Advanced Statistical Mechanics
Part III Physics, Unversity of Cambridge

Lawrence Wu

April 17, 2023

Contents

Chapter 1 **Lecture 1: Revision of equilibrium stochastic physics** — Page 4

1.1 Prerequisite: Canonical coordinates and Phase space 4

1.2 Micro-canonical ensemble 4

1.3 Canonical ensemble 5

1.4 Different types of systems (Hohenberg-Halperin 1977) 6

Model A: No conservation laws — 6 • Model B: Conserved fields — 7 • Model C-F... — 7

Chapter 2 **Lecture 2: Basic definitions of probability theory** — Page 8

2.1 Random Variables and Probabilty 8

2.2 Random Walk 9

2.3 Random walk in time 11

probabilities — 11 • Normalisation — 11 • Reduction — 11 • Correlation functions — 12 • Stationary process — 12 • Conditional probability — 12 • Markov Process — 12 • Evolution relation — 13 • Kolmogorov-Chapman relation — 13

Chapter 3 **Lecture 3:The Poisson Process** — Page 14

3.1 Poisson Process 14

3.2 The Poisson distribution 16

3.3 Evolution of the Master Equation 17

Chapter 4 **Lecture 4: The Wiener process** — Page 19

4.1 The Wiener process 19

Characteristic function — 19 • Scaling — 20 • Time inversion — 20 • Time reversal — 20 • Probability distribution — 20

4.2 Stochastic differential equations 21

4.3 Example: Geometric Brownian Motion 22

4.4 Itoh Lemma 22

Chapter 5 **Lecture 5: Black-Scholes equation** Page 25

5.1 Black-Scholes Equation 25

5.2 A simple case 26

Chapter 6 **Lecture 6: Ornstein-Uhlenbeck process** Page 28

6.1 Ornstein-Uhlenbeck process 28

6.2 O-U process with multiple variables 30

Chapter 7 **Lecture 7: More on Ornstein-Uhlenbeck** Page 32

7.1 Kinetic equation of the O-U process 32

Chapter 8 **Lecture 8: O-U process continued** Page 37

8.1 Discussion before continuing 37

8.2 O-U process key examples 38

Particle in linear potential — 38 • Particle in a quadratic potential — 38

8.3 Convection diffusion 39

Chapter 9 **Lecture 9: Mean first passage time (MFPT)** Page 40

9.1 Examples: 1D simple diffusion 40

9.2 MFTP of the Wiener Process 41

Chapter 10 **Lecture 10: MFPT continued** Page 43

10.1 Examples 45

1D diffusion — 45 • Kramers problem: — 45 • Diffusion in arbitrary periodic potential — 46

Chapter 11 **Lecture 11: Multiplicative noise** Page 47

11.1 MFPT final part 47

11.2 Multiplicative noise 47

Itoh version — 48 • Stratonovich version — 48 • Lau, Lubensky version — 48

Chapter 12 **Lecture 12: Last lecture** Page 50

12.1 Multiplicative Noise continued 50

12.2 Non-Markov processes 51

Chapter 1

Lecture 1: Revision of equilibrium stochastic physics

1.1 Prerequisite: Canonical coordinates and Phase space

In mathematics and classical mechanics, canonical coordinates are sets of coordinates on phase space which can be used to describe a physical system at any given point in time. In dynamical system theory, a phase space is a space in which all possible states of a system are represented, with each possible state corresponding to one unique point in the phase space.

1.2 Micro-canonical ensemble

In statistical mechanics, the microcanonical ensemble is a statistical ensemble that represents the possible states of a mechanical system whose total energy is exactly specified. The system is assumed to be isolated in the sense that it cannot exchange energy or particles with its environment, so that (by conservation of energy) the energy of the system does not change with time. In the Micro-canonical ensemble, the Hamiltonian for the system is usually defined as a function of the canonical coordinates momentum $\boldsymbol{p}$ and position $\boldsymbol{q}$:

$$\mathcal{H} = \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}) \tag{1.1}$$

which allows the different states of the system to be represented in phase space $\mathbb{P}$ and $\mathbb{Q}$. The energy of system E would be the weighted mean of $\mathcal{H}$ in phase space. Rather than solving for the behaviour of a particular set of initial conditions, let us hypothesise that the energy is all we need to describe the equilibrium state. This leads us to a statistical mechanical description of the equilibrium state of our system as an ensemble of all possible states with energy E—the microcanonical ensemble. Now consider the scenario where, for a given E, we introduce a small positive change dE in energy and the volume of the allowed states in phase space is changed. We define:

$$\Omega(E)\, dE = \int d\mathbb{P} d\mathbb{Q} \quad \text{(for states with } E < \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}) < E + dE) \tag{1.2}$$

$$= \int d\mathbb{P} d\mathbb{Q}\ \Theta(E + dE - H) - \Theta(E - H) \quad (\Theta \text{ is the Heaviside step function}) \tag{1.3}$$

$$= dE \cdot \int d\mathbb{P} d\mathbb{Q} \left[\frac{\partial \Theta(\mathrm{E} - H)}{\partial E} \right] \tag{1.4}$$

$$\tag{1.5}$$

and we reach the conclusion:

Definition 1.2.1: Number of states with hamiltonian $\mathcal{H}$ equal to given energy E.

$$\Rightarrow \Omega(E) = \int d\mathbb{P}d\mathbb{Q}\ \delta(E - \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}))$$

We then define the "entropy" S to be:

Definition 1.2.2: Entropy

$$S(E) = k_B \ln \Omega(E)$$

Now let's talk about probability $\rho(E)$. But how do we define it and how is it useful? Let's consider a quantity of the system $X(\boldsymbol{p}, \boldsymbol{q})$ and we're trying to deduce its expectation value $\langle X \rangle$. We'll want to integrate over all states with the constraint $\mathcal{H} = E$:

$$\langle X(\boldsymbol{p}, \boldsymbol{q}) \rangle = \frac{1}{\Omega(E)} \int d\mathbb{P}d\mathbb{Q}\ X(\boldsymbol{p}, \boldsymbol{q}) \delta(E - \mathcal{H}(\boldsymbol{p}, \boldsymbol{q})) \tag{1.6}$$

where we can read out the normalised probability $\rho(E)$ to be:

Definition 1.2.3: Normalised probability

$$\rho(E) = \frac{1}{\Omega(E)} \delta(E - \mathcal{H}(\boldsymbol{p}, \boldsymbol{q}))$$

There is a useful spooky relation when it comes to relating Ωs for different energies:

Question 1

Check that for a system consisting of two subsystems with total energy E:

$$\Omega(E) = \int \Omega_1 (E_1) \Omega_2 (E - E_1) \, dE_1$$

as should be intuitively clear

1.3 Canonical ensemble

In comparison with the micro-canonical ensemble, a canonical ensemble is a statistical ensemble that represents the possible states of a mechanical system in thermal equilibrium with a heat bath at a fixed temperature. The system can exchange energy with the heat bath so that the states of the system will differ in total energy. Since the energy is not fixed, we assign a probability P to each distinct state that's related to their energy:

$$P \propto e^{-\beta E} \tag{1.7}$$

and for P to be normalised, we define the normalisation factor Z aka the partition function to be:

Definition 1.3.1: The partition function

$$Z = \sum_{All\ states} e^{-\beta E} = \int d\mathbb{P}d\mathbb{Q}\ e^{\beta - \mathcal{H}(p,q)} \tag{1.8}$$

$$= \sum_{Microstates(E)} \Omega(E) e^{-\beta E} \tag{1.9}$$

$$= \sum_{Microstates(E)} e^{-\beta(E - k_B T \ln \Omega(E))} \tag{1.10}$$

$$\tag{1.11}$$

Where we often call the value in the exponent the "Free energy" of a microstate.

Definition 1.3.2: Free energy

$$F(E) = E - TS(E)$$

After normalisation, our normalised probability becomes:

Definition 1.3.3: Probabilty for the system to have energy E

$$\rho(E) = \frac{1}{Z} e^{\beta(E - TS)}$$

From the relation above, we can see that the probability is maximised when the free energy F is minimised. And for a large system, this gives the equilibrium state due to the central limit theorem. (We will go further into this in the later lectures.)

In conclusion, minimising the free energy allow us to find the balance between the two incompatible demands in statistical physics: 1. Minimising energy E 2. Maximising entropy S(E), and ignoring the fluctuation in the system about this preferred microstate. But in this course, we aim to look at two non-equilibrium questions: 1. How long does it take to reach such an equilibrium state from an initial state? 2. Are there barriers and metastable states?

1.4 Different types of systems (Hohenberg-Halperin 1977)

1.4.1 Model A: No conservation laws

An order parameter ϕ is defined to be a variable that can be used to describe a system and gives a complete description of the hamiltonian. In other words, the Hamiltonian can be written in the form $\mathcal{H}(\phi)$. And when there are no conservation laws in a system for this parameter ϕ, we call it a Model A system. For example, the magnetisation of a magnetic system is an order parameter of Model A since it does not have to be conserved at all. For models like this we can make a classical mechanical analogy (Think of ϕ as the cartesian coordinates of a particle):

$$\text{"driving force"} = \text{"friction force"} \tag{1.12}$$

$$-\frac{\partial \mathcal{H}(\phi)}{\partial \phi} = \gamma \frac{\partial \phi}{\partial t} \tag{1.13}$$

And this will be the principal which controls the evolution of the system towards equilibrium. This is in contrast with another type of system that we call Model B:

1.4.2 Model B: Conserved fields

For this type of system, we have a conservation law of the form:

$$\int \phi(\boldsymbol{x}) d\boldsymbol{x} = constant \tag{1.14}$$

and for this type of system, since we can only change the distribution of this parameter but not the total quantity, the characteristic equation will be the classical continuity equation:

$$\frac{\partial \phi}{\partial t} = -D\nabla^2 \frac{\partial \mathcal{H}(\phi)}{\partial \phi} \tag{1.15}$$

where the flux J is:

$$\boldsymbol{J} = -D\nabla \frac{\partial \mathcal{H}}{\partial \phi} \tag{1.16}$$

1.4.3 Model C-F...

Model C describes a field ϕ that is coupled to a conserved auxiliary field and such analysis goes down all the way to Model F. There's no need to know all of them and the above sections are just raising awareness that these mechanical analysis exists.

Chapter 2

Lecture 2: Basic definitions of probability theory

2.1 Random Variables and Probabilty

As the title suggests, this lecture covers the basic of probability theory so there is a common language that we can use to describe stochastic systems. Let's start with the definition of "Random variables" or "Stochastic variables". A random variable x with a microscopic probability $P(x)$ has the definitions:

Definition 2.1.1: EV and Variance

$$\text{Expectation value: } \langle x \rangle = \int x \cdot P(x) dx \tag{2.1}$$

$$\text{Variance: } \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{2.2}$$

We then introduce the concept of characteristic functions $\phi_x(k)$ that most of you probably have not seen before, which is the expectation value of e^{ikx}:

Definition 2.1.2: Characteristic function

$$\phi_x(k) = \langle e^{ikx} \rangle_{P(x)} = \int e^{ikx} P(x) dx \tag{2.3}$$

where we can see is just simply the Fourier Transform of $P(x)$! But if we expand the exponential in terms of the Taylor series:

$$\phi_x(k) = \int e^{ikx} P(x) dx \tag{2.4}$$

$$= \int \sum_{n=0}^{\infty} \frac{1}{n!} (ikx)^n P(x) dx \tag{2.5}$$

$$= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle \tag{2.6}$$

$$\tag{2.7}$$

where we call $\langle x^n \rangle$ the nth moment of the probability $P(X)$.

Definition 2.1.3: Moments of P(x)

$$n^{th}\ moment : \langle x^n \rangle = \int x^n P(x) dx \tag{2.8}$$

We can also express the characteristic equation in terms of an exponential of a sum:

$$\phi_x(k) = \exp\left[\sum_{n=0}^{\infty} \frac{(ik)^n}{n!} C_n\right] \tag{2.9}$$

where we call the coefficient C_n the cumulants of P(x) and we can calculate them directly from a probability distribution $P(x)$:

Definition 2.1.4: Cumulants

$$C_n = \frac{d^n}{d(ik)^n} \ln \phi_x(k) \mid_{k=0} \tag{2.10}$$

Example 2.1.1 (C_1)

$$C_1 = \frac{d}{d\,(ik)} \ln\left(\sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle\right)_{k=0} \tag{2.11}$$

$$= \left[\frac{1}{i} \frac{i \sum_m \frac{1}{(m-1)!} (ik)^{m-1} \langle x^m \rangle}{\sum_n (ik)^n / m! \langle x^n \rangle}\right]_{k=0} \tag{2.12}$$

$$= \langle x \rangle \tag{2.13}$$

Question 2: C_2

Check that $C_2 = \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2$

due to the Fourier transform nature of the characteristic function, we can see that if $P(x)$ is a Gaussian, then so is $\phi_x(k)$:

$$P = exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right) \Leftrightarrow \phi_x(k) = e\left(-\frac{\sigma^2 k^2}{2} + ikx_0\right) \tag{2.14}$$

2.2 Random Walk

Let's now consider a linked sequence y_N of random variables x_i with identical probability distributions $P(x)$ (For example the length of a polymer chain with N identical links randomly folded), we call this a "random walk" or a "stochastic process":

Definition 2.2.1: Random walk

$$y_N = \sum_{i=1}^{N} x_i \tag{2.15}$$
$$\langle y_N \rangle = N\langle x \rangle \tag{2.16}$$
$$\sigma_y^2 = \left\langle y_N^2 \right\rangle - \langle y_N \rangle^2 = N\sigma^2 \tag{2.17}$$
$$\tag{2.18}$$

where the variance relationship came from:

$$\left\langle y^2 \right\rangle - \langle y \rangle^2 = \sum_i \sum_j \left\langle x_i x_j \right\rangle - \sum_i \langle x_i \rangle \sum_j \left\langle x_j \right\rangle \tag{2.19}$$
$$= N\left\langle x^2 \right\rangle + N(N-1)\langle x \rangle^2 - N^2 \langle x \rangle^2 \tag{2.20}$$
$$= N\left(\left\langle x^2 \right\rangle - \langle x \rangle^2\right) = N\sigma^2 \tag{2.21}$$

Now we introduce the normalised random walk $S_N = y_N/N$ where $\langle S_N \rangle = \langle x_N \rangle$ and $\sigma_S^2 = \sigma_x^2/N$, and the central limit theorem suggests that no matter what $P(x)$ is, $P(S_N)$ will always converge to a Gaussian:

Theorem 2.2.1 Central Limit theorem

For a normalised random walk S_N where:

$$S_N = \frac{y_N}{N} \tag{2.22}$$
$$\langle S_N \rangle = \langle x \rangle \tag{2.23}$$
$$\sigma_S^2 = \frac{\sigma_x^2}{N} \tag{2.24}$$
$$\tag{2.25}$$

The probability distribution for S_N: $P(S)$ converges to a Gaussian with the same EV and Variance for large N.

