Lecture 18. Multiscale Analysis of SDEs

Tiejun Li[1,2]

[1]School of Mathematical Sciences (SMS),
&
[2]Center for Machine Learning Research (CMLR),
Peking University,
Beijing 100871,
P.R. China
tieli@pku.edu.cn

Office: No. 1 Science Building, Room 1376E

Table of Contents

Deterministic multiscale models

The multiscale is very common in different fields of science and engineering.

- Consider the toy model

$$\frac{dx}{dt} = f(x, y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon}(g(x) - y), \quad \epsilon \ll 1, \epsilon > 0.$$

We call x the slow variable and y the fast variable.

Deterministic multiscale models

The multiscale is very common in different fields of science and engineering.

- Consider the toy model

$$\frac{dx}{dt} = f(x, y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon}(g(x) - y), \quad \epsilon \ll 1, \epsilon > 0.$$

  We call x the slow variable and y the fast variable.

- The exact solution of y given x has the form

$$y(t) = e^{-t/\epsilon} y_0 + (1 - e^{-t/\epsilon}) g(x) \to g(x)$$

  as $t \to \infty$. That is, y will relax to $y = g(x)$ fast in $O(\epsilon)$ timescale. $y = g(x)$ is called the slow manifold.

Deterministic multiscale models

The multiscale is very common in different fields of science and engineering.

- Consider the toy model

$$\frac{dx}{dt} = f(x, y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon}(g(x) - y), \quad \epsilon \ll 1, \epsilon > 0.$$

  We call x the slow variable and y the fast variable.

- The exact solution of y given x has the form

$$y(t) = e^{-t/\epsilon} y_0 + (1 - e^{-t/\epsilon}) g(x) \to g(x)$$

  as $t \to \infty$. That is, y will relax to $y = g(x)$ fast in $O(\epsilon)$ timescale. $y = g(x)$ is called the slow manifold.

- Finally we get the adiabatic approximation:

$$\frac{dx}{dt} = f(x, g(x))$$

  as $\epsilon \to 0$.

Simple Stochastic Multiscale models

- A slight generalization is

$$\frac{dx}{dt} = f(x, y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon}(g(x) - y) + \sqrt{\frac{2}{\epsilon}}\dot{W}, \quad \epsilon \ll 1, \epsilon > 0.$$

Simple Stochastic Multiscale models

- A slight generalization is

$$\frac{dx}{dt} = f(x,y)$$
$$\frac{dy}{dt} = \frac{1}{\epsilon}(g(x)-y) + \sqrt{\frac{2}{\epsilon}}\dot{W}, \quad \epsilon \ll 1, \epsilon > 0.$$

- Given x, $y(t)$ has an invariant distribution

$$y(t) \sim N(g(x),1) := \mu_{g(x)}(y)dy$$

The effective dynamics is

$$\frac{dx}{dt} = \langle f(x,y)\rangle_{\mu_{g(x)}} = \int_{\mathbb{R}} f(x,y)\mu_{g(x)}(y)dy$$

as $\epsilon \to 0$.

Table of Contents

Diffusive Limit of a Simple Eample

For SDEs or ODEs, the presence of a small parameter usually means that the system has some disparate time scales. Our task is to eliminate the fast time scales in the system and derive effective equations that govern the dynamics on the slow time scale.

- Let us start with a simple example. Let $Y_t = y(t)$ be a stationary two-state Markov jump process taking values $\pm\alpha$ with jump rate β between these two states. With matrix notation, the infinitesimal generator for Y has the form

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

Diffusive Limit of a Simple Eample

For SDEs or ODEs, the presence of a small parameter usually means that the system has some disparate time scales. Our task is to eliminate the fast time scales in the system and derive effective equations that govern the dynamics on the slow time scale.

- Let us start with a simple example. Let $Y_t = y(t)$ be a stationary two-state Markov jump process taking values $\pm\alpha$ with jump rate β between these two states. With matrix notation, the infinitesimal generator for Y has the form

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

- Let $y^\epsilon(t) = y(t/\epsilon^2)$ where ϵ is a small parameter. Consider the SDE

$$\frac{dx^\epsilon(t)}{dt} = \frac{1}{\epsilon} y^\epsilon(t), \quad x^\epsilon(0) = x.$$

Backward Operator Analysis

- Let $u^\epsilon(x,y,t) = \mathbb{E}^{(x,y)}\Big(f(x^\epsilon(t), y^\epsilon(t))\Big)$, where f is any given smooth function. Then u^ϵ satisfies the backward Kolmogorov equation:

