Reconstruction Deconstruction:

A Brief History of Building Models of Nonlinear Dynamical Systems

Jim Crutchfield

Center for Computational Science & Engineering

Physics Department

University of California, Davis

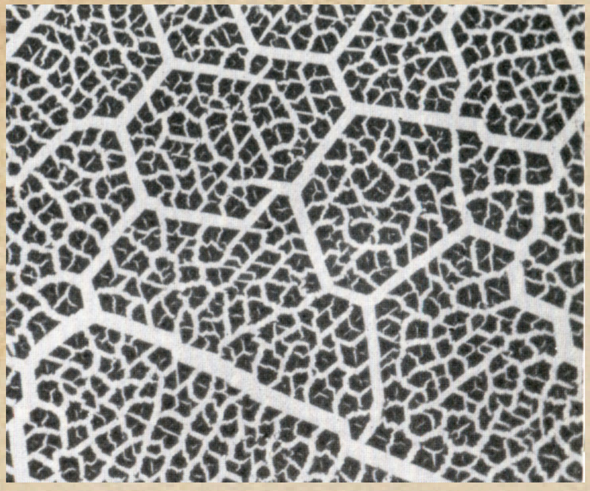
Neural Information Processing 2006 Workshop on

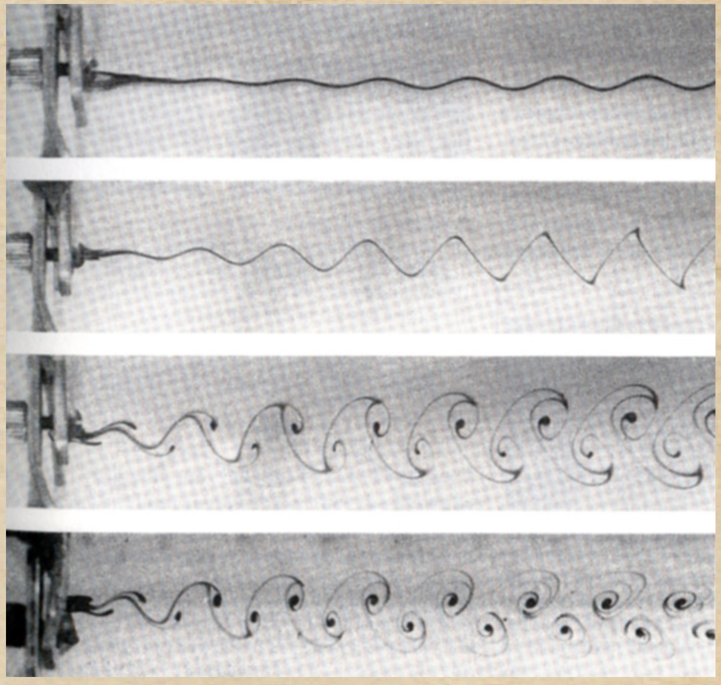
Revealing Hidden Elements of Dynamical Systems

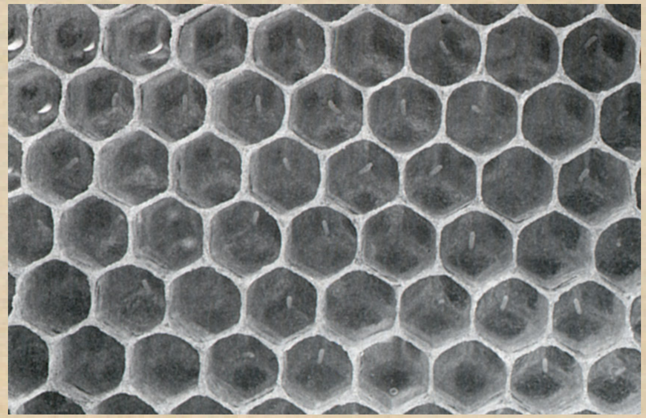
8 December 2006

Whistler, British Columbia

The Problem







The Goal:
Artificial Science

Can we automate scientific discovery and theory building?

The Guess:

Yes!

Recent advances in nonlinear dynamics

+

Recent advances in machine learning

+

Lots of computational power

The Guess:

Yes!

Recent advances in nonlinear dynamics:

Why?

Better sense of the manifestations of nonlinearity

(info. gen., self-similarity, hierarchy, ...)

The Guess:

Yes!

Recent advances in machine learning:

Why?

Information-theoretic view of learning

+

Algorithms: Structure v. error

The Guess:

Yes!

Lots of computational power:

Why?

We don't yet know how the brain does it so quickly

Agenda

- Nonlinear Dynamics:
  - History & Geometry
- Reconstruction and Nonlinear Model Building
  - Continuous-time & -value
  - Discrete-time & -value

Nonlinear Dynamics: Selected Historical Highlights

- Poincare (1892): Discovery of deterministic chaos
- van der Pol (Nature 1927): Expt'l discovery of chaos
- Soviet school (1910-1960): Lyapunov and students
- Kolmogorov (1957): Info theory + dynamical systems
- Modern era: 1960s onwards ...
    - Ulam-Lorenz (1962-63): Experimental mathematical studies
    - 1970s-1980s: Simple systems can be complicated
    - 1980s-1990s: Complicated systems can be simple (pattern formation)

Nonlinear Dynamics: Mathematical Tools

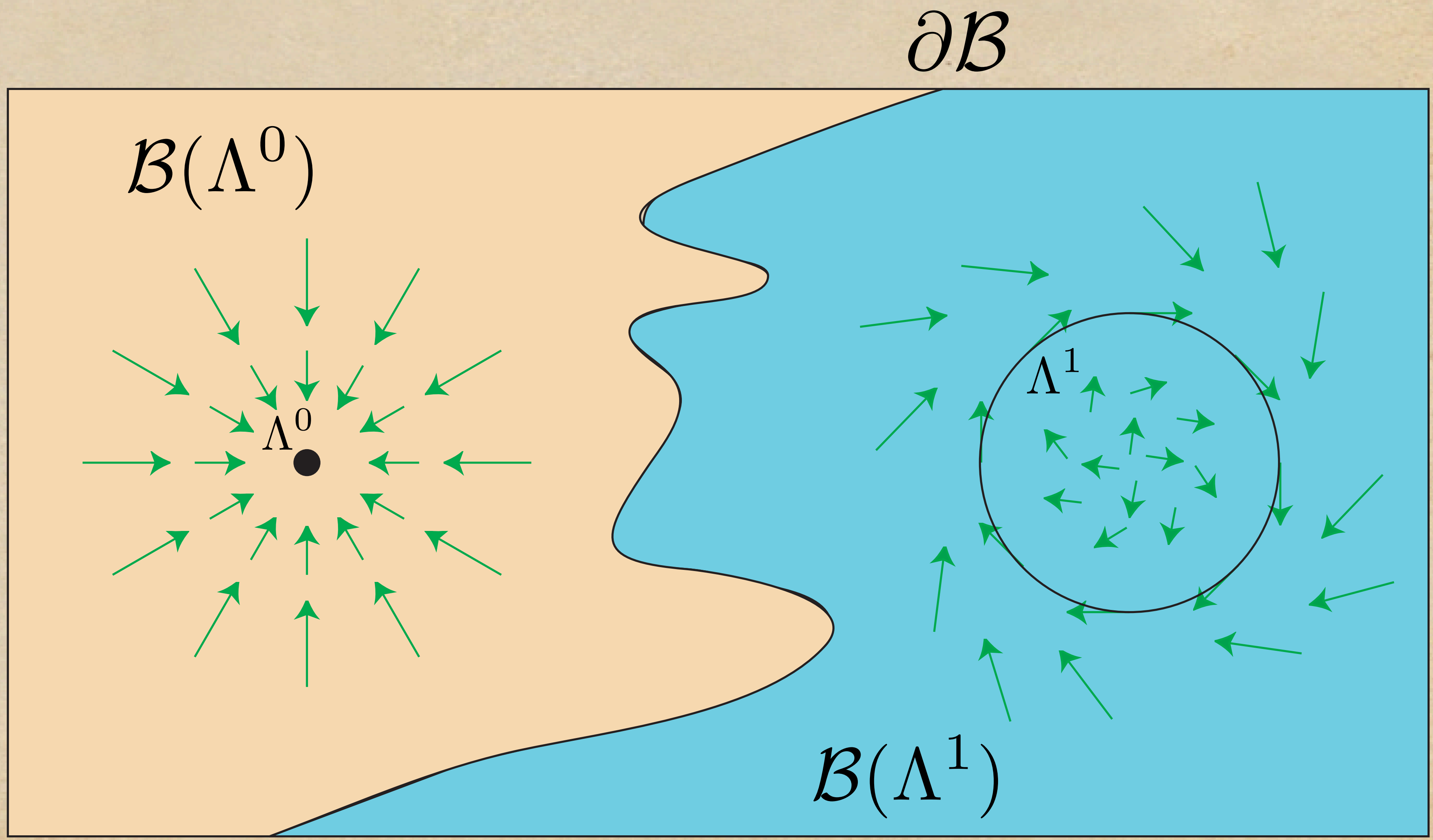
- Statistical mechanics; incl. phase transitions
- Pattern formation: Center manifold theory
- Nonlinear dynamics:
  - Qualitative dynamics
  - Bifurcations: Singularity theory

