Chaos, Complexity, and Inference (36-462)

Lecture 3

Cosma Shalizi

22 January 2008

Long-run behavior and inference

In most systems, long-run behavior is dominated by **attractors**

Aspects:

Geometry	Where do trajectories go?
Probability	What is the long-run distribution?

Attractors: Geometry

Generalize from stable fixed points and limit cycles

Attractor $\approx$ stable invariant set which cannot be split into smaller stable invariant sets

"Attractor" because nearby points move closer towards it

Requires that the map compress state space (on balance; can expand in some directions)

so Arnold cat map has no attractors

Examples: The points in the bifurcation diagram are $\approx$ the attractors for different values of r

(Why only $\approx$?)

More examples: Henon map, Lorenz system

Henon Map

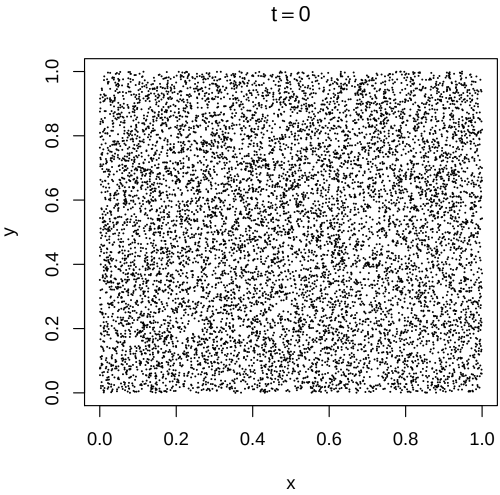
$$
\begin{aligned}
x_{t+1} &= a - x_t^2 + b y_t \\
y_{t+1} &= x_t
\end{aligned}
$$

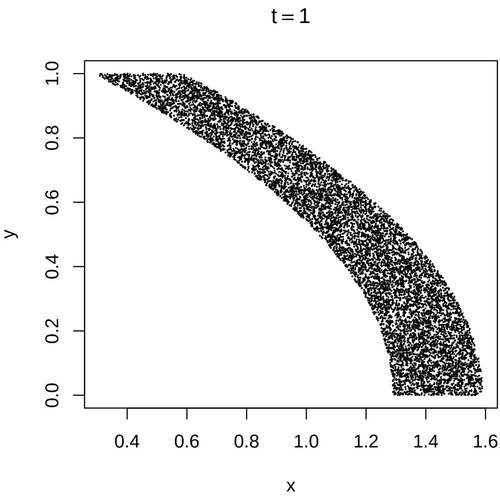
Two dimensions — *y* acts like a memory for *x*

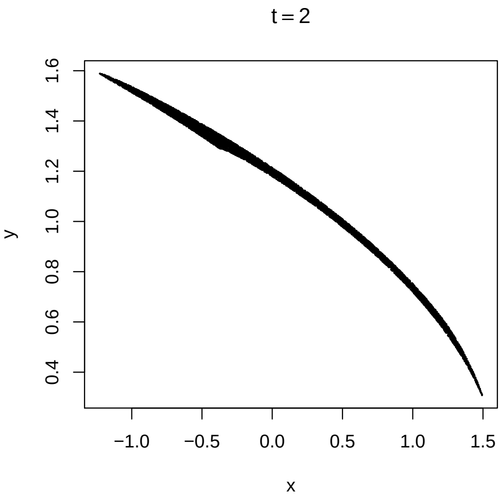
Important general fact about multidimensional dynamics

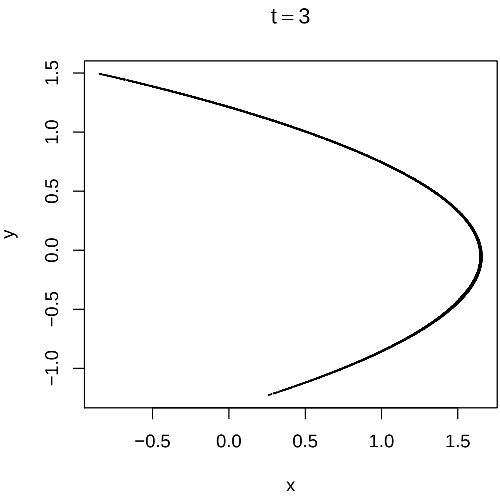
n dimension is equivalent to having a memory going back *n* time-steps

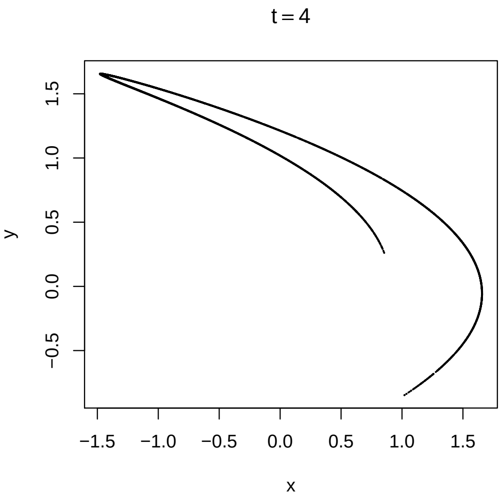
What happens with the Henon map? $a = 1.29$, $b = 0.3$

