Advanced Statistical Mechanics
Part III Physics, Unversity of Cambridge

Lawrence Wu

February 2, 2023

Contents

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:

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

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

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

$$\tag{1.5}$$

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 metastabilities 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

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

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

We then introduce the concept of characteristic function $\phi_x(k)$ 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 n^{th} 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)

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

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

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

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

$$y_N = \sum_{i=1}^{N} x_i \tag{2.15}$$

$$\langle y_N \rangle = N \langle x \rangle \tag{2.16}$$

$$\sigma_y^2 = \left\langle y_N^2 \right\rangle - \langle y_N \rangle^2 = N\sigma^2 \tag{2.17}$$

where the variance relationship came from:

$$\left\langle y^2 \right\rangle - \langle y \rangle^2 = \sum_i \sum_j \left\langle x_i x_j \right\rangle - \sum_i \langle x_i \rangle \sum_j \left\langle x_j \right\rangle \tag{2.18}$$

$$= N \left\langle x^2 \right\rangle + N(N-1)\langle x \rangle^2 - N^2 \langle x \rangle^2 \tag{2.19}$$

$$= N \left(\left\langle x^2 \right\rangle - \langle x \rangle^2 \right) = N\sigma^2 \tag{2.20}$$

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:

$$S_N = \frac{y_N}{N} \tag{2.21}$$

$$\langle S_N \rangle = \langle x \rangle \tag{2.22}$$

$$\sigma_S^2 = \frac{\sigma_x^2}{N} \tag{2.23}$$

$$\tag{2.24}$$

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 $\phi_S(k)$:

$$\phi_s(k) = \left\langle e^{iks} \right\rangle = \left\langle exp\left(i\frac{k}{N} \sum_m x_m \right) \right\rangle \tag{2.25}$$

$$= \left\langle e^{\frac{ik}{N}x} \right\rangle_{P(x)}^N = \left[\phi_x(\frac{k}{N}) \right]^N \tag{2.26}$$

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

$$\phi_s(k) = \exp\left[\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} C_m(s) \right] \tag{2.27}$$

$$= \left[\phi_x(k/N) \right]^N \tag{2.28}$$

$$= \left[\exp\left(\sum_{m=0}^{N} \frac{(ik)^m}{m!N^m} C_m(x) \right) \right]^N \tag{2.29}$$

$$= \exp\left(\sum_{m=0}^{N} \frac{(ik)^m}{m!N^{m-1}} C_m(x) \right) \tag{2.30}$$

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:

$$\phi_S(k) \approx \exp\left(N + ikC_1(x) - \frac{k^2}{2N} C_2(x) \right) \tag{2.31}$$

$$= \exp\left(N + ik\langle x \rangle - \frac{k^2}{2N} \sigma_x^2 \right) \tag{2.32}$$

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

$$P_1(y,t) : \textit{the probability of being at value } y \textit{ at time } t \tag{2.33}$$
$$P_2(y_1,y_2|t_1,t_2) : \textit{probability of being at } (y_1,t_1) \textit{ and then } (y_2,t_2) \tag{2.34}$$
$$\vdots \textit{ and so on} \tag{2.35}$$

2.3.2 Normalisation

These probabilities are normalised by:

$$1 = \int dy_1 \dots dy_N \; P_N\left(y_1, y_2 \dots y_N \mid t_1, t_2 \dots t_N\right) \tag{2.36}$$
$$\tag{2.37}$$

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

$$1 = \int dy P_1(y,t) \tag{2.38}$$

2.3.3 Reduction

Theorem 2.3.1 Reduction

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

$$\int dy_N P_N\left(y_1 y_2 \dots y_N \mid \cdots\right) = P_{N-1}\left(y_1, \cdots y_{N-1}\right) \tag{2.39}$$

Example 2.3.2

$$\int P_2\left(y_1 y_2 \mid t_1 t_2\right) dy_2 = P_1\left(y_1 \mid t_1\right) \tag{2.40}$$

2.3.4 Correlation functions

Correlation functions are defined in the following ways:

Definition 2.3.2: Correlation functions

$$\langle y(t)\rangle = \int y \cdot P(y,t)dy \tag{2.41}$$

$$\langle y_1(t_1)\, y_2(t_2)\rangle = \int y_1 y_2 p_2(y_1 y_2 \mid t_1 t_2)\, dy_1 dy_2 \tag{2.42}$$

$$etc\cdots \tag{2.43}$$

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 $\langle y_1(t_1)\, y_2(t_2)\rangle$ 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 $\langle y_1(t_1)\, y_2(t_2)\rangle$ 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

$$P_2(y_2 y_1 \mid t_2 t_1) = G(2 \mid 1) P_1(y_1 \mid t_2) \tag{2.44}$$

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.(2.39) with the propagator Eq.(2.44) to reach something called the evolution relation:

$$P_1(y_2 t_2) = \int G(y_2 y_1 \mid t_2 t_1)\, P_1(y_1 t_1)\, dy_1 \tag{2.45}$$

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:

$$G(y_2, y_0 \mid t_2, t_0) = \int G(y_2, y_1 \mid t_2, t_1)\, G(y_1, y_0 \mid t_1, t_0)\, dy_1 \tag{2.46}$$

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. 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 these N steps (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 n into m-1 and regroup the terms

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

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

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':

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

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. 2.45), 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:

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

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

Where $w_f i$ is the transition probabilities between 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.