Patterns of emergence

PhD Seminar by Nicolas Brodu,

Concordia University PhD student,

Under the supervision of Peter Grogono,

March 14th, 2005

2007 Warning: This presentation is outdated and lacks references.
Kept on the internet for the records, but use at your own risks!

Outline

Different approaches

Bottom-up approach

- Work on local interactions (ex: Newton laws)
- Integrate for global scale effects
- Problem: Easier to say than to do!

Top-down approach

- Define functional parts & recurse
- Assemble them like a well-crafted clock
- Problem: The inevitable grain of sand

So, what to do?

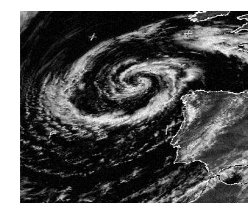
So, what to do?

"Everything is information" approach?

- Leibniz, Shannon, Church-Turing, Zuse [1]...
- But we can only crunch so many numbers!
- And not more anyway (Godel, Chaitin [2])

A solution?

- Study persistent patterns (ex: whirlpool)
- And their relations (how they emerge)
- Scale & nature are irrelevant



Welcome to self-organization theory!

[1] Konrad Zuse, 1969 [2] Gregory Chaitin, 2005

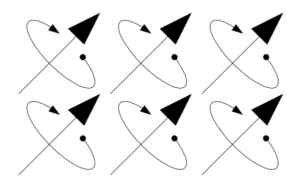
Self-Organization

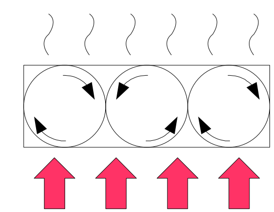
Study of “natural” modes of a system

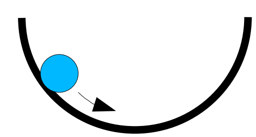
- If left alone, what can happen (static modes)
- Relation with environment (dynamic modes)

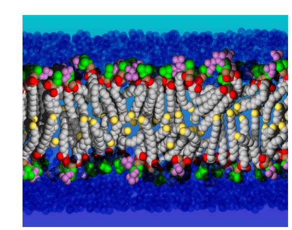
A very general concept [3], a few examples:

- Magnetization
- Rayleigh-Bénard rolls
- A ball thrown in a bowl
- Lipid bilayers [4]









But is this interesting or trivial?

[3] Hermann Haken, 1977 [4] Mueller & Robin, 1962

Open dissipative systems

Open System

- Turn over of substrate (cells, atoms…)
- External flow of energy (heat, sun-light…)

Dissipation

- Allows for far from equilibrium states
- Differential persistance [5] (ex: Bénard Rolls)

Reconciliate order & thermodynamics

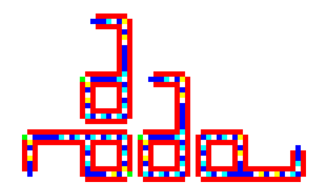
- Local entropy reduction, global dissipation
- Nobel Prize Ilya Prigogine [6]

[5] John Holland, 1998 [6] Ilya Prigogine, 1977

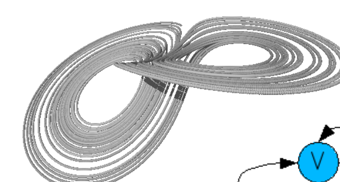
A framework

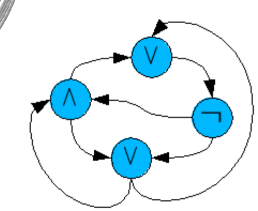
Dynamical systems

- The best tool we have… for lack of a better!
- Continuous or discrete
  - In time: differential equations, iterated functions
  - In space: continuous parameters or set of values
- Examples:
  - Iterated function systems
  - Cellular automata
  - Differential equations
  - Graphs, neural & boolean networks.



$$\begin{cases} \frac{dx}{dt} = 10(y-x) \\ \frac{dy}{dt} = 28x - y - xz \\ \frac{dz}{dt} = xy - \frac{8}{3}z \end{cases}$$



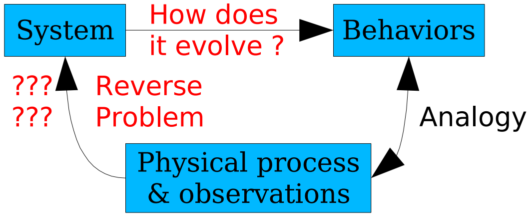


Extensions: Probabilities, noise, forcing…

Order and Chaos

Beware of meaningless generalizations!

