The martingale approach after Varadhan and Dolgopyat

Carlangelo Liverani

University of Rome *Tor Vergata*

Paris, 16 September 2015

downloadable at

http://www.mat.uniroma2.it/∼liverani/sem.html

Let $F_\varepsilon \in \mathcal{C}^4(\mathbb{T}^2, \mathbb{T}^2)$ be defined by

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

This is a family of fast-slow systems (in discrete time, for simplicity). The connection with Freidlin-Wentzell theory it is not immediately apparent but it will became clear as we proceed.

Since we want the fast dynamics to be $chaotic$ we assume

$$\inf_{x,\theta} \partial_x f(x, \theta) = \lambda > 1.$$

We are interested in the trajectories of such a dynamical system: let $(x_n, \theta_n) = F_\varepsilon^n(x_0, \theta_0)$.

As discussed before we will consider random initial conditions defined by the measure μ_0 on $\mathbb{T}^2$ by

$$\mathbb{E}_{\mu_0}(\varphi(x_0, \theta_0)) = \int_{\mathbb{T}^1} \varphi(x, \theta^*)\rho(x)dx$$

for some probability distribution $\rho \in \mathcal{C}^3(\mathbb{T}^1, \mathbb{R}_\geq)$ and $\theta^* \in \mathbb{T}^1$.

Define the polygonalization (since we interpolate between close points the procedure is uniquely defined in $\mathbb{T}$.)

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

with $t \in [0, T]$.

Note that Θ_ε is a random variable on $\mathbb{T}^2$ with values in $\mathcal{C}^0([0,T], \mathbb{T})$. Let $\mathbb{P}^\varepsilon$ be the law of Θ_ε as a random variable in $\mathcal{C}^0([0,T], \mathbb{T})$.

Here is the first result we want to establish in our journey:

Theorem 1 (Anosov, Kifer)**.** *The measures* $\{\mathbb{P}^\varepsilon\}$ *have a weak limit* $\mathbb{P}$*. Moreover,* $\mathbb{P}$ *is supported on the trajectory determined by the O.D.E.*

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

$$\overline{\Theta}(0) = \theta_0$$

where $\bar{\omega}(\theta) = \int_M \omega(x,\theta)\rho_\theta(x)dx$*.*

(Actually, working more it is possible to prove almost sure convergence, but we will not worry about this)

The next natural question concerns the speed of convergence. For each $t \in [0, T]$, let

$$\zeta_\varepsilon(t) = \varepsilon^{-\frac{1}{2}} \left[\Theta_\varepsilon(t) - \overline{\Theta}(t) \right].$$

Note that ζ_ε is a random variable on $\mathbb{T}^2$ with values in $\mathcal{C}^0([0, T], \mathbb{R})$, $\mathbb{R}$ being the universal cover of $\mathbb{T}$, which describes the fluctuations around the average.

Let $\widetilde{\mathbb{P}}^\varepsilon$ be the path measure describing ζ_ε when (x_0, θ_0) are distributed according to the measure μ_0. That is, $\widetilde{\mathbb{P}}^\varepsilon = (\zeta_\varepsilon)_* \mu_0$.

Theorem 2 (Kifer, Dolgopyat). *The measures $\{\widetilde{\mathbb{P}}^{\varepsilon}\}$ have a weak limit $\widetilde{\mathbb{P}}$. Moreover, $\widetilde{\mathbb{P}}$ is the zero average Gaussian process defined by the SDE*

$$d\zeta = \bar{\omega}'(\overline{\Theta})\zeta dt + \sigma(\overline{\Theta})dB$$
$$\zeta(0) = 0,$$

where B is the standard Brownian motion and the diffusion coefficient σ is given by

$$\sigma(\theta)^2 = m_\theta\left(\hat{\omega}(\cdot,\theta)^2\right) + 2\sum_{m=1}^{\infty} m_\theta\left(\hat{\omega}(f_\theta^m(\cdot),\theta)\hat{\omega}(\cdot,\theta)\right).$$

where $\hat{\omega} = \omega - \bar{\omega}$ and we have used the notation $f_\theta(x) = f(x,\theta)$ and $\frac{dm_\theta}{dx} = \rho_\theta$. In addition, $\sigma(\theta)$ is strictly positive, unless $\hat{\omega}(\theta,\cdot)$ is a coboundary for f_θ.

