AM216 Stochastic Differential Equations

Lecture 08
Copyright by Hongyun Wang, UCSC

List of topics in this lecture

- Ornstein-Uhlenbeck Process (continued), solution of particle position $X(t)$
- Behavior of $X(t)$, diffusion coefficient, converging to $W(t)$
- Going backward in time using Bayes' theorem
- Time reversibility of an equilibrium system
- Different interpretations of stochastic integrals

Recap

Ornstein-Uhlenbeck process (OU):

$$m dY = \underbrace{-bY dt}_{\text{dissipation}} + \underbrace{q dW}_{\text{fluctuation}}$$

Four goals of the discussion

Goal 1: Solve for $Y(t)$, the particle velocity

$$\left(Y(t)\middle|Y(0)=y_0\right) \sim N\left(e^{-\beta t}y_0,\ \frac{\gamma^2}{2\beta}(1-e^{-2\beta t})\right) \qquad \text{for } t>0$$

$$Y(\text{large } t) \sim N\left(0,\ \frac{\gamma^2}{2\beta}\right) \qquad \text{for } t>0$$

Goal 2A: $Y(t)$ is a colored noise

Goal 2B: $Y(t)$ converges to a white noise as m converges to zero

Goal 3: Fluctuation-dissipation theorem (relating q to b)

Goal 4: Study the behavior of $X(t)$, the particle position

$$Y(t)=e^{-\beta t}Y(0)+e^{-\beta t}G(t), \qquad G(t)\equiv\int_0^t \gamma e^{\beta s}\,dW(s)$$

$$X(t)-X(0)=\int_0^t Y(s)ds=\frac{1}{\beta}(1-e^{-\beta t})Y(0)+\frac{\gamma}{\beta}G_2(t)$$

where $G_2(t) = \int_0^t (1-e^{-\beta(t-s)})dW(s) \sim$ Gaussian .

We calculate the mean and variance of $G_2(t)$.

$$E\left(G_2(t)\right) = \int_0^t (1-e^{-\beta(t-s)})E\left(dW(s)\right) = 0$$

$$\text{var}\left(G_2(t)\right) = \int_0^t (1-e^{-\beta(t-s)})^2 ds = t - \frac{2}{\beta}(1-e^{-\beta t}) + \frac{1}{2\beta}(1-e^{-2\beta t})$$

==> $(X(t) - X(0))$ is a Gaussian.

$$\left(X(t)-X(0)\right) \sim \frac{(1-e^{-\beta t})}{\beta}Y(0) + \underbrace{N\left(0, \left(\frac{\gamma}{\beta}\right)^2\left(t - \frac{2(1-e^{-\beta t})}{\beta} + \frac{(1-e^{-2\beta t})}{2\beta}\right)\right)}_{\text{containing } dW\text{'s in } [0,t]} \qquad \text{(E01)}$$

<u>Remark:</u>

We cannot integrate $G(t) \sim$ Gaussian directly because $G(t_1)$ and $G(t_2)$ are not independent. We need to write the integral as a sum of dW's, which are independent for disjoint time intervals.

In Goal 4, we discuss two cases for $X(t)$.

<u>Goal 4A:</u> <u>case 1 for $X(t)$</u>: finite m

We show that over <u>long time</u>, $(X(t) - X(0))$ demonstrates a diffusion coefficient.

It follows from (E01) that

$$D \equiv \lim_{t\to\infty} \frac{1}{2t}\text{var}\left(X(t)-X(0)\right) = \frac{1}{2}\left(\frac{\gamma}{\beta}\right)^2$$

Substituting $\beta = \frac{b}{m}$, $\gamma = \frac{q}{m}$, and $q = \sqrt{2k_B T b}$, we have

$$\left(\frac{\gamma}{\beta}\right)^2 = \frac{q^2}{b^2} = \frac{2k_B T b}{b^2} = \frac{2k_B T}{b} \qquad \text{(E02)}$$

which is independent of m. Thus, we arrive at

$$\boxed{D = \frac{k_B T}{b}}$$

This is called the Einstein-Smoluchowski relation.

It relates the drag coefficient to the diffusion coefficient.

