Noise and Chaos

Liverani Carlangelo

Università di Roma *Tor Vergata*

Ravello, Settembre 2016

12 September 2016

Lecture 1

Reductionism

The traditional reductionist approach to study nature consists in identifying the phenomenon one is interested in and then consider it as an isolated system. A classical and successful example is provided by Hamiltonian mechanics.

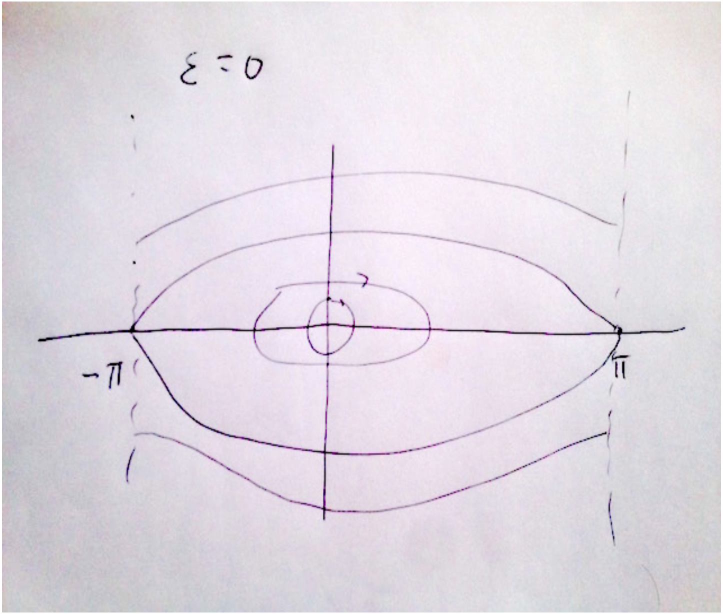
However at a more attentive exam the idea of isolated system seem to rest on very shaky grounds.

Isolated?

Let us consider a pendulum subject to a (small) periodic forcing

$$H_\varepsilon(\theta, p, t) = \frac{1}{2l^2 m} p^2 - mgl\cos\theta - \varepsilon m\omega^2 l \cos\omega t \cos\theta.$$

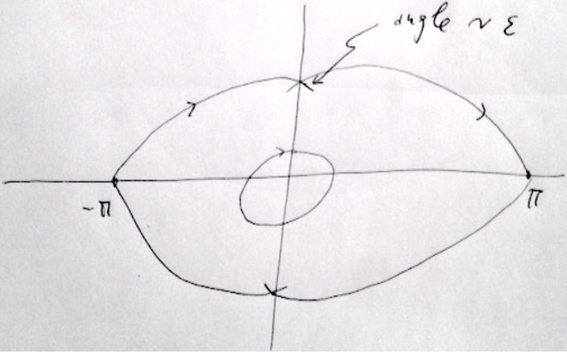
$l = 1m$, $m = 1kg$, $\varepsilon \sim 10^{-6}m$, $\omega = 1hz$ and with initial conditions such that $|p| + |\theta - \pi| \sim 10^{-3}m$.
For details see
http://www.mat.uniroma2.it/ liverani/SysDyn15/sd.html, fifth note.

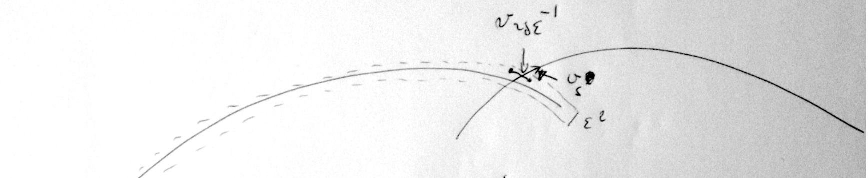


ε = 0
−π
π

$\varepsilon \neq 0$

Let us look at the poincaré map at time $2\pi/\omega$





$$(v, v_s^\perp) \sim \varepsilon |v| \sim \delta$$

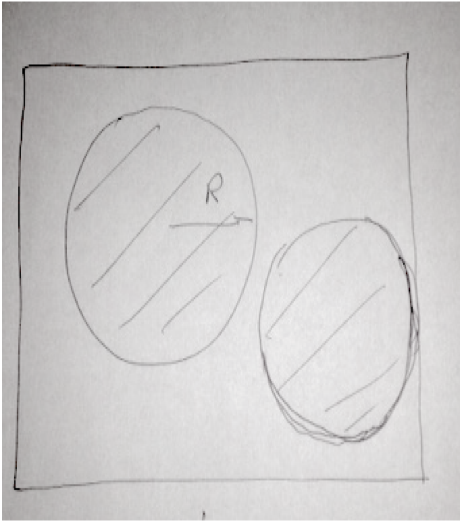
So when x, y get close to $(\pi, 0)$ their distance will be

$$\sim \delta \varepsilon^{-1}$$

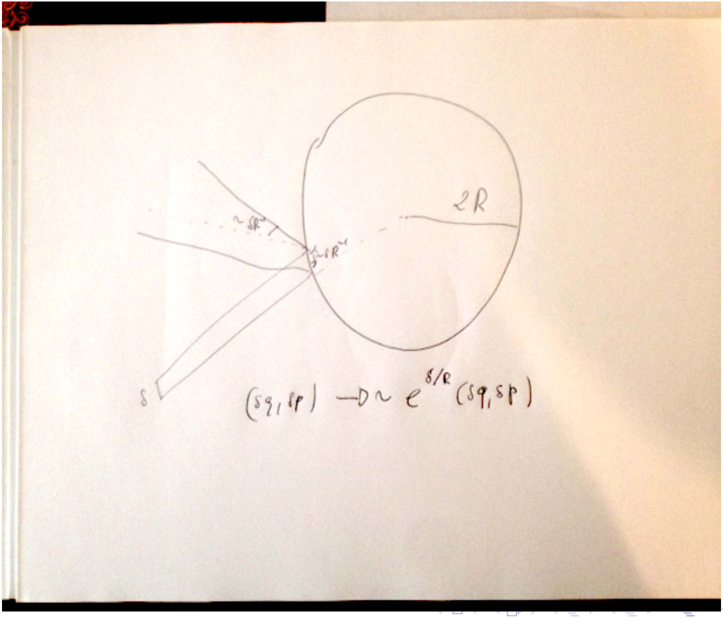
So, if we consider points that go back to our set of initial conditions (so we can see what happens next), then to predict if the pendulum will rotate or oscillate in about 6 seconds we need to know the initial condition with a precision of δ such that $\delta\varepsilon^{-1} \sim \varepsilon$, that is or order 10^{-6}, for 12 seconds of 10^{-9} and so on.

