UCSB, Physics 129L, Computational Physics
Lecture notes, Week 6

Zihang Wang (UCSB), zihangwang@ucsb.edu

February 12, 2025

Contents

1 Central Limit Theorem (CLT)

The CLT states that when a sufficiently large number of independent and identically distributed (i.i.d.) random variables, $X_1, X_2, \dots X_N$ from a given (unknown) population, as the sample szie $N \to \infty$, the distribution of the sample mean (also the sum) $\bar{X} = \sum_j X_j/N$ will approximate a normal (Gaussian) distribution, regardless of the original population's distribution,

$$f(\bar{X}) \sim \frac{1}{\sqrt{2\pi\sigma^2/N}} \mathrm{Exp}\left(-\frac{(\bar{X}-\mu)^2}{2\sigma^2/N}\right), \tag{1}$$

where $f(\bar{X})$ is the probability density function (pdf) of the sample mean, μ is the mean of each X_j, and σ^2 is the variance of each X_j. Remember, the variance adds as quadrature. Therefore, an extra N in the denominator. We should also

note that it only applies to those distributions who have well-defined mean and variance, (e.g. Lorentzian).

CLT has important applications in physics. For example, let's say we want to track the how far and how fast on average a particle move over time T. This phenomenon is known as the **random walk** or **Brownian motion**, a stochastic process with the **Markov** property, and the corresponding evolutions in the phase space density will follow the deterministic diffusion equation. We record the random variable X_{t_i} (relative position with respect to $X_{t_{i-1}}$) at each time interval t_i,

$$\bar{X}_T = \sum_j X_{t_i}, \quad \bar{V}_T = \frac{\bar{X}_T}{T}. \tag{2}$$

The CLT states that as $T \to \infty$, the distribution of both $\bar{X}_T$ and $\bar{V}_T$ (average position and velocity of the particle) will approach a normal distribution.

It is worth noting the distinction between the solution of the diffusion equation and CLT. The diffusion equation is derived from the principle of conservation of particle number and provides information about the local density of particles at a specific time. In contrast, the random variable used in the CLT above characterizes the (average) positions of a single particle. These two concepts, while related, serve distinct purposes in modeling stochastic processes. As discussed previously, **ergodicity** is a concept that relates the behavior of a single-particle time average to many-particle ensemble average. In an ergodic system, the behavior of a single trajectory (e.g., the motion of a single particle) over time is representative of the behavior of the entire ensemble of particles.

2 Statistical Processes

2.1 Bernoulli process

A single trial in a **Bernoulli process** is called a Bernoulli trial, which has two outcomes: success, denoted as 1, with probability p, and failure, denoted as 0, with probability $1-p$. Tossing a coin is a simple example of a Bernoulli process since each toss of the coin can result in one of two possible outcomes: "Heads" ($X = 1$) or "Tails" ($X = 0$) with a probability mass function of p and $1-p$,

$$P_\delta(p, X) = \begin{cases} p, & X = 1 \\ 1-p, & X = 0 \end{cases}. \tag{3}$$

In quantum mechanics, the Stern-Gerlach experiment serves as a classical illustration of a Bernoulli process. An electron passes through an inhomogeneous magnetic field, yielding two distinct outcomes: the particles are deflected either "up" or "down", based on their intrinsic angular momentum, known as spin. In particular, the initial spin quantum state of a single electron is given by,

$$|\Psi\rangle_e = \sqrt{p}\,|\uparrow\rangle + e^{i\phi}\sqrt{1-p}\,|\downarrow\rangle\,, \tag{4}$$

and ϕ is a general phase factor. The probability is then given by the overlap, e.g. $|\langle\uparrow|\Psi\rangle_e|^2 = p$ and $|\langle\downarrow|\Psi\rangle_e|^2 = 1-p$. We emphasize that the trials in a Bernoulli process are independent.

