

BOLTZMANN'S LAST SYMPHONY
THE THREE ENSEMBLES AND THEIR CONDUCTORS
PHY 301 – THERMAL PHYSICS
RAVIN RAJ

Contents

1	Ensembles	2
2	The Microcanonical Ensemble (MCE)	5
3	The Canonical Ensemble (CE)	10
4	The Grand Canonical Ensemble (GCE)	17
5	The Continuous Ensembles	27
	References	41

Preface

Please note that this set of notes is **NOT** meant to be a substitute for any course material, and is instead complementary to it at best. This content is not endorsed by Bill or the other AIs in any way.

The following review is mostly a transcription of some handwritten notes I kept as an undergraduate when I took my first and second statistical mechanics courses. It is hard for me to place exact citations throughout this, as I do not include those in my own notes. As far as possible, I have tried to compile a list of the references [1–8] which I used when writing these and taking my statistical mechanics course in the past. Specifically, my primary reference for thermodynamics is Callen [1], while I tend to lean on Reif and Pathria [2, 5] for statistical mechanics. The professor who taught the second statistical mechanics course which I took as an undergraduate has also published his notes as a full book now [4]. I also highly recommend reading Gibbs' original treatise [3], to better appreciate the decades-long process of refining statistical mechanics that we take for granted today. This document may also be updated periodically, and the latest version will be put on Canvas as well as my [personal webpage](#).

My first (and second and third) statistical mechanics course(s) brought me plenty of joy, in that it/they really did change the way I viewed physical systems to the extent that I came out of it understanding far more than I ever did with just the tools of *vanilla* classical mechanics, quantum mechanics and electrodynamics alone. My aim with this review is to (hopefully) pass some of that excitement on to you.

1 Ensembles

1.1 Microstates and Macrostates

In thermodynamics, one typically analyses the behaviour of a large system by considering its macroscopic properties, such as temperature T and pressure p . In statistical mechanics, we instead consider the microscopic constituents of the large system and study their properties (such as positions $\{\mathbf{q}_i\}$ or momenta $\{\mathbf{p}_i\}$ of the molecules in a gas), hoping to reconcile the statistical averages or fluctuations of these microscopic properties with the measurable macroscopic quantities. For a system of N particles, we expect it to be completely characterised by $6N$ unique values (corresponding to the $3N$ position coordinates and $3N$ conjugate momentum coordinates, up to the Heisenberg uncertainty principle). Thus, the microscopic state of a system, known as a **microstate**, can be fully specified by simply storing a list of $6N$ numbers¹. Such a set of coordinates is known as a **phase space coordinate**, and the \mathbb{R}^{6N} space to which it belongs is known as **phase space**.

On the other hand, the thermodynamic properties of the system as a whole are typically characterised by very few numbers, known as the **macrostate** of the system. For instance, an ideal gas of 10^{23} molecules can be described by three numbers: p, V and T . Naïvely, one might think that we just lost $(10^{23} - 3)$ degrees of freedom in this characterisation, but that is not true. Consider a single microstate α of this gas, and now flip the signs of all the momenta to obtain some new microstate α' . It should be obvious that the values of $\{p, V, T\}$ were left unchanged by this transformation, so we have just shown that multiple microstates can correspond to a single macrostate².

1.2 The Ergodic Hypothesis

We consider a single system (collection of particles subject to some Hamiltonian governing its evolution) to be the object of interest, which is itself comprised of microscopic elements that give us a set of accessible microstates. The evolution of this system over time is determined entirely by the Hamiltonian, specifically with the evolution of each coordinate given by Hamilton's equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{1.1}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1.2}$$

This will also impose various conservation laws depending on its symmetries. For instance, if the Hamiltonian exhibits time-translational symmetry ($\frac{\partial H}{\partial t} = 0$), then the energy of the system must be conserved over time. Over the course of its evolution, we may be interested in specific properties of the system which vary for the different microstates, and hence macrostates available to us. To determine their expectation values, we need to somehow average these properties over the course of the system's evolution.

This is where we can make a powerful assumption, one which fails us only in a small (but important) class of problems³. Typically, we will have a system evolving in phase space over time, moving on a surface of constant E while accessing a large set of microstates over time, with a path denoted by $\alpha(E, t)$ as shown

¹Uh yeah good luck. For a thermodynamically large system where $N \sim 10^{23}$, storing these numbers with machine precision at the picometer (10^{-12} pm) scale would require the same amount of information as the current size of the Internet.

²Another example would be to take one microstate, pluck a single molecule from one spot and place it randomly in a different spot. Intuition tells us that this will not affect any of the macroscopic properties of the system.

³See the Fermi–Pasta–Ulam–Tsingou problem. Interestingly, this is one of the first examples of a many-body physics problem solved on a computer, and it was done by Mary Tsingou (a pioneer programmer on the MANIAC computer at Los Alamos). Her contributions went unrecognised for decades until 2020 when Los Alamos first publicly acknowledged her role in the project.

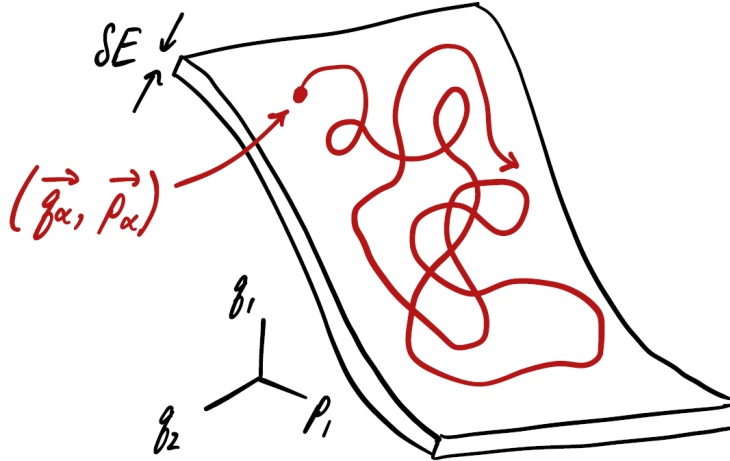


FIGURE 1: Hamiltonian Evolution on a Hypersurface. A system in some microstate α will evolve according to Hamilton's equations, but is typically constrained by conservation laws to reside on some high-dimensional surface (known as a hypersurface) in phase space. In this case, the system is constrained to move on a surface of constant energy E . The surface is actually an infinitesimally thick sheet of thickness δE , but this can be ignored for most purposes. The ergodic hypothesis allows us to replace a single system's time evolution over this surface with a large number of systems at uniformly random points on this path.

in Fig. 1. The **ergodic hypothesis** allows us to assume that over long periods of time, the time spent by the system in a region of phase space⁴ is proportional to the volume of said region, meaning that one is more likely to re-visit larger regions more frequently. A more useful formulation of this statement is as follows – suppose we initialise a large number of identical systems with N particles of total energy E , governed by the same Hamiltonian H . If we were to compute the probability distribution of microstates in the ensemble of identical systems, the ergodic hypothesis implies that this would return the same distribution as the temporal distribution obtained by monitoring the evolution of a single system over long times. Simply put, the ergodic hypothesis tells us that the temporal average is equivalent to the ensemble average (in the $t \rightarrow \infty, N \rightarrow \infty$ limit).

With this, we formally define the **statistical ensemble** as the collection of identical systems, with copies in every accessible microstate $\{\alpha(E)\}$. The ensemble average of a physical property X is then given by:

$$\langle X \rangle = \sum_{\alpha} p_{\alpha} X_{\alpha} \quad (1.3)$$

where p_{α} is the probability of occupying a single microstate α , subject to the normalisation condition on the total probability $\sum_{\alpha} p_{\alpha} = 1$.

1.3 Ensembles and Phase Space

With the idea of an ensemble now in place, we can specify the composition of an ensemble by invoking the idea of **phase space density**. With our statistical ensemble $\{\alpha(E)\}$ distributed around phase space along the hypersurface of constant E , the phase space density $\rho(\mathbf{q}, \mathbf{p})$ tells us how many copies of our system we expect to find in some infinitesimal region of phase space. Note that the vectors \mathbf{q} and \mathbf{p} are each $3N$ -dimensional

⁴When we say region here, we are technically identifying this as a macrostate, though there is some debate about the semantics of this in the mathematical community.

objects which represent the positions and momenta of all particles in the system. Mathematically, this is written as:

$$dN(\mathbf{q}, \mathbf{p}) = \rho(\mathbf{q}, \mathbf{p})d\mathbf{q}d\mathbf{p} \quad (1.4)$$

Specifically, $dN(\mathbf{q}, \mathbf{p})$ is the number of copies of the system found in the set of states between phase space coordinates (\mathbf{q}, \mathbf{p}) and $(\mathbf{q} + d\mathbf{q}, \mathbf{p} + d\mathbf{p})$. While we initiated this discussion of phase space by considering a single system evolving over time, then used the ergodic hypothesis to obtain an ensemble in phase space, let us now consider the evolution of the *ensemble* itself over time. Demanding that the number of copies in the ensemble must be conserved during evolution, we can write a continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (1.5)$$

where \mathbf{v} is the velocity of the flowing phase space density. Mathematically, we can write this velocity as:

$$\begin{aligned} d\mathbf{r} &= dq_i \hat{\mathbf{q}}_i + dp_i \hat{\mathbf{p}}_i \\ \implies \frac{d\mathbf{r}}{dt} &= \mathbf{v} = \dot{q}_i \hat{\mathbf{q}}_i + \dot{p}_i \hat{\mathbf{p}}_i \end{aligned} \quad (1.6)$$

where the implicit summation over i is implied (Einstein notation), and $\hat{\mathbf{q}}_i, \hat{\mathbf{p}}_i$ are the unit vectors corresponding to movement along those respective axes in phase space. We now use this in the continuity equation:

$$\begin{aligned} 0 &= \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \\ &= \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \\ &= \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i + \rho \left(\frac{\partial}{\partial q_i} \dot{q}_i + \frac{\partial}{\partial p_i} \dot{p}_i \right) \\ &= \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i + \rho \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} \right) \xrightarrow{0} \\ &= \left(\frac{\partial}{\partial t} + \dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i} \right) \rho \\ &= \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \right) \rho = \frac{\mathcal{D}\rho}{\mathcal{D}t} \end{aligned} \quad (1.7)$$

In the fourth line, we recognise that the time derivatives of the positions and momenta are provided by Hamilton's equations (Eq. 1.1,1.2), and use Schwarz' theorem to commute the partial derivatives. The final result $\mathcal{D}\rho/\mathcal{D}t$ is known as the **Lagrangian derivative**⁵. Unlike the standard time derivative $\partial/\partial t$ taken in a stationary coordinate system, the Lagrangian derivative accounts for the fact that the infinitesimal element under consideration may exist within an underlying flow, as shown in Fig. 2. That is, it represents the time derivative of the density of some volume which itself moves according to some velocity field \mathbf{v} . Physically, the vanishing of the Lagrangian derivative as shown in Eq. 1.7 implies that the density of a phase space volume element does not change as it moves through phase space, representing an incompressible flow in phase space!

Now, we also demand that the ensemble at equilibrium must have a constant phase space density over time, since any changes in ρ at any point would imply that the system is still undergoing the equilibration process. Mathematically, this is equivalent to setting $\partial\rho/\partial t = 0$, which then gives:

⁵Also known as the material derivative or convective derivative, depending on what department you're in.

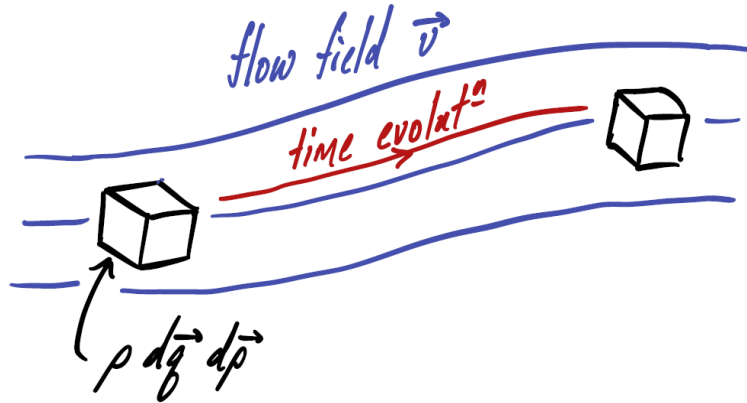


FIGURE 2: Lagrangian derivative with a flow field. The Lagrangian derivative differs from the standard time derivative in that it follows a specific region of the phase space as it flows according to the flow field (or velocity) \mathbf{v} .

$$\mathbf{v} \cdot (\nabla \rho) = 0 \quad (1.8)$$

This finally implies that the flow of phase space density occurs along a surface of constant ρ . Physically, this means that if ρ is a function of some quantity like the energy E of the system, then the phase space density is only able to flow along the hypersurface of constant E , implying that an equilibrium ensemble will remain at equilibrium for all times! With these underlying ideas of statistical ensembles and their evolution in place, we can start doing some real physics with them.

2 The Microcanonical Ensemble (MCE)

The first ensemble we encounter is the simplest in formulation, yet most often touted as the hardest to work with. We first need to introduce the quantity $\Omega(E, V, N)$, which I will frequently call the **multiplicity**⁶. The multiplicity represents the number of microstates of a system which have a total energy E , total volume V , total particle number N , in addition to any other quantities that may be relevant. Using this quantity, we state one of the most fundamental relations in statistical mechanics, the **Boltzmann entropy**⁷:

$$S(E, V, N) = k_B \log \Omega(E, V, N) \quad (2.1)$$

Thus, we now have a link between the macroscopic properties of a system (such as S, E, V, N), and the microscopic properties (such as Ω). Alongside his entropy postulate, Boltzmann also postulated that all microstates of a given energy E_0 are equally likely to be occupied. Mathematically, this is written as:

$$P_\alpha^{\text{mc}}(E, V, N) = \begin{cases} \frac{1}{\Omega(E_0, V, N)}, & E_0 - \delta E < E < E_0 \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

for some small energy window $\delta E \ll E_0$. Imposing this uniform probability distribution on the various microstates in the statistical ensemble then produces the **microcanonical ensemble** (MCE). Since the MCE

⁶I believe Bill uses the symbol \mathcal{N} in his notes for this same quantity. Boltzmann used W for *Wahrscheinlichkeit* (the German word for probability), and that is the notation used in the inscription of his entropy postulate on his gravestone.

