

# Topics in Computational Chemistry, CHEM\*7500/CHEM 750 Fall 2021

Week 2: Statistical Thermodynamics

---

Prof. Leanne D. Chen

September 22, 2021

Department of Chemistry



# Mathematics

---

# Probability Distributions

Let  $X$  be a variable which can assume the  $M$  discrete values  $X_1, X_2, \dots, X_M$  with corresponding probabilities  $P(X_1), P(X_2), \dots, P(X_M)$ . The variable  $X$  is said to be a discrete random variable, and  $P(X)$  is said to be a discrete distribution. The mean, or average, value of  $X$  is

$$\bar{X} = \frac{\sum_{j=1}^M X_j P(X_j)}{\sum_{j=1}^M P(X_j)} \quad (1)$$

Since  $P(X_j)$  is a probability,  $P(X_j)$  must be normalized. That is, the summation in the denominator in Equation 1 must equal unity. In other words, the sum of the probabilities of all possible outcomes is 1.

# Probability Distributions

Since we are dividing by 1, then we do not need to include the denominator. So the mean of any function of  $X$ ,  $f(X)$ , is given by

$$\overline{f(X)} = \sum_{j=1}^M f(X_j)p(X_j) \quad (2)$$

If  $f(X) = X^m$ ,  $\overline{f(X)}$  is called the  $m$ th moment of the distribution  $P(X)$ . If  $f(X) = (X - \bar{X})^m$ ,  $f(X)$  is called the  $m$ th central moment, that is, the  $m$  moment about the mean. In particular, the mean of  $(X - \bar{X})^2$  is called the variance, and is a measure of the spread of the distribution. The square root of the variance is the standard deviation.

# The Poisson Distribution

A very commonly occurring and useful discrete distribution is the Poisson distribution:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad k = 0, 1, 2, \dots \quad (3)$$

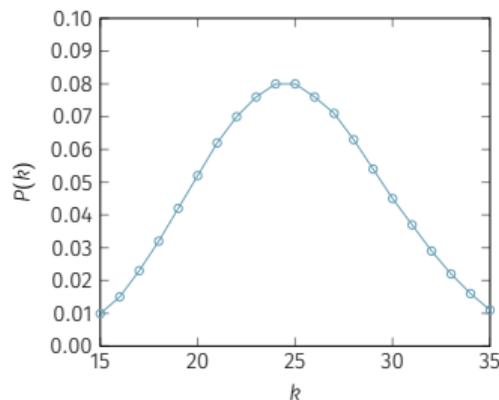
The Poisson distribution gives the probability of a given number of events  $k$  occurring in a fixed interval of time or space, if these events occur with a known constant mean rate  $\lambda$  and independently of the time since the last event.

In other words, the Poisson distribution describes events such as “if an event is expected to occur  $\lambda$  times on average, what is the probability that the event occurs exactly  $k$  times over some time interval?”

# The Poisson Distribution

**Example:** Suppose you are a prominent social media influencer who receives, on average, 600 “likes” per day, so that your rate of “likes” per hour is  $600/24 = 25$ . What is the probability that over the next hour, you will receive exactly 20 “likes”?

$$P(k = 20) = \frac{25^{20} e^{-25}}{20!} = 0.052 \quad (4)$$



# Probability Distributions

If the random variable  $x$  is continuous rather than discrete, then we interpret  $P(x)dx$  as the probability that the random variable  $x$  lies between the values of  $x$  and  $x + dx$ . The mean of any function of  $x$  is

$$\overline{f(x)} = \int f(x)P(x)dx \quad (5)$$

The limits of the integral are over the entire range of  $x$ . One of the most important continuous probability distributions is the Gaussian distribution,

$$P(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right) \quad -\infty \leq x \leq \infty \quad (6)$$

# The Gaussian Distribution

$$P(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right) \quad -\infty \leq x \leq \infty \quad (6)$$

- In Equation 6, the variance  $\sigma^2$  controls the width of the Gaussian distribution.
- The smaller the  $\sigma$ , the narrower the Gaussian distribution becomes.
- In the limit  $\sigma \rightarrow 0$ , Equation 6 becomes a delta function.

