Noise is your friend

P. Cvitanović, J.M. Heninger and D. Lippolis

November 2015

Outline

1. **what this talk is about**
   - a geophysical prelude
2. **deterministic partitions**
   - idea #1: partition by periodic points
3. **dynamicist's view of noise**
   - idea #2: evolve densities, not noisy trajectories
   - idea #3: for unstable directions, look back
4. **optimal partition hypothesis**

should you listen to the weatherman?

Brian Farrell

"Traditionally, a statistical quantity is obtained from an ensemble average of sample realizations of the turbulence"

instead:

"statistical state dynamics (SSD) takes probability density function (pdf) as a state variable. Its dynamics has one and only one fixed point ($\cdots$polar jet$\cdots$), and only SSD equations reveal it" (huh?)

"the Fokker-Planck equation is intractable for representing complex system dynamics" (that is what we'll use in this talk :)

- B F Farrell and P J Ioannou, *Statistical State Dynamics: a new perspective on turbulence in shear flows*; arXiv.org:1412.8290

statistical state dynamics

Brian Farrell

- take 2-layer baroclinic turbulence model
- truncate SSD to the first two cumulants: mean flow, perturbation state covariance C
- closure (drop higher cumulants) by stochastic forcing Q
- Lyapunov equation

$$\frac{dC}{dt} = \mathrm{A}\,\mathrm{C} + \mathrm{C}\,\mathrm{A}^{\dagger} + \epsilon \mathrm{Q}\,, \quad \mathrm{A} = \text{linearized flow}$$

makes entry into climate science (at a baby level, *enfin*!) as the attractive fixed point of SSD

fixed point SSD

works on Jupiter

see the breakfast talk min 37:00 to 44:00

beyond fixed points : quasi-bilinear oscillation

works on Earth

the same video, continued :

- every 13.5 months equatorial winds reverese direction
- SSD explanation: Hopf bifurcation to an attractive limit cycle

noise rules the state space

> Science originates from curiosity and bad eyesight.
>
> — Bernard de Fontenelle,
> *Entretiens sur la Pluralité des Mondes Habités*

in practice

every physical problem is coarse partitioned by noise

noise rules the state space

> Science originates from curiosity and bad eyesight.
>
> — Bernard de Fontenelle,
> *Entretiens sur la Pluralité des Mondes Habités*

in practice

every physical problem is coarse partitioned by noise

- any physical system experiences (some kind of) noise
- any numerical computation is ‘noisy’
- any prediction only needs a desired finite accuracy

dynamics + noise: unique coarse-grained partition

reasonable to assume that the noise

is uniform,
leading to a uniform grid partition of the state space

dynamics + noise: unique coarse-grained partition

reasonable to assume that the noise

is uniform,
leading to a uniform grid partition of the state space

in dynamics, this is wrong!

noise always has memory

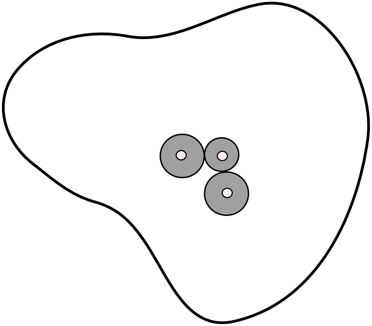
dynamics + noise: unique coarse-grained partition

noise memory

accumulated noise along dynamical trajectories
always coarsens the partition nonuniformly

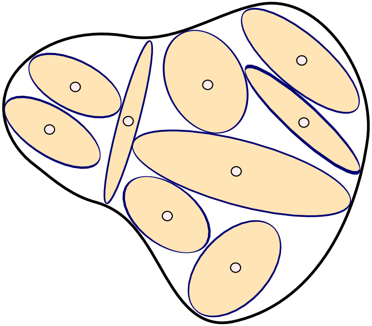
noise limited state space partitions

noise limited cell



a resolvable neighborhood is no smaller than a ball whose radius is the noise amplitude

noise limited partition grid



state space noise-partitioned into neighborhoods indicated by their centers

dynamical system

state space

a manifold $\mathcal{M} \in \mathbb{R}^d$: d numbers determine the state of the system

representative point

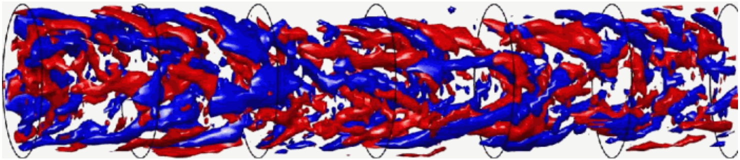
$x(t) \in \mathcal{M}$
a state of physical system at instant in time

today's experiments

example of a representative point

$x(t) \in \mathcal{M}$, $d = \infty$
a state of turbulent pipe flow at instant in time

