Persistence of dynamics in high-dimensional dynamical systems

D. J. Albers

January 9, 2006

Goals of this talk

i. complex natural phenomena (often) have “consistent” dynamics, we wish to understand this geometrically;

ii. an intuitive picture of the language an framework of mathematical dynamics;

iii. an understanding of how this language expresses characterization of dynamics and defines it’s problems;

iv. present previous mathematical results and computational results (done with Clint and Jim);

Poincaré's dream

i. scientists study nature via mathematical models (ODEs, PDEs, discrete time maps);

ii. the solutions of these mathematical models are built from two components: existence of solutions (existence and uniqueness theorems) and approximation theorems;

iii. the approximation theorems say (often) that a particular function space contains a mapping that can approximate the solution to arbitrary accuracy;

the dream: study nature via a qualitative study of the dynamical behaviors of mappings in the function spaces that form solutions to our models (for systems where *time* is a parameter) then characterized common behaviors and what assumptions or geometric structures constrain these common behaviors;

Linking the qualitative mathematical framework with the natural world: inconsistencies between science and mathematics

i. diversity in natural systems — how in the heck do we characterized "common" dynamics;

ii. example: turbulence and multiple attractors;

iii. high-entropy, high-dimensional, complex dynamics — turbulence versus spatially extended dynamics;

iv. devil is in the details — C^r mappings may not reflect natural systems;

v. measures in a post-modern world: empirical science is cast the language of probability (which requires a measure, can be dependent on the observer) and C^r function space does not admit a measure (no probability language);

vi. many "computational models" are not persistent to perturbations of the mapping or geometry;

The world is not Euclidean?

i. manifolds;

ii. embedding theorems;

iii. orthogonality (splitting or partitioning of the space);

iv. metric or distance function (yard stick);

C^r function space — and perturbations

i. continuous, differentiable r times with a continuous derivative.

ii. perturbations are of the *GRAPH* of the mapping.

iii. "open balls" in the C^r Whitney topology — just think "nearby graphs;"

iv. no distinction between "functional" or "parameter" perturbations because the key identifying object is the graph, not a particular mapping;

v. requires *infinitely* many parameters.

vi. these are perturbations of the geometric structure of the mapping;

Attractors, invariant sets and basins of attraction

i. attractor: largest set of the state space that is invariant under the mapping and composed of iterations of the mapping (meaning it is locally attracting) — only exists for dissipative (area-contracting) dynamical systems;

ii. area-preserving case — there is no “attractor,” just an invariant set;

iii. the set of initial conditions that lead to the same attractor is called the basin of attraction;

Stable, center, and unstable manifolds

i. Split the (derivative) manifold into orthogonal, contracting, neutral, and expanding “parts” relative to a metric (you need a metric to define the derivative or a rate).

Examples:

i. Pendulum

ii. Anosov diffeomorphisms

Hyperbolicity and partial hyperbolicity - quantification of stables, centers and unstables

i. invariant splitting of the "tangent space;"

ii. bounds on expansion, contraction;

iii. can be uniform along the orbit (uniform hyperbolicity);

iv. can be non-uniform along the orbit (non-uniform hyperbolicity);

v. partial versus strict hyperbolicity; centers versus no centers;

Lyapunov exponents, invariant and SRB measures, and attractors

i. think about the deformation of a frame and associated “ball” along trajectory;

ii. Lyapunov exponents (LCEs) require a measure; why? “integration” of a volume element;

iii. non-dissipative — absolute continuity with respect to Riemannian volume (f is absolutely continuous) or Lebesgue measure (logistic map at ($a = 4$)); there can only one measure;

iv. dissipative — absolutely continuous with respect to Riemannian volume on the *unstables* only; referred to as SRB measures; there can be many different such measures;

v. every attractor (manifold if non-dissipative) is associated with a measure; a dynamical systems with multiple (SRB) measures means multiple attractors.

vi. every non-zero Lyapunov exponent is associated with a stable or unstable manifold;

vii. K-S entropy, some complexity measures and be associated with sums of the positive LCEs;

Notions of equivalence and associated conjectures

i. C^r, $r = 0$, $r = 1$, etc equivalence.

ii. Ergodicity.

iii. Statistical stability.