⁷Some call this a postulate, others call it a formula. I choose to stick with the term *postulate* since it seems almost divined to me, but it's entirely possible that I just don't know enough about the origin of the statement.

demands that each system in the ensemble is confined to the hypersurfaces of constant E , V and N , it represents all possible states of a system that cannot exchange its energy (and other variables like volume, number, etc.). This makes the MCE the ideal ensemble for studying **isolated systems**, which are sufficiently decoupled from their environments that it justifies the assumption of constant E . Macroscopic properties are then easily obtained as ensemble averages using:

$$\langle X \rangle = \frac{1}{\Omega(E_0)} \sum_{\alpha} X_{\alpha} \quad (2.3)$$

We can also consider the characteristic size of fluctuations ΔX in the observed values of these macroscopic properties, which we expect to scale as $\Delta X/X \sim 1/\sqrt{N} \ll 1$. This is ridiculously small for a thermodynamically large system ($N \sim 10^{23}$), so the expectation values obtained from the MCE average are a good characterisation of the macroscopic properties at all times.

2.1 Microcanonical Thermodynamics

To study the thermodynamic properties of the microcanonical characterisation of a system, we begin with the fundamental thermodynamic relation⁸:

$$dE = TdS - pdV + \mu dN \quad (2.4)$$

$$\implies dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN \quad (2.5)$$

Identifying each of the differential coefficients with the respective partial derivatives, we thus see that the equations of state for this system are obtained as:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V,N} \quad (2.6)$$

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N} \quad (2.7)$$

$$\frac{\mu}{T} = - \left(\frac{\partial S}{\partial N} \right)_{E,V} \quad (2.8)$$

where the subscripts explicitly reflect the quantities kept constant in taking the partial derivative. We will now use this to establish the idea of thermodynamic equilibrium in the MCE. For this, we consider a composite system of total energy E , with two weakly interacting and statistically independent subsystems (E_1, V_1, N_1) and (E_2, V_2, N_2) , subject to the constraints that their energies, volumes and numbers must add as $E = E_1 + E_2$, $V = V_1 + V_2$, $N = N_1 + N_2$. Now, the multiplicity Ω factorises into a product over the multiplicities of the two subsystems:

$$\Omega(E, V, N) = \Omega_1(E_1, V_1, N_1) \times \Omega_2(E_2, V_2, N_2) \quad (2.9)$$

$$\implies S(E, V, N) = S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2) \quad (2.10)$$

which ensures the total entropy is additive. Furthermore, the maximisation of S at equilibrium also implies the maximisation of Ω , since S is a monotonically increasing function of Ω . We will use this to demonstrate

⁸This is really just a combination of the first and second laws of thermodynamics, and many would just call this the first law of thermodynamics. I choose not to do so, since the first law in its original form retains the fact that heat is a path variable and not a state variable, while the second law provides us the definition of entropy to turn the heat term into a well-defined state variable.

that the maximisation of the subsystems' multiplicities results in the two subsystems attaining thermodynamic equilibrium with one another. We start with Eq. 2.9:

$$d\Omega = \Omega_2 d\Omega_1 + \Omega_1 d\Omega_2 = 0 \quad (\text{at equilibrium}) \quad (2.11)$$

$$\frac{d\Omega_1}{\Omega_1} + \frac{d\Omega_2}{\Omega_2} = 0 \quad (2.12)$$

$$\implies d(\log \Omega_1) + d(\log \Omega_2) = 0 \quad (2.13)$$

Since the multiplicity is a function of E , V and N , we write its total differential as:

$$d(\log \Omega_i) = \left(\frac{\partial \log \Omega_i}{\partial E} \right)_{V,N} dE + \left(\frac{\partial \log \Omega_i}{\partial V} \right)_{E,N} dV + \left(\frac{\partial \log \Omega_i}{\partial N} \right)_{E,V} dN$$

Plugging this into Eq. 2.13 returns:

$$\left[\left(\frac{\partial \log \Omega_1}{\partial E} \right)_{V,N} - \left(\frac{\partial \log \Omega_2}{\partial E} \right)_{V,N} \right] dE_1 + \left[\left(\frac{\partial \log \Omega_1}{\partial V} \right)_{E,N} - \left(\frac{\partial \log \Omega_2}{\partial V} \right)_{E,N} \right] dV_1 + \left[\left(\frac{\partial \log \Omega_1}{\partial N} \right)_{E,V} - \left(\frac{\partial \log \Omega_2}{\partial N} \right)_{E,V} \right] dN_1 = 0 \quad (2.14)$$

where the imposition of the constraints forces $dE_1 = -dE_2$, and so on for the other variables. Since the right hand side has no dependence on any of the differential variables, each term in the square brackets must identically vanish. Using the equations of state from Eqs. 2.6–2.8, we thus arrive at:

$$\frac{1}{T_1} - \frac{1}{T_2} = 0 \implies T_1 = T_2 \quad (2.15)$$

$$\frac{p_1}{T_1} - \frac{p_2}{T_2} = 0 \implies p_1 = p_2 \quad (2.16)$$

$$\frac{\mu_2}{T_2} - \frac{\mu_1}{T_1} = 0 \implies \mu_1 = \mu_2 \quad (2.17)$$

Jointly, these equations imply the attainment of thermodynamic equilibrium between the two subsystems! We thus see that the Boltzmann entropy postulate acts as a bridge to connect the microscopic ideas of statistical mechanics with the macroscopic ideas of thermodynamics, whilst maintaining a consistent picture of thermodynamic equilibrium!

2.2 Two-Level Systems in the MCE

We now turn our attention to some physical examples of constructing and utilising the microcanonical ensemble for studying real systems. In the discrete case, there are two toy models which are quintessential for any statistical mechanics course. The first of these is the collection of **two-level systems**. These are sometimes also known as classical spins, in that they represent a point-like degree of freedom, occupying either of two states at a given time. The two states have energies ϵ_1 and ϵ_2 , which can always be transformed using an arbitrary linear shift in the energy to:

$$\epsilon_n = (-1)^n \epsilon \quad (2.18)$$

With a large N number of such classical spins, the microstates of the entire system are characterised are completely characterised by specifying the set of values $\alpha \equiv \{n_1, n_2, \dots, n_N\}$ which contains the state of each

spin. We also assume that the spins in the large set are able to exchange energy with one another, through some weak coupling interaction. Now that we have constructed the microstates, and hence the ensemble, we proceed by attempting to construct the multiplicity corresponding to a given energy E and number N . We first note that these are constrained by the conservation of energy and particle number as:

$$\begin{cases} N = N_1 + N_2 \\ E = (N_2 - N_1)\epsilon \end{cases} \quad (2.19)$$

We then write the multiplicity as:

$$\Omega(E, N; N_1, N_2) = \frac{N!}{N_1!N_2!} \quad (2.20)$$

Using this, we proceed to compute the entropy:

$$\begin{aligned} S(E, N; N_1, N_2) &= k_B \log \Omega(E, N; N_1, N_2) \\ &\approx -Nk_B \left(\frac{N_1}{N} \log \frac{N_1}{N} + \frac{N_2}{N} \log \frac{N_2}{N} \right) \end{aligned} \quad (2.21)$$

where I have used Stirling's approximation to simplify the factorials. Knowing that the answer should ultimately depend only on the total quantities E and N , we should invert Eq. 2.19 to obtain N_1 and N_2 as functions of E and N . These are easily obtained as:

$$\begin{cases} N_1 = \frac{N}{2} \left(1 - \frac{E}{N\epsilon} \right) \\ N_2 = \frac{N}{2} \left(1 + \frac{E}{N\epsilon} \right) \end{cases} \quad (2.22)$$

Plugging these into Eq. 2.21, we thus obtain the entropy:

$$S(E, N) = -\frac{Nk_B}{2} \left(1 + \frac{E}{N\epsilon} \right) \log \left[\frac{1}{2} \left(1 + \frac{E}{N\epsilon} \right) \right] - \frac{Nk_B}{2} \left(1 - \frac{E}{N\epsilon} \right) \log \left[\frac{1}{2} \left(1 - \frac{E}{N\epsilon} \right) \right] \quad (2.23)$$

Using the equation of state Eq. 2.6, we can then compute the thermodynamical properties of the system such as its temperature⁹:

$$T(E, N) = \frac{2\epsilon}{k_B} \left[\log \left(\frac{N\epsilon - E}{N\epsilon + E} \right) \right]^{-1} \quad (2.24)$$

In this manner, we can obtain all necessary thermodynamic properties by simply taking the appropriate derivatives and multiplying by the temperature (where applicable).

2.3 Harmonic Oscillators

We now consider the second toy model of interest, the **quantum harmonic oscillator**. In this setup, the system once again consists of a large number of copies of a point-like degree of freedom which can be in any state on an energy ladder. These energies are labelled by:

$$E_i = n_i \hbar \omega \quad (2.25)$$

⁹These derivatives and inversions are usually where a ton of mathematical gymnastics takes place, and is typically the reason why the microcanonical ensemble is harder to work with for studying thermodynamic properties. It also makes for amazing exam questions.

where E_i is the energy of the i -th oscillator, which is in the n_i -th excited state. The microstates are now once again labelled by $\alpha = \{n_1, n_2, \dots, n_N\}$, except n_i is now allowed to take any non-negative integer value. The energy constraint is now simply given by:

$$E = \hbar\omega \sum_{i=1}^N n_i \implies M = \frac{E}{\hbar\omega} = \sum_{i=1}^N n_i \quad (2.26)$$

where we define M to be the total number of energy quanta ($\hbar\omega$'s) present in the system. Now, we are left with the simple combinatorics problem of distributing M objects into N bins. As it turns out, it is a little easier to instead consider this equivalent problem – suppose we have M objects and wish to place $(N - 1)$ dividers between them. The dividers thus partition the set of objects into N distinct regions, each corresponding to the number of energy quanta placed in a given oscillator. This is a far easier combinatorial problem, as illustrated in Fig. 3, and it avoids having to impose an awkward constraint through a Kronecker delta.

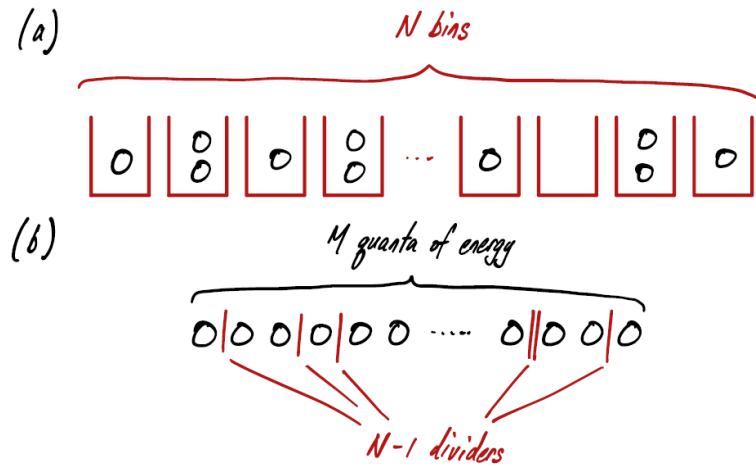


FIGURE 3: Quantum Harmonic Oscillators in the MCE. (a) With a fixed energy E , and hence a fixed number of energy quanta M , we have to distribute these quanta amongst N oscillators. (b) An equivalent problem is to instead place $(N - 1)$ dividers between M quanta, with the number of quanta between two adjacent dividers representing the number of quanta placed in a given oscillator. We are also allowed to place two (or more) dividers immediately adjacent to one another, which simply implies that no quanta are given to some oscillators.

In this modified problem, the multiplicity is then easily written as:

$$\Omega(E, N; M) = \frac{(N - 1 + M)!}{(N - 1)!M!} \quad (2.27)$$

From this, we can likewise compute the entropy:

$$\begin{aligned} S(E, N; M) &= k_B \log \Omega(E, N; M) \\ &\approx -Nk_B \left[\frac{M}{N} \log \left(\frac{M}{N + M} \right) + \log \left(\frac{N}{N + M} \right) \right] \end{aligned} \quad (2.28)$$

where we use Stirling's approximation to tame the factorials. Finally, the entropy is obtained in standard form by replacing M with E :

$$S(E, N) = -Nk_B \left[\frac{E}{N\hbar\omega} \log \left(\frac{E/\hbar\omega}{N + E/\hbar\omega} \right) + \log \left(\frac{N}{N + E/\hbar\omega} \right) \right] \quad (2.29)$$

In a similar manner to the previous example, all necessary thermodynamic properties can be extracted from this equation of state through the appropriate derivatives and subsequent gymnastics.

3 The Canonical Ensemble (CE)

Now that we have studied the microcanonical ensemble and the thermodynamic potentials which may be useful in non-isolated situations, we can start trying to construct new ensembles for these. While the MCE is a powerful construction for studying the statistical and thermodynamical properties of isolated systems, we rarely (or never, as I would argue) come across systems that are truly isolated. In fact, (maybe) the only true isolated system which exists is the entire universe as a whole. In fact, the vast majority of systems that are experimentally studied by physicists in the lab are well approximated by **closed systems**. These constitute systems of interacting particles which obey the constraint of fixed particle number, but allow for the exchange of energy with some external environment. As a result, it is common for us to characterise a system not by the variables (E, V, N) , but rather the variables (T, V, N) ¹⁰.

