Averaging of Noisy Hamiltonian Systems and Stratified Spaces

Richard B. Sowers
Department of Mathematics
University of Illinois
Urbana-Champaign
http://www.math.uiuc.edu/~r-sowers

Research supported by NSF DMS 0071484
and ONR N00040900064.

Joint work with N. S. Namachchivaya of UIUC (AAE).

May, 2001
Outline

• Classical problems
• Graph-valued problems
• Whiskered sphere problem
• Motivation for Glueing Conditions
• Stability and related results
Classical problems

Let’s start with the equation of a 1-dimensional particle in a force field
\[ \ddot{x}_t = F(x_t) \]
where \( F \in C^\infty(\mathbb{R}) \). This can be written as an \( \mathbb{R}^2 \)-valued ODE
\[ \dot{Z}_t^0 = \begin{pmatrix} \dot{x}_t^0 \\ \dot{y}_t^0 \end{pmatrix} = \begin{pmatrix} y_t^0 \\ F(x_t^0) \end{pmatrix}. \]

We can use the idea of conservation of energy to define
\[ H(x, y) \overset{\text{def}}{=} U(x) + \frac{y^2}{2} \]
where \( \dot{U}(x) = -F(x) \). Then
\[ \dot{Z}_t^0 = \begin{pmatrix} \dot{x}_t^0 \\ \dot{y}_t^0 \end{pmatrix} = \left( \frac{\partial H}{\partial y}(x_t^0, y_t^0), -\frac{\partial H}{\partial x}(x_t^0, y_t^0) \right) = \nabla H(Z_t^0) \]
and the energy \( H(Z_t^0) \) is conserved since
\[ \frac{d}{dt} H(Z_t^0) = \left( \frac{\partial H}{\partial x} \frac{\partial H}{\partial y} - \frac{\partial H}{\partial y} \frac{\partial H}{\partial x} \right) (x_t^0, y_t^0) \]
\[ = \{ H, H \}(x_t^0, y_t^0) = 0. \]
In reality, the world is noisy; the acceleration is subject to random jerk. A more accurate model for the particle is

\[ dx^\varepsilon_t = \frac{\partial H}{\partial y}(x^\varepsilon_t, y^\varepsilon_t)dt \]

\[ dy^\varepsilon_t = -\frac{\partial H}{\partial x}(x^\varepsilon_t, y^\varepsilon_t)dt + \varepsilon dW_t, \]

where \( W \) is a Wiener process, and \( \varepsilon \) is a small parameter. Now energy is not conserved, but is slowly-varying; by Ito's formula,

\[ dH(x^\varepsilon_t, y^\varepsilon_t) = \varepsilon \frac{\partial H}{\partial y}(x^\varepsilon_t, y^\varepsilon_t)dW_t + \varepsilon^2 \frac{1}{2} \frac{\partial^2 H}{\partial y^2}(x^\varepsilon_t, y^\varepsilon_t)dt \]

Ito correction term

Thus, we will see a change in energy in time of order \( \varepsilon^{-2} \); let's look at \( Z_t^\varepsilon \equiv (X_{t/\varepsilon^2}^\varepsilon, Y_{t/\varepsilon^2}^\varepsilon) \). This rescaling gives us a new SDE:

\[ dZ_t^\varepsilon = \frac{1}{\varepsilon^2} \nabla H(Z_t^\varepsilon)dt + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t \]

Then

\[ dH(X_t^\varepsilon, Y_t^\varepsilon) = \frac{\partial H}{\partial y}(X_t^\varepsilon, Y_t^\varepsilon)dW_t + \frac{1}{2} \frac{\partial^2 H}{\partial y^2}(X_t^\varepsilon, Y_t^\varepsilon)dt. \]
What are the dynamics of the conserved quantity (energy) as $\varepsilon \to 0$?

Let's consider for simplicity a simple harmonic oscillator, viz.

$$U(x) = \frac{k}{2} x^2;$$

thus the force is $F(x) = -kx$.