iv. Persistent chaos of degree$-p$

Notions of commonality

i. topological (open-dense) — no notion of probability;

ii. measure theoretic — has a notion of probability but is dependent on the measure (observer);

iii. robust — property persistent on an open set in the ambient space;

Structural stability

i. "topological" equivalence (C^0 equivalence) — coffee cup versus donut;

ii. f is C^r structurally stable if small perturbations (of the graph) of f and it's derivatives up to order r remain C^0 equivalent to f.

iii. neutral directions (centers) are not allowed;

Conjecture/theorem

A C^1 diffeomorphism is structurally stable if and only if it satisfies axiom A (it is hyperbolic and has dense periodic points) and satisfies the so called strong transversality condition.

This has been shown to not be dense in C^r; for $r > 1$ proof of the theorem is an open problem;

Stable ergodicity

i. Ergodicity: indivisibility of the attractor or state space; time average = space average; is with respect to a measure;

ii. Irrational rotations of the circle; Anosov diffeomorphisms; broken with "stacked" Anosovs;

iii. *key* — allows for center manifolds;

Conjecture/theorem

Given f is C^2, volume preserving, accessible, and center-bunched, then f is ergodic.

The point, in the non-dissipative case, we know what is needed to guarantee ergodicity — the arguments completely fail for the dissipative case.

Statistical stability

If there exists an attractor that supports a unique SRB measure everywhere (Lebesgue a.e.) on the attractor, f is statistically stable if C^k perturbations of f do not change this feature and if the SRB measures are C^k close;

No one knows how general this is, the notion is only 2 years old.

Progress?

i. we know the requirements for preserving topological structure;

ii. preservation of topological structure is not dense, but possibly it is robust;

iii. no real characterizations for dissipative dynamical systems;

iv. we know the requirements for preserving ergodicity;

v. no one knows anything about ergodicity in dissipative systems;

vi. lots of geometric intuition with regards geometric constraints on dynamics;

vii. connection to reality missing, characterizations are not useful for most natural systems;

viii. question of persistence of complex behavior — continuity between natural world and mathematical dynamics — is still unresolved;

Our approach

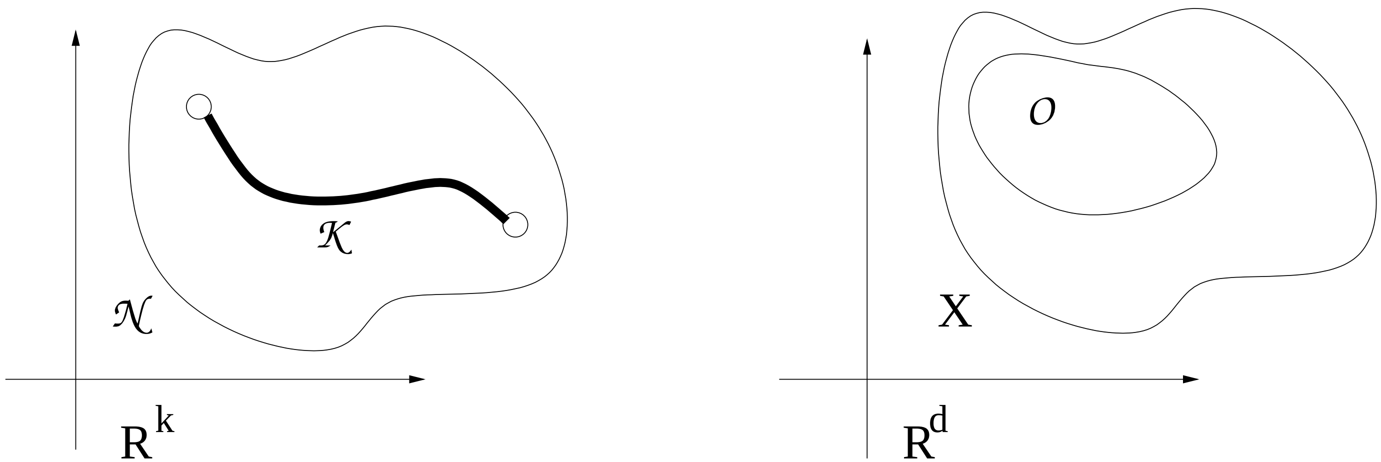
i. select a practically minded function space that can: (a) approximate C^r mappings and their derivatives (thus has infinitely many parameters); (b) have a measure associated with it (C^r function space does *not* admit a measure); (c) be fit to computational and real-world data;

ii. perform a Monte Carlo study on this function space over the parameters relative to a measure;

iii. characterize the stability of the mappings with respect to perturbations;

vi. characterize geometry and dynamics relative to the imposed measure — helps to categorize the diversity in nature;