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{\epsilon} y \frac{\partial u^\epsilon}{\partial x} + \frac{1}{\epsilon^2} A u^\epsilon, \quad u^\epsilon(x,y,0) = f(x,y).$$

Backward Operator Analysis

- Let $u^\epsilon(x,y,t) = \mathbb{E}^{(x,y)}\Big(f(x^\epsilon(t), y^\epsilon(t))\Big)$, where f is any given smooth function. Then u^ϵ satisfies the backward Kolmogorov equation:

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{\epsilon} y \frac{\partial u^\epsilon}{\partial x} + \frac{1}{\epsilon^2} A u^\epsilon, \quad u^\epsilon(x,y,0) = f(x,y).$$

- Since y can only take two values, by defining

$$u_\pm(x,t) = u^\epsilon(x, \pm\alpha, t), \;\; f_\pm(x,t) = f(x, \pm\alpha),$$

we can rewrite the above equation as

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

with initial condition $u_\pm(x,0) = f_\pm(x)$.

Asymptotic expansion

- Let $w = u_+ + u_-$, we have

$$\epsilon^2 \frac{\partial^2 w}{\partial t^2} = \alpha^2 \frac{\partial^2 w}{\partial x^2} - 2\beta \frac{\partial w}{\partial t}, \;\; w|_{t=0} = f_+ + f_-, \; \partial_t w|_{t=0} = \frac{\alpha}{\epsilon} \partial_x (f_+ - f_-).$$

Asymptotic expansion

- Let $w = u_+ + u_-$, we have

$$\epsilon^2 \frac{\partial^2 w}{\partial t^2} = \alpha^2 \frac{\partial^2 w}{\partial x^2} - 2\beta \frac{\partial w}{\partial t}, \;\; w|_{t=0} = f_+ + f_-, \; \partial_t w|_{t=0} = \frac{\alpha}{\epsilon} \partial_x (f_+ - f_-).$$

- Consider the case when $f_+ = f_- = f$. In this case the time derivative of w vanishes at $t = 0$, hence we avoid the extra complication coming from the initial layer.

Asymptotic expansion

- Let $w = u_+ + u_-$, we have

$$\epsilon^2 \frac{\partial^2 w}{\partial t^2} = \alpha^2 \frac{\partial^2 w}{\partial x^2} - 2\beta \frac{\partial w}{\partial t}, \quad w|_{t=0} = f_+ + f_-, \; \partial_t w|_{t=0} = \frac{\alpha}{\epsilon} \partial_x (f_+ - f_-).$$

- Consider the case when $f_+ = f_- = f$. In this case the time derivative of w vanishes at $t = 0$, hence we avoid the extra complication coming from the initial layer.
- Following the standard approach in asymptotic analysis, we make the ansatz: $w = w_0 + \epsilon w_1 + \epsilon^2 w_2 + \cdots$. To leading order, this gives:

$$\frac{\partial w_0}{\partial t} = \frac{\alpha^2}{2\beta} \frac{\partial^2 w_0}{\partial x^2}, \qquad w_0|_{t=0} = 2f.$$

Asymptotic expansion

- Let $w = u_+ + u_-$, we have

$$\epsilon^2 \frac{\partial^2 w}{\partial t^2} = \alpha^2 \frac{\partial^2 w}{\partial x^2} - 2\beta \frac{\partial w}{\partial t}, \quad w|_{t=0} = f_+ + f_-, \; \partial_t w|_{t=0} = \frac{\alpha}{\epsilon} \partial_x (f_+ - f_-).$$

- Consider the case when $f_+ = f_- = f$. In this case the time derivative of w vanishes at $t = 0$, hence we avoid the extra complication coming from the initial layer.
- Following the standard approach in asymptotic analysis, we make the ansatz: $w = w_0 + \epsilon w_1 + \epsilon^2 w_2 + \cdots$. To leading order, this gives:

$$\frac{\partial w_0}{\partial t} = \frac{\alpha^2}{2\beta} \frac{\partial^2 w_0}{\partial x^2}, \qquad w_0|_{t=0} = 2f.$$

- This means that to leading order, x^ϵ behaves like Brownian motion with diffusion constant $D = \alpha/\sqrt{\beta}$. This is not surprising since it is what the central limit theorem tells us about

$$x^\epsilon(t) = \frac{1}{\epsilon} \int_0^t y^\epsilon(s) ds \quad \text{as } \epsilon \to 0.$$