<u>Goal 4B:</u> <u>case 2 for $X(t)$:</u> $m \to 0$ (while b and q stay unchanged)

We show that $(X(t) - X(0))$ converges to $\sqrt{2D}\,W(t)$ on any finite resolution time grid.

Specifically, we show that for $t_2 > t_1 > 0$, as $m \to 0$,

- $X(t_1) - X(0) \to \sqrt{2D}\,N(0, t_1)$
- $X(t_1 + t_2) - X(t_1) \to \sqrt{2D}\,N(0, t_2)$
- $(X(t_1) - X(0))$ and $(X(t_1+t_2) - X(t_1))$ are independent.

As $m \to 0$, we have

$$\beta = \frac{b}{m} = O(m^{-1}), \qquad \gamma = \frac{q}{m} = O(m^{-1}), \qquad \frac{\gamma^2}{\beta} = O(m^{-1})$$

$$2D \equiv \left(\frac{\gamma}{\beta}\right)^2 = O(1) \quad \text{and} \quad \frac{1}{\beta}(1 - e^{-\beta \Delta t}) \to 0 \ \text{ for any finite } \Delta t > 0$$

The Maxwell-Boltzmann distribution gives

$$Y(t) \sim N\left(0, \frac{\gamma^2}{\beta}\right) = O\left(\sqrt{\frac{\gamma^2}{\beta}}\right) = O(m^{-0.5})$$

$$\text{==>} \qquad \frac{Y(t)}{\beta} = O(m^{0.5}) \to 0$$

Using (E01), we write $(X(t_1) - X(0))$ as

$$\left(X(t_1) - X(0)\right) \sim (1 - e^{-\beta t_1})\frac{Y(0)}{\beta} + \underbrace{N\left(0,\ 2D\left(t_1 - \frac{2(1 - e^{-\beta t_1})}{\beta} + \frac{(1 - e^{-2\beta t_1})}{2\beta}\right)\right)}_{\text{containing } dW\text{'s in } [0, t_1]}$$

- $$\text{==>} \qquad \left(X(t_1) - X(0)\right) \xrightarrow{\text{as } m \to 0} \sqrt{2D}\ \underbrace{N(0, t_1)}_{\substack{\text{containing } dW\text{'s} \\ \text{in } [0, t_1]}}$$

Similarly, we have

$$\left(X(t_1 + t_2) - X(t_1)\right) \sim (1 - e^{-\beta t_2})\frac{Y(t_1)}{\beta} + \underbrace{N\left(0,\ 2D\left(t_2 - \frac{2(1 - e^{-\beta t_2})}{\beta} + \frac{(1 - e^{-2\beta t_2})}{2\beta}\right)\right)}_{\text{containing } dW\text{'s in } [t_1, t_1 + t_2]}$$

- ==> $\Big(X(t_1+t_2)-X(t_1)\Big) \xrightarrow{\text{as } m\to 0} \sqrt{2D}\ \underbrace{N(0,t_2)}_{\substack{\text{containing } dW\text{'s} \\ \text{in}[t_1,t_1+t_2]}}$

Notice that $(X(t_1)–X(0))$ and $(X(t_1+t_2) –X(t_1))$ contain dW's from disjoint intervals.

- ==> $(X(t_1)–X(0))$ and $(X(t_1+t_2) –X(t_1))$ are independent.

Therefore, we conclude that as $m \to 0$, $(X(t) – X(0))$ converges to $\sqrt{2D}\,W(t)$ on any finite resolution time grid.

Remarks:

1. The diffusion coefficient of the standard Wiener process is 1/2 (not 1).

$$D_{\text{Wiener}} \equiv \frac{1}{2t}\text{var}\Big(W(t)\Big)=\frac{1}{2}$$

2. In the limit of $m \to 0$, $(X(t) – X(0))$ exhibits the behavior of a scaled Wiener process, called the Brownian motion, named after Scottish botanist Robert Brown.
3. The derivation above is for the "simplified story". The real story where radius $a \to 0$ is presented in Appendix A.