Schematic of degree-p Persistent Chaos



p positive LCEs (i.e. the unstable manifolds) persist for:

i. an open neighborhood in parameter space $\mathcal{N} \subset R^k$ of a set (curve/interval in parameter space) $\mathcal{K} \subset R^k$;

ii. open set in state-space, $\mathcal{O} \subset X$, for which the variation across SRB measures is small;

Artificial neural networks

Definition 1 *A neural network is a C^r mapping $\gamma : R^d \to R$. The set of feedforward networks with a single hidden layer, $\Sigma(G)$, can be written:*

$$\Sigma(G) \equiv \{\gamma : R^d \to R | \gamma(x) = \sum_{i=1}^{n} \beta_i G(\tilde{x}^T \omega_i)\} \tag{1}$$

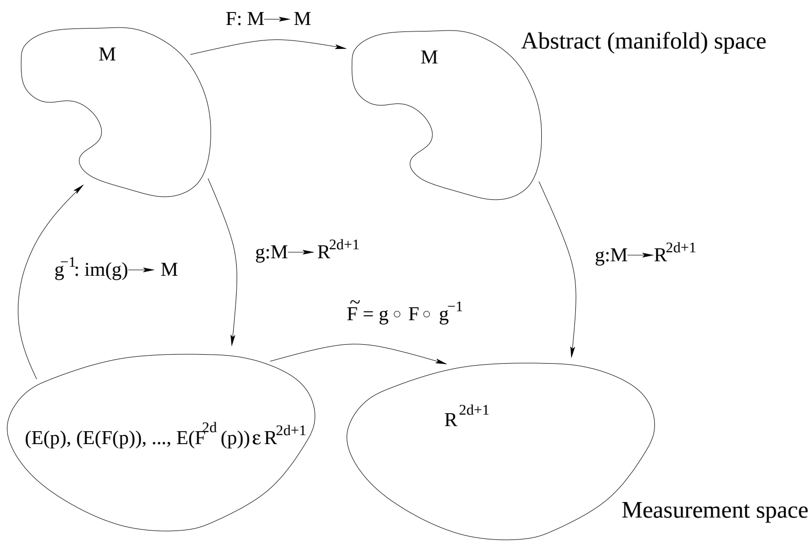
where $x \in R^d$, is the $d-$vector of networks inputs, $\tilde{x}^T \equiv (1, x^T)$ (where x^T is the transpose of x), n is the number of hidden units (neurons), $\beta_1, \ldots, \beta_n \in R$ are the hidden-to-output layer weights, $\omega_1, \ldots, \omega_n \in R^{d+1}$ are the input-to-hidden layer weights, and $G : R^d \to R$ is the hidden layer activation function (or neuron).

$$x_t = \beta_0 + \sum_{i=1}^{n} \beta_i \tanh\left(s\omega_{i0} + s \sum_{j=1}^{d} \omega_{ij} x_{t-j} \right) \tag{2}$$

$\omega_{ij} \in N(0, s)$, β_i uniform on $[0, 1]$, $d =$ number of inputs, $n =$ number of neurons. We will denote a neural network as $f_{s,\beta,\omega}$.

The probability distributions on the weights impose a measure on the parameter space R^k ($k = (n(d+2)+1)$) and thus a measure on the space of neural networks we are considering. Our results are relative to this measure.

Why time-delay neural networks?



Here M is a C^r manifold, $F \in C^r$ maps M to itself, and g is an embedding (the Takens embedding), and $\tilde{F} = g(F(g^{-1}))$; **we are studying mappings that approximate $\tilde{F}$**.