The main questions are:



- How does the system evolve?
- How to solve the reverse problem?

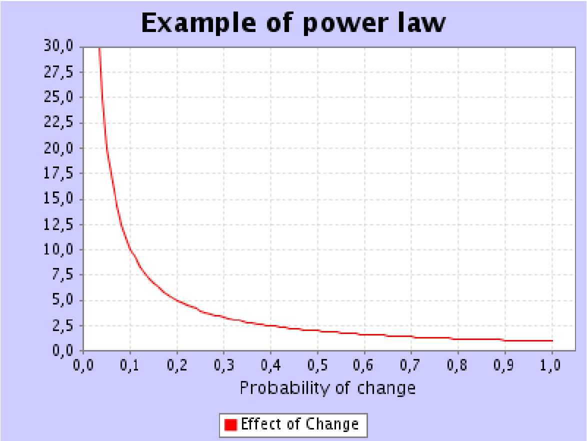
Self-organization of a dynamical system:

- It wanders without consistency.
  - Ex: $x_{n+1} = 4x_n(1-x_n)$ for x_0 in [0..1].
- It stays forever in some states: an attractor
  - Ex: Magnet, Ball in the bowl, Lorenz attractor
- Somewhere in between: Edge of chaos!

Edge of Chaos

Where “interesting” phenomena occur

- Global and local effects for perturbations.
- Some structures persist in apparent chaos.
- Neither to simple, nor completely “random”.
- Cellular automata can compute [7].
- Main properties for self-organization:
  - Some changes are allowed, unlike total order
  - The changes may persist, unlike total chaos
  - Usually there is a “Power law”

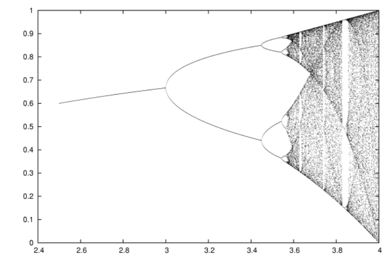


[7] Chris Langton, 1990

How to get there?

Path to chaos

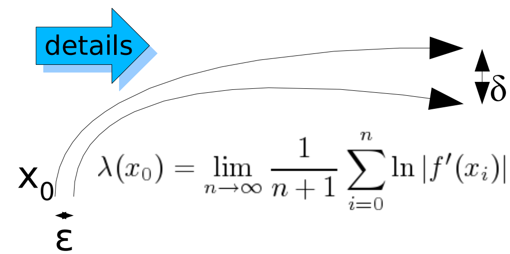
- Bifurcations & period doubling



- Phase transitions (ex: C. Langton λ parameter [7])

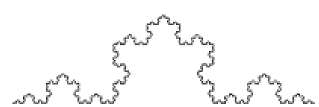
Indicators:

- Lyapunov exponents



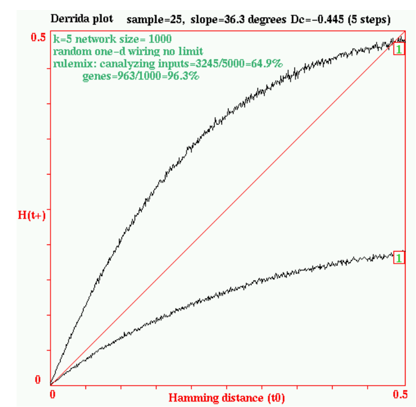
$$\lambda(x_0) = \lim_{n\to\infty} \frac{1}{n+1} \sum_{i=0}^{n} \ln |f'(x_i)|$$

- Derrida plots
- Fractal dimension [18]