General Framework

We turn now to the general case. Suppose the stochastic process $\boldsymbol{X}_t^\epsilon$ possess the backward equation for $u^\epsilon(\boldsymbol{x}, t) = \mathbb{E}^{\boldsymbol{x}} f(\boldsymbol{X}_t^\epsilon)$ as

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{\epsilon^2}\mathcal{L}_1 u^\epsilon + \frac{1}{\epsilon}\mathcal{L}_2 u^\epsilon + \mathcal{L}_3 u^\epsilon, \quad u^\epsilon(0) = f,$$

where $\mathcal{L}_1, \mathcal{L}_2$ and $\mathcal{L}_3$ are differential operators defined on some Banach space B, whose properties will be specified below. As a general framework we assume that the following conditions hold.

(a) $\mathcal{L}_1$ is an infinitesimal generator of a stationary Markov process, and the semi-group $\exp(\mathcal{L}_1 t)$ generated by $\mathcal{L}_1$ converges to a projection operator to the null space of $\mathcal{L}_1$, which we will denote as P.

$$\exp(\mathcal{L}_1 t) \to P, \quad t \to \infty.$$

General Framework

We turn now to the general case. Suppose the stochastic process $\boldsymbol{X}_t^\epsilon$ possess the backward equation for $u^\epsilon(\boldsymbol{x}, t) = \mathbb{E}^{\boldsymbol{x}} f(\boldsymbol{X}_t^\epsilon)$ as

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{\epsilon^2}\mathcal{L}_1 u^\epsilon + \frac{1}{\epsilon}\mathcal{L}_2 u^\epsilon + \mathcal{L}_3 u^\epsilon, \quad u^\epsilon(0) = f,$$

where $\mathcal{L}_1, \mathcal{L}_2$ and $\mathcal{L}_3$ are differential operators defined on some Banach space B, whose properties will be specified below. As a general framework we assume that the following conditions hold.

(a) $\mathcal{L}_1$ is an infinitesimal generator of a stationary Markov process, and the semi-group $\exp(\mathcal{L}_1 t)$ generated by $\mathcal{L}_1$ converges to a projection operator to the null space of $\mathcal{L}_1$, which we will denote as P.

$$\exp(\mathcal{L}_1 t) \to P, \quad t \to \infty.$$

(b) Solvability condition: $P\mathcal{L}_2 P = 0$.

General Framework

We turn now to the general case. Suppose the stochastic process $\boldsymbol{X}_t^\epsilon$ possess the backward equation for $u^\epsilon(\boldsymbol{x}, t) = \mathbb{E}^{\boldsymbol{x}} f(\boldsymbol{X}_t^\epsilon)$ as

$$\frac{\partial u^\epsilon}{\partial t} = \frac{1}{\epsilon^2}\mathcal{L}_1 u^\epsilon + \frac{1}{\epsilon}\mathcal{L}_2 u^\epsilon + \mathcal{L}_3 u^\epsilon, \quad u^\epsilon(0) = f,$$

where $\mathcal{L}_1, \mathcal{L}_2$ and $\mathcal{L}_3$ are differential operators defined on some Banach space B, whose properties will be specified below. As a general framework we assume that the following conditions hold.

(a) $\mathcal{L}_1$ is an infinitesimal generator of a stationary Markov process, and the semi-group $\exp(\mathcal{L}_1 t)$ generated by $\mathcal{L}_1$ converges to a projection operator to the null space of $\mathcal{L}_1$, which we will denote as P.

$$\exp(\mathcal{L}_1 t) \to P, \quad t \to \infty.$$

(b) Solvability condition: $P\mathcal{L}_2 P = 0$.

(c) Consistency condition for the initial value: $Pf = f$.

Asymptotic expansion

- Assume that u^ϵ can be expressed in the following form:

$$u^\epsilon = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \dots.$$

Substituting it into the above PDE and collecting terms of the same order in ϵ, we get

$$\begin{aligned} O(\epsilon^{-2}): &\quad \mathcal{L}_1 u_0 = 0, \\ O(\epsilon^{-1}): &\quad \mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0, \\ O(\epsilon^{0}): &\quad \mathcal{L}_1 u_2 = -\mathcal{L}_2 u_1 - \mathcal{L}_3 u_0 + \frac{\partial u_0}{\partial t}, \end{aligned}$$

and $u_0(0) = f$ from the initial condition.