Proof: (This is a very rough proof)
We start with the characteristic function $\phi_S(k)$:

$$\phi_s(k) = \left\langle e^{iks} \right\rangle = \left\langle exp\left(i\frac{k}{N}\sum_m x_m\right)\right\rangle \tag{2.26}$$
$$= \left\langle e^{\frac{ik}{N}x} \right\rangle_{P(x)}^N = \left[\phi_x(\frac{k}{N})\right]^N \tag{2.27}$$

On the other hand, we can also express $\phi_S(k)$ and $\phi_x(k)$ as:

$$\phi_s(k) = \exp\left[\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} C_m(s)\right] \tag{2.28}$$
$$= \left[\phi_x(k/N)\right]^N \tag{2.29}$$
$$= \left[\exp\left(\sum_{m=0}^{N} \frac{(ik)^m}{m!N^m} C_m(x)\right)\right]^N \tag{2.30}$$
$$= \exp\left(\sum_{m=0}^{N} \frac{(ik)^m}{m!N^{m-1}} C_m(x)\right) \tag{2.31}$$

From here, it is not hard to convince yourself that as N increases the terms in the sum with $m > 1$ will slowly disappear. So if we consider only the first two terms in the sum:

$$\phi_S(k) \approx \exp\left(N + ikC_1(x) - \frac{k^2}{2N}C_2(x)\right) \tag{2.32}$$

$$= \exp\left(N + ik\langle x\rangle - \frac{k^2}{2N}\sigma_x^2\right) \tag{2.33}$$

This is just the Fourier Transform of a Gaussian with the mean and variance mentioned above! Therefore, FTing it back to $P(S)$ gives us our desired result. 😄

2.3 Random walk in time

If we replace N in the random walk with time t, we reach a stochastic process $y(t)$ where we're randomly drifting around over continuous time. In this section, we go through a few key properties of this stochastic process.

2.3.1 probabilities

We can then define a set of probabilities:

Definition 2.3.1: Probabilities for random walks in time

$$P_1(y,t) : the\ probability\ of\ being\ at\ value\ y\ at\ time\ t \tag{2.34}$$

$$P_2(y_1, y_2|t_1, t_2) : probability\ of\ being\ at\ (y_1, t_1)\ and\ then\ (y_2, t_2) \tag{2.35}$$

$$\vdots\ and\ so\ on \tag{2.36}$$

2.3.2 Normalisation

These probabilities are normalised by:

$$1 = \int dy_1 \dots dy_N\ P_N\left(y_1, y_2 \dots y_N \mid t_1, t_2 \dots t_N\right) \tag{2.37}$$

$$\tag{2.38}$$

which is saying if we consider all possible y_i at the given times $t = t_i$, we should have a total probability of 1.

Example 2.3.1

$$1 = \int dy P_1(y,t) \tag{2.39}$$

2.3.3 Reduction

Theorem 2.3.1 Reduction

The probability P_N can be reduced into P_{N-1} by reduction:

$$\int dy_N P_N\left(y_1 y_2 \dots y_N \mid \cdots\right) = P_{N-1}\left(y_1, \cdots y_{N-1}\right) \tag{2.40}$$

Example 2.3.2

$$\int P_2\left(y_1 y_2 \mid t_1 t_2\right) dy_2 = P_1\left(y_1 \mid t_1\right) \tag{2.41}$$

2.3.4 Correlation functions

Correlation functions are defined in the following ways:

Definition 2.3.2: Correlation functions

$$\langle y(t)\rangle = \int y \cdot P(y,t) dy \tag{2.42}$$

$$\langle y_1\left(t_1\right) y_2\left(t_2\right)\rangle = \int y_1 y_2 p_2\left(y_1 y_2 \mid t_1 t_2\right) dy_1 dy_2 \tag{2.43}$$

$$etc \cdots \tag{2.44}$$

2.3.5 Stationary process

A stationary process is defined when $P_1(y,t)$ is invariant under time shift, in other words, not time dependent. For such process, the correlation function $\langle y(t)\rangle$ has no time dependence and the second correlation $\langle y_1\left(t_1\right) y_2\left(t_2\right)\rangle$ will only depend on $\Delta t = t_1 - t_2$.

2.3.6 Conditional probability

For the case where $t_2 > t_1$ it is very intuitive to think the probability $\langle y_1\left(t_1\right) y_2\left(t_2\right)\rangle$ can be written in terms of the product of $P_1(y_1|t_1)$ and some other function G, where G is the conditional probability of $y = y_2$ when $t = t_2$ given that $y = y_1$ when $t = t_1$.

Definition 2.3.3: The propagator G

$$P_2\left(y_2 y_1 \mid t_2 t_1\right) = G(2 \mid 1) P_1\left(y_1 \mid t_2\right) \tag{2.45}$$

2.3.7 Markov Process

A Markov process is a stochastic process in which the probability of each event depends only on the state attained in the previous event, or we say it's "memoryless". In our case, the probability of all P_i depends only on the single probability P_1 and our propagator G since the process has no "memory" of what happened before. Therefore if we want to calculate for example P_4, all we need is P_1 and then apply the propagator 3 more times.

Theorem 2.3.2 Markov Process

Any higher level probability P_i can be constructed by using just the single probability P_1 and the propagator G.

2.3.8 Evolution relation

If we consider the setting mentioned in the propagator section, we can actually combine the reduction relation Eq.(2.40) with the propagator Eq.(2.45) to reach something called the evolution relation:

$$P_1\left(y_2 t_2\right) = \int G\left(y_2 y_1 \mid t_2 t_1\right) P_1\left(y_1 t_1\right) dy_1 \tag{2.46}$$

2.3.9 Kolmogorov-Chapman relation

The last relation we're gonna talk about in this section is the Kolmogorov-Chapman relation. Now imagine we start with state 0 (y_0, t_0), go through a transition state 1 (y_1, t_1) at a given time $t = t_1$ and end up in state 2 (y_2, t_2). The propagator between state 0 and state 2 can be expressed by:

$$G\left(y_2, y_0 \mid t_2, t_0\right) = \int G\left(y_2, y_1 \mid t_2, t_1\right) G\left(y_1, y_0 \mid t_1, t_0\right) dy_1 \tag{2.47}$$

Chapter 3

Lecture 3:The Poisson Process

As we mentioned in the last chapter, a stationary Markov process is a process where 1. its higher probabilities P_i can all be derived from the single probability P_1 and the propagator G 2. The probabilities are all invariant with respect to time shifts. There are two stationary Markov processes are particularly important - the Poisson Process and the Wiener Process:

1. Poisson Process:
   - Independent steps
   - Only forward steps
   - Either go forward or not move after each step.
2. Wiener Process:
   - Independent steps
   - Can go both forward and backward
   - Must go either forward or backwards after each step, must move!

3.1 Poisson Process

Let's give some examples so we can have a clearer understanding. Imagine shooting a target with a pistol, let's call every time a bullet is fired a "step". So the number of bullets that hit the target is a Poisson Process since after each step the number can only increase by one (go forward 1 step) or not change if we miss (not move).

Formulating the above example in maths, we have a status of the system $N(t)$ and a parameter ν which is the rate of a single forward step (N=n-1 to n) to happen (or else we don't move). The probability P_+ to make a step in time dt is $P_+ = \nu \cdot dt$. And $N(t)$ is our stochastic process. And the most common answers that we're looking for when using this model are: 1. How far forward would we go in time t? 2. How much time would it take for us to reach a certain point n?

Definition 3.1.1: Poisson Process

1. Status of the system (integer): $N(t)$
2. Rate of going forward: ν
3. Probability to go forward a step in time dt: $P_+ = \nu \cdot dt$

If the system is in the state $N(t) = n - 1$ and we look into this single step going from (n-1) to (n), we can ask two questions: 1. What is the total probability $W(t)$ that the system makes this step after time t? 2.

Equivalently, what is the "survival probability" $S(t) = 1 - W(t)$ that the system still remains in the state $N = n - 1$ after time t? To obtain $S(t)$, we look at the relation between $S(t)$ and $S(t + dt)$:

$$S(t + dt) = S(t) \cdot [1 - \nu dt] \tag{3.1}$$

$$\frac{dS}{dt} = -v \cdot S \tag{3.2}$$

by using $S(t = 0) = 1$ we get:

$$S(t) = e^{-\nu t} \tag{3.3}$$

Definition 3.1.2: Survival Probability

$$S(t) = e^{-\nu t}$$

We can then get the average time it takes for the process to make this single step $\langle t_1 \rangle$ by using $W(t)$ and the probability density $w(t)$:

First define the probability density $w(t)$ of making the step:

$$w(t) = \frac{dW(t)}{dt} = -\frac{\partial S}{\partial t} = \nu e^{-\nu t} \tag{3.4}$$

Then the average time it takes to go forward this single step would be:

$$\langle t_1 \rangle = \int t \cdot w(t) dt = \frac{1}{\nu} \tag{3.5}$$

and with a similar method, we can also get:

$$\sigma_t^2 = \left\langle t_1^2 \right\rangle - \langle t_1 \rangle^2 = \frac{1}{\nu^2} \tag{3.6}$$

Question 3: Simulation excercise

We are constantly flipping a biased coin and the stochastic process is the total number of heads obtained. The constant step time Δt is the time it takes for each flip and the biased probability of flipping a head from the coin is p_+, we can express it in terms of the usual Poisson process rates:

$$p_+ = \nu \Delta t \quad \text{and} \quad p_0 = 1 - \nu \Delta t \tag{3.7}$$

What's the probability $P(k, N)$ of having exactly k heads after exactly N flips (or after time $T = N\Delta t$)? (k is just like our position in a Poisson process, and $N = T/\Delta t$ is effectively a type of expression for total time T.)

We can see trivially that this is a binomial process without the need for simulation where:

$$p(k, N) = \frac{N!}{k!(N - b)!} \cdot p_+^k \cdot (1 - p_+)^{N-k} \tag{3.8}$$

The interesting thing is that if you run a simulation with large enough N (small enough time step), p(k,N) reduces to:

$$p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{3.9}$$

Where $\lambda = Np_+ = \nu T$ is the expectation value of heads in the overall session. This is exactly the Poisson distribution!

Definition 3.1.3: Poisson distribution

Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time if these events occur with a known constant mean rate and independently of the time since the last event.

$$p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{3.10}$$

Where $\lambda = \nu T$ is the expected number of events in the given time interval.

3.2 The Poisson distribution

But how did we reach this result? Let's try to derive it using a standard Poisson process with external parameters position and time (n,t) and the rate ν. First, let's start with the probability of being at position n at time t+dt:

$$P(n, t + dt) = P(n, t) \cdot [1 - \nu dt] + P(n - 1, t) \cdot \nu dt \tag{3.11}$$

The first term on the RHS represents the process where it was already at position n and carry on staying at the same spot after this small time interval, and the second term corresponds to it being at position n-1 and hopping forward into our desired state n in the small time interval. Using this relation, we can find:

$$\frac{dP(n, t)}{dt} = [P(n, t + dt) - P(n, t)] / dt \tag{3.12}$$

$$= \nu \cdot [P(n - 1, t) - P(n, t)] \tag{3.13}$$

In the square brackets of the second line, the first term and second terms can be considered as the "flux in" and "flux out" of the position n. This is often called the "Master Equation". Recall from the previous chapter the characteristic function $\phi_x(k) = \left\langle e^{ikx} \right\rangle$, we can define a "generating function" $g(k, t)$ using a similar logic:

$$g(k, t) = \sum_{n=0}^{\infty} p(n, t) k^n \tag{3.14}$$

If we differentiate it in respect of t:

$$\frac{\partial g(k, t)}{\partial t} = \sum_{n=0}^{\infty} k^n \cdot v[p(n - 1, t) - p(n, t)] \tag{3.15}$$

and reindex the n in the second p into m-1 and regroup the two sums

$$= \sum_{m=0}^{\infty} v \left(k^{m+1} - k^m\right) P(m, t) \tag{3.16}$$

$$= v(k - 1) \sum_{n=0}^{\infty} k^n P(n, t) \tag{3.17}$$

we can see the sum is the same as $g(k, t)$! So we have:

$$\frac{\partial g}{\partial t} = v(k - 1) \cdot g \tag{3.18}$$

By using the following initial conditions, we obtain the exponential form of g:

$$P(n, t = 0) = \delta_{n,0} \Rightarrow g(t = 0) = 1 \tag{3.19}$$

$$g(k, t) = e^{(k-1)vt} \tag{3.20}$$

Now expand the exponential and compare it with the definition of g (Eq.3.14):

$$g(k,t) = e^{-\nu t} \sum_{n=0}^{\infty} \frac{1}{n!} (\nu t)^n \cdot k^n \tag{3.21}$$

$$= \sum_{n=0}^{\infty} p(n,t) k^n \tag{3.22}$$

We reach our conclusion:

$$P(n,t) = \frac{1}{n!} (\nu t)^n e^{-\nu t} \tag{3.23}$$

Which is the same as Eq. 3.10. We can then ask some simple questions about the expected values or probability of the process:

Example 3.2.1

What is the expectation value of the occurring time of the 10th event?