D = ln 4 / ln 3

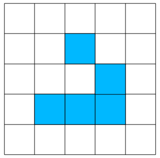
details



Derrida Plot [11]

Others? Still active research

[18] Michael F. Barnsley, 1993 [11] Andrew Wuensche, 2002



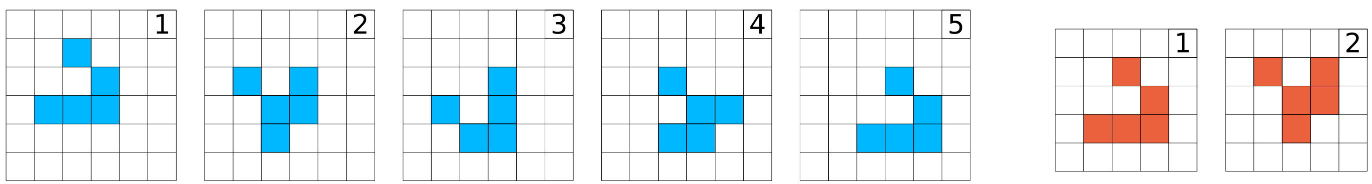
An example: the glider

Let's note X the grid space, $x \in X$ a possible grid state.

Let note $f : X \to X$ the game of life rules.

Then let $T = \{t : X \to X \ / \ t \circ f = f \circ t,\ t \text{ invertible},\ t \text{ preserve grid size}\}$.

T defines an equivalence relation (proof available on request). Let's note $C(x)$ the class of x: $C(x) = \{y \in X \ / \ \exists t \in T, y = t(x)\}$.



Glider sequence in grid space   Glider sequence in equivalence class space

Note that $\forall x, y = t(x)$, we have $f(y) = f \circ t(x) = t \circ f(x)$ by definition of T. Thus $f(x)$ and $f(y)$ are in the same class. By extension, we can now define $f(C)$: f acting on a whole class. Now consider S the space of all classes. Then $f : S \to S$ has a cycle of length 2 corresponding to the glider.

Thus the glider is a persistant feature (cycle) in S-space, and time.

Where does the translation we see in grid space comes from? From the fact we give a particular significance to the initial grid x_0. Since $f^2(C) = C$ for $C = C(x_0)$ class of x_0, there is a $t \in T$ such that $f^2(x_0) = t(x_0)$. Actually it is a composition of a rotation and a translation, and we find the same glider in another orientation at step 3. t^2 is the translation we see from step 1 to 5.

Discussion

Locality in space & time

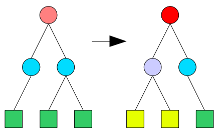
- Example of colliding gliders
- How to refine “sufficiently persistent”
- Locality, both in space and time

Stability on perturbation:

- How to quantify stability?
  - Lyapunov exponent, derrida plots, etc...
  - Statistics about perturbation effects?
- Relation to autopoiesis [17]
  - Consider glider as autonomous entity

map

[17] Randall D. Beer, 2004



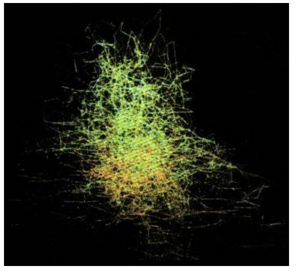
Structural stability

Structural stability

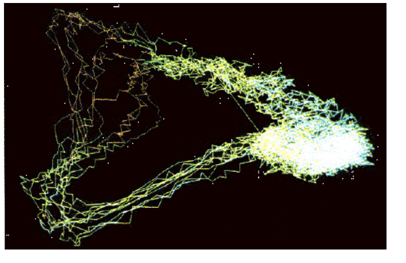
- Self-organized order & explicitly built order
- Think about biological & mechanical systems

Sensitivity to perturbations

- Environment, noise, natural instability…
- Dynamic mode change
- Structural change
  - Hard: rupture point
  - Soft: dynamical mode gone permanent



Normal rat brain pattern [8]: chaotic dynamical mode



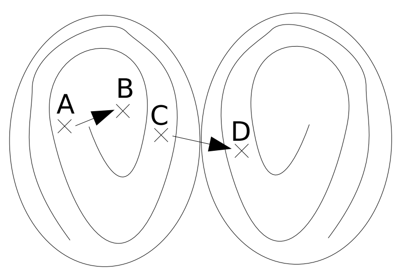
Epileptic rat brain pattern [8]: Attractor mode

[8] Walter J. Freeman, 1998

System behavior

With perturbations / environment

- The system organizes into patterns
- It jumps from one mode to another



A perturbation from A to B does not change the dynamic mode of the system

A perturbation from C to D cause an attractor change: little cause, great effects!

