Lecture 13: Asymptotic analysis of SDEs

Readings

Recommended:

- Pavliotis [2014] 5.1

Optional:

- Pavliotis and Stuart [2008] Chapter 11 (homogenization), other chapters for other types of convergence and many applications.
- Pavliotis [2014] (overdamped & underdamped limits of Langevin equation)
- Papanicolaou [1977] – a nice review article covering several aspects of asymptotic analysis, with hints at some of the more rigorous techniques and results
- Gardiner [2009] Ch 8, 9 (for a slightly different approach)

Sometimes an SDE will contain a parameter which is much smaller or larger than the others, for example due to widely separate time or space scales. It is often of interest to average over these fast scales to obtain dynamics on longer scales, which are often simpler to analyze. Today we will consider one technique to do this, based on performing formal singular perturbation analysis on the generator of the process.

13.1 Limit of non-white noise

Consider the equation

$$\frac{dx^\varepsilon}{dt} = h(x^\varepsilon) + \frac{1}{\varepsilon} f(x^\varepsilon) y^\varepsilon(t), \tag{1}$$

where $y^\varepsilon(t)$ is a stationary, *non-white* stochastic process. If the correlation time of $y^\varepsilon(t)$ is very short (i.e. the covariance function decreases very rapidly), then we might expect it to behave as a white noise, so that (1) becomes an SDE.

To investigate this, suppose y^ε is found from changing the timescale of a regular Ornstein-Uhlenbeck process, as $y^\varepsilon = y(t/\varepsilon^2)$, where $\varepsilon \ll 1$, and $y(t)$ is an OU process:

$$dy = -\alpha y + \sigma dW_t. \tag{2}$$

Then, using the fact that $W_{t/\varepsilon^2} \overset{d}{=} \frac{1}{\varepsilon} W_t$, we obtain

$$dy^\varepsilon = -\alpha \frac{y^\varepsilon}{\varepsilon^2} dt + \frac{\sigma}{\varepsilon} dW_t. \tag{3}$$

The covariance function in steady-state, and corresponding power spectrum, are

$$C^\varepsilon(t) = \frac{\sigma^2}{2\alpha} e^{-\frac{\alpha}{\varepsilon^2}|t|}, \qquad \hat{C}^\varepsilon(\omega) = \frac{\sigma^2}{2\pi\alpha} \frac{\alpha\varepsilon^2}{(\alpha^2 + \varepsilon^4\omega^2)} \quad \sim \quad \frac{\sigma^2}{2\pi\alpha^2}\varepsilon^2 \quad \text{as } \varepsilon \to 0.$$

This covariance function loses “power” as $\varepsilon \to 0$, namely $\int_{-\infty}^{\infty} C^\varepsilon(t)dt \to 0$. However, if we re-scale the process y^ε so its variance increases as ε^2, then we can get something interesting in the limit:[1]

$$\frac{1}{\varepsilon^2}C^\varepsilon(t) = \mathbb{E}\frac{y^\varepsilon(0)}{\varepsilon}\frac{y^\varepsilon(t)}{\varepsilon} \quad \xrightarrow{\varepsilon\to 0} \quad \frac{\sigma^2}{\alpha^2}\delta(t).$$

Therefore, heuristically we would expect $\frac{y^\varepsilon}{\varepsilon}$ to converge in some sense to $\frac{\sigma}{\alpha}\eta$, where η is a white noise.

However, just substituting this heuristic into (1), to get $dX_t = h(X_t)dt + \frac{\sigma}{\alpha}f(X_t)dW_t$, is WRONG! The correct result is the Stratonovich interpretation:

$$dX_t = h(X_t) + \frac{\sigma}{\alpha}f(X_t)\circ dW_t. \tag{4}$$

We will show this using singular perturbation theory.

13.1.1 Limit of (1), (3) as $\varepsilon \to 0$

To derive this limit, we consider the generator of the process $(x^\varepsilon, y^\varepsilon)$, and show that, formally at least, this converges to the generator of the limit process (4). This is a very weak convergence result: it is a statement about the asymptotic equivalence of generators, which is not as strong as weak convergence of the processes (convergence of the transition probabilities), which in turn is not as strong as strong convergence (convergence of paths.) Nevertheless, this is a very useful technique, frequently used in applications, and often these results can be strengthened to a stronger form of convergence. See Pavliotis and Stuart [2008], for a wide range of applications of this technique, as well as stronger convergence results.