Nonlinear Dynamics: Geometric Review

Dynamical System:
- State space
- Dynamic
- Initial condition

The Attractor-Basin Portrait: The Big Picture

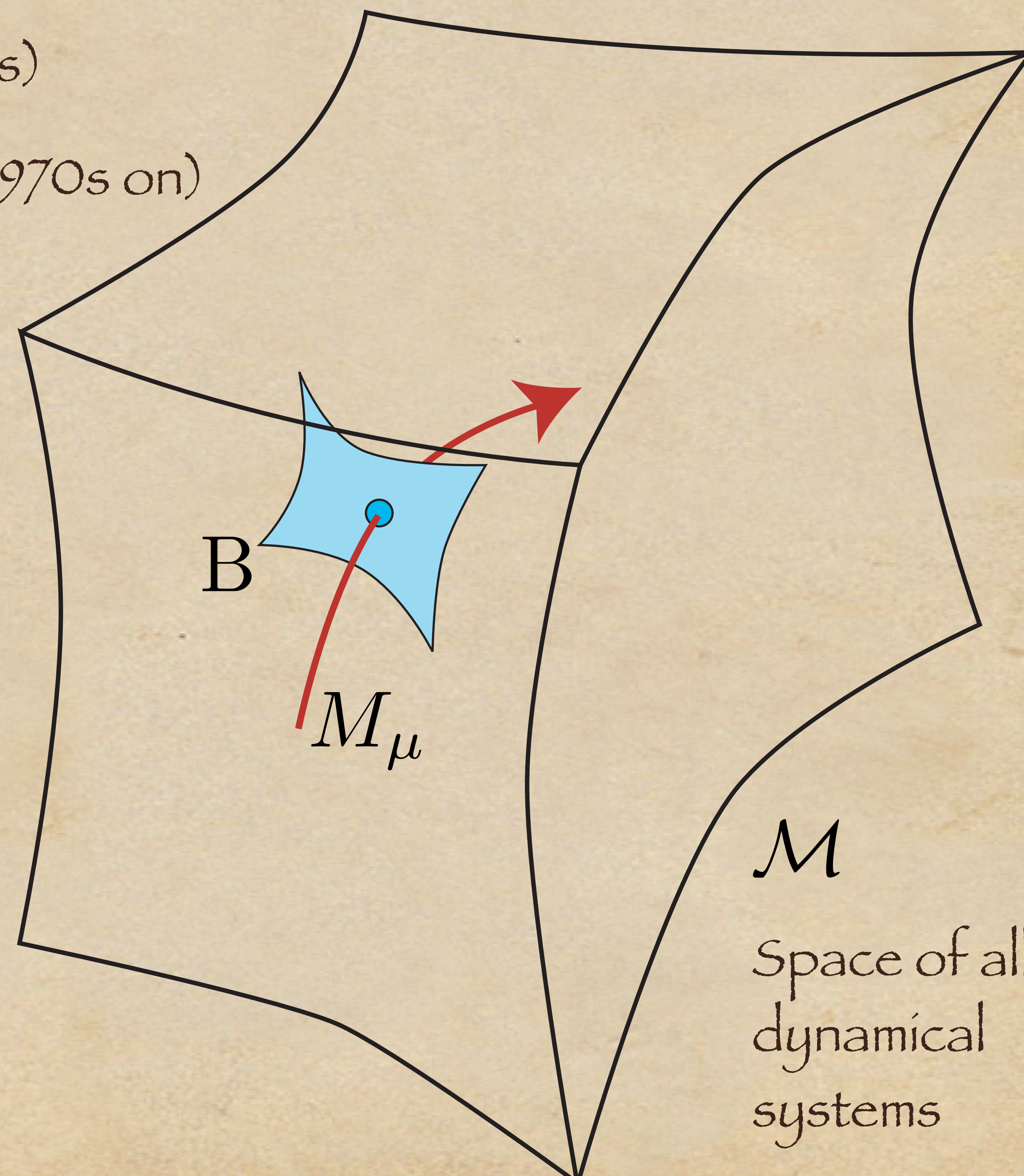
Basins, Attractors, Separatrices



Nonlinear Dynamics: Review ...

- Bifurcations: The Big, Big Picture
  - Catastrophe theory (Thom, 1960s)
  - Singularity theory (Arnold et al, 1970s on)

What happens when you change control parameters?



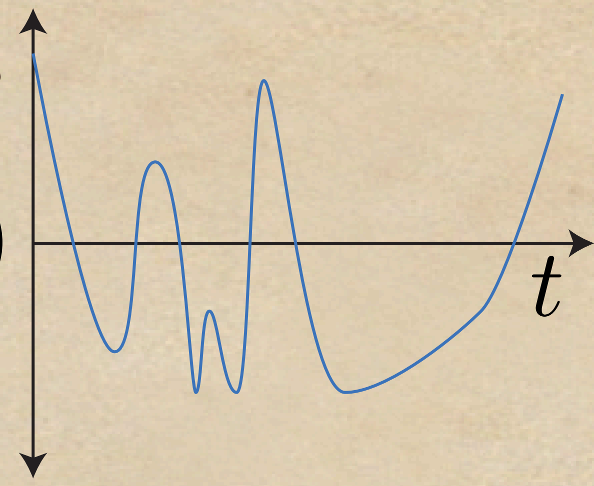
Reconstruction:

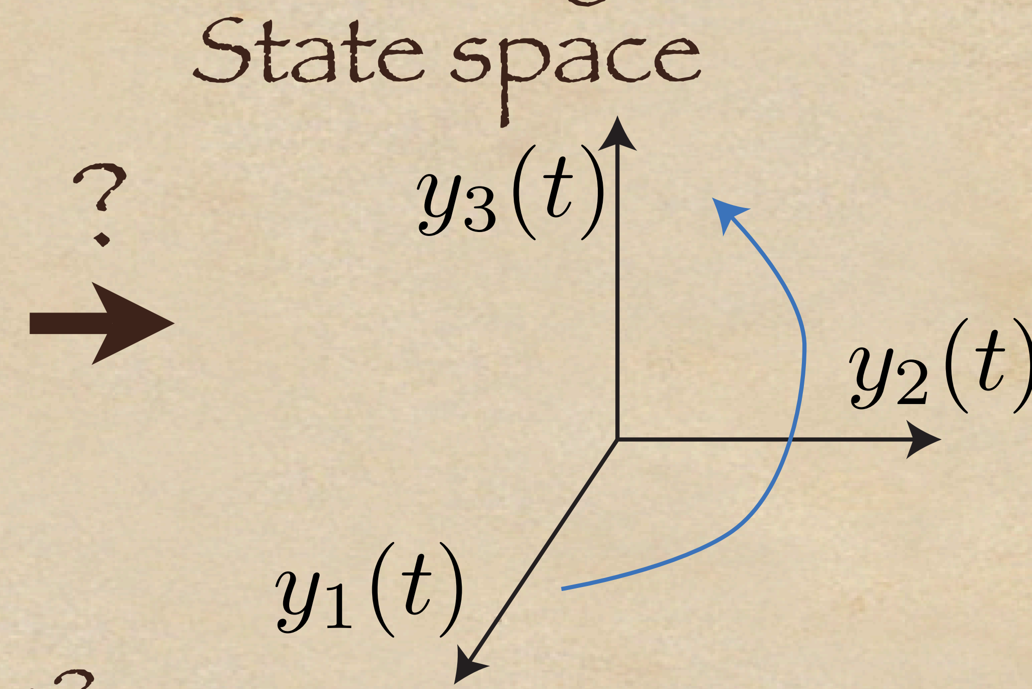
How to build a model of nonlinear system?

- Goals:
  - What are the states?
  - What are the equations of motion?
- Two cases:
  - Continuous measurements
  - Discrete measurements

Reconstruction: Continuous Dynamics

- Time series





- Reconstructed state space?
  - Derivative embedding: $\vec{y}(t) = (x(t), \dot{x}(t), \ddot{x}(t), \ldots)$
  - Time-delay embedding:

$$\vec{y}(t) = (x(t), x(t-\tau), x(t-2\tau), \ldots)$$

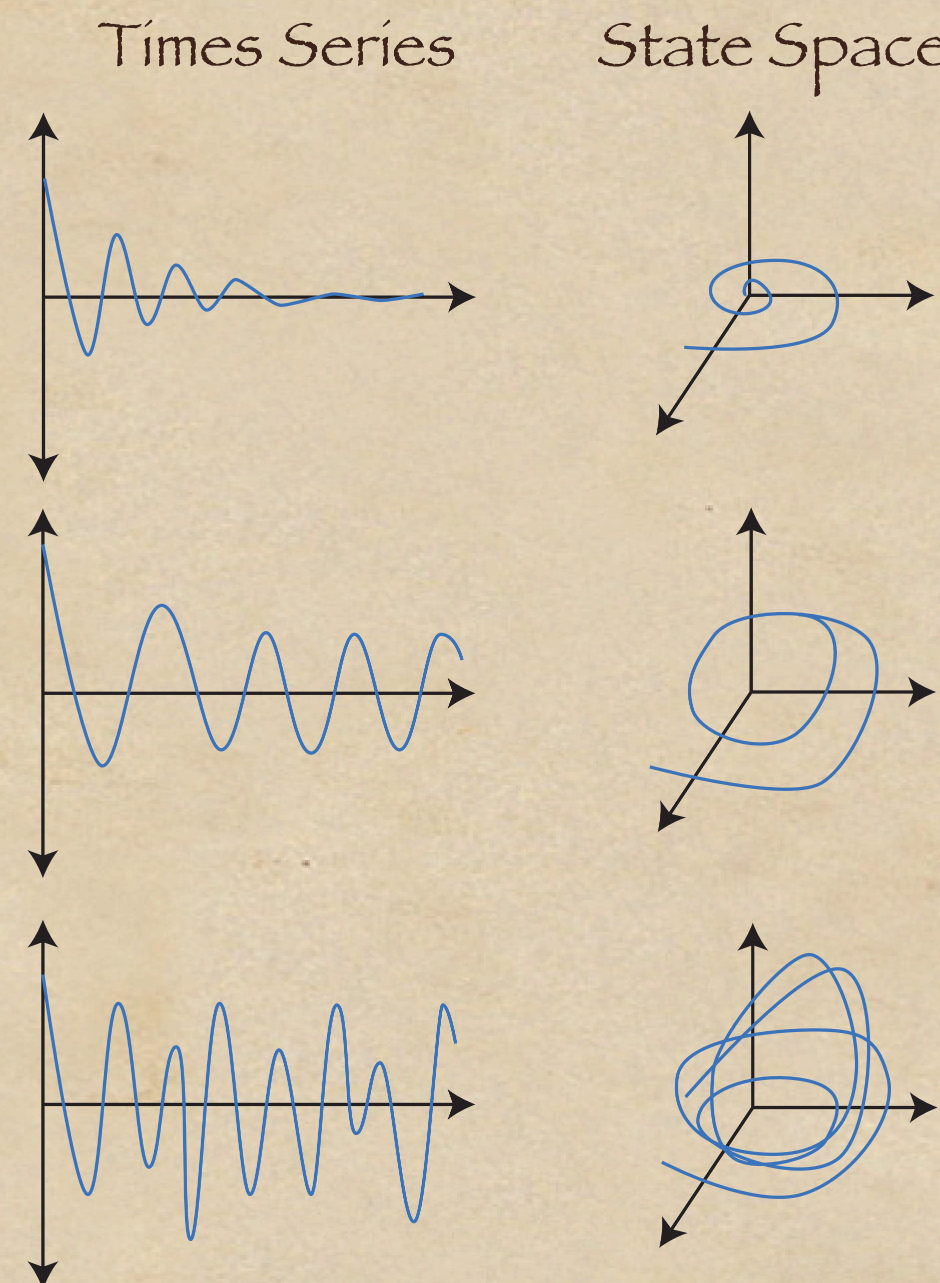
  - Embedding dimension: Number of active degrees of freedom

N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a Times Series", Physical Review Letter 45 (1980) 712-715.

F. Takens, "Detecting Strange Attractors in Fluid Turbulence", in *Symposium on Dynamical Systems and Turbulence*", D. A. Rand and L. S. Young, editors, Lect. Notes Math. 898, Springer-Verlag (Berlin 1981).