Cognitive domain

- Regions the system may explore
- Limited by structural modifications (rupture)

But how to use this in practice?

Patterns in practice

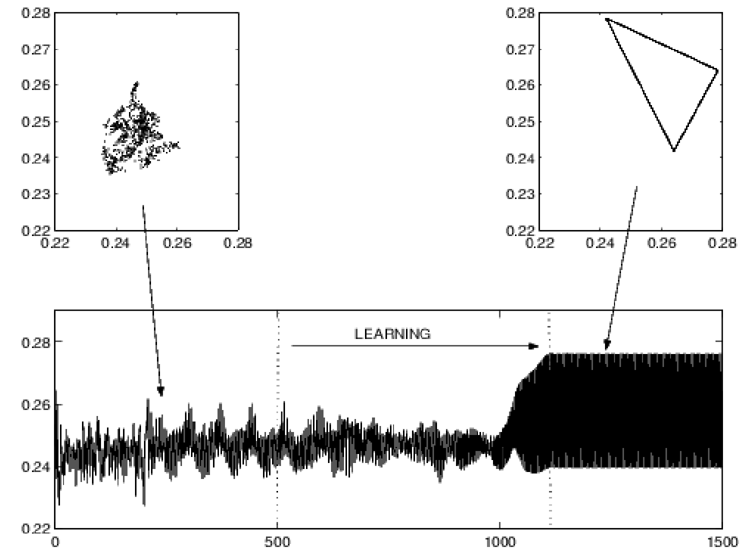
forward

3 levels to consider:

- Physical implementation, substrate.
- Its organization, including attractors.
- Associating internal states to features.

Example with recurrent neural networks

- Measure the network behavior:
  - Use statistics [9]
  - Symbolic dynamics [10]
- Adapt it to produce desired attractors
- Associate attractors to concepts
- Extension: learn mapping, not concepts.
- Run-time: detect cycles, no convergence



Learning process from [9]

[9] Daucé & Quoy, 2000 [10] Molter, Salihoglu, Bersini, 2004

Learning & Evolution

Structure defines possible modes

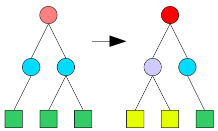
- Some are “natural”, self-organized ⇒ innate
- Other are reachable through learning only.
- There is only so much a structure can learn

Learning as mode exploration.

- Previous example [10]:
  - Reliably stored up to 50 patterns with 3 “neurons”
  - Use an input to network translation layer to overcome structure limitations.
- Coupling with environment

Evolution*: changing the structure itself.

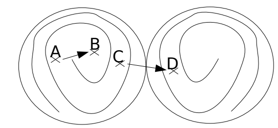
* This is not Darwinian evolution, yet!

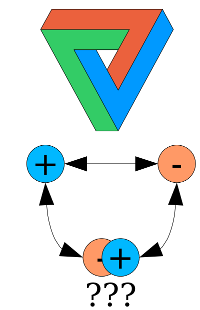


Hierarchies

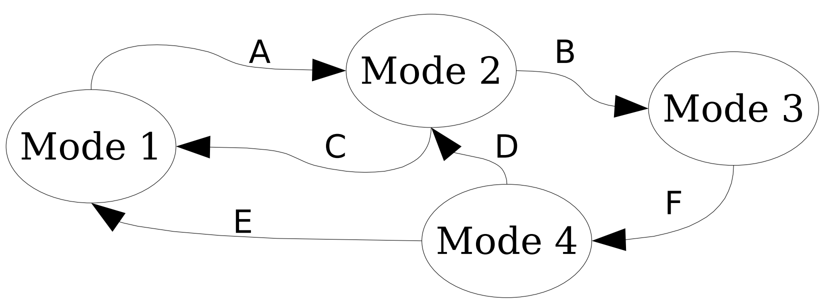
Suppose "modes" have transition rules

- Attractor shift (noise, etc).
- Transient "stability" (frustrated chaos, etc.)