Let us try again

Consider two billiards ball of radius R is a square table of size L with mass one and kinetic energy K. To simplify matters consider the case in which $2\sqrt{2}R < L < 4R$ so that there exists a length ℓ_0 such that the distance between two collisions of the balls is, at most, ℓ_0.


R

Then a change of the initial condition by ε will create a change of velocities after the next collision proportional to R^{-1}. Thus the change in the trajectories will grow, at least, like $e^{N/R}\varepsilon$, where N is the number of collisions among balls.


2R
δ
$(\delta q, \delta p) \longrightarrow \sim e^{\delta/R}(\delta q, \delta p)$

Note that $N\ell_0 \leq T\sqrt{2K}$, where T is the time. Thus the change in the trajectory due to a small initial perturbation grows, at least, like $e^{T\sqrt{2K}/R^2}\varepsilon$.
If $R = .1$ meters e $\sqrt{2K} = 10$ meters per second (36 Km/h), then the perturbation of the trajectory grows, at least, like $e^{10^3 T}$.
If you observe two identical systems and, at a certain point, only on one of them, acts, for 10^{-10} seconds, a force of size 10^{-90} newtons, then this will create a change in velocity of size 10^{-100} meters per second which, after a tenth of a second, will create a difference in the coordinates of the same size of the box.

Noise

A standard way of taking into account all the above issues is to add to the system a small random perturbation: if you have

$$\dot{x} = F(x),$$

you might add to it a noise writing

$$dx = F(x)dt + \varepsilon\Sigma(x)dB$$

where B is a d dimensional Brownian motion and $\Sigma(x)$ is a positive symmetric matrix.

Consider, for example, the the noisy Hamiltonian system

$$
\begin{aligned}
dq &= pdt \\
dp &= -V'(q)dt + \varepsilon\sigma dB
\end{aligned}
$$

where we put the noise only on the second equation because we think of it as a random force acting on the system.

By Ito's formula $\left[df(B) = f'(B)dB + \frac{1}{2}f''(B)dt\right]$

$$dH = \frac{\varepsilon^2\sigma^2}{2}dt + \varepsilon\sigma p dB.$$

Hence,

$$\mathbb{E}(H) = \frac{\varepsilon^2\sigma^2}{2}t.$$

The system keeps heating up. Not what we see !!!!!

The usual fix for this problem is to consider the equations

$$\begin{aligned} dq &= pdt \\ dp &= -V'(q)dt - \gamma pdt + \varepsilon\sigma dB, \end{aligned}$$

where we have added a *friction* to the system. The above is called a Langevin equation or an Ornstein-Uhlenbeck process.

Such a process does now have an invariant measure:

$$
\begin{aligned}
&\frac{d}{dt}\int e^{-\beta H(q,p)}\mathbb{E}(\varphi(q(t,q,p),p(t,q,p)))\Big|_{t=0} \\
&= \int e^{-\beta H(q,p)}\left\{p\partial_q\varphi - [V'(q)+\gamma p]\partial_p\varphi + \frac{\varepsilon^2\sigma^2}{2}\partial_p^2\varphi\right\} \\
&= \int e^{-\beta H(q,p)}\left\{-\beta\gamma p^2 + \gamma + \varepsilon^2\sigma^2\beta^2 p^2 - \beta\varepsilon^2\sigma^2\right\}\varphi.
\end{aligned}
$$

Thus the derivative is zero provided $\gamma = \varepsilon^2\sigma^2\beta$ (this is some sort of Einstein relation).

But how can noise emerge from a deterministic system. Where does the probability comes from?

13 September 2016

Lecture 2

A super simple example

The perturbation of a trivial (macroscopic) dynamics by an external (microscopic) degree of freedom:

let $F_\varepsilon \in \mathcal{C}^r(\mathbb{T}^2, \mathbb{T}^2)$, $r > 1$, defined as

$$
\begin{aligned}
&F_\varepsilon(x,\theta) = (f(x,\theta), \theta + \varepsilon\omega(x,\theta) \mod 1)\\
&\partial_x f \geq \lambda > 1; \quad \|\omega\|_{\mathcal{C}^r} = 1.
\end{aligned}
$$

A first question

We are interested in the behaviour of $\theta_\varepsilon \in \mathcal{C}^0([0,T],\mathbb{T})$ defined by

$$\theta_\varepsilon(t) = \theta_{\lfloor \varepsilon^{-1} t\rfloor} + (\varepsilon^{-1} t - \lfloor \varepsilon^{-1} t\rfloor)(\theta_{\lfloor \varepsilon^{-1} t\rfloor+1} - \theta_{\lfloor \varepsilon^{-1} t\rfloor}),$$

and, first, we should ask ourselves:

Does θ_ε has some limiting behaviour for $\varepsilon \to 0$?

Even simpler

To further simplify the problem, let us start with the case $\partial_\theta \omega = \partial_\theta f = 0$. This is called a skew product. Then

$$\left| \theta_\varepsilon(t) - \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0) \right| \le C_\# \varepsilon.$$

Thus our variable is described by an ergodic average.
By Birkhoff ergodic theorem

$$\lim_{\varepsilon \to 0} \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0)$$

exists for almost every point with respect to any invariant measure of f . . . . invariant measures?

Before discussing the issue of invariant measures let us comment on Birkhoff theorem. It is essentially a law of large numbers for deterministic systems. Let us recall the simple case for independent i.i.d. random variables $\{X_k\}$.

