Large deviations in Averaging

Yuri Kifer

Hebrew University of Jerusalem

Centre Bernoulli, EPFL, March 2013

Averaging setup: continuous time

We consider a system of differential equations for $X^\varepsilon = X^\varepsilon_{x,y}$ and $Y^\varepsilon = Y^\varepsilon_{x,y}$,

$$\frac{dX^\varepsilon(t)}{dt} = \varepsilon B(X^\varepsilon(t), Y^\varepsilon(t)), \quad \frac{dY^\varepsilon(t)}{dt} = b(X^\varepsilon(t), Y^\varepsilon(t))$$

with initial conditions $X^\varepsilon(0) = x$, $Y^\varepsilon(0) = y$ on the product $\mathbb{R}^d \times \mathbf{M}$ where $\mathbf{M}$ is a compact $n_{\mathbf{M}}$-dimensional C^2 Riemannian manifold and $B(x,y)$, $b(x,y)$ are smooth in x, y families of bounded vector fields on $\mathbb{R}^d$ and on $\mathbf{M}$, respectively, so that y serves as a parameter for B and x for b. The solutions of the above equations determine the flow of diffeomorphisms Φ^t_ε on $\mathbb{R}^d \times \mathbf{M}$ acting by $\Phi^t_\varepsilon(x,y) = (X^\varepsilon_{x,y}(t), Y^\varepsilon_{x,y}(t))$. Taking $\varepsilon = 0$ we arrive at the flow $\Phi^t = \Phi^t_0$ acting by $\Phi^t(x,y) = (x, F^t_x y)$ where F^t_x is another family of flows given by $F^t_x y = Y_{x,y}(t)$ with $Y = Y_{x,y} = Y^0_{x,y}$ being the solution of

$$\frac{dY(t)}{dt} = b(x, Y(t)), \quad Y(0) = y.$$

It is natural to view the flow Φ^t as describing an idealized physical system where parameters $x = (x_1, ..., x_d)$ are assumed to be constants (integrals) of motion while the perturbed flow Φ^t_ε is regarded as describing a real system where evolution of these parameters is also taken into consideration.

Averaging setup: discrete time

In the discrete time case we deal with difference equations for sequences $X^\varepsilon(n) = X^\varepsilon_{x,y}(n)$ and $Y^\varepsilon(n) = Y^\varepsilon_{x,y}(n)$, $n = 0, 1, ...$ so that

$$X^\varepsilon(n+1) - X^\varepsilon(n) = \varepsilon B(X^\varepsilon(n), Y^\varepsilon(n)),$$
$$Y^\varepsilon(n+1) = F_{X^\varepsilon(n)} Y^\varepsilon(n),\ \ X^\varepsilon(0) = x,\, Y^\varepsilon(0) = y$$

where $B : \mathcal{X} \times \mathbf{M} \to \mathbb{R}^d$ is Lipschitz in both variables and the maps $F_x : \mathbf{M} \to \mathbf{M}$ are smooth and depend smoothly on the parameter $x \in \mathbb{R}^d$. Introducing the map

$$\Phi_\varepsilon(x, y) = (X^\varepsilon_{x,y}(1), Y^\varepsilon_{x,y}(1)) = (x + \varepsilon B(x, y), F_x y).$$

This setup is viewed as a perturbation of the map $\Phi(x, y) = (x, F_x y)$ describing an ideal system where parameters $x \in \mathbb{R}^d$ do not change. Assuming that F_x, $x \in \mathbb{R}^d$ are C^2 depending on x families of either C^2 expanding transformations or C^2 Axiom A diffeomorphisms in a neighborhood of an attractor Λ_x we will derive large deviations estimates for the difference $X^\varepsilon_{x,y}(n) - \bar{X}^\varepsilon_x(n)$ where $\bar{X}^\varepsilon = \bar{X}^\varepsilon_x$ solves the equation

$$\frac{d\bar{X}^\varepsilon(t)}{dt} = \varepsilon \bar{B}(\bar{X}^\varepsilon(t)),\ \ \bar{X}^\varepsilon(0) = x$$

where $\bar{B}(x) = \int B(x, y) d\mu^{\text{SRB}}_x(y)$ and μ^{SRB}_x is the corresponding SRB invariant measure of F_x on Λ_x. The discrete time results are obtained, essentially, by simplifications of the corresponding arguments in the continuous time case.

Assume that the limit

$$\bar{B}(x)=\bar{B}_y(x)=\lim_{T\to\infty} T^{-1}\int_0^T B(x,F_x^t y)dt$$

exists and it is the same for "many" y's. For instance, suppose that μ_x is an ergodic invariant measure of the flow F_x^t then this limit exists for μ_x-almost all y and is equal to

$$\bar{B}(x)=\bar{B}_{\mu_x}(x)=\int B(x,y)d\mu_x(y).$$

If $b(x,y)$ does not, in fact, depend on x then $F_x^t=F^t$ and $\mu_x=\mu$ are also independent of x and we arrive at the classical uncoupled setup. In this case Lipschitz continuity of B implies already that $\bar{B}(x)$ is also Lipshitz continuous in x, and so there exists a unique solution $\bar{X}=\bar{X}_x$ of the averaged equation

$$\frac{d\bar{X}^\varepsilon(t)}{dt}=\varepsilon\bar{B}(\bar{X}^\varepsilon(t)),\quad \bar{X}^\varepsilon(0)=x.$$

In this case the standard averaging principle says that for μ-almost all y,

$$\lim_{\varepsilon\to 0}\sup_{0\le t\le T/\varepsilon}|X^\varepsilon_{x,y}(t)-\bar{X}^\varepsilon_x(t)|=0.$$

In the fully coupled case the averaging principle in the form above usually does not hold true both for nearly integrable Hamiltonian systems in the presence of resonances and for hyperbolic systems (Anosov flows depending on parameters).

