Nonequilibrium Field Theories and Stochastic Dynamics

Zehong Liao

September 25, 2025

1 From Random Walks to Physical Laws

A stochastic process is a mathematical framework for describing systems that evolve probabilistically over time. Among the many stochastic processes, the random walk is undoubtedly the most typical and fundamental model. Its applications are ubiquitous, ranging from the diffusion of molecules inside cells to the fluctuations of prices in financial markets, where its basic concept plays a central role. This note aims to systematically organize and elucidate the one-dimensional simple random walk model, based on the lectures of Professor Erwin Frey.

1.1 Simple Random Walk Model

Basic elements of the model: Imagine a particle (or "walker") moving along a one-dimensional line. Its motion is not continuous but occurs in discrete time steps $t_k = k\tau$, where $k = 1, 2, 3, \ldots$ denotes the number of steps and τ represents the time interval of each jump. The particle's position observed at any time t_k is denoted by X_k.

Core random variable: Step length ξ_k — the displacement of each jump, which is treated as a random variable. The key assumption of this model is that all step lengths are independent and identically distributed random variables. "Identically distributed" means that each jump follows the same probabilistic rule, while "independent" means that the outcome of any single step does not influence any other step.

Discrete and continuous step lengths: The distribution of step lengths can be either discrete or continuous.

Discrete case: The step length can only take a finite set of specific values. The simplest case is that, at each step, the particle can move either a distance a to the right or a distance a to the left, that is, $\xi_k \in \{\pm a\}$. The probability of moving to the right is p, while the probability of moving to the left is q, with $p + q = 1$. This is the central topic of this lecture — Simple Random Walk.

Continuous case: The step length can take any real value, with its probability described by a probability density function $W(\xi)$. Specifically, the probability that the step length falls within the interval $[\xi, \xi + d\xi]$ is given by $W(\xi)d\xi$.

The dynamic evolution of this model can be described by a simple recursive relation:

$$X_k = X_{k-1} + \xi_k \tag{1}$$

This equation expresses the model's "local" rule of motion: the new position equals the old position plus a random step length. If we assume that the particle starts from the origin, i.e., $X_0 = 0$, then by iteratively expanding the above recursive relation, we can obtain an explicit expression for the particle's position at the k-th step.

$$X_k = \sum_{j=1}^{k} \xi_j \tag{2}$$

This equation reveals the deepest mathematical structure of the model: the macroscopic physical quantity we observe (the position X) is, in essence, the sum of a large number of microscopic, independent random variables (the step lengths ξ). It is precisely this "additive" structure that allows us to employ powerful statistical tools, such as the law of large numbers and the central limit theorem, to analyze this seemingly entirely random process and uncover deterministic, predictable macroscopic

laws. The "Identically distributed" assumption is the cornerstone of all such analyses, as it ensures that the statistical properties of the process remain unchanged over time, thereby enabling the derivation of time-independent physical constants such as drift velocity and diffusion coefficient.

1.2 Statistical Averages and the Emergence of Physical Laws

For a stochastic process, we cannot precisely predict the walker's exact position in a single realization. However, by analyzing the statistical average behavior over many such realizations, we can extract deterministic physical information. Instead of asking, "Where is the particle?" we ask, "Where is the particle on average?" and "How dispersed are the particle's possible positions?" These questions can be addressed by computing the moments of the random variable, such as the mean and variance.

1.2.1 Mean Position: Revealing the Drift Phenomenon

We first calculate the expected value (or mean) of the particle's position after k steps, denoted as $\langle X_k \rangle$. By exploiting the linearity of expectation, the expected total displacement equals the sum of the expected displacements at each individual step.

$$\langle X_k \rangle = \langle \sum_{k}^{j=1} \xi_j \rangle = \sum_{k}^{j=1} \langle \xi_j \rangle \tag{3}$$

The expected value of a single-step displacement, $\langle \xi \rangle$, can be calculated. By definition of expectation, it equals the sum of all possible values multiplied by their corresponding probabilities:

$$\langle \xi_j \rangle = (+a) \cdot P(\xi_j = a) + (-a) \cdot P(\xi_j = -a) = ap - aq = a(p - q) \tag{4}$$

This value represents the average displacement per step.

Since all step lengths are identically distributed, the expectation is the same for each step. Therefore, the expected total displacement is given by:

$$\langle X_k \rangle = \sum_{k}^{j=1} a(p - q) = k \cdot a(p - q) \tag{5}$$

Physical Interpretation: This mathematical result indicates that the particle's mean position is proportional to the number of steps k. By introducing the physical time $t_k = k\tau$, we can define an average velocity, known as the drift velocity v:

$$v = \frac{\langle X_k \rangle}{t_k} = \frac{k \cdot a(p - q)}{k\tau} = \frac{a(p - q)}{\tau} \tag{6}$$

This means that, although the trajectory of an individual particle is erratic and unpredictable, the particle ensemble as a whole appears to move in a directed manner at a constant velocity v_{d}. This macroscopic directed motion, arising from a probabilistic asymmetry, is referred to as drift.

Symmetric and Asymmetric Walks: The expression for the mean position is entirely determined by the properties of (p-q). This reveals two fundamental behavioral patterns of the model:

- Symmetric Random Walk: In this case, (p = q = 0.5), and thus $\langle X_k \rangle$. The probabilities of moving left and right are equal, with no directional preference. On average, the particle remains near the origin. This represents a pure diffusion process.

- Biased/Asymmetric Random Walk: In this case $p \neq q$, $\langle X_k \rangle \neq 0$, resulting in a nonzero drift velocity. On average, the particle moves toward the direction with higher probability. the dashed line showing the mean position $\langle x_k \rangle$ increasing linearly with time clearly illustrates a biased random walk with $p > q$.