Our job now is to work within the framework of the MCE and find a way to adequately construct a statistical ensemble being held at constant T , V and N . The resultant statistical ensemble is known as the **canonical ensemble** (CE).

3.1 Obtaining the CE from the MCE

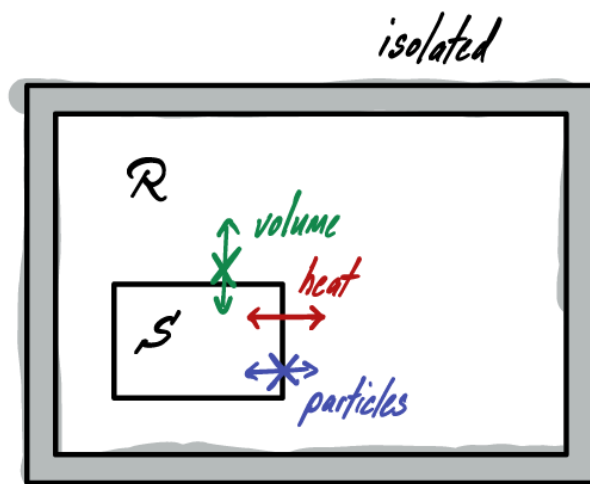


FIGURE 4: Construction of the Canonical Ensemble. We consider a small subsystem S placed amidst a much larger reservoir \mathcal{R} , such that the enclosure containing \mathcal{R} and S can be considered an isolated system. The subsystem S is then considered a good approximation to a closed system, with only the exchange of heat allowed across the barrier between itself and the reservoir.

In constructing this new ensemble from the MCE, our first job is to assign the right probability distribution to the microstates. To do this, we first consider a large isolated system with a much smaller subsystem S , which will be our system of interest. The remainder of the isolated system will be denoted by \mathcal{R} , which denotes the

¹⁰In the chemical sciences, it further turns out that the ideal variables are actually (T, p, N) . This is because controlled reactions and syntheses typically occur under environments of well-known temperature and pressure. In a physics experiment, however, it is far more common for us to enclose the system of interest within a rigid box, forbidding the exchange of volume with the environment.

thermal reservoir or thermal bath. Now, let \mathcal{S} be in a specific microstate α of energy E_α . The total energy of the composite isolated system is: $E_T = E_{\mathcal{R}} + E_\alpha$. Holding \mathcal{S} in a specific microstate α , let's now count the number of microstates available to the composite isolated system:

$$\begin{aligned}\Omega_T(E_T; E_\alpha) &= \Omega_{\mathcal{R}}(E_T - E_\alpha) \underbrace{\Omega_{\mathcal{S}}(E_\alpha)}_{=1} \\ &= \Omega_{\mathcal{R}}(E_T - E_\alpha)\end{aligned}\quad (3.1)$$

where we note that the multiplicity of the subsystem \mathcal{S} is trivially equal to 1 by virtue of our imposition that \mathcal{S} is in a specific microstate. To construct the probability distribution, we then use the fact that the composite isolated system is well represented by a MCE. Thus, we can say that all microstates of the composite system with energy E_T have equal probability, and we would just like to determine the conditional probability that the subsystem \mathcal{S} is also in microstate α , while the composite system has energy E_T . This probability is given by:

$$\begin{aligned}P_\alpha &= \frac{\Omega_T(E_T; E_\alpha)}{\Omega_T(E_T)} \\ &= \frac{\exp\left[\frac{S_{\mathcal{R}}(E_T - E_\alpha)}{k_B}\right]}{\exp\left[\frac{S_T(E_T)}{k_B}\right]}\end{aligned}\quad (3.2)$$

We now have the rough beginnings of a probability for the microstate α , but we need to obtain expressions for $S_T(E_T)$ and $S_{\mathcal{R}}(E_T - E_\alpha)$ in terms of the subsystem \mathcal{S} , since the characterisation of the reservoir \mathcal{R} and composite isolated system are usually unknown to us. One possible avenue is to use the fact that microcanonical fluctuations are incredibly small, so the typical energy $\langle E \rangle$ of the subsystem \mathcal{S} is a good characterisation of the measurable energy. We can use the additivity of entropy to write:

$$S_T(E_T) = S_{\mathcal{R}}(E_T - \langle E \rangle) + S(\langle E \rangle)\quad (3.3)$$

where $\langle E \rangle$ is the average internal energy of \mathcal{S} in any unrestricted state (it can be in any of its accessible microstates) at equilibrium. We now invoke the idea that any changes in the energy of the subsystem \mathcal{S} will never significantly change the energy of the reservoir \mathcal{R} , to make the approximation:

$$\frac{|\langle E \rangle - E_\alpha|}{|E_T - \langle E \rangle|} \ll 1\quad (3.4)$$

This is a key tool for performing a Taylor expansion on the entropy of the reservoir in Eq. 3.2, since this will now relate the entropy of \mathcal{R} to the energies of the subsystem \mathcal{S} and the total composite system. For this, we take the entropy of the reservoir and perform the standard trick of inserting a vanishing term:

$$S_{\mathcal{R}}(E_T - E_\alpha) = S_{\mathcal{R}}[(E_T - \langle E \rangle) + (\langle E \rangle - E_\alpha)]\quad (3.5)$$

With this expression, we can invoke a Taylor expansion around $(E_T - \langle E \rangle)$ with the perturbation as the term $(\langle E \rangle - E_\alpha)$. This produces a first-order expansion:

$$\begin{aligned}S_{\mathcal{R}}(E_T - E_\alpha) &\approx S_{\mathcal{R}}(E_{\mathcal{R}}) \Big|_{E_{\mathcal{R}}=E_T-\langle E \rangle} + \frac{dS_{\mathcal{R}}(E_{\mathcal{R}})}{dE_{\mathcal{R}}} \Big|_{E_{\mathcal{R}}=E_T-\langle E \rangle} (\langle E \rangle - E_\alpha) \\ &= S_{\mathcal{R}}(E_T - \langle E \rangle) + \frac{1}{T_{\mathcal{R}}} (\langle E \rangle - E_\alpha)\end{aligned}\quad (3.6)$$

Thus, we have the temperature $T_{\mathcal{R}}$ of the reservoir as a relatively fixed value, which sets the temperature of the subsystem if the reservoir is sufficiently large compared to the subsystem. We can finally write the probability of occupying a specific microstate α from Eq. 3.2 as:

$$\begin{aligned} P_{\alpha} &= \frac{\exp\left[\frac{S_{\mathcal{R}}(E_{\mathcal{R}} - \langle E \rangle)}{k_B}\right] \exp\left[\frac{\langle E \rangle - E_{\alpha}}{k_B T_{\mathcal{R}}}\right]}{\exp\left[\frac{S_{\mathcal{R}}(E_{\mathcal{R}})}{k_B}\right]} \\ &= \exp\left[-\frac{S(\langle E \rangle)}{k_B}\right] \exp\left(\frac{\langle E \rangle}{k_B T_{\mathcal{R}}}\right) \exp\left(-\frac{E_{\alpha}}{k_B T_{\mathcal{R}}}\right) \end{aligned} \quad (3.7)$$

In this expression for the probability of a single microstate α of the subsystem \mathcal{S} , we note that the only appearance of a temperature is in the temperature $T_{\mathcal{R}}$ of the reservoir. This agrees with our intuition that a sufficiently large reservoir at a fixed temperature will set the temperature of the smaller system which it is in contact with. We thus make the identification that $T_{\mathcal{R}} \equiv T$ is the only relevant temperature in the composite system and rewrite the probability as:

$$\begin{aligned} P_{\alpha} &= \exp\left[\frac{\langle E \rangle - TS(\langle E \rangle)}{k_B T}\right] \exp\left(-\frac{E_{\alpha}}{k_B T}\right) \\ &= \exp(\beta F) \exp(-\beta E_{\alpha}), \quad \text{where } \beta = \frac{1}{k_B T} \end{aligned} \quad (3.8)$$

where we use the definition of the Helmholtz free energy $F = \langle E \rangle - TS(\langle E \rangle)$. Thus, we finally have the probability of occupancy in each microstate α of the subsystem \mathcal{S} , in terms of measurable observables of the subsystem. Interestingly, we see that the macroscopic Legendre transformation from internal energy to the Helmholtz free energy was automatically performed by the coupling of the system to the heat bath \mathcal{R} , with the appearance of F in the canonical probability distribution now. To complete the procedure, we just need to normalise this to ensure that all probabilities sum to unity. We thus obtain:

$$\begin{aligned} \sum_{\alpha} \exp(\beta F) \exp(-\beta E_{\alpha}) &= 1 \\ \implies \sum_{\alpha} \exp(-\beta E_{\alpha}) &= \exp(-\beta F) = Z(T, V, N) \end{aligned} \quad (3.9)$$

Here, we have defined the quantity $Z(T, V, N)$ which is obtained by summing all of the individual terms $\exp(-\beta E_{\alpha})$ for all subsystem microstates. This is arguably the most important object in practical statistical mechanics, and is known as the **canonical partition function**. With the canonical partition function, we have a normalisation factor for our probabilities and can write:

$$P_{\alpha} = \frac{\exp(-\beta E_{\alpha})}{\sum_{\alpha} \exp(-\beta E_{\alpha})} = \frac{1}{Z} \exp(-\beta E_{\alpha}) \quad (3.10)$$

The canonical probability distribution thus applies a weight to the probability of a microstate α that is exponentially linked to its energy E_{α} and the inverse temperature β . The canonical partition function $Z(T, V, N)$ enters this expression by acting as a normalisation factor, being defined as the sum over all canonical weights $\exp(-\beta E_{\alpha})$. In some cases, it turns out to be easier to sum the partition function over energy levels E_{α} instead of microstates α . However, we must then occur for the degeneracy $g(E_{\alpha})$ of each energy level:

$$Z(T, V, N) = \sum_{E_{\alpha}} g(E_{\alpha}) \exp(-\beta E_{\alpha}) \quad (3.11)$$

As with the microcanonical ensemble, this assumes that all microstates with the same energy are equally likely to be occupied.

3.2 Canonical Thermodynamics

While the canonical partition function looks like a simple normalisation factor, it turns out to be an incredibly powerful link between the microscopic and macroscopic worlds of statistical mechanics and thermodynamics respectively. Just as the Boltzmann entropy postulate provides such a link in studying an isolated system, the canonical partition function plays a similar role in the case of a closed system. We start by looking at the thermodynamic quantities associated with this ensemble. The simplest and most straightforward is the energy:

$$\begin{aligned}
 \langle E \rangle &= \sum_{\alpha} P_{\alpha} E_{\alpha} \\
 &= \frac{1}{Z} \sum_{\alpha} E_{\alpha} \exp(-\beta E_{\alpha}) \\
 &= -\frac{1}{Z} \left(\frac{\partial Z}{\partial \beta} \right)_{V,N} \\
 &= -\left(\frac{\partial \log Z}{\partial \beta} \right)_{V,N}
 \end{aligned} \tag{3.12}$$

We can also consider the characteristic size of energy fluctuations through:

$$\begin{aligned}
 \langle (\Delta E)^2 \rangle &= \langle (E - \langle E \rangle)^2 \rangle \\
 &= \langle E^2 \rangle - \langle E \rangle^2 \\
 &= \frac{1}{Z} \left(\frac{\partial^2 Z}{\partial \beta^2} \right)_{V,N} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)_{V,N}^2 \\
 &= \left(\frac{\partial^2 \log Z}{\partial \beta^2} \right)_{V,N}
 \end{aligned} \tag{3.13}$$

$$= -\left(\frac{\partial \langle E \rangle}{\partial \beta} \right)_{V,N} \tag{3.14}$$

From the fundamental relation in energy representation (Eq. ??), we observe that:

$$\left(\frac{\partial E}{\partial T} \right)_{V,N} = T \left(\frac{\partial S}{\partial T} \right)_{V,N} = C_V \tag{3.15}$$

where C_V is the heat capacity of the system at constant volume. In the thermodynamic setting, we clearly have $\langle E \rangle = E$ since the thermodynamic variables are defined at equilibrium. Finally, we convert the derivative $\partial/\partial\beta = -k_B T^2 \partial/\partial T$ to obtain:

$$\langle (\Delta E)^2 \rangle = k_B T^2 C_V \tag{3.16}$$

Rather strangely, this relates the typical size of energy fluctuations in the canonical ensemble to the temperature and constant-volume heat capacity, which are both measurable quantities! On the right, we see that T is intensive, while C_V is extensive, so we expect the fluctuations to scale as $\langle (\Delta E)^2 \rangle \sim N$, and thus the relative size of the fluctuations scale as:

$$\frac{\sqrt{\langle (\Delta E)^2 \rangle}}{\langle E \rangle} \sim \frac{1}{\sqrt{N}} \ll 1 \tag{3.17}$$

Thus, we see that while the energy fluctuates in the canonical ensemble, the fluctuations are so absurdly small relative to the energy (suppressed by a factor of approximately 10^{-12}) that only a small range of energies around the expected equilibrium $\langle E \rangle$ are important for any computation. As we recall, the microcanonical ensemble also uses a thin energy shell of small size $\delta E \ll E$ for restricting the microcanonical probability distribution under consideration. These restrictions are, in a good sense, the same consideration – we thus consider both the canonical and microcanonical ensemble equivalent for practical computational purposes.