Neural networks have the following convenient properties:

a. they can approximate any C^r mapping on a compact, metrizible set;

b. they admit a measure;

c. they are used for fitting real-world data sets;

d. they are approximate the time-delay mapping Takens and Sauer et. al. proved can represent a C^r mapping (subject to certain techincal requirements) in R^d;

The experiment

1. because stable and unstable manifolds are identified with negative and positive Lyapunov exponents respectively—the geometric structure of the attractor is associated with LCEs—studying LCE variation with parameter variation helps understand the dynamics geometrically

2. for our networks the s parameter is particularly important; it controls the variance of the connectivity matrix and thus the degree of non-linearity of the map; this it is a unique and important bifurcation parameter;

3. for each network in the ensemble:

   i. fix weights

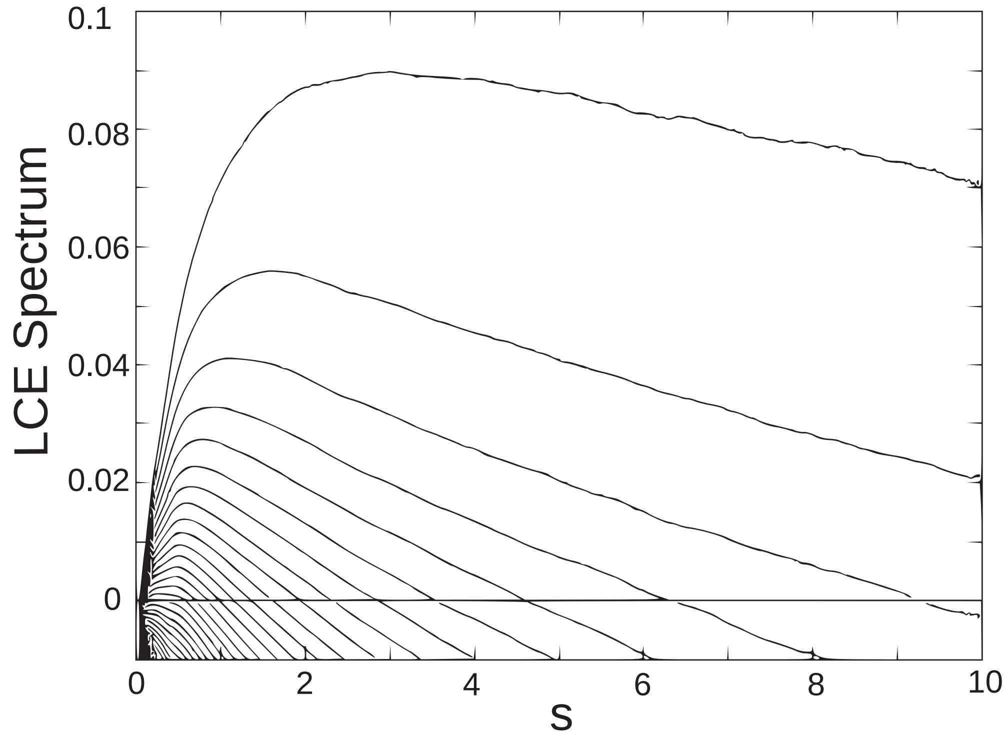
   ii. vary s (R^1 parameter curve)

   iii. fix s, vary weights R^k ball in parameter space

   iv. carve out a tube in parameter space — an R^k ball in parameter space along the R^1 s-interval.

4. study the variation of the LCEs with respect to these parameter curves.

Persistent geometric structure — one realization of persistent chaos



This is a plot of the LCE variation along an s-interval (s is functioning as the bifurcation parameter). The *main claim* in our work is that this picture is persistent to parameter and functional variation — the remainder of this poster involves precisely quantifying the persistence of this plot for the space of mappings and the respective measures on that space of mappings we are employing.