Thus, the value of the parameter p connects the abstract mathematical model to concrete physical scenarios. A dye molecule diffusing in a stationary solvent can be described by a symmetric model, whereas a charged particle moving under an external electric field (electrophoresis) or a particle suspended in a flowing liquid requires a biased model for accurate description.

1.2.2 Variance: The Mathematical Fingerprint of Diffusion

Next, we analyze the variance of the particle's position, $\mathrm{Var}(x_k)$, which quantifies the dispersion or spread of the particle's positional distribution. By definition, the variance is given by $\mathrm{Var}(X_k) = \langle X_k^2 \rangle - \langle X_k \rangle^2$. We have already obtained $\langle X_k \rangle$; the next task is to calculate the second moment $\langle X_k^2 \rangle$.

$$\langle X_k^2 \rangle = \langle (\sum_{j=1}^{k} \xi_j)^2 \rangle = \langle (\sum_{i=1}^{k} \xi_i)(\sum_{j=1}^{k} \xi_j) \rangle = \sum_{i=1}^{k} \sum_{j=1}^{k} \langle \xi_i \xi_j \rangle \tag{7}$$

Separate diagonal and off-diagonal terms. This is a key technique in the calculation process. We split the double summation into two parts:

- Diagonal terms ($i = j$): $\sum_{j=1}^{k} \langle \xi_j^2 \rangle$
- Off-diagonal terms ($i \neq j$): $\sum_{i \neq j} \langle \xi_i \xi_j \rangle$

For the off-diagonal terms, since the step lengths ξ_i and ξ_j are independent, the expectation of their product equals the product of their individual expectations:

$$\langle \xi_i \xi_j \rangle = \langle \xi_i \rangle \langle \xi_j \rangle = (a(p-q))^2 \tag{8}$$

For the diagonal terms, we calculate the expectation of the square of the step length.

$$\langle \xi_j^2 \rangle = (+a)^2 \cdot P(\xi_j = a) + (-a)^2 \cdot P(\xi_j = -a) = a^2 p + a^2 q = a^2 (p+q) = a^2 \tag{9}$$

Combine the calculations: there are k diagonal terms and $k(k-1)$ off-diagonal terms. Therefore,

$$\langle X_k^2 \rangle = k \cdot a^2 + k(k-1) \cdot (a(p-q))^2 \tag{10}$$

By definition, the variance is given by $\mathrm{Var}(X_k) = \langle X_k^2 \rangle - \langle X_k \rangle^2$:

$$\mathrm{Var}(X_k) = (k \cdot a^2 + k(k-1) \cdot (a(p-q))^2) - (ka(p-q))^2 = ka^2(1 - (p-q)^2) \tag{11}$$

Using $p + q = 1$, we have $1 - (p-q)^2 = (p+q)^2 - (p-q)^2 = 4pq$. Thus, we finally obtain:

$$\mathrm{Var}(X_k) = 4ka^2pq \tag{12}$$

Physical Interpretation: The most important feature of this result is that the variance is proportional to the number of steps k, and hence also proportional to the physical time t_k. This linear growth of variance with time is the hallmark of a diffusive process.

We can define the diffusion coefficient D through the relation $\mathrm{Var}(x) = 2Dt_k$. By comparing with our derived result, we obtain:

$$2D = \frac{4ka^2pq}{t_k} = \frac{4ka^2pq}{k\tau} = \frac{4a^2pq}{\tau} \tag{13}$$

The diffusion coefficient D is a physical quantity that measures the rate of stochastic diffusion.

The physical meaning of the linear growth of variance with time ($\mathrm{Var}(x) \propto t$) is that the width of the particle's position distribution, i.e., the standard deviation $\sigma = \sqrt{\mathrm{Var}(x)}$, grows with the square root of time ($\sigma \propto \sqrt{t}$). This sub-linear growth is a universal feature of diffusive processes and is fundamentally different from ballistic motion, where the position grows linearly with time ($x \propto t$). In the PPT, the cone-shaped region formed by $\pm\sigma$ (also called the "funnel of uncertainty") provides an excellent visual representation of this basic scaling law. This scaling explains why diffusion is efficient at microscopic scales, such as inside cells, but extremely slow at macroscopic scales.

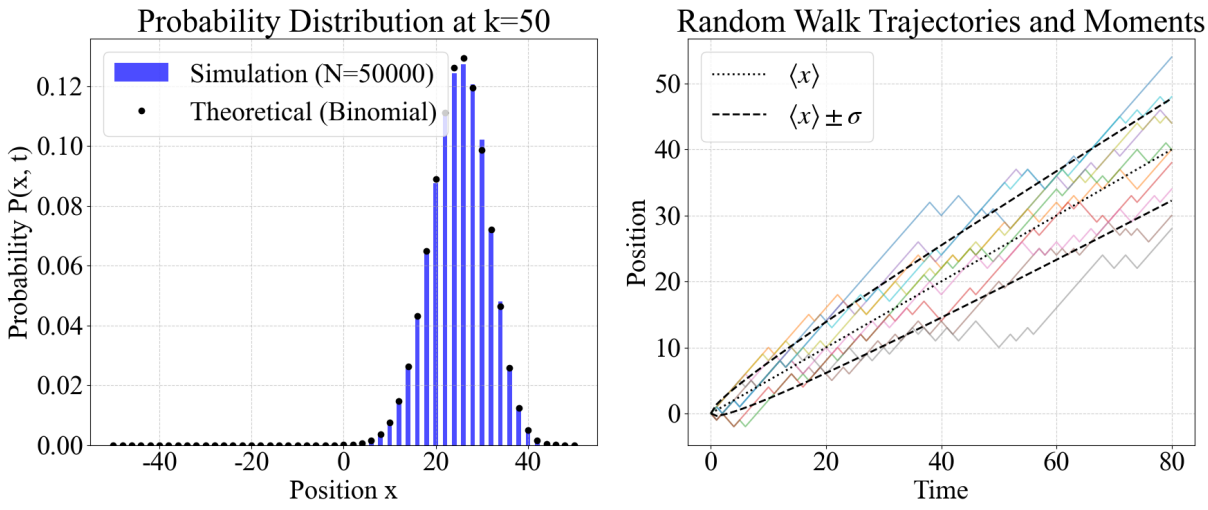

Figure 1:

1.2.3 Complete Probability Distribution

The mean and variance describe the center and the width of the distribution, but to obtain complete information about the particle's position, we need to determine its full probability distribution function $P(x, k)$. This function answers the question: "After k steps, what is the exact probability of finding the particle at position x?"