Finally, we can also link Eq. ?? and Eq. 3.9 to write expressions for the three equations of state: (using the fact that the canonical partition function is linked to the Helmholtz free energy as $F = -k_B T \log Z(T, V, N)$)

$$S(T, V, N) = - \left(\frac{\partial F}{\partial T} \right)_{V, N} \quad (3.18)$$

$$p(T, V, N) = - \left(\frac{\partial F}{\partial V} \right)_{T, N} \quad (3.19)$$

$$\mu(T, V, N) = \left(\frac{\partial F}{\partial N} \right)_{T, V} \quad (3.20)$$

3.3 Two-Level Systems in the CE

We now return to the example of two-level systems, but instead study this in the canonical ensemble held at fixed temperature T . The setup here is similar to before, in that we have a large N number of identical, weakly-interacting classical spins. Each particle once again occupies a non-degenerate energy level given by Eq. 2.18. The microstates of the system are then described by the set of values $\{n_i\} \equiv \alpha$. The goal in the MCE formalism was to find an expression for the entropy, from which all other thermodynamic quantities could be obtained using the appropriate derivatives. Here, we have a similar goal, except our keystone now comes in the form of the canonical partition function. We write this as:

$$\begin{aligned} Z(T, N) &= \sum_{\alpha} \exp(-\beta E_{\alpha}) \\ &= \sum_{n_1=1}^2 \sum_{n_2=1}^2 \cdots \sum_{n_N=1}^2 \exp \left[-\beta \sum_{i=1}^N (-1)^{n_i} \epsilon \right] \\ &= \prod_{i=1}^N \sum_{n_i=1}^2 \exp [(-1)^{n_i+1} \beta \epsilon] \\ &= [\exp(\beta \epsilon) + \exp(-\beta \epsilon)]^N \\ &= [2 \cosh(\beta \epsilon)]^N \end{aligned} \quad (3.21)$$

In the second line, we convert the sum over microstates to a product of N sums over the N classical spins, each of which can occupy one of two possible states. In the third line, we recognise that all N sums are identical, so this can be written as a product of N copies of a single sum. In the fourth line, we explicitly evaluate the sum since only two states are in consideration for a single classical spin, and finally convert it to the cleaner hyperbolic expression. This also allows us to evaluate the Helmholtz free energy simply as:

$$\begin{aligned} F(T, N) &= -k_B T \log Z \\ &= -N k_B T \left[\log 2 + \log \cosh \left(\frac{\epsilon}{k_B T} \right) \right] \end{aligned} \quad (3.22)$$

3.4 Harmonic Oscillators in the CE

We also return to the example of the quantum harmonic oscillators, with our setup similar to before except the entire system is now held at some given temperature T . The energy of each oscillator is once again provided by Eq. 2.25. We then proceed by finding the canonical partition function:

$$\begin{aligned}
 Z(T, N) &= \sum_{\alpha} \exp(-\beta E_{\alpha}) \\
 &= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \exp\left(-\beta \sum_{i=1}^N n_i \hbar \omega\right) \\
 &= \prod_{i=1}^N \sum_{n_i=0}^{\infty} \exp(-\beta \hbar \omega n_i) \\
 &= \left[\sum_{n=0}^{\infty} \exp(-\beta \hbar \omega n) \right]^N \\
 &= \left[\frac{1}{1 - \exp(-\beta \hbar \omega)} \right]^N \tag{3.23}
 \end{aligned}$$

In the last line, we use the formula for the infinite geometric sum to obtain the compact final expression. We can go a small step further and obtain the Helmholtz free energy in this representation as:

$$\begin{aligned}
 F(T, N) &= -k_B T \log Z \\
 &= N k_B T \log \left[1 - \exp\left(\frac{\hbar \omega}{k_B T}\right) \right] \tag{3.24}
 \end{aligned}$$

3.5 Factorisation of the Canonical Partition Function

Before ending our study of the canonical partition function, we should also consider the special cases where the Hamiltonian of the system is composed of separable parts. In this situation, the energy of each microstate can be broken into several contributions which are decoupled from one another. Suppose we have the separable Hamiltonian of the form:

$$H = \sum_i H_i \implies E_{\alpha} = \sum_i E_i \tag{3.25}$$

This would allow for the canonical partition function to be broken up as:

$$Z = \sum_{\alpha} \exp(-\beta E_{\alpha}) = \prod_i \sum_{\alpha_i} \exp(-\beta E_i) = \prod_i Z_i \tag{3.26}$$

Thus, the partition function is multiplicatively separable for an additively separable Hamiltonian! This means we can study the various Hamiltonian contributions individually to form partition functions for each degree of freedom, before obtaining the overall partition function by simply multiplying the contributions together. This is also why a system with N copies of some identical smaller system has its partition function easily broken up as N copies of the single-particle partition function¹¹. There is one small problem here, which turns out to be a huge problem for the case of identical/indistinguishable particles – this grossly overcounts

¹¹This is one of the main reasons why the study of interacting systems is incredibly difficult, as the introduction of interactions between constituent particles will mix the degrees of freedom together and no longer allow us to decouple the various parts of the Hamiltonian.

the number of microstates present. For theatics, let me show you the solution now but demonstrate a case where this returns really stupid results later on. The correct expression for identical particles is given by:

$$Z_N = \frac{1}{N!} Z_1^N \quad (3.27)$$

3.6 The Gibbs Entropy Formula

Previously, we had the Boltzmann entropy postulate in Eq. 2.1 which told us the entropy associated with an isolated system at some fixed energy, volume and particle number (in addition to any other known extensive variables). In the canonical ensemble, we found that the entropy could be obtained as a simple derivative of the partition function, as dictated by the equations of state for the Helmholtz representation of the fundamental relation. The question now is whether we can find a statement for the entropy that is similar to the Boltzmann postulate, in that it describes the entropy based on the microscopic probability distribution, rather than the macroscopic equilibrium free energy. Indeed, such an expression exists and is not hard to obtain. We begin with the entropy from Eq. ?? as:

$$\begin{aligned} S &= \frac{1}{T} (\langle E \rangle - F) \\ &= k_B \beta (\langle E \rangle - F) \\ &= k_B (\log Z + \beta \langle E \rangle) \end{aligned} \quad (3.28)$$

where in the last line, we use the relation Eq. 3.9 to convert the appearance of the macroscopic Helmholtz free energy to the microscopic partition function. To proceed, we must somehow thrust ourselves back into the microscopic world, and this is easily achieved through the second term, which is a probabilistic sum over all microscopic energies of the various microstates. We thus write:

$$\begin{aligned} S &= k_B \sum_{\alpha} P_{\alpha} (\log Z + \beta E_{\alpha}) \\ &= k_B \sum_{\alpha} P_{\alpha} \log (Z e^{\beta E_{\alpha}}) \\ &= k_B \sum_{\alpha} P_{\alpha} \log \frac{1}{P_{\alpha}} \\ &= -k_B \sum_{\alpha} P_{\alpha} \log P_{\alpha} \end{aligned} \quad (3.29)$$

where in the third line, we invoke the definition of the canonical probability from Eq. 3.10. This final expression is known as the **Gibbs entropy formula**, and is an incredibly general and powerful way to express the entropy of a system in terms of its microstate probabilities¹². As a simple example, we can show that Gibbs' entropy formula reduces to that of Boltzmann in the MCE where $P_{\alpha}^{\text{mce}} = 1/\Omega$:

¹²This also provides a beautiful connection between statistical mechanics and information theory, though it's funny you never see Claude Shannon mention Gibbs' name even once in his original paper where he introduces his famous entropy formula (which is nearly identical to Gibbs').

$$\begin{aligned}
S &= -k_B \sum_{\alpha} P_{\alpha} \log P_{\alpha} \\
&= -k_B \sum_{\alpha} \frac{1}{\Omega} \log \frac{1}{\Omega} \\
&= k_B \sum_{\alpha} \frac{1}{\Omega} \log \Omega \\
&= k_B \log \Omega
\end{aligned}$$

where we simply sum over all Ω microstates to cancel the factor of $1/\Omega$. In general, the question of which entropy formula is more useful ultimately depends on the provided information. If we are handed a ton of microscopic information, such as the probability distribution amongst the microstates¹³, then the Gibbs formula is ideal. If we are instead given the Hamiltonian of the system and know to use the canonical ensemble since it is being held at some fixed temperature, it may be easier to work out the partition function and take the appropriate derivative of the macroscopic equilibrium Helmholtz free energy. If we are given an isolated system, then we should turn to counting arguments and determine an expression for Ω itself, obtaining the entropy using the Boltzmann entropy postulate.

4 The Grand Canonical Ensemble (GCE)

We finally turn to the last statistical ensemble of great interest in this course, which arises from relaxing the restriction of particle conservation. That is, we had previously considered systems which always maintained a constant particle number in both the microcanonical and canonical ensembles. Real physical systems which attain chemical equilibrium almost definitely exhibit some form of particle exchange with an external reservoir, and this means the set of variables (T, V, N) are no longer ideal for characterising the equilibrium state of the system since we would expect N to fluctuate. Instead, we now expect that the conjugate variable μ , known as the chemical potential, would be an ideal characterisation of the equilibrium state, leaving us with the canonical variables (T, V, μ) . This is obtained by taking the Legendre transform of the Helmholtz free energy on the conjugate variable pair (μ, N) to obtain:

$$\Phi = \langle E \rangle - TS - \mu \langle N \rangle \implies d\Phi = -S dT - p dV - N d\mu \quad (4.1)$$

where Φ is known as the **grand potential**, or **Landau potential** depending on your school, or the schools of your educators. This potential will return later, though it is good to have the expectation beforehand that the grand potential will make an appearance, given that we are going to couple our system of interest to a heat and particle reservoir. We now begin again from the framework of the MCE, and (in a similar vein to the CE), construct a statistical ensemble being held at constant T, V and μ . The resultant statistical ensemble is known as the **grand canonical ensemble** (GCE).

4.1 Obtaining the GCE from the MCE

Similar to the derivation of the CE from the MCE, we consider a subsystem \mathcal{S} coupled to a much larger reservoir \mathcal{R} , with the entire system $\mathcal{R} + \mathcal{S}$ considered isolated. The entire system has total energy E_T and total particle number N_T , while the subsystem \mathcal{S} in a given microstate α has energy $E_{N_S, \alpha}$, with the subscript denoting

¹³Nature (and Bill) will probably never do this for you, it's too nice and unrealistic.

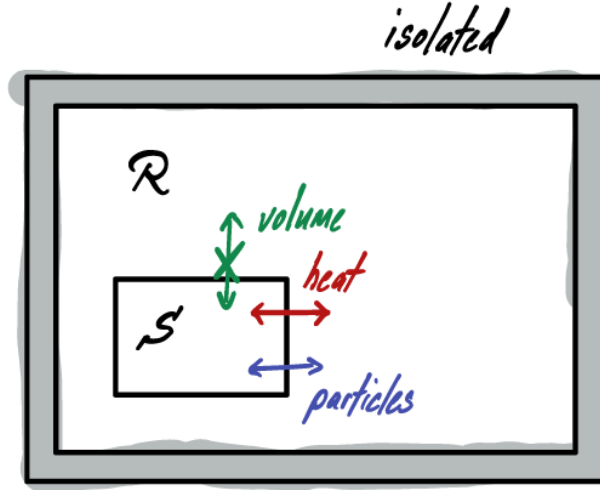


FIGURE 5: Construction of the Grand Canonical Ensemble. We consider a small subsystem S placed amidst a much larger reservoir \mathcal{R} , such that the enclosure containing \mathcal{R} and S can be considered an isolated system. The subsystem S is then considered an open system, with the exchange of heat and particles allowed across the barrier between itself and the reservoir.

that this energy is for a definite microstate with N_S particles. Then, the total energy is: $E_T = E_{\mathcal{R}} + E_{N_S, \alpha}$, where the energy of a single microstate can be interpreted as $E_{N_S, \alpha} = E_{\alpha}(N_S)$. We now count the multiplicity for the total system given that the subsystem S is in a definite microstate α :

$$\begin{aligned} \Omega_T(E_T, N_T; E_{N_S, \alpha}) &= \Omega_{\mathcal{R}}(E_T - E_{N_S, \alpha}, N_T - N_S) \underbrace{\Omega(E_{N_S, \alpha}; N_S)}_{=1} \\ &= \Omega_{\mathcal{R}}(E_T - E_{N_S, \alpha}, N_T - N_S) \end{aligned} \quad (4.2)$$

where the multiplicity of the subsystem is trivially set to 1 by virtue of our construction. Now, we invoke the Boltzmann entropy to get the probability of the subsystem S being in a microstate α with energy $E_{N_S, \alpha}$:

$$\begin{aligned} P_{N_S, \alpha} &= \frac{\Omega_T(E_T, N_T; E_{N_S, \alpha})}{\Omega_T(E_T, N_T)} \\ &= \frac{\exp\left[\frac{S_{\mathcal{R}}(E_T - E_{N_S, \alpha}, N_T - N_S)}{k_B}\right]}{\exp\left[\frac{S_T(E_T, N_T)}{k_B}\right]} \end{aligned} \quad (4.3)$$

To remove the dependence of the entropy expression on characteristics of the reservoir (which are unmeasurable), we once again hint at expanding around the equilibrium expectation values of E and N , for which we need to make the following assumptions:

$$\frac{|\langle E \rangle - E_{N_S, \alpha}|}{|E_T - \langle E \rangle|} \ll 1 \quad (4.4)$$

$$\frac{|\langle N \rangle - N_S|}{|N_T - \langle N \rangle|} \ll 1 \quad (4.5)$$