The generator of the process $(x^\varepsilon(t), y^\varepsilon(t))$ is

$$\mathscr{L}^\varepsilon = \frac{1}{\varepsilon^2}(-\alpha y\partial_y + \frac{1}{2}\sigma^2\partial_{yy}) + \frac{1}{\varepsilon}f(x)y\partial_x + h(x)\partial_x.$$

The backward equation is

$$\frac{\partial u^\varepsilon}{\partial t} = \left(\frac{1}{\varepsilon^2}\mathscr{L}_0 + \frac{1}{\varepsilon}\mathscr{L}_1 + \mathscr{L}_2\right)u^\varepsilon, \tag{5}$$

where $\mathscr{L}_0 = -\alpha y\partial_y + \frac{\sigma^2}{2}\partial_{yy}$, $\mathscr{L}_1 = f(x)y\partial_x$, $\mathscr{L}_2 = h(x)\partial_x$.

Let's expand u^ε in an asymptotic expansion, as

$$u^\varepsilon(x,y,t) = u_0(x,y,t) + \varepsilon u_1(x,y,t) + \varepsilon^2 u_2(x,y,t) + \cdots$$

Now we substitute this ansatz into the backward equation (5), and equate terms of each order:

$$\begin{array}{rrcl}
O(\frac{1}{\varepsilon^2}): & -\mathscr{L}_0 u_0 & = & 0 \\
O(\frac{1}{\varepsilon}): & -\mathscr{L}_0 u_1 & = & \mathscr{L}_1 u_0 \\
O(1): & -\mathscr{L}_0 u_2 & = & \mathscr{L}_1 u_1 + \mathscr{L}_2 u_2 - \frac{\partial u_0}{\partial t}
\end{array}$$

Now we solve these, order-by-order.

[1] We are being very heuristic about the use of convergence, but you can define a particular kind of convergence to make these statements more rigorous.

At $O(\frac{1}{\varepsilon^2})$, we have

$$-\mathscr{L}_0 u_0 = 0 \qquad \Rightarrow \qquad -\alpha y \partial_y u_0 + \frac{\sigma^2}{2}\partial_{yy} u_0 = 0.$$

This is an operator only in y, the "fast" variable.[2] The solution is $u_0 =$ constant (in y), so $u_0 = u_0(x,t)$. This is the only solution ($\mathscr{L}_0$ is ergodic.)

At $O(\frac{1}{\varepsilon})$, we have

$$-\mathscr{L}_0 u_1 = \mathscr{L}_1 u_0 \qquad \Rightarrow \qquad -\alpha y \partial_y u_1 + \frac{\sigma^2}{2}\partial_{yy} u_1 = y f(x) \partial_x u_0.$$

We can solve this, for example by separation of variables, to get

$$u_1(x,y,t) = \frac{y}{\varepsilon} f(x)\partial_x u_0 + \Psi_1(x,t),$$

where Ψ_1 is some unknown function. Note that we can drop it, since it is in the null space of $\mathscr{L}_0$ so we could absorb it in u_0, but we'll keep it anyways and show it doesn't matter in the end.

At $O(1)$, we have

$$-\mathscr{L}_0 u_2 = \mathscr{L}_1 u_1 + \mathscr{L}_2 u_0 - \frac{\partial u_0}{\partial t}. \tag{6}$$

To be able to solve for u_2, the RHS must satisfy the *Fredholm Alternative*: if π is in the null space of $\mathscr{L}_0^*$, i.e. $\mathscr{L}_0^* \pi = 0$, then $\langle \pi, RHS \rangle = 0$.

To see why, take the inner product of π with (6):

$$\langle \pi, -\mathscr{L}_0 u_2 \rangle = \langle \mathscr{L}_0^* \pi, -u_2 \rangle = 0 \quad \Rightarrow \quad \langle \pi, RHS \rangle = 0.$$

Therefore

$$\int \pi(y) \frac{\partial u_0}{\partial t} dy = \int \pi(y) (\mathscr{L}_1 u_1 + \mathscr{L}_2 u_0) dy,$$

where $\pi(y)$ is the stationary distribution for the process with generator $\mathscr{L}_0$. Writing this out explicitly, we have

$$\begin{aligned} \frac{\partial u_0}{\partial t} &= \int \pi(y) \left[y f(x) \partial_x \left(\frac{y}{\alpha} f(x) \partial_x u_0(x,t) + \Psi_1(x,t) \right) + h(x) \partial_x u_0(x,t) \right] dy \\ &= \frac{f(x)}{\alpha} \partial_x (f(x) \partial_x u_0) \underbrace{\int y^2 \pi(y) dy}_{\mathbb{E}^\pi y^2} + [\cdots \Psi_1 \cdots] \underbrace{\int y \pi(y) dy}_{\mathbb{E}^\pi y} + h(x) \partial_x u_0. \end{aligned}$$