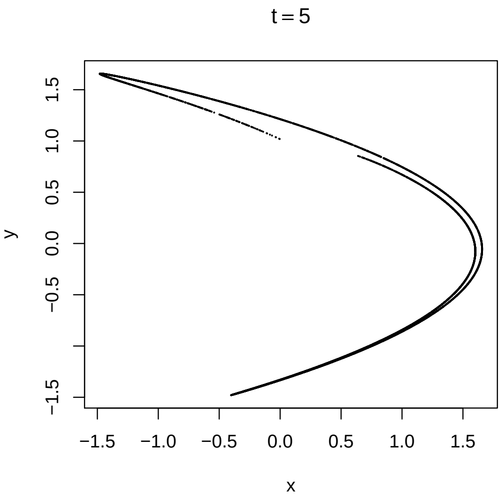


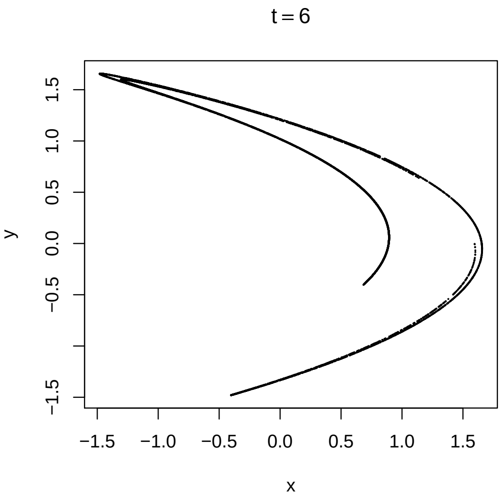

t = 1
y
1.0
0.8
0.6
0.4
0.2
0.0
0.4
0.6
0.8
1.0
1.2
1.4
1.6
x

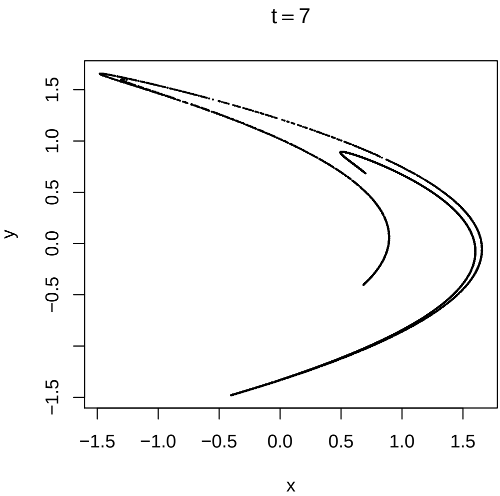

t = 2
1.6
1.4
1.2
1.0
0.8
0.6
0.4
y
−1.0
−0.5
0.0
0.5
1.0
1.5
x

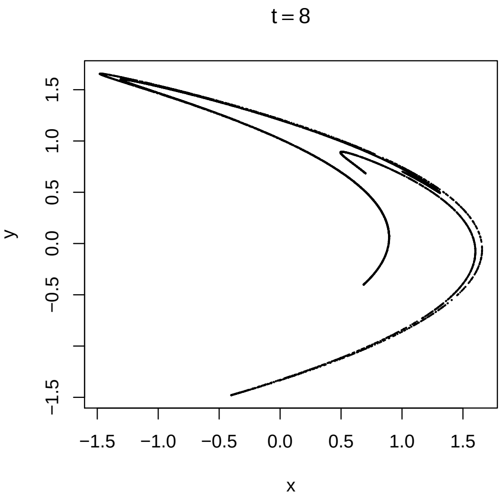

t = 3
1.5
1.0
0.5
0.0
−0.5
−1.0
y
−0.5
0.0
0.5
1.0
1.5
x

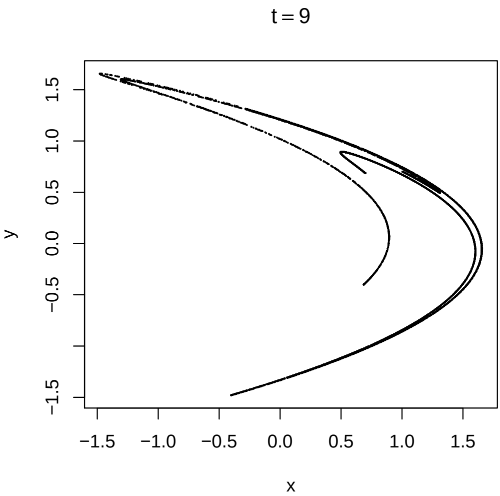

t = 4
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

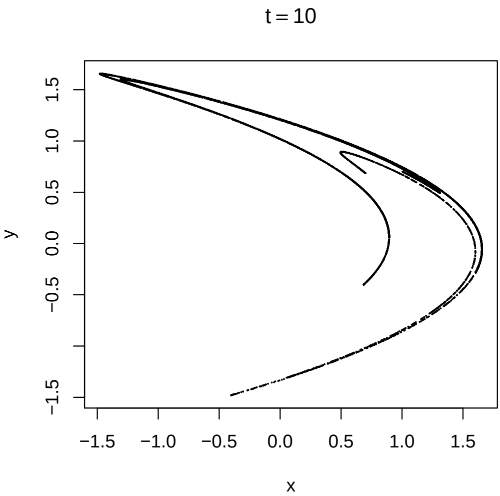

t = 5
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

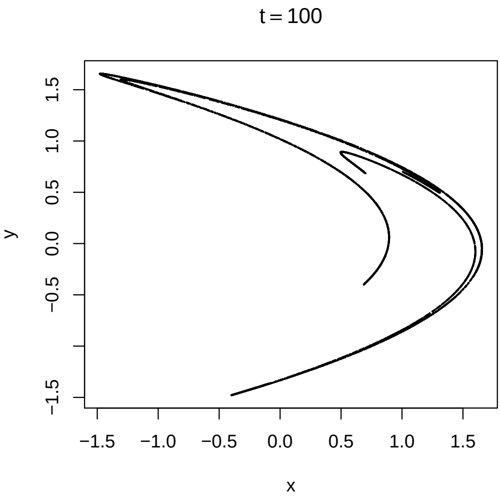

t = 6
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x


t = 7
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

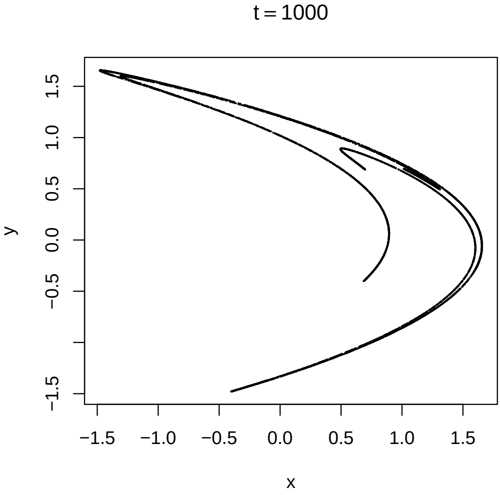

t = 8
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