Stereoscopic Particle Image Velocimetry $\rightarrow$ 3-d velocity field over the entire pipe[1]

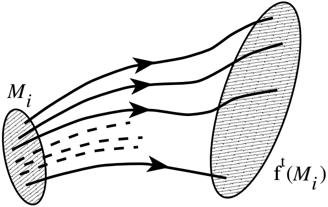


[1] Casimir W.H. van Doorne (PhD thesis, Delft 2004)

dynamics

map $f^t(x_0)$ = representative point time t later

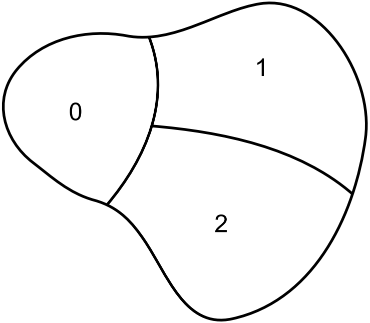
evolution in time



f^t maps a region $\mathcal{M}_i$ of the state space into the region $f^t(\mathcal{M}_i)$

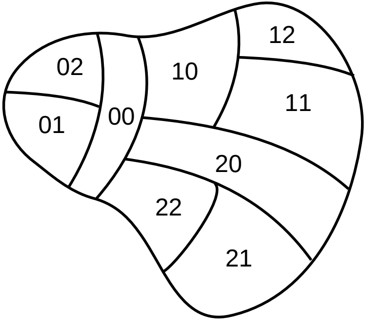
deterministic partition into regions of similar states

1-step memory partition



$\mathcal{M} = \mathcal{M}_0 \cup \mathcal{M}_1 \cup \mathcal{M}_2$
ternary alphabet
$\mathcal{A} = \{1, 2, 3\}$.

2-step memory refinement



$\mathcal{M}_i = \mathcal{M}_{i0} \cup \mathcal{M}_{i1} \cup \mathcal{M}_{i2}$
labeled by nine 'words'
$\{00, 01, 02, \cdots, 21, 22\}$.

deterministic partitions are no good

deterministic dynamics: partitioning can be arbitrarily fine

requires exponential # of exponentially small regions

deterministic partitions are no good

deterministic dynamics: partitioning can be arbitrarily fine

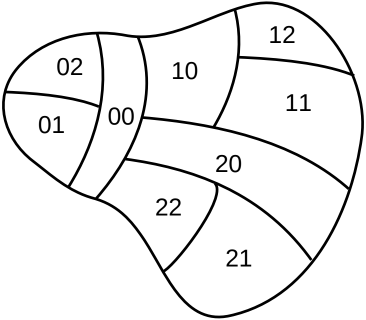
requires exponential # of exponentially small regions

yet

in practice

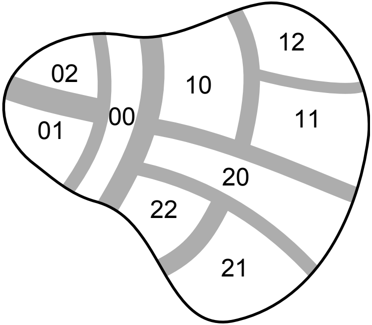
every physical problem must be coarse partitioned

deterministic vs. noisy partitions



deterministic partition

can be refined
ad infinitum



noise blurs the boundaries

when overlapping, no further refinement of partition

periodic points instead of boundaries

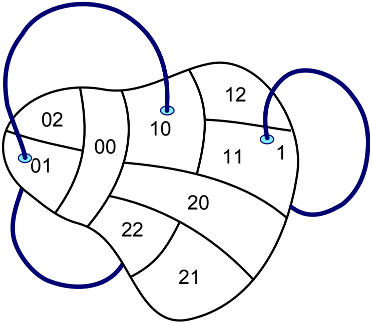
- mhm, do not know how to compute boundaries...

periodic points instead of boundaries

- mhm, do not know how to compute boundaries...
- however, each partition contains a short periodic point