We can now make the connection with Freidlin-Wentzell theory: consider the SDE

$$\begin{aligned} dz &= \bar{\omega}(z)dt + \sqrt{\varepsilon}\sigma(z)dB \\ z(0) &= \theta^* \end{aligned}$$

One can show that (Freidlin-Wentzell, more details later)

$$z(t) = \overline{\Theta}(t) + \sqrt{\varepsilon}\zeta(t) + \mathcal{O}_{L^2}(\varepsilon).$$

Our model is thus a natural candidate to develop a Freidlin-Wentzell theory in a completely deterministic setting.

We have thus the right setting to investigate the questions posed at the beginning. In general the answers are not known, yet for the simple situation we are considering some preliminary result do exist.

Consider the (classical Freidlin-Wentzell) case in which $\bar{\omega}$ has finitely many non degenerate zeroes. Let n_Z be the number of zeroes and assume $n_Z > 0$. So the differential equation for $\overline{\Theta}$ is of a very simple form and has $2n_Z$ point masses as unique ergodic measures.

Also we assume

1) that the map is not too distant from a cocycle: for some appropriate, fixed, $\delta_0 > 0$ we have

$$\|\partial_\theta f\|_\infty \leq \delta_0.$$

2) for any $\theta_1, \theta_2 \in \mathbb{T}$ and any neighbourhoods U_i of such points the exists an admissible path connecting such two neighbourhoods.

An admissible path is defined as follows:

for each $\theta \in \mathbb{T}^1$, define the convex set $\Omega(\theta) = \{\mu(\omega(\cdot, \theta)) | \mu$ is a f_θ-invariant probability$\}$. A path $h \in \mathcal{C}^1([0, T], \mathbb{T}^1)$ is said to be *admissible* if it satisfies the differential inclusion $h'(s) \in \operatorname{int} \Omega(h(s))$.

Theorem 3 (De Simoi, L.). *Given the previous hypotheses, there exists* $\varepsilon_0 > 0$ *such that, for all* $\varepsilon \in (0, \varepsilon_0)$*, the system has a unique* physical measure μ_ε*. Moreover, for each* $A, B \in \mathcal{C}^1$*,*

$$|\mathrm{Leb}(A \cdot B \circ F_\varepsilon^n) - \mathrm{Leb}(A)\mu_\varepsilon(B)| \leq C_1 \|A\|_{\mathcal{C}^1} \|B\|_{\mathcal{C}^1} \exp(-c_\varepsilon n),$$

where the rate of decay of correlations is given by:

$$c_\varepsilon = \begin{cases} C_2 \varepsilon / \log \varepsilon^{-1} & \textit{if } n_Z = 1, \\ C_3 e^{-C_4 \varepsilon^{-1}} & \textit{otherwise.} \end{cases}$$

In the case of a $\bar{\omega}$ with several zeroes, one has a Freidlin–Wentzell theory like situation (but now for a completely deterministic system), hence metastable states and a much longer time ($\sim e^{-ne^{-c/\varepsilon}}$) of convergence to the unique equilibrium state. A hint of such a theory can be found in Kifer's work, but only for exponentially long times.

So far for general consideration and ideas: let us go into some details.

Let us start with Theorem 1.

Let us start with a trivial exercise: let z_ε be the solution of

$$dz = \bar{\omega}(z)dz + \sqrt{\varepsilon}\sigma(z)dB$$
$$z(0) = \theta^*$$

$$\mathbb{E}(|z_\varepsilon(t) - \overline{\Theta}(t)|^2) \leq 2\mathbb{E}\left(\left|\int_0^t [\bar{\omega}(z_\varepsilon(s)) - \bar{\omega}(\overline{\Theta}(s))]ds\right|^2\right) + 2\varepsilon\mathbb{E}\left(\left|\int_0^t [\sigma(z_\varepsilon(s))dB(s)\right|^2\right)$$

$$\begin{aligned}\mathbb{E}(|z_\varepsilon(t) - \overline{\Theta}(t)|^2) \leq & C\mathbb{E}\left(\left|\int_0^t [z_\varepsilon(s) - \overline{\Theta}(s)]ds\right|^2\right) \\ & + C\varepsilon\mathbb{E}\left(\int_0^t \sigma(z_\varepsilon(s))^2 ds\right) \\ & \leq C\int_0^t \mathbb{E}\left([z_\varepsilon(s) - \overline{\Theta}(s)]^2\right) + C\varepsilon\end{aligned}$$

Then, for each $t \leq T$, we have (by Gronwall inequality)

$$\|z_\varepsilon - \overline{\Theta}\|_{L^2} = \mathbb{E}(|z_\varepsilon(t) - \overline{\Theta}(t)|^2)^{\frac{1}{2}} \leq Ce^{CT}\sqrt{\varepsilon}.$$

We have used

1. knowledge of the limit;
2. Ito's formula (hence martingales and conditioning).