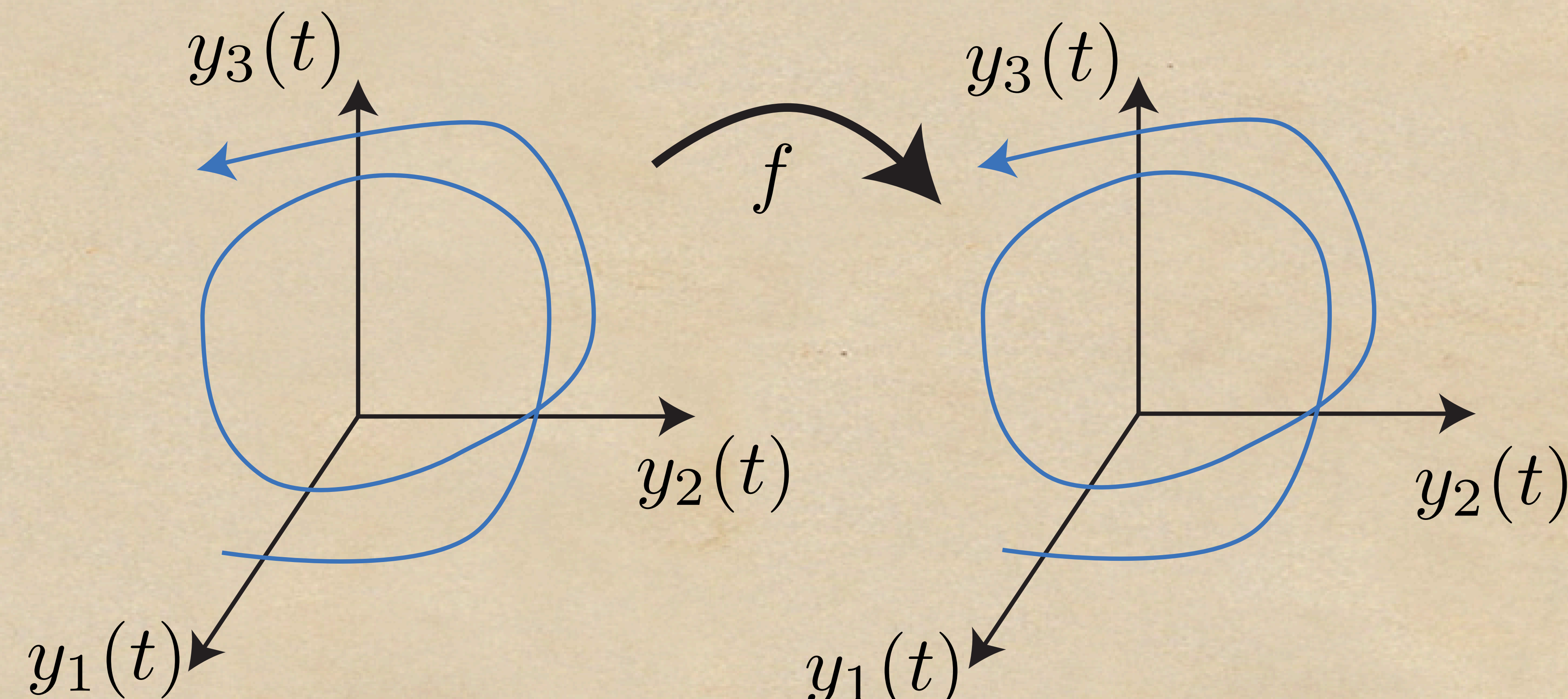
Reconstruction: Continuous Dynamics

- Fixed point
- Limit Cycle
- Chaotic Attractor



Reconstruction: Continuous Dynamics

- Equations of motion? $\dot{\vec{y}} = \vec{f}(\vec{y})$
- Find $f : \vec{y}(t) \to \vec{y}(t + dt)$



J. P. Crutchfield and B. S. McNamara, "Equations of Motion from a Data Series", Complex Systems 1 (1987) 417 - 452.

Reconstruction: Symbolic Dynamics

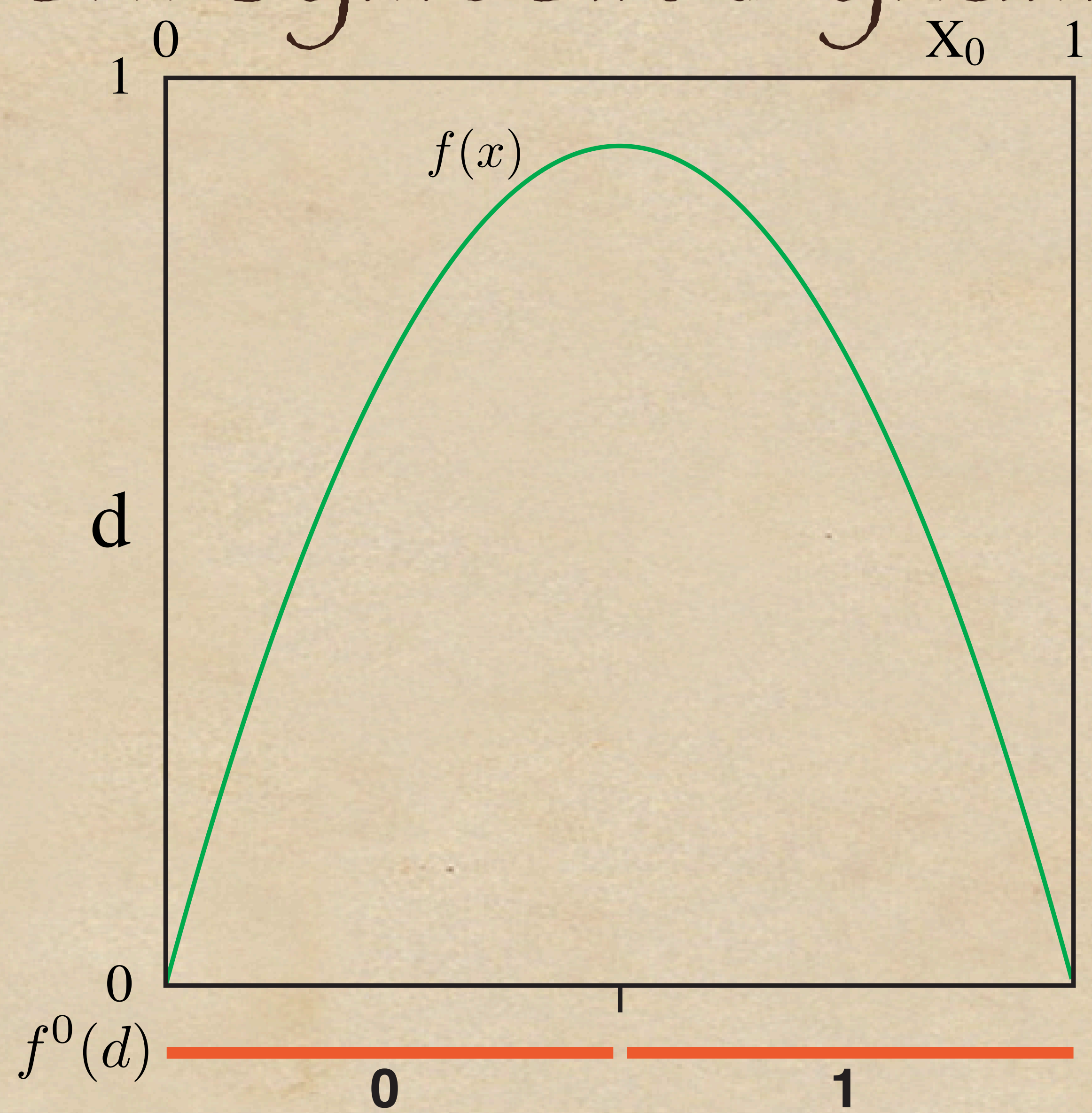
Simplest Dynamical System:

State space:

$$x \in [0, 1]$$

Equations of motion:

$$x_{n+1} = f(x_n)$$



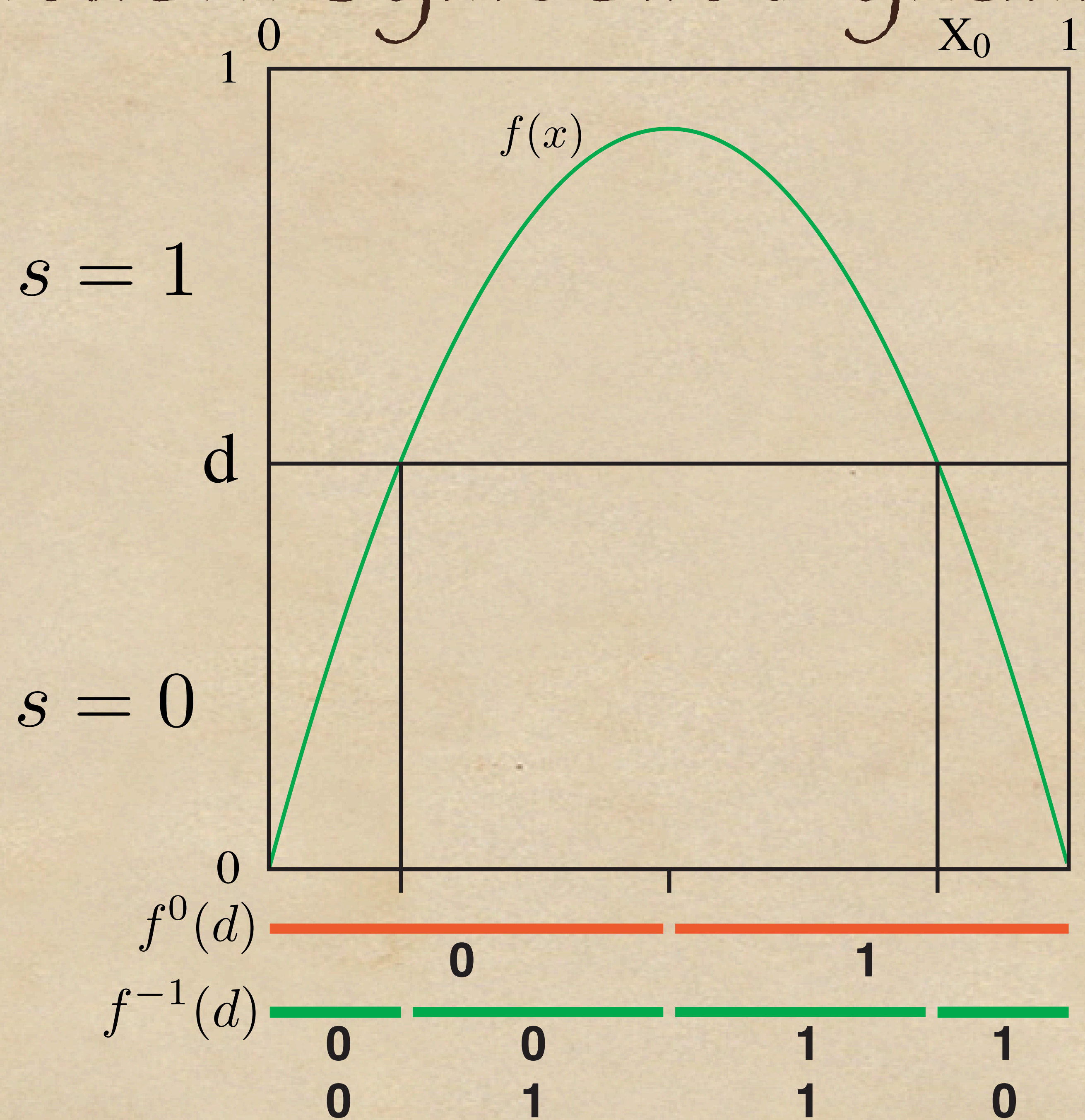
Binary instrument:

$$\mathcal{P} = \{0 \sim x \in [0, d], 1 \sim x \in (d, 1]\}$$

Decision point: $d \in [0, 1]$

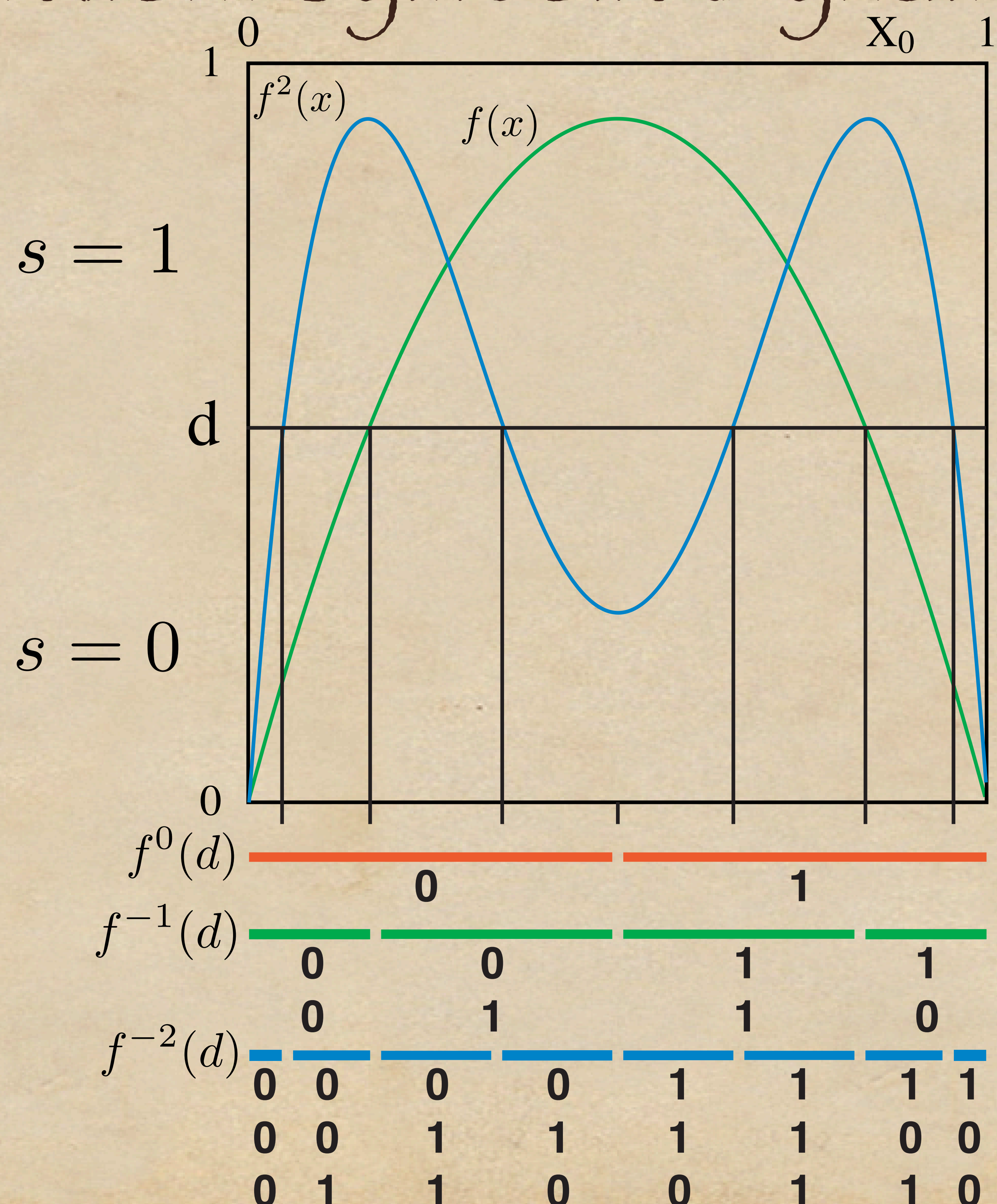
Reconstruction: Symbolic Dynamics

One time step



Reconstruction: Symbolic Dynamics

Two time steps



Reconstruction: Symbolic Dynamics

When are partitions good?

When symbol sequences encode orbits

$$\begin{tikzcd} M \arrow[r, "\mathcal{T}"] & M \\ \mathcal{A}^{\mathbb{Z}} \arrow[u, "\Delta"] \arrow[r, "\sigma"] & \mathcal{A}^{\mathbb{Z}} \arrow[u, "\Delta"] \end{tikzcd}$$

Diagram commutes:

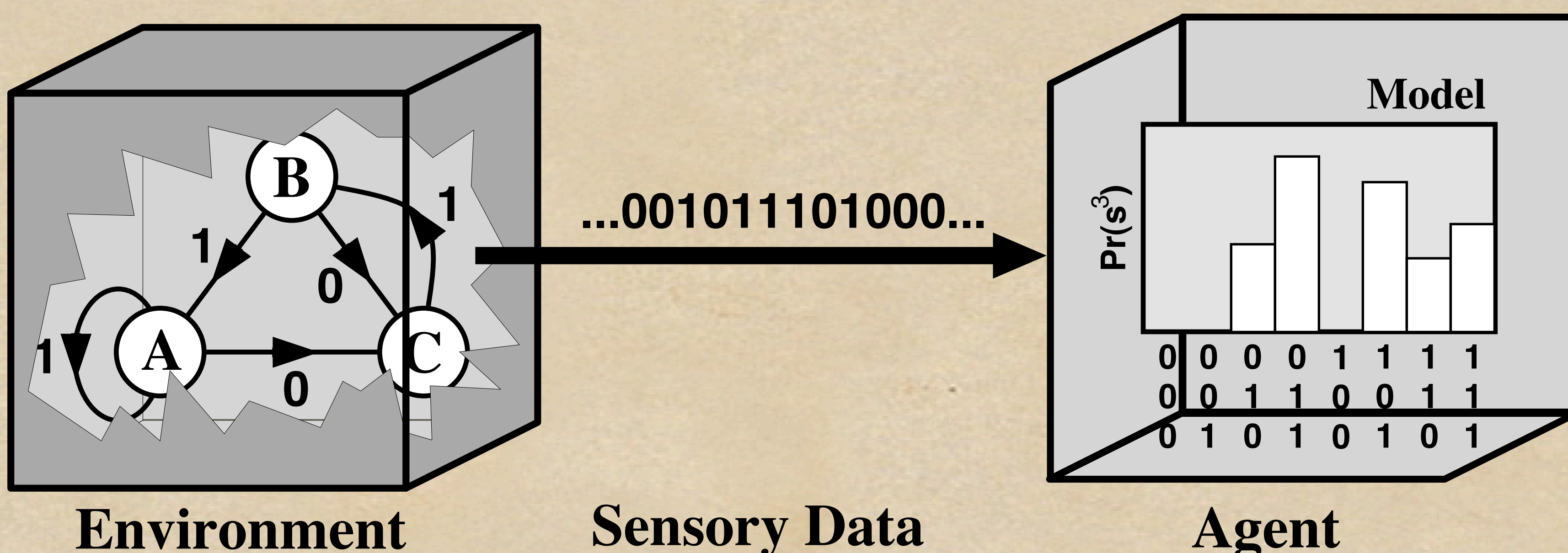
$$\mathcal{T}(x) = \Delta \circ \sigma \circ \Delta^{-1}(x)$$

Good kinds of instruments:

Markov partitions

Generating partitions

Measurement Channel



Goals: (i) States? (ii) Dynamic?

Computational Mechanics

How to circumvent representation choice?

Is there a preferred representation?

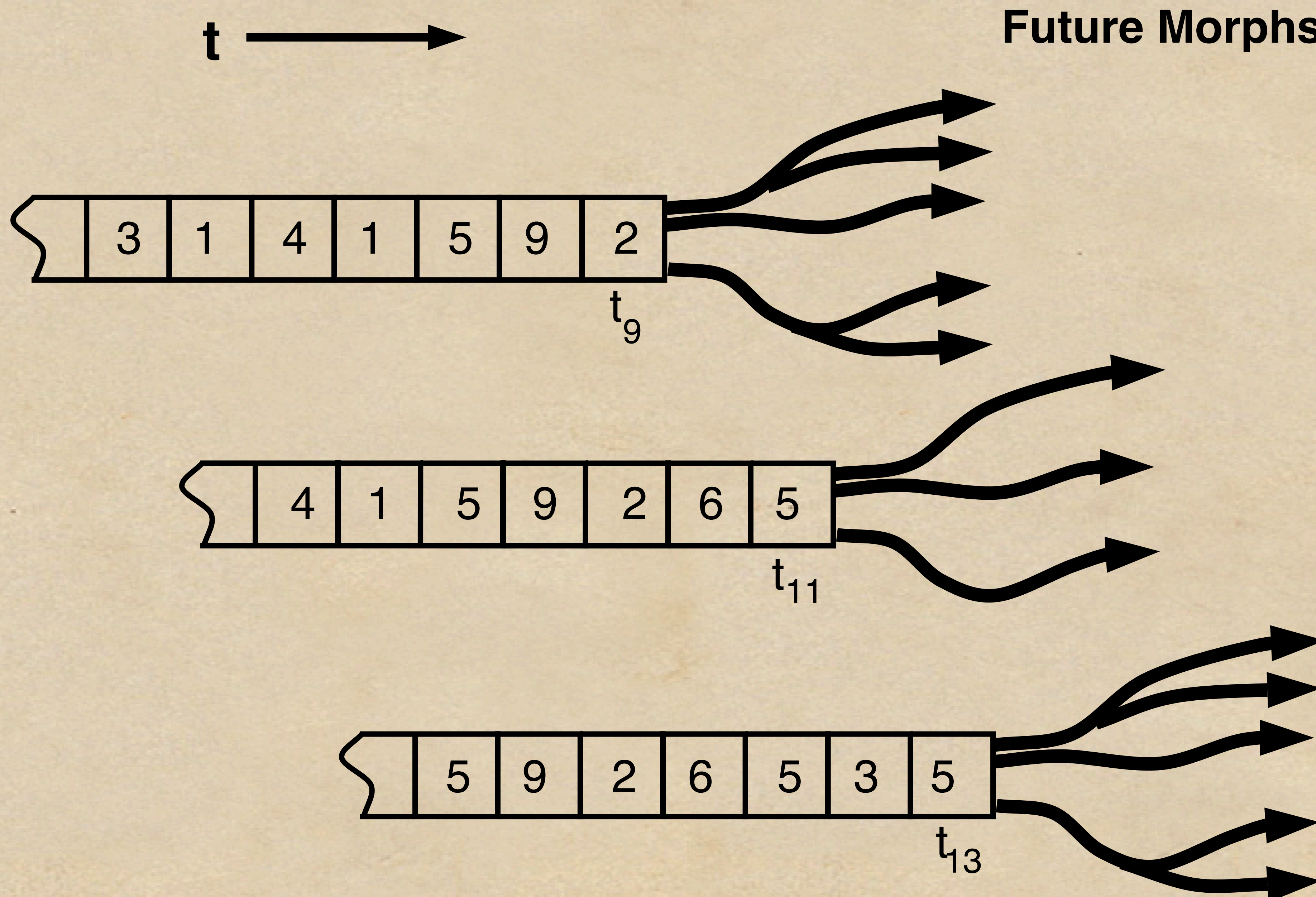
J. P. Crutchfield and K. Young , "Inferring Statistical Complexity",
Physical Review Letters 63 (1989) 105-108.