Solution: Considering each step is independent, the answer is simply:

$$\langle t_{10} \rangle = 10 \langle t_1 \rangle = 10/v \tag{3.24}$$

Example 3.2.2

What is the probability that the 10th event occurs after time t?

Solution:
There are two ways of doing this: 1. Consider the probability $P(t_{10} > t)$: Which we want to integrate the probability that the 10th event happens at exactly $t = t'$ over $t' = t \sim \infty$:

$$P(t_{10} > t) = \int_t^{\infty} P(t_{10} = t') dt' \tag{3.25}$$

And we write this probability as a product of having exactly 9 events and having another one during the time interval dt':

$$= \int_t^{\infty} \frac{(\nu t')^9}{9!} e^{-vt'} \cdot v dt' \tag{3.26}$$

2. The second way is much simpler, we just consider $P(n(t) < 10)$, which is:

$$P(n(t) < 10) = \sum_{j=0}^{9} \frac{(\nu t)^j}{j!} e^{-\nu t} \tag{3.27}$$

3.3 Evolution of the Master Equation

Using the evolution relation (Eq. 2.46), we can write out the evolution relation for any stationary Markov process distribution (not just Poisson):

(The integral has been replaced by a sum since m is discrete)

$$P(n, t + \Delta t) = \sum_m G(n, t + \Delta t \mid m, t) \cdot P(m, t) \tag{3.28}$$

And if we subtract $P(n,t)$ from both sides:

$$\frac{\partial P(n,t)}{\partial t} = \frac{1}{\Delta t}\left[\sum_m G(n,t+\Delta t m,t)P(m,t) - P(n,t)\cdot\sum_m G(m,t+\Delta t \mid n,t)\right] \tag{3.29}$$

where:

$$\sum_m G(m,t+\Delta t \mid n,t) = 1 \tag{3.30}$$

we get:

$$= \sum_m \frac{G\left(n,t+\Delta t \mid m,t\right)}{dt}P(m,t) - \sum_m \frac{G(m,t+\Delta t \mid u,t)}{dt}P(u,t) \tag{3.31}$$

$$= \sum_m \left[w_{nm}P(m,t) - w_{mn}P(n,t)\right] \tag{3.32}$$

Where w_{fi} is the transition rate from initial state $\boldsymbol{i}$ to final state $\boldsymbol{f}$.

$$w_{fi} = \text{limit}_{\ dt\to 0}\frac{1}{dt}G(j,t+dt \mid i,t) \tag{3.33}$$

If we are considering the Poisson process, we need to imply the restrictions of $n > m$ and $n < m$ respectively on the two sums, but what we derived here is more general and did not include this restriction.

Chapter 4

Lecture 4: The Wiener process

As the title suggests, we will be looking at the other stationary Markov process that we mentioned at the beginning of the last chapter, the Wiener process.

4.1 The Wiener process

So what is the Wiener process? Let's start with the definition of a random walk (Section 2.2), where you have a random variable $\boldsymbol{x}$ (a single step) with probability density $P(\boldsymbol{x})$ and a linked sequence $\boldsymbol{y}_N$, which can be converted into a time-dependent function $\boldsymbol{y}(t)$ with a stepping rate. We call this the Wiener process when the probability distribution of independent step $P(\boldsymbol{x})$ is the Gaussian probability. And we usually use the expression $W(t)$ to represent a Wiener process (rather than $y(t)$).

Definition 4.1.1: Wiener process

A random walk $W(t)$ where the increment ΔW has the Gaussian probability of variance $\sigma^2 = \Delta t$

$$\Delta W = W(t + \Delta t) - W(t) \tag{4.1}$$

$$P(\Delta w) = \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{\Delta w^2}{2\Delta t}\right) \tag{4.2}$$

And then there are a few important properties of the Wiener process:

4.1.1 Characteristic function

It is easy to check (Fourier Transform) that the characteristic function ϕ of ΔW is:

Theorem 4.1.1 Wiener Characteristic Function

$$\phi(k, t_0 + \Delta t) = \left\langle e^{ik(w_0 + \Delta w)} \right\rangle_{P(\Delta w)} \tag{4.3}$$

$$= \exp\left(ikw_0 - \frac{1}{2}k^2 (\Delta t - t_0)\right) \tag{4.4}$$

where $w(t_0) = w_0$.

4.1.2 Scaling

If we speed up (or slow down) the rate of the process by a scaling factor c, this process can be linearly scaled into a Wiener-Process:

Theorem 4.1.2 Scaling invariance

$$V(t) = \frac{1}{\sqrt{c}} W(ct) \tag{4.5}$$

is a Wiener process. The scaling factor $1/\sqrt{c}$ came from the exponent in Eq. (4.2) since we need to keep the variance of Δv as Δt. (If time flows faster and we're taking more steps per time, the steps must be smaller to keep the overall variance the same.)

4.1.3 Time inversion

This is a bit of a weird one to be interpreted physically.

Theorem 4.1.3 Time inversion invariance

$$V(t) = t \cdot W(1/t) \tag{4.6}$$

is also a Wiener process.

4.1.4 Time reversal

If we reverse the time and relabel $W(t = A)$ to be the new origin, this is still a wiener process:

Theorem 4.1.4 Time reversal invariance

$$V(t) = W(A) - W(A - t) \tag{4.7}$$

is a Wiener process.

4.1.5 Probability distribution

Using the propagator $G = P(\Delta w)$ from the definition of a Wiener process:

$$P(w + \Delta w, t + \Delta t) = \int_{-\infty}^{\infty} G(\Delta w, \Delta t) P(w, t) dw \tag{4.8}$$

And using the relation $\Delta w = w(t + \Delta t) - w(t)$, we can change the variables and change the integral $dw = -d\Delta w$ (think $w(t + \Delta t)$ as a constant)

$$= -\int_{\infty}^{-\infty} G(\Delta w, \Delta t) P(w(t + \Delta t) - \Delta w, t) d\Delta w \tag{4.9}$$

flipping the interval of integration back

$$= \int_{-\infty}^{\infty} G(\Delta w, \Delta t) P(w(t + \Delta t) - \Delta w, t) d\Delta w \tag{4.10}$$

taking out the $-\Delta w$ term since it's $<< w(t)$

$$= \int_{-\infty}^{\infty} G(\Delta w, \Delta t) P(w(t+\Delta t), t) d\Delta w \tag{4.11}$$

$$- \int G(\Delta w, \Delta t) \Delta w \frac{\partial P}{\partial w} d\Delta w \tag{4.12}$$

$$+ \frac{1}{2} \int G(\Delta w, \Delta t) \Delta w^2 \frac{\partial^2 P}{\partial w} d\Delta w \tag{4.13}$$

The integral in the first line over G is 1 (P doesn't depend on Δw), due to normalisation over all possible Δw. The second line is 0 since it's a product of an even function G and odd function Δw:

$$P(w+\Delta w, t+\Delta t) - P(w(t+\Delta t), t) = \frac{1}{2} \int G(\Delta w, \Delta t) \Delta w^2 d\Delta w \frac{\partial^2 p}{\partial w^2} \tag{4.14}$$

by using $\Delta w = w(t+\Delta t) - w(t)$ and the definition of variance:

$$\frac{\partial P(w,t)}{\partial t} = \frac{1}{2} \frac{\sigma_{\Delta w}^2}{\Delta t} \cdot \frac{\partial^2 P(w,t)}{\partial w^2} \tag{4.15}$$

wrapping the terms up with D, we reach the diffusion equation:

$$\frac{\partial P(w,t)}{\partial t} = D \frac{\partial^2 P(w,t)}{\partial w^2} \tag{4.16}$$

From the derivation above, we conclude that the diffusion equation (4.16) is how $P(w,t)$ evolves over time. There is no universal solution to this equation since it depends on the initial boundary condition $P(w, t=0)$. Although there is a special case where we start from a single point w_0, setting the boundary condition $P(w, t=0)$ to be $\delta(w = w_0)$. We get a Gaussian probability distribution:

Theorem 4.1.5 Probabilty distribution of W(t)

$$P(w,t) = w_0 + \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)} \tag{4.17}$$

with variance $\sigma_w^2 = t$.

It's easy to check that Eq. (4.17) is a solution of the diffusion differential equation and converges to a delta function $P(w) = \delta(w - w_0)$ at t=0. (Plug it in yourself!) Also, by using the variance $\sigma_{\Delta w}^2$ of Δt and Eq. (2.17), it's not hard to be convinced after time t ($N = t/\Delta t$ steps), the probability distribution of W(t) : $P(w,t)$ will still be a Gaussian distribution if we start from a single point.

The equation for the global functions (such as the diffusion equation for the global probability distribution $P(w,t)$) are called "kinetic" or Macroscopic. And on a microscopic scale of $W(t)$, they're governed by "Stochastic differential equations", SDE.

4.2 Stochastic differential equations

Here's a fun little fact. With so many "Physics-inspired" algorithms/models nowadays, the solution of Brownian motion was actually "trading-inspired". In 1900, Louis Bachelier presented the method of stochastic analysis of stock and option markets in his PhD thesis "The theory of speculation". Then in 1905 Albert Einstein and Marian Smoluchowski brought the solution of Brownian motion to the attention of physicists and presented it as a way to indirectly confirm the existence of atoms and molecules. Later in 1911, Langevin proposed a more general approach to formulating and analysing SDEs.

Let's start with something that we're all familiar with, the stochastic differential equation for Brownian Motion:

$$m\frac{du}{dt} = -\gamma u + \xi(t) \tag{4.18}$$

where u is the velocity of the particle and ξ is a stochastic force. More generally, any SDE can be written like this:

$$\frac{dx}{dt} = \underbrace{F(x)}_{\text{drift term}} + \underbrace{\sigma(x)\Delta W(t)}_{\text{diffusion term}} \tag{4.19}$$

where W is the normalised Wiener process, and $\sigma(x)$ is the amplitude which may depend on x. The reason why the second term is called the diffusion term is that if we have only the second term, the differential equation can be written in the diffusion equation form with diffusion equation $D = \frac{1}{2}\sigma^2$.

Mathematically (mathematician uses this notation because they don't trust the physicists with maths), the same SDE above can be written as:

$$dx_t = \mu(x,t)dt + \omega(x,t)dw_t \tag{4.20}$$

since we have a wiener term dw_t, if the steps dt is small enough, we expect dx_t to be Gaussian as well (have gaussian probability distribution).

4.3 Example: Geometric Brownian Motion

The difference between geometric and typical Brownian motion is that for each step Δt, the value of the motion is multiplied by the step x rather than added. Just like a geometric vs an arithmetic series, hence the "geometric" in the name.

Definition 4.3.1: Geometric Brownian Motion

A stochastic process such that:

$$dS_t = \mu \cdot S_t \cdot dt + \sigma \cdot S_t dw_t \tag{4.21}$$

which is also known as the "percentage drift" or the "percentage volatility".

To solve this we rearrange the equation above into:

$$\frac{dS_t}{S_t} = \mu \cdot dt + \sigma \cdot dw_t \tag{4.22}$$

which you will very intuitively write into:

Wrong!

$$d\ln S_t = \cancel{d(\mu t + \sigma w_t)} \tag{4.23}$$

But this second line above is actually WRONG!!! $\int \sigma dw_t$ is not actually σw_t! The problem stems from the cheeky term $\int dw_t$. Although in our notation we wrote the infinitesimal interval dw_t in the integral, the function w_t is not actually a smooth continuous differentiable function (The mathematicians lied to us rip). The solution that we deal with this has a name called "Itoh Calculus".

4.4 Itoh Lemma

In this section, it's all mathematician BS, so we used some big words to blend in. So the Itoh lemma was formed to solve the following question:

Question 4: Itoh Lemma

If we have a stochastic process $x(t)$ with SDE:

$$dx_t = \mu dt + \sigma dw_t \tag{4.24}$$

and a fucntion $f(x)$ that depends on x. What is the SDE for f?