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1} X_k - \mathbb{E}(X_k)\right]^2 = \frac{1}{N^2}\sum_{k=0}^{N-1}\mathbb{E}([X_k - \mathbb{E}(X_k)]^2) = \frac{1}{N}\mathbb{E}([X - \mathbb{E}(X)]^2)$$

Thus $\frac{1}{N}\sum_{k=0}^{N-1} X_k$ converges to $\mathbb{E}(X)$ in L^2.
Can we have stronger convergence?

Interlude

By Chebyshev inequality

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{k=0}^{N-1} X_k - \mathbb{E}(X_k)\right| \geq \delta\right) \leq \delta^{-2}\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1} X_k - \mathbb{E}(X_k)\right]^2 \leq \frac{C_\#}{N}.$$

Unfortunately, summing on N we get infinity, so one cannot directly get almost sure convergence.

Interlude

Yet, if we assume $\|X\|_{L^\infty} = K < \infty$, then for each $N \geq M > 0$

$$\left\| \frac{1}{N} \sum_{k=0}^{N-1} X_k - \frac{1}{M} \sum_{k=0}^{M-1} X_k \right\|_{L^\infty} \leq C_\# \frac{N-M}{M^2}$$

Thus setting $N = 2^m + j$, then for each $j \in \{0, \dots, 2^m - 1\}$,

$$\left\| \frac{1}{N} \sum_{k=0}^{N-1} X_k - \frac{1}{2^m} \sum_{k=0}^{2^m-1} X_k \right\|_{L^\infty} \leq C_\# 2^{-m}.$$

Hence for all δ and $2^{-N_*} \leq \delta^3$

$$\mathbb{P}\left(\left\{sup_{N\geq N_*}\left|\frac{1}{N}\sum_{k=0}^{N-1} X_k - \mathbb{E}(X_k)\right| \geq \delta\right\}\right)$$
$$\leq \mathbb{P}\left(\left\{sup_{m\geq \ln N_*}\left|2^{-m}\sum_{k=0}^{2^m-1} X_k - \mathbb{E}(X_k)\right| \geq \delta/2\right\}\right)$$
$$\leq C_\# \sum_{m\geq \ln N_*} 2^{-m}\delta^{-2} \leq C_\#\delta$$

hence we have almost sure convergence.

14 September 2016

Lecture 3

Let us consider the map $f \in \mathcal{C}^\infty(\mathbb{T},\mathbb{T})$ defined by

$$f(x) = x + \omega \quad \mod 1$$

where $\omega \notin \mathbb{Q}$. Then, for all $\phi \in \mathcal{C}^2(\mathbb{T},\mathbb{R})$,

$$\frac{1}{N}\sum_{n=0}^{N-1} \phi \circ f^n(x) = \frac{1}{N}\sum_{n=0}^{N-1} \phi(x + n\omega) = \sum_{k\in\mathbb{Z}} \hat{\phi}_k e^{2\pi i k x} \frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i k n\omega}$$

$$= \hat{\phi}_0 + \sum_{\substack{k\in\mathbb{Z}\\ k\neq 0}} \hat{\phi}_k e^{2\pi i k x} \frac{1}{N} \frac{1 - e^{2\pi i k N\omega}}{1 - e^{2\pi i k\omega}}$$

Unique ergodicity

Thus

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \phi \circ f^n(x) = \hat{\phi}_0 = \int_{\mathbb{T}} \phi(x) dx$$

Note that the convergence is uniform!
This is way too much to expect in general and is related to the fact that Leb is the **unique invariant measure** for f.
This false in general.
Solve the case $\omega \in \mathbb{Q}$ to see which kind of catastrophes can happen.

Interlude

Before discussing the issue of invariant measures let us comment on Birkhoff theorem. It is essentially a law of large numbers for deterministic systems. Let us recall the simple case for independent i.i.d. random variables $\{X_k\}$.

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=0}^{N-1} X_k - \mathbb{E}(X_k)\right]^2 = \frac{1}{N^2}\sum_{k=0}^{N-1}\mathbb{E}([X_k - \mathbb{E}(X_k)]^2) = \frac{1}{N}\mathbb{E}([X - \mathbb{E}(X)]^2)$$

Thus $\frac{1}{N}\sum_{k=0}^{N-1} X_k$ converges to $\mathbb{E}(X)$ in L^2.
Can we have stronger convergence?

To further simplify the problem, let us start with the case $\partial_\theta \omega = \partial_\theta f = 0$. This is called a skew product. Then

$$\left| \theta_\varepsilon(t) - \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0) \right| \leq C_\# \varepsilon.$$

Thus our variable is described by an ergodic average.
By Birkhoff ergodic theorem

$$\lim_{\varepsilon \to 0} \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0)$$

exists for almost every point with respect to any invariant measure of f . . . . invariant measures?

Invariant measures (too much of a good thing)

A standard way to construct invariant measures (Krylov-Bogoliubov) is to start with an arbitrary measure μ and take the average of the pushforward $f_*^k\mu(\varphi) =: \mu(\varphi \circ f^k)$. Indeed,

$$\frac{1}{n}\sum_{k=0}^{n-1} f_*^k \mu$$

is a weakly compact set, hence it has accumulation points.
It is easy to check that such accumulation points are invariant measures (i.e. $f_*\nu = \nu$).

Invariance

Indeed, let ν be an accumulation point, then for all $\varphi \in \mathcal{C}^0$ we have

$$\begin{aligned} f_*\nu(\varphi) &= \nu(\varphi \circ f) = \lim_{j\to\infty} \frac{1}{n_j} \sum_{k=0}^{n_j-1} \mu(\varphi \circ f^{k+1}) \\ &= \lim_{j\to\infty} \frac{1}{n_j} \sum_{k=0}^{n_j-1} \mu(\varphi \circ f^k) + \frac{1}{n_j}\left[\mu(\varphi \circ f^{n_j}) - \mu(\varphi)\right] \\ &= \nu(\varphi). \end{aligned}$$

Physical measures

Suppose that the initial measure is a.c.w.r.t. Lebesgue: $d\mu = hd\text{Leb}$. A simple change of variables (do it!) shows that $\frac{d(f_*\mu)}{d\text{Leb}} = \mathcal{L}h$ where

$$\mathcal{L}h(x) = \sum_{f(y)=x} \frac{h(y)}{f'(y)}.$$

The operator $\mathcal{L}$ is called a (Ruelle) transfer operator. Of course an operator, to be properly defined, must have a well specified domain.