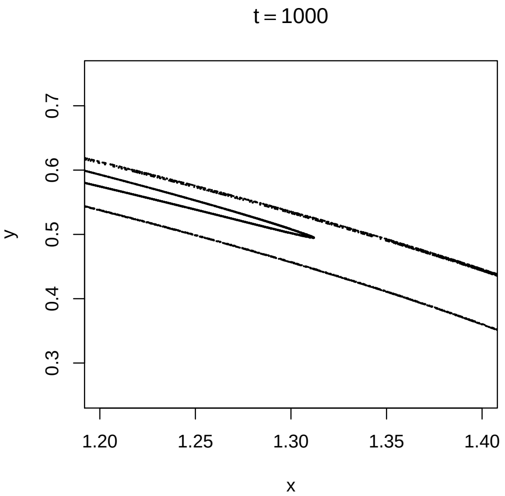

t = 9
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

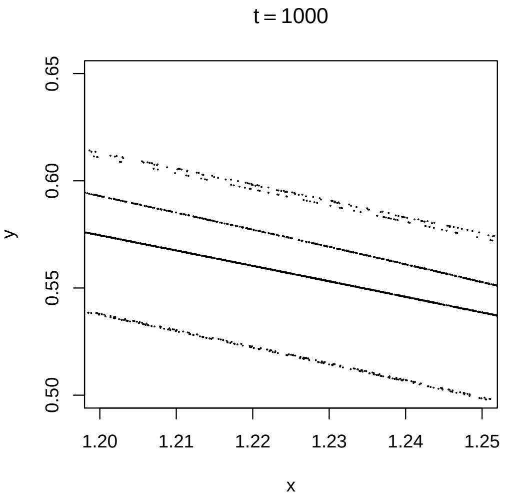

t = 10
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5


t = 100
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

take the attractor and zoom in




t = 1000
y
0.3
0.4
0.5
0.6
0.7
x
1.20
1.25
1.30
1.35
1.40


t = 1000
0.65
0.60
0.55
0.50
y
1.20
1.21
1.22
1.23
1.24
1.25
x

contraction $\Rightarrow$ typically, attractors occupy a vanishingly small fraction of the state space

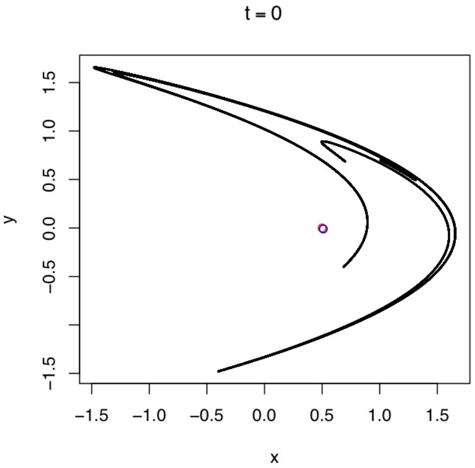
sometimes a well-behaved geometric object (points, curves)

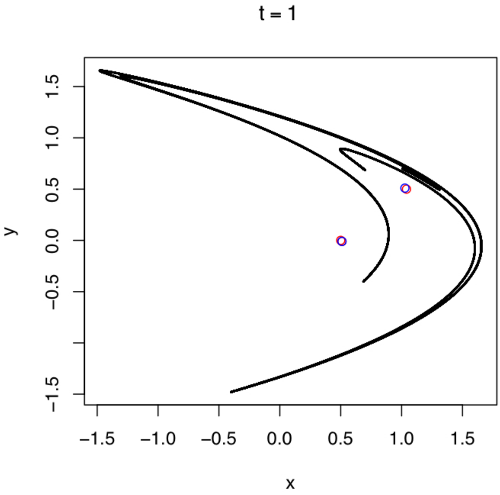
this attractor is “strange”: fractal (=fractional-dimensional), self-similar

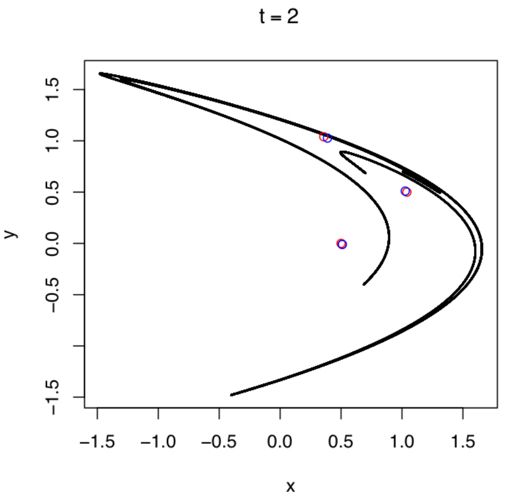
chaotic attractors are typically strange stretch and fold