The first was trivial for us but in general may be difficult. The second is not available nor seems to make any sense for deterministic systems .... so?

Let us try a different approach: First of all let $\mathbb{P}_\varepsilon$ be the measure on $\mathcal{C}^0([0,T],\mathbb{R})$ induced by the process z_ε. It is easy to verify that the sequence of measures $\{\mathbb{P}_\varepsilon\}$ is tight (an application of Ito's formula and Kolmogorov criteria), thus it has accumulation points. If we can show that the accumulation points must satisfy some equation that has a unique solution, than we have proven that the limit (in the sense of weak convergence) exists.

Next, notice that, for each $A \in \mathcal{C}^2$ and $\varepsilon > 0$

$$\begin{aligned} M_{\varepsilon,t} = & A(z(t)) - f(\theta^*) - \int_0^t \bar{\omega}(z(s))A'(z(s))ds \\ & - \frac{\varepsilon}{2}\int_0^t \sigma(z(s))^2 A''(z(s))ds \end{aligned}$$

is a martingale under $\mathbb{P}_\varepsilon$ with respect to the filtration $\mathcal{F}_t$ generated by $\{z(s)\}_{s<t}$ (it follows from Ito's formula). Moreover $\mathbb{E}_\varepsilon(M_{\varepsilon,0}) = 0$.

In fact, $\mathbb{P}_\varepsilon$ is uniquely determined by the above properties (uniqueness of the Martingale problem)

Let $\mathbb{P}_{\varepsilon_j}$ be a converging sequence, and let $\mathbb{P}$ be the limit. Then, setting

$$N_t = A(z(t)) - A(\theta^*) - \int_0^t \bar{\omega}(z(s))A'(z(s))ds$$

we have

$$\lim_{j\to\infty} \mathbb{E}_{\varepsilon_j}(M_{\varepsilon_j,t} \mid \mathcal{F}_s) = \mathbb{E}\left(N_t \mid \mathcal{F}_s\right)$$
$$\lim_{j\to\infty} \mathbb{E}_{\varepsilon_j}(M_{\varepsilon_j,t} \mid \mathcal{F}_s) = \lim_{j\to\infty} M_{\varepsilon_j,s} = N_s$$

Thus N_t is a martingale under $\mathbb{P}$ and $\mathbb{E}(A(z(t))) = A(\theta^*)$.

If we can prove that $\mathbb{P}$ is uniquely determined by the above properties (which is actually true and not hard to prove), then we are done.

This time we have used

1. tightness of the sequence of probability measures

2. the fact that each limiting measure satisfies a Martingale Problem that admits a unique solution

This seems much more flexible than before (in particular there is no need to have Martiangales before the limit, only after they play a fundamental role) although the result is weaker. Our goal is to see how such a strategy can be implemented in the case of deterministic systems.

Conditioning and right class of measures

The basic tool that takes the pace of conditioning is standard pair and, in its final version, is due to Dolgopyat.

A standard pair is a couple $\ell = (G, \rho)$. $G \in \mathcal{C}^r([a,b], \mathbb{T})$, $|b-a| \in [\delta, 2\delta]$, $|G'| \leq c\varepsilon$ and $\rho \in \mathcal{C}^r([a,b], \mathbb{R}_{\geq})$, $|\frac{d}{dx} \ln \rho(x)| \leq c$. To each standard pair is associated the measure

$$\mu_\ell(\phi) = \int_a^b \rho(x)\phi(x, G(x))dx.$$

Also we require that μ_ℓ is a probability measure.

The Key property of standard pairs is their invariance:

Given a standard pail ℓ consider the push forward $(F_\varepsilon)_*\mu_\ell$. Then there exists a finite collection of indices $\mathcal{A}$, a probability measure ν on $\mathcal{A}$ and standard pairs $\{\ell_\alpha\}_{\alpha\in\mathcal{A}}$ such that

$$(F_\varepsilon)_*\mu_\ell = \sum_{\alpha\in\mathcal{A}} \nu(\{\alpha\})\mu_{\ell_\alpha}.$$

This allows to condition with respect to the past and hence treat a deterministic Dynamical Systems in analogy to a random Markov process.