Since

$$\int |\mathcal{L}h(x)|dx \le \int \mathcal{L}|h|(x)dx = \int |h(x)|dx$$

it follows that $\mathcal{L}$ is a contraction on $L^1(\mathbb{T}, \text{Leb})$. However, the spectrum of $\mathcal{L}$ on L^1 turns out to be the full unit disk, not a very useful fact.

Following Lasota-Yorke, we look then at the action of $\mathcal{L}$ on $W^{1,1}$.

$$\frac{d}{dx}\mathcal{L}h = \mathcal{L}\left(\frac{h}{f'}\right) - \mathcal{L}\left(h\frac{f''}{(f')^2}\right).$$

The above implies the so called Lasota-Yorke inequalities

$$\begin{aligned}
&\|\mathcal{L}h\|_{L^1} \le \|h\|_{L^1}\\
&\|(\mathcal{L}h)'\|_{L^1} \le \lambda^{-1}\|h'\|_{L^1} + D\|h\|_{L^1}.
\end{aligned}$$

Such inequalities imply that $\mathcal{L}$, when acting on $W^{1,1}$, has a spectral gap. To give an idea of the why, let us consider the simple case in which $D = \|\frac{f''}{(f')^2}\|_{L^\infty}$ is small, more precisely $\lambda^{-1} + D < 1$.

If $\mathrm{Leb}(h) = 0$, then $\mathrm{Leb}(\mathcal{L}h) = 0$, hence the space $\mathbb{V} = \{h \in L^1 \; : \; \mathrm{Leb}(h) = 0\}$ is invariant.
Indeed

$$\mathrm{Leb}(\mathcal{L}h) = \int_{\mathbb{T}} \mathcal{L}h(x)dx = \int_{\mathbb{T}} 1 \circ f \cdot h(x)dx = \int_{\mathbb{T}} h(x)dx = \mathrm{Leb}(h) = 0.$$

Spectral gap II

If $h \in \mathbb{V}$, then

$$\|h\|_{L^1} = \int_{\mathbb{T}} |h(x)| \leq \|h'\|_{L^1}.$$

Define the norm $\|h\|_{W^{1,1}} = \|h'\|_{L^1} + a\|h\|_{L^1}$ for some $a > 0$. Then, for $h \in \mathbb{V}$,

$$\begin{aligned}\|\mathcal{L}h\|_{W^{1,1}} &\leq \lambda^{-1}\|h'\|_{L^1} + (D + a)\|h\|_{L^1} \leq (\lambda^{-1} + D + a)\|h'\|_{L^1} \\ &\leq (\lambda^{-1} + D + a)\|h\|_{W^{1,1}}.\end{aligned}$$

Choosing a such that $\lambda^{-1} + D + a < 1$ it follows that $\mathcal{L}$ is a strict contraction on $\mathbb{V}$.

Spectral gap III

Since $\mathcal{L}'\text{Leb} = \text{Leb}$, $1 \in \sigma(\mathcal{L})$ and we have that $\exists!\ h_* \in L^1$ such that

$$\mathcal{L}h = h_*\text{Leb}(h) + Qh,$$

where $\|Q\|_{W^{1,1}} < 1$ and $\text{Leb}Q = Qh_* = 0$.

Hence, $h_*(x)dx$ is the only invariant measure of f absolutely continuous with respect to Lebesgue. In fact it is equivalent to Lebesgue, i.e. $h_* > 0$.

15 September 2016

Lecture 4

What was the problem again?

Recall that we are discussing the simple case $\partial_\theta \omega = \partial_\theta f = 0$.

$$\left| \theta_\varepsilon(t) - \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0) \right| \leq C_\# \varepsilon.$$

Thus our variable is described by an ergodic average.
By Birkhoff ergodic theorem

$$\lim_{\varepsilon \to 0} \varepsilon \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \omega \circ f^k(x_0)$$

exists for almost every point with respect to any invariant measure of f . . . . invariant measures?

Averaging

Thus, for lebesgue almost all x,

$$\lim_{\varepsilon\to 0}\theta_\varepsilon(t) = \bar\theta(t) = t\int \omega(x)h_*(x)dx =: t\bar\omega.$$

That is, the limit satisfies the differential equation

$$\frac{d}{dt}\bar\theta = \bar\omega.$$

This is a simple example of *averaging*, first done by Anosov '60.

You do not want to use Birkhoff theorem?
No problem: argue as we did for independent random variables.
Let $\rho \in \mathcal{C}^1$ be the density of a probability measure, $\hat\omega = \omega - \bar\omega$,

$$
\begin{aligned}
&\mathrm{Leb}\left(\rho\left[\varepsilon\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\hat\omega\circ f^k\right]^2\right)\\
&=\varepsilon^2\sum_{k,j=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\mathrm{Leb}\left(\rho\hat\omega\circ f^k\hat\omega\circ f^j\right)\\
&=\varepsilon^2\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\mathrm{Leb}\left(\hat\omega^2\mathcal{L}^k\rho\right)+2\varepsilon^2\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\sum_{j=1}^{\lfloor\varepsilon^{-1}t\rfloor-1-k}\mathrm{Leb}\left(\hat\omega\circ f^k\,\hat\omega\circ f^{j+k}\rho\right).
\end{aligned}
$$

Decay of correlations

$$
\begin{aligned}
&\mathrm{Leb}\left(\hat\omega\circ f^k\,\hat\omega\circ f^{j+k}\rho\right)=\mathrm{Leb}\left(\hat\omega\circ f^j\,\hat\omega\,\mathcal{L}^k\rho\right)=\mathrm{Leb}\left(\hat\omega\mathcal{L}^j\left[\hat\omega\mathcal{L}^k\rho\right]\right)\\
&=\mathrm{Leb}\left(\hat\omega h_*\right)\mathrm{Leb}\left(\hat\omega\mathcal{L}^k\rho\right)+\mathcal{O}\left(\|Q^j\|\|\hat\omega\mathcal{L}^k\rho\|_{W^{1,1}}\right).
\end{aligned}
$$