First we start with the definition of df:

$$df = f(x + dx) - f(x) \tag{4.25}$$

which the terms may all have explicit time dependence. We Taylor expand them:

$$= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2 + \cdots\cdots \tag{4.26}$$

then using $dx_t = \mu(x,t)dt + \omega(x,t)dw_t$ from Eq(4.20) and only keeping terms linear in dt:

(Note that $\left\langle dw_t^2 \right\rangle = dt$ for fast steps with small time interval)

$$= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(\mu\delta t + \sigma dw) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\left(\mu^2 \cancelto{0}{dt^2} + 2\mu\sigma \cancelto{0}{dtdw_t} + \sigma^2 dw_t^2\right) \tag{4.27}$$

Now:

$$df = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\right)\cdot dt + \underbrace{\sigma\frac{\partial f}{\partial x}\cdot dw_t}_{\text{diffusion term}} \tag{4.28}$$

Lenma 4.4.1 Itoh Lemma

If we have a stochastic process $x(t)$ with SDE:

$$dx_t = \mu dt + \sigma dw_t \tag{4.29}$$

and a fucntion $f(x)$ that depends on x. The SDE for f is:

$$df = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\right)\cdot dt + \sigma\frac{\partial f}{\partial x}\cdot dw_t \tag{4.30}$$

Now applying the same analysis to Geometric Brownian Motion Eq.(4.21) and consider the function $f(S) = \ln S_t$:

$$d(\ln S) = \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}dS^2 \tag{4.31}$$

$$= \frac{dS}{S} - \frac{1}{2S^2}dS^2 \tag{4.32}$$

$$= \frac{dS}{S} - \frac{1}{2S^2}\left(\mu^2 S^2 \cancelto{0}{dt^2} + 2\mu\sigma S^2 \cancelto{0}{dtdw} + \sigma^2 S^2 dw^2\right) \tag{4.33}$$

$$= \frac{dS}{S} - \frac{\sigma^2}{2}dt \tag{4.34}$$

$$= \mu\cdot dt + \sigma\cdot dw_t - \frac{\sigma^2}{2}dt \tag{4.35}$$

which gives:

$$S = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma w_t\right) \tag{4.36}$$

Definition 4.4.1: Solution of Geometric Brownian Motion

$$S(t) = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma w_t\right) \tag{4.37}$$

From this we can see two interesting things:

1. Since dw_t is a small step Wiener process, we have a diffusive growth $e^{\sigma w_t}$
2. If we solved the differential equation without the diffusion term we would've reached the solution:

$$dS = \mu s dt \longrightarrow S = S_0 e^{\mu t} \tag{4.38}$$

   But instead we have a shift in the deterministic drift term, giving us a lower growth rate:

$$S \propto \exp\left(\mu - \frac{\sigma^2}{2}\right)t \tag{4.39}$$

Try using the code on the GBM Wikipedia page to produce GBM graphs in Python!

Chapter 5

Lecture 5: Black-Scholes equation

In this lecture, we will be using the two things we introduced in the last lecture, Etoh's lemma and Geometric Brownian Motion, to study the Black-Scholes equation, which won the Nobel prize for economics in 1997. So what is the Black-Scholes equation and how is it useful? It is a mathematical analysis of dynamics and prediction of option prices in volatile markets that are stochastic. The term "option" might not be familiar to some readers, in simple terms, the B-S equation gives investment managers a tool to allocate their risk and cash. For example, buying more stocks is allocating more risk and less cash.

Interestingly, there are theories that suggest that the 1987 financial crisis was caused by a group of traders that conspired to punish the early adopters of the B-S analysis. Even more interestingly, one of the main causes of the 2008 financial crisis is actually too much trust in the B-S analysis and incorrect use of it (Since they didn't pick the ASM option in Lent like you obviously lol).

5.1 Black-Scholes Equation

So first we use a stochastic variable $S(t)$ to present the value of a given stock, which is governed by the SDE of Geometric Brownian Motion (GBM):

$$dS = \mu S dt + \sigma S dw \tag{5.1}$$

since in stocks we usually describe the growth in percentage growth and not linear growth. An "Option" is a contract when you (the seller) agree to sell me (the buyer) a certain stock $\mathcal{S}$, at a future time t, at a "strike price" E. So you see the seller hopes the stock to depreciate and then the opposite for the buyer. You might think, why would the seller agree to a such weird contract? They already have the stock in hand and can sell them at any time that they want! The only reason that a seller will be interested in such a contract is if the buyer (me) pays a small extra fee for this. We denote this fee as another stochastic variable $V(s,t)$, which depends on the value of the stock S and time t. A "Portfolio" Π is a combination of assets (cash, stocks, options etc.) that I currently have:

$$\Pi = -V(s_t) + \alpha S \tag{5.2}$$

where the $-V$ came from the amount that you have already paid for the options contract, and αS from selling a portion α of that same stock at the same time to "hedge" aka decrease risk exposure (Just like the hedge in "Hedge fund"). In continuous trading, we have small increments of time, therefore:

$$\Delta\Pi = -\Delta V(s) + \alpha \Delta S \tag{5.3}$$

Now, we can use the Itoh Lemma 4.30:

$$dV(s) = \left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial s} ds + \frac{1}{2}\frac{\partial^2 V}{\partial s^2} ds^2\right) \tag{5.4}$$

And subbing in:

$$ds = \mu s dt + \sigma s dw \tag{5.5}$$

We get:

$$dV(s) = \left(\frac{\partial v}{\partial t} + \mu s \frac{\partial v}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2}\right) dt + \sigma s \frac{\partial V}{\partial s} dW \tag{5.6}$$

Then substitute dV and dS into $d\Pi$:

$$d\Pi = -\frac{\partial v}{\partial t} dt - \mu s \frac{\partial v}{\partial s} dt - \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v}{\partial s^2} dt + \alpha \mu s dt \underbrace{- \sigma s \frac{\partial v}{\partial s} dW + \alpha \sigma s dW}_{\text{stochastic part}} \tag{5.7}$$

With our stochastic description of $d\Pi$, the key step of reaching the BS equation is to choose the factor α to be $\partial V / \partial S$ so the volatility terms above cancel out and we have no volatility in $d\Pi$. Then, we make a decision on what we want to achieve for $d\Pi$ (The goal of our investments), a steady geometric growth:

$$d\Pi = r \cdot \Pi dt \tag{5.8}$$

Getting us the relation with rate r:

$$r \cdot \overbrace{(\alpha S - V)}^{\Pi} dt = \left(-\frac{\partial V}{\partial t} - \cancel{\mu s \frac{\partial V}{\partial s}} - \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial S^2} + \cancel{\alpha \mu s}\right) dt \tag{5.9}$$

$$\tag{5.10}$$

Finally we have the B-S equation:

Definition 5.1.1: Black-Scholes Equation

$$\frac{\partial V(s,t)}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} = r\left(V - s\frac{\partial V}{\partial s}\right) \tag{5.11}$$

It allows us to solve for $V(s,t)$ for any given r and σ. The lack of dependence on μ makes it even more powerful since if we look back to our SDE Eq.5.1 for $S(t)$, it is μ that makes the growth of each stock different and the dW term is just geometric stochastic term that represents the market volatility and does not relate to the growth.

What about r? What are the restrictions on this risk-free geometric growth? The "desired legislation" should have this rate of risk-free growth fixed (limited). We will not go into too much detail about the impact of r in the overall economy since we're physicists not econ students. In reality, we use the rate of US Treasury Bonds r_0 (effectively the interest rate you get by lending money to the US government) since they are such a global superpower and the capitalists believe lending money to them is risk-free lol.

5.2 A simple case

Let's consider the case where $r = 0$ and see how the Black-Scholes equation is useful if we were a hedge fund manager back in the 1980s skrskr:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 V}{\partial S^2} = 0 \tag{5.12}$$

Later in the course, you will see that this is actually the "super diffusion" equation (Another great naming from the physicists ofc) since the diffusion constant $D = \frac{\sigma^2}{2} S^2$ grows as S^2, the effective position (non-linear diffusion).

Using this equation, we can solve for the fair value of the contract V at the current time and buy all options that are underpriced and sell all the overpriced ones.

(Note: There's another interesting question to ask: Do we really know the value of volatility σ? To this day, the measurement or modelling of σ is still a very competitive field between hedge funds and a lot of different analyses are still being proposed.)

With all that waffling above, how do we actually solve for $V(s,t)$ using this equation? Using the current price($t = 0$) of a stock S and the strike price E of the stock at the end of the contract ($t = T$), we define:

$$x = \ln S/E \tag{5.13}$$
$$S = Ee^x \tag{5.14}$$
$$\tau = T - t \tag{5.15}$$

re-express V into:

$$V\left(Ee^x, T - \tau\right) = Z(x, \tau) \tag{5.16}$$

we get:

$$\frac{\partial z}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 z}{\partial x^2} + \left(\frac{\sigma^2}{2} - r\right)\frac{\partial z}{\partial x} + r = 0 \tag{5.17}$$

Which is a homogeneous differential equation! In order to solve this, there's a beautiful solution by introducing a new variable u:

$$u(x,\tau) = z(x,\tau) \cdot e^{\alpha x + \beta \tau} \tag{5.18}$$

plugging it into the PDE above:

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + A(\alpha,\beta) \cdot \frac{\partial u}{\partial x} + \underbrace{B(\alpha,\beta)}\cdot u = 0 \tag{5.19}$$

choose an α and β that yields $A = B = 0$: (try it yourself!)

$$\alpha = r/\sigma^2 - 1/2 \tag{5.20}$$
$$\beta = r/2 + \sigma^2/8 + r^2/2\sigma^2 \tag{5.21}$$

and all we have left is:

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} = 0 \tag{5.22}$$

Amazing right? The last step is to simply solve this by using Green's function and its initial condition:

$$u(x,\tau) = \int \underbrace{\frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left(-\frac{(x-w)^2}{2\sigma^2\tau}\right)}_{\text{Green's function}} \cdot \underbrace{u(w,0)}_{\text{initial condition}} \quad dw \tag{5.23}$$

and hence we can backroll and find Z, and $V(s,t)$.

Chapter 6

Lecture 6: Ornstein-Uhlenbeck process

After all that discussion about financial maths, in this lecture let's go back to physical systems and talk about the Ornstein-Uhlenbeck process.

6.1 Ornstein-Uhlenbeck process

Once again, let's start with the most general form of an SDE:

$$dx_t = \mu(x,t)dt + \sigma(x,t)dw \tag{6.1}$$

The O-U process is when μ and σ are chosen to be:

Definition 6.1.1: Ornstein-Uhlenbeck process SDE

$$dx_t = \Theta \cdot (x_0 - x)\, dt + \sigma dw \tag{6.2}$$

where Θ, x_0, σ are constants. This might seems like something completely new but actually, we have seen this before! Let's look at this equation:

$$m\dot{v} = -\gamma v + \sqrt{2k_B T \cdot \gamma}\xi(t) \tag{6.3}$$

this is the brownian motion equation in the Part II TSP course, also known as the free diffusion equation. If we divide both sides by m and times by dt:

$$dv = -\frac{\gamma}{m} v \cdot dt + \sqrt{\frac{2k_B T \cdot \gamma}{m}} dW \tag{6.4}$$

In comparison to Eq 6.2, this is the same equation with $x = v$, $\Theta = \gamma/m$, $x_0 = 0$, $\sigma = \sqrt{\frac{2k_B T \cdot \gamma}{m}}$, and $dW = \xi(t)$.

Or similarly, the diffusion equation in a spring potential:

$$\gamma \dot{x} = -k\,(x - x_0) + \sqrt{2k_b T \cdot \gamma}\xi(t) \tag{6.5}$$

rearrange:

$$dx = -\frac{k}{\gamma}\,(x - x_0)\, dt + \sqrt{\frac{2k_B T}{\gamma}} dw \tag{6.6}$$

where we can see this is just another Ornstein-Uhlenbeck SDE with a different choice of the constants. So why is the O-U important? Seems like just another SDE with a niche form? The reason lies in the linear term

$\Theta \cdot (x_0 - x)$. Most physical processes near equilibrium can be approximated with a linear restoring force, or a parabolic potential (think about Taylor-expansion). So the O-U process is what describes the dynamic of a stochastic system near equilibrium.