Nonconvergence examples in the fully coupled case

1st example could be viewed as a perturbation of circle rotations but describes also the motion of a pendulum with a small friction and it has the form

$$\frac{dX^\varepsilon(t)}{dt} = \varepsilon\big(4 + 8\sin Y^\varepsilon(t) - X^\varepsilon(t)\big), \quad \frac{dY^\varepsilon(t)}{dt} = X^\varepsilon(t)$$

with the corresponding averaged equation $\frac{d\bar{X}^\varepsilon(t)}{dt} = \varepsilon(4 - \bar{X}^\varepsilon(t))$. Due to the "capture into resonance" if $X^\varepsilon(0) = \bar{X}^\varepsilon(0) = x \in (-2, -1)$ then

$$\limsup_{\varepsilon\to 0} |X^\varepsilon(1/\varepsilon) - \bar{X}^\varepsilon(1/\varepsilon)| > 3/2.$$

2nd example is determined by the system of difference equations for sequences $X^\varepsilon(n) = X^\varepsilon_{x,y}(n) \in \mathbb{R}$ and $Y^\varepsilon(n) = Y^\varepsilon_{x,y}(n) \in \mathbb{R}$, $n = 0, 1, ...$ such that

$$X^\varepsilon(n+1) - X^\varepsilon(n) = \varepsilon \sin(Y^\varepsilon(n)), \; X^\varepsilon(0) = x$$

$$Y^\varepsilon(n+1) = 2Y^\varepsilon(n) + X^\varepsilon(n) + c\xi_n \; (mod\; 2\pi), \; Y^\varepsilon(0) = y$$

where $\{\xi_n\}_{n=0}^\infty$ is an arbitrary sequence with $\sup_n |\xi_n| \le 1$ and $c \ge 0$ is a small number. Here $F_x y = F_{x,n} y = 2y + x + c\xi_n \,(mod\, 2\pi)$ and all $F_{x,n}$ act on the circle $\mathbb{T}^1$ (of the length 2π) preserving the Lebesgue measure there and since $\int_0^{2\pi} \sin y dy = 0$ we obtain that the averaged motion stays forever at the initial point. Then for any $x \in \mathbb{R}$ and $\xi = \{\xi_n\}_{n=0}^\infty$ as above $\exists\, \Gamma_{x,\xi}$ with full Lebesgue measure on the circle $\mathbb{T}^1$ such that for each $y \in \Gamma_{x,\xi}$,

$$\limsup_{\varepsilon\to 0} \max_{0\le n\le 1/\varepsilon} |X^\varepsilon_{x,y}(n) - x| \ge \delta > 0.$$

Averaging principle in the fully coupled case

Set $\mathcal{X}_t = \{x \in \mathcal{X} : X^\varepsilon_{x,y}(s) \in \mathcal{X},\ \bar{X}^\varepsilon_x(s) \in \mathcal{X}$ for all $y \in \mathbf{M}$ and $s \in [0, t/\varepsilon]\}$ and

$$E_\varepsilon(t,\delta) = \{(x,y) \in \mathcal{X}_t \times \mathbf{M} : |\frac{1}{t}\int_0^t B(x, Y^\varepsilon_{x,y}(u))du - \bar{B}(x)| > \delta\}.$$

Theorem

Suppose that vector fields b and B are Lipschitz continuous and bounded and that $\bar{B}(x) = \int B(x,y)d\mu_x(y)$ is Lipschitz, as well. Then

$$\lim_{\varepsilon\to 0}\int_{\mathcal{X}_T}\int_{\mathbf{M}} \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)|d\mu(x,y) = 0, \quad d\mu(x,y) = d\mu_x(y)d\nu(x)$$

if and only if there exists an integer valued function $n = n(\varepsilon) \to \infty$ as $\varepsilon \to 0$ such that for any $\delta > 0$,

$$\lim_{\varepsilon\to 0}\max_{0\le j<n(\varepsilon)} \mu\{(\mathcal{X}_T \times \mathbf{M}) \cap \Phi_\varepsilon^{-jt(\varepsilon)} E_\varepsilon(t(\varepsilon),\delta)\} = 0,$$

where $t(\varepsilon) = \frac{T}{\varepsilon n(\varepsilon)}$ and, recall, $\Phi^t_\varepsilon(x,y) = (X^\varepsilon_{x,y}(t), Y^\varepsilon_{x,y}(t))$.

Conditions hold true if μ ~Lebesgue on a compact and either each μ_x is a F_x-invariant and equivalent to the volume on $\mathbf{M}$ (Anosov theorem) or when F^t_x-is a C^2-dependent on x family of Anosov flows (or Axiom A) and $\mu_x = \mu_x^{SRB}$.