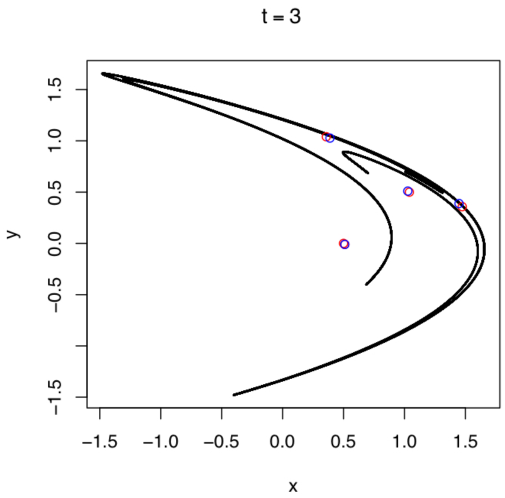
directions away from the attractor: stable; move back towards attractor

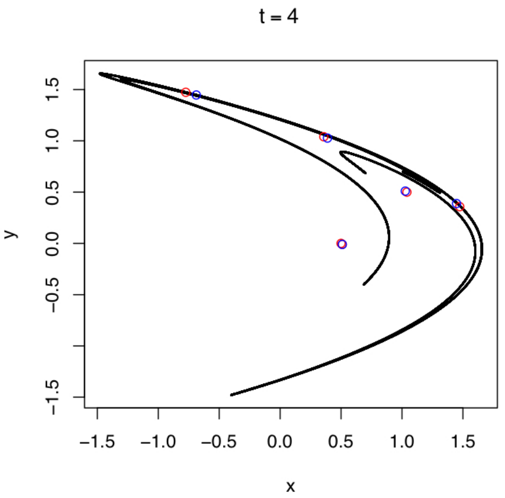
directions along the attractor: (possibly) unstable red: $(x_0, y_0) = (0.5, 0.0)$; blue: $(x_0, y_0) = (0.51, -0.01)$

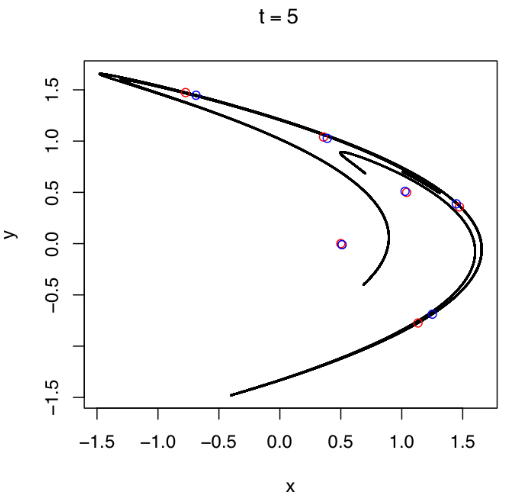


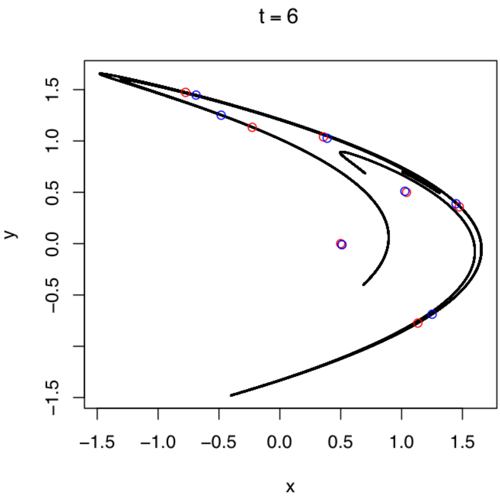

t = 1
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

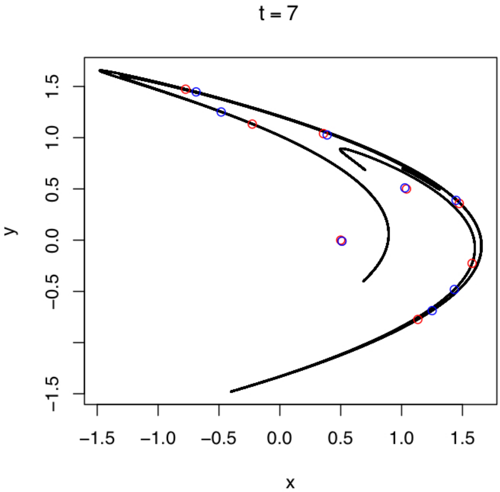

t = 2
1.5
1.0
0.5
0.0
−0.5
−1.5
y
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