Asymptotic expansion

- Assume that u^ϵ can be expressed in the following form:

$$u^\epsilon = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \dots.$$

Substituting it into the above PDE and collecting terms of the same order in ϵ, we get

$$\begin{aligned} O(\epsilon^{-2}): &\quad \mathcal{L}_1 u_0 = 0, \\ O(\epsilon^{-1}): &\quad \mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0, \\ O(\epsilon^{0}): &\quad \mathcal{L}_1 u_2 = -\mathcal{L}_2 u_1 - \mathcal{L}_3 u_0 + \frac{\partial u_0}{\partial t}, \end{aligned}$$

and $u_0(0) = f$ from the initial condition.

- We obtain from the $O(\epsilon^{-2})$ term that u_0 is in the null space of $\mathcal{L}_1$, which is the same as the range of P (see the following pages for reason), i.e.

$$P u_0 = u_0.$$

Asymptotic expansion

- Assume that u^ϵ can be expressed in the following form:

$$u^\epsilon = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \dots.$$

Substituting it into the above PDE and collecting terms of the same order in ϵ, we get

$$\begin{aligned} O(\epsilon^{-2}): &\quad \mathcal{L}_1 u_0 = 0, \\ O(\epsilon^{-1}): &\quad \mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0, \\ O(\epsilon^{0}): &\quad \mathcal{L}_1 u_2 = -\mathcal{L}_2 u_1 - \mathcal{L}_3 u_0 + \frac{\partial u_0}{\partial t}, \end{aligned}$$

and $u_0(0) = f$ from the initial condition.

- We obtain from the $O(\epsilon^{-2})$ term that u_0 is in the null space of $\mathcal{L}_1$, which is the same as the range of P (see the following pages for reason), i.e.

$$P u_0 = u_0.$$

- The consistency condition $Pf = f$, i.e. $Pu_0(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}, 0)$ allows us to avoid the initial layer problem.

Supplemental: Fredholm Alternative

- Fredholm Alternative — Finite dimensional case.

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}$$

  - Either: $Ax = b$ has a solution x
  - Or: $A^T y = 0$ has a nontrivial solution y with $\langle y, b \rangle = y^T b \neq 0$.

  Key point: $\mathbb{R}^m = R(A) \oplus N(A^T)$

Supplemental: Fredholm Alternative

- Fredholm Alternative — Finite dimensional case.

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}$$

  - Either: $Ax = b$ has a solution x
  - Or: $A^T y = 0$ has a nontrivial solution y with $\langle y, b \rangle = y^T b \neq 0$.

Key point: $\mathbb{R}^m = R(A) \oplus N(A^T)$

Either b is in the range of A or it has a nontrivial projection in the null space of A^T.

Supplemental: Fredholm Alternative

- Fredholm Alternative — Finite dimensional case.

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}$$

  - Either: $Ax = b$ has a solution x
  - Or: $A^T y = 0$ has a nontrivial solution y with $\langle y, b \rangle = y^T b \neq 0$.

  Key point: $\mathbb{R}^m = R(A) \oplus N(A^T)$

  Either b is in the range of A or it has a nontrivial projection in the null space of A^T.

- It can be also stated as

$$Ax = b \text{ has a solution if and only if}$$

for any $y \in N(A^T)$, we have $\langle y, b \rangle = 0$.

Supplemental: Fredholm Alternative

- Fredholm Alternative — Extension to infinite dimensional space

$$\mathcal{L}u = f, \quad \mathcal{L} : V \to H$$

  - Either: $\mathcal{L}u = f$ has a solution u
  - Or: $\mathcal{L}^* v = 0$ has a nontrivial solution v with $\langle v, f \rangle \neq 0$.

[1] L.C. Evans, PDE, AMS.

Supplemental: Fredholm Alternative

- Fredholm Alternative — Extension to infinite dimensional space

$$\mathcal{L}u = f, \quad \mathcal{L} : V \to H$$

  - Either: $\mathcal{L}u = f$ has a solution u
  - Or: $\mathcal{L}^* v = 0$ has a nontrivial solution v with $\langle v, f \rangle \neq 0$.

- It can be also stated as

$$\mathcal{L}u = f \text{ has a solution iff } \forall v \in N(\mathcal{L}^*), \exists \langle v, f \rangle = 0$$

Typical Examples:[1]

- $\mathcal{L} = \mathcal{I} - \mathcal{K}$, where $\mathcal{K}$ is a compact operator.
  Fact: $R(\mathcal{L}) = N(\mathcal{L}^*)^\perp$
- Elliptic operator with coercive condition for the diffusion matrix.