Assumption

The family $b(x,\cdot)$ consists of C^2 vector fields on a compact $n_{\mathbf{M}}$-dimensional Riemannian manifold $\mathbf{M}$ with uniform C^2 dependence on the parameter x belonging to a neighborhood of the closure $\bar{\mathcal{X}}$ of a relatively compact open connected set $\mathcal{X} \subset \mathbb{R}^d$. Each flow F_x^t, $x \in \bar{\mathcal{X}}$ on $\mathbf{M}$ given by

$$\frac{dF_x^t y}{dt} = b(x, F_x^t y), \quad F_x^0 y = y$$

possesses a basic hyperbolic attractor Λ_x (topologically transitive hyperbolic attractor with periodic orbits dense there) with a splitting $T_{\Lambda_x}\mathbf{M} = \Gamma_x^s \oplus \Gamma_x^0 \oplus \Gamma_x^u$ satisfying hyperbolicity assumptions with the same exponent and there exists an open set $\mathcal{W} \subset \mathbf{M}$ and $t_0 > 0$ such that

$$\Lambda_x \subset \mathcal{W},\ F_x^t \bar{\mathcal{W}} \subset \mathcal{W}\, \forall t \geq t_0,\ \text{and}\ \cap_{t>0} F_x^t \mathcal{W} = \Lambda_x\ \forall x \in \bar{\mathcal{X}}.$$

Let $J_x^u(t,y)$ be the absolute value of the Jacobian of the linear map $DF_x^t(y) : \Gamma_x^u(y) \to \Gamma_x^u(F_x^t y)$ with respect to the Riemannian inner products and set

$$\varphi_x^u(y) = -\frac{dJ_x^u(t,y)}{dt}\Big|_{t=0}$$

which is a Hölder continuous function.

Denote by $\mathcal{M}_x$ the space of F_x^t-invariant probability measures on Λ_x then we have the variational principle for the topological pressure $\Pi_x(\psi) = \sup_{\mu\in\mathcal{M}_x}(\int\psi d\mu + h_\mu(F_x^1))$. If $\Pi_x(\varphi_x^u + q) = \int(\varphi_x^u + q)d\mu_x^q + h_{\mu_x^q}(F_x^1)$ then a F_x^t-invariant measure μ_x^q on Λ_x is called the equilibrium state for $\varphi_x^u + q$ while $\mu_x^0 = \mu_x^{\mathrm{SRB}}$ is called the Sinai–Ruelle–Bowen (SRB) measure. Since Λ_x are attractors $P_x(\varphi_x^u) = 0$. For any probability measure ν on $\bar{\mathcal{W}}$ set $I_x(\nu) = -\int\varphi_x^u d\nu - h_\nu(F_x^1)$ if $\nu\in\mathcal{M}_x$ and $=\infty$, otherwise. Then $\Pi_x(\varphi_x^u + q) = \sup_\nu(\int q d\nu - I_x(\nu))$. The functional $I_x(\nu)$ is lower semi-continuous in ν and it is also convex (and affine on $\mathcal{M}_x$), and so by the duality theorem $I_x(\nu) = \sup_{q\in\mathcal{C}(\mathbf{M})}(\int q d\nu - \Pi_x(\varphi_x^u + q))$. Set

$$L(x, x', \alpha) = \inf\{I_x(\nu) : \int B(x', y)d\nu(y) = \alpha\}$$

if $\exists\,\nu\in\mathcal{M}_x$ satisfying the condition in brackets and $L(x, x', \alpha) = \infty$, otherwise. Let C_{0T} be the space of continuous curves $\gamma = \gamma_t$, $t\in[0, T]$ in $\mathcal{X}$. For each absolutely continuous $\gamma\in C_{0T}$ denote by $\dot{\gamma}_t$ its velocity. Now set

$$S_{0T}(\gamma) = \int_0^T L(\gamma_t, \dot{\gamma}_t)dt = \int_0^T \inf\{I_{\gamma_t}(\nu) : \dot{\gamma}_t = \bar{B}_\nu(\gamma_t),\ \nu\in\mathcal{M}_{\gamma_t}\}dt,$$

where $\bar{B}_\nu(x) = \int B(x, y)d\nu(y)$, provided for Lebesgue almost all $t\in[0, T]$ there exists $\nu_t\in\mathcal{M}_{\gamma_t}$ for which $\dot{\gamma}_t = \bar{B}_{\nu_t}(\gamma_t)$, and $S_{0T}(\gamma) = \infty$ otherwise.

Let γ_t^u be the unique solution of the equation $\dot{\gamma}_t^u = \bar{B}(\gamma_t^u), \quad \gamma_0^u = x$ where $\bar{B}(z) = \bar{B}_{\mu_z^{\rm SRB}}(z)$. Define the uniform metric on C_{0T} by $\mathbf{r}_{0T}(\gamma, \eta) = \sup_{0 \le t \le T} |\gamma_t - \eta_t|$ for any $\gamma, \eta \in C_{0T}$. Set $\Psi_{0T}^a(x) = \{\gamma \in C_{0T} : \gamma_0 = x,\ S_{0T}(\gamma) \le a\}$. Then S_{0T} is a lower semi-continuous functional on C_{0T} with respect to the metric $\mathbf{r}_{0T}$, and so $\Psi_{0T}^a(x)$ is a closed set.

Set $\mathcal{X}_t = \{x \in \mathcal{X} : X_{x,y}^\varepsilon(s) \in \mathcal{X}$ and $\bar{X}_x^\varepsilon(s) \in \mathcal{X}$ for all $y \in \bar{\mathcal{W}},\ s \in [0, t/\varepsilon],\ \varepsilon > 0\}$. Clearly, $\mathcal{X}_t \supset \{x \in \mathcal{X} : \inf_{z \in \partial\mathcal{X}} |x - z| \ge 2Kt\}$.