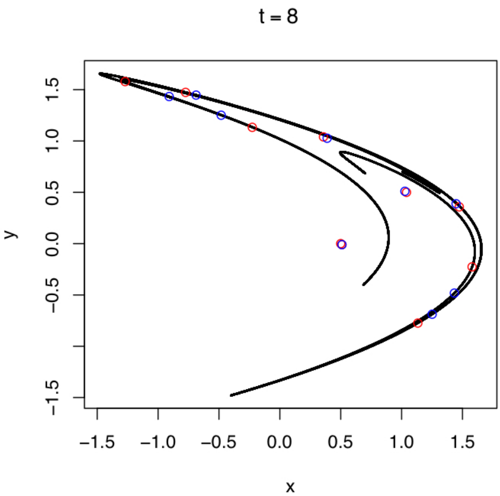

t = 3
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

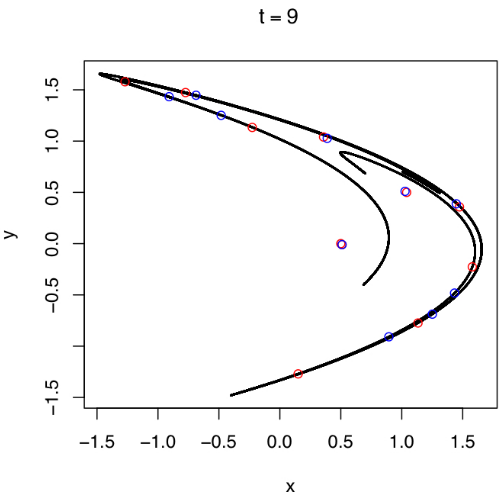

t = 4
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

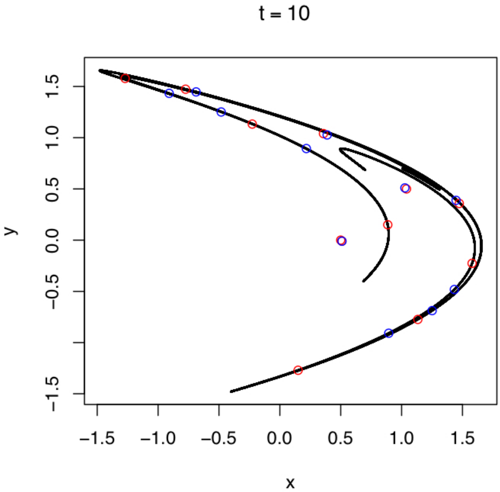

t = 5
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

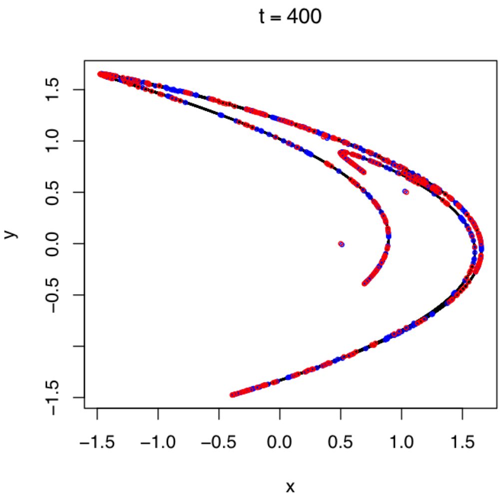

t = 6
y
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5


t = 7
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5


t = 8
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5


t = 9
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x


t = 10
y
1.5
1.0
0.5
0.0
−0.5
−1.5
x
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

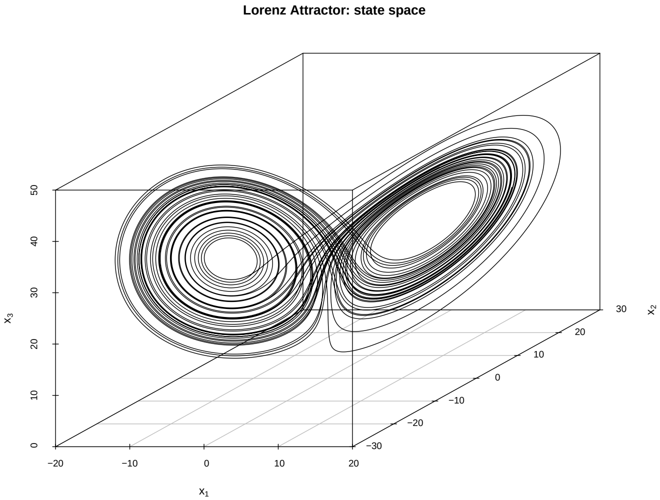

t = 400
y
1.5
1.0
0.5
0.0
−0.5
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
x

LORENZ SYSTEM

$$\frac{dx}{dt} = ay - ax$$
$$\frac{dy}{dt} = bx - y - xz$$
$$\frac{dz}{dt} = xy - cz$$

(more usual notation $a = \sigma$, $b = \rho$, $c = \beta$)
Crude approximation to tricky nonlinear model of fluid flow
Sensitive dependence on initial conditions discovered due to truncation

use $a = 10$, $b = 28$, $c = 8/3$



Basin of attraction: all the initial conditions which converge on an attractor

We have only seen systems with one attractor but there can be many

another kind of unpredictability: points very near the boundary between 2 basins of attraction

Lyapunov Exponents

p dimensional state-space
Stable and unstable directions locally at any point
Also rates of exponential contraction/expansion along them
so starting at x an initial separation of δ along some direction becomes a separation of $\delta e^{\Lambda(x)t}$ — not necessarily along that direction
There is always a most unstable direction $e_1(x)$, rate $\Lambda_1(x)$
Then a next most unstable direction $\perp e_1(x)$, rate $\Lambda_2(x)$
... finally a most stable direction $e_p(x)$, rate $\Lambda_p(x)$
Pick a trajectory; $\lambda_i \equiv$ time-average of $\Lambda_i(x)$
(How do we know that's well-defined?)
λ_i are the **Lyapunov exponents**
Practically, "chaos" means an attractor, and $\lambda_1 > 0$

Attractors: Probability

probability of states (invariant distributions)
probability of trajectory segments (correlations)
linked to geometry via instabilities in the attractor

Invariant distributions

Invariant distributions are (generally) confined to invariant sets — stable or unstable

Natural or **physical** invariant distributions are confined to attractors

Periodic attractors have uniform invariant distributions

Strange attractors generally have non-uniform invariant distributions

If multiple attractors then multiple physical invariant distributions

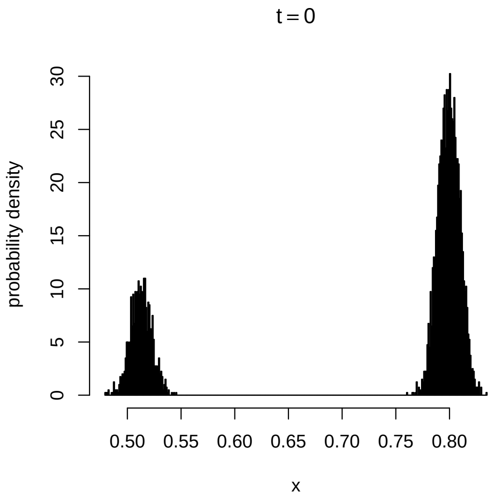
Volumes in state-space keep shrinking $\Rightarrow$ attractor is actually infinitely small compared to whole state space
(points vs. line, lines vs. plane, etc., or weird fractal shapes)
Invariant distributions are generally *not smooth at all*
Usually no simple parametric form ($r = 1$ logistic map is an exception)
Kernel density estimation can work but too much smoothing is very misleading!
(ditto Gaussian kernels)
For a known 1D map, see [1]