2.2 Binomial Process

A binomial process consists of a sequence of N Bernoulli trials, and the probability mass function for having M successes is determined by the binomial distribution,

$$P_B(p, M) = \binom{N}{M} p^M (1-p)^{N-M}. \tag{5}$$

Instead of having a single electron in the Stern-Gerlach experiment, a beam (or a pack) of electrons (with fixed number N) is utilized, and the outcome intensity (snapshot) is characterized by the binomial distribution. The many-body quantum state can be written as,

$$|\Psi\rangle_{\text{e,beam}} = \bigotimes_j^N \left(\sqrt{p}\,|\uparrow_j\rangle + e^{i\phi_j}\sqrt{1-p}\,|\downarrow_j\rangle\right), \tag{6}$$

and the probability of observing a particular N-body quantum state, e.g.

$|\Psi\rangle_{\text{observe}} = |\uparrow\uparrow\downarrow\uparrow\ldots\downarrow\rangle$, with M spin up electrons is given by the **binomial distribution**. The key difference is that a Bernoulli process involves independent trials with two outcomes, while a binomial process counts successes in a fixed number of these trials.

For example, the probability of the first occurrence of an event after k Bernoulli trials with success probability p,

$$P_G(p, k) = (1-p)^{k-1}p, \tag{7}$$

which is the **geometric distributio**n.

2.3 Poisson Process

The **Poisson process** assumes that events are rare, independent, and isolated*. The probability of M events occurring in a time interval T is described by the probability mass function $P(M,T) = (\gamma T)^M/M!e^{-\gamma T}$, where γ represents a general rate function. The Poisson distribution can be derived from the binomial distribution in the limit $dt \to 0$,

$$P(M,T) = \lim_{dt\to 0} \frac{N!}{M!(N-M)!}(\gamma dt)^M(1-\gamma dt)^{N-M} = \frac{(\gamma T)^M}{M!}e^{-\gamma T}, \tag{8}$$

where $N = T/dt$, and the limits are,

$$\lim_{dt\to 0} \frac{N!}{M!(N-M)!}(\gamma dt)^M = \frac{(\gamma N dt)^M}{M!} = \frac{(\gamma T)^M}{M!}, \quad \lim_{dt\to 0}(1-\gamma dt)^{N-M} = e^{-\gamma T}. \tag{9}$$

In the second equality, we use the series expansion when $N \to \infty$,

$$\lim_{N\to\infty} \frac{N!}{M!(N-M)!} \approx N^M \left(1 - \frac{M(M-1)}{N} + \mathcal{O}(\frac{1}{N^2})\right), \tag{10}$$

and only kept the zero-th order.

The Poisson distribution is particularly useful in various applications, including astronomy, quantum optics, and telecommunications, where the detection of individual photons/particles is essential, and their arrival can be modeled as a Poisson process. For example, Poisson process is used to study atomic decay, photon detection, and quantum tunneling. Let's recall the Stern-Gerlach experiment of a beam of electrons. When $p \to 0$, the binomial distribution becomes the Poisson distribution. We should note that the number of total events goes to infinity.

Let's consider an example in astrophysics: Let's assume stars randomly distributed around us with density n, what is probability that the nearest star is at distance R ?

We first consider the total number of stars within a differential volume,

$$N = ndV = 4\pi R^2 ndR. \tag{11}$$

The probability of have no stars within $r < R$ is given by a Poisson process,

$$p(r < R, 0 \text{ star}|n) \sim \frac{(ndV)^0}{0!} e^{-n\frac{4\pi}{3}R^3}, \tag{12}$$

and it is worth noting that it only tells a proportionality without normalization. Similarly, the probability of having one star within the radius is proportional to the following,

$$p(r \approx R, 1 \text{ star}|n) \sim \frac{(ndV)^1}{1!} e^{-ndV} = \frac{(n4\pi R^2 dR)^1}{1!} e^{-n4\pi R^2 dR}. \tag{13}$$

Therefore, the joint probability has the product form, i.e., no stars are within the volume enclosed by R, and exactly one star at radius $R \pm dR$,

$$\begin{aligned} &p(r < R, 0 \text{ star}, r \approx R, 1 \text{ star}|n) = p(r < R, 0 \text{ star}|n)p(r \approx R, 1 \text{ star}|n) \\ &\sim n4\pi R^2 dR e^{-n4\pi R^2 dR} e^{-n\frac{4\pi}{3}R^3} \approx n4\pi R^2 dR e^{-n\frac{4\pi}{3}R^3}. \end{aligned} \tag{14}$$