Theorem

Set $Z_{x,y}^\varepsilon(t) = X_{x,y}^\varepsilon(t/\varepsilon)$ then for any $x \in \mathcal{X}_T$, $a, \delta, \lambda > 0$ and every $\gamma \in C_{0T}$, $\gamma_0 = x$ there exists $\varepsilon_0 = \varepsilon_0(x, \gamma, a, \delta, \lambda) > 0$ such that for $\varepsilon < \varepsilon_0$,

$$m\left\{y \in \mathcal{W} : \mathbf{r}_{0T}(Z_{x,y}^\varepsilon, \gamma) < \delta\right\} \ge \exp\left\{-\frac{1}{\varepsilon}(S_{0T}(\gamma) + \lambda)\right\} \text{ and}$$

$$m\left\{y \in \mathcal{W} : \mathbf{r}_{0T}(Z_{x,y}^\varepsilon, \Psi_{0T}^a(x)) \ge \delta\right\} \le \exp\left\{-\frac{1}{\varepsilon}(a - \lambda)\right\}$$

where, recall, m is the normalized Riemannian volume on $\mathbf{M}$. The functional $S_{0T}(\gamma)$ for $\gamma \in C_{0T}$ is finite if and only if $\dot{\gamma}_t = \bar{B}_{\nu_t}(\gamma_t)$, $\nu_t \in \mathcal{M}_{\gamma_t}$ for Lebesgue almost all $t \in [0, T]$. Finally, $S_{0T}(\gamma)$ achieves its minimum 0 only on γ^u.

Let $V \subset \mathcal{X}$ be a connected open set and put $\tau^\varepsilon_{x,y}(V) = \inf\{t \geq 0 : Z^\varepsilon_{x,y}(t) \notin V\}$ where we take $\tau^\varepsilon_{x,y}(V) = \infty$ if $X^\varepsilon_{x,y}(t) \in V$ for all $t \geq 0$.

Corollary

Under the conditions of the above theorem for any $T > 0$ and $x \in V$,

$$\lim_{\varepsilon \to 0} \varepsilon \log m\left\{y \in \mathcal{W} : \tau^\varepsilon_{x,y}(V) < T\right\}$$
$$= -\inf\left\{S_{0t}(\gamma) : \gamma \in C_{0T},\, t \in [0, T],\, \gamma_0 = x,\, \gamma_t \notin V\right\}.$$

Next, we will study "very long", i.e. exponential in $1/\varepsilon$, time "adiabatic" behaviour of the slow motion which cannot be described usually in the traditional theory of averaging where only perturbations of integrable Hamiltonian systems are considered which are not chaotic enough. Namely, we will describe such long time behavior of Z^ε in terms of the function

$$R(x, z) = \inf_{t \geq 0, \gamma \in C_{0t}} \{S_{0t}(\gamma) : \gamma_0 = x,\, \gamma_t = z\}$$

under various assumptions on the averaged motion $\bar{Z}$. Observe that R satisfies the triangle inequality $R(x_1, x_2) + R(x_2, x_3) \geq R(x_1, x_3)$ for any $x_1, x_2, x_3 \in \mathcal{X}$ and it determines a semi metric on $\mathcal{X}$ which measures "the difficulty"' for the slow motion to move from point to point in terms of the functional S.

Introduce the averaged flow Ψ^t on $\mathcal{X}_t$ by

$$\frac{d\Psi^t x}{dt} = \bar{B}(\Psi^t x), \ x \in \mathcal{X}_t$$

where, recall, $\bar{B}(z) = \bar{B}_{\mu_z^{\text{SRB}}}(z)$ and $\bar{B}_\nu(z) = \int B(z,y)d\nu(y)$ for any probability measure ν on $\mathbf{M}$. Call a Ψ^t-invariant compact set $\mathcal{O} \subset \mathcal{X}$ an S-compact if $\forall\, \eta > 0$, $\exists\, T_\eta \geq 0$ and $\exists$ open $U_\eta \supset \mathcal{O}$ such that whenever $x \in \mathcal{O}$ and $z \in U_\eta$ there are $t \in [0, T_\eta]$ and $\gamma \in C_{0t}$ with $\gamma_0 = x$, $\gamma_t = z$ and $S_{0t}(\gamma) \leq \eta$. Then $R(x,z) = 0$ for any x, z in an S-compact $\mathcal{O}$ and $R(x,z) \equiv$const when $z \in \mathcal{X}$ is fixed and x runs over $\mathcal{O}$. A vector field B on $\mathcal{X} \times \mathbf{M}$ is called complete at $x \in \mathcal{X}$ if the convex set of vectors $\{\beta \bar{B}_\nu(x) : \beta \in [0,1],\ \nu \in \mathcal{M}_x\}$ contains an open neighborhood of the origin in $\mathbb{R}^d$. It turns out that if $\mathcal{O} \subset \mathcal{X}$ is a compact Ψ^t-invariant set such that B is complete at each $x \in \mathcal{O}$ and either $\mathcal{O}$ contains a dense orbit of the flow Ψ^t (i.e. Ψ^t is topologically transitive on $\mathcal{O}$) or $R(x,z) = 0$ for any $x, z \in \mathcal{O}$ then $\mathcal{O}$ is an S-compact. Moreover, $\mathcal{O}$ is an S-compact if B is complete only at some point of $\mathcal{O}$ and the flow Ψ^t on $\mathcal{O}$ is minimal, i.e. the Ψ^t-orbits of all points are dense in $\mathcal{O}$.

A compact Ψ^t-invariant set $\mathcal{O} \subset \mathcal{X}$ is called an attractor for Ψ^t if $\exists$ open $U \supset \mathcal{O}$ and $t_U > 0$ such that $\Psi^{t_U} \bar{U} \subset U$ and $\lim_{t\to\infty} \text{dist}(\Psi^t z, \mathcal{O}) = 0$ for all $z \in U$. The set $V = \{z \in \mathcal{X} : \lim_{t\to\infty} \text{dist}(\Psi^t z, \mathcal{O}) = 0\}$, which is clearly open, is called the basin of $\mathcal{O}$. An attractor which is an S-compact is called S-attractor.