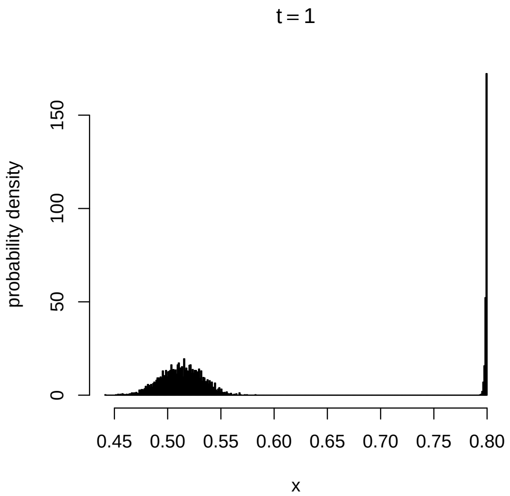
Ensembles converge to distributions on the attractor, not necessarily invariant ones

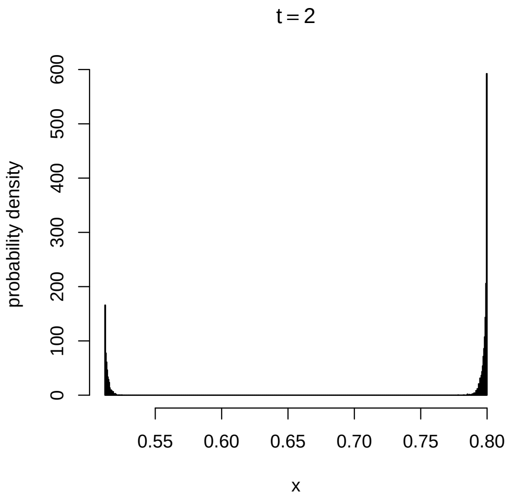
Example: stable 2-cycle; 3/4 of ensemble near one point, 1/4 near the other; will converge to a distribution on the cycle, but won't re-balance the probability

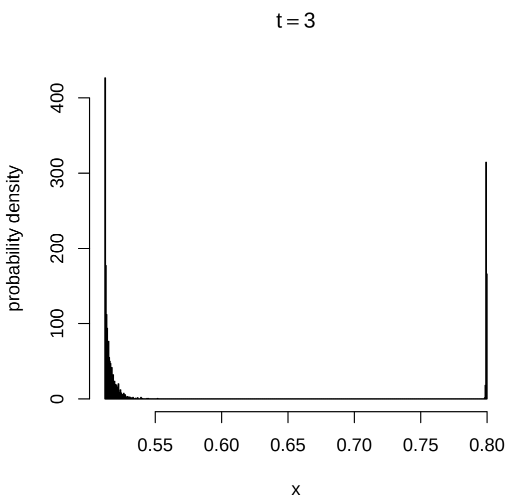
logistic map, $r = 0.8$, cycle from 0.5130445 to 0.7994555 and back

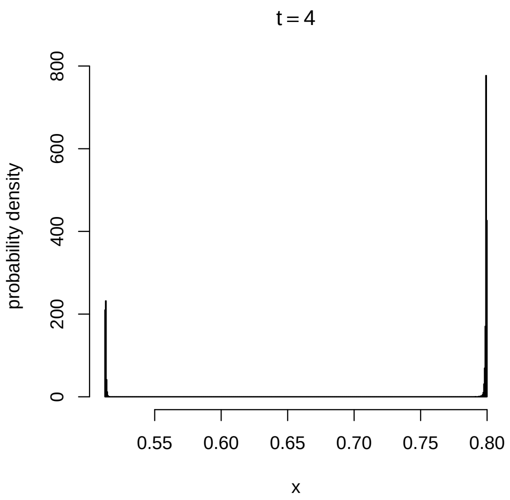
initial ensemble: two Gaussians centered at the cycle points, $\sigma = 0.01$, 4000 point, 3/4 near higher point

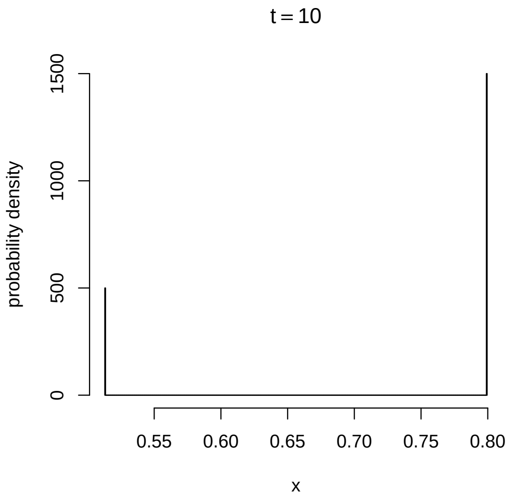

t = 0
probability density
0
5
10
15
20
25
30
0.50
0.55
0.60
0.65
0.70
0.75
0.80
x