Binomial Distribution: We can obtain an exact solution to this problem using combinatorial methods. Assume that during k steps the particle makes k_+ steps to the right and k_- steps to the left. Then $k = k_+ + k_-$. The final position is $X_k = (k_+ + k_-)\, a$. For convenience, we introduce the dimensionless position $n \equiv k_+ + k_-$, so that $X = n\, a$.

We can express k_+ and k_- in terms of the known quantities k and n:

$$k_+ = \frac{k+n}{2}, k_- = \frac{k-n}{2} \tag{14}$$

It is important to note that, in order for k_+ and k_- to be integers, k and n must have the same parity (both even or both odd). Otherwise, the particle cannot reach position $x = na$, and the probability is zero.

Probability of a single path: For a given sequence of k_+ right steps and k_- left steps (each step is independen), the probability is

$$p^{k_+} q^{k_-} = p^{k_+} (1-p)^{k_-} \tag{15}$$

Total number of paths: The number of distinct sequences that we can go to the final position $x = na$ (for total steps k, and the right step k_+) is given by the binomial coefficient:

$$\frac{k!}{k_+!(1-k_+)!} = \frac{k!}{k_+!k_-!} \tag{16}$$

By multiplying the probability of a single path by the total number of such paths, we obtain the total probability of finding the particle at position $x = na$ after k steps. This is exactly the Binomial Distribution:

$$P(X_k = na) = \binom{k}{k_+} p^{k_+} (1-p)^{k_-}, k_+ = \frac{k+n}{2},\ k_- = \frac{k-n}{2} \tag{17}$$

Connection between Microstates and Macrostates: The structure of this formula profoundly reflects the fundamental ideas of statistical mechanics. Each specific sequence of steps represents a microstate, while the resulting net displacement (e.g., $n = k_+ - k_-$) corresponds to a macrostate. The formula shows that the probability of a macrostate is equal to the probability of a single microstate ($p^{k_+} q^{k_-}$) multiplied by the degeneracy or multiplicity ($\binom{k}{k_+}$), i.e., the number of distinct microstates that lead to the same macro outcome. The most probable macrostate is the one corresponding to the largest number of microstates. In this context, the combinatorial coefficient plays a role analogous to entropy.

Emergence of the Gaussian Distribution: Central Limit Theorem (CLT)

The binomial distribution is exact, but when the number of steps k becomes very large, its computation becomes cumbersome. In physics, we are often more interested in the macroscopic behavior after many steps. When k is very large ($k \gg 1$), the binomial distribution can be approximated by a Gaussian distribution, also known as the normal distribution:

$$P(X_k = na) \approx \frac{1}{\sqrt{2\pi \mathrm{Var}(n)}} \exp(\frac{(n - \langle n \rangle)^2}{2\mathrm{Var}(n)}) \tag{18}$$

This transition is not merely a mathematical approximation; it reflects a deeper physical principle—the Central Limit Theorem. The theorem states that the sum of a large number of independent and identically distributed random variables (in our model, the total displacement $X_k = \sum \xi_j$) tends toward a Gaussian distribution, regardless of the distribution of the individual variables (here, the single-step displacement follows a simple two-point Bernoulli distribution). This explains why, in nature, processes involving the accumulation of many random factors often result in Gaussian-distributed outcomes.

Figure. 1 perfectly demonstrates the self-consistency of the model. The results obtained by directly calculating the moments coincide exactly with those derived by solving the full distribution and then taking the large-step limit. This process clearly illustrates how the macroscopic diffusion law, whose solution is a Gaussian function, emerges from the microscopic rules of random walks.

2 Gaussian Random Walk and Poisson Process

2.1 Gaussian Distribution: A Universal Attractor

When the the step number is very large, the binomial distribution that describes the particle's position can be very well approximated by a Gaussian (normal) distribution. This also is the core content of Central Limit Theorem, CLT.

$$P(X_k = na) = \binom{k}{k_+} p^{k_+}(1-p)^{k_-} \approx \frac{1}{\sqrt{2\pi \mathrm{Var}(n)}} \exp(\frac{(n - \langle n \rangle)^2}{2\mathrm{Var}(n)}) \tag{19}$$

Here, $\langle n \rangle = k(p - q)$, $\mathrm{Var}(n) = 4kpq$. $P(x, t)$ is the probability density of finding the walker at position x at time t. The peak of the distribution moves with velocity v, which corresponds to the drift we discussed earlier. The width of the distribution is determined by the standard deviation σ, which grows with the square root of time. This indicates that the particle is spreading outward through diffusion, as shown in the slides: over time, the blue curve becomes broader and flatter.

The Gaussian random walk is the *universal attractor* of all random walks with weak local correlations. This concept comes from dynamical systems and statistical physics, and its physical meaning is far deeper than a mere mathematical approximation. The physical premise of this approximation is that our observational scale (the final displacement x) is much larger than the single-step length a, and the total time t is much longer than the single-step duration τ. It is precisely under this macroscopic perspective that microscopic discreteness ceases to matter, and a continuous, smooth Gaussian distribution emerges as an effective macroscopic description.