Thus

$$
\mathrm{Leb}\left(\rho\left[\varepsilon\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\hat\omega\circ f^k\right]^2\right)\le C_\#\varepsilon
$$

The general case

In general, any map $f_\theta(x) = f(x, \theta)$ has a unique invariant physical measure μ_θ with density h_θ. Define

$$\bar{\omega}(\theta) = \mu_\theta(\omega(\cdot, \theta)) \; ; \quad \hat{\omega} = \omega - \bar{\omega}.$$

Then the accumulation points of θ_ε satisfy,

$$\frac{d}{dt}\bar{\theta} = \bar{\omega}(\bar{\theta}).$$

For details see:

Jacopo De Simoi, Carlangelo Liverani, *The Martingale approach after Varadhan and Dolpogpyat* . In "Hyperbolic Dynamics, Fluctuations and Large Deviations", Dolgopyat, Pesin, Pollicott, Stoyanov editors, Proceedings of Symposia in Pure Mathematics, **89**, AMS (2015).

OK, but what about the noise?

We have just seen that an isolated dynamics can arise from the interaction from a suppressed degree of freedom, but can we see the presence of such a degree of freedom if we look closer?
To this end consider the variable (fluctuations)

$$\zeta_\varepsilon = \frac{1}{\sqrt{\varepsilon}}(\theta_\varepsilon(t) - \bar{\theta}(t))$$

Initial conditions

To continue we have to be more precise on the initial conditions: we consider the system with random initial conditions such that, for each $\varphi \in \mathcal{C}^0$,

$$\mathbb{E}(\varphi(x_0, \theta_0)) = \int_{\mathbb{T}} \varphi(x, \theta_*)\rho(x)dx$$

where $\rho \in W^{1,1}$ and $\theta_* \in \mathbb{T}$.

Remark

Technically these initial conditions are a special case of the measures called standard pairs *introduced by Dolgopyat and that are a basic tool to investigate the statistical properties of systems with some chaoticity.*

With such an initial condition the ζ_ε are random variables and if they have a limiting behaviour then they could be responsible for the appearance of the noise.
To this end the standard strategy is to study the characteristic function

$$\Phi(\xi) = \mathbb{E}\left(e^{i\xi\zeta_\varepsilon(t)}\right).$$

Let us consider again the case of a skew product (i.e. $\partial_\theta \omega = \partial_\theta f = 0$).

$$\left| \zeta_\varepsilon(t) - \sqrt{\varepsilon} \sum_{k=0}^{\lfloor \varepsilon^{-1} t \rfloor - 1} \hat{\omega} \circ f^k(x_0) \right| \leq C\sqrt{\varepsilon},$$

where $\hat{\omega} = \omega - \bar{\omega}$. So, up to a precision of order ε, our problem is equivalent to the one of studying the characteristic function of the sum.

Brute force

That is we aim at computing

$$\mathbb{E}\left(\exp\left[i\xi\sqrt{\varepsilon}\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\hat{\omega}\circ f^{k}\right]\right).$$

Interlude

How one does that for i.i.d. random variables? Let $\hat{X} = X - \mathbb{E}(X)$, then

$$
\begin{aligned}
&\mathbb{E}\left(\exp\left\{\frac{\xi i}{\sqrt{N}}\sum_{k=0}^{N-1}\hat{X}_k\right\}\right) = \left[\mathbb{E}\left(\exp\left\{\frac{\xi i}{\sqrt{N}}\hat{X}\right\}\right)\right]^N \\
&= \left[1 + \frac{\xi i}{\sqrt{N}}\mathbb{E}(\hat{X}) - \frac{\xi^2}{2N}\mathbb{E}(X^2) + \mathcal{O}(N^{-\frac{3}{2}})\right]^N \\
&= \exp\left[N\ln\left\{1 + \frac{\xi i}{\sqrt{N}}\mathbb{E}(\hat{X}) - \frac{\xi^2}{2N}\mathbb{E}(X^2) + \mathcal{O}(N^{-\frac{3}{2}})\right\}\right] \\
&= \exp\left[-\frac{1}{2}\xi^2\mathbb{E}(X^2) + \mathcal{O}(N^{-\frac{1}{2}})\right]
\end{aligned}
$$

Interlude

Hence

$$\lim_{N\to\infty}\mathbb{E}\left(\exp\left\{\frac{\xi i}{\sqrt{N}}\sum_{k=0}^{N-1}\hat{X}_k\right\}\right)=\exp\left[-\frac{1}{2}\xi^2\mathbb{E}(X^2)\right]$$

Setting $\sigma^2=\mathbb{E}(X^2)$, we thus have, for all smooth φ,

$$\lim_{N\to\infty}\mathbb{E}\left(\varphi\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\hat{X}_k\right)\right)=\int_{\mathbb{R}}d\xi\,\hat{\varphi}(\xi)e^{-\frac{1}{2}\xi^2\sigma^2}$$
$$=\int\varphi(x)\frac{1}{\sqrt{2\sigma^2\pi}}e^{-\frac{x^2}{2\sigma^2}}dx.$$

This is the Central Limit Theorem. OK, but the deterministic case?

16 September 2016

Lecture 5

Define the transfer operator, for each $\phi \in L^1$,

$$\mathcal{L}_\nu \phi(x) = \sum_{f(y)=x} \frac{e^{i\nu\hat{\omega}(y)}}{f'(y)} \phi(y)$$

and notice that (this is a nice exercise)

$$\mathbb{E}\left(\exp\left[i\xi\sqrt{\varepsilon}\sum_{k=0}^{\lfloor \varepsilon^{-1}t\rfloor - 1} \hat{\omega} \circ f^k\right]\right) = \int_{\mathbb{T}} \mathcal{L}_{\xi\sqrt{\varepsilon}}^{\lfloor \varepsilon^{-1}t\rfloor} \rho.$$

Note that $\mathcal{L}_0 = \mathcal{L}$. Since $\mathcal{L}$ has a spectral gap on $W^{1,1}$, it makes sense to try to apply perturbation theory.