We then use this to expand the entropy of the reservoir to first order in small deviations from equilibrium, in both the energy and particle number:

$$\begin{aligned}
S_{\mathcal{R}}(E_T - E_{N_S, \alpha}, N_T - N_S) &= S_{\mathcal{R}}[(E_T - \langle E \rangle) + (\langle E \rangle - E_{N_S, \alpha}), (N_T - \langle N \rangle) + (\langle N \rangle - N_S)] \\
&\approx S_{\mathcal{R}}(E_T - \langle E \rangle, N_T - \langle N \rangle) + \frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}} \Big|_{E_{\mathcal{R}}=E_T - \langle E \rangle} (\langle E \rangle - E_{N_S, \alpha}) \\
&\quad + \frac{\partial S_{\mathcal{R}}}{\partial N_{\mathcal{R}}} \Big|_{N_{\mathcal{R}}=N_T - \langle N \rangle} (\langle N \rangle - N_S) \\
&= S_{\mathcal{R}}(E_T - \langle E \rangle, N_T - \langle N \rangle) + \frac{\langle E \rangle - E_{N_S, \alpha}}{T_{\mathcal{R}}} - \frac{\mu_{\mathcal{R}}(\langle N \rangle - N_S)}{T_{\mathcal{R}}} \tag{4.6}
\end{aligned}$$

where in the last line, we use the equations of state for the grand potential (as is easily visible from Eq. 4.1) to obtain the chemical potential $\mu_{\mathcal{R}}$ of the reservoir. Since we expect that in chemical equilibrium, the chemical potential of the reservoir will set the chemical potential of the subsystem as well, we can make the identification that $\mu_{\mathcal{R}} = \mu$. Finally, we rid ourselves of the reservoir entropy by using the additivity of entropy as:

$$S_T(E_T, N_T) = S_{\mathcal{R}}(E_T - \langle E \rangle, N_T - \langle N \rangle) + S(\langle E \rangle, \langle N \rangle) \tag{4.7}$$

Combining this with our new expansion for the entropy of the reservoir, we return to compute the grand canonical probability as:

$$\begin{aligned}
P_{N_S, \alpha} &= \frac{\exp \left[\frac{S_{\mathcal{R}}(E_T - \langle E \rangle, N_T - \langle N \rangle)}{k_B} \right] \exp \left[\frac{\langle E \rangle - E_{N_S, \alpha}}{k_B T} \right] \exp \left[\frac{-\mu(\langle N \rangle - N_S)}{k_B T} \right]}{\exp \left[\frac{S_{\mathcal{R}}(E_T - \langle E \rangle, N_T - \langle N \rangle)}{k_B} \right] \exp \left[\frac{S(\langle E \rangle, \langle N \rangle)}{k_B} \right]} \\
&= \exp \left(\frac{\langle E \rangle - TS - \mu \langle N \rangle}{k_B T} \right) \exp \left(-\frac{E_{N_S, \alpha}}{k_B T} \right) \exp \left(\frac{\mu N_S}{k_B T} \right) \\
&= \exp(\beta \Phi) \exp(-\beta E_{N_S, \alpha}) \exp(\beta \mu N_S) \tag{4.8}
\end{aligned}$$

Finally, we impose the normalisation condition on this grand canonical probability:

$$\begin{aligned}
\sum_{N_S} \sum_{\alpha} \exp(\beta \Phi) \exp[-\beta(E_{N_S, \alpha} - \mu N_S)] &= 1 \\
\implies \sum_{N_S} \sum_{\alpha} \exp[-\beta(E_{N_S, \alpha} - \mu N_S)] &= \exp(-\beta \Phi) = \mathcal{Z}(T, V, \mu) \tag{4.9}
\end{aligned}$$

which finally returns us the object $\mathcal{Z}(T, V, \mu)$ ¹⁴, known as the **grand canonical partition function**, or grand partition function. In an incredibly theatric and humble manner, I wish to convey that *this* is truly the most powerful function in statistical mechanics (in my honest opinion). While it looks like we just repeated the procedure for the derivation of the canonical ensemble, now allowing for a new type of exchange between the system and bath, the physical implications run far deeper. Let's explore some of that first before we delve into the thermodynamics of things.

The grand canonical probability provides the relative probability of occupying a microstate α with energy E_{α} and particle number N_{α} . In writing the grand partition function which sums over all of these terms, it is wiser to interpret these two sums as follows:

1. Fix N first.

¹⁴This is why I typically don't cross my Z 's, because I reserve that for the grandest of partition functions. A good number of texts (and the way I was taught) uses the symbol $Q(T, V, N)$ for the canonical partition function so that we have $Z(T, V, \mu)$ for the grand partition function, but I am trying to stick with the same terminology as the class as best I can.

2. Sum over all microstates α which are available for that N .
3. Iterate over all values of N .

Mathematically, this is written as:

$$\begin{aligned}
\mathcal{Z}(T, V, \mu) &= \sum_{\alpha} e^{-\beta E_{\alpha} + \beta \mu N_{\alpha}} \\
&= \sum_N \sum_{\alpha} e^{-\beta E_{\alpha} + \beta \mu N} \delta_{N, N_{\alpha}} \\
&= \sum_N e^{\beta \mu N} \sum_{\alpha} e^{-\beta E_{\alpha}} \delta_{N, N_{\alpha}} \\
&= \sum_N e^{\beta \mu N} Z(T, V, N)
\end{aligned} \tag{4.10}$$

where I now include the Kronecker delta in the canonical partition function to explicitly demand that particle numbers are fixed in the canonical ensemble. This form tells us that the grand partition function is really just a weighted sum of canonical partition functions! We can also go the other way, extracting a canonical partition at some fixed particle number from a grand partition function, though this requires some careful work. We first define the quantity $z = e^{\beta \mu}$, known as the **fugacity**. With this, the grand partition function is written as a power series expansion:

$$\begin{aligned}
\mathcal{Z}(T, V, \mu) &= \sum_N e^{\beta \mu N} Z(T, V, N) \\
&= \sum_N z^N Z(T, V, N)
\end{aligned} \tag{4.11}$$

where the canonical partition functions $Z(T, V, N)$ are the coefficients, and $\mathcal{Z}(T, V, \mu)$ is expanded in powers of the fugacity. We thus use the standard method of extracting the coefficients of a Taylor series to obtain the canonical partition function for some desired particle number N_0 :

$$Z(T, V, N_0) = \frac{1}{N_0!} \left[\left(\frac{\partial}{\partial z} \right)^{N_0} \mathcal{Z}(T, V, \mu) \right] \Bigg|_{z=0} \tag{4.12}$$

To truly demonstrate why I prefer the grand partition function, we need to consider the fact that a large number of the systems we have studied (and will study) have discrete energy levels¹⁵. Furthermore, when we work in the regime of indistinguishability whereby the constituent particles which take on those discrete energies cannot be distinguished from one another, counting microstates becomes tough. That is, when we cannot tell particles apart from one another, then assigning specific energy levels to specific particles loses its meaning as a process. Instead, it is physically more meaningful to count the number of particles occupying a given energy level, from which the total energy of the system is easily obtained. This is pictorially represented in Fig. 6.

To convey the difficulty of the particle labelling procedure in the indistinguishable case, suppose we have a collection of N identical particles, each of which can have some energy ϵ_i . We also assume that this gas of particles is held at some temperature T through coupling to a heat reservoir, but there is no exchange of particles allowed, so N is constant. The canonical partition function is:

¹⁵This argument would probably break down if this were Bill's biophysics class, so please take all of this with an entire shaker of salt.

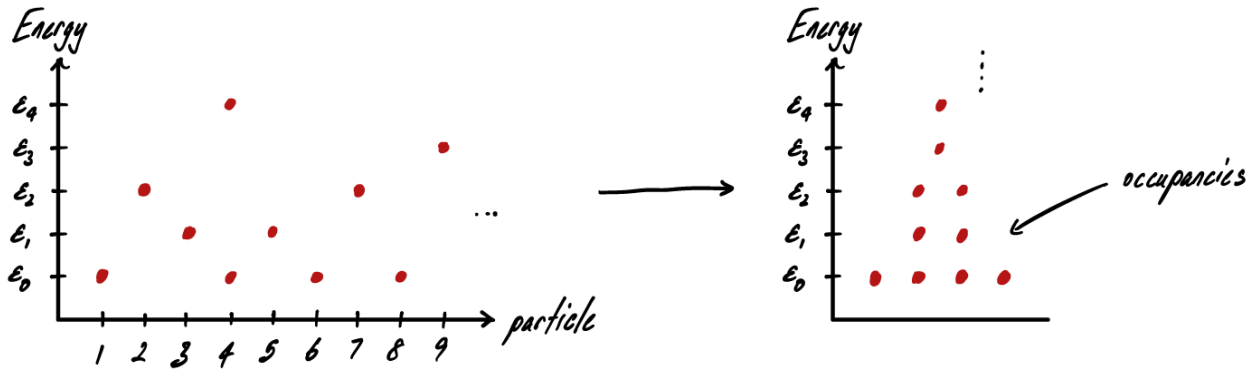


FIGURE 6: Counting Distinguishable vs. Indistinguishable Particles. (Left) When the constituent particles are distinguishable from one another, we can assign specific energies to each particle. (Right) When the particles are identical (or sufficiently indistinguishable), it is more justified to instead count the *number* of particles occupying each energy level.

$$\begin{aligned}
 Z(T, V, N) &= \sum_{\{n_i\}} e^{-\beta \sum_j \epsilon_j n_j} \delta_{N, \sum_j n_j} \\
 &= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots \sum_{n_j=0}^{\infty} \dots e^{-\beta \sum_j \epsilon_j n_j} \delta_{N, \sum_j n_j}
 \end{aligned} \tag{4.13}$$

The first sum is over all possible occupancy distributions $\{n_i\}$ of the various energy levels, and the Kronecker delta ensures that all of these occupancies sum up to N . Thus, when the configurational sum for $\{n_i\}$ is expanded in the second line, not all of the sums are independent since the Kronecker delta will eliminate one of them. This situation is made even worse if there are limits on the number of particles within each energy level, such as when the particles are fermions (then only a maximum of one particle can occupy a single state). Instead, if we view each energy level as an open system which exchanges particles (and heat) with a larger reservoir that is our closed system, and have our closed system fixed at some temperature through heat exchange with a reservoir (all within an isolated system) the situation becomes a lot simpler!¹⁶ Suppose, for now, that there is no restriction on the number of particles which can occupy a given energy level (so these particles are bosons), the grand partition function of each energy level is then:

$$Z_i = \sum_{n_i=0}^{\infty} e^{-\beta n_i (\epsilon_i - \mu)} \tag{4.14}$$

Knowing that the partition function for a system comprised of subsystems is simply the product of the individual partition functions, we can easily combine all the energy levels' partition functions to obtain:

$$\begin{aligned}
 Z(T, V, \mu) &= \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \dots \sum_{n_i=0}^{\infty} \dots e^{-\beta \sum_i n_i (\epsilon_i - \mu)} \\
 &= \prod_i \sum_{n_i=0}^{\infty} e^{-\beta n_i (\epsilon_i - \mu)} \\
 &= \prod_i \frac{1}{1 - e^{-\beta (\epsilon_i - \mu)}}
 \end{aligned} \tag{4.15}$$

¹⁶Sorry this sentence was a massive chunk, I hope the illustration helps. This whole process of nesting systems feels a lot like the SNL Taco Town skit from 2005.

where we are able to evaluate the sum since all energy levels are coupled to the same particle bath, so the chemical potentials of all energy levels at equilibrium must all be the same. Note that we do not need to impose the particle number constraint here since N is no longer fixed in the grand canonical ensemble. Instead, the particle number N is determined using the equation of state for the grand potential, through which a fixed particle number constraint can later be imposed by tuning the new canonical variables (T, V, μ) .

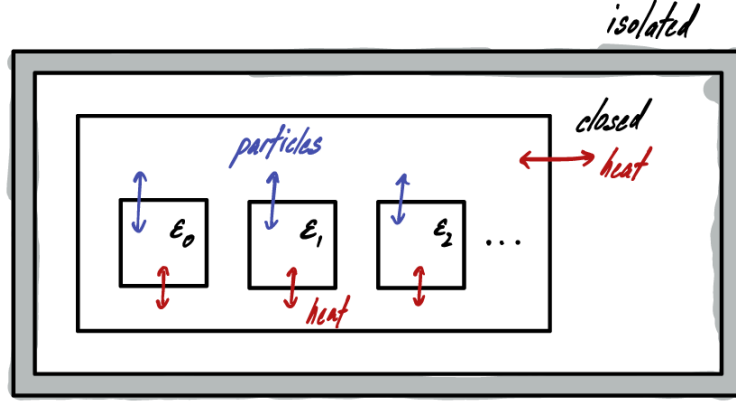


FIGURE 7: Closed Systems as Interacting Open Systems. When the constituent particles of a closed system are indistinguishable from one another, it is preferable to work in the basis of occupation numbers for the various energy levels. Here, we consider each energy level as an open system, able to exchange heat and particles with the closed system as a reservoir. In this manner, the partition function of the entire system is better studied by considering the grand partition functions of the constituent energy levels, rather than the single canonical partition function of the closed system.