This means that at long times and large scales, the macroscopic statistical behavior of a random process (ultimately manifesting as a Gaussian distribution) does not depend on its microscopic details (for example, whether each step follows a Bernoulli distribution, a uniform distribution, or something else). As long as the single-step random variable satisfies certain basic conditions (such as having a finite variance), the cumulative outcome of many steps will always be "attracted" to the same endpoint—the Gaussian distribution.

An intuitive interpretation of the Central Limit Theorem (CLT) is that the sum of many independent and identically distributed random variables (here, our steps) will approach a Gaussian (normal) distribution, regardless of the precise form of the single-step distribution (so long as its variance is finite). Therefore, the CLT serves as a bridge connecting the discrete microscopic world and the continuous macroscopic world, and reveals a profound principle in nature: *universality*. It is precisely this universality that allows us to ignore the complexity of microscopic details and instead capture the common laws governing the macroscopic behavior of systems. This idea also lays the groundwork for our later understanding of the more general framework of the renormalization group theory.

The Gaussian Random Walk (GRW) describes a process that is continuous in space but evolves in time with fixed steps. It is defined as a sequence of random variables, whose evolution rule is as follows:

$$X_t = X_{t-1} + \xi_t \tag{20}$$

Here, X_t is the position at moment t, and ξ_t is the random length this moment. Compared with the Stochastic random walk model, the ξ_t is a continuous random variable, which is samlped from a Gaussian distribution $W(\xi_t)$.

$$W(\xi_t) = \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \exp(-\frac{\xi_t^2}{2\sigma_\xi^2}) \tag{21}$$

For this distribution, the mean value is 0, and the variance is σ_ξ^2. Assume

Assume that starting from $X_0 = 0$, after t steps, the total displacement is the sum of all independent and identically distributed (i.i.d.) single-step displacements:

$$X_t = \sum_{t'=1}^{t} \xi_{t'} \tag{22}$$

We can directly use the linearity of expectation and variance to compute its statistical properties:

Mean movement: Since the mean displacement per step is zero, the mean total displacement is also zero.

$$\langle X_t \rangle = \langle \sum_{t'=1}^{t} \xi_{t'} \rangle = \sum_{t'=1}^{t} \langle \xi_{t'} \rangle = 0 \tag{23}$$

Variance: Since each step's displacement is independent, the variance of the total displacement equals the sum of the variances of each step. If each step has the same variance, denoted as σ_ξ^2 , then

$$\langle X_t^2 \rangle = \mathrm{Var}(X_t) = \langle (\sum_{t'=1}^{t} \xi_{t'})(\sum_{t''=1}^{t} \xi_{t''}) \rangle = \sum_{t'=1}^{t} \sum_{t''=1}^{t} \langle \xi_{t'} \xi_{t''} \rangle \tag{24}$$

Since $\langle \xi_{t'} \xi_{t''} \rangle = \delta_{t't''} \sigma_\xi^2$ (only when $t' = t''$), so

$$\langle X_t^2 \rangle = \sum_{t'=1}^{t} \langle \xi_{t'}^2 \rangle = \sum_{t'=1}^{t} \langle \sigma_\xi^2 \rangle = t\sigma_\xi^2 \tag{25}$$

This once again confirms the diffusive characteristic that the variance grows linearly with time.

2.2 Characteristic Function: A Powerful Tool for Handling Sums of Random Variables

Although directly calculating the mean and variance is straightforward, obtaining the complete probability distribution function is much more complicated. For problems involving the summation of independent random variables, the characteristic function is an extremely powerful mathematical tool.

2.2.1 Definition and Role

The characteristic function of a random variable X is defined as the expectation of its exponential form $\exp(isX)$:

$$\chi(X(s)) = \langle \exp(isX) \rangle = \int_{-\infty}^{\infty} \exp(isX) p(x) dx \tag{26}$$

Here, i is the imaginary unit, and s is a real variable. From the definition, it is clear that the characteristic function is essentially the Fourier transform of the probability density function (PDF).

Uniqueness: A probability distribution is uniquely determined by its characteristic function. If two random variables have the same characteristic function, their probability distributions must also be identical.

Summation Property: This is its most fundamental advantage. For two independent random variables X and Y, the characteristic function of their sum equals the product of their individual characteristic functions:

$$\chi_{X+Y}(k) = \chi_X(k) \cdot \chi_Y(k). \tag{27}$$

This property transforms a complex convolution operation in real space into a simple multiplication in Fourier space, greatly simplifying the analytical process.

2.2.2 Derivation of the Characteristic Function for the Gaussian Random Walk (GRW)

Now, we use the characteristic function to derive the complete probability distribution of X_t

Single-step characteristic function: First, we compute the characteristic function $\chi_{\xi_t}(s)$ of a single Gaussian displacement ξ_t.

$$\chi_{\xi_t}(s) = \langle \exp(is\xi_t) \rangle = \int_{-\infty}^{\infty} \exp(is\xi_t) \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \exp(-\frac{\xi^2}{2\sigma_\xi^2}) d\xi = \exp(-\frac{1}{2}s^2\sigma_\xi^2) \tag{28}$$

This is a standard result of Gaussian integration, showing that the Fourier transform of a Gaussian distribution is still Gaussian in form.

t-step characteristic function: Since $X_t = \sum_{t'=1}^{t} \xi_{t'}$ is the sum of independent random variables, its characteristic function is the product of the characteristic functions of all single steps.

$$\chi_{\xi_t}(s) = \langle \exp(is \sum_{t'=1}^{t} \xi_{t'}) \rangle = \langle \prod_{t'=1}^{t} \exp(is\xi_t) \rangle \tag{29}$$

Since the steps are independent, the expectation of the product equals the product of the expectations.