# Stirling's Approximation

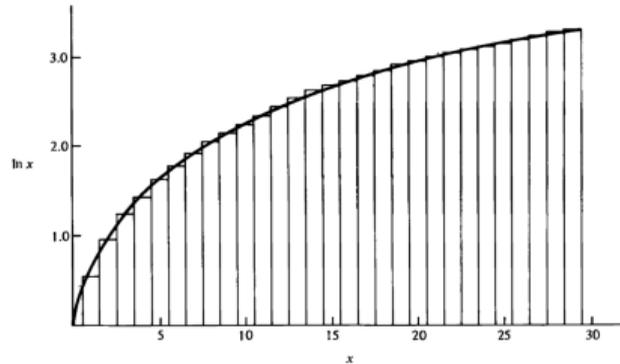
- In statistical thermodynamics, we often encounter factorials of very large numbers, such as Avogadro's number.
- The calculation and mathematical manipulation of factorials become awkward for large  $N$ .
- Therefore, it is desirable to find an approximation of  $N!$  for large  $N$ .
- Problems of this sort occur often in mathematics and are called asymptotic approximations.
- An asymptotic approximation of a function improves as the argument of that function increases.

# Stirling's Approximation

- Since  $N!$  is a product, it is more convenient to deal with  $\ln(N!)$  as this is a sum.
- The asymptotic approximation to  $\ln(N!)$  is called Stirling's approximation, which we will now derive.
- Since  $N! = N(N - 1)(N - 2) \dots (2)(1)$ ,  $\ln(N!)$  is

$$\ln(N!) = \sum_{m=1}^N \ln(m) \quad (7)$$

# Stirling's Approximation



The above figure shows  $\ln(x)$  plotted versus  $x$ . The sum of the areas under these rectangles up to  $N$  is  $\ln(N!)$ . It can also be seen in the figure that the continuous curve  $\ln(x)$  forms an envelope around the upper edges of the rectangles, which steadily becomes a smoother approximation to the rectangles as  $x$  increases.

# Stirling's Approximation

We can approximate the total area of the rectangles by the integral of  $\ln(x)$ . The area under  $\ln(x)$  will poorly approximate the rectangles only in the beginning when  $N$  is small. If  $N$  is large enough, the error will make a negligible contribution to the total area. Then, we may write

$$\ln(N!) = \sum_{m=1}^N \ln(m) \approx \int_1^N \ln(x) dx = N \ln(N) - N \quad (8)$$

which is Stirling's approximation to  $\ln(N!)$ . The lower limit could just as well have been taken as 0 in Equation 8, since  $N$  is large, and  $x \ln(x) \rightarrow 0$  as  $x \rightarrow 0$ .

# The Binomial Distribution

When we begin our discussion of the canonical ensemble later in this lecture, a common theme is to determine how many ways we can divide  $N$  distinguishable systems into groups such that there are  $n_1$  systems in the first group,  $n_2$  systems in the second group, and so on, so that  $n_1 + n_2 + \dots = N$ , that is, all the systems are accounted for. There are multiple ways to think about this problem. We will go through a couple of different ways in what follows.

# The Binomial Distribution

- Let's think about the colours of the rainbow, as in red, orange, yellow, green, blue, indigo, violet.
- How many different unique three-colour combinations can we pick out?
- To answer this, let's think in terms of three spots, and how many possibilities can go into each spot, once the previous one (if applicable) has been determined.

\_\_\_\_\_

# The Binomial Distribution

\_\_\_\_\_

- Because we have seven colours total, there are seven different possibilities for the first spot.
- Then, once the first spot has been taken, there are six different choices left, so the second spot has six different possibilities.
- Similarly, the last spot has five different possibilities once the first two spots have been determined.
- So the total number of arrangements is

$$7 \times 6 \times 5 = 210$$

# The Binomial Distribution

- We are not yet finished, because the 210 different arrangements count the same group of three colours but in different orders as unique groupings (e.g. **rgb** and **grb**), which is not what we are looking for.
- So to eliminate the effect of ordering in our groups, we have to further divide by a factor, and now the question is: how many different ways can three different objects be ordered?
- This is similar to the previous question, except now the total number of choices has changed; we are simply arranging three objects into three spots, which means the factor is

$$3 \times 2 \times 1 = 6$$

# The Binomial Distribution

- We have arrived at the answer (35) through considering the number of ways to choose three object from a total of seven, where the order does not matter.
- We need a general formula to describe this, so let us consider the problem in a slightly different way.
- Now, let's begin by calculating the number of permutations of  $N$  distinguishable projects, i.e., the number of possible ways to *order*  $N$  distinguishable objects.
- It is not difficult to see that this is simply

$$N(N - 1)(N - 2) \dots (2)(1) \equiv N!$$

# The Binomial Distribution

- Next, we calculate the number of ways of dividing  $N$  distinguishable objects into two groups, one group containing  $N_1$  objects, and the other containing the remaining  $N - N_1$  objects.
- There are  $N(N - 1) \dots (N - N_1 + 1)$  ways to form the first group, and  $N_2! = (N - N_1)!$  ways to form the second group.
- The total number is, then, the product

$$N(N - 1) \dots (N - N_1 + 1) \times (N - N_1)! = \frac{N!}{(N - N_1)!} \times (N - N_1)! = N!$$