Going backward in time in an equilibrium system

In the discussion of Goals #1–4 above, we avoided the issue of going backward in time.

$$E\Big(Y(t)\Big|Y(0)\Big)=e^{-\beta t}Y(0) \qquad \text{for } t > 0$$

Question:

What happens for $–t < 0$? Do we have

$$E\Big(Y(-t)\Big|Y(0)\Big)=e^{+\beta t}Y(0)\ ?$$

which diverges to infinity as $t \to +\infty$. That seems unreasonable.

Answer:

It is more complicated than simply setting $t_{\text{new}} = –t_{\text{old}}$ in the equation.

Recall that when we scale dW, it is best to work with $\dfrac{dW}{\sqrt{dt}}$

$$dW(t)=\sqrt{dt}\cdot\frac{dW(t)}{\sqrt{dt}}, \qquad \frac{dW(t)}{\sqrt{dt}} \sim N(0,1) \quad \text{independent of } t \text{ and } dt$$

It is clear that this works only for $dt > 0$, not for $t_{\text{new}} = –t_{\text{old}}$.

Key point:

In stochastic differential equations, scaling $t_{\text{new}} = –t_{\text{old}}$ does not work!

Recall that Bayes' theorem describes Pr(A | B) in terms of Pr(B | A). We use Bayes' theorem to calculate the backward time evolution based on the forward time evolution.

Bayes' theorem for densities:

$$\rho\left(Y(-t)=y_1\middle|Y(0)=y_2\right) \propto \rho\left(Y(0)=y_2\middle|Y(-t)=y_1\right)\cdot\rho\left(Y(-t)=y_1\right)$$

We assume that the equilibrium has been reached long time ago (at $t = -\infty$) and $Y(t)$ is a stationary process for all t (including negative t). In particular, the unconstrained $Y(t)$ has the equilibrium distribution for all t.

$$\rho\left(Y(-t)=y_1\right) \sim N\left(0,\frac{\gamma^2}{2\beta}\right) \propto \exp\left(\frac{-y_1^{\,2}}{2\gamma^2/(2\beta)}\right)$$

For the forward time evolution, we already derived

$$\left(Y(t_1+t)\middle|Y(t_1)=y_1\right) \sim N\left(e^{-\beta t}y_1,\ \frac{\gamma^2}{2\beta}\left(1-e^{-2\beta t}\right)\right) \qquad \text{for } t>0 \text{ and any } t_1$$

$$\Longrightarrow \qquad \rho\left(Y(0)=y_2\middle|Y(-t)=y_1\right)\propto\exp\left(\frac{-(y_2-e^{-\beta t}y_1)^2}{2(1-e^{-2\beta t})\gamma^2/(2\beta)}\right)$$

Substituting into Bayes' theorem, we obtain

$$\rho\left(Y(-t)=y_1\middle|Y(0)=y_2\right)\propto\exp\left(\frac{-(y_2-e^{-\beta t}y_1)^2}{2(1-e^{-2\beta t})\gamma^2/(2\beta)}\right)\cdot\exp\left(\frac{-y_1^{\,2}}{2\gamma^2/(2\beta)}\right)$$

Note that here y_1 is the independent variable of PDF and we only need to keep track factors that depend on y_1.

$$\rho\left(Y(-t)=y_1\middle|Y(0)=y_2\right)\propto\exp\left(\frac{-\left[e^{-2\beta t}y_1^{\,2}-2e^{-\beta t}y_2\cdot y_1+(1-e^{-2\beta t})y_1^{\,2}\right]}{2(1-e^{-2\beta t})\gamma^2/(2\beta)}\right)$$

$$\propto\exp\left(\frac{-\left[y_1^{\,2}-2e^{-\beta t}y_2\cdot y_1\right]}{2(1-e^{-2\beta t})\gamma^2/(2\beta)}\right)\propto\exp\left(\frac{-(y_1-e^{-\beta t}y_2)^2}{2(1-e^{-2\beta t})\gamma^2/(2\beta)}\right)$$

It follows that the backward time evolution is described by

$$\left(Y(-t)\middle|Y(0)=y_2\right) \sim N\left(e^{-\beta t}y_2,\ \frac{\gamma^2}{2\beta}\left(1-e^{-2\beta t}\right)\right) \qquad \text{for } t>0$$

