1$ which have to be distinguished to achieve unique orbit, see 2.11.

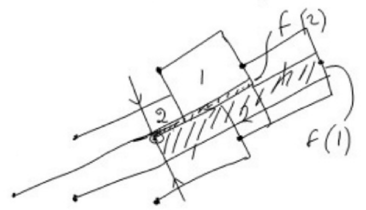

Fig. 2.10: An almost Markov partition for Arnold's cat map.

Then f is topologically conjugate to the shift σ on the set of doubly infinite paths in Γ

$$\sigma : \quad \cdots e_{-1}.e_0 e_1 \cdots \quad \mapsto \quad \cdots e_{-1} e_0 . e_1 \cdots \ ,$$

where e_i label edges in Γ, with product topology (two paths are close if they agree on a long finite portion $e_M, \ldots, e_N$) modulo some identifications of paths which correspond to the same orbit of f.

A general way to obtain symbolic dynamics is to show that for any locally maximal, compact uniformly hyperbolic set, there exists a Markov partition. This can be proved by using the

Theorem 2.2.1 ("Shadowing theorem".). *Given a uniformly hyperbolic set* Λ *there exist* $K, \varepsilon_0 > 0$ *such that for all* $\varepsilon \leq \varepsilon_0$ *every* ε*-pseudo-orbit is* $K\varepsilon$*-shadowed by a true orbit of* f*.*

Here, $\underset{\sim}{y}$ is an $\underline{\varepsilon\text{-pseudo-orbit}}$ for f if $d(fy_n, y_{n+1}) \leq \varepsilon$ for all n, i.e. $d(F\underset{\sim}{y}, \underset{\sim}{y}) \leq \varepsilon$ in supremum metric, and $\underline{x}$ $\underline{\delta\text{-shadows}}$ $\underset{\sim}{y}$ if $d(x_n, y_n) \leq \delta$ for all n, i.e. $d(\underline{x}, \underset{\sim}{y}) \leq \delta$ in supremum metric. See Fig. 2.12.

One can prove the shadowing theorem by constructing a good approximation J to $(I - DF_{\underset{\sim}{y}})^{-1}$ and then showing $\underline{x} \mapsto \underline{x} - JF(\underline{x})$ is a contraction in a neighbourhood of $\underset{\sim}{y}$ (in some local coordinate system about $\underset{\sim}{y}$).

To make a Markov partition, choose "enough" periodic orbits of f on Λ as the "symbols" σ and use the shadowing theorem to make the whole of Λ from true orbits shadowing a sequence of segments of the chosen periodic orbits.

One can also show that $x^t(\underline{\sigma})$ depends exponentially weakly on σ^s, for $|s - t| > N$ large.

Fig. 2.11: The graph Γ of allowed transitions.

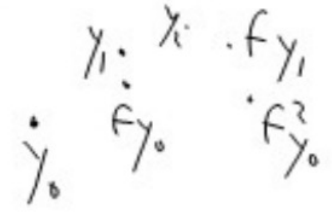

Fig. 2.12: An ε-pseudo-orbit.

We are now ready to sketch why there is a natural measure on a uniformly hyperbolic attractor and how it corresponds to a Gibbsian process. Suppose $f \in C^{1+\text{Hölder}}$ and Λ is a uniformly hyperbolic, bounded topologically mixing attractor (*topologically mixing* means there exists an N such that for all $A, B \neq \emptyset$ open, $f^n A \cap B \neq \emptyset$ for all $n \geq N$). Then there exists a probability ρ on Λ (called *SRB measure*, after Sinai, Ruelle and Bowen) such that for ν any absolutely continuous[1] probability on the basin $B(\Lambda)$ of attraction of Λ, $\nu f_*^t \to \rho$ as $t \to +\infty$ in weak-* topology.

The simplest way to understand weak-* topology on $\mathcal{P}(\lambda)$, Λ compact, is to metrise it by

$$d(\mu, \nu) = sup_{g \in \text{Lip} \setminus \text{C}} \frac{\mu(g) - \nu(g)}{\text{Lip}(g)},$$

C denoting the constant functions, and Lip the best Lipschitz constant.