# The Binomial Distribution

- But this has overcounted the situation drastically, since the order in which we place  $N_1$  members in the first group and  $N_2$  in the second group is immaterial to the problem as stated.
- All  $N_1!$  orders of the first group and  $N_2!$  orders of the second group correspond to just one division of  $N$  objects into  $N_1$  and  $N_2$  objects.
- Therefore, the desired result is

$$\frac{N!}{N_1!(N - N_1)!} = \frac{N!}{N_1!N_2!} \quad (9)$$

# The Binomial Distribution

Writing it in terms of the “choose” notation, we have

$$\binom{N}{N_1} = \frac{N!}{N_1!(N - N_1)!} \quad (10)$$

which is also known as the binomial coefficient as it occurs in the binomial expansion,

$$(x + y)^N = \sum_{N_1=0}^N \frac{N!}{N_1!(N - N_1)!} x^{N-N_1} y^{N_1} = \sum_{N_1 N_2}^* \frac{N!}{N_1! N_2!} x^{N_1} y^{N_2} \quad (11)$$

where the asterisk on the second summation signifies the restriction  $N_1 + N_2 = N$ .

# The Multinomial Distribution

- The multinomial distribution is the generalization of the binomial distribution, where instead of describing the probability distribution of an event with two outcomes (e.g. flipping a coin), we now describe the probability distribution of an event with multiple outcomes (e.g. rolling a die).
- The multinomial coefficient can be thought of as the number of ways to divide  $N$  objects into  $r$  groups, where the first group contains  $N_1$  objects, the second group contains  $N_2$  objects, and so on.

$$\frac{N!}{N_1!N_2!\dots N_r!} = \frac{N!}{\prod_{j=1}^r N_j!} \quad (12)$$

where  $N_1 + N_2 + \dots + N_r = N$ .

# The Multinomial Distribution

Equation 12 is the formula for the multinomial coefficient, which occurs in the multinomial distribution formula,

$$(x_1 + x_2 + \dots + x_r)^N = \sum_{N_1=0}^N \sum_{N_2=0}^N \dots \sum_{N_r=0}^N * \frac{N!}{\prod_{j=1}^r N_j!} x_1^{N_1} \dots x_r^{N_r} \quad (13)$$

Here, the asterisk signifies the restriction

$$N_1 + N_2 + \dots + N_r = N \quad (14)$$

# Pascal's Triangle

$n = 0$					1								
$n = 1$				1		1							
$n = 2$			1		2		1						
$n = 3$		1		3		3		1					
$n = 4$		1		4		6		4		1			
$n = 5$		1		5		10		10		5		1	
$n = 6$	1		6		15		20		15		6		1

# Binomial Distribution for Large Numbers

The maximum of the binomial coefficient for a given  $N$  can be approximated in the form of a Gaussian curve

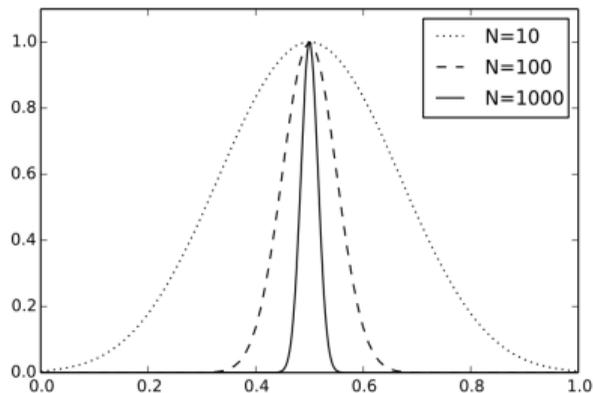
$$f(N_1) = f(N_1^*) \exp \left( -\frac{2(N_1 - N_1^*)^2}{N} \right) \quad (15)$$

Compare this with the standard form of the Gaussian function,

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left( -\frac{(x - x^*)^2}{2\sigma^2} \right) \quad (16)$$

So here  $N_1^*$  is the mean of the distribution (equal to  $N/2$ ), and  $N^{1/2}$  (multiplied by a small integer factor) is the standard deviation.