Theorem

Let $\mathcal{O} \subset \mathcal{X}$ be an S-attractor whose basin contains the closure $\bar{V}$ of a connected open set V with a piecewise smooth boundary ∂V such that $\bar{V} \subset \mathcal{X}$ and assume that for each $z \in \partial V$ there exists $\varpi = \varpi(z) > 0$ and an F_z^t-invariant probability measure $\nu = \nu_z$ on Λ_z such that $z + s\bar{B}(z) \in V$ but $z + s\bar{B}_\nu(z) \in \mathbb{R}^d \setminus \bar{V}$ $\forall s \in (0, \varpi]$, i.e. $\bar{B}(z) \neq 0$, $\bar{B}_\nu(z) \neq 0$ and the former vector points out into the interior while the latter into the exterior of V. Set $R_\partial(z) = \inf\{R(z,\tilde{z}) : \tilde{z} \in \partial V\}$ and $\partial_{\min}(z) = \{\tilde{z} \in \partial V : R(z,\tilde{z}) = R_\partial(z)\}$. Then $R_\partial(z) \equiv R_\partial$ and $\partial_{\min}(z) \equiv \partial_{\min}$ $\forall z \in \mathcal{O}$, $R_\partial(x) \leq R_\partial$ $\forall x \in V$ and

$$\lim_{\varepsilon \to 0} \varepsilon \log \int_{\mathcal{W}} \tau^\varepsilon_{x,y}(V) dm(y) = R_\partial > 0 \ \forall x \in V.$$

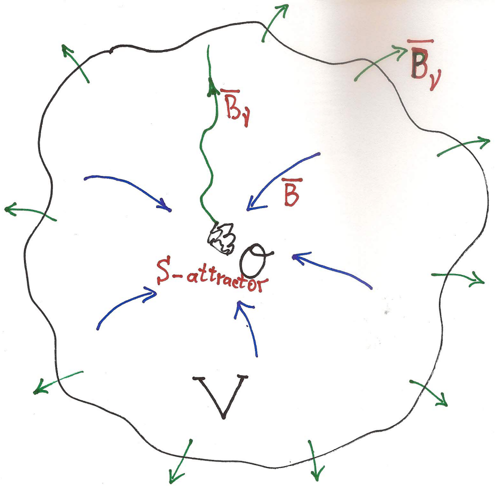
Next, $\forall \alpha > 0 \ \exists \lambda(\alpha) = \lambda(x,\alpha) > 0$ such that $\forall \varepsilon > 0$ small,

$$m\{y \in \mathcal{W} : \ e^{(R_\partial - \alpha)/\varepsilon} > \tau^\varepsilon_{x,y}(V) \text{ or } \tau^\varepsilon_{x,y}(V) > e^{(R_\partial + \alpha)/\varepsilon}\} \leq e^{-\lambda(\alpha)/\varepsilon} \text{ and}$$

$$\lim_{\varepsilon \to 0} m\{y \in \mathcal{W} : \ dist(Z^\varepsilon_{x,y}(\tau^\varepsilon_{x,y}(V)), \partial_{\min}) \geq \delta\} = 0 \ \forall x \in V, \ \delta > 0$$

provided $R_\partial < \infty$ and the latter holds true if and only if for some $T > 0$ there exists $\gamma \in C_{0T}$, $\gamma_0 \in \mathcal{O}$, $\gamma_T \in \partial V$ such that $\dot{\gamma}_t = \bar{B}_{\nu_t}(\gamma_t)$ for Lebesgue almost all $t \in [0, T]$ with $\nu_t \in \mathcal{M}_{\gamma_t}$ then $R_\partial < \infty$.

Exit from a domain: picture



$$\frac{d\Psi_x^t}{dt} = \bar{B}(\Psi_x^t), \quad \bar{B} = \bar{B}_{\mu^{SRB}}$$

$$\bar{B}_\nu(x) = \int B(x,y)\, d\nu_x(y), \quad \nu_x \text{ - some } F_x\text{-invariant measure}$$

Next, we want to describe transitions of the slow motion Z^ε between basins of attractors of the averaged flow Ψ^t.

Assumption

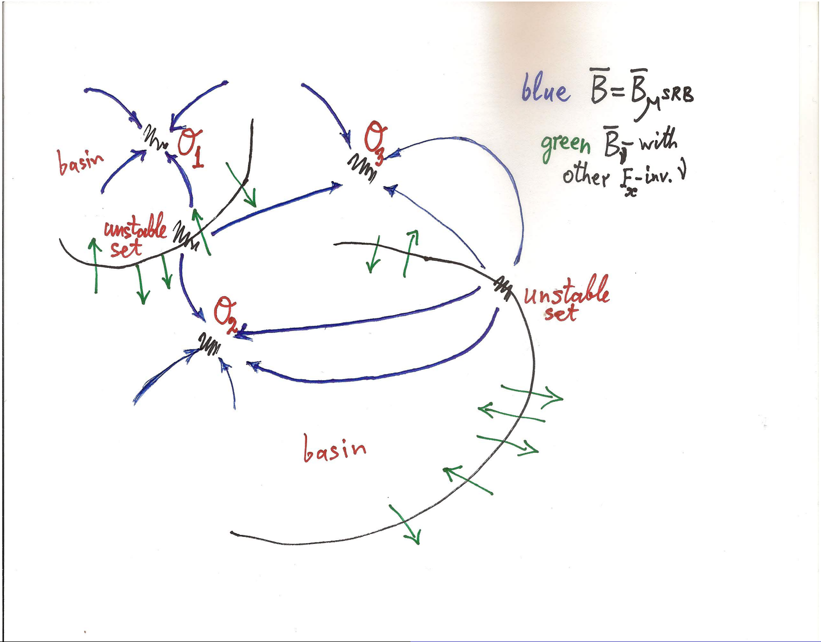
The family $\{F_x^t,\ t \le 1,\ x \in \mathbb{R}^d\}$ is a compact set of diffeomorphisms in the C^2 topology, $\|B(x,y)\|_{C^2(\mathbb{R}^d \times \mathbf{M})} \le K$ for some $K > 0$ independent of x, y and there exists $r_0 > 0$ such that

$$(x, B(x,y)) \le -K^{-1} \text{ for any } y \in \mathcal{W} \text{ and } |x| \ge r_0.$$