Persistent properties:

a. lack of periodic windows with respect to (s, β, ω);

b. the LCEs vary continuously with s;

c. the LCEs have a single maximum (up to statistical fluctuations, with respect to s);

d. $f_{s,\beta,\omega}$ has SRB measure(s) that yields a distribution of LCEs whose variance obeys $\sigma^2_{\chi_i} < \inf_{j=\pm 1}(|\chi_i - \chi_j|)$ at fixed s;

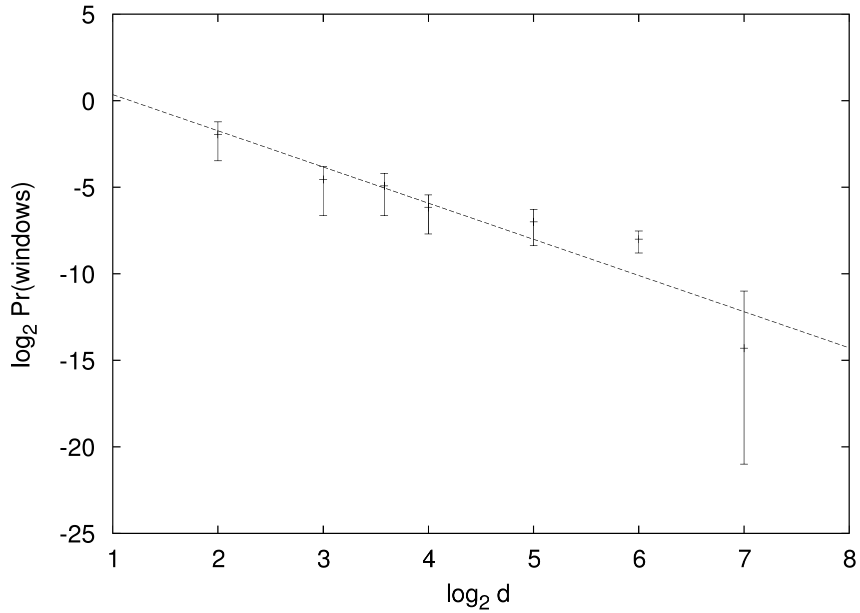
e. as d increases, the length of the s-intervals between LCE zero-crossings decreases as $\sim d^{-1.92}$;

f. the maximum number of positive LCEs increases monotonically $\sim d/4$ and the attractor's Kaplan-Yorke dimension scales as $\sim d/2$;

Persistent geometric mechanisms that lead to persistent chaos as the dimension of a dynamical system becomes large.

Conjecture 1 *Given $f_{s,\beta,\omega}$, if k and d are large enough, the probability with respect to $m_{\beta} \times m_{\omega}$ of the set (β, ω) with the properties (a)-(f) is large and approaches 1 as $k, d \to \infty$.*

Property (a)



Log probability of non-chaotic behavior versus log dimension for 700 cases per d with s fixed at 3, k-ball with radius 10^{-3}, 100 perturbations per network. The best-fit line is $\sim 1/d^2$. These findings are robust for $s \in (0.1, 8)$ and to perturbation sizes ranging from 10^{-10} to 0.1.

Properties (b)-(d)

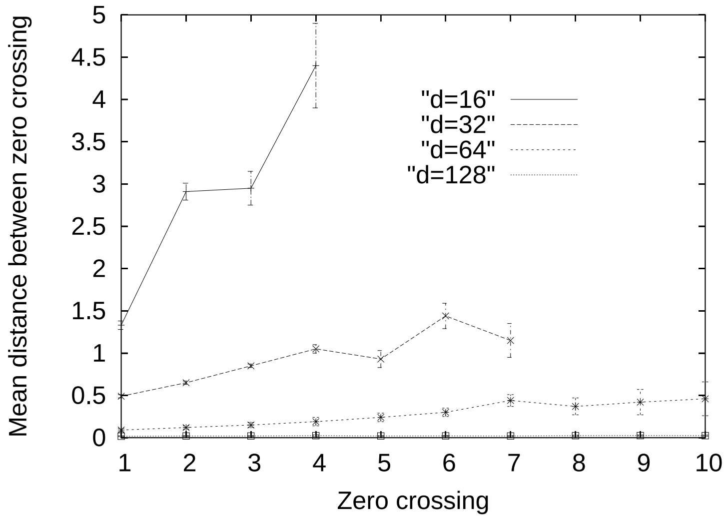
b. that the LCEs vary continuously with s is implicit upon considering the variation of the LCEs with s;