Single-Well Potential
Thus

\[ H(x, y) = \frac{kx^2 + y^2}{2} \]

Look at the \textbf{graph} of the energy \( H \)

Energy contour for single-well potential

Let \( \dot{Z}_t = \nabla H(Z_t) \); the solutions of this simply move around the level sets of \( H \) and if \( H(Z_0) = h \), then

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T f(Z_s^0(z))ds = \langle f, \mu_h \rangle
\]

where

\[
\mu_h(dx) = C_h \int_{z': H(z') = h} f(z') \| \nabla H(z') \|^{-1} ds(dz')
\]

(where \( C_h \) normalizes so that \( \mu_h(\mathbb{R}^2) = 1 \)); note that depends only on \( h \); i.e., \( \mu_z = \mu_H(z) \) (\( \mu_h \) is a \textit{Liouville measure}).
Graph-valued Problems

The foregoing is essentially classical. A natural next problem, solved by Freidlin and Wentzell (1996) and Freidlin and Weber (1998) is that of a **double-well**

\[
U(x) = x(x - \alpha_1)(x + \alpha_2)
\]

(where the \(\alpha_i\)'s are positive and in general not equal).

Double-Well Potential

The graph of the total energy thus has **two** wells.
Again we want to consider the SDE

\[
dZ_t^\varepsilon = \frac{1}{\varepsilon^2} \nabla H(Z_t^\varepsilon) dt + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t
\]

Heuristically, we again want to divide by the fast motion and find a reduced state space. However, the behavior of \( Z^\varepsilon \) at the zero energy level (the homoclinic orbit) is very sensitive to the noisy perturbations. We should reduce \( \mathbb{R}^2 \) to a graph representing the topology of \( H \); the Liouville measures can be parametrized by such graphs.
Freidlin-Wentzell and Freidlin-Weber identified a limiting graph-valued process \( \xi_t \overset{\text{def}}{=} \lim_{\varepsilon \to 0} [Z_t^\varepsilon] \) (where \([z]\) is the equivalence class of \(z\) under \(\sim\)). To mathematically get the graph, we can use \textbf{chain equivalent} orbit space. Define \( \mathcal{M} \overset{\text{def}}{=} \mathbb{R}^2 / \sim \).
Theorem. [Freidlin and Wentzell] The process $[Z^\varepsilon_t]$ tends to a Markov process $\xi$ on the graph with a computable generator.

The dynamics of $\xi$ are given by the classical calculations as long as it does not touch the vertex. When it does, its behavior is governed by some glueing conditions, which are restrictions on the domain of the generator of the limiting process. These glueing conditions are roughly and imprecisely as follows. If $\xi$ starts at the vertex, one flips a three-sided coin (with statistics given by the glueing gonditions) to decided which leg to go to. Then $\xi$ evolves according to the classical dynamics until it again hits the vertex, where the process is repeated again (with a new coin). This is similar to skew Brownian motion. It is also related to spider martingales. We note that $\xi$ is diffusive (e.g. continuous), but not representable by an SDE.
Whiskered sphere problems

Let’s make two changes to the system.

- In both the single-well and double-well problems, the critical points of $H$ are nondegenerate; i.e., if $dH(x, y) = 0$, then $D^2H(x, y)$ is full rank. We want to violently remove this assumption. We will assume that $H$ has a flat.

- Let’s assume that the noise is in both components (to avoid complications).

Our “canonical” Hamiltonian will be

$$H(z) \overset{\text{def}}{=} (\max\{\|z\| - 1, 0\})^n \quad z \in \mathbb{R}^2$$

where $n > 2$ (so that $H$ has a locally Lipshitz derivative). Then

$$\mathcal{A} \overset{\text{def}}{=} \{z \in \mathbb{R}^2 : H(z) = 0\} = \{z \in \mathbb{R}^2 : \|z\| \leq 1\}$$
Flattened Hamiltonian (from below)

As usual, define

\[ \nabla H(x, y) \overset{\text{def}}{=} \left( \frac{\partial H}{\partial y}(x, y), -\frac{\partial H}{\partial x}(x, y) \right) \]

and consider the SDE

\[ dZ_t^\varepsilon = \frac{1}{\varepsilon^2} \nabla H(Z_t^\varepsilon) dt + dW_t \]

where \( W \) is now a two-dimensional Brownian motion.
Again, we define

\[ \mathcal{M} \overset{\text{def}}{=} \mathbb{R}^2 / \sim \]

where \( \sim \) denotes chain equivalence.