The probability ρ can be obtained as a Gibbs phase for symbol sequences with "energy" contribution from time t

$$\phi_-^t(\underline{\sigma}) = \log|\det Df|_{E^-}(x^t(\underline{\sigma}))|,$$

i.e.

$$\rho\{\sigma^{-T} \cdots \sigma^R | \left(\sigma^t\right)_{t<-T, t>+R}\} \propto \exp\{-\sum_{t \in \mathbb{Z}} \left(\phi_-^t\left(\underline{\sigma}\right) - \phi_-^t(\underline{\bar{\sigma}})\right)\},$$

for any allowed reference sequence $\bar{\sigma}$ satisfying given the past and future.

The Gibbs phase is unique because Λ is topologically mixing, the dependence of ϕ_-^t on σ^s decays exponentially with $|s-t|$ and $\mathbb{Z}$ is $1D$ so one can apply a $1D$ statistical mechanics result to deduce unique phase [21].

Idea of proof. We sketch a proof that $\nu f_*^t \to \rho$ as $t \to \infty$, given by Gibbs potential $\log|\det Df|_{E^-}\left(x^t\left(\underline{\sigma}\right)\right)|$.

Given an absolutely continuous initial probability ν on $B(\Lambda)$ in the distant past $t = -M$, what are the relative probabilities of seeing allowed paths $\sigma^0 \cdots \sigma^N$? (We adopt the convention that σ^t labels the edge in Γ taken from time t to $t+1$).

We illustrate the relevant subsets for a nonlinear distortion of the solenoid map, see Fig. 2.13. So $\sigma_0 = 0 \to 1$ means that at $t = 0$ we are in the shaded region in Fig. 2.13. Then $\sigma^0 \cdots \sigma^N$ equal to a specified sequence corresponds to a thin slice at $t = 0$, see Fig. 2.14. Go back to $t = -M$, but without specifying σ^t, $t < 0$. $P(\sigma^0 \cdots \sigma^N) = \nu(\text{set at } t = -M \text{ giving rise to this sequence}) = \nu(\text{union of } 2^M \text{ disks})$, as shown in Fig. 2.15.

[1]with respect to Lebesgue class, i.e. ifλ denotes a Lebesgue class measure then $\lambda(A) = 0 \Rightarrow \nu(A) = 0$. Alternatively ν has a density h in $L^1(\lambda)$.

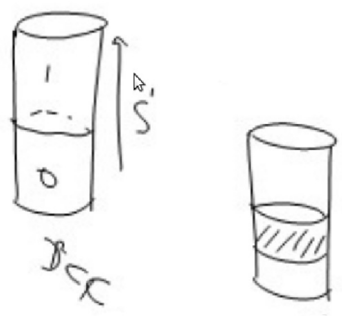

Fig. 2.13: The set which will make transition $0 \to 1$.

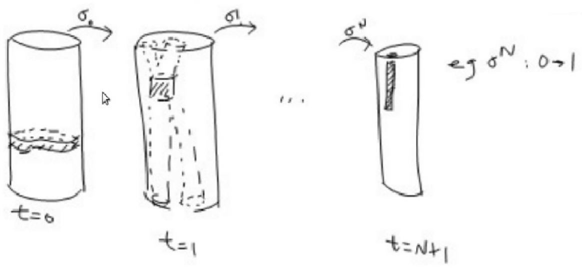

Fig. 2.14: The sets of points at times $0, 1, \ldots, N+1$ corresponding to symbol sequence $\sigma^0, \ldots, \sigma^N$.

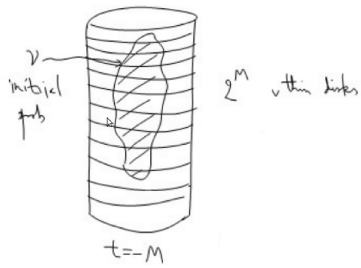

Fig. 2.15: The 2^M thin disks corresponding to initial conditions at $t = -M$ which will perform sequence $\sigma^0 \ldots \sigma^N$.