4.2 Grand Canonical Thermodynamics

We first consider the average/expected particle number in the grand canonical ensemble through:

$$\begin{aligned}
 \langle N \rangle &= \sum_{\alpha} P_{\alpha} N_{\alpha} \\
 &= \frac{\sum_{\alpha} N_{\alpha} e^{-\beta E_{\alpha} + \beta \mu N_{\alpha}}}{\mathcal{Z}} \\
 &= \frac{\sum_N N z^N Z(T, V, N)}{\sum_N z^N Z(T, V, N)} \\
 &= \frac{1}{\beta} \left(\frac{\partial \log \mathcal{Z}}{\partial \mu} \right)_{T, V} = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V}
 \end{aligned} \tag{4.16}$$

The particle number fluctuations can similarly be computed:

$$\begin{aligned}
 \langle (\Delta N)^2 \rangle &= \langle (N - \langle N \rangle)^2 \rangle \\
 &= \langle N^2 \rangle - \langle N \rangle^2 \\
 &= \frac{1}{\beta^2} \frac{1}{\mathcal{Z}} \left(\frac{\partial^2 \mathcal{Z}}{\partial \mu^2} \right)_{T, V} - \frac{1}{\beta^2} \frac{1}{\mathcal{Z}^2} \left(\frac{\partial \mathcal{Z}}{\partial \mu} \right)_{T, V}^2 \\
 &= \frac{1}{\beta^2} \left(\frac{\partial^2 \log \mathcal{Z}}{\partial \mu^2} \right)_{T, V} = \frac{1}{\beta} \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T, V}
 \end{aligned} \tag{4.17}$$

In a similar fashion to the energy fluctuations in the CE, we see that the relative fluctuations scale as:

$$\frac{\sqrt{\langle(\Delta N)^2\rangle}}{\langle N\rangle} \sim \frac{1}{\sqrt{N}} \quad (4.18)$$

which is incredibly small for a thermodynamically large system, and so we expect only a very small range of particle numbers meaningfully contribute to the actual computation of expectation values. Interestingly, this is not always true in the grand canonical ensemble, specifically close to a phase transition. In the vicinity of a phase transition, large (comparable to N) numbers of particles can suddenly enter or leave a given energy level, resulting in macroscopic particle number fluctuations. It also turns out that this has a significant effect on the energy fluctuations. For this, we revisit the expression from Eq. 3.14, but now take the partial derivative at constant T, V, μ instead:

$$\begin{aligned} \langle(\Delta E)^2\rangle &= -\left(\frac{\partial\langle E\rangle}{\partial\beta}\right)_{V,\mu} = k_B T^2 \left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,\mu} \\ &= k_B T^2 \left[\left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,N} + \left(\frac{\partial\langle E\rangle}{\partial\langle N\rangle}\right)_{T,V} \left(\frac{\partial\langle N\rangle}{\partial T}\right)_{V,\mu} \right] \end{aligned} \quad (4.19)$$

where we interchangeably use $N = \langle N \rangle$ since it is understood that these are generally the same quantity away from critical points. Now, we note that the opposite derivatives of the mean particle number and energy produce:

$$\left(\frac{\partial\langle N\rangle}{\partial T}\right)_{V,\mu} = \frac{1}{T} \left(\frac{\partial\langle E\rangle}{\partial\mu}\right)_{T,V} \quad (4.20)$$

using Schwarz' theorem. Plugging this into the expression for the energy fluctuations, we arrive at:

$$\begin{aligned} \langle(\Delta E)^2\rangle &= k_B T^2 \left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,N} + k_B T \left(\frac{\partial\langle E\rangle}{\partial\langle N\rangle}\right)_{T,V} \left(\frac{\partial\langle E\rangle}{\partial\mu}\right)_{T,V} \\ &= k_B T^2 \left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,N} + k_B T \left[\left(\frac{\partial\langle E\rangle}{\partial\langle N\rangle}\right)_{T,V} \right]^2 \left(\frac{\partial\langle N\rangle}{\partial\mu}\right)_{T,V} \\ &= k_B T^2 C_V + k_B T \left[\left(\frac{\partial\langle E\rangle}{\partial\langle N\rangle}\right)_{T,V} \right]^2 \langle(\Delta N)^2\rangle \end{aligned} \quad (4.21)$$

where in the second line, we use the chain rule to re-obtain the partial derivative of the particle number, before re-invoking Eq. 4.17 in the last line. This tells us that the typical energy fluctuations in the grand canonical ensemble contain the same heat capacity contribution from the canonical ensemble, but now also include a contribution from the fluctuations in particle numbers. Under normal circumstances, this fluctuation would scale as $1/\sqrt{N}$ and be highly suppressed, but this can show macroscopic deviations in the vicinity of a phase transition through the second term¹⁷. Finally, we can also link Eq. 4.1 and Eq. 4.9 to write expressions for the three equations of state:

¹⁷While we did interchangeably use N and $\langle N \rangle$ in this derivation, this relation turns out to be general, and works even in the vicinity of a critical point (near a phase transition). The proof of this is too elusive for me for now.

$$S(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial T} \right)_{V, \mu} \quad (4.22)$$

$$p(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial V} \right)_{T, \mu} \quad (4.23)$$

$$N(T, V, \mu) = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V} \quad (4.24)$$

4.3 Two-Level Systems in the GCE

Let us now return to the toy model of the two-level systems again¹⁸, and attempt to construct the problem in the grand canonical ensemble. The setup is essentially the same as before, only now we can relax the restriction on the fixed number of classical spins. We thus simply have a collection of identical, weakly-interacting classical spins. Each particle occupies a non-degenerate energy level given by Eq. 2.18. The microstates are still described by a set of values $\{n_i\} \equiv \alpha$, only now the length of this set is unspecified. We construct the grand partition function by simply using the result from the CE for a fixed number N , and write a series expansion in the fugacity as:

$$\begin{aligned} \mathcal{Z}(T, \mu) &= \sum_{N=0}^{\infty} z^N Z(T, N) \\ &= \sum_{N=0}^{\infty} e^{\beta \mu N} [e^{\beta \epsilon} + e^{-\beta \epsilon}]^N \\ &= \frac{1}{1 - 2e^{\beta \mu} \cosh(\beta \epsilon)} \end{aligned} \quad (4.25)$$

This gives us the grand potential simply as:

$$\begin{aligned} \Phi(T, \mu) &= -k_B T \log \mathcal{Z}(T, \mu) \\ &= k_B T \log [1 - 2e^{\beta \mu} \cosh(\beta \epsilon)] \end{aligned} \quad (4.26)$$

4.4 Three Ensembles Playing the Same Piece

Through the course of this review, we constructed three different (but related) statistical ensembles for solving various types of statistical mechanics problems. As examples, I demonstrated how we could approach the situations of two-level systems and harmonic oscillators in all three ensembles. At this point, you are probably sick of those examples (as am I) and may have had the natural question of how they could possibly be related given how different the expressions look and how we are holding different quantities constant in each ensemble. In this section, I aim to convince you that we have really been telling three versions of the exact same story, and we can mathematically show (in a simple enough situation, such as the two-level systems) that this is the case. For convenience, we restate the primary expression in the natural potential for all three ensembles:

¹⁸I have chosen not to revisit the harmonic oscillators in the GCE as it is a simple exercise, and explicitly typing it out does not add much more value to the discussion than simply revisiting some algebra.

$$S(E, N) = -\frac{Nk_B}{2} \left(1 + \frac{E}{N\epsilon}\right) \log \left[\frac{1}{2} \left(1 + \frac{E}{N\epsilon}\right)\right] - \frac{Nk_B}{2} \left(1 - \frac{E}{N\epsilon}\right) \log \left[\frac{1}{2} \left(1 - \frac{E}{N\epsilon}\right)\right] \quad (4.27)$$

$$F(T, N) = -Nk_B T [\log 2 + \log \cosh(\beta\epsilon)] \quad (4.28)$$

$$\Phi(T, \mu) = k_B T \log [1 - 2e^{\beta\mu} \cosh(\beta\epsilon)] \quad (4.29)$$

Firstly, the statistical ensembles gain their true power in the thermodynamic limit, so we should expect that these constructions will only agree in the $N \rightarrow \infty$ limit. To put them on common ground, we will turn our attention to the **specific entropy**, or the entropy per particle, in each of these ensembles. Mathematically, this is written as:

$$s \equiv \lim_{N \rightarrow \infty} \frac{S}{\langle N \rangle} \quad (4.30)$$

We now define the quantity $x = \frac{E}{N\epsilon}$ as the dimensionless energy, and start with the microcanonical ensemble:

$$\begin{aligned} S(E, N) &= -\frac{Nk_B}{2} (1+x) \log \left[\frac{1}{2}(1+x)\right] - \frac{Nk_B}{2} (1-x) \log \left[\frac{1}{2}(1-x)\right] \\ s &= k_B \left[\log 2 - \frac{1}{2}(1+x) \log(1+x) - \frac{1}{2}(1-x) \log(1-x) \right] \end{aligned} \quad (4.31)$$

This is the expression we should obtain from the other two ensembles. We now proceed with the canonical ensemble. Since we have to obtain the entropy, which is easily obtained from F , we first get an expression for the mean energy:

$$\begin{aligned} \langle E \rangle &= -\left(\frac{\partial \log Z}{\partial \beta} \right)_N = -N\epsilon \tanh(\beta\epsilon) \\ &\implies x = -\tanh(\beta\epsilon) \\ &\implies \beta\epsilon = \frac{1}{2} [\log(1-x) - \log(1+x)] \end{aligned} \quad (4.32)$$

Plugging this back into the mean energy, we get a self-consistent equation:

$$\langle E \rangle = \frac{Nk_B T}{2} x [\log(1-x) - \log(1+x)] \quad (4.33)$$

This gives us our bridge to the entropy expression, which we (with a little intermediate work that I am skipping)¹⁹ write as:

$$\begin{aligned} S &= \frac{\langle E \rangle - F}{T} = Nk_B \left[\log 2 - \frac{1}{2}(1+x) \log(1+x) - \frac{1}{2}(1-x) \log(1-x) \right] \\ &\implies s = k_B \left[\log 2 - \frac{1}{2}(1+x) \log(1+x) - \frac{1}{2}(1-x) \log(1-x) \right] \end{aligned}$$

which agrees with the microcanonical result! We finally turn to the grand canonical ensemble, and note that we now also have to find the mean particle number since this is no longer fixed. The mean energy and particle number are thus found through the appropriate derivatives of the grand potential:

¹⁹I promise it's not difficult, it's just a bit of manipulation using logarithms and hyperbolic functions. Typing it out is incredibly cumbersome.

$$\langle N \rangle = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V} = \frac{e^{\beta \mu} (e^{\beta \epsilon} + e^{-\beta \epsilon})}{1 - e^{\beta \mu} (e^{\beta \epsilon} + e^{-\beta \epsilon})} \quad (4.34)$$

$$\langle E \rangle = \beta \left(\frac{\partial \Phi}{\partial \beta} \right)_{V, \mu} = \frac{e^{\beta \mu} (e^{\beta \epsilon} - e^{-\beta \epsilon})}{1 - e^{\beta \mu} (e^{\beta \epsilon} + e^{-\beta \epsilon})} \epsilon \quad (4.35)$$

From this, we solve for $\beta \epsilon$ and $\beta \mu$ (this is probably the hardest step, and requires a couple pages of mathematical gymnastics which I'd recommend skipping if you're short on patience):

$$\begin{cases} \beta \epsilon &= \frac{1}{2} [\log(1-x) - \log(1+x)] \\ \beta \mu &= \log \left(\frac{\langle N \rangle}{\langle N \rangle + 1} \right) - \log 2 + \frac{1}{2} [\log(1-x) - \log(1+x)] \end{cases} \quad (4.36)$$

$$\implies \begin{cases} \langle E \rangle &= \langle N \rangle \epsilon x \\ \Phi &= -k_B T \log(\langle N \rangle + 1) \end{cases} \quad (4.37)$$

Finally, we insert this back into the expression for the entropy as obtained from the definition of the grand potential in Eq. 4.1:

$$\begin{aligned} S &= \frac{\langle E \rangle - \mu \langle N \rangle - \Phi}{T} \\ &= \langle N \rangle k_B \left[\frac{\log(\langle N \rangle + 1)}{\langle N \rangle} + \log \left(1 + \frac{1}{\langle N \rangle} \right) + \log 2 - \frac{1}{2} (1+x) \log(1+x) - (1-x) \log(1-x) \right] \\ \implies s &= k_B \left[\log 2 - \frac{1}{2} (1+x) \log(1+x) - \frac{1}{2} (1-x) \log(1-x) \right] \end{aligned}$$

which agrees with the microcanonical result in the thermodynamic limit! We thus see that all three ensembles were describing the exact same physical picture, and this is best reconciled through the entropy per particle. In my personal opinion, this demonstration is beautiful in that it shows why entropy is pretty much the most fundamental thermodynamic potential (despite the others being equivalent to it), in that it carries a great deal of easily extractable meaning.

Now that we've shown the equivalence of the three ensembles in producing the same entropy per particle (in the thermodynamic limit), let's unpack the expression a little further. From Eq. 4.31, we first regroup the constant $\log 2$ term with the other two, which brings it into better agreement with the original form from the microcanonical ensemble. This is easily done by writing:

$$\begin{aligned} \log 2 &= (1+x) \log 2 + (1-x) \log 2 \\ \implies s &= -k_B \left[\frac{1+x}{2} \log \left(\frac{1+x}{2} \right) + \frac{1-x}{2} \log \left(\frac{1-x}{2} \right) \right] \end{aligned} \quad (4.38)$$

Now, the setup of the problem tells us that E is bounded from below by $-N\epsilon$ and bounded above by $N\epsilon$. This implies that $x \in [-1, 1]$. Enacting a simple variable transformation of the form $p = (1+x)/2$, we obtain:

$$s = -k_B [p \log p + (1-p) \log(1-p)] \quad (4.39)$$

This is just the Gibbs entropy of a system with two states of probability p and $(1-p)$ respectively! The probability $p(E, N)$ here is explicitly stated in terms of the energy E and particle number N , while we are

typically more used to seeing this in terms of the Boltzmann factor $\exp(-\beta\Delta)$, where Δ is the splitting between the energy levels. This is simply reconciled by considering the relative probability:

$$\frac{p}{1-p} = \frac{1+x}{1-x} = \frac{1-\tanh(\beta\epsilon)}{1+\tanh(\beta\epsilon)} = e^{-2\beta\epsilon} \quad (4.40)$$

which is exactly the Boltzmann factor we expect for two states of energies $\pm\epsilon$! Incredible stuff Boltzmann and Gibbs, your work continues to astound us even today.