2.4 Lorentzian

Resonance behavior is critical for understanding the energy dissipation and absorption of a system. As an example, let's consider a general dissipative, driven harmonic oscillator in both time and frequency domain,

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \gamma\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = Fe^{i\omega_f t}, \quad -\omega^2\tilde{x} + i\gamma\omega\tilde{x} + \omega_0^2\tilde{x} = F\delta(\omega - \omega_f), \tag{15}$$

where $\tilde{x}$ and x are related via Fourier transform. You can see that the first order term has an imaginary coefficient, capturing the dissipative nature of the

system. The natural and driven frequency are labeled by ω_0 and ω_f. γ is the damping factor. Then, we have the solution in frequency and time domain,

$$\tilde{x} = \frac{F\delta(\omega - \omega_f)}{\omega_0^2 - \omega^2 + i\gamma\omega}, \quad x = \mathcal{F}^{-1}(\tilde{x}) = \frac{e^{i\omega_f t}}{\omega_0^2 - \omega_f^2 + i\gamma\omega_f} = Ae^{i\omega_f t + i\Sigma}, \tag{16}$$

where,

$$A = \frac{1}{\sqrt{(\omega_0^2 - \omega_f^2)^2 + \gamma^2\omega_f^2}}, \quad \tan(\Sigma) = \frac{-\gamma\omega_f}{\omega_0^2 - \omega_f^2}. \tag{17}$$

We finally have the energy absorption per cycle as a Lorentzian,

$$\begin{aligned} E &= AF\mathrm{Re}\left\{\int e^{-i\omega_f t}\frac{\mathrm{d}}{\mathrm{d}t}e^{i\omega_f t + i\Sigma}\mathrm{d}t\right\} = i\omega_f AF\mathrm{Re}\left\{\int_0^{\pi/\omega_f} e^{i\Sigma}\mathrm{d}t\right\} \\ &= -\frac{\pi}{\omega_f}AF\omega_f\sin(\Sigma) = \pi A^2 F\gamma\omega_f = F\pi\frac{\gamma\omega_f}{(\omega_0^2 - \omega_f^2)^2 + \gamma^2\omega_f^2}, \end{aligned} \tag{18}$$

and we use the fact $\sin(\Sigma) = -A\gamma\omega_f$. The universal absorption profile of the driven harmonic oscillator, characterized by the Lorentzian distribution, holds profound implications across various fields of physics. Although we will not discuss it in detail, it's worth mentioning that this represents a specific case of the spectral function, a fundamental idea in many-body physics and modern quantum field theory.

3 Stochastic process

As a sequence of random variables, a stochastic (or random) process describes the time evolutions of a collection of random variables within a probability space. Stochastic processes serve as essential mathematical models for physical systems and phenomena that have deep connections via the Fluctuation–dissipation theorem. For example:

- The random movement of microscopic particles suspended in a fluid, caused by collisions with fluid molecules, can be modeled as **Brownian motion** (or a **Markov process**), a type of stochastic process that is considered one of the foundational processes in statistical mechanics. In continuous mechanics, the time evolution of phase space density (a statistical quantity) is probabilistic, governed by **classical master equations**. You should notice the difference between the classical master equation and the **Liouville's equation**: the latter captures the deterministic evolution of phase space density, which does not involve any stochastic process.

- The probabilistic nature of particle positions and momenta is a direct manifestation of the stochastic processes underlying quantum mechanics. The time evolution of quantum eigenstates (or density matrices) is governed by the **quantum master equations**. Let's look at the difference

between **Von Neumann Equation** and the quantum master equation: The former describes a closed quantum system where the evolution is unitary. On the other hand, quantum master equations describe the evolution of the density matrix of an open quantum system, which interacts with an external environment. They are typically non-unitary due to dissipative and decoherence effects. The general form is given by,

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + L(\rho), \tag{19}$$

where the first term is the Von Neumann Equation, and $L(\rho)$ is a dissipative term accounting for interaction with the environment. The most common form of QME is the **Lindblad Master Equation**.