# Binomial Distribution for Large Numbers



Source: Peter Eastman

Thus, the binomial coefficient (and by extension, the multinomial coefficient) shows a very sharp peak for very large  $N$ . In these cases, we can approximate the overall distribution of  $N_j$ 's with a single value of  $N_j^* = N/2$ .

# Physical Chemistry

---

- Now we will begin to introduce the basic concepts and assumptions of statistical thermodynamics, and then apply them to a system with fixed values of  $V$ ,  $N$ , and is in thermal equilibrium with its environment.
- We will derive the fundamental connection between the quantum mechanical energy levels available to an  $N$ -body system and its thermodynamic functions.
- The link is provided by the partition function, which is a function of central importance in statistical thermodynamics.

# Ensemble Averages

- Our goal is to calculate thermodynamic properties in terms of molecular properties.
- Given the structure of the individual molecules of our system and the form of the intermolecular potential, we wish to be able to calculate thermodynamic properties, such as entropy and free energy.
- Consider some macroscopic system of interest, such as a litre of water or a salt solution
- From a macroscopic point of view, we can completely specify a system by a few parameters, e.g. the volume, concentration, density, temperature, etc.

# Ensemble Averages

- Regardless of the complexity of the system, it requires only a small number of parameters to describe it.
- From a microscopic point of view, on the other hand, there will be an enormous number of quantum states consistent with the fixed macroscopic properties.
- The degeneracy of an isolated  $N$ -body system is of the order  $10^N$  for all but the very lowest energies, so our macroscopic system could be in any one of  $10^N$  possible quantum states.
- It would be impossible for us to determine which of the  $10^N$  possible states the system is in, given the sheer size of  $N$ .

# Ensemble Averages

- However, we must be able to determine the state of the system in order to calculate a thermodynamic property such as pressure, since the values of that property in each of the possible quantum states would be different in general.
- How can we do this?
- We need to take the *average* of a property of interest, which is calculated from *all* quantum states consistent with a few macroscopic parameters that we have used to define the system.
- For example, we postulate that the average energy corresponds to the thermodynamic energy, and that the average pressure corresponds to the thermodynamic pressure.

# The Microcanonical Ensemble

- Let us begin considering an *ensemble* of systems, which is a concept first introduced by Gibbs.
- An ensemble is a **collection of a very large number of systems**, say  $\mathcal{A}$ , each constructed to be a replica on a macroscopic level of the particular system of interest.
- For example, suppose the system has a volume  $V$ , contains  $N$  molecules of a single component, and is known to have energy  $E$ .
- In other words, it is an *isolated* system with  $N$ ,  $V$ , and  $E$  fixed.

# The Microcanonical Ensemble

- Then, the entire ensemble would have a volume  $\mathcal{A}V$ , contain  $\mathcal{A}N$  molecules, and have a total energy  $\mathcal{E} = \mathcal{A}E$ .
- Each of the individual systems in this ensemble is a quantum mechanical system of  $N$  interacting atoms or molecules in a container of volume  $V$ .
- The values of  $N$  and  $V$ , along with the force law between the molecules, are sufficient to determine the energy eigenvalues  $E_j$  of the Schrödinger equation along with their associated degeneracies  $\Omega(E_j)$ .

# The Microcanonical Ensemble

- These energies are the only energies available to the  $N$ -body system.
- Thus, the fixed energy  $E$  must be one of these  $E_j$ 's, and, consequently, there is a degeneracy  $\Omega(E)$ .
- Said another way, there are  $\Omega(E)$  different quantum states consistent with the only things we know about our macroscopic system of interest, determined by  $N$ ,  $V$ , and  $E$ .
- Although all the systems in the ensemble are identical from a thermodynamic or macroscopic point of view, they are not necessarily identical on a molecular level!
- We need to think about the *distribution* of the members of the ensemble with respect to the  $\Omega(E)$  possible quantum states.

# The Microcanonical Ensemble

- We shall further restrict our ensemble to obey the *principle of equal a priori probabilities*.
- In other words, we require that each and every one of the  $\Omega(E)$  quantum states is represented an equal number of times in the ensemble.
- Since we have no information to consider any one of the  $\Omega(E)$  quantum states to be more important than any other, we must treat each of them equally, i.e., we must invoke the principle of equal a priori probabilities.
- All of the  $\Omega(E)$  quantum states are consistent with the given values of  $N$ ,  $V$ , and  $E$ , the only information we have about the system.