Then we can consider a "higher level"



A to F: Transition rules
Modes: Stable regions

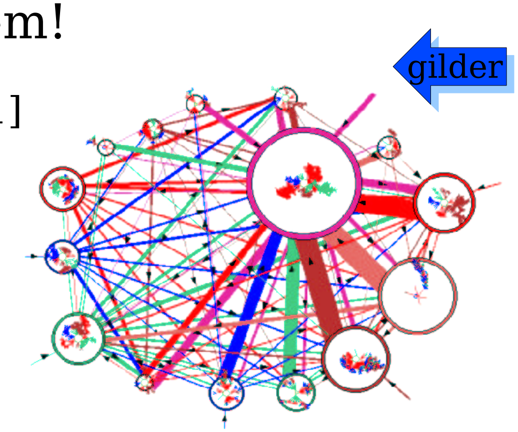
This defines an oriented graph (network), which is itself a dynamical system!

Ex: Random boolean networks [11]

Attractors are perturbed

Probability map for each attractor shift

Larger basins and links are scale accordingly



[11] Andrew Wuensche, 2002

Warning: outdated & incomplete

Emergence

A tentative definition

- Higher-level features or relations
- Sufficiently persistent in space & time
- Achieved by attractors, but not only

Counter-examples: temperature, color...

- Pro: Global effects not present at lower-scale
- Con: Is it an artifact from the observer?

Formal framework often not applicable, and incomplete.

No consensus! What are common criteria?

Common criteria

Downward causation

- The higher levels constrain the lower ones
- Ex: Bénard rolls, brain deciding motion, etc.

Whole is more than sum of parts

- Intuitive, but may be trivial [5]

Creative or combinatorial [13]?

- Creative: higher level not describable with lower levels concepts.
- Combinatorial: global effects are distributed

Computational, or physical [15]? (or both [1])

[5] Holland, 1997 [13] Kubík, 2003 [15] Cariani, 1989 [1] Zuse, 1969

Definitions review

back

Syntactical vs Semantic Howard Pattee, 1989	• Uses the 3 levels of consideration • Syntactical is level 1 to 2 • Semantic is level 3 when it is not reducible to formal descriptions in 1 & 2
Emergence relative to a Model Peter Cariani, 1989	• Uses the 3 levels of consideration • Level 3 = model of the world = internal representation, is necessarily incomplete • Emergence when physical observation differs
Basic emergence Aleš Kubík, 2003	• Reducible to lower level interactions • Passive environment • Explicit definition for “sum of the parts”
Weak emergence M. Bedeau, 1997	• Emergence as phenomenon that can only be described by the full length of a simulation. • Equivalent to the notion of algorithmic incompressibility, and randomness [2]

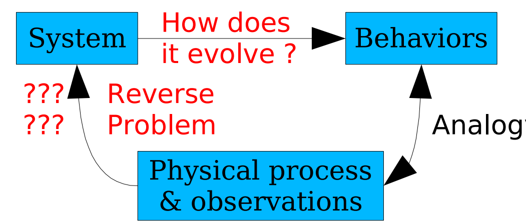
Note: Many authors use the word emergence, not that many would risk to give a definition. Hence the cautious terms above. Some others, like John Holland [5], prefer to describe a framework and give a set of properties emergence should have in it.

[2] Gregory Chaitin, 2005

And now, what?

Challenges

- Framework independence, genericity
- Danger: too generic is inapplicable!
- The reverse problem



Current work, state of art

- Formalization of emergence
- Mathematical developments
- More frameworks
- Numerical experiments becoming tractable

References

[1] Konrad Zuse, Rechnender Raum, Friedrich Vieweg & Sohn, Braunschweig, 1969. English translation: Calculating Space, MIT Technical Translation AZT-70-164-GEMIT, MIT (Proj. MAC), Cambridge, Mass. 02139, Feb. 1970.