- Due to an exponential sensitivity on initial conditions (e.g. the Logistic Map), such as **bifurcations and chaos** in Turbulent flows can be understood as stochastic processes where energy dissipation and scale interactions are untractable in contrast to the in laminar flow, yet both are governed deterministically via the **Navier–Stokes equations**.

3.1 Time series

A time series is defined as a sequence of random variables (e.g. observations) $\{X_t\}$ with discrete time or continuous time t,

$$X_t = S_t + \epsilon_t, \tag{20}$$

where S_t and ϵ_t are random variables representing signal and noise such that,

$$S_t \sim f(t), \quad \epsilon_t \sim g(t) \tag{21}$$

where $f(t), g(t)$ are probability density functions (PDFs) at time t. The PDF of the random variable X_t is given by the convolution between the signal and noise,

$$h(t) = [f * g](t) = \int_{-\infty}^{\infty} d\tau f(t-\tau)\Theta(t-\tau)g(\tau), \tag{22}$$

where the time order is enforced via the step function $\Theta(t-\tau)$. In frequency space, it becomes the product,

$$\tilde{h}(\omega) = i\tilde{g}(\omega)\int_{-\infty}^{\infty} d\omega' \frac{\tilde{f}(\omega')}{\omega - \omega' + i\eta} = i\tilde{g}(\omega)PV\left(\int_{-\infty}^{\infty} d\omega' \frac{\tilde{f}(\omega')}{\omega - \omega'}\right) + \pi\tilde{g}(\omega)\tilde{f}(\omega), \tag{23}$$

where $\tilde{g}(\omega), \tilde{f}(\omega)$ are Fourier transforms, and the integration axis is shifted by η, and the last line uses the residue theorem,

$$\int_{-\infty}^{\infty} dt\Theta(t)e^{i[\omega-\omega'+i\eta]t} = \lim_{\eta\to 0^+} \frac{i}{\omega - \omega' + i\eta} = PV\left(\frac{i}{\omega - \omega'}\right) + i(-i)\pi\delta(\omega - \omega'). \tag{24}$$

Let's consider a general case, $\eta = \eta_0 + \eta'$. η_0 is finite and positive while $\eta' \to 0^+$,

$$\tilde{h}(\omega) = i\tilde{g}(\omega) \int_{-\infty}^{\infty} d\omega' \frac{\tilde{f}(\omega')}{\omega - \omega' + i\eta} = i\tilde{g}(\omega)(\omega - i\eta_0) \int_{-\infty}^{\infty} d\omega' \frac{\tilde{f}(\omega')}{(\omega - \omega')^2 + \eta_0^2}, \tag{25}$$

The δ term becomes zero because $\omega + i\eta_0 \neq \omega'$ since $\omega, \omega' \in \mathbb{R}$.

The **linear response function** is defined as the ratio between the observation and noise in frequency space,

$$\tilde{h}(\omega)/\tilde{g}(\omega) \equiv \chi(\omega) = \int_{-\infty}^{\infty} d\omega' \frac{\tilde{f}(\omega')}{\omega - \omega' + i\eta}, \tag{26}$$

Its imaginary part is called the **spectrum** of the signal,

$$\mathrm{Im}(\chi(\omega)) = \begin{cases} \int_{-\infty}^{\infty} d\omega' \frac{\eta_0}{(\omega-\omega')^2+\eta_0^2} \tilde{f}(\omega'), & \eta_0 \neq 0 \\ \pi \tilde{f}(\omega), & \eta_0 = 0 \end{cases} \tag{27}$$

When $\eta_0 = 0$, only the delta function term contributes.