J. P. Crutchfield , "The Calculi of Emergence: Computation, Dynamics, and Induction",
Physica D 75 (1994) 11-54.

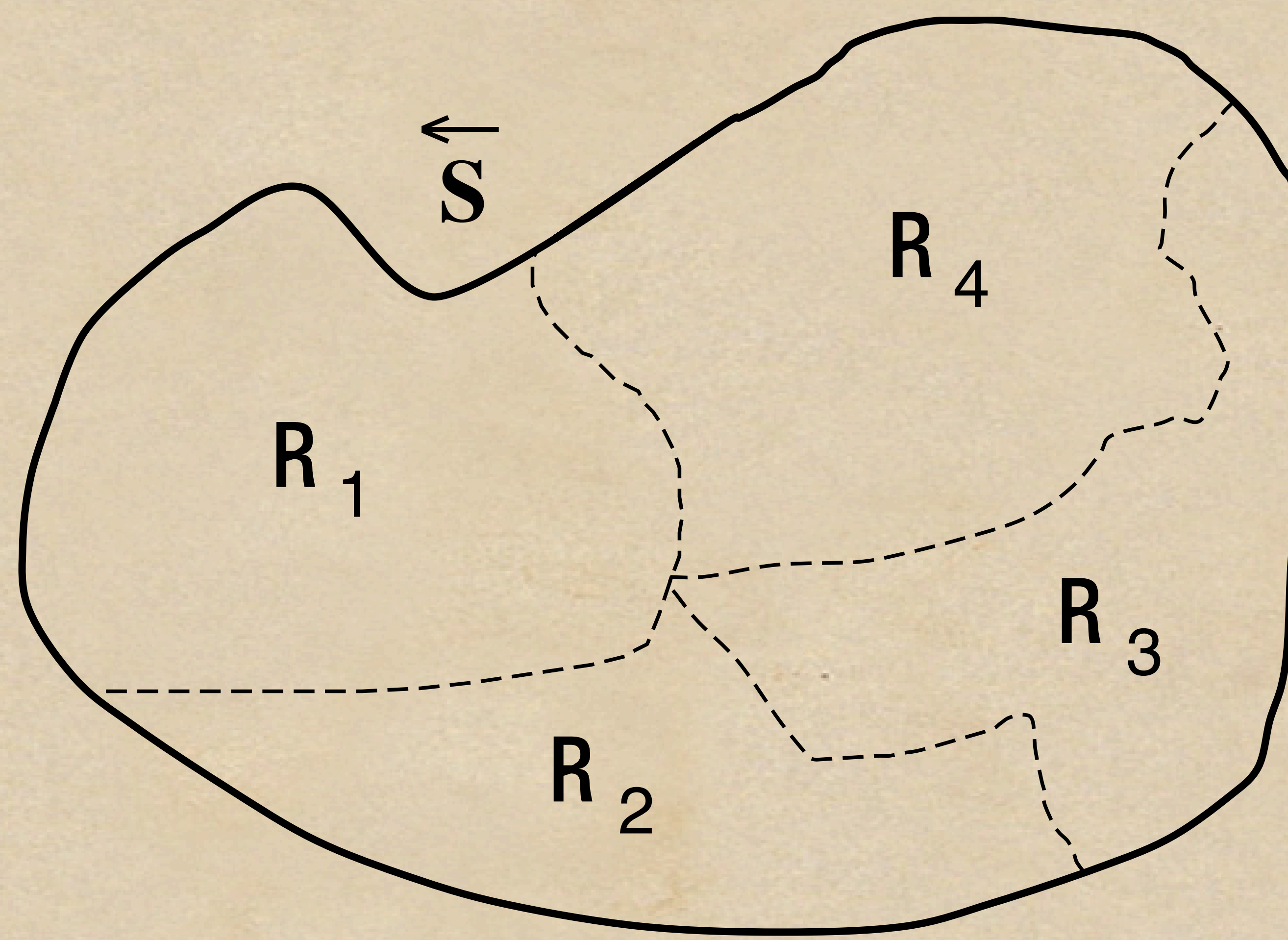
C. R. Shalizi and J. P. Crutchfield, "Computational Mechanics: Pattern and Prediction, Structure and Simplicity",
Journal of Statistical Physics 104 (2001) 817-879.

N. Ay and J. P. Crutchfield, Reductions of Hidden Information Sources",
Journal of Statistical Physics 210:3-4 (2005) 659-684.

Effective States

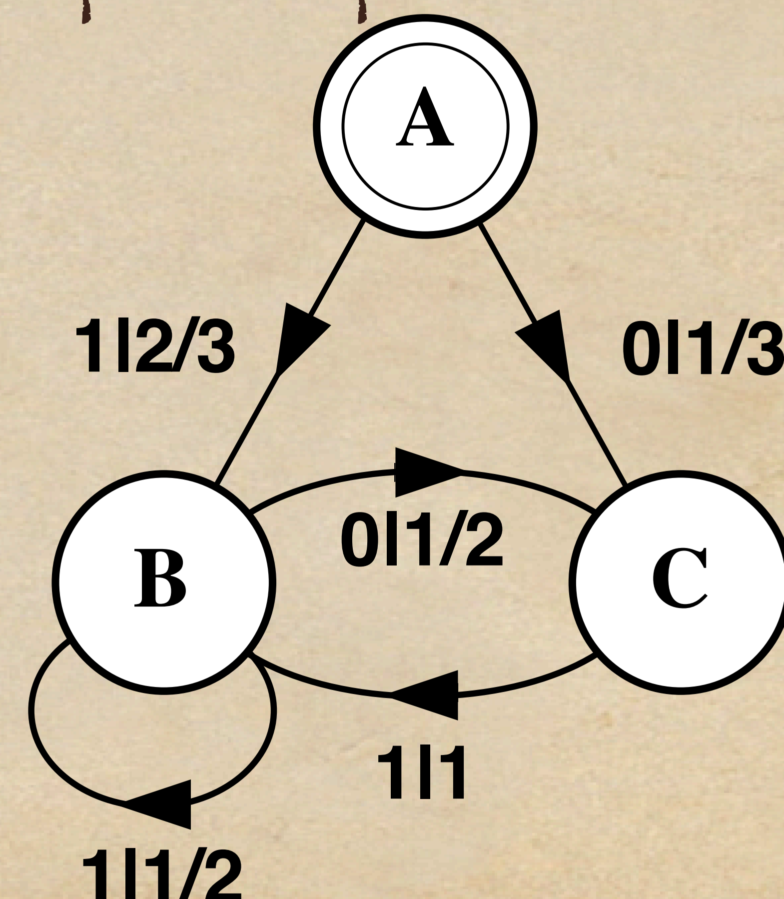


Space of Histories Partitioned



Causal States

- $\overleftrightarrow{S} = \overleftarrow{S}\,\overrightarrow{S}$
- $\overleftarrow{S} \sim \overleftarrow{S}' \Leftrightarrow \Pr(\overrightarrow{S} \mid \overleftarrow{S}) = \Pr(\overrightarrow{S} \mid \overleftarrow{S}')$
- ε-Machine = $\{\mathcal{S}, \mathcal{T}\}$
- Theorems: Unique, minimal, & optimal predictor
- Form = ε-Machine semigroup



Stored Information (1989)

$$C_\mu = -\sum_{\mathcal{S}\in\mathbf{S}} Pr(\mathcal{S})\log_2 Pr(\mathcal{S})$$

Information Production (1948)

$$h_\mu = -\sum_{\mathcal{S}\in\mathbf{S}} Pr(\mathcal{S}) \sum_{\mathcal{S}'\in\mathbf{S}} Pr(\mathcal{S}'|\mathcal{S})\log_2 Pr(\mathcal{S}'|\mathcal{S})$$

Excess Entropy (1982)

(all-point mutual information)

$$\mathbf{E} = I(\overleftarrow{S}, \overrightarrow{S})$$

(Thm: $\mathbf{E} \leq C_\mu$)