Lemma

There exists $\nu_0 > 0$ and continuous functions $C_\nu > 0$ and $\rho_\nu \in (0,1)$ such that, for all $|\nu| \le \nu_0$, $\mathcal{L}_\nu = e^{\alpha_\nu}\Pi_\nu + Q_\nu$, $\Pi_\nu Q_\nu = Q_\nu \Pi_\nu = 0$, $\|Q_\nu^n\|_{W^{1,1}} \le C_\nu \rho_\nu^n e^{\alpha_\nu n}$. Also $\Pi_\nu(g) = h_\nu \ell_\nu(g)$, $\ell_\nu(h_\nu) = 1$, $\ell_\nu(h_\nu') = 0$. In addition, everything is analytic in ν.

Lemma

Fot all $|\nu| \leq \nu_0$*, the function* α_ν *satisfies* $\alpha_0 = \alpha_0' = 0$ *and* $|\alpha_\nu|_{\mathcal{C}^3} \leq C_\#$, $\alpha_0'' \leq 0$*. Finally,* $\alpha_0'' = 0$ *iff there exists* $g \in \mathcal{C}^0$ *such that* $\hat{\omega} = g - g \circ f$*; that is, only if* $\hat{\omega}$ *is a* $\mathcal{C}^0$-coboundary.

So what?

Given the above Lemmata

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\exp\left[i\xi\sqrt{\varepsilon}\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\hat{\omega}\circ f^k\right]\right) = \lim_{\varepsilon\to 0}\int_{\mathbb{T}}\mathcal{L}_{\xi\sqrt{\varepsilon}}^{\lfloor\varepsilon^{-1}t\rfloor}\rho$$
$$= \lim_{\varepsilon\to 0} e^{-\frac{1}{2}\alpha_0''\xi^2 t+\mathcal{O}(\xi^3)\sqrt{\varepsilon}} = \exp\left[-\frac{1}{2}\alpha_0''\xi^2 t\right]$$

which is the Fourier transform of a Gaussian, thus we have convergence to a Gaussian random variable.

"Proof of Lemma"

We have seen that

$$\mathcal{L}_\nu h_\nu = e^{\alpha_\nu} h_\nu \,; \quad \ell_\nu(\mathcal{L}_\nu \phi) = e^{\alpha_\nu} \ell_\nu(\phi)$$

where $\alpha_0 = 1$, $h_0 = h_*$ and $\ell_0 = \text{Leb}$. Differentiating

$$\mathcal{L}_\nu' h_\nu + \mathcal{L}_\nu h_\nu' = \alpha_\nu' e^{\alpha_\nu} h_\nu + e^{\alpha_\nu} h_\nu'.$$

Applying ℓ_ν yields

$$\frac{d\alpha_\nu}{d\nu} = i\ell_\nu(\hat{\omega} h_\nu) =: i\mu_\nu(\hat{\omega}).$$

Thus $\alpha_0' = 0$.

"Proof of Lemma"

Setting $\hat{\mathcal{L}}_\nu := e^{-\alpha_\nu}\mathcal{L}_\nu$, $\hat{Q}_\nu := e^{-\alpha_\nu}Q_\nu$ and $\omega_\nu = \omega - \ell_\nu(\omega h_\nu)$ we have

$$h'_\nu = i(\mathbb{1} - \hat{Q}_\nu)^{-1}\hat{\mathcal{L}}_\nu(\omega_\nu h_\nu).$$

Doing similar considerations on the equation $\ell_\nu(\mathcal{L}_\nu) = \alpha_\nu \ell_\nu(g)$, we obtain

$$\begin{aligned}
\alpha''_\nu &= -\ell_\nu(\omega_\nu(\mathbb{1} - \hat{Q}_\nu)^{-1}(\mathbb{1} + \hat{Q}_\nu)(\omega_\nu h_\nu)) \\
&= -\sum_{n=1}^{\infty} \ell_\nu(\omega_\nu \hat{\mathcal{L}}_\nu^n(\mathbb{1} + \hat{\mathcal{L}}_\nu)(\omega_\nu h_\nu)) \\
&= -\mu_\nu(\omega_\nu^2) - 2\sum_{n=1}^{\infty} \ell_\nu(\omega_\nu \hat{\mathcal{L}}_\nu^n(\omega_\nu h_\nu)).
\end{aligned}$$

Finally, notice that

$$\ell_\nu(\omega_\nu \hat{\mathcal{L}}_\nu^n(\omega_\nu h_\nu)) = \ell_\nu(\hat{\mathcal{L}}_\nu^n(\omega_\nu \circ f^n \omega_\nu h_\nu)) = \mu_\nu(\omega_\nu \circ f^n \omega_\nu)$$

and

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\mu_\nu\left(\left[\sum_{k=0}^{n-1}\omega_\nu \circ f^k\right]^2\right) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k,j=0}^{n-1}\mu_\nu(\omega_\nu \circ f^k \omega_\nu \circ f^j)\\
&= \mu_\nu(\omega_\nu^2) + \lim_{n\to\infty}\frac{2}{n}\sum_{k=1}^{n-1}(n-k)\mu_\nu(\omega_\nu \circ f^k \omega_\nu)\\
&= \mu_\nu(\omega_\nu^2) + 2\sum_{k=1}^{\infty}\mu_\nu(\omega_\nu \circ f^k \omega_\nu).
\end{aligned}$$

As a result we get the so called *Green-Kubo formula*.

$$-\alpha_0'' = \lim_{n\to\infty}\frac{1}{n}\mu_0\left(\left[\sum_{k=0}^{n-1}\omega_0 \circ f^k\right]^2\right) \geq 0.$$

L^2 Coboundaries

Finally, if $\alpha_0'' = 0$, then

$$\mu_0\left(\left[\sum_{k=0}^{n-1}\omega_0\circ f^k\right]^2\right) = -2\sum_{k=1}^{n-1}k\text{Leb}(\hat\omega\circ f^k\hat\omega) \le C_\#.$$

Accordingly, $\sum_{k=0}^{n-1}\omega_0\circ f^k$ is weakly compact, in L^2.
Let $g\in L^2$ be an accumulation point. Then, for each $\phi\in W^{1,1}$,

$$\begin{aligned}\text{Leb}\left(\phi(g\circ f-g)\right) &= \lim_{j\to\infty}\sum_{k=0}^{n_j-1}\text{Leb}\left(\phi[\hat\omega\circ f^{k+1}-\hat\omega\circ f^k]\right)\\ &= -\text{Leb}(\phi\hat\omega)+\lim_{j\to\infty}\text{Leb}(\hat\omega\mathcal{L}^{n_j}\phi) = -\text{Leb}(\phi\hat\omega).\end{aligned}$$

Since $W^{1,1}$ is dense in L^2, it follows that $g-g\circ f=\hat\omega$.
Thus $\hat\omega$ is an L^2 coboundary.