[1] L.C. Evans, PDE, AMS.

Supplemental: Properties of the limit of Markov semigroups

Consider a Markov semigroup

$$S(t) = \exp(\mathcal{L}t), \quad t \geq 0,$$

where $\mathcal{L}$ is the infinitesimal generator. Denote

$$P = \lim_{t \to +\infty} S(t).$$

We have the following important properties [2]

1. $P^2 = P, \ \|P\| \leq 1.$
2. $S(t)P = PS(t) = P$
3. $\mathcal{L}P = P\mathcal{L} = 0$
4. $R(P) = N(\mathcal{L})$
5. $N(P) = \overline{R(\mathcal{L})}$

[2] Ethier and Kurtz, Markov processes: Characterization and Convergence, Wiley.

Supplemental: Solvability of the equation in $O(\epsilon^{-1})$ term

- The equation in $O(\epsilon^{-1})$ term reads

$$\mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0.$$

  We assume the solvability condition $P\mathcal{L}_2 P = 0$.

- Note that from the $O(\epsilon^{-2})$ term: $\mathcal{L}_1 u_0 = 0$, we have $u_0 \in N(\mathcal{L}_1) = R(P)$.

Supplemental: Solvability of the equation in $O(\epsilon^{-1})$ term

- The equation in $O(\epsilon^{-1})$ term reads

$$\mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0.$$

  We assume the solvability condition $P\mathcal{L}_2 P = 0$.
- Note that from the $O(\epsilon^{-2})$ term: $\mathcal{L}_1 u_0 = 0$, we have $u_0 \in N(\mathcal{L}_1) = R(P)$.
- From the Fredholm alternative, the $O(\epsilon^{-1})$ term has a solution iff for any $f \in N(\mathcal{L}_1^*)$, we have $\langle f, \mathcal{L}_2 P u_0 \rangle = 0$.

Supplemental: Solvability of the equation in $O(\epsilon^{-1})$ term

- The equation in $O(\epsilon^{-1})$ term reads

$$\mathcal{L}_1 u_1 = -\mathcal{L}_2 u_0.$$

  We assume the solvability condition $P\mathcal{L}_2 P = 0$.
- Note that from the $O(\epsilon^{-2})$ term: $\mathcal{L}_1 u_0 = 0$, we have $u_0 \in N(\mathcal{L}_1) = R(P)$.
- From the Fredholm alternative, the $O(\epsilon^{-1})$ term has a solution iff for any $f \in N(\mathcal{L}_1^*)$, we have $\langle f, \mathcal{L}_2 P u_0 \rangle = 0$.
- This is true since

$$f \in N(\mathcal{L}_1^*) \Rightarrow \langle \mathcal{L}_1^* f, g \rangle = 0, \ \forall g \Rightarrow \langle f, \mathcal{L}_1 g \rangle = 0$$

$$\Rightarrow f \in \overline{R(\mathcal{L}_1)}^{\perp} = N(P)^{\perp} \Rightarrow \forall g \in N(P), \ \langle f, g \rangle = 0$$

  while $\mathcal{L}_2 P u_0 \in N(P)$ since $P\mathcal{L}_2 P = 0$. Done!

Effective equation

- To solve u_1, we assume the Fredholm alternative holds for the operator $\mathcal{L}_1$, which should be rigorously proved for each concrete problem.

Effective equation

- To solve u_1, we assume the Fredholm alternative holds for the operator $\mathcal{L}_1$, which should be rigorously proved for each concrete problem.
- With the solvability condition and the Fredholm alternative, we denote a solution as

$$u_1 = -\mathcal{L}_1^{-1}\mathcal{L}_2 P u_0.$$

Effective equation

- To solve u_1, we assume the Fredholm alternative holds for the operator $\mathcal{L}_1$, which should be rigorously proved for each concrete problem.
- With the solvability condition and the Fredholm alternative, we denote a solution as

$$u_1 = -\mathcal{L}_1^{-1}\mathcal{L}_2 P u_0.$$

- Substituting this into the last equation and applying P on both sides, we obtain the effective equation for the leading order u_0

$$\frac{\partial u_0}{\partial t} = (P\mathcal{L}_3 P - P\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2 P)u_0 := \bar{\mathcal{L}}u_0, \quad u_0(0) = f$$

in the range of P.