Note the Lorentzian function (distribution) in the integrand,

$$A(\omega'|\omega) = A(\omega - \omega') = \frac{\eta}{(\omega - \omega')^2 + \eta^2}, \tag{28}$$

which captures the probability density of "exciting" a mode with frequency ω' to a mode with frequency frequency ω. In the language of probability theory, it can be understood as the **likelihood function** or the conditional probability $A(\omega'|\omega)$, conditioning on the mode ω, leading to the following **marginal probability**,

$$\mathrm{Im}(\chi(\omega)) = \int_{-\infty}^{\infty} d\omega' A(\omega'|\omega) \tilde{f}(\omega'). \tag{29}$$

The physical meaning of the linear response function is somewhat clear now: it reflects the **marginal probability** of "observing" a mode ω generated from **all** other modes ω' with transition amplitude $A(\omega - \omega')$ scaled by the "strength" $\tilde{f}(\omega')$. You might notice the similarity with the response function in resonances or quasiparticle spectral function in Fermi-liquid theory.

The probability density function of observations (i.e. time series) in frequency space is a simple product of the response function and the noise,

$$\tilde{h}(\omega) = \chi(\omega)\tilde{g}(\omega), \tag{30}$$

and the inverse Fourier transform gives back the convolution between the response function and the noise,

$$h(t) = \int_{-\infty}^{\infty} d\tau \chi(t - \tau) g(\tau), \tag{31}$$

induced via time dependent processes (e.g. external perturbation). The response function $\chi(\omega)$ is also a correlation measure, which has profound consequence in modern physics, enabling various computational approaches.

3.2 Markov Chain

A Markov chain is a sequence of random variables $X_0, X_1, X_2, \ldots$, where each X_n represents the state of a system at time n. The sequence $\{X_n\}_{n=0}^{\infty}$ is called a Markov chain on a continuous (discrete) **state space** $\boldsymbol{\mathcal{S}}$ if it satisfies the **Markov property**: for all $n \geq 0$ and all states $s_0, s_1, \ldots, s_{n+1} \in \boldsymbol{\mathcal{S}}$, the conditional probability,

$$\Pr(X_{n+1} = s_{n+1} \mid X_n = s_n, \ldots, X_0 = s_0) = \Pr(X_{n+1} = s_{n+1} \mid X_n = s_n). \quad (32)$$

This "memory-less" property states that the probability of transitioning to the next state depends only on the current state and not on the sequence of states that preceded it. The random walk (or Brownian motion) is a classical example of a Markov chain:

In modern quantum mechanics, concepts of Markov chain has been investigated intensively in modeling thermal environments, e.g. open quantum systems. An open quantum system interacts with its surrounding environment, leading to a complex state evolution that can no longer be fully described in a isolated and separable form (e.g. quantum system's Schrodinger equation).

Open quantum systems lose information to their environment, causing effects like decoherence and dissipation, which are key research challenges in current quantum computation, quantum information, and quantum thermodynamics. These interactions between the system and its environment are often modeled using Markovian and non-Markovian processes. In the Markovian approach, the evolution of the open quantum system is typically assumed to be "memoryless".

A **finite-state, discrete-time Markov chain** has a finite set of possible states, often denoted by $S = \{s_1, s_2, \ldots, s_k\}$. The probabilities of moving from one state to another are specified by a **transition matrix** P, where each element $p_{ij} = \Pr(X_{n+1} = s_j \mid X_n = s_i)$ represents the probability of transitioning from state s_i to state s_j in a single time step. The matrix P is a square $k \times k$ matrix, with each row summing to 1, since these rows represent probability distributions:

$$\sum_{j=1}^{k} p_{ij} = 1, \quad \text{for all } i = 1, 2, \ldots, k. \quad (33)$$

Example of Transition Matrix: For a system with three states $S = \{A, B, C\}$, a possible transition matrix P could look like the following,

$$P = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.1 & 0.6 & 0.3 \\ 0.4 & 0.1 & 0.5 \end{bmatrix}. \quad (34)$$

Each entry p_{ij} in this matrix represents the probability of transitioning from state i to state j.

The **n-step transition probabilities** give the probability of being in a particular state after n steps. They are given by the n-th power of the transition

matrix, P^n, where $[P^n]_{ij}$ represents the probability of being in state s_j after n steps, starting from state s_i.

The probability at step n and at a given state s_i is given by $\pi_i(n) = \Pr(X_n = s_i)$, and the system at a particular step is represented by the row vector, $\pi(n) = (\pi_1, \pi_2, \ldots, \pi_k)$.