We can solve explicitly for the stationary density, since $\mathscr{L}_0$ is just the generator of an Ornstein-Uhlenbeck process: $\pi(y) = \frac{1}{\sqrt{2\pi\sigma^2/2\alpha}} e^{-\frac{1}{2}\frac{y^2}{\sigma^2/2\alpha}}$. Therefore $\mathbb{E}^\pi y^2 = \frac{\sigma^2}{2\alpha}$, $\mathbb{E}^\pi y = 0$. This gives an evolution equation for u_0 as

$$\frac{\partial u_0}{\partial t} = \frac{\sigma^2}{2\alpha^2} f(x) \partial_x (f(x) \partial_x u_0) + h(x) \partial_x u_0 \ = \ \left(h(x) + \frac{\sigma^2}{2\alpha^2} f(x) \partial_x f(x) \right) \partial_x u_0 + \frac{\sigma^2}{2\alpha^2} f^2(x) \partial_{xx} u_0.$$

[2] We call y the "fast" variable because it evolves very quickly, whereas x is expected to change on much slower timescales . We are interested in the dynamics of x, on the slow timescale, so we want to "average over" the fast variables in some way.

This is the backward equation for the Stratonovich SDE

$$dX_t = h(X_t) + \frac{\sigma}{\alpha} f(X_t) \circ dW_t.$$

Notes

- In general, a coloured noise $y(t)$ will converge to the Stratonovich increment when scaled as $\frac{1}{\varepsilon} y(t/\varepsilon^2)$, provided y is scalar.
- In higher dimensions, this will not necessarily converge to the Stratonovich SDE, but rather an extra drift term could arise – this is called the Levy area correction. See Pavliotis [2014], section 5.1, for a detailed example. See also Pavliotis and Stuart [2008], section 11.7.7, p. 176.
- The Fokker-Planck operator can also be used to do the asymptotics – see e.g. Gardiner [2009].
- A similar derivation works provided $y^\varepsilon(t)$ is *any* Markov process – we simply use the generator $\mathscr{A}$ of y^ε, instead of the OU generator. We still obtain the form $\mathscr{L} = \frac{1}{\varepsilon^2}\mathscr{L}_0 + \frac{1}{\varepsilon}\mathscr{L}_1 + \mathscr{L}_2$.
- To make these asymptotics work in general, we need the following assumptions Papanicolaou [1977]:
  - (i) $\mathscr{L}_0$ is the generator of a stationary Markov process, and it depends only on the fast variables, y.
  - (ii) $\mathscr{L}_0$ is ergodic: it has only constants in its null space, and the semigroup $T_t = e^{\mathscr{L}_0 t}$ converges (in some sense) to $P =$ the projection operator onto the null space of $\mathscr{L}_0$: $e^{\mathscr{L}_0 t} \to P$, $Pu = \int u(x,y)dy$.
  - (iii) The Fredholm Alternative holds for $\mathscr{L}_1$: $P\mathscr{L}_1 P = 0 \iff \langle \pi, \mathscr{L}_1 \rangle = 0$.
  - (iv) Consistency in the initial condition: $Pu|_{t=0} = u|_{t=0}$ (otherwise we must consider the initial layer problem)

  If these hold, we can solve for u_0 as

  $$\frac{\partial u_0}{\partial t} = (P\mathscr{L}_2 P - P\mathscr{L}_1 \mathscr{L}_0^{-1} \mathscr{L}_1 P) u_0,$$

  where $\mathscr{L}_0^{-1} = -\int_0^\infty (e^{\mathscr{L}_0 t} - P) dt$.
- Similar results hold for non-Markov processes $y(t)$. See papers by Papanicolaou, Khasminskii, etc.

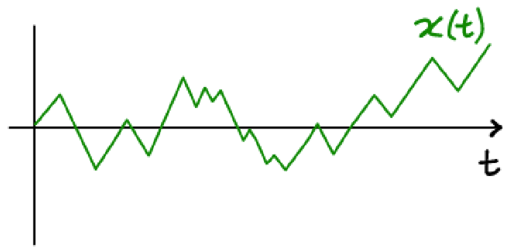
13.2 Markov chain → Diffusion process

Consider the equation

$$\frac{dx^\varepsilon}{dt} = \frac{1}{\varepsilon} y^\varepsilon(t), \tag{7}$$

where $y^\varepsilon = y(t/\varepsilon^2)$ is a 2-state continuous-time Markov chain. It takes values $y = \pm\alpha$, and jumps between them with rate β.