Effective equation

- To solve u_1, we assume the Fredholm alternative holds for the operator $\mathcal{L}_1$, which should be rigorously proved for each concrete problem.
- With the solvability condition and the Fredholm alternative, we denote a solution as

$$u_1 = -\mathcal{L}_1^{-1}\mathcal{L}_2 P u_0.$$

- Substituting this into the last equation and applying P on both sides, we obtain the effective equation for the leading order u_0

$$\frac{\partial u_0}{\partial t} = (P\mathcal{L}_3 P - P\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2 P)u_0 := \bar{\mathcal{L}}u_0, \quad u_0(0) = f$$

  in the range of P.
- One can also derive effective equations for the higher order terms u_1, u_2, etc. But it is more complicated and usually not very useful.

Applied to the Simple Example

To see this abstract framework actually works, we use it for the simple model introduced at the beginning of this section. We have

$$\mathcal{L}_1 = A,\ \mathcal{L}_2 = \begin{pmatrix} +\alpha & 0 \\ 0 & -\alpha \end{pmatrix} \frac{\partial}{\partial x},\quad \mathcal{L}_3 = 0.$$

Thus the projection operator P is given by

$$P = \lim_{t\to\infty} \exp(\mathcal{L}_1 t) = \lim_{t\to\infty} \frac{1}{2} \begin{pmatrix} 1+e^{-2\beta t} & 1-e^{-2\beta t} \\ 1-e^{-2\beta t} & 1+e^{-2\beta t} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

- In the current example, we can simply pick a version of $\mathcal{L}_1^{-1}$ as

$$\mathcal{L}_1^{-1} = -\int_0^\infty (\exp(\mathcal{L}_1 t) - P)dt = -\frac{1}{4\beta} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

It is easy to verify that the solvability condition $P\mathcal{L}_2 P = 0$ is satisfied. The consistency condition $Pf = f$ gives $f_+ = f_-$, which we still denote as f.

Applied to the Simple Example

- Finally the effective operator

$$-P\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2P = \frac{\alpha^2}{4\beta}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\partial^2}{\partial x^2}.$$

Applied to the Simple Example

- Finally the effective operator

$$-P\mathcal{L}_2\mathcal{L}_1^{-1}\mathcal{L}_2P = \frac{\alpha^2}{4\beta}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\partial^2}{\partial x^2}.$$

- Combining these we obtain the effective equation

$$\frac{\partial}{\partial t}(u_0^+ + u_0^-) = \frac{\alpha^2}{2\beta}\frac{\partial^2}{\partial x^2}(u_0^+ + u_0^-), \quad (u_0^+ + u_0^-)|_{t=0} = f_+ + f_- = 2f,$$

where $u_0 = (u_0^+, u_0^-)$. Set $w_0 = u_0^+ + u_0^-$, we recover the above result.

Table of Contents

Averaging for Chemical Reaction Kinetics

- Consider a well-stirred system of N molecular species $\{S_1, S_2, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, R_2, \ldots, R_M\}$.

Averaging for Chemical Reaction Kinetics

- Consider a well-stirred system of N molecular species $\{S_1, S_2, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, R_2, \ldots, R_M\}$.
- The following quantities are used to characterize the discrete reaction dynamics.

Averaging for Chemical Reaction Kinetics

- Consider a well-stirred system of N molecular species $\{S_1, S_2, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, R_2, \ldots, R_M\}$.
- The following quantities are used to characterize the discrete reaction dynamics.
  (a) *States* $\boldsymbol{X}_t$.

$$\boldsymbol{X}_t = (X_t^1, X_t^2, \ldots, X_t^N) \in \mathbb{N}^N,$$

where the kth component X_t^k is the number of molecules of species S_k at time t.

Averaging for Chemical Reaction Kinetics

- Consider a well-stirred system of N molecular species $\{S_1, S_2, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, R_2, \ldots, R_M\}$.
- The following quantities are used to characterize the discrete reaction dynamics.
  (a) *States* $\boldsymbol{X}_t$.

$$\boldsymbol{X}_t = (X_t^1, X_t^2, \ldots, X_t^N) \in \mathbb{N}^N,$$

  where the kth component X_t^k is the number of molecules of species S_k at time t.
  (b) *Reactions* $\{R_j\}$. Each reaction channel R_j is characterized by its propensity function $a_j(\boldsymbol{x})$ and its state change vector

$$\boldsymbol{\nu}_j = (\nu_j^1, \nu_j^2, \ldots, \nu_j^N) \in \mathbb{Z}^N,$$

  where $a_j(\boldsymbol{x})dt$ gives the probability that the system will experience a R_j reaction in the next infinitesimal amount of time dt when the current state is $\boldsymbol{X}_t = \boldsymbol{x}$.