Suppose that the ω-limit set of Ψ^t is compact and it consists of a finite number of S-attractors $\mathcal{O}_1, ..., \mathcal{O}_\ell$ whose basins $V_1, ..., V_\ell$ have piecewise smooth boundaries $\partial V_1, ..., \partial V_\ell$ and of the remaining part which is contained in $\cup_{1 \le j \le \ell} \partial V_j$. Assume also that for any $z \in \cap_{1 \le i \le k} \partial V_{j_i}$, $k \le \ell$ there exist $\varpi = \varpi(z) > 0$ and an F_z^t-invariant measures $\nu_1, ..., \nu_k$ such that $z + s\bar{B}_{\nu_i}(z) \in V_{j_i}\ \forall\, s \in (0, \varpi]$ and $i = 1, ..., k$, i.e. $\bar{B}_{\nu_i}(z) \ne 0$ and it points out into the interior of V_{j_i} which means that from any boundary point it is possible to go to any adjacent basin along a curve with an arbitrarily small S-functional. Let $\delta > 0$ be so small that each $U_\delta(\mathcal{O}_i) = \{z \in \mathcal{X} :\ \mathrm{dist}(z, \mathcal{O}_i) < \delta\}$ is contained with its closure in the basin V_i. For any $x \in V_i$ set

$$\tau_{x,y}^\varepsilon(i) = \inf\left\{t \ge 0 :\ Z_{x,y}^\varepsilon(t) \in \cup_{j \ne i} U_\delta(\mathcal{O}_j)\right\}.$$

Several S-attractors: picture



Theorem

The function $R_{ij}(x) = \inf_{z \in V_j} R(x,z)$ *takes on the same value* R_{ij} *for all* $x \in \mathcal{O}_i$, $i \neq j$. *Let* $R_i = \min_{j \neq i, j \leq \ell} R_{ij}$. *Then*

$$\lim_{\varepsilon \to 0} \varepsilon \log \int_{\mathcal{W}} \tau^{\varepsilon}_{x,y}(i) dm(y) = R_i > 0 \; \forall \, x \in V_i$$

and $\forall \, \alpha > 0 \; \exists \, \lambda(\alpha) = \lambda(x, \alpha) > 0$ *such that* $\forall \, \varepsilon > 0$ *small,*

$$m\{y \in \mathcal{W} : \; e^{(R_i - \alpha)/\varepsilon} > \tau^{\varepsilon}_{x,y}(i) \text{ or } \tau^{\varepsilon}_{x,y}(i) > e^{(R_i+\alpha)/\varepsilon}\} \leq e^{-\lambda(\alpha)/\varepsilon}.$$

Suppose that for all i, B *is complete on* ∂V_i *and the restriction of the* ω*-limit set of* Ψ^t *to* ∂V_i *consists of a finite number of* S*-compacts. Assume also* $\forall \, i$ $\exists ! \, \iota(i) \leq \ell$, $\iota(i) \neq i$ *such that* $R_i = R_{i\iota(i)}$. *Define* $\iota_0(i) = i$, $\tau^{\varepsilon}_v(i,1) = \tau^{\varepsilon}_v(i)$ *and recursively,*
$\iota_k(i) = \iota(\iota_{k-1}(i))$ *and* $\tau^{\varepsilon}_v(i,k) = \tau^{\varepsilon}_v(i, k-1) + \tau^{\varepsilon}_{v_\varepsilon(k-1)}\big(j(v_\varepsilon(k-1))\big)$, *where* $v_\varepsilon(k) = \Phi_\varepsilon^{\varepsilon^{-1}\tau^{\varepsilon}_v(i,k)} v$, $j((x,y)) = j$ *if* $x \in V_j$. *Then* $\forall \, x \in V_i$ *and* $\forall \, \alpha > 0$ $\exists \, \lambda = \lambda(x) > 0$ *such that* $\forall \, n \in \mathbb{N}$,

$$m\{y \in \mathcal{W} : \; Z^{\varepsilon}_{x,y}(\tau^{\varepsilon}_{x,y}(i,k)) \notin V_{\iota_k(i)} \text{ for some } k \leq n\} \leq n e^{-\lambda/\varepsilon}.$$

Discrete time case, Markov fast motions

Similar results hold true for difference equations where fast motions satisfy

Assumption

The family $F_x = \Phi(x, \cdot)$ consists of C^2-diffeomorphisms or endomorphisms of a compact $n_{\mathbf{M}}$-dimensional Riemannian manifold $\mathbf{M}$ with uniform C^2 dependence on the parameter x belonging to a relatively compact open connected set $\mathcal{X} \subset \mathbb{R}^d$. All F_x, $x \in \bar{\mathcal{X}}$ are either expanding maps of $\mathbf{M}$ or diffeomorphisms possessesing basic hyperbolic attractors Λ_x with (uniform in x) hyperbolic splittings and one open set $\mathcal{W} \subset \mathbf{M}$ which contains all Λ_x but is contained in their basins.