Before going into more details about O-U, let's revise some basic mechanics:

relation between position and velocity:

$$x(t) = \int_0^t v(s)ds \tag{6.7}$$

using the product of this and taking the ensemble average(just think of it as the expected value):

$$\left\langle x^2(t) \right\rangle = \int_0^t ds_1 \int_0^t ds_2 \left\langle v\left(s_1\right) v\left(s_2\right) \right\rangle \tag{6.8}$$

taking the time derivative, we should have 2 identical integrals since $s1, s2$ is symmetrical under exchange:

$$\frac{d}{dt} \left\langle x^2(t) \right\rangle = 2 \int_0^t ds \langle v(t)v(s) \rangle \tag{6.9}$$

Using the relation we got above and comparing it to the diffusion equation $\left\langle x^2 \right\rangle = 2Dt$, we see that:

$$D = \int_0^t ds \langle v(t)v(s) \rangle \quad \text{(1D case)} \tag{6.10}$$

$$\text{or} \quad D = \frac{1}{N} \int_0^t ds \langle v(t)v(s) \rangle \quad \text{(ND case)} \tag{6.11}$$

in the end we shift the time by $\tilde{s} = t - s$ so it looks neater:

$$D = \int_0^t d\tilde{s} \langle v(\tilde{s})v(0) \rangle \tag{6.12}$$

Let's go back to the free diffusion equation 6.3. If we times both sides by $v(0)$ and take the average over the whole ensemble:

$$m \frac{d}{dt} \langle v(t) \cdot v(0) \rangle = -\gamma \langle v(t)v(0) \rangle + C \cancelto{0}{\langle \xi(t) \cdot v(0) \rangle} \tag{6.13}$$

$$= -\gamma \langle v(t)v(0) \rangle \tag{6.14}$$

The last term canecls out due to the stochastic nature of ξ, so:

$$\langle v(t)v(0) \rangle = \langle v(0)^2 \rangle e^{-\frac{\gamma}{m}t} \tag{6.15}$$

then subbing it back into Eq 6.12:

$$D = \left\langle v_0^2 \right\rangle \int_0^t ds e^{-\frac{\gamma}{m}s} = \left\langle v_0^2 \right\rangle \frac{m}{\gamma} \left(1 - e^{-\frac{\gamma}{m}t}\right) \tag{6.16}$$

We have obtained a time dependent value of D instead of the usual constant form. How do we physically interpret this exponential decay in time? The exponential term in the self correlation of v in Eq. 6.15 can be interpreted as a "memory" with relaxation time scale m/γ. If our time t is much greater than this "time of velocity relaxation", which the give the system enough time to forget its initial conditions and reach its equilibrium, we can obtain a contant form of D:

$$D \approx \frac{m}{\gamma} \left\langle v_0^2 \right\rangle = \frac{k_B T}{\gamma} \quad \text{(we used the 2D case where } m \left\langle v_0^2 \right\rangle = k_B T\text{)} \tag{6.17}$$

Definition 6.1.2: Green-Kubo formula

For a time that's much longer than the relaxation time, we define the usual diffusion constant that we use to be:

$$D = \int_0^\infty \langle v(t)v(0) \rangle dt \tag{6.18}$$

6.2 O-U process with multiple variables

By choosing the equilibrium point $x_0 = 0$ we have the usual single variable O-U SDE:

$$dx = -\theta x dt + \sigma dw \tag{6.19}$$

Now if we want to make it multi-variable:

$$dx_i = -\theta_{ik} x_k dt + \sigma_{ik} dw_k \tag{6.20}$$

$$\text{or} \quad \frac{d}{dt} x_i = -\theta_{ik} x_k + \sigma_{ik} \xi_k \tag{6.21}$$

The reason of the matrices θ_{ik}, σ_{ik} is so variables can be correlated to each other so we don't just boringly have one independent ODEs for each variables. To solve this, we convolve the initial condition with the Green's function:

$$x_i(t) = \int_0^t \underbrace{e^{-\theta_{ik}(t-s)}}_{G(t-s)} \cdot \sigma_{ke} \xi_l^{(s)} ds \tag{6.22}$$

If you are not quite convinced by this, you can check it by differentiation and it gives you the same as Eq.6.21.

Now we construct a matrix M_{ik} to represent the correlation functions:

$$M_{ik} = \langle x_i(t) x_k(t) \rangle \tag{6.23}$$

subbing in Eq. 6.22:

$$M_{ik} = \int_0^t ds_1 \int_0^t ds_2 \;\; e^{-\theta_{im}(t-s_1)} e^{-\theta_{kp}(t-s_2)} \cdot \sigma_{ml} \sigma_{pq} \cdot \langle \xi_l(s_1) \xi_q(s_2) \rangle \tag{6.24}$$

since Wiener processes are not related to one another, the average term can be written into:

$$\left\langle \xi_e(s_1) \xi_q(s_2) \right\rangle = \delta_{qe} \delta(s_1 - s_2) \tag{6.25}$$

therefore:

$$M_{ik} = \int_0^t ds e^{-\theta_{im}(t-s)} \sigma_{ml} \cdot \sigma_{pl} e^{-\theta_{kp}(t-s)} \tag{6.26}$$

$$\underline{\underline{M}} = \int_0^t ds e^{-\underline{\underline{\theta}}(t-s)} \underline{\underline{\sigma}} \cdot \underline{\underline{\sigma}}^\top \cdot e^{\left(-\underline{\underline{\theta}}^\top t - s\right)} \tag{6.27}$$

taking t to infinity (to be in the equilibrium state) and shifting time $(t-s) \longrightarrow \tilde{t}$ to make it look nicer:

$$\underline{\underline{M}} = \int_0^\infty ds \exp\left(-\underline{\underline{\theta}} \tilde{t}\right) \underline{\underline{\sigma}} \cdot \underline{\underline{\sigma}}^\top \cdot \exp\left(-\underline{\underline{\theta}}^\top \tilde{t}\right) \tag{6.28}$$

Example 6.1 (Single variable case)
or the 1 variable case, the equation above becomes:

$$M = \left\langle x^2 \right\rangle = \int_0^\infty dt e^{-2\theta t} \cdot \sigma^2 = \frac{\sigma^2}{2\theta} \tag{6.29}$$

which gives the value of $\left\langle x^2 \right\rangle$ in equilibrium state. If we use it on the Eq.6.4, we get:

$$\left\langle v^2 \right\rangle = \frac{2 k_B T \gamma}{2 \gamma m} = \frac{kT}{m} \tag{6.30}$$

Let's now construct another matrix that's related to the fluctuation-dissapation relationship(we love matrices rip):

$$\underline{\underline{\theta}} \cdot \underline{\underline{M}} + \underline{\underline{M}} \cdot \underline{\underline{\theta}}^\top \tag{6.31}$$

remember, neither θ nor σ need to be symmetric (I hate these double underlines):

$$\underline{\underline{\theta M}} + \underline{\underline{M\theta^\top}} = \int_0^\infty dt \left[\underline{\underline{\theta}} \cdot e^{-\underline{\underline{\theta}}t} \underline{\underline{\sigma\sigma^\top}} e^{-\underline{\underline{\theta}}^\top t} + e^{\underline{\underline{\theta}}t} \underline{\underline{\sigma\sigma^\top}} \cdot e^{-\underline{\underline{\theta}}^\top t} \cdot \underline{\underline{\theta}}^\top \right] \tag{6.32}$$

noticing that the square bracket can be written into a time derivative:

$$= -\int_0^\infty dt \frac{d}{dt} \left[e^{-\underline{\underline{\theta}}t} \underline{\underline{\sigma\sigma^\top}} e^{-\underline{\underline{\theta}}^\top t} \right] \tag{6.33}$$

and the square bracket goes to 0 for t towards ∞ so the integral becomes:

$$= \underline{\underline{\sigma\sigma^\top}} \tag{6.34}$$

In the end, this gives us:

$$\theta_{ik} \langle x_k x_l \rangle + \langle x_i x_k \rangle \theta_{e_k} = \sigma_{ik}\sigma_{lk} \tag{6.35}$$

So why do we care about this weirdly looking matrix? It gives us a direct way of solving for all correlator in equilibrium between variables $\langle x_i x_j \rangle$ using the constant matrices σ and θ.

Chapter 7

Lecture 7: More on Ornstein-Uhlenbeck

Last lecture we talked about the O-U process and its multivariable form. In reality, SDEs are very hard to solve, instead, it would be better if we can find its probability distribution $P(x,t)$. To achieve that, we transform the SDE to something called the "kinetic equation" for P, which is an equation that can be differentiable properly just like P.

7.1 Kinetic equation of the O-U process

Let's start with the evolution equation 2.46 (rewriting it a little):

$$P(x,t+\Delta t) = \int G(x,t+\Delta t \mid y,t)P(y,t)dy \tag{7.1}$$

Normally when we think of a propagator $G(x \mid y)$, we have a fixed y (starting point) and propagate to a variable target x but in this equation, x is the fixed target and y is the variable starting point. Let's introduce another variable $\Delta x = x - y$ so $y = x - \Delta x,\ dy = -\Delta x$, and rewrite the equation above:

$$P(x,t+\Delta t) = \int G(x-\Delta x+\Delta x, t+\Delta t \mid x-\Delta x, t)P(x-\Delta x,t)d[-\Delta x] \tag{7.2}$$

Taylor expanding in $(-\Delta x)$:

$$P(x,t+\Delta t) = \int d(\Delta x)\sum_{n=0}^{\infty}\frac{(-\Delta x)^n}{n!}\frac{\partial^n}{\partial x^n}\left[G(x+\Delta x, t+\Delta t \mid x,t)P(x,t)\right] \tag{7.3}$$

Let's define the Moments to be:

Definition 7.1.1: nth-Moments of G

$$\langle \Delta x^n \rangle = \int d(\Delta x)(\Delta x)^n G(x+\Delta x \mid x) \tag{7.4}$$

For the O-U process, the moments are: (from now on we only keep terms linear in Δt for the same reasoning in Itoh analysis)

$$\left\langle \Delta x^0 \right\rangle = 1 \quad \text{(Normalisation of G)} \tag{7.5}$$

$$\left\langle \Delta x^1 \right\rangle = -\theta x \Delta t + \cdots \tag{7.6}$$

$$\left\langle \Delta x^0 \right\rangle = \sigma^2 \Delta t + \cdots \tag{7.7}$$

$$\tag{7.8}$$

Therefore the Taylor expansion above becomes:

$$P(x,t+\Delta t) = P(x,t) - \frac{\partial}{\partial x}\left(\langle \Delta x \rangle P(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\left\langle \Delta x^2 \right\rangle P(x,t)\right)\cdots \tag{7.9}$$

$$= P(x,t) - \frac{\partial}{\partial x}\left(-\theta x \Delta t P(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 \Delta t P(x,t)\right) \tag{7.10}$$

and we reach our kinetic equation:

Definition 7.1.2: Kinetic equation of the O-U process

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(\theta x P(x,t)\right) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}\left(P(x,t)\right) \tag{7.11}$$

where you can see it has an external force term and a diffusion coefficient term

Theorem 7.1.1 Digression: Kramers-Moyal expansion

Now let's go on a little digression. We got the kinetic equation above from the key properties of the Wiener process (keeping only the Δt terms) but what if don't even have a Wiener process in the first place? Let's try to approach it generally, which is called the Kramers-Moyal expansion. Let's once again start with the Evolution equation 2.46:

$$P(x,t+\Delta t) = \int G(x,t+at \mid y,t)P(y,t)dy \tag{7.12}$$

(We could also use the Kolmogorov-Chapman equation 2.47 if we have the initial conditions at x_0, t_0 but it doesn't matter because the analysis is the same) Substituting variables $\Delta x = x - y \quad , \quad y = x - \Delta x, \quad dy = -\Delta x$:

$$P(x,t+\Delta t) = \int G(x-\Delta x+\Delta x, t+\Delta t \mid x-\Delta x, t)P(x-\Delta x,t)d[-\Delta x] \tag{7.13}$$

Taylor expanding in $(-\Delta x)$:

$$P(x,t+\Delta t) = \int d(\Delta x)\sum_{n=0}^{\infty}\frac{(-\Delta x)^n}{n!}\frac{\partial^n}{\partial x^n}\left[G(x+\Delta x, t+\Delta t \mid x,t)P(x,t)\right] \tag{7.14}$$

(so far it's all the same as what we did above) Then we define the Kramers-Moyal coefficients:

Definition 7.1.3: Kramers-Moyal coefficients

$$D^{(n)}(x) = \lim_{\Delta t \to 0}\frac{1}{\Delta t}\int \frac{(\Delta x)^n}{n!}G(x+\Delta x, t+\Delta t \mid x,t)d(\Delta x) \tag{7.15}$$

This is very similar to the "Moments" that we have just seen but it's not quite the same for various reasons: There's an n factorial term and a division of Δt in the limit. If we now substitute these coefficients into the equation above and rearrange them a little:

$$\frac{\partial P(x,f)}{\partial t} = \sum_{n=1}^{\infty}(-1)^n\frac{\partial^n}{\partial x^n}\left[D^{(n)}_{(x)}P(x,t)\right] \tag{7.16}$$

we define this as the K-M operator:

$$\equiv \hat{L}_{kM}P(x,t) \tag{7.17}$$

Definition 7.1.4: K-M operator

$$\hat{L}_{kM} \equiv \sum_{n=1}^{\infty} (-1)^n \frac{\partial^n}{\partial x^n} D^{(n)}_{(x)} \tag{7.18}$$

As this analysis doesn't specify which stochastic process it is, the coefficients $D^{(n)}_{(x)}$ carry all the information we need to form the kinetic equation. For the Wiener process case, there are only two coefficients $D^{(1)}$ and $D^{(2)}$.

Example 7.1.1 (Brownian Motion)

Let's look at one example of kinetic analysis. Recall the Brownian motion equation:

$$\dot{v} = -\underbrace{\frac{\gamma}{m}}_{\theta} v + \underbrace{\frac{\sqrt{2k_BT}}{m}}_{\sigma} \xi(t) \tag{7.19}$$

And if we apply our analysis above, we obtain its kinetic equation:

$$\frac{\partial P(v,t)}{\partial t} = \frac{\partial}{\partial v}\left(\frac{\gamma}{m} vP\right) + \frac{2k_BT\gamma}{2m^2} \cdot \frac{\partial^2 P}{\partial v^2} \tag{7.20}$$

To solve for the steady state:

$$0 = \frac{\partial}{\partial v}\left[\frac{\gamma}{m} vP + \frac{k_bT\gamma}{m^2}\frac{\partial P}{\partial v}\right] \tag{7.21}$$

$$\Rightarrow \frac{dP}{P} = -\frac{m}{k_BT} v dv \tag{7.22}$$

$$P_{\text{eq}}(v) \propto \exp\left(-\frac{mv^2}{2k_BT}\right) \tag{7.23}$$

Which is the thermostat we're very familiar with! The next question is how long does it take for an arbitrary initial setup to reach this equilibrium state? Recall from the last lecture, we expect some sort of exponential decay with the time scale $\tau = m/\gamma$.