$$\chi_{\xi_t}(s) = \langle \prod_{t'=1}^{t} \exp(is\xi_t) \rangle = \prod_{t'=1}^{t} \langle \exp(is\xi_t) \rangle = \prod_{t'=1}^{t} \chi_{\xi_t}(s) \tag{30}$$

Substituting the single-step result, we obtain:

$$\chi_{\xi_t}(s) = \prod_{t'=1}^{t} \chi_{\xi_t}(s) = \prod_{t'=1}^{t} \exp(-\frac{1}{2}s^2\sigma_\xi^2) = \exp(-\frac{1}{2}s^2 \sum_{t'=1}^{t} \sigma_\xi^2) = \exp(-\frac{1}{2}s^2 t\sigma_\xi^2) \tag{31}$$

Obtaining the Probability Distribution Function: We compare the final result with the form of the single-step characteristic function. We found $\chi_{\xi_t}(s)$ is exactly a characteristic function corresponding to a Gaussian distribution whose mean value equals 0 and variance equals $t\sigma_\xi^2$. By performing the inverse Fourier transform, we can obtain its probability density function.

$$p(x,t) = \frac{1}{\sqrt{2\pi\sigma_{X_i}^2}} \exp(-\frac{x^2}{2\sigma_{X_i}^2}) \tag{32}$$

This derivation perfectly demonstrates the power of characteristic functions: a complex problem involving n-fold convolutions is transformed into simple exponential multiplication in Fourier space, ultimately yielding an analytical solution with ease.

To gain an intuitive understanding of the properties of the Gaussian Random Walk (GRW), we can perform simulations using Python code. The following code simulates multiple GRW trajectories and displays the particle position distributions at different times. The simulation results clearly demonstrate that as time increases, the distribution range (variance) of the particle positions continuously expands, while the distribution shape consistently remains Gaussian, in perfect agreement with our theoretical derivation.

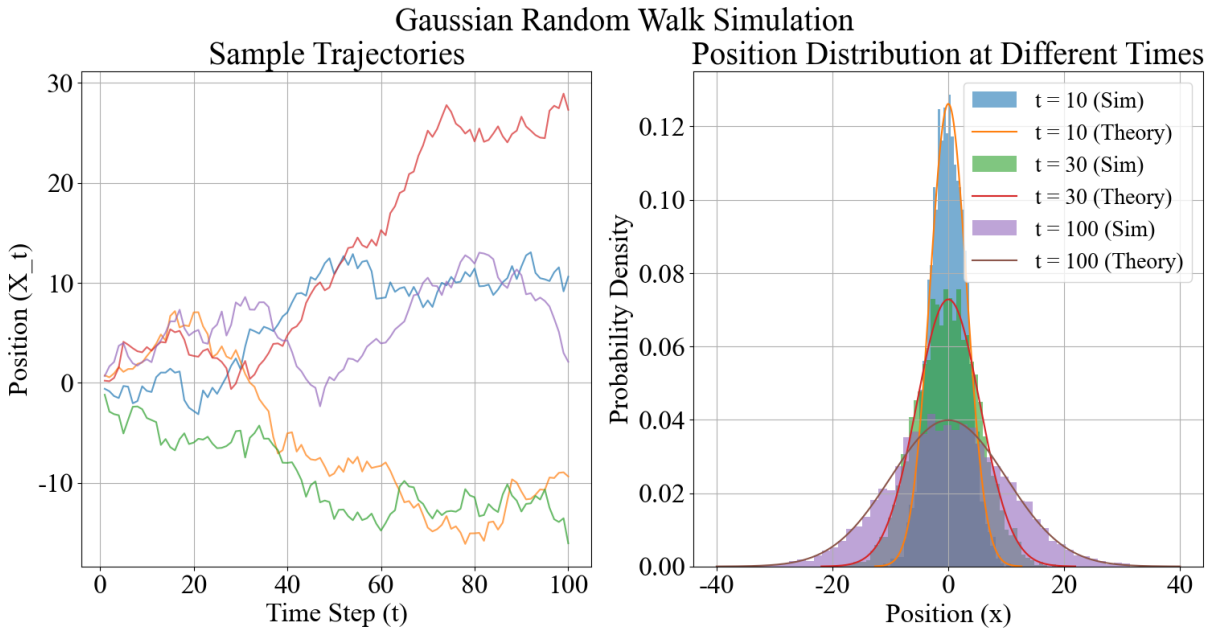

Figure 2:

2.3 Universality, Scaling, and the Concept of the Renormalization Group

First, it is important to clarify that the Central Limit Theorem we previously discussed—and the emergence of the Gaussian attractor—holds under one crucial condition: the variance of the single-step random variables must be finite. However, in the physical world, there exist many **heavy-tailed distributions** whose variance is infinite. For example, in certain glassy materials or financial markets, extreme events (very large fluctuations) occur with probabilities far higher than those predicted by a Gaussian distribution. For such systems, the limiting distribution of the sum of many random variables is no longer Gaussian.

This naturally raises a question: Is there a more universal theoretical framework that can describe all these cases in a unified way?

2.3.1 The Renormalization Group Perspective

The renormalization group (RG) theory was originally developed to address divergence problems in quantum field theory, but its core ideas have since permeated nearly every branch of physics, especially the theory of phase transitions in statistical mechanics.

Core Idea of RG: Coarse-graining and Scale Transformation. At its heart, RG is a systematic method of *coarse-graining*. Imagine a magnetic system composed of microscopic spins. We can group together a small block of spins (for example, a 2×2 cluster) and replace them with a single "effective" block spin that captures their average behavior. In doing so, we ignore microscopic details and obtain a new system with fewer degrees of freedom but defined at a larger scale. Repeating this procedure again and again—zooming out while adjusting parameters to preserve the form of the theory—constitutes an RG transformation.