We compare it with the forward time evolution

$$\left(Y(t)\middle|Y(0)=y_2\right) \sim N\left(e^{-\beta t}y_2,\ \frac{\gamma^2}{2\beta}\left(1-e^{-2\beta t}\right)\right) \qquad \text{for } t>0$$

Conclusions/remarks:

- At equilibrium, the stochastic evolution of going backward in time is statistically the same as the evolution of going forward in time. This is called the time reversibility of equilibrium.
- The time reversibility of equilibrium is a universal law applicable to all thermodynamic systems.
- The intuitive meaning of time reversibility is that if we are given a time series of a system in equilibrium, we won't be able to tell the direction of the time series no matter how long and how detailed the time series is.
- Bayes' theorem is very powerful in expressing the backward time evolution in terms of the forward time evolution.

Going backward in time in a non-equilibrium system (skip in lecture)

Suppose the system starts with $Y(0) = 0$.

For $t_1 > 0$ and $t_2 > 0$, we use Bayes' theorem to calculate $\rho\left(Y(t_1)=y_1 \middle| Y(t_1+t_2)=y_2\right)$.

Bayes' theorem for densities:

$$\rho\left(Y(t_1)=y_1 \middle| Y(t_1+t_2)=y_2\right) \propto \rho\left(Y(t_1+t_2)=y_2 \middle| Y(t_1)=y_1\right)\cdot\rho\left(Y(t_1)=y_1\right)$$

We already derived

- $$\left(Y(t_1) \middle| Y(0)=0\right) \sim N\left(0,\ \frac{\gamma^2}{2\beta}(1-e^{-2\beta t_1})\right) \qquad \text{for } t_1>0$$

==> $$\rho\left(Y(t_1)=y_1\right) \propto \exp\left(\frac{-y_1^{\,2}}{2(1-e^{-2\beta t_1})\gamma^2/(2\beta)}\right)$$

- $$\left(Y(t_1+t_2) \middle| Y(t_1)=y_1\right) \sim N\left(e^{-\beta t_2}y_1,\ \frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2})\right) \quad \text{for } t_1>0,\ t_2>0$$

==> $$\rho\left(Y(t_1+t_2)=y_2 \middle| Y(t_1)=y_1\right) \propto \exp\left(\frac{-(y_2-e^{-\beta t_2}y_1)^2}{2(1-e^{-2\beta t_2})\gamma^2/(2\beta)}\right)$$

Substituting into Bayes' theorem, we obtain

$$\rho\left(Y(t_1)=y_1 \middle| Y(t_1+t_2)=y_2\right) \propto \exp\left(\frac{-(y_2-e^{-\beta t_2}y_1)^2}{2(1-e^{-2\beta t_2})\gamma^2/(2\beta)}\right)\cdot\exp\left(\frac{-y_1^{\,2}}{2(1-e^{-2\beta t_1})\gamma^2/(2\beta)}\right)$$

$$\propto \exp\left(\frac{-\left[(1-e^{-2\beta(t_1+t_2)})y_1^{\,2}-2(1-e^{-2\beta t_1})e^{-\beta t_2}y_2\cdot y_1\right]}{2(1-e^{-2\beta t_1})(1-e^{-2\beta t_2})\gamma^2/(2\beta)}\right)$$

$$\propto \exp\left(\frac{-\left(y_1 - \frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})} e^{-\beta t_2} y_2 \right)^2}{2\frac{(1-e^{-2\beta t_1})(1-e^{-2\beta t_2})}{(1-e^{-2\beta(t_1+t_2)})}\gamma^2/(2\beta)} \right)$$

It follows that

$$\left(Y(t_1)\middle|Y(t_1+t_2)=y_2\right) \sim N\left(\frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})} e^{-\beta t_2} y_2,\ \ \frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})}\frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2}) \right)$$

We discuss two special cases.

