0.1 Summary

Summary of the previous lectures. We considered a more general stochastic process, a *Markov Process*, when the future only depends on the present. We wrote a Master Equation, and taking the continuum limit we get a second order partial differential equation, with two coefficients depending on the first two moments of the transition rate: f and D. We would want them to represent the *force* and *diffusion rate*, but we can't find their physical meaning. So we consider the Langenvin equation, reaching the desired physical meaning.

There, the increment depends on a *deterministic term* f and a *noise term*:

$$\mathrm{d}x(t) = f(x(t), t)\, \mathrm{d}t + \sqrt{2D(x(t), t)}\, \mathrm{d}B(t) \qquad f = \frac{F_{\text{ext}}}{\gamma}$$

If we discretize this equation, passing to finite differences, we get:

$$\Delta x(t) = f(x(t), t)\Delta t + \sqrt{2D(x(t), t)}\Delta B(t) \qquad \Delta B(t) \sim \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{\Delta B^2}{2\Delta t}\right)$$

This is needed because $\mathrm{d}x(t)\,/\,\mathrm{d}t$ is ill-defined (as we saw in the previous lecture). Note that $\Delta x(t) = x(t + \Delta t) - x(t)$.

We want to show that this kind equation leads to the same Fokker-Planck equation that we saw previously, and that was derived from the Master Equation. Then we would like to examine how much the stochastic amplitude (coefficient of $\mathrm{d}B(t)$) is related to *temperature*. In fact, we know already that f depends on F_{ext}, with $\boldsymbol{F}_{\text{ext}} = -\boldsymbol{\nabla}V$. We would like that, at constant temperature, the pdf of the stationary state will tend to the *Maxwell-Boltzmann distribution*:

$$\mathbb{P}(x, t) \xrightarrow[t\to\infty]{} \frac{1}{z} \exp\left(-\frac{V(x)}{k_B T}\right)$$

0.2 Stochastic integral

Now the main problem is that the Langenvin equation is a **stochastic equation**, and we do not (yet) know how to integrate a stochastic term:

$$x(t) = x(0) + \int_0^t f(x(\tau), \tau)\, \mathrm{d}\tau + \int_0^t \sqrt{2D(x(\tau), \tau)}\, \mathrm{d}B(\tau)$$

Note that we cannot "multiply" by $\mathrm{d}\tau\,/\,\mathrm{d}\tau$ to change variables, as this would require the *existence of the first derivative*!

So, let's consider a stochastic integral:

$$S = \int_0^t G(\tau)\, \mathrm{d}B\,(\tau)$$

This is a **random variable**. To compute it, we first discretize it:

$$S_n = \sum_{i=0}^n G(\tau_i)[B(t_i) - B(t_{i-1})]$$

where $t_{i-1} \leq \tau_i \leq t_i$. If $B(t)$ is well-behaving (has a first derivative), then how we choose τ_i inside $[t_{i-1}, t_i]$ does not matter, as the limit will be always the same. However, this is not the case in this expression.

First, we ask what does it mean to take the limit $n \to \infty$. Suppose S_n is a generic random variable, and we consider $S_n \xrightarrow[n\to\infty]{} S$. We can consider the following definitions:

1. **Standard (stochastic) limit**: $S_n \xrightarrow[n\to\infty]{\text{a.s.}} S$, i.e. S_n tends to S *almost surely*. More precisely:

$$\forall \epsilon > 0 \quad \exists N : n \geq N \, |S_n - S| < \epsilon$$

   In other words, the fractions of sequences S_n that *do not come ϵ-close to S* tends to 0 as the number N of considered sequences goes to infinity.

2. $S_n \xrightarrow[n\to\infty]{P} S$, meaning that:

$$\lim_{n\to\infty} \mathbb{P}(|S - S_n| > \epsilon) = 0 \quad \forall \epsilon > 0$$

   This definition is closely related to (1), but here we focus on the *probability* of getting a "bad sequence", and not on the occurrences when this does indeed happen.