This describes a random walk, with exponentially distributed jump times. If the jump times and sizes are scaled in the right way, then we would expect this to look like a Brownian motion as $\varepsilon \to 0$. Let's make some estimates.



- Tme between jumps is $\Delta t = \beta^{-1}\varepsilon^2$.
- Size of jumps is $\Delta x = \frac{\alpha}{\varepsilon}\Delta t = \alpha\beta^{-1}\varepsilon$.
- Therefore $\frac{\Delta x^2}{\Delta t} = \alpha^2\beta^{-1}$.

This is consistent with the scaling of a Brownian motion, so we would expect the limiting process to behave as a Brownian motion with diffusivity $\alpha\beta^{-1/2}$ as $\varepsilon \to 0$. Let's show this using an asymptotic analysis of the generator.

The generator of the unscaled jump process $y(t)$ is

$$A = \begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix}.$$

The backward equation for $u(x,y,t) = \mathbb{E}^{(x,y)} f(x(t),y(t))$ is

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon} y \frac{\partial u}{\partial x} + \frac{1}{\varepsilon^2} Au.$$

Define

$$u_+(x,t) = u(x,\alpha,t), \quad u_-(x,t) = u(x,-\alpha,t).$$

The backward equation can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \frac{1}{\varepsilon}\begin{pmatrix} \alpha & 0 \\ 0 & -\alpha \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} u_+ \\ u_- \end{pmatrix} + \frac{1}{\varepsilon^2}\begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix}\begin{pmatrix} u_+ \\ u_- \end{pmatrix}, \tag{8}$$

with initial condition $u_\pm(x,0) = f_\pm(x)$. We could now perform the asymptotic analysis on (8) in exactly the same way as in section 13.1, see e.g. E et al. [2014]. We won't do this, but rather pursue an alternative strategy which is to change variables, to make (8) simpler to analyze. Let

$$w = u_+ + u_-, \quad v = u_+ - u_-.$$

Take one more time derivative of (8), and substitute the change of variables. The equation for w is

$$\varepsilon^2 \frac{\partial^2 w}{\partial t^2} = \alpha^2 \frac{\partial^2 w}{\partial x^2} - 2\beta\frac{\partial w}{\partial t}, \qquad w(x,0) = f_+(x) + f_-(x), \quad \frac{\partial w}{\partial t}(x,0) = \frac{\alpha}{\varepsilon}\frac{\partial}{\partial x}(f_+ - f_-). \tag{9}$$

This equation is almost decoupled from that of v; the only coupling enters through the initial condition. Suppose $f_+ = f_- = f(x)$. Then $\frac{\partial w}{\partial t}(x,0) = 0$, so there is no initial transient layer; the equations are fully decoupled.

Now we can consider an asymptotic analysis of (9). Let $w = w_0 + \varepsilon w_1 + \varepsilon^2 w_2 + \ldots$. Substitute this ansatz into (9) and equate terms of different orders. At $O(1)$, we get

$$\frac{\partial w_0}{\partial t} = \frac{\alpha^2}{2\beta}\frac{\partial^2 w_0}{\partial x^2}, \qquad w_0(x,0) = 2f(x). \tag{10}$$

This is the generator of the diffusion

$$dX_t = \frac{\alpha}{\sqrt{\beta}} dW_t. \tag{11}$$

Therefore $x^{\varepsilon}(t)$ behaves asymptotically as the scaled Brownian motion $\frac{\alpha}{\sqrt{\beta}} W_t$. See Papanicolaou [1977], E et al. [2014].

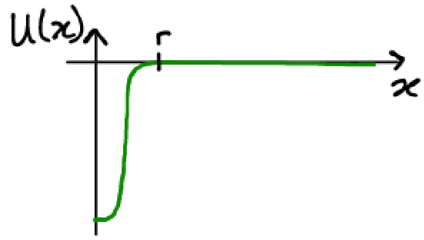
13.3 "Sticky" boundary condition

We now show how a sticky boundary condition can come from Brownian dynamics, when there is a very narrow, deep potential near a boundary (based on Holmes-Cerfon et al. [2013].)