[2] "Meta math! The quest for Omega", Gregory Chaitin, to be published in sept 2005.

[3] "Synergetics", H. Haken, Physics Bulletin, London, Sept. 1977

[4] "Reconstitution of a cell membrane structure in vitro and its transformation into an excitable system". Nature 194:979-980. P. Mueller, D.O. Rudin, et al. 1962.

[5] "Emergence: From chaos to order". John Holland, 1998

[6] "Ilya Prigogine" homage presentation by Professor Minati in the Plenary in Memoriam of Past Presidents of ISSS, The Forty-seventh Meeting of the International Society for the System Sciences July 7th - 11th, 2003. Available at http://www.isss.org/lumprig.htm

[7] "Computation at the edge of chaos: phase transitions and emergent computation", Chris Langton, Physica D, 42, 12, 1990. 42

[8] "Strange Attractors that Govern Mammalian Brain Dynamics Shown by Trajectories of Electroencephalographic (EEG) Potential", Walter J. Freeman, IEEE transactions on circuits and systems, Vol. 35, No. 7, July, 1988

[9] "Random recurrent neural networks for autonomous system design", E. Daucé, M. Quoy, 2000

[10] "How chaos boosts the encoding capacity of small recurrent neural networks: learning consideration". Colin Molter, Utku Salihoglu and Hugues Bersini, IJCNN 2004.

[11] "Basin of attraction in network dynamics", A. Wuensche, in: "Modularity in Development and Evolution", 2002

[12] L.M. Rocha, "Exploring Uncertainty, Context, and Embodiment in Cognitive and Biological Systems" PhD Dissertation in Systems Science. State University of New York at Binghamton, 1997

[13] "Toward a formalization of emergence", Aleš Kubík, in Artificial Life 9:41-65, 2003

[14] "Weak emergence", Bedeau, 1997. In Philosophical perspectives: Mind, causation and world, vol. 11.

[15] "On the design of devices with emergent semantic functions", Cariani, 1989, PhD dissertation.

[16] "Simulations, realizations, and theories of life", Howard Pattee, 1989, in Artificial Life, C. Langton ed., SFI series in the sciences of complexity, Addison-Wesley.

[17] "Autopoiesis and cognition in the game of life", Randall D. Beer, Artificial Life 10: 309-326, 2004

[18] "Fractals Everywhere", second edition, Michael F. Barnsley, ISBN 0-12-079061-0, 1993

Note: The photos p3, the lipid bilayer model p4, the bifurcation map p9, and the Von Koch curve p9 and in appendix, are in the public domain. The brain patterns p12 are under Creative Commons license Attribution, Share-Alike, v2.0, US. The learning process graph p14, the Derrida plot p9, and the basin of attraction schema p16, are citations from their respective articles in reference, under fair use. The "emergence" and "perspectives" logos were made by Valérie Dagrain for this document. All remaining logos, images, and texts are my own creation.

Lyapunov exponents

Suppose we want to know how "stable" a point is on a trajectory. If you perturb it a little by $\varepsilon > 0$, does this error get amplified?

Formally, if $x_{n+1} = f(x_n)$, then what is the value of $\delta = |f(x_0 + \varepsilon) - f(x_0)|$ with respect to ε ?

If f is differentiable at x_0, then we have $\lim_{\varepsilon \to 0} \frac{\delta}{\varepsilon} = |f'(x_0)|$. So by comparing $|f'(x_0)|$ to 1, we can decide if d gets larger or smaller than ε.

Equivalently, we can also compare $\ln|f'(x_i)|$ with 0. This has 3 advantages:

- Mathematicians like linear things, and $f'(x) = \alpha f(x)$ for exponentials. Taking $\ln|f'(x_i)|$ allows to consider $\frac{\delta}{\varepsilon}$ as an exponent, hence the name: Lyapunov exponents.
- Linear algebra can be applied. This leads to nice interpretations of Lyapunov exponents in terms of eigenvalues of a Jacobian matrix.
- On the practical side, a sum is easier to compute than a product. This point is detailed in the next part.