Example 7.1.2 (Quadratic potential)

Let's look at another example. Recall again the stochastic overdamped equation of motion for a particle moving in a quadratic potential:

$$\gamma \dot{x} = -a\,(x - x_0) + \sqrt{2k_BT\gamma}\xi(t) \tag{7.24}$$

which is just O-U with constants:

$$\theta = a/\gamma, \quad \sigma = \sqrt{2k\pi/\gamma} \tag{7.25}$$

applying the kinetic equation analysis:

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(\frac{a}{\gamma}(x - x_0)P\right) + \frac{k_BT}{\gamma} \cdot \frac{\partial^2 P}{\partial x^2} \tag{7.26}$$

solve for steady state:

$$\frac{dP}{P} = -\frac{a}{k_B T}(x - x_0)\,dx \tag{7.27}$$

$$P_{eq} = \exp\left(-\frac{a(x - x_0)^2}{2kT}\right) \tag{7.28}$$

Same as the reason above, we expect the system to reach equilibrium on the time scale of $\tau = \gamma/a$.

Example 7.1.3 (General Case of "Brownian Motion")

Now we go even one step further and look at the case where we are no longer in the overdamped limit and also allow the driving force to be a general force. We'll take $x = 0$ as our equilibrium position (You can put it back in if you want but I'm lazy lol):

$$m\dot{v} = -\gamma v - ax + \sqrt{2k_B T \cdot r}\,\xi(t) \tag{7.29}$$

since we have two unknowns, we need another equation and it is simply:

$$\dot{x} = v \tag{7.30}$$

recall that they can be combined into Eq. 6.22:

$$\dot{x}_i = -\theta_{ik} x_k + \sigma_{ik}\xi_k \tag{7.31}$$

with matrices:

$$\underline{\underline{\theta}} = \begin{pmatrix} \gamma/m & a/m \\ -1 & 0 \end{pmatrix}, \quad and \quad \underline{\underline{\sigma}} = \begin{pmatrix} \frac{\sqrt{2k_B T\gamma}}{m} & 0 \\ 0 & 0 \end{pmatrix} \tag{7.32}$$

Now let's try to check Fluctuation-Dissipation Relation. If given the 3 correlators: $\langle v^2 \rangle$ $\langle vx \rangle$ $\langle x^2 \rangle$. Can we write the full kinetic equation for $P(v, x, t)$?
Let's try to rewrite the kinetic equation of the O-U into a multi-variable form:

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x_i}(\theta_{ik} x_k \cdot P) + \frac{1}{2}\left(\sigma\sigma^\top\right)\frac{\partial^2 P}{\partial x_{ik}\partial x_k} \tag{7.33}$$

and subbing in the matrices σ and θ above and expand:

$$= \frac{\partial}{\partial v}\left(\frac{\gamma}{m} v \cdot P\right) + \frac{\partial}{\partial v}\left(\frac{a}{m} x \cdot P\right) + \frac{\partial}{\partial x}(-1 \cdot v \cdot p) \tag{7.34}$$

$$+ \frac{\partial}{\partial x}(0) + \frac{1}{2}\frac{2k_B T\gamma}{m^2}\frac{\partial^2 P}{\partial V^2} \tag{7.35}$$

taking the $\frac{\partial}{\partial x}(-1 \cdot v \cdot p)$ to the left:

$$\frac{\partial P}{\partial t} + \frac{\partial}{\partial x}(vP) = \frac{\partial}{\partial v}\left(\frac{\gamma v + ax}{m} \cdot P\right) + \frac{k_B T\gamma}{m^2}\frac{\partial^2 P}{\partial v^2} \tag{7.36}$$

We can see that the terms on the LHS are a convective derivative and the $\gamma v + ax$ term is a total force between the drag and linear restoring force. This is the general Fokker-Planck equation for $P(v, x, t)$

Definition 7.1.5: Fokker-Planck equation

$$\frac{\partial P}{\partial t} + \frac{\partial}{\partial x}(vP) = \frac{\partial}{\partial v}\left(\frac{\gamma v + ax}{m} \cdot P\right) + \frac{k_B T\gamma}{m^2}\frac{\partial^2 P}{\partial v^2} \tag{7.37}$$

(This is a very very hard equation to solve.)

It is interesting to point out that the time scale for v to reach equilibrium is much smaller than it takes for x. Therefore, the velocity-wise Maxwell distribution develops wayyy before the position-wise Boltzmann distribution does. So this equation actually reduces to the diffusion equation for $P(x,t)$ at t much greater than the velocity relaxation time $t \gg m/\gamma$.

Chapter 8

Lecture 8: O-U process continued

8.1 Discussion before continuing

In the last chapter, we talked about the O-U process where the forces are linear. What happens if the force field is arbitrary: $f(x) = -\frac{\partial}{\partial x}V(x)$?

In the overdamped limit we have the SDE:

$$\gamma \dot{x} = f(x) + \sqrt{2k_B T \cdot \gamma} \cdot \xi(t) \tag{8.1}$$

and via the Kramers-Moyal expansion in D^1 and D^2 we get:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{f(x)P}{\gamma}\right) + \frac{1}{2}\frac{2kT}{\gamma}\frac{\partial^2 P}{\partial x^2} \tag{8.2}$$

which can be written in terms of the current/flux of P:

$$= -\frac{\partial}{\partial x}(J(x,t)) \tag{8.3}$$

with the continuity equation:

$$\dot{P} = -\operatorname{div} J \tag{8.4}$$

By expanding J, here we have:

$$J = \frac{f(x)}{\gamma} \cdot P(x,t) - \frac{k_B T}{\gamma}\frac{\partial P(x,t)}{\partial x} \tag{8.5}$$

There's also another definition of J:

$$J = -De^{-\beta V(x)}\frac{\partial}{\partial x}\left[e^{\beta V(x)} \cdot P(x,t)\right] \tag{8.6}$$

which is the same if you expand it with the product rule and treat $D = \frac{k_B T}{\gamma}$

After adding in the force, let's make it even more general. What if we make the diffusion term ξ x-dependent?

We'll have an SDE where σ is x-dependent:

$$dx = \mu(x,t)dt + \sigma(x,t)dw \tag{8.7}$$

If we take an educated guess of the solution, we will probably want to write something like Eq8.2:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}(\mu(x,t)P(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma(x,t)\cdot P(x,t)] \tag{8.8}$$

But a couple of questions arise regarding the last term:

Why can't σ be completely outside of the derivatives?:

$$\frac{1}{2}\sigma^2(x,t)\frac{\partial^2}{\partial x^2}P \tag{8.9}$$

How about partially inside?

$$\frac{1}{2}\sigma\frac{\partial}{\partial x}\left(\sigma\frac{\partial p}{\partial x}\right) \tag{8.10}$$

which is called the "multiplicative noise" but we're not dealing with it today.

8.2 O-U process key examples

8.2.1 Particle in linear potential

The most common case is gravity, where the particle experience a constant force $f = mg$.

Let's start with a linear potential:

$$V(x) = fx \tag{8.11}$$

where we can get the kinetic equation for our probability P:

$$\dot{P} = \frac{f}{\gamma}P' + \frac{kT}{\gamma}P'' \tag{8.12}$$

Which can be solved to get: (remember $D = \frac{kT}{\gamma}$)

$$P(x,t) = \frac{1}{\sqrt{4\pi D\,(t-t_0)}}\cdot\exp\left[-\frac{\left(x - x_0 + \frac{f_0}{\gamma}\,(t-t_0)\right)^2}{4D\,(t-t_0)}\right] \tag{8.13}$$

8.2.2 Particle in a quadratic potential

Similar to the linear potential case, we start with the kinetic equation by using the potential $V = \frac{1}{2}\alpha x^2$ and $f = \alpha x$:

$$\dot{P} = \frac{\alpha}{\gamma}(x\cdot P)' + \frac{k_s T}{\gamma}P'' \tag{8.14}$$

where we can solve to get:

$$P\,(x,t) = \frac{1}{\sqrt{2\pi k_B T S(t)/\alpha}}\cdot\exp\left[-\frac{\alpha\left(x - x_0\exp\left(-\frac{\alpha}{\gamma}\,(t-t_0)\right)\right)}{2k_B T S(t)}\right] \tag{8.15}$$

where:

$$S(t) = 1 - e^{-\frac{2\alpha}{\gamma}(t-t_0)} \tag{8.16}$$

Note: In both examples, we solve for the solutions for the initial condition $P(x,t) = \delta\,(x-x_0)|_{t=t_0}$. And in the $t \to \infty$ limit, P follows the Boltzmann distribution.

8.3 Convection diffusion

(Sometimes it's called advection-diffusion.) If we have an arbitrary background flow $u(x)$ of fluid on top of the thermal Brownian motion of the particle:

The original flux equation:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x} J(x,t) \tag{8.17}$$

$$J = -De^{-\beta v(x)} \frac{\gamma}{\partial x} \left[e^{\beta v(x)} P(x,t) \right] \tag{8.18}$$

after adding on an extra convection:

$$J_{\text{extra}} = u \cdot P\,(x,t) \tag{8.19}$$

$$\frac{\partial P(x,t)}{\partial t} = \nabla \left[De^{-\beta V} \nabla \left(e^{\beta V} \cdot P \right) \right] - \nabla(u \cdot P) \tag{8.20}$$

rearranging:

$$\frac{\partial P(x,t)}{\partial t} + \frac{\partial}{\partial x}(u(x) \cdot P) = \nabla \left[De^{-\beta V} \nabla \left(e^{\beta V} \cdot P \right) \right] \tag{8.21}$$

$$= -\frac{\partial}{\partial x} \left(\frac{f(x)}{\gamma} p \right) + D \frac{\partial^2 p}{\partial x^2} \tag{8.22}$$

Now if $\nabla \dot{u} = 0$, the term above becomes a proper convective derivative:

$$\frac{\partial}{\partial x}(u \cdot P) = (u \cdot \nabla) P\,(x,t) \tag{8.23}$$

And since u adds extra forces onto the force field, it's like adding extra structures to the potential V:

$$\frac{\partial P(x,t)}{\partial t} = -\nabla \left[\left(\frac{f(x)}{\gamma} + u(x) \right) P \right] + P \frac{\partial^2 P}{\partial x^2} \tag{8.24}$$

where the term in the bracket is the new effective potential V_{eff}.

Now let's try to solve problems with the most simple constant convection-diffusion.

We have a constant flow to the right:

$$u = \text{constant} \tag{8.25}$$

$$\frac{\partial P}{\partial t} + u \cdot \frac{\partial P}{\partial x} = D \frac{\partial^2 P}{\partial x^2} \tag{8.26}$$

If the initial concentration of the particles are (concentration is interchangeable with probability in this case):

$$C = C_0 \delta(x) \tag{8.27}$$

We have the trivial solution of a Gaussian diffusion drifting with velocity u:

$$C(x,t) = \frac{1}{\sqrt{4\pi D t}} e^{-\frac{(x-ut)^2}{4Dt}} \tag{8.28}$$

Now let's consider the case where the initial condition is no longer a delta function but a constant source of concentration $c = c_0$ at $x = 0$ at all times. In the large t limit, it's not hard to be convinced that the solution will look like a shock wave propagating to the right with a near-vertical shock front. The analytical solution looks like this:

$$C(x,t) = \frac{1}{2} C_0 \left[\text{Erfc} \left(\frac{x - ut}{\sqrt{4Dt}} \right) + e^{\frac{ux}{D}} \text{Erfc} \left(\frac{x - ut}{\sqrt{4Dt}} \right) \right] \tag{8.29}$$

The exponential term $\exp(ux/D)$ is called the dispersivity. (There's a very long question regarding this in the example sheet but the lecturer said that there will never be questions that long and that annoying in the actual exam.)

Chapter 9

Lecture 9: Mean first passage time (MFPT)

The idea is simple, imagine having a particle travelling at a high speed, bouncing around in a small spherical container. There is a small opening in the container with which the particle can escape. What is the probability of the particle escaping in time t? And what is its meantime of escape? These questions can be asked in all kinds of setting for all types of stochastic processes.