Given

$$dX_t = -\nabla U(X_t)dt + \sqrt{2}dW_t, \qquad X_t \in [0,1] \quad \text{(reflecting boundary)}$$

where the potential $U(x)$ is deep near $x = 0$, but $= 0$ beyond some cutoff at $x = r$:



Let's scale $U(x)$ to be very narrow, and moderately deep:

$$U^{\varepsilon}(x) = C_{\varepsilon} U(x/\varepsilon), \qquad \varepsilon \ll 1.$$

Here the parameter ε scales the width of the potential, and C_{ε} scales the depth. It is chosen so that

$$\frac{e^{-C_{\varepsilon}U(0)}}{\sqrt{C_{\varepsilon}U''(0)}} \sqrt{\frac{\pi}{2}} = \kappa,$$

where κ is an $O(1)$ constant. Roughly, it goes as $C_{\varepsilon} \sim |\log \varepsilon|$. Physically, what this means is the partition function (integral of the Boltzmann factor) associated with the minimum of the potential is constant.

The Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = \partial_x(\partial_x U^{\varepsilon}(x)p + \partial_x p). \tag{12}$$

We will analyze this asymptotically as $\varepsilon \to 0$ using boundary-layer theory.

First consider the "outer" solution for $x > \varepsilon r$: this solves $\partial_t \tilde{p} = \partial_{xx}\tilde{p}$, with a reflecting boundary condition at $x = 1$. The boundary condition at $x = \varepsilon r \to 0$ will be determined by the matching condition with the solution near the wall.

Next consider the "inner" solution, near $x = 0$. Let $X = x/\varepsilon$ describe the variable near the wall. The FP equation is

$$\frac{\partial p}{\partial t} = \frac{1}{\varepsilon^2}\partial_X(\partial_X C_\varepsilon U(X)p + \partial_X p).$$

Let $p^{wall} = p_0(X,t) + \varepsilon p_1(X,t) + \ldots$. Substitute the ansatz into the FPE and equate terms of the same order.

The $O(\varepsilon^{-2})$ equation shows

$$p_0(X,t) = p_0(t)e^{-C_\varepsilon U(X)}.$$

The inner solution p^{wall} must be matched to the outer solution $\tilde{p}$. The matching conditions are the same as those near a discontinuity:

(a) probability is continuous: $p^{wall}|_{X=r} = \tilde{p}|_{x=0}$.
(b) flux is continuous: $\varepsilon j^{wall}|_{X=r} = \tilde{j}|_{x=0}$.

Matching condition (a) shows that $\tilde{p}(0,t) = p_0(t)$.

Applying (b) directly is hard, since it requires evaluate p_1, p_2. Instead, we use an alternate form: the total probability is conserved. This requires that

$$\begin{aligned} &\frac{d}{dt}\int_0^{\varepsilon r} p^{wall} dx = \tilde{j}|_{x=0} \qquad \text{(to leading order in } \varepsilon) \\ \Rightarrow \quad &\int_0^{\varepsilon r} p_0'(t)e^{-C_\varepsilon U(x/\varepsilon)}dx = \tilde{p}_x|_{x=0} \\ \Rightarrow \quad &\kappa p_0'(t) = \tilde{p}_x|_{x=0}. \end{aligned}$$

Combining with (a) gives a boundary condition for $\tilde{p}$:

$$\tilde{p}_x|_{x=0} = \kappa\partial_t\tilde{p}|_{x=0} \quad \Leftrightarrow \quad \tilde{p}_x|_{x=0} = \kappa\tilde{p}_{xx}|_{x=0}.$$

This is the sticky boundary condition $\tilde{j}\cdot\hat{n} = \mathscr{L}\tilde{p}$.

References

W. E, T. Li, and E. Vanden-Eijnden. *Applied Stochastic Analysis*. In preparation., 2014.

C. Gardiner. *Stochastic methods: A handbook for the natural sciences*. Springer, 4th edition, 2009.

M. C. Holmes-Cerfon, S. J. Gortler, and M. P. Brenner. A geometrical approach to computing free energy landscapes from short-ranged potentials. *Proc. Natl. Acad. Sci.*, 110(1), 2013.

G. C. Papanicolaou. Introduction to asymptotic analysis of stochastic equations. In *Modern modeling of continuum phenomena*, volume 16 of *Lectures in Appl. Math.*, pages 109–147. Amer. Math. Soc., 1977.

G. A. Pavliotis. *Stochastic Processes and Applications*. Springer, 2014.

G. A. Pavliotis and A. M. Stuart. *Multiscale methods: Averaging and homogenization*. Springer, 2008.