Also $g \in L^2(\mu_0)$ and without loss of generality, we can assume $\mu_0(g) = 0$. Then, multiplying by h_0 and applying $\mathcal{L}$

$$\mathcal{L}\hat{\omega} h_0 = \mathcal{L} g h_0 - \mathcal{L}(g \circ f h_0) = \mathcal{L}(g h_0) - g h_0 = (Q_0 - \mathbb{1})(g h_0).$$

That is $g h_0 = -(\mathbb{1} - Q_0)^{-1} \mathcal{L}\hat{\omega} h_0 \in W^{1,1} \subset \mathcal{C}^0$. On the other hand it is easy to show that $h_0 > 0$ (another nice exercize). Hence it must be $h_0 > 0$ and then $g \in \mathcal{C}^0$.

17 September 2016

Lecture 6

Recapping

We have thus seen that, provided ν_0 is small enough, for all $|\xi| \le \nu_0 \varepsilon^{-\frac{1}{2}}$, we have

$$\mathbb{E}\left(\exp\left[i\xi\sqrt{\varepsilon}\sum_{k=0}^{\lfloor\varepsilon^{-1}t\rfloor-1}\hat{\omega}\circ f^k\right]\right) = \int_{\mathbb{T}} \mathcal{L}_{\xi\sqrt{\varepsilon}}^{\lfloor\varepsilon^{-1}t\rfloor}\rho$$
$$= \exp\left[-\frac{1}{2}\alpha_0''\xi^2 t + \mathcal{O}(\xi^3\sqrt{\varepsilon}t)\right] + \mathcal{O}(e^{-\varepsilon^{-1}}).$$

Milking the cow

Consider a measure device represented by the function $\psi_{\varepsilon,z}(\zeta_\varepsilon) = \psi((\zeta_\varepsilon - z)\varepsilon^{-\alpha})$, where $\psi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}_+)$ is symmetric, has support in the interval $[-1,1]$ and $\alpha \in [0, 1/2)$. For simplicity put $t = 1$.

$$
\begin{aligned}
\mathbb{E}(\psi_{\varepsilon,z}(\zeta_\varepsilon)) &= \mathbb{E}\left(\frac{1}{2\pi}\int \hat{\psi}_{\varepsilon,z}(\xi)e^{i\xi\zeta_\varepsilon}\right) = \frac{1}{2\pi}\int \hat{\psi}_{\varepsilon,z}(\xi)\mathbb{E}(e^{i\xi\zeta_\varepsilon}) \\
&= \frac{1}{2\pi}\int \hat{\psi}_{\varepsilon,z}(\xi)\int_{\mathbb{T}} \mathcal{L}_{\xi\sqrt{\varepsilon}}^{\lfloor\varepsilon^{-1}\rfloor}\rho + \mathcal{O}(\varepsilon^{\frac{1}{2}-\alpha}) \\
&= \frac{1}{2\pi}\int_{\sqrt{\varepsilon}|\xi|\le\nu_0} \varepsilon^\alpha e^{i\xi z}\hat{\psi}(\xi\varepsilon^\alpha)\int_{\mathbb{T}} \mathcal{L}_{\xi\sqrt{\varepsilon}}^{\lfloor\varepsilon^{-1}\rfloor}\rho + \mathcal{O}(\varepsilon^{\frac{1}{2}-\alpha}) \\
&\quad + \mathcal{O}\left(\int_{|\eta|\ge\nu_0\varepsilon^{\alpha-1/2}} |\hat{\psi}(\xi)|\right).
\end{aligned}
$$

Milking the cow

Thus, if we set $\beta = \min\{2\alpha, \frac{1}{2} - \alpha\} > 0$ and $\sigma^2 = -\alpha_0''$, we have

$$\begin{aligned}
\mathbb{E}(\psi_{\varepsilon,z}(\zeta_\varepsilon)) &= \frac{1}{2\pi}\int_{\sqrt{\varepsilon}|\xi|\le\nu_0} \varepsilon^\alpha e^{i\xi z}\hat\psi(\xi\varepsilon^\alpha)e^{-\frac{\sigma^2}{2}\xi^2+\mathcal{O}(\sqrt{\varepsilon}\xi^3)}d\xi + \mathcal{O}(\varepsilon^\beta) \\
&= \frac{\hat\psi(0)}{2\pi}\int_{\mathbb{R}} \varepsilon^\alpha e^{i\xi z}e^{-\frac{\sigma^2}{2}\xi^2}d\xi + \mathcal{O}(\varepsilon^\beta) \\
&= \mathrm{Leb}(\psi_{\varepsilon,z})\frac{e^{-\frac{z^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} + \mathcal{O}(\varepsilon^\beta) \\
&= \int_{\mathbb{R}} \psi_{\varepsilon,z}(x)\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} + \mathcal{O}(\varepsilon^\beta).
\end{aligned}$$

Since $\text{Leb}(\psi_{\varepsilon,z}) = \mathcal{O}(\varepsilon^{\alpha})$ the above formula is useful only if $\beta > \alpha$, thus we can explore the distribution only till intervals of size $\varepsilon^{\frac{1}{4}}$. To have informations on smaller scales one must investigate the operators $\mathcal{L}_{\nu}$ for values of ν beyond the perturbative regime. This is indeed possible, but outside the scopes of the present note.

Noise (linear)

This gives us informations for a single time, about the full process? A computations "similar" to the previous one yields

$$\mathbb{E}([\zeta_\varepsilon(t) - \zeta_\varepsilon(s)]^4) \le C|t-s|^2.$$

Hence, by Kolmogorow criteria, the sequence is tight.