Bottom view of is

Contour Plot of Flattened Hamiltonian

Each point inside the disk is equivalent only to itself. Each circle outside the disk (including the edge of the disk) can be collapsed into a single point.
We get a **whiskered sphere**.

![Diagram]

becomes

\[ \delta^\circ \]

\[ \cup \text{Collapsed edge of disk} \]

\[ \cup \text{Collapsed circles} \]

\[ = \]

Geometry of Flattened Hamiltonian
**Theorem. [Sowers]** The process $[Z^\varepsilon_t]$ tends to a Markov process $\xi$ on the whiskered sphere with a computable generator.

This is geometrically interesting; the whiskered sphere is a **stratified space** (with a dimensional discontinuity). This is the general framework of such averaging problems.

![Stratified Space](image)

Locally, the whiskered sphere can look like a line, a plane or the union of a line and plane (a flagpole)—related recent work by **Burdzy and Bass** on a “fiber” Brownian motion. The limiting process executes a standard Brownian motion while it is inside the disk (on the ball), and when it is on the line it executes a diffusion whose coefficients stem from averaging the original coefficients. At the vertex, a ”coin toss” is again needed (work of **Evans and Sowers**).
Stochastic averaging in general leads to processes on Stratified Spaces
Motivation for Glueing Conditions

Where do the glueing conditions come from? We note that by looking at the Fokker-Planck equation, they correspond to a continuity equation for densities.

i) Essentially, we want to understand the relative likelihood that the limiting process changes strata. Since the strata are defined by level sets of the Hamiltonian $H$ (the slowly-varying quantity), such transitions are defined by the quadratic variation $\langle dH, dH \rangle$ (the bracket of the generator of the slow motion).

ii) We must compare such transitions to the fast drift caused by $\varepsilon^{-2} \tilde{\nabla} H$.

iii) Essentially, we must make a boundary expansion (a singular perturbation expansion) near orbits which collapse into points between strata (not the original technique by Freidlin and Wentzell). Motivated by some work of Khasminskii, we can use an angle $\Theta$ defined by

$$ (\tilde{\nabla} H, \nabla \Theta) = \langle dH, dH \rangle $$

This normalizes the picture so that equal amounts of angle correspond to equal “likelihood” of jumping between strata.
Stability and related results

• How does averaging interact with ”stability”? Consider the cylinder-valued SDE

\[
dZ_t^\varepsilon(z) = a(Z_t^\varepsilon(x)) \frac{1}{\varepsilon^2} \frac{\partial}{\partial \theta}(Z_t^\varepsilon(z)) dt + \sigma(Z_t(z)) dW_t
\]

\[
Z_0(z) = z
\]

where \(a\) depends only on the height on the cylinder and \(a\) is increasing (a twist map). Consider the effect of a small perturbation on the initial conditions.

It turns out that (infinitesimally for \(z\) and \(z'\) close), we have

\[
\mathbb{E} \left[ \| Z_t^\varepsilon(z) - Z_t^\varepsilon(z') \| \right] \asymp \exp \left[ \varepsilon^{-4/3} I_t \right]
\]
(due to Sowers; the main idea comes from some work of Pinsky and Wihstutz). Relevant work has been also done by Baxendale and Goukasian, and Imkeller and Namachchivaya.
Higher-Dimensional Problems

In the classical and graph-valued problem, we were essentially looking at Newtonian dynamics of a single quantity. Often, several Newtonian quantities can be coupled. Energy can then be transferred between the variables. Usually there is an obvious energy function which is slowly-varying. Often there is also another conserved quantity (i.e., a symmetry). In this case, the reduced state space will locally have dimension

\[
\underbrace{4} - \underbrace{2} = 2.
\]

dimension of original state space  number of conserved quantities