3. **Convergence in Distribution sense**. $S_n \xrightarrow[n\to\infty]{P} S$, that is:

$$\mathbb{P}(S_n \in A) \to \mathbb{P}(S \in A)$$

4. $S_n \xrightarrow[n\to\infty]{L^q} S$, where:

$$\langle |S_n - S|^q \rangle \xrightarrow[n\to\infty]{} 0$$

One can show that $(1) \Rightarrow (2) \Rightarrow (3)$, and $(4) \Rightarrow (2)$. This last implication can be easily derived. First, we note that the probability of a sequence S_n not ϵ-close to S is just the expected value of the chracteristic function of the set where an inequality holds:

$$\langle \mathbb{I}_{|S-S_n|>\epsilon} \rangle = \mathbb{P}(|S - S_n| > \epsilon) \leq \langle \underbrace{\mathbb{I}_{|S-S_n|>\epsilon}}_{\leq 1} \cdot \underbrace{\left| \frac{S - S_n}{\epsilon} \right|^q}_{\leq 1} \rangle \leq \langle |S - S_n|^q \rangle \frac{1}{\epsilon^q} \xrightarrow[n\to\infty]{} 0$$

In our case, we consider the L^2 convergence definition. So we say that $S_n \to S$ as $n \to \infty$ iff:

$$\lim_{n\to\infty} \langle (S_n - S)^2 \rangle = 0$$

This implies the *standard convergence* of the expected value:

$$\langle S_n \rangle \to \langle S \rangle$$

In fact:

$$|\langle S_n \rangle - \langle S \rangle| = |\langle S_n - S \rangle| \le \langle |S_n - S| \rangle$$

Also, the Hölder inequality holds:

$$|\langle fg \rangle| \le \langle |fg| \rangle \le \langle f^2 \rangle^{1/2} \langle g^2 \rangle^{1/2}$$

that, for this particular instance, is also called the *Schwartz inequality*.

If we apply it to the previous expression:

$$\langle |S_n - S| \rangle \le \langle (S_n - S)^2 \rangle^{1/2} \langle 1^2 \rangle^{1/2} = \langle (S_n - S)^2 \rangle^{1/2}$$

which implies our thesis: $\lim_{n \to \infty} \langle S_n \rangle = \langle S \rangle$. In the case of q=2, the convergence is called "mean square" (m.s.)

Reference: Chapter 3 of book by Gardiner, *Handbook of Stochastic Processes.*

Example 1 (a):

Suppose $G(\tau) = B(\tau)$, and consider the following integral:

$$S = \int_0^t B(\tau) \, dB(\tau)$$

If $B(\tau)$ where differentiable, then we could simply change variables and solve:

$$S = \int_0^t B(\tau) \frac{dB(\tau)}{d\tau} \, d\tau = \frac{1}{2} B^2(\tau) \Big|_0^t = \frac{B^2(t) - B^2(0)}{2} \quad \text{if } \exists \frac{dB}{d\tau}$$

But $\frac{d\beta}{d\tau}$ does not exist!
So, we first discretize:

$$S_n = \sum_{i=1}^n B(\tau_i)[B(t_i) - B(t_{i-1})] \qquad t_0 = 0; t_n = t \qquad (1)$$

Then:

$$\langle S_n \rangle = \sum_{i=1}^n \langle B(\tau_i)(B(t_i) - B(t_{i-1})) \rangle$$

Recall that:

$$\langle B(t)B(t') \rangle = \min(t, t') \qquad (2)$$

And so, as $t_{i-1} \le \tau_i \le t_i$:

$$(1) = \langle S_n \rangle = \sum_{i=1}^n [\langle B(\tau_i)B(t_i) \rangle - \langle B(\tau_i)B(t_{i-1}) \rangle] \underset{(2)}{=} \sum_{i=1}^n (\tau_i - t_{i-1})$$

3

Then suppose τ_i lies in "the same place" at every interval $[t_{i-1}, t_i]$, so that the following holds:

$$\tau_i = \lambda t_i + (1 - \lambda) t_{i-1}$$

Substituting in the previous expression:

$$\langle S_n \rangle = \lambda \sum_{i=1}^{n} (t_i - t_{i-1}) = \lambda t$$

And finally:

$$\lim_{n \to \infty} \langle S_n \rangle = \lambda t$$

So the limit *depends* on how we choose the τ_i inside the intervals $[t_{i-1}, t_i]$ - which is **not** true in ordinary calculus (where the result is independent on this choice).

There are many possibilities for λ:

$$\lambda = \begin{cases} 0 & \text{Ito's prescription} \\ \frac{1}{2} & \text{Stratonovich's prescription (or middle-point prescr.)} \end{cases}$$

Leading to:

$$S_n \xrightarrow[n \to \infty]{\text{m.s.}} S = \begin{cases} \frac{B^2(t) - B^2(0)}{2} & \lambda = 1/2 \\ \frac{B^2(t) - B^2(0)}{2} - \frac{t}{2} & \lambda = 0 \end{cases}$$

The Stratonovich prescription gives exactly the same result as ordinary calculus. However, note that it involves a dependence *on the future*, i.e. the next step of a path depends on the point that is a *half-step* later. This has no a real physical meaning (in a certain sense, it "violates causality"). That's why many physicists prefer the Ito's prescription.