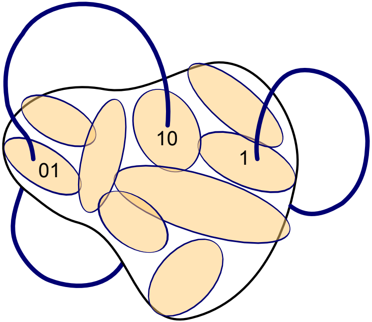
periodic orbit partition

deterministic partition



some short periodic points:
fixed point $\overline{1} = \{x_1\}$
two-cycle $\overline{01} = \{x_{01}, x_{10}\}$

noisy partition



periodic points blurred by noise
into cigar-shaped densities

periodic points and their cigars

- each partition contains a short periodic point smeared into a ‘cigar’ by noise

periodic points and their cigars

- each partition contains a short periodic point smeared into a ‘cigar’ by noise
- compute the size of a noisy periodic point neighborhood!

how big is the neighborhood blurred by the accumulated noise?

the (well known) **key formula** that we now derive:

$$Q_{n+1} = M_n Q_n M_n^T + \Delta_n$$

density covariance matrix at time n: Q_n
noise covariance matrix: Δ_n
Jacobian matrix of linearized flow: M_n

Lyapunov equation, doctoral dissertation 1892
Ornstein-Uhlenbeck 1930
Kalman filter 'prediction' 1960

Langevin, Fokker-Planck ...

continuous time stochastic dynamical system $(\mathcal{M}, v, \sigma)$

$$dx = v(x)\,dt + \sigma(x)\,d\hat{\xi}(t)$$

x a point in state space $\mathcal{M}$
$v(x)$ the deterministic velocity field or 'drift'
$d\hat{\xi}(t)$ the standard Brownian noise, uncorrelated in time

$$\langle d\hat{\xi}_i(t')\,d\hat{\xi}_j^\top(t)\rangle = \delta_{ij}\,\delta(t-t')dt$$

the noise

anisotropic, state dependent and multiplicative
strength given by
diffusion matrix $\sigma(x)$, or
noise covariance matrix is $\Delta(x) = \sigma\,\sigma^\top$

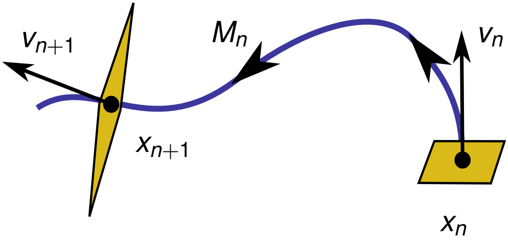
strategy

assume the noise is weak
(i.e., deterministic dynamics dominates for short times)

focus on behavior in the vicinity of an equilibrium point
(the argument is valid for any orbit of the system)

1. consider the action of the deterministic dynamics in a neighborhood of a periodic orbit
2. consider the action of the noise
   as if the dynamics were absent
3. the noise and deterministic dynamics combined describe the noisy flow

linearized deterministic flow



$$x_{n+1} + z_{n+1} = f(x_n) + M_n\, z_n\,, \quad M_{ij} = \partial f_i / \partial x_j$$

in one time step a linearized neighborhood of x_n is

(1) advected by the flow

(2) transported by the Jacobian matrix M_n into a neighborhood given by the M eigenvalues and eigenvectors

covariance advection

let the initial density of deviations z from the deterministic center be a Gaussian whose covariance matrix is

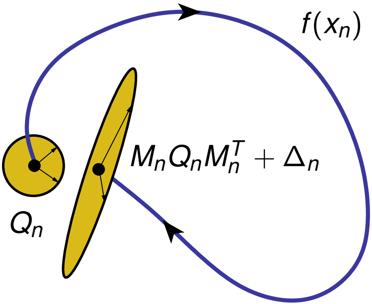
$$Q_{jk} = \langle z_j z_k^T \rangle$$

a step later the Gaussian is advected to

$$\begin{aligned} \langle z_j z_k^T \rangle &\rightarrow \langle (M\,z)_j\,(M\,z)_k^T \rangle \\ Q &\rightarrow M\,Q\,M^T \end{aligned}$$

next: add noise

roll your own cigar



in one time step
a Gaussian density distribution with covariance matrix Q_n is

(1) advected by the flow

(2) smeared with additive noise

into a Gaussian 'cigar' whose widths and orientation are given by the singular values and vectors of Q_{n+1}

