Deterministic noise (beyond averaging and fluctuations)

Liverani Carlangelo
Università di Roma *Tor Vergata*

Supported by ERC *MALADY*
Work in Collaboration with Jacopo De Simoi

Lyon, 10 July 2015

Noise, what is it?

It is commonplace that any real dynamical system is subject to noise, which can be reduced but not eliminated.

Often the noise is *phenomenologically* described a by a small diffusion. A simple example is

$$dx = V(x)dt + \varepsilon dB$$

or, looking at the the evolutions of densities,

$$\partial_t u = -\operatorname{div}(Vu) + \varepsilon \Delta u.$$

But from where does the noise come from?

Interaction

The usual answer is that any system interacts with the rest of the world: **isolated systems are only an idealisation**

For example consider the simple case

$$\dot{x} = V(x)$$

A weak interaction between x with a free field (infinitely many degrees of freedom) may lead to a stochastic differential equation of the type

$$dx = V(x)dt + \varepsilon\sigma dB$$

where dB is standard brownian motion.

The hard truth

This is rather superficial since, in reality, the external degrees of freedom are neither infinitely many, nor have a simple dynamics, nor the dynamics is independent on x.
Most of all, the very concept of "isolated systems" rests on shaky ground.

To better understand let us consider an **interaction with degrees of freedom having a complex fully coupled dynamics**.

It turns out that such a study is far from trivial, hence we will consider the absolutely simplest possible situation.

Simplifying life

To this end, we will make the following assumptions

1. We consider the case of discrete, rather than continuous time.
2. We consider as few degrees of freedom as possible: one for the "system", one for the "exterior".
3. We realise the "complex dynamics" of the external degree of freedom via the simplest possible example of "chaotic" map.
4. We assume that the dynamics takes place in a compact space.
5. We assume a strong time scale separation.

A "super simple" model

This brings us to the one parameter family of maps $F_\varepsilon \in \mathcal{C}^4(\mathbb{T}^2, \mathbb{T}^2)$

$$F_\varepsilon(x,\theta) = (f(x,\theta), \theta + \varepsilon\omega(x,\theta)),$$

and the dynamics $(x_n, \theta_n) = F_\varepsilon^n(x_0, \theta_0)$ with initial conditions

$$\mathbb{E}(g(x_0,\theta_0)) = \int_{\mathbb{T}^1} \rho(x) g(x,\bar{\theta}) dx \quad \bar{\theta} \in \mathbb{T}^1.$$

1. $\partial_x f(x,\theta) \geq \lambda > 1$ (expanding map)
2. F_0 has θ as a *conserved quantity*
3. $\rho \in \mathcal{C}^2(\mathbb{T}^1, \mathbb{R}_+)$

Averaging

Note that we have $\theta_n = \theta_0 + \varepsilon \sum_{k=0}^{n-1} \omega(x_k, \theta_k)$, thus

$$\theta_n - \theta_m = \mathcal{O}(\varepsilon(n-m)).$$

It is then natural to introduce the macroscopic time $t = \varepsilon n$, this is the time in which the variable θ may have a change of order one. It is also covenient to introduce the continuous paths

$$\Theta_\varepsilon(t) = \theta_{\lfloor \varepsilon^{-1} t \rfloor} + (\varepsilon^{-1} t - \lfloor \varepsilon^{-1} t \rfloor)(\theta_{\lfloor \varepsilon^{-1} t \rfloor + 1} - \theta_{\lfloor \varepsilon^{-1} t \rfloor}), \quad t \in [0, T].$$

Since the Θ_ε are uniformly Lipschitz, they belong to a compact set in $\mathcal{C}^0([0,T],\mathbb{R})$, hence they have convergent subsequences. It is possible to show that there exists $\bar{\omega} \in \mathcal{C}^3$ such that all the accumulation points $\overline{\Theta}$ must satisfy the ODE

$$\dot{\overline{\Theta}} = \bar{\omega}(\overline{\Theta})$$
$$\overline{\Theta}(0) = \bar{\theta}.$$

We have thus an isolated (autonomous) dynamics which emerges, in first approximation, from the complex interaction with the external degrees of freedom.
This type of results goes back, at least, to Anosov (1960) and Bogolyubov-Mitropolskii (1961).

Basic intuition

If it were $\partial_\theta f = \partial_\theta \omega = 0$, then

$$\theta_n = \theta_0 + \varepsilon n \left[\frac{1}{n} \sum_{k=0}^{n-1} \omega(f^k(x_0)) \right].$$

Note that:

1. the quantity in the square brackets is an ergodic average
2. an expanding map of the circle has a unique invariant measure μ absolutely continuous w.r.t. Lebesgue (let h be the density)
3. the dynamical system $(\mathbb{T}^1, f, \mu)$ is exponentially mixing for Hölder observables

A computation

Thus, setting $\bar{\omega} = \mu(\omega)$ and $\hat{\omega} = \omega - \bar{\omega}$,

$$\mathbb{E}(|\theta_n - \theta_0 - \varepsilon n\bar{\omega}|^2) = \varepsilon^2 \sum_{k,j=0}^{n-1} \mathbb{E}(\hat{\omega}(x_k)\hat{\omega}(x_j))$$

$$\leq \varepsilon^2 C \sum_{k,j=0}^{n-1} e^{-c|k-j|} \leq C\varepsilon[\varepsilon n].$$

With some more work it is possible to show that the same holds in the general case $\partial_\theta f \neq 0$, $\partial_\theta \omega \neq 0$ where now

$$\bar{\omega}(\theta) = \mu_\theta(\omega(\cdot, \theta))$$

and μ_θ is the unique a.c.i.m. of $f(\cdot, \theta)$.