Let's bring back the concept of survival probability $S(t)$, which is the probability that the particle has not escaped at time t. If the probability of the position of the particle is $P(x,t)$, then S would be:

$$S(t) = \int_{\text{vol. inside}} P(x,t)dx \tag{9.1}$$

Let's try to derive S. Consider $S(t) - S(t+dt)$ to be the rate of escape:

$$S(t) - S(t+dt) = f(t)dt \tag{9.2}$$

Which we can see that f is the probability density of escaping in time t:

$$f(t) = -\frac{\partial S(t)}{\partial t} \tag{9.3}$$

$$\langle t \rangle = \int_0^\infty t f(t)dt \tag{9.4}$$

and if intergrate by parts:

$$\langle t \rangle = \int S(t)dt \tag{9.5}$$

9.1 Examples: 1D simple diffusion

Consider the simple 1D diffusion where the particle starts from $x = 0$ and diffuses through Brownian motion:

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}} \tag{9.6}$$

And imagining adding a little door at $x = L$ where as soon as the particle touches it, it escapes and no longer exists in the system. We call this the absorbing boundary.

$$P(x = L, t) = 0 \tag{9.7}$$

To achieve this, we try using the method of images (just like in electric fields!) by introducing another identical but negative diffusing probability distribution at $x = 2L$. Overall, we'll have:

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}}\left(e^{-\frac{x^2}{4Dt}} - e^{-\frac{(x-2L)^2}{4Dt}}\right) \tag{9.8}$$

Now S becomes:

$$S(t) = \int_{-\infty}^{L} P\left(x_c t\right) dx = \text{Erf}\left[\frac{L}{\sqrt{4Dt}}\right] \tag{9.9}$$

And we can find $f(t)$:

$$f(t) = -\frac{\partial S}{\partial t} = \frac{L}{\sqrt{4\pi Dt^3}} e^{-\frac{L^2}{4Dt}} \tag{9.10}$$

However... This gives us a diverging MFPT!

$$\text{MFPT } = \int_0^{\infty} t f(t) dt = \int \frac{dt}{\sqrt{t}} \to \infty \tag{9.11}$$

The problem is that we're modelling this with two 1D diffusions where particles can freely move to the left and right freely without actually being "absorbed". The probability of $P(x = L) = 0$ is just a delusion! To fix this, we add a reflective wall at $x = -a$ to confine the particle. Now we have two reflections: the reflective wall creates symmetric images, and the absorption boundary creates antisymmetric images. And since images of images will also be created by these reflections, we now have an infinite set of images! Of course that we can just express them as infinite series but the simpler way is just to solve the diffusion equation with the boundary conditions:

$$P' = 0 \mid x = -a \tag{9.12}$$

$$P = 0 \mid x = L \tag{9.13}$$

$$P = \delta(x) \text{ at } t = 0 \tag{9.14}$$

where the solution will be the infinite series of the form (C is a constant):

$$P = \sum_n^{\theta} \sin\left(\frac{n\pi x - C}{a + L}\right) e^{-(nDt)} \tag{9.15}$$

Once we have obtained $P(x,t)$ we can quickly obtain S:

$$S(t) = \int_{-a}^{L} P(x,t) dx \tag{9.16}$$

And the MFPT will be:

$$\text{MFPT } = \frac{L(L+2a)}{2D} \tag{9.17}$$

So in this case where the diffusion starts from x=0: if $a \to 0$ (only diffusion to the right is allowed at the origin), we can easily obtain MFPT $= L^2/2D$. But if we start at some $x = x_0$ which is not the origin, MFPT will depend on x_0 as well! (Imagine if we start very very close to x=L, we will expect a much shorter MFTP.)

9.2 MFTP of the Wiener Process

For a motion in an arbitrary potential $V(x)$, we have the Fokker-Planck equation:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{f(x)}{\gamma} P(x,t)\right) + D\frac{\partial^2 P}{\partial x^2} \tag{9.18}$$

$$= -\frac{\partial}{\partial x}(\mu(x) \cdot P) + \frac{1}{2}\sigma^2 \frac{\partial^2 P}{\partial x^2} \tag{9.19}$$

We define an F-P operator to simplify things:

$$\hat{L}_x = D\frac{\partial}{\partial x}e^{-\beta V(x)}\left(\frac{\partial}{\partial x}e^{\beta V(x)}\right) \tag{9.20}$$

$$\frac{\partial P}{\partial t} = -\hat{L}_x P \tag{9.21}$$

and the solution can be written in the form:

$$P(x,t) = e^{-\hat{L}_x \cdot t}P(x,0) \tag{9.22}$$

If we now want to calculate MFPT:

$$MFPT = \int_0^\infty S(t)dt \tag{9.23}$$

$$= \int_0^\infty dt \int_{\text{vol}} dx P\left(x,t \mid x_0, t=0\right) \tag{9.24}$$

we can see that this value will only depend on x_0

$$= \tau(x_0) \tag{9.25}$$

From this, could we find an equation that would determine τ? Sneak Peak: For this, we need operators that act on x_0 (initial condition), which is an "adjoint F-P operator". We need the Kolmogorov-Chapman relationship:

$$P\left(x,t \mid x_0,t_0\right) = \int G\left(x,t \mid y,t'\right) P\left(y,t' \mid x_0 t_0\right) dy \tag{9.26}$$

The rest will be in the next lecture.

Chapter 10

Lecture 10: MFPT continued

At the end of the last chapter, we mentioned the Kolmogorov-Chapman equation:

$$P\left(x,t \mid x_0,t_0\right) = \int G\left(x,t \mid y,t'\right) P\left(y,t' \mid x_0 t_0\right) dy \tag{10.1}$$

And since: (K-M Operator)

$$\frac{\partial P}{\partial t} = -\hat{L}_x p \tag{10.2}$$

We can get this relation since the only entry of t and x is in G:

$$\frac{\partial G}{\partial t} = -\hat{L}_x G \tag{10.3}$$

But what if we differentiate w.r.t. t'? We get:

$$\frac{\partial}{\partial t'} P\left(x,t \mid x_{0,0}\right) = 0 \tag{10.4}$$

$$= \int P\left(y,t' \mid x_0,0\right) \frac{\partial}{\partial t'} G\left(x,t \mid y,+'\right) dy + \int G\left(x,t \mid y,t'\right) \frac{\partial}{\partial t'} P\left(y,t' \mid x_0,0\right) dy \tag{10.5}$$

where with the K-C equation, the time derivative of P in the second integral can be written in:

$$\frac{\partial}{\partial t'} P\left(y,t' \mid x_0,0\right) = -\hat{L}_y P\left(y,t' \mid x_0,0\right) \tag{10.6}$$

Let's now define what an adjoint operator is:

if:

$$\dot{P} = -\hat{L}_x P \tag{10.7}$$

where:

$$\hat{L}_x = \partial_x \mu(x) - D\partial_x^2 \tag{10.8}$$

then the adjoint operator $L^\dagger$ is defined to satisfy:

$$\int dx g(x) \hat{L}_x f(x) = \int dx f(x) L_x^\dagger + g(x) \tag{10.9}$$

$$\langle g \hat{L} f \rangle = \left\langle f \hat{L}^\dagger g \right\rangle \tag{10.10}$$

And we can find it by plugging in the definition of L_x and integrating it by parts:

$$\int dx g(x) \frac{\partial}{\partial x}(\mu(x) f(x)) = \cancel{[g\mu f]}^{0} - \int dx f(x) \mu(x) \frac{\partial g(x)}{\partial x} \tag{10.11}$$

Therefore find the adjoint:

$$L_x^\dagger = \mu(x) \frac{\partial}{\partial x} - D \frac{\partial^2}{\partial x^2} \tag{10.12}$$

So if:

$$\dot{P} = -\hat{L}_x P \tag{10.13}$$

The adjoint operator to be:

$$\hat{L}_x^\dagger = D e^{\beta V(x)} \frac{\partial}{\partial x} \left(e^{-\beta V(x)} \frac{\partial}{\partial x} \right) \tag{10.14}$$

Now if we go back to:

$$\frac{\partial}{\partial t'} P(x, t \mid x_0, 0) = 0 \tag{10.15}$$

$$= \int P(y, t' \mid x_0, 0) \frac{\partial}{\partial t'} G(x, t \mid y, t') \, dy \tag{10.16}$$

$$- \int G(x, t \mid y, t') \hat{L}_y P(y, t' \mid x_0, 0) \, dy \tag{10.17}$$

using the definition of the adjoint operator:

$$\int G(x, t \mid y, t') \hat{L}_y P(y, t' \mid x_0, 0) \, dy = \int P(y, t' \mid x_0, 0) \hat{L}_y^\dagger G(x, t \mid y, t') \, dy \tag{10.18}$$

we can write it like:

$$0 = \int \left[\frac{\partial}{\partial t'} G(x, t \mid y, t') - \hat{L}_y^\dagger G(x, t \mid y, t') \right] P(y, t' \mid x_0, 0) \, dy \tag{10.19}$$

and simplify it by letting the square bracket equal 0:

$$\frac{\partial}{\partial t'} G = L_y^\dagger G \tag{10.20}$$

And since we can obtain P by applying this propagator onto the initial conditions, P will obey the same equation:

$$\frac{\partial}{\partial t_0} P(x, t \mid x_0, t_0) = \hat{L}_{x_0}^\dagger P(x, t \mid x_0, t_0) \tag{10.21}$$

So with the adjoint operator, we can go back to the definition of $\tau(x_0)$:

$$\hat{L}_{x_0}^\dagger \left[\tau(x_0) = \int dt \int_{\text{vol.}} P(x, t \mid x_0, t_0) \, dx \right] \tag{10.22}$$

$$\tag{10.23}$$

$$\hat{L}_{x_0}^\dagger \tau(x_0) = \int dt \int dx \left(\hat{L}_{x_0}^\dagger p \right) \tag{10.24}$$

$$= \int dt \int dx \frac{\partial}{\partial t_0} P(x, t \mid x_0 t_0) \tag{10.25}$$

using the fact that we know P is a function of $\tilde{t} = t - t_0$, we rewrite it:

$$= \int d\tilde{t} \int dx \left[\frac{\partial}{\partial \tilde{t}} P\left(x, \tilde{t} \mid x_0, t_0\right) \right] \tag{10.26}$$

$$= \int d\tilde{t} f(\tilde{t}) = -1 \tag{10.27}$$

we have simplified everything into:

$$\hat{L}_{x_0} \tau\left(x_0\right) = -1 \tag{10.28}$$

or :

$$D e^{\beta v(x)} \frac{\partial}{\partial x} \left(e^{-\beta v(x)} \frac{\partial \tau(x)}{\partial x} \right) = -1 \tag{10.29}$$

With this, we can solve for τ:

$$e^{-\beta V(x)} \frac{\partial \tau(x)}{\partial x} = -\frac{1}{D} \int^{x} e^{-\beta V(y)} dy \tag{10.30}$$

Note: the lower limit is to be determined. Either we have a reflective boundary and use that as the lower limit, or use $-\infty$ if $V(y)$ is well behaved. And then perform a second integration:

$$\tau\left(x_0\right) = -\int_{D}^{x_0} \frac{1}{D} e^{\beta V(x)} dx \int^{x} e^{-\beta V(y)} dy \tag{10.31}$$

Note: Same with the previous case, this lower limit is to be determined as well. We can see that this lower bound should be the absorbing boundary.

In the end we achieve our final expression of the MFPS:

$$\tau\left(x_0\right) = \frac{1}{D} \int_{x_0}^{L} e^{\beta V(x)} dx \int_{-a}^{x} e^{-\beta V(y)} dy \tag{10.32}$$

where L is the location of the absorbing boundary and $-a$ is the location of the reflecting boundary or $-\infty$.

10.1 Examples

10.1.1 1D diffusion

Let's go look at our original example where the initial condition is a delta function at $x = x_0$ and an absorbing and reflective boundary at $x = L$ and $x = -a$.

$$\tau\left(x_0\right) = \frac{1}{D} \int_{x_0}^{L} dx \int_{-a}^{x} dy = \frac{1}{D} \int_{x_0}^{L} dx(x+a) = \frac{1}{D} \left(\frac{1}{2} L + aL \right) = \frac{1}{2D} \left[L(L+2a) - x_0(x_0+2a) \right] \tag{10.33}$$

10.1.2 Kramers problem:

A potential $V = \frac{1}{2}\omega^2 x^2$ with abosrobing boundary at $x = L$. Please look at the lecture notes for the detailed shape of the potential V.

$$MFPT = \frac{1}{D} \int_{x_0}^{L} dx e^{\beta V(x)} \int_{-\infty}^{x} e^{-\beta V(y)} dy \tag{10.34}$$

$$\approx \frac{1}{D} \cdot e^{\beta V_{\max}} \sqrt{\frac{2\pi k T}{\omega_B^2}} \sqrt{\frac{2\pi \cdot L\pi}{\omega_0^2}} \tag{10.35}$$

$$MFPT \approx \frac{2\pi\gamma}{\omega_0 \omega_\beta} e^{\beta V_{\max}} \tag{10.36}$$

And the escape rate will be half of the reciprocal (a half chance of being absorbed after reaching L, the other half steps to the left)

$$\text{rate} = \frac{\omega_0 \omega_\beta}{4\pi\gamma} e^{-\beta V_{\max}} \tag{10.37}$$

10.1.3 Diffusion in arbitrary periodic potential

A potential $V(x)$ that's periodic over space with the period of L. We want to find the MFPT of a particle going from 0 to L (travelling across a whole period). Without losing generality, let's set the local minimum to be at x and L (and 2L and so on).

start with the standard eq:

$$\tau = \frac{1}{D} \int_{x_0}^{L} dx e^{\beta V(x)} \int_{-a}^{x} e^{-\beta V(y)} dy \tag{10.38}$$

x_0 is obviously 0 and $-a = 0$ as well (I'm not sure why but the lecturer said so). In the case of over a long time and a long distance travelled the total deviation from the original position can be written into an effective diffusion:

$$L^2 = 2D_{\text{eff}} \cdot \tau \tag{10.39}$$

so:

$$D_{\text{eff}} = \frac{L^2 D}{2} \frac{1}{\int_0^L e^{\beta v(x)} dx \int_0^x e^{-\beta V(y)} dy} \tag{10.40}$$

and since the denominator is always quite a large value, we know that the effective diffusion will be:

$$D_{\text{eff}} \ll D \tag{10.41}$$

Chapter 11

Lecture 11: Multiplicative noise

11.1 MFPT final part

We've studied the MFPT in the last 2 lectures and we have some conclusions:

$$\langle t \rangle = \int_0^\infty S(t)dt \tag{11.1}$$

$$\langle t\,(x_0)\rangle = \frac{1}{D}\int_{x_0}^{L} dx e^{\beta V(x)} \int^{x} e^{-\beta V(y)} dy \tag{11.2}$$

These conclusions were strictly in 1D but in higher dimensions, the particle might miss the target altogether and never be absorbed. In this case, the MFPT must depend on the size of the target since it might be unbounded if the target is too small! **Examples:** 1. Forming a ring: The MFPT of one end of a long polymer chain makes contact with its other end. 2. Narrow escape: The very first example that we gave of a shell with a small opening and a particle trapped inside.