Noise (linear)

In general, with considerable more work, it is possible to prove that the accumulation points ζ of ζ_ε satisfy

$$\begin{aligned} d\zeta &= \bar{\omega}'(\bar{\theta}(t))\zeta(t)dt + \sigma(\bar{\theta}(t))dB \\ \zeta(0) &= 0 \end{aligned}$$

where $\sigma > 0$ is given by an appropriate Green-Kubo formula.
This type of results are much more recent and, in the above form, have been obtained by Dolgopyat (2004).

Noise

We have thus seen that $z_\varepsilon \sim \bar{\theta} + \sqrt{\varepsilon}\zeta$. On the other hand it is possible to show that $\bar{\theta} + \sqrt{\varepsilon}\zeta \sim \tilde{z}_\varepsilon$ where

$$d\tilde{z}_\varepsilon = \bar{\omega}(\tilde{z}_\varepsilon)dt + \sqrt{\varepsilon}\sigma(\tilde{z}_\varepsilon)dB.$$

Thus the motion is described by an ODE with a small random noise of the type introduced by Hasselmann (1976) and extensively studied by Wentzell–Freidlin and Kifer in the 70's-80's.
But what $\sim$ really means? For which times does it hold?

Noise (quantitative)

There exists $\alpha \in (0,1)$ and a coupling $\mathbb{P}_c$ such that, for all $\varepsilon > 0$ and $t \leq \varepsilon^{-\alpha}$, we have
[De Simoi-Liverani +De Simoi-Liverani-Poquet-Volk, (w.i.p.)]

$$\mathbb{P}_c(|z_\varepsilon(t) - \tilde{z}_\varepsilon(t)| \geq \varepsilon)| \leq C\varepsilon^{\alpha}.$$

In other words, up to the scale ε, the stochastic and deterministic process are indistinguishable for a very long time.
But what happens for even longer times ?
Are the random and the deterministic processes really the same?

Invariant measures (random)

Consider the case in which $\bar{\omega}$ has $2N$ non-degenerate zeroes $\{z_i\}$ with $\bar{\omega}'(z_{2i}) < 0$. Then the equation

$$\dot{\bar{\theta}} = \bar{\omega}(\bar{\theta})$$

has $\{\delta_{z_i}\}$ as invariant measures. On the contrary

$$d\tilde{z}_\varepsilon = \bar{\omega}(\tilde{z}_\varepsilon)dt + \sqrt{\varepsilon}\sigma(\tilde{z}_\varepsilon)dB$$

has only one invariant measure that is essentially of the form $\sum_i p_i N_{i,\varepsilon}(z_{2i})$ where $N_{i,\varepsilon}$ is a Gaussian variable centred at z_{2i} and of variance $\sim \sqrt{\varepsilon}$.

Invariant measures (deterministic)

The deterministic system has infinitely many invariant measures, yet the *Physical Measures* must be (essentially) of the form [De Simoi-Liverani, to appear in Inventiones]

$$\nu_p = \sum_i p_i \mu_{z_{2i}} \times N_{i,\varepsilon}(z_{2i}).$$

More precisely, for each initial measure μ as described, we have

$$\inf_p D((F_\varepsilon^n)_*\mu, \nu_p) \le C \max\{\varepsilon^\alpha, e^{-c\frac{\varepsilon}{\ln \varepsilon^{-1}} n}\}$$

where D is the Wasserstein distance.

Yet,

$$D((F_\varepsilon^n)_*\nu_p, \nu_p) \le \varepsilon^\alpha \quad \forall n \le e^{-c\varepsilon^{-1}},$$

Thus we have metastable states.
Similar results hold for the purely stochastic model: the distribution

$$\sum_i p_i(t) N_{i,\varepsilon}(z_{2i})$$

evolves on the time scale $e^{-c\varepsilon^{-1}t}$. The evolution of the p_i are determined by the large deviation [Wentzell-Freidlin 1979].

Not the same

The large deviation functional of the random and the deterministic case are very different.
They agree only in a neighbourhood of the minimum.
In the deterministic case one can have
several invariant physical measures.
If the invariant measure is unique, then it will be typically concentrated on one sink i_1, that is $1 - p_{i_1} \leq e^{-c\varepsilon^{-1}}$, but that may not be the sink where is concentrated the invariant measure of the random model.

But, really.... noise?

In what we have done there is a problem: the results are averages with respect to an initial distribution, but a deterministic system starts form some initial conditions, it does not care that we cannot measure it! This poses the general problem: if you see just one trajectory, how can you talk about randomness?

There will be only one

If you see only one trajectory $B(t)$ from a Brownian motions, then choose a time interval h and different times $\{t_i\}_{i=1}^N$, $t_{i+1} - t_i \geq h$ and study the quantities

$$\frac{1}{N}\sum_{i=1}^{N}\varphi(B(t_i+h)-B(t_i))$$
$$\frac{1}{N}\sum_{i=1}^{N}\varphi(B(t_{i+1}+h)-B(t_{i+1}))g(B(t_i+h)-B(t_i)).$$

the first quantity, for $N \to \infty$, should converge to the average of φ with respect to a Gaussian, and the second should converge to the product of the averages of φ and g, for almost all the trajectories.

Hence we would like to show that, Lebesgue almost surely,

$$\lim_{N\to\infty}\lim_{\varepsilon\to 0}\frac{1}{N}\sum_{i=1}^{N}\exp\left[i\xi(\zeta_\varepsilon(t_i+h)-\zeta_\varepsilon(t_i))\right]=\lim_{\varepsilon\to 0}\mathbb{E}(e^{i\xi(\zeta_\varepsilon(h)-\zeta_\varepsilon(0))})$$
$$=\exp\left[-\frac{\xi^2\sigma^2 h}{2}\right].$$

How to prove it? Compute

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N}\exp\left[i\xi(\zeta_\varepsilon(t_i+h)-\zeta_\varepsilon(t_i))\right]-\exp\left[-\frac{\xi^2\sigma^2 h}{2}\right]\right|^2\right]$$