The dependence of $P(\sigma^0 \cdots \sigma^N)$ on sequence $\sigma^0 \cdots \sigma^N$ is via the thicknesses of these disks. So

$$P(\sigma^0 \cdots \sigma^N) \underset{\sim}{\propto} \prod \left|\det Df|_{E^-}(x^t)\right|^{-1}$$

along associated backwards orbits. More precisely, letting $\mathcal{C}$ be any choice of $\sigma^{-M/2} \cdots \sigma^{-1}, \sigma^{N+1} \cdots \sigma^{N+M/2}$,

$$\lim_{M\to\infty} P(\sigma^0 \cdots \sigma^N | \mathcal{C}, \nu \text{ at } t=-M) = \frac{1}{Z(\mathcal{C},\bar{\sigma})} \prod_{t=-\infty}^{+\infty} \frac{|\det Df_-(x^t(\bar{\sigma}))|}{|\det Df_-(x^t(\underline{\sigma}))|},$$

where $\bar{\sigma}$ is some particular choice of $\sigma^0 \cdots \sigma^N$ completed by the chosen conditions $\mathcal{C}$ and Z is a normalisation constant. We choose to write this as

$$\frac{1}{Z} \exp\{-\sum_t (\phi^t(\underline{\sigma}) - \phi^t(\underline{\bar{\sigma}}))\}$$

with $\phi^t(\underline{\sigma}) = \log|\det Df_-(x^t(\bar{\sigma}))|$.

The function ϕ^t depends exponentially weakly on σ^s with respect to $|s-t|$, so the infinite sum makes sense and standard statistical mechanics (e.g. [21]) implies that there exists a unique probability with these conditionals. $\square$

2.3 Uniformly Hyperbolic Dynamics on Networks

Now we move to the spatially extended context. A *network* is a countable metric space (S,d). For all $s \in S$, we suppose as local state space a finite-dimensional manifold M_s with a norm $|\cdot|_s$ on tangent vectors and the induced Finsler metric. We let $M = \prod M_s$ with supremum metric, and hence tangent vectors $\underset{\sim}{\xi}$ satisfying $\|\underset{\sim}{\xi}\| = \sup_{s\in S}|\xi_s| < \infty$. The state $x^t \in M$ evolves by $x^{t+1} = f(x^t)$ (f is a CML). Suppose $f \in C^1$, in particular $\sup_{r\in S}\sum_{s\in S}|\frac{\partial f_r}{\partial x_s}| < \infty$. If all $M_s = V$ a vector space, then a standard class of examples is

$$x'_s = f^\varepsilon_s(x_s) = f^0_s(x_s) + \varepsilon \sum_{r\in S} C_{sr}(x_r - x_s)$$

with the matrix C satisfying $\sup_s \sum_r |C_{sr}| < \infty$. But one can also couple maps of nontrivial manifolds, like cat maps and solenoid maps.

The orbits of f are precisely the fixed points of the super-map $F: M^{\mathbb{Z}} \to M^{\mathbb{Z}}$ defined by $\underset{\sim}{x} = (x^t_s) \mapsto f_s(\underset{\sim}{x}^{t-1})$ for all $(s,t) \in S \times \mathbb{Z}$. We endow $M^{\mathbb{Z}}$ with supremum metric and then tangent vectors are those $\underset{\sim}{\xi}$ with finite supremum norm. F is C^1 for $\underset{\sim}{x}$ in bounded subsets.

Definition. An orbit $\underset{\sim}{x}$ of the CML f is uniformly hyperbolic if it is a non-degenerate fixed point of F (using supremum norm on $M^{\mathbb{Z}}$), i.e. $\|(I - DF)^{-1}\| < \infty$.

One could apply the same results as in the previous section (existence of a splitting, robustness, shadowing, Markov partition) except for deducing unique natural probability if $|S| = \infty$ because $M = \prod_{s\in S} M_s$ is infinite dimensional and so $\det Df_-$ requires interpretation. Also one would not deduce anything useful about dependence on $s \in S$, e.g. spatial correlations of ρ.

So we do a space-time version of uniform hyperbolicity theory (cf. [19, 14]). Assume that Df is *exponentially local* (which includes the case of finite range interaction), i.e.

$$\exists \zeta > 1,\ \phi : [0,\zeta) \to \mathbb{R} \text{ s.t. } \sup_r \sum_s \left|\frac{\partial f_r}{\partial x_s}\right| z^{d(r,s)} \le \phi(z)\ \forall z \in [0,\zeta).$$

Then the uniformly hyperbolic splitting $TM_x = E^+_{\underset{\sim}{x}} \oplus E^-_{\underset{\sim}{x}}$ constructed as in the previous section is given by exponentially local projections $P^{\pm}_{\underset{\sim}{x}}$.

One can prove this via the following theorem (the application will be to $S = S \times \mathbb{Z}$).