Lyapunov exponents

What we want to know is whether the trajectories will diverge, not just for x_1 but for future points x_n, for $n \geq 1$. But by chain rule $\frac{d}{dx}f(x_1) = \frac{d}{dx}f(f(x_0)) = f(x_0) * f'(f(x_0)) = f'(x_0) * f'(x_1)$. Recursively, this means we have to compute $\prod_{i=0}^{n} |f'(x_i)|$ and checking whether this converges to 0 or diverge to $+\infty$, when each derivative is repectively < 1 or > 1.

As said previously, it's much easier to take the log of this product and convert it to a sum: $s = \sum_{i=0}^{n} \ln |f'(x_i)|$. We can also normalize s by $\frac{1}{n+1}$, to get a result comparable with the value at the single point x_0, independently of the $n+1$ number of points considered along the trajectory. Then we can compare $\frac{s}{n+1}$ with $\ln 1 = 0$ for our stability condition.

The Lyapunov exponent at x_0 is just the limit of this normalized sum, when it exists:

$$\lambda(x_0) = \lim_{n \to \infty} \frac{1}{n+1} \sum_{i=0}^{n} \ln |f'(x_i)|$$

For a multidimentional vector $X = \{x_0 \dots x_n\}$, we get one λ_i per dimension i. Since the system gets unstable as soon as it is unstable along any one dimension, we usually pay attention only to the largest λ_i for the stability condition.

Fractal Dimension

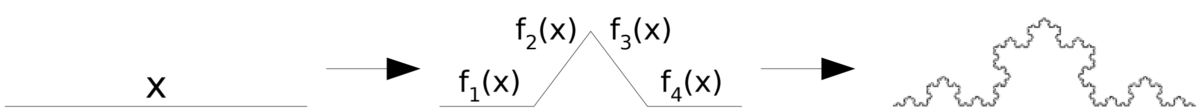
Fractal dimension: $D = \lim_{\varepsilon \to 0} \left\{ \frac{\ln \mathcal{N}(\mathcal{A}, \varepsilon)}{\ln(\frac{1}{\varepsilon})} \right\}$ with $\mathcal{N}(\mathcal{A}, \varepsilon)$ the smaller number of balls of radius ε necessary to cover the set $\mathcal{A}$.

Box counting theorem: $D = \lim_{n \to \infty} \left\{ \frac{\ln \mathcal{N}_n(\mathcal{A})}{\ln(2^n)} \right\}$ with $\mathcal{N}_n(\mathcal{A})$ the number of boxes of side length 2^{-n} intersecting the set $\mathcal{A}$.

Iterated function systems: $\sum_{i=1}^{n} |s_i|^D = 1$ with $s_1 \dots s_n$ the scaling factors of the n mappings of the IFS.

Definitions from “Fractals Everywhere”

Example on Von Koch curve:



$$\begin{cases} f_1(x) = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} x \\ f_2(x) = \begin{bmatrix} \frac{1}{3}\cos\frac{\pi}{3} & -\frac{1}{3}\sin\frac{\pi}{3} \\ \frac{1}{3}\sin\frac{\pi}{3} & \frac{1}{3}\cos\frac{\pi}{3} \end{bmatrix} x + \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix} \\ f_3(x) = \begin{bmatrix} \frac{1}{3}\cos\frac{\pi}{3} & \frac{1}{3}\sin\frac{\pi}{3} \\ -\frac{1}{3}\sin\frac{\pi}{3} & \frac{1}{3}\cos\frac{\pi}{3} \end{bmatrix} x + \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3}\sin\frac{\pi}{3} \end{bmatrix} \\ f_4(x) = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} x + \begin{bmatrix} \frac{2}{3} \\ 0 \end{bmatrix} \end{cases}$$

Each transformation f_i has a scaling factor of 1/3.

Therefore $4*(1/3)^D=1$ and $D = \ln 4 / \ln 3$.