# The Microcanonical Ensemble

- Clearly, the number of systems in the ensemble must be an integral multiple of  $\Omega(E)$  (since we have to represent each one of these states at least once, or twice/thrice etc.).
- The number of systems in an ensemble is a very large number and can be made arbitrarily large by simply doubling, tripling... etc. the size of the ensemble.
- An alternate statement of the principle of equal a priori probabilities is that **an isolated system, where  $N$ ,  $V$ , and  $E$  are all fixed, is equally likely to be in any of its  $\Omega(E)$  possible quantum states.**

# The Microcanonical Ensemble

- To summarize, we have begun our discussion of an ensemble, with the goal of obtaining macroscopic properties, which can be calculated from the *averages* of a specific property (e.g. pressure) in all systems of an ensemble.
- For the microcanonical ensemble, the energy of each system in an ensemble is fixed to a single value of  $E$ , which is helpful for a theoretical discussion but not as useful in practice.
- The most commonly used ensemble in statistical thermodynamics is the *canonical ensemble*, where the individual systems have  $N$ ,  $V$ , and  $T$  (instead of  $E$ ) fixed.
- We will continue our discussion with the canonical ensemble.

# The Canonical Ensemble

- Consider an experimental system with  $N, V, T$  as its independent thermodynamic variables.
  - We can construct an ensemble of such systems in the following way.
1. Each system is enclosed in a container of volume  $V$ , with walls that are heat conducting but impermeable to the passage of molecules.

# The Canonical Ensemble

2. The entire ensemble of systems is then placed in a very large heat bath at temperature  $T$ .
3. When equilibrium is reached, the entire ensemble is at a uniform temperature  $T$ , and each system has the same temperature  $T$ .
4. Since the entire ensemble is surrounded by thermal insulation, the ensemble itself is an isolated system with volume  $\mathcal{A}V$ , number of molecules  $\mathcal{A}N$ , and some total energy  $\mathcal{A}E$ —not the same for each system in the ensemble!

# The Canonical Ensemble

- The number of states for the canonical ensemble becomes more tricky to count, since the energy of each system is not fixed to some value  $E$ .
- We will need to consider the entire spectrum of energy states for each member of the canonical ensemble.
- Let the energy eigenvalues of the quantum states of a system be  $E_1(N, V)$ ,  $E_2(N, V)$ ..., ordered such that  $E_{j+1} \geq E_j$ .
- It is very important to note that any particular energy, say  $E_i$ , is *repeated according to its degeneracy*, i.e.,  $E_i$  occurs  $\Omega(E_i)$  times.
- Any particular system might be found in any of these quantum states.

# The Canonical Ensemble

We can specify a state of the entire ensemble by saying that  $a_1, a_2, a_3, \dots$  of the systems are in states  $1, 2, 3, \dots$ , respectively, with energies  $E_1, E_2, E_3, \dots$

State Number	1	2	3	...	$l$	...
Energy	$E_1$	$E_2$	$E_3$	...	$E_l$	...
Occupation Number	$a_1$	$a_2$	$a_3$	...	$a_l$	...

Here, occupation number means the number of systems of the ensemble in that particular state. The set of occupations numbers is called a distribution.

# The Canonical Ensemble

The occupation numbers must satisfy the two conditions of the canonical ensemble:

$$\sum_j a_j = \mathcal{A} \quad (17)$$

and

$$\sum_j a_j E_j = \mathcal{E} \quad (18)$$

The first condition simply accounts for all the members of the ensemble, and the second condition enforces a total fixed energy  $\mathcal{E}$ ; in other words, only distributions for which the total energy is  $\mathcal{E}$  are allowed.

# The Canonical Ensemble

- Since the canonical ensemble has been isolated from its surroundings by thermal insulation, we can apply the principle of equal a priori probabilities to this isolated system.
- In the form that we wish to use here, the principle of equal a priori probabilities states that every possible state of the canonical ensemble, that is, every distribution of occupation numbers  $a_1, a_2, \dots$  consistent with the two constraints is *equally probable* and must be given equal weight in performing ensemble averages.