covariance evolution

$$Q_{n+1} = M_n Q_n M_n^T + \Delta_n$$

(1) advect deterministically
local density covariance matrix $Q \to MQM^T$

(2) add noise covariance matrix Δ

covariances add up as sums of squares

cumulative noise along a trajectory

iterate $Q_{n+1} = M_n Q_n M_n^T + \Delta_n$ along a trajectory

if M is contracting, $|\Lambda_j| < 1$,

the memory of the covariance Q_0 of the starting density is lost, with iteration leading to the limit distribution

$$Q_n = \Delta_n + M_{n-1}\Delta_{n-1}M_{n-1}^T + M_{n-2}^2\Delta_{n-2}(M_{n-2}^2)^T + \cdots .$$

example : noise and a single attractive fixed point

if all eigenvalues of M are strictly contracting, all $|\Lambda_j| < 1$

any initial compact measure converges to the unique invariant Gaussian measure $\rho_0(z)$ whose covariance matrix satisfies

Lyapunov equation: time-invariant measure condition

$$Q = MQM^T + \Delta$$

[A. M. Lyapunov doctoral dissertation 1892]

example : Ornstein-Uhlenbeck process

width of the natural measure concentrated at the attractive deterministic fixed point $z = 0$

$$\rho_0(z) = \frac{1}{\sqrt{2\pi Q}} \exp\left(-\frac{z^2}{2Q}\right), \qquad Q = \frac{\Delta}{1 - |\Lambda|^2},$$

- is balance between contraction by Λ and noisy smearing by Δ at each time step
- for strongly contracting Λ, the width is due to the noise only
- As $|\Lambda| \to 1$ the width diverges: the trajectories are no longer confined, but diffuse by Brownian motion

example : statistical state dynamics

Brian Farrell

- assume SSD has a single attractive equilibrium
- truncate SSD to the first two cumulants: mean flow, perturbation state covariance C
- closure (drop higher cumulants) by Lyapunov equation

$$\frac{dC}{dt} = A\,C + C\,A^{\dagger} + \epsilon Q\,, \quad A = \text{linearized flow}$$

- works for Jupiter

example : 2D Brusselator limit cycle

214105-7 Nakanishi, Sakaue, and Wakou J. Chem. Phys. **139**, 214105 (2013)

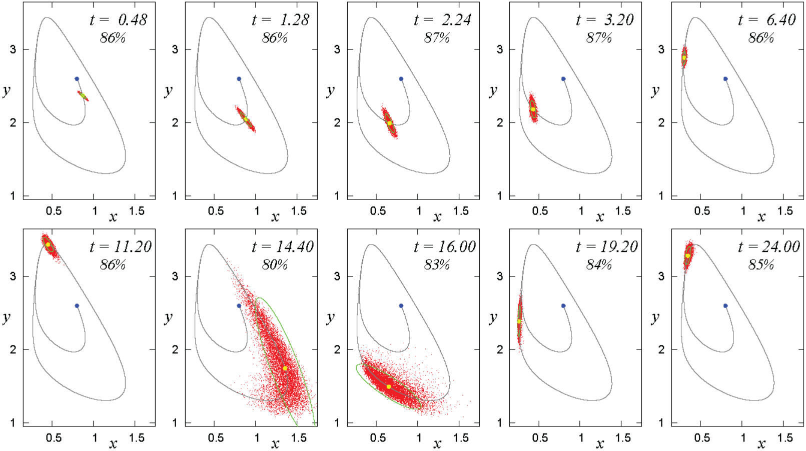


FIG. 2. Time development of distribution for Brusselator. 10 000 samples of Monte Carlo simulations are plotted by the red dots along with the covariance matrix $\hat{M}$ estimated by Eq. (E7); $\hat{M}$'s are represented by the green ellipses given by $\delta \boldsymbol{x}^T \hat{M}^{-1} \delta \boldsymbol{x} = 4/\Omega$, where $\delta \boldsymbol{x}^T \equiv (x - x^*(t), y - y^*(t))$. The percentages of the samples that fall within the ellipses are shown in each panel. The gray curves represent the trajectory by the rate equation starting from the initial point marked by the blue circles. The system parameters are $k_1 = 0.5$, $k_2 = 1.5$, $k_3 = 1.0$, $k_4 = 1.0$, and $\Omega = 10^4$. The initial point is $(x_0^*, y_0^*) = (0.8, 2.6)$.

remembrance of things past

noisy dynamics of a nonlinear system is fundamentally different from Brownian motion, as the flow ALWAYS induces a local, history dependent effective noise

things fall apart, centre cannot hold

but what if M has *expanding* eigenvalues?

