

Statistical Mechanics deals with many-body physical problems, and its main roots derive from thermodynamics.

There are three main concepts in mechanical statistics: *microstates*, *macrostates* and *statistical ensembles*

The microstate is defined by the coordinates and momenta of the system (in classical mechanics) and by the wavefunction of the system (in quantum mechanics). The collection of all microstate is referred as statistical ensemble.

§ 0.1 Probability Theory

Before diving deeply into statistical mechanics we need to grasp a basis of probability theory.

Definition 1 (Random Variable). *A random variable is a quantity X which takes a value x depending upon the elements $e \in E$ called events.*

In each observation X is uncertain, and only the probability of the occurrence of a given result x_i .

Definition 2 (Probability Density Function). *Let X be a random variable, we define the probability density function $w(x)$ as the probability that X assumes a value x in a certain interval of values given by the following integral*

$$\int_a^b w(x) \, dx$$

By definition $w(x)$ is normalized.

$$\int_{-\infty}^{\infty} w(x) \, dx = 1$$

Definition 3 (Mean Value). *Let $w(x)$ be the probability density of a random variable X , the expectation or mean value is defined as the integral*

$$\mathbb{E}(X) = \langle X \rangle = \int_{\mathbb{R}} xw(x) \, dx$$

Defining a random function as the function of a random variable $F(X)$, one can also define

$$\langle F(X) \rangle = \int_{\mathbb{R}} F(x)w(x) \, dx$$

Definition 4 (Moment). *We define the n -th moment of a random value as follows*

$$\mu_n = \langle X^n \rangle$$

Definition 5 (Characteristic Function). *The characteristic function of a random variable is defined as the Fourier transform of the probability density function*

$$\chi(k) = \hat{\mathcal{F}}[w](k) = \frac{1}{2\pi} \int_{\mathbb{R}} w(x)e^{-ikx} \, dx = \langle e^{-ikX} \rangle$$

If all moments are well defined, we can write the characteristic function as a power series of moments

$$\chi(k) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle X^n \rangle$$

For a multidimensional random function $F(\mathbf{X})$ which can take values f corresponding to a density function $w_F(f)$. $w_F(f)$ is defined as follows

$$w_F(f) = \langle \delta(F(\mathbf{X}) - f) \rangle$$

One important theorem that should be defined is the *central limit theorem*:

Theorem 1 (Central Limit). Let X_1, \dots, X_n be a set of independent random variables with probability distributions $w(x_i)$, suppose that there exist the n -th moment of every variable, and define a second random variable Y as follows

$$Y = \sum_{i=1}^n X_i$$

Then the probability density function of Y will be a Gaussian distribution, when $n \rightarrow \infty$

Proof. Let's define a third random variable Z as follows

$$Z = \sum_{i=1}^n \frac{1}{\sqrt{n}} (X_i - \langle X \rangle) = \frac{1}{\sqrt{n}} (Y - n \langle X \rangle)$$

Where by definition we have $\langle X_1 \rangle = \langle X_2 \rangle = \dots = \langle X_n \rangle$.

The probability density function of Z is then defined as follows

$$\begin{aligned} w_Z(z) &= \int \prod_{i=1}^n w(x_i) \delta \left(z - \frac{1}{\sqrt{n}} \left(\sum_{k=1}^n x_k \right) + \sqrt{n} \langle x \rangle \right) dx_i = \\ &= \frac{1}{2\pi} \int e^{ikz} dk \int \prod_{i=1}^n w(x_i) e^{\frac{-ik \sum_{k=1}^n x_i}{\sqrt{n}} + ik\sqrt{n}\langle X \rangle} dx_i = \\ &= \frac{1}{2\pi} \int e^{ikz + ik\sqrt{n}\langle X \rangle} \left(\chi \left(\frac{k}{\sqrt{n}} \right) \right)^N \end{aligned}$$

Where $\chi(q)$ is the characteristic function of $w(x_i)$. We can then reformulate the probability density by defining in an alternative way the characteristic function, as follows

$$\chi(q) = \exp \left(\sum_{n=1}^{\infty} \frac{(-iq)^n}{n!} C_n \right)$$

Where C_n are called the *cumulants*. They're related to the momenta of the variable, and it can be seen by comparing the Taylor series expansion of $\chi(q)$ with the previous

equation.