Similar results hold also true when fast motions are Markov processes satisfying Doeblin conditions with the rate functional I given by the Donsker-Varadhan formula. In the continuous time take $X^\varepsilon(t), Y^\varepsilon(t)$ determined by an ordinary differential equation for the slow motion X^ε together with a non degenerated stochastic differential equation with coefficients dependent on the slow x-variable. In the discrete time we can start with a parametric family of Markov chains $Y_{x,y}(n)$, $n \geq 0$, $Y_{x,y}(0) = y$ on a compact C^2 Riemannian manifold M with transition probabilities $P^x(y, \Gamma) = P^x_y\{Y_{x,y}(1) \in \Gamma\}$ having positive C^1 densities $p^x(y, z) = P^x(y, dz)/m(dz)$, m-volume. Now, define $X^\varepsilon(n)$ and $Y^\varepsilon(n)$ adding to the difference equation for the slow motion X^ε another equation $P\{Y^\varepsilon(n+1) \in \Gamma \big| X^\varepsilon(n) = x,\ Y^\varepsilon(n) = y\} = P^x(y, \Gamma)$.

In both examples $F_x y = 3y + x \pmod 1$ where $x \in \mathbb{R}^1$ and $y \in [0,1]$, F_x are expanding maps of the circle $\mathbb{T}^1$. Next,

$$B(x,y) = x(x^2-4)(1-x)(a+x) + 50\sin 2\pi y$$

where $a = 1$ in the 1st example and $a = 1.5$ in the 2nd example. Thus, $X^\varepsilon(n+1) = X^\varepsilon(n) + \varepsilon B(X^\varepsilon(n), Y^\varepsilon(n))$, $Y^\varepsilon(n+1) = 3Y^\varepsilon(n) + X^\varepsilon(n)$ (mod 1) and we have maps $\Phi_\varepsilon : \mathbb{R}^1 \times \mathbb{T}^1 \to \mathbb{R}^1 \times \mathbb{T}^1$ defined by

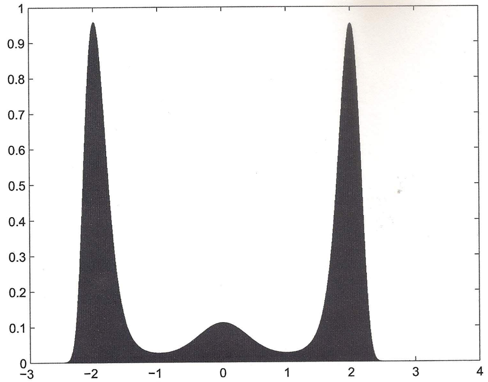
$$\Phi_\varepsilon(x,y) = \big(x + \varepsilon(x(x^2-4)(1-x)(a+x) + 50\sin 2\pi y),\, 3y + x \pmod 1\big).$$

All maps F_x preserve the normalized Lebesgue measure Leb on $\mathbb{T}^1$ which is the SRB measure μ_x^{SRB} for each F_x in this case. The averaged equation for $\bar{Z}(t) = \bar{X}^\varepsilon(t/\varepsilon)$ has here the form

$$\frac{d\bar{Z}(t)}{dt} = \bar{B}(\bar{Z}(t)),$$

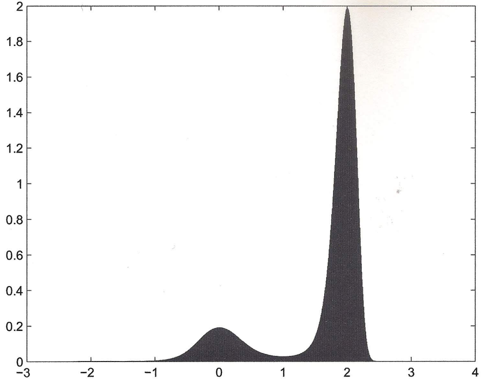
where $\bar{B}(x) = x(x^2-4)(1-x)(a+x)$. When $a = 1$ the one dimensional vector field $\bar{B}(x)$ has 3 attracting fixed points $\mathcal{O}_1 = 2, \mathcal{O}_2 = 0, \mathcal{O}_3 = -2$ and two repelling fixed points 1 and -1. When $a = 1.5$ it has the same attracting fixed points but one of two repelling fixed points moves from -1 to $-3/2$ making the basin of -2 smaller while the left interval of the basin of 0 becomes larger, so it is more difficult to exit from there to the left than to the right.