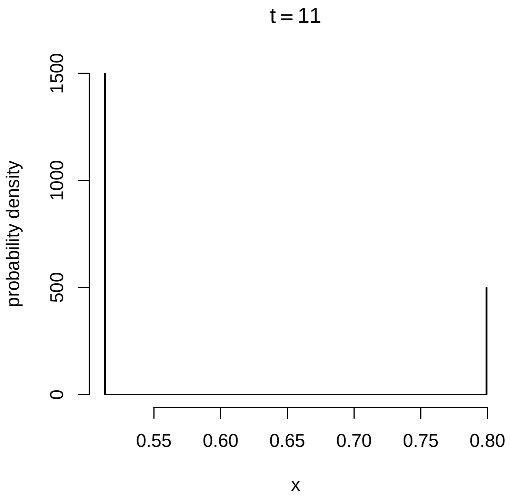

t = 1
probability density
0
50
100
150
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
x

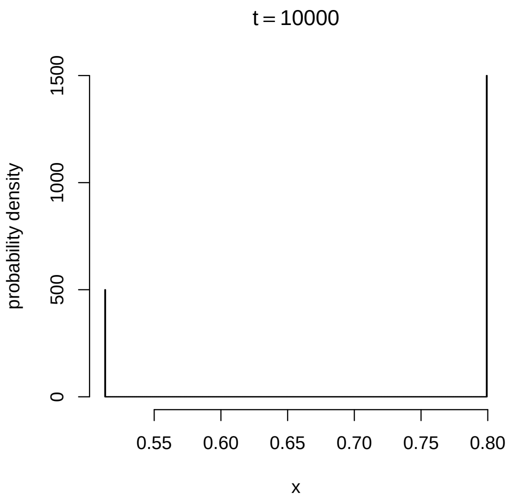

t = 2
probability density
0
100
200
300
400
500
600
0.55
0.60
0.65
0.70
0.75
0.80
x

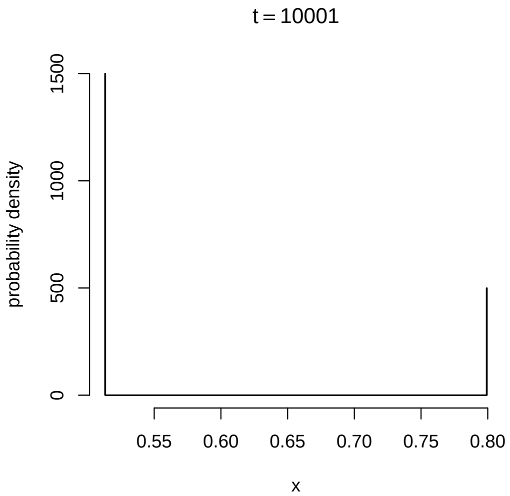

t = 3
probability density
0
100
200
300
400
0.55
0.60
0.65
0.70
0.75
0.80
x


t = 4
probability density
0
200
400
600
800
0.55
0.60
0.65
0.70
0.75
0.80
x


t = 10
probability density
0
500
1000
1500
0.55
0.60
0.65
0.70
0.75
0.80
x


t = 11
probability density
0
500
1000
1500
0.55
0.60
0.65
0.70
0.75
0.80
x


t = 10000
probability density
0
500
1000
1500
0.55
0.60
0.65
0.70
0.75
0.80
x


t = 10001
probability density
0
500
1000
1500
0.55
0.60
0.65
0.70
0.75
0.80
x

Ensemble: lopsided, alternates forever, oscillations never diminish

Individual trajectory's distribution: balanced, any initial bias diminishes steadily over time

Time-averaged ensemble looks like time-averaged single-trajectory distribution

Notice that there is no instability *within* the attractor here

If there was, we might expect ensembles to look more like individual trajectories

Mixing and decay of correlations

Mixing: As $\tau \to \infty$, X_t becomes independent of $X_{t+\tau}$

Equivalent to

Decay of correlations: for any reasonable functions g, h,

$$\mathrm{cov}[g(X_t), h(X_{t+\tau})] \xrightarrow[\tau \to \infty]{} 0$$

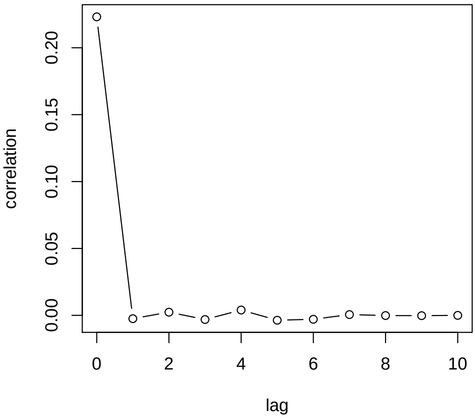
Equivalent to

Statistical stability: any initial ensemble converges to the natural invariant distribution

Probabilists: this is weak convergence

```
plot.decay.of.correlations = function(from=0,to=10,ts,...) {
  n=length(ts)
  rho = vector(mode="numeric",length=(to-from+1))
  for (t in from:to) {
    rho[t] = cor(cos(ts[1:(n-t)]),sin(ts[(1+t):n]))
  }
  plot(from:to,rho,xlim=c(from,to),xlab="lag",ylab="correlatio
}
```

Logistic map with $r = 1$, $g(x) = \cos x$, $h(x) = \sin x$, time series of length 10^5



Mixing $\Rightarrow$ ergodicity

Terminological confusion: an "ergodic Markov chain" is really a *mixing* Markov chain

Attractors can be non-mixing (e.g., periodic cycles)

Non-attractors can be mixing (e.g., cat map)