Anyway the relationship between the size of the target ϵ and the MFPT is:

$$1D : \propto \frac{L^2}{D} \tag{11.3}$$

$$2D : \propto \frac{L^2}{D}\ln(\frac{1}{\epsilon}) \tag{11.4}$$

$$1D : \propto \frac{L^3}{D\epsilon} \tag{11.5}$$

11.2 Multiplicative noise

We have mentioned this in a previous lecture but the overall idea is that when we solve the SDE:

$$dx = \mu dt + \sigma dW \tag{11.6}$$

where in the overdamped limit for Brownian motion we have:

$$\sigma = \sqrt{\frac{2k_B T}{\gamma}} \tag{11.7}$$

But when σ is a function of x, this adds a degree of freedom in the last term of F-P.

The usual F-P equation is:

$$\frac{\partial P(x,t)}{\partial f} = -\frac{\partial}{\partial x}(\mu(x,t)\cdot P) + \frac{\partial^2}{\partial x^2}\left(\frac{\sigma^2(x)}{2}P\right) \tag{11.8}$$

But how do we know the last term is not in the forms of:

$$\frac{\sigma^2(x)}{2}\frac{\partial^2 P}{\partial x^2} \tag{11.9}$$

$$\frac{1}{2}\frac{\partial}{\partial x}\left[\sigma(x)\frac{\partial}{\partial x}(\sigma(x)p)\right] \tag{11.10}$$

We need to re-examine the K-M expansion. The evaluation of $\langle \Delta x \rangle$ is as followed:

$$\langle \Delta x \rangle = \int_t^{t+\Delta t} ds(\mu(x,s) + \sigma(x)\xi(s)) \tag{11.11}$$

$$= \mu(x,t)\Delta t + \sigma(x)\int_t^{t+\Delta t} \xi(s)ds \tag{11.12}$$

The whole problem stems from the question that at which point of time between t and $t + \Delta t$ do we take the value x to determine σ?

11.2.1 Itoh version

Itoh's lemma suggests that we evaluate σ at the start of the time interval (at time t):

$$\sigma(x[t])\int_t^{t+\Delta t} \xi(s)ds \tag{11.13}$$

since x(t) is statistically independent from $\xi(s)$, this makes the most sense mathematically. And this is the only version that mathematicians believes.

11.2.2 Stratonovich version

But Stratonovich suggests that maybe we should evaluate σ in the middle of this small time interval instead:

$$\sigma(x[t + \frac{1}{2}\Delta t])\int_t^{t+\Delta t} \xi(s)ds \tag{11.14}$$

11.2.3 Lau, Lubensky version

From the two examples above, we see that recently people (physicists) have been trying to sort out this problem since they've encountered problems in using the Itoh version. They take the approach by taking an arbitrary point in the interval for evaluating σ:

$$\sigma(x[t + \alpha\Delta t]) \tag{11.15}$$

where α is between 0 to 1. And then decide what value is the "best" experimentally. Lau and Lubensky concluded that $\alpha = 1$ is the best value.

For all the 3 versions that we've mentioned above, the corresponding F-P equations are:

Itoh:($\alpha = 0$)

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(\mu(x,t)P) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x)\cdot P\right) \tag{11.16}$$

Stratonovich:($\alpha = 0.5$)

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(\mu(x,t)P) + \frac{1}{2}\frac{\partial}{\partial x}\left(\sigma(x)\frac{\partial}{\partial x}(\sigma(x)P)\right) \tag{11.17}$$

Lau-Lubensky:($\alpha = 1$)

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(\mu(x,t)P) + \frac{1}{2}\frac{\partial}{\partial x}\left(\sigma^2(x)\frac{\partial p}{\partial x}\right) \tag{11.18}$$

How do we test which one is correct experimentally? If we have a long diffusion channel where there are no potentials and a linear temperature distribution from T_1 on one end to T_2 on the other end. The initial probability P (concentration) is constant across the whole channel. Physicists have observed something called the Thermophoresis or Soret effect:

$$J = -D\nabla P + \alpha P(1-p)\nabla T \tag{11.19}$$

at low concentrations, this can be approximated into:

$$J \approx -D\nabla P + \alpha P\nabla T \qquad = T\frac{\partial p}{\partial x} + p\frac{\partial T}{\partial x} \tag{11.20}$$

From the 3 versions of solutions that we mentioned above, these are their predictions for J:

Itoh:($\alpha = 0$)

$$J = T\nabla P + P\nabla T \tag{11.21}$$

Stratonovich:($\alpha = 0.5$)

$$J = T\nabla P + P\nabla T \tag{11.22}$$

Lau-Lubensky:($\alpha = 1$)

$$J = T\nabla P(\text{only???}) \tag{11.23}$$

We see that L-L is wrong here but the issue is with the "physics" behind the SDE:

$$dx = \mu\,(x,t)\,dt + G(x)dW \tag{11.24}$$

The "Wiener" process dW is always an **approximation** of the Brownian motion. The wiener process is built on a Markov assumption that the process is memoryless but do we actually know that? Are we sure that the time intervals between two impulses are actually much greater than the relaxation time for the system to "not remember" what happened? We will dive further into this issue in the next later but for the moment we have not gotten enough evidence to disprove L-L.

Chapter 12

Lecture 12: Last lecture

12.1 Multiplicative Noise continued

From the discussion in the last lecture, we talked about the multiplicative noise which stems from the σ term in this integral.

$$\sigma(x[t]) \cdot \int_t^{t+\Delta t} \xi(s) ds \tag{12.1}$$

If we're talking about a mathematically "correct" statistical process, in which $\sigma(x)$ is statistically independent of this wiener step (the integral), then Itoh calculus follows and we should have no debate about this "multiplicative problem". Which gives us an F-P equation of:

$$\frac{\partial P(x,t)}{\partial f} = -\frac{\partial}{\partial x}(\mu(x,t) \cdot P) + \frac{\partial^2}{\partial x^2}\left(\frac{\sigma^2(x)}{2}P\right) \tag{12.2}$$

But if the underlying stochastic process $W(t)$ is a sequence of deterministic events averaged over Δt, then the Wong-Zakai theorem applies and we must add an extra dt term and rewrite our SDE: (We don't have time to derive this so sadly we'll have to take it as granted)

$$dx = \left(\mu(x,t) + \frac{1}{2}\sigma(x)\frac{\gamma\sigma}{2x}\right)dt + \sigma(x)dW \tag{12.3}$$

And use this SDE to rederive the F-P equation, using mathematically rigorous Itoh assumptions about statistical independence:

Notice that the second term on the RHS is the new correction term:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\mu(x,t)P] - \frac{\partial}{\partial x}\left[\frac{1}{2}\sigma\frac{\partial\sigma}{\partial x} \cdot P\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2 P\right] \tag{12.4}$$

Rearranging it:

$$= -\frac{\partial}{\partial x}(\mu(x,t)P) + \frac{1}{2}\frac{\partial}{\partial x}\sigma(x)\left(\frac{\partial}{\partial x}[\epsilon(x)P)\right) \tag{12.5}$$

Which is exactly the Stratonovich form! This is very surprising since Stratonovich's result was proposed 50 years ago with a very simple mid-point assumption. Now with this, we can rewrite our old probability flux J:

$$J = -De^{-\beta V(x)}\frac{\partial}{\partial x}\left(e^{\beta V(x)}P(x,t)\right) \tag{12.6}$$

where we have a constant σ:

$$\sigma = \sqrt{2D} \tag{12.7}$$

into

$$J = -\frac{1}{2}\sigma(x)e^{-\beta V(x)} \cdot \frac{\partial}{\partial x}\left(\sigma(x)e^{\beta V(x)} \cdot P(x,t)\right) \tag{12.8}$$

by using the definition of J:

$$\dot{P} = -\nabla J \tag{12.9}$$

And as long as we identify the "friction constant" as an x-dependent variable:

$$\gamma(x) = \frac{2k_B T}{\sigma(x)^2} \tag{12.10}$$

we can still write the new diffusion constant as:

$$D(x) = \frac{\sigma^2(x)}{2} = \frac{k_B T}{\gamma(x)} \tag{12.11}$$

12.2 Non-Markov processes

Could we re-examine the Markov assumption that we have been using throughout the whole course?

Example: Viscoelasticity (Retarded linear response)

The idea of viscoelasticity is simply a non-memoryless stress-strain relationship (which we all learned in A-level), therefore the relationship will look something like this:

$$\text{Stress }(t) = \int_{-\infty}^{t} G\left(t-t'\right) \cdot \frac{\partial \,\text{Strain}\,(\text{t}')}{\partial t'} dt' \tag{12.12}$$

For a general retarded linear response:

If F and x are conjugate thermodynamic variables: (Think about internal energy in thermodynamics)

$$dU = -Fdx \tag{12.13}$$

We can write:

$$F(t) = -\int_{-\infty}^{t} g\left(t-t'\right)\dot{x}\left(t'\right)dt' \tag{12.14}$$

Or:

$$x(t) = \int_{-\infty}^{t} \alpha\left(t-t'\right)\dot{F}\left(t'\right)dt' \tag{12.15}$$

and we call g and α the memory kernels.

Now let's look at the two extreme forms of g.

If $g(t-t')$ is a constant:

$$g\left(t-t'\right) = \text{ const.}(k) \tag{12.16}$$

We get a memoryless instant response:

$$F(t) = -kx \tag{12.17}$$

And if $g(t-t')$ is a delta function:

$$g\left(t-t'\right)=\gamma\delta\left(t-t'\right) \tag{12.18}$$

We get an always dissipative drag:

$$F=-\gamma\dot{x} \tag{12.19}$$

and usually a generic retarded response function g is a decaying function that's somewhere in between.

So in comparison to a generic Brownian motion with Markov noise

where we have the equation of motion like this:

$$m\dot{v}=-\gamma v-kx+[\text{Noise}]\quad\left(\sigma\xi(t),\ \sigma\cdot\frac{dW}{dt},\ \text{etc.}\right) \tag{12.20}$$

Now we have:

$$m\dot{v}=-\int_{-\infty}^{t}g\left(t-t'\right)\dot{x}(t)dt'+[\text{ Noise }] \tag{12.21}$$

where we can no longer say the noise is Markov. Let's quickly look at the correlation function of this noise$\eta(t)$ without deriving it:

$$\langle\eta(t)\eta\left(t'\right)\rangle=k_BTg\left(t-t'\right) \tag{12.22}$$

which is the Fluctuation-Dissipation relation.

Example: Brownian motion with inertia

$$\dot{x}=v \tag{12.23}$$

$$\dot{V}=-\frac{\gamma}{m}v-\frac{k}{m}x+\frac{\sqrt{2k_BT\gamma}}{m}\xi(t) \tag{12.24}$$

Without even mentioning any memory kernel, there is already a problem with our Markov assumption that we have not noticed before:

First integrate to find v:

$$v(t)=\int_{-\infty}^{t}e^{-\frac{\gamma}{m}(t-t')}\cdot\left[-\frac{k}{m}x\left(t'\right)+\frac{\sqrt{2k_T\cdot\gamma}}{m}\xi\left(t'\right)\right]dt' \tag{12.25}$$

And sub it into x:

$$\dot{x}=-\frac{k}{m}\int_{-\infty}^{t}e^{-\frac{\gamma}{m}(t-t')}x\left(t'\right)dt'+\frac{\gamma}{m}\sqrt{\frac{2k_BT}{\gamma}}\cdot\int_{-\infty}^{t}e^{-\frac{\gamma}{m}(t-t')}\xi(t')dt' \tag{12.26}$$

We can see that in both integrals, there is already a "memory-like" exponential term due to the particle's own inertia with a memory time scale $\tau=m/\gamma$.

By using this definition of the delta function:

$$\delta(t)=\lim_{t\to 0}\frac{1}{\tau}e^{-t/\tau} \tag{12.27}$$

We can fix the Markov assumption by going to the overdamped limit ($t\ll\tau$):

$$\dot{x}=-\frac{k}{\gamma}x(t)+\sigma\xi(t) \tag{12.28}$$

Last Exercise:

By using $v(t),x(t)$ at time $\leqslant m/\gamma$, derive:

$$\langle\eta(t)\eta(s)\rangle=\ldots\frac{k_BT}{k}\left(\tau^2e^{-\frac{t-s}{\tau}}\right) \tag{12.29}$$