# The Canonical Ensemble

- The number of ways  $W(\mathbf{a}) \equiv W(a_1, a_2, a_3, \dots)$  that any particular distribution of the  $a_j$ 's can be realized is the number of ways that  $\mathcal{A}$  distinguishable objects can be arranged into groups, such that  $a_1$  are in the first group,  $a_2$  in the second, and so on:

$$W(\mathbf{a}) = \frac{\mathcal{A}!}{a_1!a_2!a_3!\dots} = \frac{\mathcal{A}!}{\prod_k a_k!} \quad (19)$$

- Note that we have denoted the set of occupation numbers  $\{a_j\}$  with the symbol  $\mathbf{a}$ .

# The Canonical Ensemble

- In general, there are very many allowed distributions, i.e. distributions that follow the constraints given by Equations 17 and 18.
- In any particular distribution,  $a_j/\mathcal{A}$  is the fraction of systems or members of the canonical ensemble in the  $j$ th energy state, with energy  $E_j$ .
- The overall probability  $P_j$  that a system is in the  $j$ th quantum state is obtained by averaging  $a_j/\mathcal{A}$  over all the allowed distributions, giving equal weight to each one according to the principle of a priori probabilities.
- Thus  $P_j$  is given by

$$P_j = \frac{\bar{a}_j}{\mathcal{A}} = \frac{1}{\mathcal{A}} \frac{\sum_{\mathbf{a}} W(\mathbf{a}) a_j(\mathbf{a})}{\sum_{\mathbf{a}} W(\mathbf{a})} \quad (20)$$

# The Canonical Ensemble

- In Equation 20, the notation  $a_j(\mathbf{a})$  signifies that the value of  $a_j$  depends upon the distribution, and summations are over all distributions that satisfy Equations 17 and 18
- Given the probability that a system with fixed values of  $N$ ,  $V$ , and  $T$ , is in the  $j$ th quantum state, one can calculate the canonical ensemble average of any mechanical property from

$$\bar{M} = \sum_j M_j P_j \quad (21)$$

# Method of the Most Probable Distribution

- It is very difficult to perform the summations in Equation 20.
- Thus, we will use the property of a multinomial distribution to become very narrow when  $N$  becomes very large, and replace the summation with a single set of  $a_j$ s that maximizes  $W(\mathbf{a})$ .

$$P_j = \frac{1}{\mathcal{A}} \frac{\sum_{\mathbf{a}} W(\mathbf{a}) a_j(\mathbf{a})}{\sum_{\mathbf{a}} W(\mathbf{a})} = \frac{1}{\mathcal{A}} \frac{W(\mathbf{a}^*) a_j^*}{W(\mathbf{a}^*)} = \frac{a_j^*}{\mathcal{A}} \quad \lim a_j \rightarrow \infty \quad (22)$$

- So we have

$$P_j = \frac{\bar{a}_j}{\mathcal{A}} = \frac{a_j^*}{\mathcal{A}} \quad (23)$$

# The Partition Function

- After replacing the summation over all possible distribution of  $a_j$ 's with a single distribution of  $a_j^*$  that maximizes  $W(\mathbf{a})$ , we can now obtain the form of the *partition function*, which allows us to describe the statistical properties of a thermodynamic system.

$$Q(N, V, T) = \sum_j e^{-E_j(N, V)/k_B T} \quad (24)$$

- This function  $Q(N, V, T)$  is the central function of the canonical ensemble, and allows us to calculate different macroscopic properties.

# The Partition Function

- For example, if we would like to calculate the average energy of an ensemble,

$$\bar{E}(N, V, T) = \frac{\sum_j E_j(N, V) e^{-E_j(N, V)/k_B T}}{\sum_j e^{-E_j(N, V)/k_B T}} = \frac{\sum_j E_j(N, V) e^{-E_j(N, V)/k_B T}}{Q} \quad (25)$$

- Or, to obtain the average pressure, where pressure is equal to the negative of the partial derivative of the energy with respect to the volume,

$$\bar{p} = \sum_j p_j P_j = - \frac{\sum_j \left( \frac{\partial E_j}{\partial V} \right) e^{-E_j/k_B T}}{\sum_j e^{-E_j/k_B T}} \quad (26)$$

- In summary, we are able to connect the average quantities, which can be obtained by considering an ensemble of microscopic systems, to macroscopic thermodynamic quantities.

$$p \leftrightarrow \bar{p}$$

and

$$E \leftrightarrow \bar{E}$$

- These two lines represent the Gibbs postulate: *the ensemble average of a property corresponds to the thermodynamic quantity.*

# Thank you!

Please feel free to unmute yourself or raise your hand to ask a question.