Fixed Points and Universality As RG transformations are iterated, the system's parameters "flow" along a trajectory in parameter space. The eventual endpoint of this flow is called a *fixed point*. Systems at a fixed point exhibit *scale invariance*: they look self-similar under changes of scale. A fixed point, along with all initial systems whose flows converge to it, defines a *universality class*. Systems within the same universality class may have vastly different microscopic details, yet they share identical macroscopic critical behavior, characterized by the same set of critical exponents.

2.3.2 Applying the Idea of RG to Random Walks

Summation of Random Variables as Coarse-Graining: Adding n independent and identically distributed (i.i.d.) random variables ξ_i to obtain a new random variable $S_n = \sum_{i=1}^{n} \xi_i$ can itself be regarded as a coarse-graining operation in RG. Starting from n microscopic variables, we obtain a macroscopic variable that describes behavior at a larger scale S_n. After RG coarse-graining, a scaling transformation is usually needed to compare the new system with the original one. In the random walk problem, this corresponds to normalizing the total displacement S_n, for example by dividing by a factor $\sqrt{n}$ (in the case of finite variance). We find that after this transformation $S_n/\sqrt{n}$, the form of the distribution becomes stable (Gaussian), which reflects the self-similarity of the system under the transformation.

Stable distributions as fixed points: What we seek is the limiting distribution under the summation operation. A distribution is called *stable* if, after summation (with appropriate translation and scaling), its form remains unchanged. In the language of RG, these stable distributions are precisely the fixed points of the RG transformation defined by summation.

RG interpretation of the Central Limit Theorem (CLT): We can now reinterpret the CLT: for all initial distributions of random variables with finite variance (defining a large "basin of attraction"), under the RG flow induced by summation (coarse-graining), they are all attracted to the same—and unique—fixed point: the Gaussian distribution. Other details of the initial distribution, such as skewness or kurtosis, are considered *irrelevant operators* in RG language; they are gradually "smoothed out" through repeated coarse-graining and have no effect on the resulting macroscopic behavior.

Generalized Central Limit Theorem (GCLT): When the initial distribution has infinite variance (e.g., heavy-tailed distributions), the system is no longer within the Gaussian fixed point's basin of attraction. In this case, the RG flow drives it toward other fixed points. These non-Gaussian stable distributions are known as *Lévy stable distributions*. Unlike Gaussian distributions, Lévy distributions often have heavy tails, meaning the probability of extreme jumps is much higher than predicted by Gaussian statistics, and their variance may even be infinite.

2.4 Poisson Process: Introducing Random Time

So far, in the models we have discussed, time progresses in fixed, deterministic steps. However, in many natural processes, the occurrence of events is itself random. Examples include the decay of radioactive nuclei and the movement of molecular motors inside cells. The Poisson process is the fundamental model describing such events that are randomly distributed in time.

2.4.1 Motivation: From "Clock" Steps to Random Events

In the SRW and GRW models, the particle behaves like a wound-up clock, moving one step at each fixed time interval dt. Now, we consider a new scenario: the particle has a fixed probability of jumping forward in any infinitesimal time interval. This probability of a jump per unit time is called the hopping rate, denoted by v

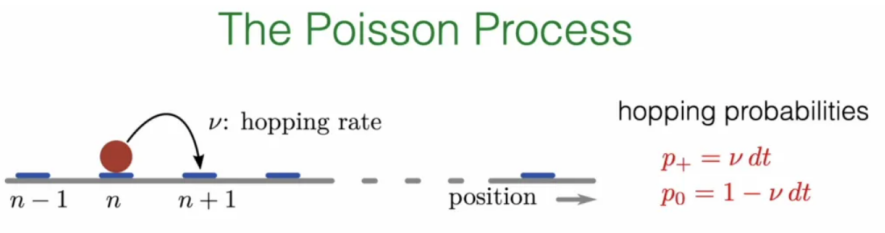

Figure 3:

- The probability of jumping forward within a time interval dt: $P_+ = vdt$
- The probability of remaining stationary within a time interval dt: $P_0 = 1 - vdt$

2.4.2 Deriving the Waiting Time Distribution

The first core question of the Poisson process is: after one event occurs, how long must one wait for the next event to happen? We introduce a key quantity: the survival probability $S(t) = \text{Prob } T > t$. It represents the probability that the event of interest (for example, a jump) has not occurred up to time t. Here, T is the waiting time random variable. By definition, the survival function is related to the cumulative distribution function (CDF) $S(t) = \text{Prob } T < t$ as follows: $S(t) = 1 - F(t)$.

For the system to "survive" until time $t + dt$, two independent conditions must be satisfied:

- It has survived up to time t (with probability $S(t)$).
- And, during the subsequent infinitesimal time interval dt, no jump occurs (with probability $1 - vdt$).

Thus

$$S(t + dt) = S(t) \cdot (1 - rdt) \tag{33}$$

Rewriting the above expression:

$$\frac{S(t + dt) - S(t)}{dt} = -vS(t) \tag{34}$$

When $dt \to 0$, the left term is the derivative of $S(t)$,giving us a first-order ordinary differential equation:

$$\frac{dS(t)}{dt} = -vS(t) \tag{35}$$

The solution to this differential equation is straightforward. With the initial condition $S(0) = 1$ (the event has certainly not occurred at time zero), solving the equation yields:$S(t) = e^{-vt}$.

We have obtained the survival probability function. The probability density function (PDF) of the waiting time can be obtained by differentiating with respect to t:$p(t) = -\frac{dS(t)}{dt} = ve^{-vt}$. This is the exponential distribution. It describes the distribution of waiting times between consecutive events when the event occurs at a constant rate.

Some important properties of the Poisson process can be derived from the exponential distribution:

Mean waiting time:$\langle T \rangle = \frac{1}{v}$. This agrees with intuition: the higher the rate, the shorter the average waiting time.