5 The Continuous Ensembles

Over the last 25 or so pages, I have demonstrated the construction and use of the various statistical ensembles for some case studies. However, I did start this by waxing poetic about Hamiltonian mechanics and phase space, though we never used any of that knowledge in the preceding discussion of the statistical ensembles. This is because our previous discussions were solely focused on discrete problems, with energies dependent on some discrete quantity (such as n_i as the energy level label). Since there is no way to define a canonical momentum conjugate for this discrete quantity²⁰, it is not possible to construct a phase space and use the ideas of Hamiltonian mechanics for such discrete problems. In this section, we will finally make use of those initial ideas, turning to the continuous formulation of all the ensembles we have previously seen. Through this, we will study the classical ideal gas in the framework of statistical mechanics and hope to reconcile the notions we are used to from thermodynamics.

5.1 Density of States

Earlier in the MCE, we defined $\Omega(E)$ as the number of microstates of a system with energy between $(E - \delta E)$ and E , where δE can be made sufficiently small compared to the precision of the instrumentation used to measure the energy. Then, we can essentially consider all states under consideration for $\Omega(E)$ to be of the same energy.

Now, we can also consider the number of microstates with energy *less than or equal to* E , denoted by $\Sigma(E)$. Intuitively, we expect that $\Sigma(E)$ is going to be a non-decreasing function of E , and any increase in $\Sigma(E)$ for increasing E represents a counting of sorts in the total number of microstates lying between E and $E + \delta E$. We thus consider the following:

$$\begin{aligned} \Omega(E) &= \Sigma(E) - \Sigma(E - \delta E) \\ &\approx \frac{\partial \Sigma(E)}{\partial E} \delta E \\ &= g(E) \delta E \end{aligned} \quad (5.1)$$

where we have defined a new quantity $g(E) = \partial \Sigma(E) / \partial E$, known as the **density of states**. The density of states effectively counts the phase space density of microstates that reside within an energy shell centred at E , and of thickness δE . That is, this represents the number of microstates we expect per unit energy, at a given energy. Since we have imposed that δE can be made arbitrarily small to lie within the precision of the probing

²⁰While it is true that continuum mechanics actually arises from discretisation in the limit of sufficiently small spacing, this limit is justified since we can insert points in position and momentum space to subdivide the space at will, eventually taking the continuum limit. This is not possible in our discrete problems such as the two-level systems and harmonic oscillators, since the energy levels are clearly defined.

instrument, the density of states can also be understood as the **degeneracy** of a given energy in a continuous setting.

5.2 The Classical MCE

To construct the microcanonical ensemble in a continuous setting, we need to take a more careful look at how the ensemble resides in phase space. Particularly, we need a way to count states in a continuous fashion, if this is possible at all. We first formalise the definition of a **phase space volume** \mathcal{V} , and distinguish it from a real space volume V as:

$$\mathcal{V} = \prod_i \int d\mathbf{q}_i d\mathbf{p}_i = V \prod_i \int d\mathbf{p}_i \quad (5.2)$$

Thus, the phase space volume is a product of a real volume and a volume in momentum space, with the region within the volume defined by the bounds placed on the canonical coordinates q and the canonical conjugate momenta \mathbf{p} ²¹. Now, as in the discrete case, we wish to obtain the microcanonical probability by selecting only a specific portion of phase space corresponding to the energy under consideration in the microcanonical ensemble. We begin by specifying some energy E_0 and selecting only the volume of phase space which corresponds to that energy:

$$\mathcal{V}(E_0) = \int d\mathbf{q} d\mathbf{p} \delta(E - E_0) \quad (5.3)$$

where the Dirac delta is enforced through the fact that the energy E is a function of (\mathbf{q}, \mathbf{p}) . For a system whose energy is conserved, we can impose that $E_0 = H(\mathbf{q}, \mathbf{p})$, where H is the Hamiltonian of the entire N -particle system. Thus, when an energy is specified for an N -particle system in continuous phase space, we have the phase space volume per unit energy as:

$$g_{\text{ph}}(E) = \int d\mathbf{q} d\mathbf{p} \delta(E - H(\mathbf{q}, \mathbf{p})) \quad (5.4)$$

The subscript ‘ph’ here distinguishes this as a phase space volume, rather than a dimensionless count of states as in the case of the density of states from before. In this continuous phase space, we know that the microcanonical probability distribution takes on the form of a probability density $\rho_{\text{mc}}(\mathbf{q}, \mathbf{p})$. Now that we have isolated the region of interest for a given energy E , we need to simply ensure that the microcanonical probability density is evenly distributed over this region to obey Boltzmann’s postulate. Knowing that the microcanonical probability should be uniform within the energy shell at E and vanish everywhere else, we expect something of the form:

$$\rho_{\text{mc}}(E; \mathbf{q}, \mathbf{p}) = \lambda \delta(E - H(\mathbf{q}, \mathbf{p})) \quad (5.5)$$

where λ is a normalisation factor to be determined. Imposing the normalisation condition tells us that:

$$\begin{aligned} 1 &= \int d\mathbf{q} d\mathbf{p} \rho_{\text{mc}}(\mathbf{q}, \mathbf{p}) = \int d\mathbf{q} d\mathbf{p} \lambda \delta(E - H(\mathbf{q}, \mathbf{p})) \\ \implies \lambda &= \left[\int d\mathbf{q} d\mathbf{p} \delta(E - H(\mathbf{q}, \mathbf{p})) \right]^{-1} = \frac{1}{g_{\text{ph}}(E)} \end{aligned}$$

²¹Recall from the first section that the vectors \mathbf{q} and \mathbf{p} are $3N$ -dimensional objects which specify the positions and momenta of all particles in the system.

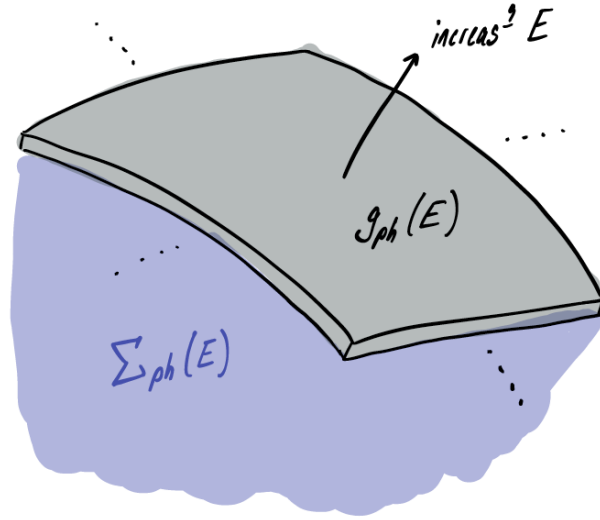


FIGURE 8: Phase Space Volume and Volume Density in the Continuous Ensembles. The grey shaded region represents the volume of phase space contained within the energy shell at E , with thickness δE (not indicated on the diagram), while the blue shaded region represents the volume of phase space with energies up to E . Note that the blue shaded region actually overlaps with the grey, but I have not done so to make it clearer to distinguish. All of phase space with energy larger than E is unshaded, since this region is not counted in either quantity. Furthermore, as all realistic Hamiltonians are convex in the canonical coordinates and momenta, we can define a directionality to the process of increasing E . That is, the shell $g_{\text{ph}}(E)$ is never expected to demonstrate any bubbles or inward bumps, so there is always a well defined outward direction corresponding to increasing energy.

This provides us with the microcanonical probability density function as²²:

$$\rho_{\text{mc}}(E; \mathbf{q}, \mathbf{p}) = \frac{1}{g_{\text{ph}}(E)} \delta(E - H(\mathbf{q}, \mathbf{p})) \quad (5.6)$$

It is also worth considering the expression for the phase space volume with energy *less than or equal to* E , analogous to $\Sigma(E)$ from before. This is:

$$\begin{aligned} \Sigma_{\text{ph}}(E) &= \int_{H(\mathbf{q}, \mathbf{p}) \leq E} d\mathbf{q} d\mathbf{p} \\ &= \int d\mathbf{q} d\mathbf{p} \theta(E - H(\mathbf{q}, \mathbf{p})) \end{aligned} \quad (5.7)$$

where we use the Heaviside theta function $\theta(x)$ to count all states with energy $H \leq E$. It is often far easier to count the number of states up to a certain energy by imposing the Heaviside theta in this integral, then differentiate $\Sigma(E)$ to obtain $g(E)$, rather than to try imposing the Dirac delta in Eq. 5.11 directly. A brief visualisation of both these quantities are shown in Fig. 8.

Now, we need to introduce some quantum mechanical ideas before things get out of hand. In phase space, we currently have a volume which satisfies certain energy requirements, and this volume is used to construct the microcanonical ensemble of interest. This volume, however, turns out to be a flawed notion due to quantum mechanics. Specifically, Heisenberg's uncertainty principle tells us that simultaneous knowledge

²²It should be noted at this point that it is quite meaningless in a physical sense to write ρ_{mc} on its own, and we should always interpret this as $\rho_{\text{mc}} d\mathbf{q} d\mathbf{p}$, which is the total probability within some infinitesimal phase space volume.

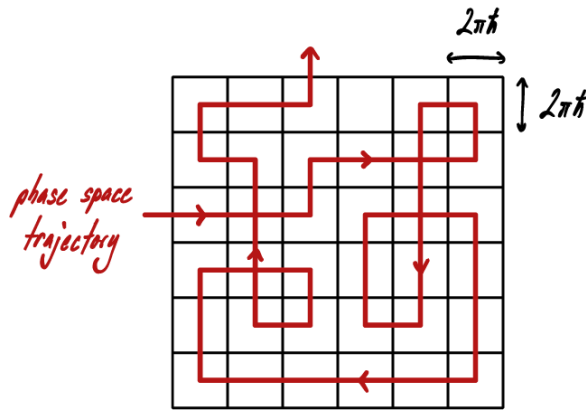


FIGURE 9: Trajectory of a System through Phase Space Cells. Subdividing phase space into unit cells of side length $(2\pi\hbar)$ along each axis (and there are $6N$ of them in total), we see that the trajectory of a system through phase space is simply a path which visits all states on an energy shell equally as often. In the ergodic equivalent, we say that any of these cells are equally likely to be occupied. Another way to view the imposition of the uncertainty principle is that along the trajectory, we have to make some measurement over a finite amount of time. In trying to keep the energy shell arbitrarily thin (δE) , we cannot make an arbitrarily short measurement, and so the system would have travelled some significant distance through phase space in that time. This ‘blurs’ out any instantaneous sampling of a phase space trajectory into a distribution. Adapted from [4].

of the exact position and momentum of a particle is impossible, and is instead constrained by the following inequality:

$$\Delta x \Delta p \geq h = 2\pi\hbar \quad (5.8)$$

with the right hand side being true up to a multiplicative factor. Thus, any set of states in phase space that are separated by coordinates and momenta with a product less than $2\pi\hbar$ ²³ correspond to the same measurable *microstate*. We can then denote the smallest meaningful volume of phase space²⁴, known as the **phase space unit cell**, by $(\Delta\mathbf{q})(\Delta\mathbf{p}) = (2\pi\hbar)^{3N}$. That is, every distinguishable microstate occupies a minimal phase space volume of $(2\pi\hbar)^{3N}$. This is our link from the continuous ensemble to the discrete ideas from before!

We now note that the phase space unit cell allows us to explicitly count multiplicities as follows:

$$\Omega(E_0) = \frac{\mathcal{V}(E_0)}{(2\pi\hbar)^{3N}} \quad (5.9)$$

This enables us to finally convert the phase space volume $\Sigma_{\text{ph}}(E)$ and volume density $g_{\text{ph}}(E)$ into a multiplicity and a multiplicity density according to:

²³One may be wondering why I keep using $2\pi\hbar$ rather than simply h in each of these appearances. My training comes from atomic physics and quantum gases, so I firmly believe that the reduced Planck’s constant is a more meaningful quantity due to its connection with oscillatory two-level systems, as in the school of I.I. Rabi. If in a position of sufficient power, I would define $\frac{1}{2}$ as the reduced Archimedes’ constant (Yes, I will die on this hill).

²⁴It might seem incredibly strange that this entire *classical* discussion has an appearance of \hbar in it at all, but this really shouldn’t surprise you. I honestly think David Tong put it best when he explained this appearance – “...it is a vestigial object, like the male nipple. It is redundant, serving only as a reminder of where we came from. And the classical world came from the quantum.”

$$\Sigma(E) = \frac{\Sigma_{\text{ph}}(E)}{(2\pi\hbar)^{3N}} = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{q} d\mathbf{p} \theta(E - H(\mathbf{q}, \mathbf{p})) \quad (5.10)$$

$$g(E) = \frac{g_{\text{ph}}(E)}{(2\pi\hbar)^{3N}} = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{q} d\mathbf{p} \delta(E - H(\mathbf{q}, \mathbf{p})) \quad (5.11)$$

which count the number of microstates of energy up to E and the number density of microstates per unit energy at E respectively. This provides us with the machinery we need to perform actual computations in the microcanonical ensemble.