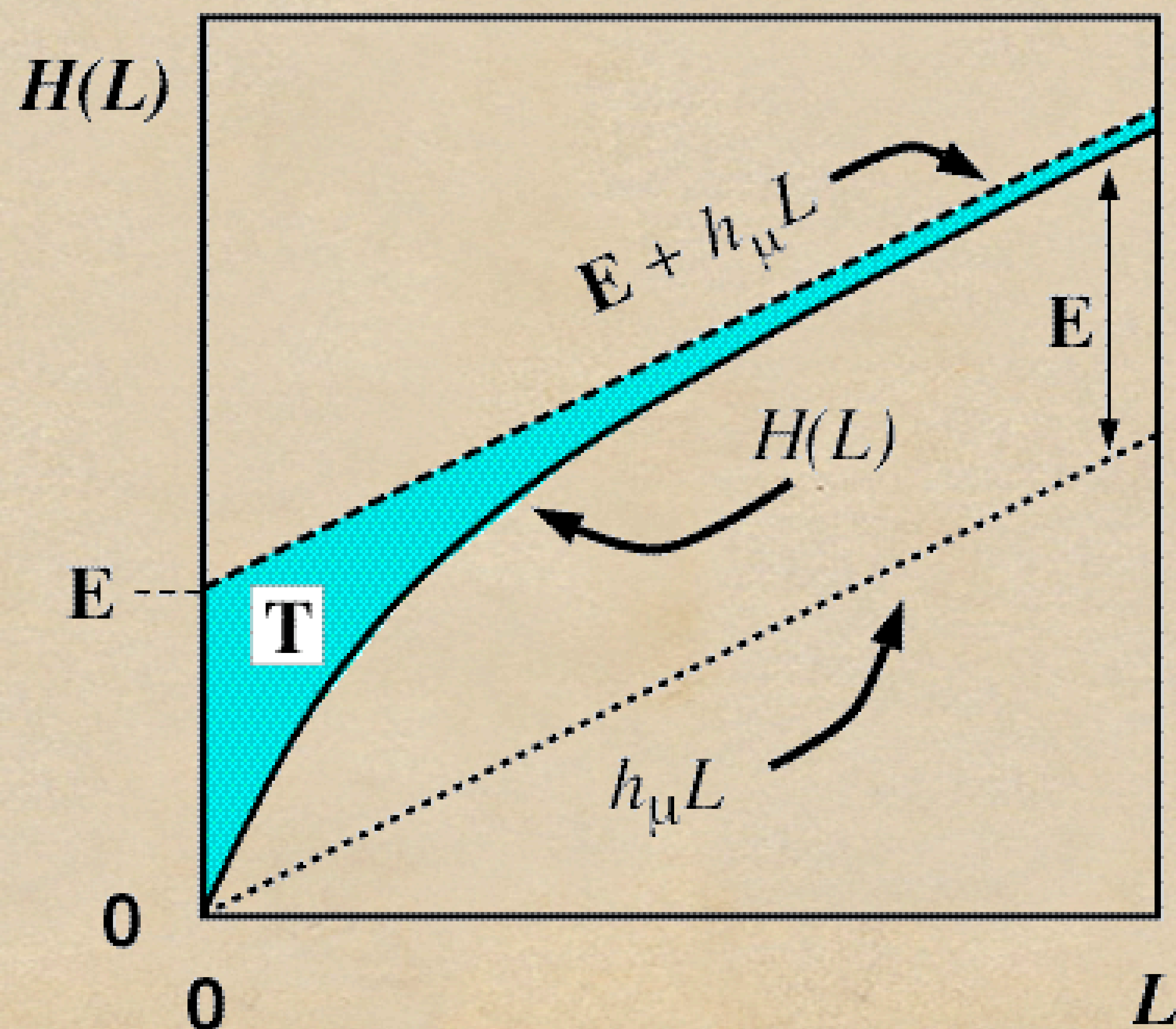
Sync-Transient Information (2005)

$$\mathbf{T} = \sum_{L=0}^{\infty} [\mathbf{E} + h_\mu L - H(L)]$$

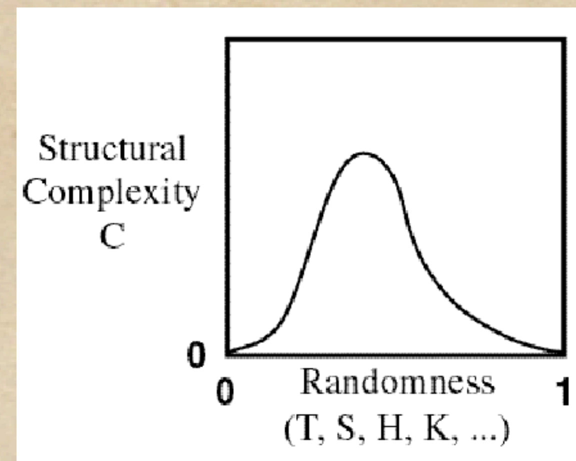
J. P. Crutchfield and D. P. Feldman, "Regularities Unseen, Randomness Observed: Levels of Entropy Convergence", CHAOS 13:1 (2005) 25-54".

Entropy Hierarchy

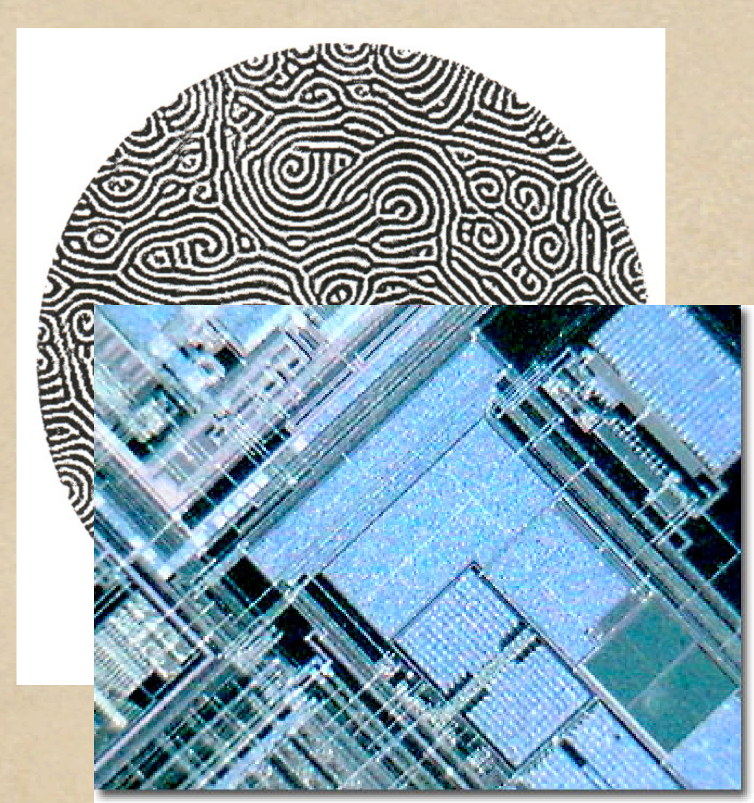
Entropy growth H(L)



One Lesson: Structure v. Randomness



Boredom Delight Confusion



Variatio Delectat

Concluding Remarks

- Yes, there is a preferred representation:

    ε-Machine: Causal architecture

- Structure can be formalized (and quantified)
- Intrinsic computation:

    Structure = how a system stores & processes info

- Question:

    How does all this affect modeling nonlinear systems?

Background

Chaos and Time-Series Analysis by J. C Sprott (2003)

Nonlinear Time Series Analysis: Second Edition by H. Kantz and T. Schreiber (2006)

Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers by R. Hilborn (2001)

Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering by S. H. Strogatz (2001)

Advert: Chris Strelioff @ Saturday Workshop

Bayesian Estimation of Information Production

Thanks!

Panel Discussion

- In what cases do you think the approach you propose would be better compared to other approaches? Where would it be worse?
- What are the special features of dynamical system learning that separate it from other learning problems?
- How can statistical and information theoretic techniques be combined with the theoretical structure of dynamical systems?
- Are there generic features that can be extracted from time-series data?
- How can we combine static and time series data for modeling dynamic systems?

Panel Discussion

- In what cases do you think the approach you propose would be better compared to other approaches?
    - Goal is scientific understanding. Learning *the* laws!
    - Results come from verified predictions, not a competition.

- Where would it be worse?
    - Do we have a choice?

Panel Discussion

- What are the special features of dynamical system learning that separate it from other learning problems?
  - Time, time, and time.
  - Please: NO MORE IID SAMPLING!
  - Use deep math'l knowledge to constrain search/estimation:

    Ergodicity, dissipation, genericity, structural stability, bifurcations, smoothness, ...

Panel Discussion

- How can statistical and information theoretic techniques be combined with the theoretical structure of dynamical systems?
    - Information theory *is* a fundamental part of dynamical systems theory
    - Historically and practically.
    - Interesting Q: How do the information storage and production properties of dynamical systems affect the learning complexity?

Panel Discussion

- Are there generic features that can be extracted from time-series data?
    - Yes.
    - System invariants: Entropy rate, embedding dimension, statistical complexity, excess entropy, ...
    - Ensembles of experiments: Equations of motion, component forces, ...

Panel Discussion

- How can we combine static and time series data for modeling dynamic systems?
- Fussy note: "Dynamic systems" is not "Dynamical systems".
- Former linear, latter unconstrained.
- Not sure about question, though: Is this about state-space averages versus time averages? (I.e., about ergodicity?)