<u>Case i)</u> $t_1 \to +\infty$ while t_2 = fixed

$$\frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})} e^{-\beta t_2} y_2 \to e^{-\beta t_2} y_2 \qquad \text{for large } t_1$$

$$\frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})}\frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2}) \to \frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2}) \qquad \text{for large } t_1$$

==> $\left(Y(t_1)\middle|Y(t_1+t_2)=y_2\right) \sim N\left(e^{-\beta t_2} y_2,\ \frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2}) \right)$ for large t_1

This is the same as the equilibrium case, not a surprise at all.

Case ii) $t_1 = t_2 = h$ and βh is not large.

$$\frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})} e^{-\beta t_2} y_2 = \frac{e^{-\beta h} y_2}{1+e^{-2\beta h}}$$

$$\frac{(1-e^{-2\beta t_1})}{(1-e^{-2\beta(t_1+t_2)})}\frac{\gamma^2}{2\beta}(1-e^{-2\beta t_2}) = \frac{\gamma^2}{2\beta}\left(\frac{1-e^{-2\beta h}}{1+e^{-2\beta h}} \right)$$

$$\left(Y(h)\middle|Y(2h)=y_2\right) \sim N\left(\frac{e^{-\beta h} y_2}{1+e^{-2\beta h}},\ \ \frac{\gamma^2}{2\beta}\left(\frac{1-e^{-2\beta h}}{1+e^{-2\beta h}} \right) \right)$$

We compare it with the forward time evolution

$$\rho\left(Y(2h)\middle|Y(h)=y_1\right) \sim N\left(e^{-\beta h} y_1,\ \frac{\gamma^2}{2\beta}\left(1-e^{-2\beta h}\right) \right)$$

When βh is not large, this case clearly demonstrates the difference between forward time evolution and backward time evolution in a non-equilibrium system.

Different interpretations of stochastic integrals

Beauty of the deterministic calculus

Consider the integral of a deterministic function $f(s)$.

$$\int_0^t f(s)ds = \lim_{N\to\infty}\sum_{j=0}^{N-1} f(\tilde{s}_j)\Delta s$$

where

$$\Delta s = \frac{t}{N}, \qquad s_j = j\Delta s, \qquad \tilde{s}_j \in [s_j\ s_{j+1}]$$

Note: When $f(s)$ is continuous, the choice of $\tilde{s}_j \in [s_j, s_{j+1}]$ does not affect the result.

In particular,

$$\lim_{N\to\infty}\sum_{j=0}^{N-1} f(s_j)\Delta s \;=\; \lim_{N\to\infty}\sum_{j=0}^{N-1} f(s_{j+1})\Delta s \;=\; \lim_{N\to\infty}\sum_{j=0}^{N-1} f(s_{j+1/2})\Delta s$$

A simple stochastic integral

$$\int_0^t f(s)dW(s) = \lim_{N\to\infty}\sum_{j=0}^{N-1} f(\tilde{s}_j)\Delta W_j$$

where

$$\Delta s = \frac{t}{N}, \qquad s_j = j\Delta s, \qquad \tilde{s}_j \in [s_j\ s_{j+1}]$$

$$\Delta W_j = W(s_{j+1}) - W(s_j)$$

Riemann sum $\lim_{N\to\infty}\sum_{j=0}^{N-1} f(\tilde{s}_j)\Delta W_j$ is a Gaussian with mean = 0 and

$$\text{variance} = \lim_{N\to\infty}\sum_{j=0}^{N-1} f(\tilde{s}_j)^2\Delta s = \int_0^t f(s)^2 ds$$

Again, when $f(s)$ is continuous, the choice of $\tilde{s}_j \in [s_j, s_{j+1}]$ does not affect the result.

A more complicated stochastic integral:

$$\int_0^t f(s, W(s))dW(s) = \lim_{N\to\infty}\sum_{j=0}^{N-1} f(\tilde{s}_j, W(\tilde{s}_j))\Delta W_j$$

where

$$\Delta s = \frac{t}{N}, \qquad s_j = j\Delta s, \qquad \tilde{s}_j \in [s_j\ s_{j+1}]$$

$$\Delta W_j = W(s_{j+1}) - W(s_j)$$

Note that

- $f(s, W(s))$ is not a deterministic function of s.
- $f(\tilde{s}_j, W(\tilde{s}_j))$ is a random variable, correlated with ΔW_j.
- As a result, different choices of $\tilde{s}_j \in [s_j, s_{j+1}]$ lead to different results.
- Thus, integral $\int_0^t f(s, W(s))dW(s)$ is subject to different interpretations.