Averaging for Chemical Reaction Kinetics

- Consider a well-stirred system of N molecular species $\{S_1, S_2, \ldots, S_N\}$ interacting through M reaction channels $\{R_1, R_2, \ldots, R_M\}$.
- The following quantities are used to characterize the discrete reaction dynamics.
  (a) *States* $\boldsymbol{X}_t$.

$$\boldsymbol{X}_t = (X_t^1, X_t^2, \ldots, X_t^N) \in \mathbb{N}^N,$$

  where the kth component X_t^k is the number of molecules of species S_k at time t.
  (b) *Reactions* $\{R_j\}$. Each reaction channel R_j is characterized by its propensity function $a_j(\boldsymbol{x})$ and its state change vector

$$\boldsymbol{\nu}_j = (\nu_j^1, \nu_j^2, \ldots, \nu_j^N) \in \mathbb{Z}^N,$$

  where $a_j(\boldsymbol{x})dt$ gives the probability that the system will experience a R_j reaction in the next infinitesimal amount of time dt when the current state is $\boldsymbol{X}_t = \boldsymbol{x}$.
- We have $a_j(\boldsymbol{x}) \geq 0$ for physically meaningful states $\boldsymbol{x}$.

Chemical Master Equation

- The Chapman-Kolmogorov equation in the time interval $[t, t+dt)$

$$P(\boldsymbol{x}, t+dt|\boldsymbol{x}_0, t_0) = \sum_{j=1}^{M} P(\boldsymbol{x}-\boldsymbol{\nu}_j, t|\boldsymbol{x}_0, t_0) a_j(\boldsymbol{x}-\boldsymbol{\nu}_j)dt + \left(1 - \sum_{j=1}^{M} a_j(\boldsymbol{x})dt\right) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0),$$

where dt is an infinitesimal time, and we have already omitted the higher order terms in $o(dt)$.

Chemical Master Equation

- The Chapman-Kolmogorov equation in the time interval $[t, t+dt)$

$$P(\boldsymbol{x}, t+dt|\boldsymbol{x}_0, t_0) = \sum_{j=1}^{M} P(\boldsymbol{x}-\boldsymbol{\nu}_j, t|\boldsymbol{x}_0, t_0) a_j(\boldsymbol{x}-\boldsymbol{\nu}_j)dt + \left(1 - \sum_{j=1}^{M} a_j(\boldsymbol{x})dt\right) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0),$$

  where dt is an infinitesimal time, and we have already omitted the higher order terms in $o(dt)$.

- With some algebra, we get the well-known *chemical master equation*.

$$\partial_t P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) = \sum_{j=1}^{M} a_j(\boldsymbol{x}-\boldsymbol{\nu}_j) P(\boldsymbol{x}-\boldsymbol{\nu}_j, t|\boldsymbol{x}_0, t_0) - a_0(\boldsymbol{x}) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)$$

  where $a_0(\boldsymbol{x}) := \sum_{j=1}^{M} a_j(\boldsymbol{x})$ is the total propensity.

Scaling of the Propensity Function

- In physics, we take the large volume scaling:

$$a_j(V\boldsymbol{x}) = \frac{\kappa_j}{V^{|\boldsymbol{\nu}_j^-|-1}}\binom{V\boldsymbol{x}}{\boldsymbol{\nu}_j^-} = V a_j^V(\boldsymbol{x}),$$

where

$$a_j^V(\boldsymbol{x}) = \tilde{a}_j(\boldsymbol{x}) + O(V^{-1}),\ \tilde{a}_j(\boldsymbol{x}) = \frac{\kappa_j}{\boldsymbol{\nu}_j^-!}\boldsymbol{x}^{\boldsymbol{\nu}_j^-} = \kappa_j \prod_{m=1}^{N} \frac{1}{\nu_j^{-,m}!} x_m^{\nu_j^{-,m}}$$

and $a_j^V(\boldsymbol{x}),\ \tilde{a}_j(\boldsymbol{x}) \sim O(1)$ if $\kappa_j, \boldsymbol{x} \sim O(1)$.