Noise (linear)

But what happens for $\varepsilon > 0$ and/or for times longer than ε^{-1}?
Let us consider the quantity $\zeta_\varepsilon(t) = \varepsilon^{-\frac{1}{2}}\left[\Theta_\varepsilon(t) - \overline{\Theta}(t)\right]$. This are the fluctuations around the average. A computations "similar" to the previous one yields

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

Hence, by Kolmogorow criteria, the sequence is tight.

Under some mild technical condition, with considerable more work, it is possible to prove that the accumulation points ζ of ζ_ε satisfy

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

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 (non-linear)

We have thus seen that $\Theta_\varepsilon \sim \overline{\Theta} + \sqrt{\varepsilon}\zeta$. On the other hand it is possible to show that $\overline{\Theta} + \sqrt{\varepsilon}\zeta \sim \widetilde{\Theta}_\varepsilon$ where

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

We have, again, a system with 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?

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

$$\mathbb{P}_c(|\Theta_\varepsilon(t) - \widetilde{\Theta}_\varepsilon(t)| \geq \varepsilon)| \leq C\varepsilon^\alpha.$$

In other words, if we do measurements up to the scale ε the stochastic and deterministic process are indistinguishable for a very long time.

Larger fluctuations

The above essentially explores fluctuations of order up to $\sqrt{\varepsilon}$, the range of validity of CLT (in fact the result is obtained by proving a Local Central Limit Theorem). One may wonder if the observation of larger fluctuation can allow to differentiate between a true noise and a noise of dynamical origin.

This turns out not to be the case for moderate deviations: let $\beta \in (0, 1/2)$ and consider a path $\bar{x} \in \mathcal{C}^0([0, T], \mathbb{R})$ such that $\|\bar{x} - \overline{\Theta}\|_{\mathcal{C}^0} = \varepsilon^\beta$ and the event

$$Q_{\varepsilon,\delta} = \{z(t) \in \mathcal{C}^0 \ : \ |z(t) - \bar{x}(t)| \le \delta\varepsilon^\beta\}$$

Then (De Simoi-Liverani 2014)

$$\mathbb{P}(Q_{\varepsilon,\delta}) \sim e^{-\frac{1}{2}\varepsilon^{-1}\inf_{z\in Q_{\varepsilon,\delta}}\int_0^T \sigma(\overline{\Theta}(s))^{-1}[z'(s)-\bar{\omega}(\overline{\Theta}(s))]^2}$$

which is exactly the probability in the purely stochastic model. Not so for large deviations. Indeed, while the rate function for the random model is always as above, the rate function for the deterministic system is rather different (Kifer 1998). In particular, certain trajectories might have zero probability, while in the random case all regular trajectories are possible. Let us be more precise.

Invariant measures (random)

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

$$\dot{\overline{\Theta}} = \bar{\omega}(\overline{\Theta})$$

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

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

has only one invariant measure that is essentially of the form $\sum_i p_i N_{i,\varepsilon}(\theta_{2i})$ where $N_{i,\varepsilon}$ is a Gaussian variable centred at θ_{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 2014)

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

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.

Metastability

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 (Wentzell-Freidlin 1979): the distribution

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

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

Not the same

What is drastically different between the random and the deterministic case are the large deviation probabilities.

In the deterministic case one can have several invariant 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.

Lyapunov exponents

Another relevant different between the random and the deterministic model emerges when one considers the evolution of two nearby initial conditions.
A way to do so is to compute the Lyapunov exponents.
One can show that for the random model the Lyapunov exponents are always negative.
Not so for the deterministic model!

For example (De Simoi-Liverani-Volk w.i.p.)

$$F_\varepsilon(x,\theta) = (20x+\sin(2\pi\theta)\left[\sin(2\pi x)+\sin(4\pi x)\right], \theta+\varepsilon\cos(2\pi x)) \mod 1.$$

has a unique physical (and SRB) measure ν and both Lyapunov exponents are positive ν-a.s..
Thus the random model fails miserably in describing the full interplay between the system and the noise.

Conclusions

It seems that the concept of deterministic noise has some reasonable meaning.
Gaussian noise provides an accurate phenomenological description in the small fluctuation regime, but fails dramatically for long times or when measuring correlations between different trajectories.

It remains thus open problem of providing a better phenomenological description of deterministic noise.