Theorem 2.3.1. *For each $s \in S$, let X_s, Y_s be Banach spaces and let $X = \{x \in \prod_{s\in S} X_s : \|x\|_\infty < \infty\}$, $Y = \{y \in \prod_{s\in S} Y_s : \|y\|_\infty < \infty\}$. Let $L : X \to Y$ be a linear map, ϕ-exponentially local, with bounded inverse, and let $y \in Y$ be (C,λ)-exponentially localised around $o \in S$, i.e. $|y_s| \le C\lambda^{d(s,o)}$. Then $x = L^{-1}y$ is (WC,μ)-exponentially localised around o for some W,μ, functions of $\|L^{-1}\|, \lambda, \phi$.*

Proof. (following Baesens and MacKay [2]).

Let $z \in [1, \zeta)$, with $z\lambda < 1$, $\tilde{y}_r = y_r z^{d(r,o)}$, $\tilde{x}_r = x_r z^{d(r,o)}$. Then $\|\tilde{y}\| < \infty$, and we want to bound $\|\tilde{x}\|$. Now $Lx = y$ implies that

$$\sum_s L_{rs} z^{d(r,o)-d(s,o)} \tilde{x}_s = \tilde{y}_r.$$

So $(L + L^o)\tilde{x} = \tilde{y}$ with $L^o_{rs} = L_{rs}\left(z^{d(r,o)-d(s,o)} - 1\right)$.

Let $\tilde{L}_{rs} = L_{rs} z^{d(r,s)}$, so $\|\tilde{L}\| < \phi(z) < \infty$. Now $d(r,o) - d(s,o) \le d(r,s)$, by the triangle inequality, so $|L^o_{rs}| \le |\tilde{L}_{rs} - L_{rs}|$, and so $\|L^o\| \le \|\tilde{L} - L\|$. For $z = 1$, $\tilde{L} = L$. $\tilde{L}(z)$ depends continuously on z because given $1 \le z_1 \le z_2 < \zeta$

$$(\tilde{L}(z_2) - \tilde{L}(z_1))_{rs} = L_{rs}(z_2^{d(r,s)} - z_1^{d(r,s)}).$$

Therefore

$$\begin{aligned} |(\tilde{L}(z_2) - \tilde{L}(z_1))_{rs}| &\le |L_{rs}|\ d(r,s) z_2^{d(r,s)} \log \frac{z_2}{z_1} \\ &\le \frac{1}{e} \frac{\log z_2/z_1}{\log z_3/z_1} z_3^{d(r,s)} |L_{rs}|, \text{ for any } z_3 > z_2, \end{aligned}$$

and so

$$\begin{aligned} \|\tilde{L}(z_2) - \tilde{L}(z_1)\| &\le \frac{1}{e} \frac{\log z_2/z_1}{\log z_3/z_1} \|\tilde{L}(z_3)\| \\ &\le \frac{1}{e} \frac{\log z_2/z_1}{\log z_3/z_1} \phi(z_3). \end{aligned}$$

In particular $\|\tilde{L}(z) - L\| \le \beta(z) = \min_{z_3} \frac{1}{e} \frac{\log z}{\log z_3/z} \phi(z_3) \to 0$ as $z \searrow 1$, and so is less than $\|L^{-1}\|^{-1}$ for z near enough to 1. Hence $\|(L + L^0)^{-1}\| \le (\|\tilde{L}\|^{-1} - \|\tilde{L} - L\|)^{-1}$ for z near 1. Finally

$$\|\tilde{x}\| \le \frac{\|\tilde{y}\|}{\|L^{-1}\|^{-1} - \|\tilde{L} - L\|}.$$

□

Apply this to the uniform hyperbolic splitting for CML by taking $S = S \times \mathbb{Z}$ and considering DF exponentially local with respect to

$$\bar{d}((s_1, t_1), (s_2, t_2)) = d(s_1, s_2) + |t_2 - t_1|,$$

and repeat the construction of the splitting and see that the resulting projections $P^\pm : TM_x \to TM_x$ are exponentially local with respect to d.

As for the case of time only, we obtain robustness of uniformly hyperbolic orbits in $S \times T$ with respect to exponentially local perturbations in space. Also one can prove $S \times T$ shadowing. Similarly one can obtain Markov partitions for uniformly hyperbolic sets of CML (cf. [19]). This means a coding $\underset{\sim}{x} = \underset{\sim}{x}(\underset{\sim}{\sigma})$ of the orbits $\underset{\sim}{x} = (x^t_s)^{t \in T = \mathbb{Z}}_{s \in S}$ of the uniformly hyperbolic set by $S \times T$ symbol tables $\underset{\sim}{\sigma}$ (see Fig. 2.16) from some allowed set Σ such that every allowed table occurs and there exists a unique $\underset{\sim}{x}$ for each $\underset{\sim}{\sigma}$.