Chaotic attractors are (generally) mixing

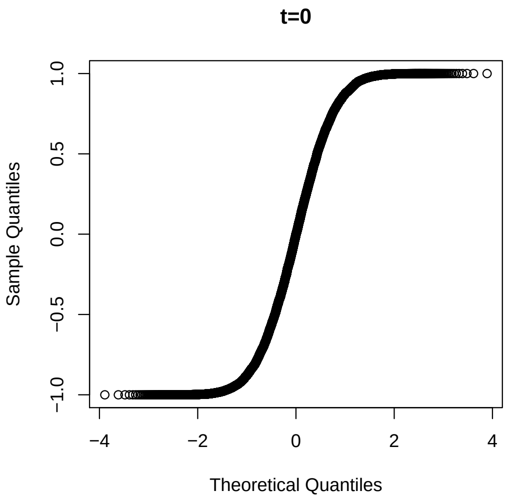
(hierarchy of ergodic properties — see [2, 3] for more)

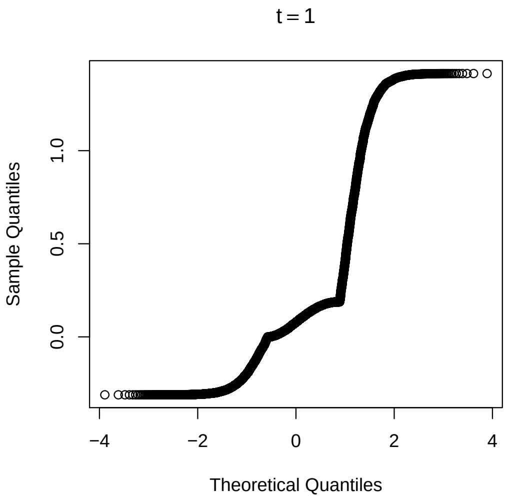
Some uses of mixing

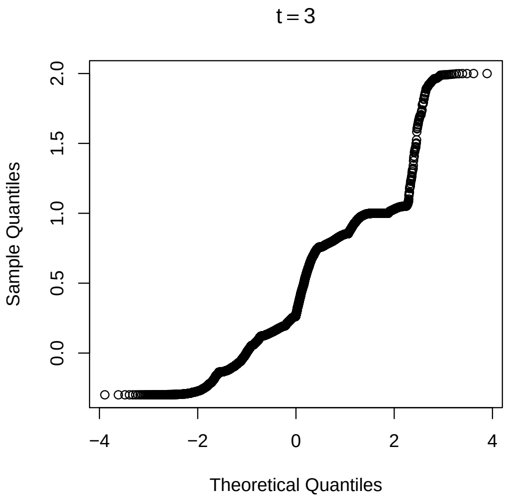
If $\tau \gg$ mixing time, X_t, $X_{t+\tau}$, $X_{t+2\tau}, \ldots$ are $\approx$ independent (sampling)

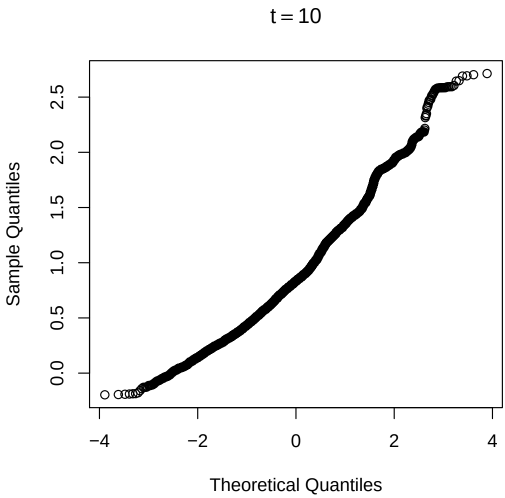
"Forgetting" of initial conditions

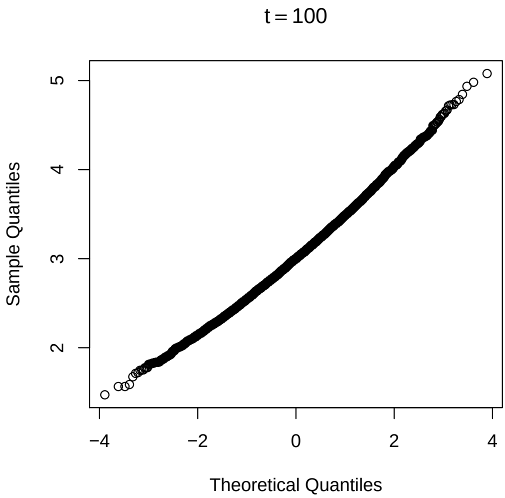
Central limit theorem: if correlations decay fast enough, time averages$*\sqrt{t}$ become Gaussian [4]


t=0
Sample Quantiles
1.0
0.5
0.0
−0.5
−1.0
−4
−2
0
2
4
Theoretical Quantiles


t = 1
Sample Quantiles
0.0
0.5
1.0
−4
−2
0
2
4
Theoretical Quantiles


t = 3
Sample Quantiles
2.0
1.5
1.0
0.5
0.0
−4
−2
0
2
4
Theoretical Quantiles


t = 10
Sample Quantiles
0.0
0.5
1.0
1.5
2.0
2.5
−4
−2
0
2
4
Theoretical Quantiles


t = 100
Sample Quantiles
2
3
4
5
−4
−2
0
2
4
Theoretical Quantiles

[1] P.-M. Binder and David H. Campos. Direct calculation of invariant measures for chaotic maps. *Physical Review E*, 53:R4259–R4262, 1996. URL `http://link.aps.org/abstract/PRE/v53/pR4259`.

[2] Michael C. Mackey. *Time's Arrow: The Origins of Thermodynamic Behavior*. Springer-Verlag, Berlin, 1992.

[3] Andrzej Lasota and Michael C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, Berlin, 1994. First edition, *Probabilistic Properties of Deterministic Systems*, Cambridge University Press, 1985.

[4] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences (USA)*, 42:43–47, 1956. URL `http://www.pnas.org/cgi/reprint/42/1/43`.