A Markov chain may have a **stationary distribution**, which is a probability distribution over states that remains unchanged under the transition dynamics. Mathematically, π is a vector that satisfies,

$$\pi P = \pi, \quad \text{and} \quad \sum_{i=1}^{k} \pi_i = 1. \tag{35}$$

This stationary distribution exists for certain types of chains (e.g., irreducible, aperiodic chains) and describes the long-term behavior of the system. For example, as $n \to \infty$, the probability of finding the chain in any particular state converges to the corresponding entry in π.

Piratically, we can iteratively multiply an initial guess π^0 by the transition matrix P, which will eventually converge to the stationary distribution. This is the **power iteration**, with the following steps,

- Initialize a random probability vector π^0 (usually with positive values and normalized so that the sum of the entries is 1).
- Iterate the following update:

$$\pi^{r+1} = \pi^r P, \tag{36}$$

  where π^r is the margional distribution at the r-th step.
- Normalize the distribution π^{r+1} (though it will remain normalized as long as the sum of its components is 1).
- Repeat steps 2 and 3 until the vector π^r converges.

3.3 Discrete-Time Transition Matrix to Continuous-Time Transition Matrix

Let's consider a discrete-time Markov process over n time steps, represented by the n-th power of the transition matrix P, P^n. Let's consider the following transition matrix,

$$P = \begin{pmatrix} 1-\lambda & \lambda \\ \lambda & 1-\lambda \end{pmatrix}, \tag{37}$$

we want to find the behavior of P^n as $n \to \infty$. We can write the matrix P as.

$$P = \begin{pmatrix} 1-\lambda & \lambda \\ \lambda & 1-\lambda \end{pmatrix} = I + \lambda \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{38}$$

To conserve the probability, the sum of each row must be 1, i.e.,

$$P_{ii} + \sum_{j \neq i} P_{ij} = 1. \tag{39}$$

Let's define a matrix Q as,

$$Q = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{40}$$

and we write P as,

$$P = I + \lambda Q. \tag{41}$$

We can see that Q does not depends on the self-transition $i \to i$ as it is directly incorporated by the normalization. We use the fact that for large n, P^n can be approximated using the **matrix exponential** form. The matrix Q has a very simple structure, so we can compute powers of P.

The power P^n for large n will behave similarly to,

$$P^n = \lim_{n \to \infty} (I + n\frac{\lambda}{n}Q)^n \approx e^{n\lambda Q}. \tag{42}$$

We can compute $e^{n\lambda A}$ using the Taylor series expansion,

$$e^{n\lambda Q} = I + n\lambda Q + \frac{(n\lambda Q)^2}{2!} + \frac{(n\lambda Q)^3}{3!} + \dots \tag{43}$$

The power iteration defines the **continuous-time Markov chain**,

$$\pi^n = \pi^0 P^n \to \pi(t) = \pi^0 P(t), \tag{44}$$

where we define the time as $t = n\lambda = n\Delta t$ such that,

$$Q \sim \frac{1}{n\Delta t} \ln(P^n). \tag{45}$$

3.4 Classical Master Equation

The classical master equation describes Markov processes, where a system transitions between discrete states with well-defined probabilities. It takes the general form:

$$\frac{d\pi_i}{dt} = \sum_j \left(Q_{ji}\pi_j - Q_{ij}\pi_i \right), \quad \frac{d\pi}{dt} = \pi Q, \tag{46}$$

where $\pi_i(t)$ is the probability of being in state i, and Q_{ij} are transition rates satisfying **detailed balance** at equilibrium,

$$\frac{\pi_j}{\pi_i} = \frac{Q_{ij}}{Q_{ji}} = e^{-(E_j - E_i)/k_B T}. \tag{47}$$

3.5 Classical Master Equation for 2-level systems

Consider a system with two states: State 1 with energy E_1 and state 2 with energy E_2. The transition rates between the states follow Boltzmann-type expressions,

$$Q_{12} = Q_0 e^{-(E_2-E_1)/k_B T}, \quad Q_{21} = Q_0 e^{-(E_1-E_2)/k_B T}. \tag{48}$$