Variance of waiting time:

$$\text{Var}(t) = \frac{1}{v^2}.$$

Memoryless property: This is the most distinctive feature of the exponential distribution. It states that the probability of a future event occurring is independent of how long one has already waited. Mathematically,

$$P(T > s + t \mid T > t) = P(T > s).$$

In other words, if we have already waited t_0 seconds and the molecular motor has not moved, the probability that it remains stationary for another s seconds is the same as for a motor that has just started waiting. The system "forgets" its history. This property directly stems from our initial physical assumption that the hopping rate is a constant and does not change over time.

Another core question of the Poisson process is: "What is the probability that exactly k events occur within a given time window T?" The answer is given by the Poisson distribution:

$$P(k; \langle n \rangle) = \frac{\langle n \rangle^k e^{-\langle n \rangle}}{k!}, \tag{36}$$

where $\langle n \rangle = vT$ is the expected number of events in that time window. Together with the exponential distribution, this distribution provides a complete description of the Poisson process: the exponential distribution describes *how long to wait between events*, while the Poisson distribution describes *how many events occur within a given time interval*.

2.5 Gillespie Algorithm

2.5.1 Definition

The theoretical derivation in the previous section not only established the statistical properties of the Poisson process, but also provided us with an exact and efficient computational simulation method: the Gillespie algorithm (also known as the Stochastic Simulation Algorithm, SSA).

Traditional simulation methods may discretize time into many small intervals Δt, and then decide whether or not an event occurs within each interval according to the probability. This method is not only computationally inefficient, but also approximate.

The essence of the Gillespie algorithm is that it makes direct use of the analytical results we have already derived, providing precise answers to two questions:

When will the next event occur? Answer: The waiting time is sampled from the exponential distribution.

Which event will occur? (when multiple events are possible) Answer: The choice is made according to the relative rates of the events. For example, if a particle can jump forward at rate λ_f or backward at rate λ_b, then the total rate is $\lambda = \lambda_f + \lambda_b$. Once the time of the next event is determined, the probability of a forward jump is λ_f/λ, while the probability of a backward jump is λ_b/λ.

In this way, the Gillespie algorithm can exactly simulate arbitrarily complex stochastic reaction networks.

2.5.2 Core Idea: Inverse Transform Sampling

The core of the Gillespie algorithm is how to generate a random number that satisfies a specific probability distribution (Here, the distribution is $w(t) = v\exp(-vt)$). Here, we need to rely on a powerful statistical technology "Inverse Transform Sampling".

Inverse Transform Sampling is a method for generating random numbers from any probability distribution. In simple terms, suppose X is a continuous random variable with probability density function $f(x)$ and cumulative distribution function $F(x)$. To generate random samples that follow the distribution of X, one can first generate a uniform random variable $U \sim \text{Uniform}(0, 1)$, and then apply the inverse of the cumulative distribution function: $X = F^{-1}(U)$. This yields random samples that follow the desired distribution.

Its core idea relies on a key property: **the cumulative distribution function (CDF) of any random variable follows a uniform distribution on the interval** $[0, 1]$**.**

Therefore, we can reverse the process:

1. **Generate a uniform random number:** Draw a random number, U, from a uniform distribution on the interval $[0, 1]$.
2. **Apply the inverse function:** Plug this random number U into the **inverse function** ($F^{-1}(U)$) of the target probability distribution's CDF.
3. **Get the result:** The resulting value, $X = F^{-1}(U)$, is a random sample drawn from your desired probability distribution.

This method is effective because it establishes a one-to-one mapping from a uniform distribution to your target distribution. Let's use the Exponential Distribution as an example, which has a CDF of $F(x) = 1 - e^{-\lambda x}$.

1. **Find the inverse function:**
   - Let $U = F(X)$, so $U = 1 - e^{-\lambda X}$.
   - Solve for X: $1 - U = e^{-\lambda X}$.
   - Take the natural logarithm of both sides: $\ln(1 - U) = -\lambda X$.
   - The inverse function is: $X = -\frac{1}{\lambda}\ln(1 - U)$.
2. **Generate samples:** Now, you can generate a random number U on $[0, 1]$. By plugging this U into the formula $X = -\frac{1}{\lambda}\ln(1 - U)$, you get a random number that follows an exponential distribution.

- **Pros:**
  - Theoretically, you can generate samples from **any** distribution whose inverse CDF is known.
  - It is an **exact** sampling method, meaning it doesn't discard any samples.
  - It's relatively simple to implement, especially when the inverse function is easy to compute.
- **Cons:**
  - The biggest challenge is that the **inverse of the CDF (F^{-1}) can be difficult or impossible to find in a closed-form analytical expression**. For example, the normal distribution's CDF has no simple analytical inverse, so inverse transform sampling is typically not used for generating normal random numbers in practice.

In short, Inverse Transform Sampling is an elegant and powerful theoretical method that serves as a highly efficient tool when you can easily compute the inverse of your target distribution's cumulative distribution function.

For the simple one-way hopping process under discussion here, the second step is trivial (the particle can only hop forward). Therefore, the core of the algorithm is to repeatedly generate the next waiting time from the exponential distribution.

Based on the principle above, the Gillespie algorithm for simulating a Poisson process can be clearly summarized in the following three-step loop:

- Set the initial time to $T = 0$ and the initial position (number of events that have occurred) to $n = 0$
- Generate a random number ξ on $[0, 1]$. According to the equation $t = -\frac{1}{v}\ln(\xi)$ to calculate to waiting time of the next event.
- Advance the current time: $T = T + \Delta t$. The event count increases by one: $n = n + 1$. Repeat the steps until the simulated time t reaches the preset stopping time $T_{\max}$.

This event-driven simulation method, where the time step Δt is itself random, perfectly reproduces the inherent uncertainty of the stochastic process. It is both accurate and efficient, as the computer only performs a calculation when an event actually occurs.