both deterministic dynamics and noise tend to smear densities away from the fixed point: no peaked Gaussian in your future

things fall apart, centre cannot hold

but what if M has *expanding* eigenvalues?

look into the past, for initial peaked distribution that spreads to the present state

for unstable directions, look back

if M has only *expanding* eigenvalues,

balance between the two is attained by iteration from the past, and the evolution of the covariance matrix $\tilde{Q}$ is now given by

$$\tilde{Q}_{n+1} + \Delta_n = M_n \tilde{Q}_n M_n^T \,,$$

[aside to control theorists: reachability and observability Gramians]

solving the Lyapunov equation

iterate $Q_{n+1} = M_n Q_n M_n^T + \Delta_n$
attractive fixed point, $Q = Q_\infty$, $M = M_n$, $Q = Q_n$:

$$Q \;=\; \Delta + M\Delta M^\top + M^2\Delta(M^\top)^2 + \cdots = \sum_{m,n=0}^{\infty} \delta_{mn} M^n \Delta (M^\top)^m$$

bring to resolvent form, $\delta_{mn} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\theta(m-n)}$

for M contracting, expanding, or hyperbolic (!)

$$Q = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{\mathbf{1} - e^{-i\theta} M} \Delta \frac{1}{\mathbf{1} - e^{i\theta} M^\top}$$

Cauchy magic

a similarity transformation S separates the expanding and contracting subspaces

$$\Lambda \equiv S^{-1} M S = \left[\begin{array}{cc} \Lambda_e & 0 \\ 0 & \Lambda_c \end{array} \right]$$

transformed noise covariance matrix

$$\hat{\Delta} \equiv S^{-1} \Delta (S^{-1})^\top = \left[\begin{array}{cc} \Delta_{ee} & \Delta_{ec} \\ \Delta_{ce} & \Delta_{cc} \end{array} \right]$$

Cauchy magic

contour integral representation

$$Q = \oint_\Gamma \frac{ds}{2\pi} (\mathbf{1} - s^{-1} M)^{-1} \Delta (\mathbf{1} - sM)^{-1}$$

separates Q into expanding and contracting covariances:

$$\tilde{Q}_e \equiv S \begin{bmatrix} Q_e & 0 \\ 0 & 0 \end{bmatrix} S^\top , \quad Q_c \equiv S \begin{bmatrix} 0 & 0 \\ 0 & Q_c \end{bmatrix} S^\top$$

two stationary ‘cigars’, one in the expanding manifold and the other in the contracting manifold (not orthogonal to each other!)

local problem solved: can compute every cigar

a periodic point of period n is a fixed point of nth iterate of dynamics

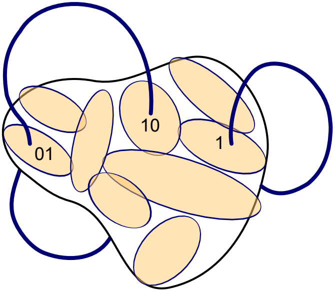
global problem solved: can compute all cigars

more algebra: can compute the noisy neighborhoods of all periodic points

noisy dynamics partitions: strategy

- use periodic orbits to partition state space
- compute local covariances at periodic points to determine their neighborhoods
- done once neighborhoods overlap

optimal partition hypothesis

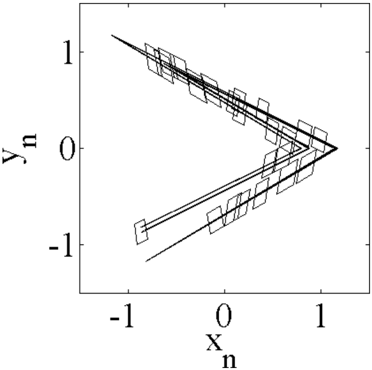


optimal partition:

the maximal set of resolvable periodic point neighborhoods

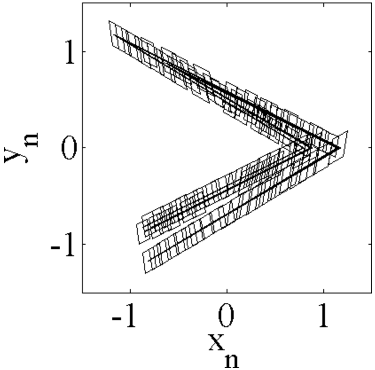
building the partition for the Lozi attractor