The simplest case is $\Sigma = (\prod_{s \in S} \Sigma_s)^{\mathbb{Z}}$ with each Σ_s finite, e.g. an uncoupled lattice of solenoid maps $\Sigma_s = \{0, 1\}$ and $x^t_s = X_s\left(\cdots \sigma^{t-1}_s . \sigma^t_s \sigma^{t+1}_s \cdots\right)$, the unique point whose orbit visits partition elements $0, 1$ of $\mathbb{S}^1 \times D^2$ in sequence $(\sigma^t_s)_{t \in \mathbb{Z}}$. Now make a C^1-small exponentially local coupling. By the robustness of uniformly hyperbolic sets, we can deform the coding $\underset{\sim}{x} = \underset{\sim}{x}(\underset{\sim}{\sigma})$ but now x^t_s depends in general on all $\sigma^{t'}_{s'}$ (not just $\sigma^{t'}_s$), though exponentially weak on distant (s', t').

The same holds if the uncoupled dynamics has a general finite graph of allowed transitions, e.g. cat map. Also one can use the $S \times T$ shadowing theorem to construct a Markov partition for locally maximal uniformly hyperbolic sets without assuming proximity to an uncoupled case.

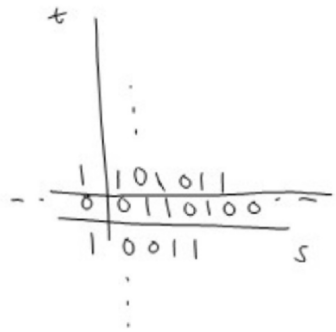

Fig. 2.16: A space-time symbol table for a CML of solenoid maps.

2.4 Natural Measures on Uniformly Hyperbolic Attractors for Coupled Map Lattices

Finally we sketch how the phases of uniformly hyperbolic attractors of CML can be analysed. The phases are the probability measures on space-time orbits which arise by starting in the distant past with a probability measure ν whose marginals on all finite subsets of S are absolutely continuous and have Hölder density (plus perhaps a little more).

They are the Gibbs phases for an "energy" with contribution

$$\phi_s^t(\underset{\sim}{\sigma}) = \mathrm{tr}[\log Df_-(\underset{\sim}{x}^t(\underset{\sim}{\sigma}))]_{ss},$$

from space-time site (s,t). By uniform hyperbolicity theory, this ϕ_s^t depends exponentially weakly on $\sigma_{s'}^{t'}$ with respect to $\bar{d}((s',t'),(s,t))$.

The connection with the theory for a single dynamical system (SRB measures) is that

$$\log|\det A| = \mathrm{tr}[\log A] = \sum_{s \in S} \mathrm{tr}[\log A]_{ss},$$

plus a constant from the choice of branch of log, see [5].

Thus phases for uniformly hyperbolic attractors of CML correspond to equilibrium statistical mechanics for a special class of spin systems on $S \times T$. When the dimension of $(S \times T)$ is greater than 2 (or even equal to 2), one can expect to make differentiable examples with non-unique Gibbs phase if sufficiently coupled (just as we did for piecewise affine CML).

Challenge: make a CML of solenoid maps with uniformly hyperbolic attractor exhibiting non-unique space-time phase, cf. [3] who came close.

In conclusion, we have seen that the space-time phases of a large class of spatially extended differentiable dynamical systems can be reduced to Gibbs phases for an associated spin system on space-time.

References

[1] R. Adler. Symbolic dynamics and markov partitions. *Bull Am Math Soc*, 35:1–56, 1998. 2.2

[2] C. Baesens and R.S. MacKay. Effect of temperature on polaronic and bipolaronic structures of the adiabatic holstein model. *Journal of Statistical Physics*, 85:471–488, 1996. 2.3

[3] J.B. Bardet and G. Keller. Phase transitions in a piecewise expanding coupled map lattice with linear nearest neighbour coupling. *Nonlinearity*, 19:2193, 2006. 2.1, 2.4