So, we want now to prove that result:

$$\sum_{i=1}^{n} B(t_{i-1})(B(t_i) - B(t_{i-1})) \xrightarrow[n \to \infty]{\text{m.s.}} \frac{B^2(t) - B(0)}{2} - \frac{t}{2}$$

Denote:

$$B(t_i) = B_i; \qquad \Delta B_i = B_i - B_{i-1}$$

and so:

$$S_n = \sum_{i=1}^{n} B_{i-1} \Delta B_i = \frac{1}{2} \sum_{i=1}^{n} \left[(B_{i-1} + \Delta B_i)^2 - B_{i-1}^2 - (\Delta B_i)^2 \right] =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \left[B_i^2 - B_{i-1}^2 - (\Delta B_i)^2 \right] = \frac{1}{2} (B_n^2 - B_0^2) - \frac{1}{2} \sum_{i=1}^{n} (\Delta B_i)^2$$

Then we want to show:

$$\frac{B^2(t) - B^2(0)}{2} - \frac{1}{2}\sum_{i=1}^{n}(\Delta B_i)^2 \xrightarrow[n\to\infty]{\text{m.s.}} \frac{B^2(t) - B(0)}{2} - \frac{t}{2} \qquad t_n = t;\ t_0 = 0$$

So we take their difference in modulus squared and take the limit of the average:

$$\left\langle \left| \frac{B^2(t) - \cancel{B^2(0)}}{2} - \frac{1}{2}\sum_{i=1}^{n}(\Delta B_i)^2 - \left[\frac{B^2(t) - \cancel{B(0)}}{2} - \frac{t}{2} \right] \right|^2 \right\rangle \xrightarrow[n\to\infty]{?} 0$$

Expanding:

$$\frac{1}{4}\left\langle \left[\sum_{i=1}^{n}(\Delta B_i)^2 + t \right]^2 \right\rangle = \frac{1}{4}\left\langle \left[t - \sum_{i=1}^{n}(\Delta B_i)^2 \right]^2 \right\rangle = \frac{1}{4}\left\langle \left[\sum_{i=1}^{n}(\Delta t_i - \Delta B_i^2) \right]^2 \right\rangle$$

Note that $t = \sum_{i=1}^{n}\Delta t_i$. Then, by expanding the square of the sum:

$$= \frac{1}{4}\left\langle \sum_{i=1}^{n}(\Delta t_i - (\Delta B_i)^2) \right\rangle \sum_{j}(\Delta t_j - (\Delta B_j)^2)$$

as $\left(\sum_i a_i\right)^2 = \sum_{ij} a_i a_j$. In this summation we isolate the contribution with $i = j$ and $i \neq j$, and use the linearity of the average (as these are finite sums):

$$= \frac{1}{4}\left[\sum_{i=1}^{n}\langle (\Delta t_i - (\Delta B_i)^2)^2 \rangle + \sum_{i\neq j}\langle (\Delta t_i - (\Delta B_i)^2)(\Delta t_j - (\Delta B_j)^2) \rangle \right]$$

The only random variable is now the ΔB_i. Recall that:

$$\langle A \rangle = \int \mathrm{d}\Delta B_i \dots \mathrm{d}\Delta B_n\, A \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\Delta t_i}} \exp\left(-\frac{(\Delta B_i)^2}{2\Delta t_i} \right)$$

The ΔB_i are gaussian and **independent**, and so the average of the product is just the product of the averages:

$$\langle (\Delta t_i - (\Delta B_i)^2)(\Delta t_j - (\Delta B_j)^2) \rangle = \langle (\Delta t_i - (\Delta B_i)^2) \rangle \langle (\Delta t_j - (\Delta B_j)^2) \rangle$$

Then, recall that:

$$\langle (\Delta B_i)^2 \rangle = \int \frac{\mathrm{d}\Delta B_i}{\sqrt{2\pi\Delta t_i}} \Delta B_i^2 \exp\left(-\frac{\Delta B_i^2}{2\Delta t_i} \right) = \Delta t_i$$

and so:

$$\langle (\Delta t_i - (\Delta B_i)^2) \rangle = 0$$

The only term left is:

$$= \frac{1}{4}\sum_{i=1}^{n}\langle (\Delta t_i - (\Delta B_i)^2)^2 \rangle = \frac{1}{4}\sum_{i=1}^{n}\left[\Delta t_i^2 - 2\Delta t_i \underbrace{\langle (\Delta B_i)^2 \rangle}_{\Delta t_i} + \langle \Delta B_i^4 \rangle \right]$$