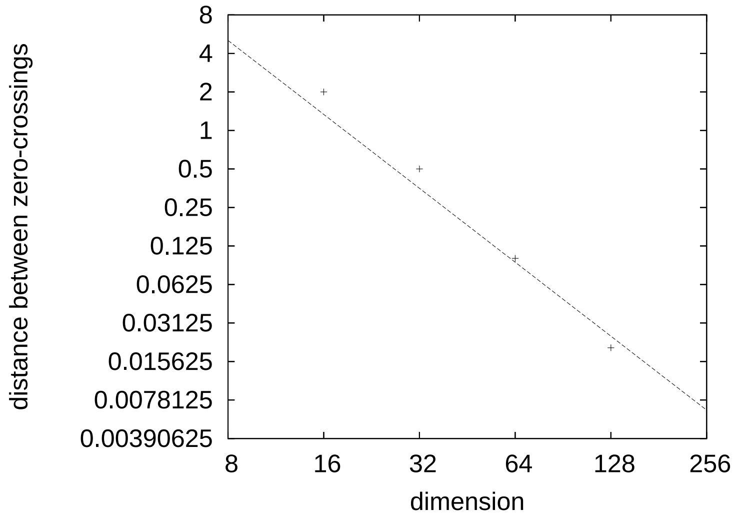
c. that the LCEs have a single maximum (up to statistical fluctuations) is implicit upon considering the variation of the LCEs with s;

d. that $f_{s,\beta,\omega}$ has SRB measure(s) that obeys property (d) can be seen by noting that for every s value in the plot of the LCE variation with s, a different initial condition with respect to m_I was used for computation.

For s values below the onset of chaos ($s < 0.1$), this condition does not apply.

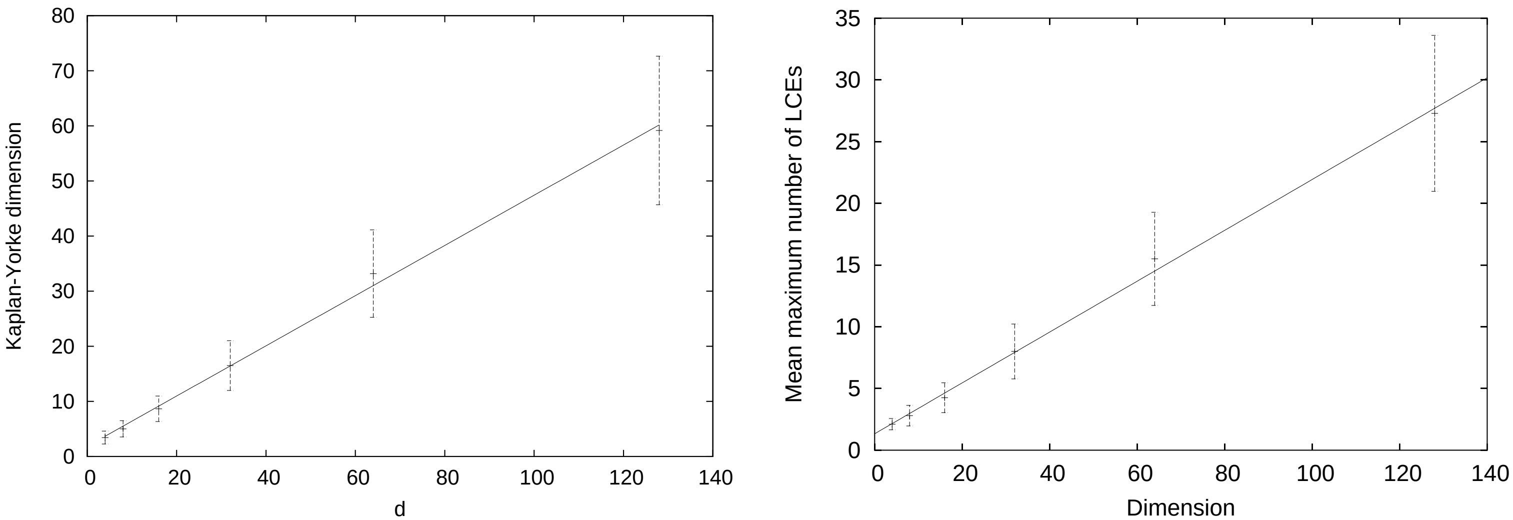
Property (e)