initial partition



periodic points of periods ≤ 5

final, optimal partition



no neighbors overlap $> 50\%$

payback

optimal partition: 10's to 100's of regions
uniform mesh: $\approx 10^6$ bins

application : long time averages of observables

if dynamics is chaotic
can predict accurately for long times only

expectation values of observable $a(x)$

$$\langle a \rangle = \int dx\, \rho(x)\, a(x)\,,$$

stationary distribution (natural measure) $\rho(x)$

=

probability of finding the system in state x

optimal partition Gaussian basis

stationary distribution is Fokker-Planck eigenfunction

$$\mathcal{L}_{FP}\, \rho(x) = \rho(x)\,.$$

stationary distribution, optimal partition basis approximation

$$\rho(x) = \sum_{a=1}^{N} h_a\, \phi_a(x)\,, \quad \phi_a = e^{-x_a^\top Q_a x_a}$$

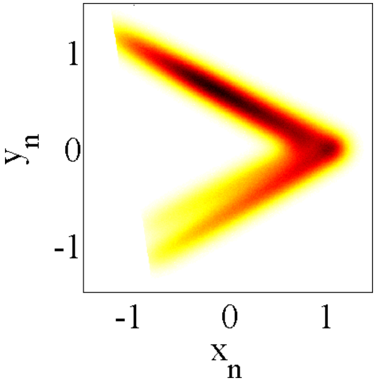
Gaussian basis functions, with “Lyapunov” covariances Q_a

coefficients $\{h_a\}$ determined by minimizing

$$\int \left(\sum_{a=1}^{N} h_a(\mathcal{L}_{FP} - 1)\phi_a(x) \right)^2 dx$$

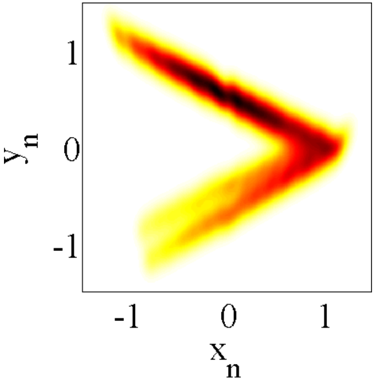
stationary probability distribution function

direct numerical calculation



uniform mesh of $\approx 10^6$ bins

optimal partition



Gaussian basis approximation

payback

L2 distance between approximation and exact $< 5\%$

better accuracy on expectation values of observables

take home message

computation of unstable periodic orbits in high-dimensional state spaces, such as Navier-Stokes,

is at the border of what is feasible numerically, and criteria to identify finite sets of the most important solutions are very much needed

we are to stop calculating these solutions when we attain

take home message

optimal partition

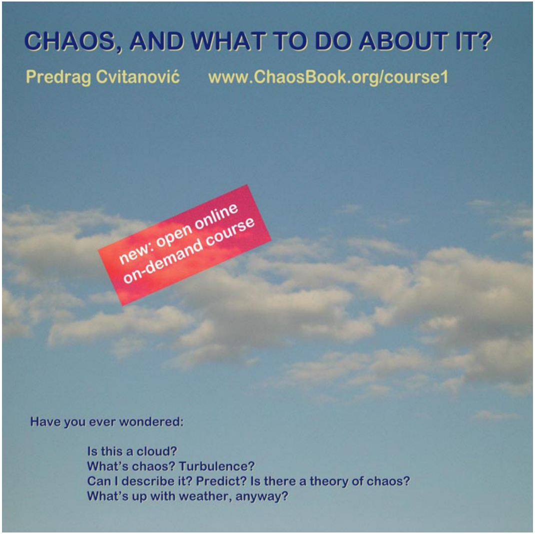
optimal partition

- the best of all possible state space partitions
- optimal for the given dynamical system, the given noise

references

- D. Lippolis and P. Cvitanović, *How well can one resolve the state space of a chaotic map?*, Phys. Rev. Lett. 104, 014101 (2010); arXiv.org:0902.4269
- P. Cvitanović and D. Lippolis, *Knowing when to stop: How noise frees us from determinism*, in M. Robnik and V.G. Romanovski, eds., *Let's Face Chaos through Nonlinear Dynamics* (Am. Inst. of Phys., Melville, New York, 2012); arXiv.org:1206.5506
- J. M. Heninger, D. Lippolis and P. Cvitanović, *Neighborhoods of periodic orbits and the stationary distribution of a noisy chaotic system*; arXiv.org:1507.00462

what next? take the course!



student raves :
...10^6 times harder than any other online course...