Reinserting it all into the definition of $w_Z(z)$, we get

$$w_Z(z) = \frac{1}{2\pi} \int e^{ikz - \frac{1}{2}k^2\mu_2^2 + \dots + k^3\sqrt{n} + \dots} dk$$

Neglecting the terms that vanish for $\mathcal{O}(\sqrt{N})$, we obtain the following

$$w_Z(z) = \frac{1}{\sqrt{2\pi\mu_2^2}} e^{-\frac{z^2}{2\mu_2^2}}$$

And since $w_Y(y) dy = w_Z dz$ we get, substituting the definition of z

$$w_Y(y) = \frac{1}{\sqrt{2\pi n\mu_2^2}} e^{-\frac{(y - n\langle X \rangle)^2}{2n\mu_2^2}}$$

Which is our searched Gaussian probability density for large n , which demonstrates the theorem. \square

§ 0.2 Quantum Statistics

For dealing with quantum statistics we need to define a deeper mathematical apparatus. We already know that in quantum mechanics, the expected value of an observable \hat{A} is defined as follows

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle$$

And the trace of this operator as follows

$$\text{Tr} \hat{A} = \sum_n \langle n | \hat{A} | n \rangle \quad (1)$$

From the previous two definitions, we can then define a new operator, called the *density operator*

$$\hat{\rho} = |\psi\rangle \langle \psi| \quad (2)$$

Which can be used to redefine the expectation value as follows

$$\langle A \rangle = \text{Tr} \hat{\rho} \hat{A} \quad (3)$$

If the system in question is all in one single state $|\psi\rangle$, we have what's called a *pure ensemble*, whereas if the system is in a mixture of states, each with a probability p_i , we are in what's called a *mixed ensemble* or a *statistical mixture*.

In the case of a statistical mixture of states $|\psi_i\rangle$, we then define our density operator as follows

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (4)$$

Another important feature we can define for the density matrix is its behavior in time evolution. For this we start by writing both the time dependent Schrödinger equation and its adjoint

$$\begin{cases} i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{\mathcal{H}} |\psi\rangle \\ -i\hbar \frac{\partial}{\partial t} \langle\psi| = \langle\psi| \hat{\mathcal{H}} \end{cases} \quad (5)$$

From the definition of $\hat{\rho}$ for statistical mixtures, inserting it into the Schrödinger equation we obtain

$$\begin{aligned} i\hbar \frac{\partial \hat{\rho}}{\partial t} &= i\hbar \sum_i p_i \left(\frac{\partial}{\partial t} |\psi_i\rangle \langle\psi_i| + |\psi_i\rangle \frac{\partial}{\partial t} \langle\psi_i| \right) = \\ &= \sum_i p_i \left(\hat{\mathcal{H}} |\psi_i\rangle \langle\psi_i| - |\psi_i\rangle \langle\psi_i| \hat{\mathcal{H}} \right) \end{aligned} \quad (6a)$$

From this, substituting again into it the definition of $\hat{\rho}$, we obtain the *Von Neumann equation*

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{\mathcal{H}}, \hat{\rho}] \quad (6b)$$

In Heisenberg representation, this equation appears easily remembering the definition of $\hat{\rho}$. Without delving into the calculations (which can be made directly using the definition itself of a time-evolved state), we get

$$\hat{\rho}(t) = \hat{\mathcal{U}}(t) \hat{\rho} \hat{\mathcal{U}}^\dagger(t) \quad (6c)$$