[4] C.H. Bennett and G. Grinstein. Role of irreversibility in stabilizing complex and nonergodic behavior in locally interacting discrete systems. *Phys Rev Lett*, 55:657–660, 1985. 1.2

[5] J. Bricmont and A. Kupiainen. High temperature expansions and dynamical systems. *Communications in Mathematical Physics*, 178:703–732, 1996. 2.4

[6] L.A. Bunimovich and Ya.G. Sinaĭ. Spacetime chaos in coupled map lattices. *Nonlinearity*, 1(4):491–516, 1988. 2.1

[7] M. Diakonova and R.S. MacKay. Mathematical examples of space-time phase. *Int J Bifurc Chaos*, to appear, 2011. 1, 1.2, 1.3.1

[8] R.L. Dobrushin. Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity. *Probl Info Transm*, 7:149–164, 1971. 1.1.1

[9] R. Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12(4):999–1040, 1984. 1.1.3

[10] G. Gielis and R.S. MacKay. Coupled map lattices with phase transition. *Nonlinearity*, 13(3):867–888, 2000. 2.1

[11] L.V. Kantorovich and G.Sh. Rubinstein. On a space of completely additive functions (in Russian). *Vestnik Leningrad Univ*, 13(7):52–59, 1958. 1.1.1

[12] J.L. Lebowitz, C. Maes, and E.R. Speer. Statistical mechanics of probabilistic cellular automata. *Journal of Statistical Physics*, 59:117–170, 1990. 1.2, 1.3.2

[13] T.M. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, Berlin, 2005. Reprint of the 1985 original. 1.1.3, 1.2, 1.2

[14] R.S. MacKay. Dynamics of networks: features which persist from the uncoupled limit. In Strien S.J. Van and Verduyn Lunel S.M., editors, *Stochastic and Spatial Structures of Dynamical Systems*, pages 81–104. N.Holland, 1996. 2.3

[15] R.S. MacKay. Indecomposable Coupled Map Lattices with Non–unique Phase. In Chazottes J.R. and Fernandez B., editors, *Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems*, Lect Notes Phys 671, pages 65–94. Springer, 2005. 1.3.3

[16] R.S. MacKay. Nonlinearity in complexity science. *Nonlinearity*, 21(12):T273–T281, 2008. 1.3.1

[17] R.S. MacKay. Robustness of markov processes on large networks. *J Difference Eqns & Applns*, iFirst, 2011. 1.1.1, 1.3.2

[18] C. Maes. Coupling interacting particle systems. *Reviews in Mathematical Physics*, 5(3):457–475, 1993. 1.1.1, 1.1.1

[19] Ya.B. Pesin and Ya.G. Sinaĭ. Space-time chaos in chains of weakly interacting hyperbolic mappings. *Adv Sov Math*, 3:165–198, 1991. 2.3, 2.3

[20] R.C. Robinson. *Dynamical systems.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, second edition, 1999. Stability, symbolic dynamics, and chaos. 2.1

[21] D. Ruelle. *Thermodynamic formalism.* Addison-Wesley, 1978. 2.2, 2.2

[22] H. Sakaguchi. Phase transitions in coupled bernoulli maps. *Prog Theor Phys*, 80:7–12, 1988. 2.1

[23] Ya.G. Sinaĭ. Gibbs measures in ergodic theory. *Russ Math Surv*, 28:21–64, 1972. 2.2

[24] O.N. Stavskaya. Gibbs invariant measures for Markov chains on finite lattices with local interaction (in Russian). *Matematicheskii Sbornik*, 134(3):402–419, 1973. 1.1

[25] A.L. Toom. Stable and attractive trajectories in multicomponent systems. In R.L. Dobrushin and Ya.G. Sinai, editors, *Multicomponent Random Systems, Adv Probab Related Topics 6*, pages 549–576. 1980. 1.2

[26] A.L. Toom, N.B. Vasilyev, O.N. Stavskaya, L.G. Mityushin, G.L. Kurdyumov, and S.A. Pirogov. Discrete local Markov systems. In R.L. Dobrushin, V.L. Kryukov, and A.L. Toom, editors, *Stochastic cellular systems: ergodicity, memory, morphogenesis*, pages 1–182. Manchester University Press, 1990. 1, 1.1.2, 1.2

[27] N.B. Vasilyev, M.B. Petrovskaya, and I.I. Piatetski-Shapiro. Simulation of voting with random errors (in Russian). *Automatica i Telemekh*, 10:103–7, 1969. 1.2