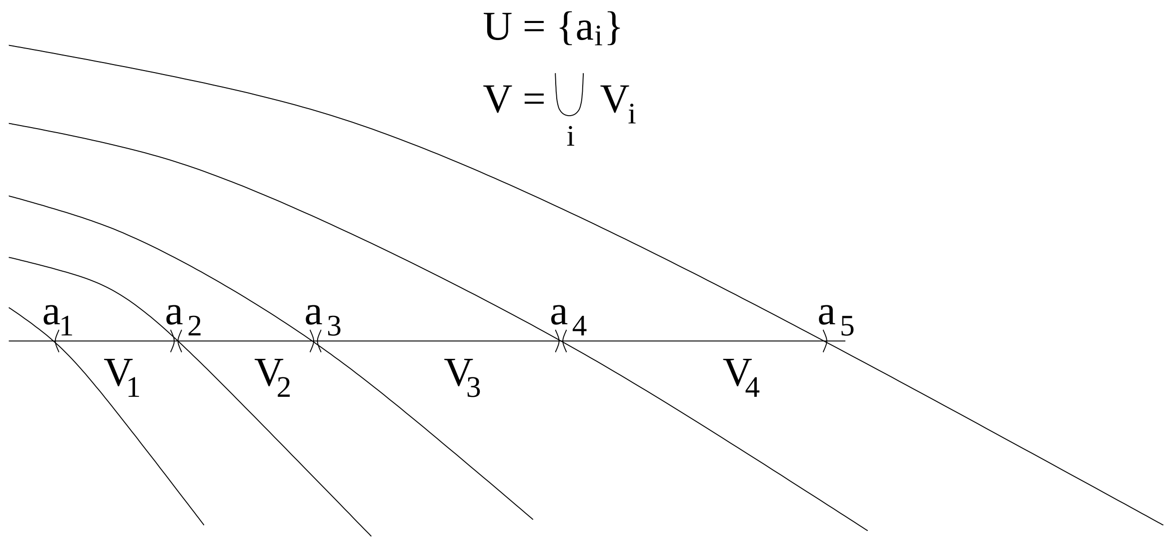
The left plot shows the mean distance between the first 10 exponent zero crossings for $s > s_{min}$ (s_{min} is the smallest s value such that the least positive LCE crosses zero and becomes positive) for networks with dimensions from $d = 16$ to $d = 128$. The right plot shows the scaling of the mean distance between exponent zero crossings with increasing dimension—it is this feature that suggests non-catastrophic topological variation with minute parameter change.

Property (f)



The left plot is of the maximum D_{KY} with increasing dimension, d. The right plot is of the maximum number of positive LCEs (MPLCE) with increasing dimension, d. Both figures document the claimed scalings MPLCE$\sim d/4$ and $D_{KY} \sim d/2$. Different scaling exist for different n and d values, however the increases in MPLCE and D_{KY} are monotonic with increases in n and d and is given approximately by MPLCE$\approx 0.11 n^{0.4} d^{0.8}$.

Intuition for geometric variation



An intuitive diagram for chain link sets, V, bifurcation link sets, V_i, and bifurcation chain sets, U. for an LCE decreasing chain link set V.

i. property (b) (LCEs vary continuously with s);

ii. property (d) (SRB measures are "boundedly" similar);

iii. property (e) distance between LCE zero-crossings decreases;

iv. property (f) D_{KY} increases monotonically with d;

Successes?

i. the geometric picture for persistence of complex (turbulent like?) dynamics;

ii. partial link between abstract dynamics and natural systems — it elevates some of the discrepancy between what is observed in natural systems and the toy models in computational dynamics;

iii. provides a language to begin to express and qualify different types of dynamics geometrically;

iv. this was but an example of what can be done abstractly, similar analysis can (and in some cases has) been done with applied systems;

v. provides hope for a positive connection between abstract geometry and it's affect on practical dynamics;

vi. provides the possibility of identifications of measures on function spaces with geometrical dynamics;

Problems, problems

i. lack of language to characterize many situations;

ii. comfort with the dependence on the "measure;"

iii. identification of categories in nature the abstract with categories of dynamics and geometry;

iv. some qualitative success, quantitative success is yet lacking in most practical examples (one partial success of this approach - Buz and Cars);