Scaling of the Propensity Function

- In physics, we take the large volume scaling:

$$a_j(V\boldsymbol{x}) = \frac{\kappa_j}{V^{|\boldsymbol{\nu}_j^-|-1}}\binom{V\boldsymbol{x}}{\boldsymbol{\nu}_j^-} = V a_j^V(\boldsymbol{x}),$$

  where

$$a_j^V(\boldsymbol{x}) = \tilde{a}_j(\boldsymbol{x}) + O(V^{-1}),\ \tilde{a}_j(\boldsymbol{x}) = \frac{\kappa_j}{\boldsymbol{\nu}_j^-!}\boldsymbol{x}^{\boldsymbol{\nu}_j^-} = \kappa_j \prod_{m=1}^{N} \frac{1}{\nu_j^{-,m}!} x_m^{\nu_j^{-,m}}$$

  and $a_j^V(\boldsymbol{x}),\ \tilde{a}_j(\boldsymbol{x}) \sim O(1)$ if $\kappa_j, \boldsymbol{x} \sim O(1)$.

- For simplicity, we will only consider the case when $a_j^V(\boldsymbol{x}) = \tilde{a}_j(\boldsymbol{x})$. The general case can be analyzed similarly.

Large Volume Limit

Now let us derive the continuum limit of a rescaled process

$$\boldsymbol{X}_t^V := \boldsymbol{X}_t/V.$$

- The backward operator for the original process $\boldsymbol{X}_t$ has the form

$$\mathcal{A}f(\boldsymbol{x}) = \sum_{j=1}^{M} a_j(\boldsymbol{x})\big(f(\boldsymbol{x}+\boldsymbol{\nu}_j) - f(\boldsymbol{x})\big).$$

Large Volume Limit

Now let us derive the continuum limit of a rescaled process

$$\boldsymbol{X}_t^V := \boldsymbol{X}_t/V.$$

- The backward operator for the original process $\boldsymbol{X}_t$ has the form

$$\mathcal{A}f(\boldsymbol{x}) = \sum_{j=1}^{M} a_j(\boldsymbol{x})\big(f(\boldsymbol{x}+\boldsymbol{\nu}_j) - f(\boldsymbol{x})\big).$$

- For the rescaled process $\boldsymbol{X}_t^V$, define $u(\boldsymbol{x},t) = \mathbb{E}^{\boldsymbol{x}} f(\boldsymbol{X}_t^V)$ for any bounded continuous function f. We have

$$\partial_t u(\boldsymbol{x},t) = \mathcal{A}^V u = \sum_{j=1}^{M} V a_j^V(\boldsymbol{x}) \left(u\left(\boldsymbol{x}+\frac{\boldsymbol{\nu}_j}{V}\right) - u(\boldsymbol{x})\right).$$

Large Volume Limit

- Under the scaling assumption and the ansatz $u(\boldsymbol{x}, t) = u_0(\boldsymbol{x}, t) + V^{-1} u_1(\boldsymbol{x}, t) + o(V^{-1})$, we get, to leading order

$$\partial_t u_0(\boldsymbol{x}, t) = \mathcal{A}_0 u_0 = \sum_{j=1}^{M} \tilde{a}_j(\boldsymbol{x}) \boldsymbol{\nu}_j \cdot \nabla u_0(\boldsymbol{x}, t)$$

Large Volume Limit

- Under the scaling assumption and the ansatz $u(\boldsymbol{x}, t) = u_0(\boldsymbol{x}, t) + V^{-1} u_1(\boldsymbol{x}, t) + o(V^{-1})$, we get, to leading order

$$\partial_t u_0(\boldsymbol{x}, t) = \mathcal{A}_0 u_0 = \sum_{j=1}^{M} \tilde{a}_j(\boldsymbol{x}) \boldsymbol{\nu}_j \cdot \nabla u_0(\boldsymbol{x}, t)$$

- This is the backward equation for the reaction rate equations

$$\frac{d}{dt}\boldsymbol{x} = \sum_{j=1}^{M} \tilde{a}_j(\boldsymbol{x}) \boldsymbol{\nu}_j.$$

Over-damped Limit: A Less Trivial Example

Question: Derive the Brownian dynamics from Langevin equations

$$d\boldsymbol{X}_t = \boldsymbol{V}_t dt$$
$$d\boldsymbol{V}_t = (-\nabla U(\boldsymbol{X}_t) - \gamma A(\boldsymbol{X}_t) \cdot \boldsymbol{V}_t)dt + \sqrt{2\gamma}\sigma(\boldsymbol{X}_t) \cdot d\boldsymbol{W}_t$$

in the time scale $t \sim O(1/\epsilon)$, where $\gamma = 1/\epsilon \to \infty$ and we set $m = 1$, $k_B T = 1$.