Appendix A case 2 for $X(t)$, the real story where radius $a \to 0$

We show that $\sqrt{a}\left(X(t_1) - X(0)\right)$ converges to $cW(t)$ on any finite resolution time grid.

Specifically, we show that for $t_2 > t_1 > 0$, as $a \to 0$,

- $\sqrt{a}\left(X(t_1) - X(0)\right) \to c\,N(0, t_1)$
- $\sqrt{a}\left(X(t_1 + t_2) - X(t_1)\right) \to c\,N(0, t_2)$
- $(X(t_1) - X(0))$ and $(X(t_1+t_2) - X(t_1))$ are independent.

As $a \to 0$, we have

$$m = O(a^3), \qquad b = O(a), \qquad q = \sqrt{2k_B T b} = O(\sqrt{a})$$

$$\beta = \frac{b}{m} = O(a^{-2}), \qquad \gamma = \frac{q}{m} = O(a^{-2.5}), \qquad \frac{\gamma}{\beta} = O(a^{-0.5}), \qquad \frac{\gamma^2}{\beta} = O(a^{-3})$$

$$c \equiv \sqrt{a}\frac{\gamma}{\beta} = O(1), \quad \text{and} \quad \frac{1}{\beta}(1 - e^{-\beta \Delta t}) \to 0 \ \text{for any finite } \Delta t > 0$$

The Maxwell-Boltzmann distribution gives

$$Y(t) \sim N\left(0, \frac{\gamma^2}{\beta}\right) = O\left(\sqrt{\frac{\gamma^2}{\beta}}\right) = O(a^{-1.5})$$

$$\Longrightarrow \qquad \frac{\sqrt{a}\,Y(t)}{\beta} = O(a) \to 0$$

Using (E01), we write $(X(t_1)–X(0))$ as

$$\sqrt{a}\Big(X(t_1)-X(0)\Big) \sim (1-e^{-\beta t_1})\frac{\sqrt{a}\,Y(0)}{\beta}+\underbrace{N\left(0,c^2\left(t_1-\frac{2(1-e^{-\beta t_1})}{\beta}+\frac{(1-e^{-2\beta t_1})}{2\beta}\right)\right)}_{\text{containing } dW\text{'s in } [0,t_1]}$$

- ==> $\sqrt{a}\Big(X(t_1)-X(0)\Big) \xrightarrow{\text{as } a\to 0} \underbrace{c^2 N(0,t_1)}_{\substack{\text{containing } dW\text{'s} \\ \text{in } [0,t_1]}}$

Similarly, we have

$$\sqrt{a}\Big(X(t_1+t_2)-X(t_1)\Big) \sim (1-e^{-\beta t_2})\frac{\sqrt{a}\,Y(t_1)}{\beta}+\underbrace{N\left(0,c^2\left(t_2-\frac{2(1-e^{-\beta t_2})}{\beta}+\frac{(1-e^{-2\beta t_2})}{2\beta}\right)\right)}_{\text{containing } dW\text{'s in } [t_1,t_1+t_2]}$$

- ==> $\sqrt{a}\Big(X(t_1+t_2)-X(t_1)\Big) \xrightarrow{\text{as } a\to 0} \underbrace{c^2 N(0,t_2)}_{\substack{\text{containing } dW\text{'s} \\ \text{in} [t_1,t_1+t_2]}}$

Again, $(X(t_1)–X(0))$ and $(X(t_1+t_2) –X(t_1))$ contain dW's from disjoint intervals.

- ==> $(X(t_1)–X(0))$ and $(X(t_1+t_2) –X(t_1))$ are independent.

Therefore, we conclude that $\sqrt{a}\Big(X(t_1)-X(0)\Big)$ converges to c$W(t)$ as $a \to 0$.