Let $\pi_1(t)$ and $\pi_2(t)$ be the probabilities of the system being in states 1 and 2 at time t. The Master equation governing the evolution of these probabilities is:

$$\frac{d\pi_1}{dt} = -Q_{12}\pi_1 + Q_{21}\pi_2 \tag{49}$$

$$\frac{d\pi_2}{dt} = -Q_{21}\pi_2 + Q_{12}\pi_1 \tag{50}$$

Since total probability is conserved:

$$\pi_1 + \pi_2 = 1 \tag{51}$$

At equilibrium ($\frac{d\pi_1}{dt} = \frac{d\pi_2}{dt} = 0$), we set:

$$Q_{12}\pi_1 = Q_{21}\pi_2 \tag{52}$$

Solving for π_1 and π_2:

$$\frac{\pi_2}{\pi_1} = \frac{Q_{12}}{Q_{21}} = e^{-(E_2-E_1)/k_B T} \tag{53}$$

we have the Steady-State Solution, which gives the Boltzmann distribution:

$$\pi_1 = \frac{1}{1 + e^{-(E_2-E_1)/k_B T}}, \quad \pi_2 = \frac{e^{-(E_2-E_1)/k_B T}}{1 + e^{-(E_2-E_1)/k_B T}} \tag{54}$$

This shows that at high temperatures ($T \to \infty$), both states are equally probable, while at low temperatures ($T \to 0$), the system strongly favors the lower-energy state.

One might notice a problem in here if you want to extend the above expression into quantum regime: no phase information is encoded, i.e. the system lacks quantum coherence or superpositions. To see this, let consider the following classical ensemble at step n,

$$\rho_{\text{classical}}(n) = \sum_i \pi_i(n) \left|\psi_i\right\rangle \left\langle\psi_i\right|, \tag{55}$$

where $|\psi_i\rangle$ are general quantum states (e.g. energy eigenstates $|E_i\rangle$). The transition rate are classical objects, e.g., in the above two-level system, we have,

$$Q_{12} \equiv \Pr(|E_1\rangle \langle E_1| \to |E_2\rangle \langle E_2|, n) \sim Q_0 e^{-(E_2-E_1)/k_B T}, \tag{56}$$

and in general, the transition rate between two diagonal elements in the density matrix has the general form,

$$Q_{ij} \equiv \Pr(|\psi_i\rangle \langle\psi_i| \to |\psi_j\rangle \langle\psi_j|, n), \tag{57}$$

and it no longer has the Boltzmann-type expression. It is clear that no quantum coherence $|\psi_i\rangle \langle\psi_j|$ (off-diagonal terms in the density matrix) can be captured by the classical master equation,

$$\rho_{\text{quantum}}(n) = \sum_j \sum_i \pi_j(n) |\psi_i\rangle \langle\psi_i| . \tag{58}$$

In most applications, quantum coherence can be corrected via some **local hidden theories**. In other words, one can approximate the quantum evolution by the evolution of diagonal elements (diagonal approximation). However, for simulating dynamics of long-coherence system, e.g. demonstrating **Rabi oscillations** or **Bell' inequality**, one must go beyond the non-equilibrium classical regime or even beyond the Markov regime. This motivates the development of theories in **non-equilibrium quantum dynamics**.

3.6 Quantum Master Equation (Lindblad Formulation)

For quantum systems, classical probabilities are insufficient, and we describe the evolution of the density matrix ρ using the Lindblad master equation,

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + \sum_j \left(L_j \rho L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, \rho\} \right). \tag{59}$$

H is the system Hamiltonian, and L_j are **jump operators** modeling decoherence and dissipation. The Lindblad equation must be used when the quantum coherence and superposition dominates the system dynamics, such as in open quantum systems, including superconducting qubits, laser-driven atoms, and quantum thermodynamics.

In summary,

- If **coherence is negligible** and only populations matter → **Classical master equation**.
- If **quantum coherence and phase evolution** are relevant → **Lindblad equation**.