Histogram of $X^{\varepsilon}_{x,y}(n)$, $n = 0, 1, \dots, 10^9$



Asymmetrical basins case

Histogram of $X^{\varepsilon}_{x,y}(n)$, $n = 0, 1, \dots, 10^9$



Discussion

- We plot histograms of a single orbit of the slow motion $X^{\varepsilon}_{x,y}(n),\ n = 0, 1, 2, ..., 10^9$ with $\varepsilon = 10^{-3}$ and the initial values in the 1st case $x = 0,\ y = 0.001$ and in the 2nd case $x = -2,\ y = 0.001$.
- In order to verify that B is complete at the fixed points $-2, -1, 0, 1, 2$ of the averaged system in the 1st example we observe that these points F_x coincides with the map $y \to 3y (\mathrm{mod}\ 1)$, and so we can take the periodic orbits $1/8, 3/8$ and $5/8, 7/8$ of the latter and notice that the average of $\sin 2\pi y$ along the former is $1/\sqrt{2}$ and along the latter $-1/\sqrt{2}$ which yields completeness of B at zeros of $\bar{B}$. For the 2nd example it remains to verify completness only for $x = -3/2$ which follows since $\sin 2\pi y$ equals 1 and -1 at two fixed points $1/4$ and $3/4$ of $F_{-3/2}$, respectively.
- By the theorem the transitions between $\mathcal{O}_1, \mathcal{O}_2$, and $\mathcal{O}_3$ are determined by $R_{ij},\ i, j = 1, 2, 3$ obtained via the functionals $S_{0t}(\gamma)$. Even here they are not easy to compute since this leads to complicated non traditional variational problems though the functionals $I_x(\nu)$ are given now by the simple formula $I_x(\nu) = \ln 3 - h_\nu(F^1_x)$ if ν is F_x-invariant and $= \infty$, otherwise, while the set of F_x-invariant measures can be reasonably described since all F_x's are conjugate to the simple map $y \to 3y\,(\mathrm{mod}\,1)$.
- It turns out that "exactly" the same histograms we obtain when we replace $Y^{\varepsilon}(n+1) = 3Y^{\varepsilon}(n) + X^{\varepsilon}(n)$ (mod 1) by $Y^{\varepsilon}(n+1) = Y^{\varepsilon}(n) + X^{\varepsilon}(n) + \xi_{n+1}$ (mod 1) where $\xi_1, \xi_2, ...$ are i.i.d. random variables on $[0, 1]$ with a Lebesgue equivalent distribution.

In the non coupled situation

$$\frac{dX^\varepsilon(t)}{dt} = \varepsilon B(X^\varepsilon(t), F^t y)$$

the basic large deviations theorem in averaging is not difficult. Then it suffices to show that for any continuous function $q_t(y) = q(t, y)$,

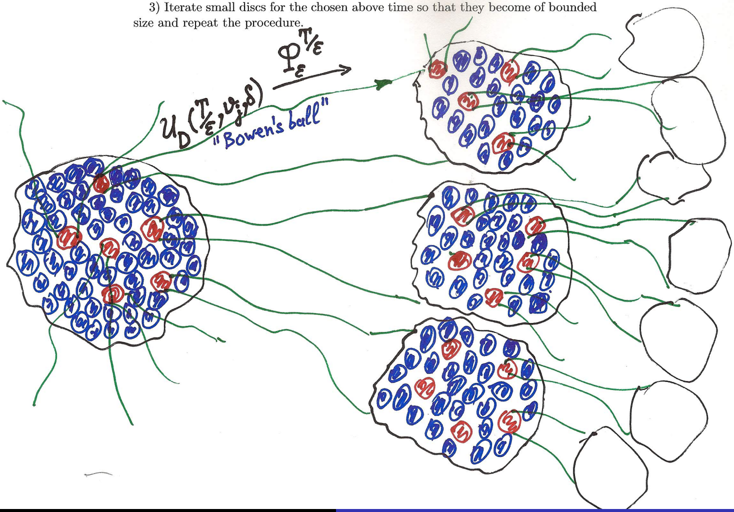
$$\lim_{\varepsilon \to 0} \varepsilon \log \int \exp\left(\varepsilon^{-1} \int_0^T q_t(F^{t/\varepsilon} y) dt\right) dm(y) = \int_0^T \Pi(\varphi^u + q_t) dt$$

which can be done splitting the interval $[0, T]$ into small subintervals. In the fully coupled case one needs more complicated technical tools, in particular, a version of the general large deviations bounds when usual assumptions hold true with errors. This allows approximate decoupling on small time intervals since the flows F_x^t change slowly in time as x is the slow variable but one has to be careful since for hyperbolic flows errors accumulate exponentially in time. The above results concerning exponential in $1/\varepsilon$ time behavior are not easy already in the non coupled situation. Actually, in the non coupled probabilistic situation when the fast motion is a non degenerated Markov chain with continuous time and finitely many states this is easier and it was done by Freidlin. In the dynamical systems setup this can be done by a kind of rough Markov property argument for unstable disks.

1) Split an unstable disk into very small disks so that for some time (long for fast variables but short for slow variables) all solutions starting on each small disk go close to each other so that the questions we are interested about (exit from a domain or transition to the basin of another S-compact) hold or not hold true simultaneously for all points in the small disk;

2) Apply approximate LD on the original disk estimating the measure of small disks satisfying our conditions.

3) Iterate small discs for the chosen above time so that they become of bounded size and repeat the procedure.



- D.B. Anosov, *Averaging in systems of ordinary differential equations with fast oscillating solutions*, Izv. Acad. Nauk SSSR Ser. Mat. **24** (1960), 731–742 (in Russian).
- V.I. Bakhtin and Yu. Kifer, *Nonconvergence examples in averaging*, Contemporary Math. **469** (2008), 1–17.
- M.I. Freidlin, *The averaging principle and theorems on large deviations,* Russ. Math. Surv., **33,** No.5 (1978), 107–160.
- Yu. Kifer, *Averaging in dynamical systems and large deviations,* Invent. Math., **110** (1992), 337–370.
- Yu. Kifer, *Averaging principle for fully coupled dynamical systems and large deviations,* Ergod. Th.& Dynam. Syst. **24** (2004), 847–871.
- Yu. Kifer, *Another proof of the averaging principle for fully coupled dynamical systems with hyperbolic fast motions*, Discrete Contin. Dyn. Syst. **13** (2005), 1187–1201.
- Yu. Kifer, *Large Deviations and Adiabatic Transitions for Dynamical Systems and Markov Processes in Fully Coupled Averaging*, Memoirs of the AMS **201**, Providence R.I., 2009.