The following Python code implements the Gillespie algorithm to simulate a simple one-way Poisson process

As we can see, no two trajectories are exactly the same: the jump times differ from one trajectory to another, and their final positions at a fixed time $t = 10s$ are also different. However, although each individual process is inherently random, all trajectories fluctuate around the black theoretical mean curve $N(t) = 2 \cdot t$. In other words, while the evolution of a single trajectory is unpredictable, the ensemble behavior strictly follows the average trend predicted by theory.

Bottom left panel (event count distribution): This figure reveals the deterministic law from a statistical perspective. We collect the final positions of thousands of independent simulations at time $t = 10s$ and display them as a histogram. The results show that, although the outcome of a single run is random, the distribution of many runs exhibits a clear and well-defined bell shape. More importantly, the simulated distribution (blue bars) coincides precisely with the theoretical Poisson distribution (red dotted line). This demonstrates that, at a fixed time point, the probability of the system being in a particular state can be predicted exactly.

Bottom right panel (waiting time distribution): This figure provides insight into the microscopic dynamics of the process. We extract the time intervals between consecutive jumps across all trajectories (that is, the horizontal segments in the step plots) and plot their distribution. The results show that short waiting times are most common, whereas long waiting times are very rare, with the probability decaying exponentially. The simulated distribution (blue) again matches the theoretical exponential distribution (red curve). This confirms that the very "engine" driving the evolution of the process—the waiting time between events—is itself a random variable that obeys a strict probabilistic law.

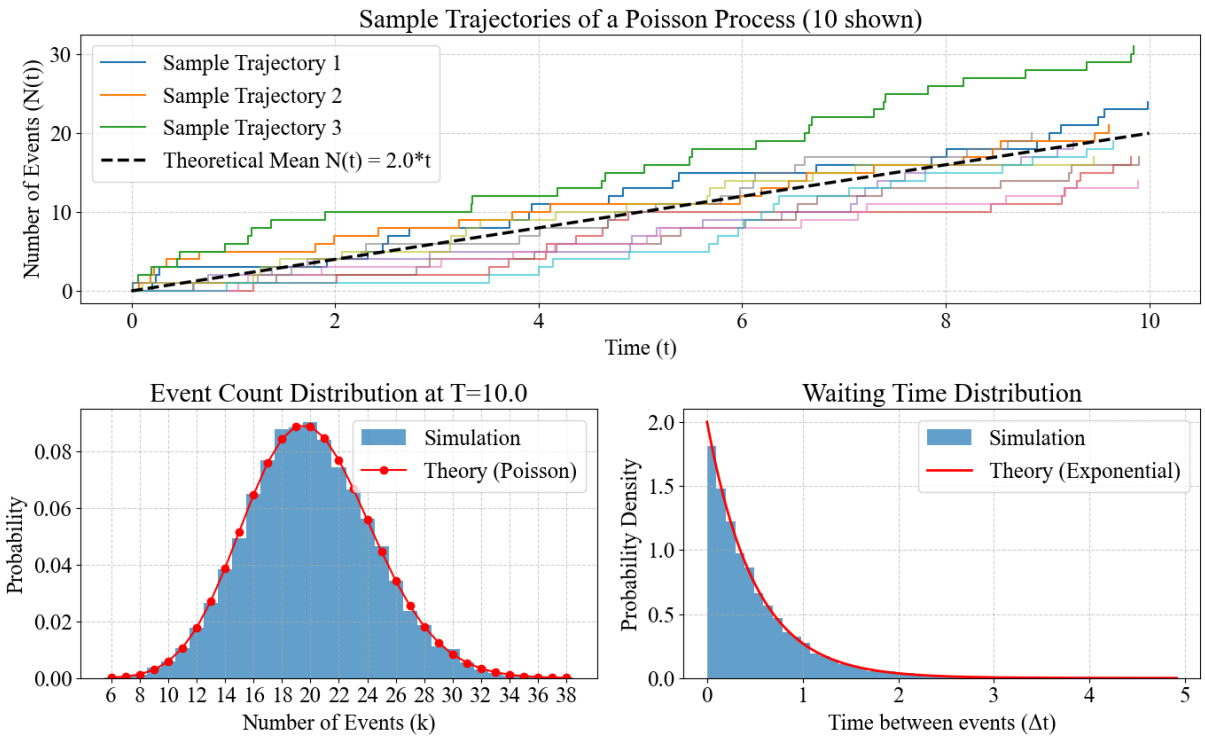

Figure 4:

2.6 Summary

This lecture has guided us through two fundamental transitions from discrete to continuous stochastic processes, thereby deepening our understanding of randomness:

From discrete space to continuous space: We observed that a simple random walk approaches a Gaussian distribution in the macroscopic limit. This convergence can not only be explained by the Central Limit Theorem, but also interpreted through the profound perspective of the renormalization group, which identifies the Gaussian distribution as a universal fixed point under summation. In this context, the characteristic function proved to be a powerful analytical tool for handling problems involving independent random variables.

From discrete time to continuous time: By introducing a constant event rate, we constructed the Poisson process model. Its essence lies in the fact that the waiting time between two events follows an exponential distribution, characterized by its unique memoryless property. The analysis of survival probability—formulated and solved through a differential equation—played a key role in deriving this result.

The overarching theme throughout this lecture is how simple, microscopic physical assumptions—such as single-step jump distributions or constant event rates—can be systematically elevated to universal, macroscopic statistical laws.

A natural next step is to unify these two ideas: studying a process that evolves in continuous space while its events occur at continuous random times. This leads us directly into the domain of Brownian motion, together with the mathematical frameworks that describe it—the Langevin equation and the Fokker–Planck equation—which will form the core topics of the subsequent lectures.

References