Recall that:

$$\langle x^{2n} \rangle = \sigma^{2n} \frac{(2n)!}{2^n n!} = \begin{cases} \sigma^2 & n = 1 \\ \sigma^4 \frac{4!}{4 \cdot 2!} = 3\sigma^4 & n = 2 \end{cases}$$

if x is sampled by a Gaussian with 0 average. So:

$$= \frac{1}{2} \sum_{i=1}^{n} \Delta t_i^2$$

When taking the limit of the mesh ($n \to \infty$), the time intervals vanish, meaning that:

$$\max_i \Delta t_i \xrightarrow[n \to \infty]{} 0$$

So:

$$\frac{1}{2} \sum_{i=1}^{n} \Delta t_i^2 \leq \frac{1}{2} \left(\max_j \Delta t_j \right) \sum_i \Delta t_i$$

as $\Delta t_i \leq \max_j \Delta t_j$. Then $\sum_i \Delta t_i = t$:

$$= \frac{t}{2} \max_j \Delta t_j \xrightarrow[n \to \infty]{} 0$$

To prove the result with the S prescription, we then must show:

$$S_n = \sum_{i=1}^{n} B\left(\frac{t_i + t_{i-1}}{2} \right) [B(t_i) - B(t_{i-1})] \xrightarrow[n \to \infty]{\text{m.s.}} \frac{B^2(t) - B^2(0)}{2}$$

Note that now we need a set of *middle points* in the mesh, which leads to some complications.

One trick is to simply *double* the discretization, and choose the *middle points* to be the *odd* indices. We then define:

$$S'_{2n} = \sum_{i=1}^{2n} B_{2i-1}(B_{2i} - B_{2(i-1)})$$

with $t_{2i-1} \equiv (t_{2i} + t_{2(i-1)})/2$, while the t_{2i} may be distributed arbitrarily.

A shorter way to compute that, but not as rigorous, is by stating that:

$$S_n = \sum_{i=1}^{n} \frac{B(t_i) + B(t_{i-1})}{2}(B(t_i) - B(t_{i-1}))$$

However it is not obvious, because $B(t_i)$ are all random variables. The two expressions *have the same distribution*, but they *are not the same*! In any way, if we do this, the thesis immediately follows:

$$= \frac{1}{2} \sum_{i=1}^{n} (B^2(t_i) - B^2(t_{i-1}))$$

In our calculations, we will be usually concerned with the following kinds of stochastic integrals $G(t)$:

1. $\int_0^t F(B(\tau)) \, dB(\tau)$

2. $\int_0^t g(\tau) \, dB(\tau)$

3. $\int_0^t g(\tau) \, d\tau$ (usual integrals)

These integrals $G(t)$ are called *non-anticipating functions*, because they are independent of $B(t') - B(t)$ for $t' > t$, meaning that they do not dependent on what happens in the Brownian motion at times later than t (the future). So, by using Ito's prescription:

$$\int F(B(\tau)) \, dB(\tau) \underset{\text{m.s. I}-\text{p}}{=} \sum_{i=1}^n F(B_{i-1}) \Delta B_i$$

Note how $F(B_{i-1})$ and ΔB_i are independent of each other, simplifying the calculations. (Note that the Stratonovich prescription here causes *troubles* during evaluation, as it introduces some interdependence between different terms).

The goal of the next lecture will be to derive the Langenvin equation from the Fokker-Planck equation, and also to show that integrals of the kind:

$$\int_0^t H(B(\tau), \tau)(dB(\tau))^k \underset{m.s.}{=} \sum_{i=1}^n H(B_{i-1}, \tau_{i-1})(\Delta B_i)^k =$$

$$= \begin{cases} \int_0^t H(B, \tau) \, dB(\tau) & k = 1 \\ \int_0^t H(B(\tau), \tau) \, d\tau & k = 2 \\ 0 & k > 2 \end{cases}$$

leading to Ito's rules:

$$(dB)^k = \begin{cases} dB & k = 1 \\ d\tau & k = 2 \\ 0 & k > 2 \end{cases}$$