5.3 The Classical Ideal Gas in the MCE

Let us now consider the simplest possible continuous ensemble construction that we can – the classical ideal gas in a box. In the ideal gas, the Hamiltonian simply has a kinetic energy term for each particle, and no potential energy (since the ideal gas is non-interacting). The Hamiltonian is thus written as:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{j=1}^{3N} \frac{p_j^2}{2m} \quad (5.12)$$

With the Hamiltonian, we can now explicitly write the total number of microstates $\Sigma(E)$ up to and including the energy E ²⁵:

$$\begin{aligned} \Sigma(E) &= \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{q} d\mathbf{p} \theta\left(E - \sum_{j=1}^{3N} \frac{p_j^2}{2m}\right) \\ &= \frac{1}{(2\pi\hbar)^{3N}} \left[\prod_{i=1}^N \int d\mathbf{q}_i \right] \int d\mathbf{p} \theta\left(E - \sum_{j=1}^{3N} \frac{p_j^2}{2m}\right) \\ &= \frac{V^N}{(2\pi\hbar)^{3N}} \int d\mathbf{p} \theta\left(E - \sum_{j=1}^{3N} \frac{p_j^2}{2m}\right) \end{aligned} \quad (5.13)$$

where we recognise that the integral of a particle's position over all coordinate values simply returns the volume of the box, for which we get N copies of V . To deal with the integrals over the momenta, we first make the variable transformation $y_j^2 = p_j^2/2mE$. This simplifies the variable in the Heaviside theta and reduces the need for any special treatment later on. We have:

$$\begin{aligned} \Sigma(E) &= \frac{V^N}{(2\pi\hbar)^{3N}} \int dp_1 \cdots dp_{3N} \theta\left(E - \sum_{j=1}^{3N} \frac{p_j^2}{2m}\right) \\ &= \frac{V^N}{(2\pi\hbar)^{3N}} (2mE)^{3N/2} \int dy_1 \cdots dy_{3N} \theta\left(1 - \sum_{j=1}^{3N} y_j^2\right) \end{aligned} \quad (5.14)$$

where we now get $3N$ copies of the factor $\sqrt{2mE}$ out front from the change of variables. The remaining integral has an interesting geometric interpretation – it is simply the volume of the unit $3N$ -dimensional hypersphere²⁶! The volume of the unit n -dimensional hypersphere is easily given by:

²⁵There's going to be a lot of subtle notational changes here, so pay attention! \mathbf{q} is a $3N$ -dimensional object detailing the three position coordinates of all N particles, \mathbf{q}_i is a three-dimensional object detailing the three position coordinates of the i -th particle only (there are N of these) and q_i is the i -th position coordinate (there are $3N$ of these in total).

²⁶It's worth taking as much time as you need to convince yourself that this is true.

$$V_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} \quad (5.15)$$

Plugging this into Eq. 5.14 quickly returns:

$$\begin{aligned} \Sigma(E, V, N) &= \frac{V^N}{(2\pi\hbar)^{3N}} (2mE)^{3N/2} \frac{\pi^{3N/2}}{(3N/2)!} \\ &= V^N \frac{E^{3N/2}}{(3N/2)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \end{aligned} \quad (5.16)$$

With this expression for $\Sigma(E)$, we proceed to obtain the density of states through a simple derivative:

$$\begin{aligned} g(E, V, N) &= \frac{\partial \Sigma(E)}{\partial E} \\ &= V^N \left(\frac{3N}{2}\right) \frac{E^{3N/2-1}}{(3N/2)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \\ &= V^N \frac{E^{3N/2-1}}{(3N/2-1)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \end{aligned} \quad (5.17)$$

This gives us the multiplicity as:

$$\begin{aligned} \Omega(E, V, N) &= g(E, V, N) \delta E \\ &= V^N \frac{E^{3N/2-1}}{(3N/2-1)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \delta E \\ &= V^N \frac{E^{3N/2}}{(3N/2)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \frac{3N}{2} \frac{\delta E}{E} \end{aligned} \quad (5.18)$$

where we quickly realise that it was a waste of time to simplify the previous derivative. Finally, with the multiplicity, we can compute the entropy:

$$\begin{aligned} S(E, V, N) &= k_B \log \Omega(E, V, N) \\ &= k_B \left[N \log V + \frac{3N}{2} \log E - \log \left(\frac{3N}{2}\right)! + \frac{3N}{2} \log \left(\frac{m}{2\pi\hbar^2}\right) + \log \left(\frac{3N}{2} \frac{\delta E}{E}\right) \right] \\ &\approx k_B \left[N \log V + \frac{3N}{2} \log E - \frac{3N}{2} \log \left(\frac{3N}{2}\right) + \frac{3N}{2} + \frac{3N}{2} \log \left(\frac{m}{2\pi\hbar^2}\right) + \log \left(\frac{3N}{2} \frac{\delta E}{E}\right) \right] \\ &\approx N k_B \left[\log V + \frac{3}{2} \log \left(\frac{E}{N}\right) + \frac{3}{2} \log \left(\frac{m}{3\pi\hbar^2}\right) + \frac{3}{2} \right] \end{aligned} \quad (5.19)$$

where in the last line, we drop the last term containing $\log(\delta E/E)$. This is because it is a factor of N smaller than all the other terms, and so it will become incredibly small in the thermodynamic limit and can be neglected.

ALERT ALERT ALERT ALERT ALERT

Now that I have your attention, allow me to tell you that Eq. 5.19 is wrong, in that it grossly overestimates the entropy of the ideal gas! A simple way to see this is to consider the extensivity of the entropy, that is, we expect that the entropy should scale up in a manner like:

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N) \quad (5.20)$$

where λ is some multiplicative factor which scales the system. In Eq. 5.19, this is mostly satisfied except in the $\log V$ term, which simply pulls out an awkward $\log \lambda$ term. It turns out that the root of our mistake lies in the initial counting of the states itself. When we write the multiplicity Ω , we need to account for the fact that the particles in the ideal gas are indistinguishable from one another, so the permutations of the particles in phase space do not contribute unique microstates. Thus, we can easily resolve this mistake by inserting the correct Boltzmann counting factor as $\Omega \rightarrow \Omega/N!$. Now, we re-evaluate the entropy with this modification:

$$\begin{aligned} S(E, V, N) &= k_B \left[N \log V - \log N! + \frac{3N}{2} \log E - \log \left(\frac{3N}{2} \right)! + \frac{3N}{2} \log \left(\frac{m}{2\pi\hbar^2} \right) + \log \left(\frac{3N}{2} \frac{\delta E}{E} \right) \right] \\ &\approx Nk_B \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{E}{N} \right) + \frac{3}{2} \log \left(\frac{m}{3\pi\hbar^2} \right) + \frac{5}{2} \right] \end{aligned} \quad (5.21)$$

where we apply Stirling's approximation to tame the new factorial term. This is known as the **Sackur-Tetrode equation**. It is extensive as required, since every appearance of E and V is paired with a reciprocal factor of N , leaving only the overall factor of N out front for the extensivity! This is our fundamental equation in the entropy representation, and all thermodynamic quantities of interest can be derived from here through appropriate derivatives. As a demonstration, let's obtain the three equations of state, starting with the inverse temperature:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V,N} = \frac{3}{2} Nk_B \frac{1}{E} \implies E = \frac{3}{2} Nk_B T \quad (5.22)$$

which is the expected result from the equipartition theorem²⁷! Next, let's obtain the pressure equation of state:

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N} = \frac{Nk_B}{V} \implies pV = Nk_B T \quad (5.23)$$

which is the famous ideal gas equation that we are used to! Finally, we can also obtain the chemical equation of state (for studying the change in particle number):

$$\frac{\mu}{T} = - \left(\frac{\partial S}{\partial N} \right)_{E,V} = -k_B \log \left[\frac{V}{N} \left(\frac{m}{3\pi\hbar^2} \frac{E}{N} \right)^{3/2} \right] \implies \mu = -k_B T \log \left[\frac{V}{N} \left(\frac{m}{3\pi\hbar^2} \frac{E}{N} \right)^{3/2} \right] \quad (5.24)$$

We've already dealt with the fact that \hbar appearing in the *classical* ideal gas fundamental equation is acceptable, but isn't it strange that \hbar then only proceeds to appear in one of the three equations of state? Specifically, the temperature and pressure equations look classical while the chemical equation of state contains the appearance of \hbar as well as m , which is a microscopic specification of the constituent gas particles. The following is speculation – I believe this may be because the chemical potential tells us about the amount of energy required to insert a new particle whilst keeping the energy and volume of the gas constant. Normally, this isn't an issue²⁸ in the microcanonical ensemble, but we start to get ridiculous results in the quantum mechanical limit. Firstly, let's plug the temperature equation of state into the chemical equation of state to write:

$$\mu = -k_B T \log \left[\frac{V}{N} \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \right] = -k_B T \log \left(\frac{V}{N\Lambda^3} \right) \quad (5.25)$$

²⁷It's alright if you haven't seen or heard of that theorem before, this should at least look familiar as the constant-volume heat capacity of an ideal gas in three dimensions.

²⁸Experimentally, this is a nightmare, but let's pretend the extent of our lab is a chalk and a chalkboard.

where we define the quantity $\Lambda = \sqrt{2\pi\hbar^2/mk_B T}$ known as the **thermal de Broglie wavelength**, or simply the thermal wavelength. This represents the typical de Broglie wavelength of a particle in an ideal gas at the temperature T . The thermal wavelength then sets a scale on the temperature T for which quantum effects become relevant, through its relative size to the mean interparticle spacing and the size of the containing box itself. Once the thermal wavelength becomes on the order of the mean free spacing between the particles ($\sqrt[3]{V/N}$), quantum mechanical effects such as the symmetrisation or antisymmetrisation of the wavefunctions become important enough to make our ideal gas approximation break down. In the regime that quantum effects start to become important, we see that the chemical potential becomes strongly negative, suggesting that the addition of new particles can lower the energy of the gas. This is clearly wrong for a gas of fermions, since the Pauli exclusion principle tells us that adding particles into occupied states should be impossible. This is not the only strange result, for instance, if we were to rewrite the entropy per particle from Eq. 5.21 using the temperature equation of state, we have:

$$s(T, V) = \frac{S(T, V, N)}{N} = k_B \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{mk_B T}{2\pi\hbar^2} \right) + \frac{5}{2} \right] \quad (5.26)$$

This clearly exhibits a logarithmic singularity in the low temperature limit, implying that the entropy per particle tends to $-\infty$ as the temperature of the gas is lowered to absolute zero. This is clearly rubbish since it violates the Nernst statement of the third law of thermodynamics, which tells us that any system being taken to absolute zero must have its entropy tend toward a finite constant. This is why the study of the classical ideal gas can only be done at reasonably high temperatures, since there is a temperature below which the entire notion of an “ideal” gas breaks down and we have to accept that reality truly is quantum.

5.4 The Classical Ideal Gas in the CE

Assuming for now that we are far from the quantum realm, we now want to revisit the classical ideal gas in the canonical ensemble. For this, we need to construct the canonical probability density function ρ_c . The probability distribution should take on a similar form to the discrete canonical probability, in that it should simply apply a Boltzmann factor to weight the probability using the energy of a microstate. We can thus immediately write the canonical probability distribution as:

$$\rho_c(\mathbf{q}, \mathbf{p}) = \frac{1}{(2\pi\hbar)^{3N}} \frac{1}{Z(T, V, N)} \exp[-\beta H(\mathbf{q}, \mathbf{p})] \quad (5.27)$$

$$\text{where } Z(T, V, N) = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{q} d\mathbf{p} \exp[-\beta H(\mathbf{q}, \mathbf{p})] \quad (5.28)$$

A simple interpretation of this distribution is that we simply take the canonical probability $e^{-\beta E}$ to be the volume of phase space occupied by a microstate in this ensemble, then divide by the phase space unit cell to obtain a multiplicity of sorts. The partition function Z then enters as a normalisation factor once again.

We now apply this to the classical ideal gas, with the same Hamiltonian as before in Eq. 5.12. The partition function is constructed as:

$$\begin{aligned}
Z(T, V, N) &= \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \left[\prod_{i=1}^N \int d\mathbf{q} \right] \left[\int d\mathbf{p} \exp \left(-\beta \sum_{i=1}^{3N} \frac{p_i^2}{2m} \right) \right] \\
&= \frac{1}{(2\pi\hbar)^{3N}} \frac{V^N}{N!} \left[\int dp_i \exp \left(-\frac{\beta p_i^2}{2m} \right) \right]^{3N} \\
&= \frac{1}{(2\pi\hbar)^{3N}} \frac{V^N}{N!} (2\pi m k_B T)^{3N/2} \\
&= \frac{V^N}{N!} \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3N/2}
\end{aligned} \tag{5.29}$$

From this, we can easily compute the Helmholtz free energy:

$$\begin{aligned}
F(T, V, N) &= -k_B T \log Z \\
&= -k_B T \left[N \log V - \log N! + \frac{3N}{2} \log \left(\frac{m k_B T}{2\pi\hbar^2} \right) \right] \\
&= -N k_B T \left\{ \log \left[\frac{V}{N} \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} \right] + 1 \right\}
\end{aligned} \tag{5.30}$$

which gives us everything we need to compute the equations of state! Let's go ahead and do so, starting with the entropy:

$$\begin{aligned}
S(T, V, N) &= - \left(\frac{\partial F}{\partial T} \right)_{V, N} \\
&= N k_B \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{m k_B T}{2\pi\hbar^2} \right) + \frac{5}{2} \right]
\end{aligned} \tag{5.31}$$

This is the Sackur-Tetrode equation in the Helmholtz representation, and is also the exact same result that we obtained above in the microcanonical ensemble! The pressure equation of state is also easily obtained as:

$$p(T, V, N) = - \left(\frac{\partial F}{\partial V} \right)_{T, N} = \frac{N k_B T}{V} \implies pV = N k_B T \tag{5.32}$$

which is the ideal gas equation that we expect! Lastly, we can also obtain the chemical equation of state:

$$\mu(T, V, N) = \left(\frac{\partial F}{\partial N} \right)_{T, V} = -k_B T \log \left[\frac{V}{N} \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} \right] \tag{5.33}$$

Thankfully, this also exactly agrees with our result from the microcanonical ensemble earlier. This confirms our intuitive view that the canonical ensemble should produce identical results to the microcanonical ensemble, with the only difference being that the canonical ensemble takes on a fluctuating energy whose mean $\langle E \rangle$ agrees with the exact microcanonical energy E to good approximation. This can be confirmed by evaluating $\langle E \rangle$ from the partition function, which will yield the same temperature equation of state as the microcanonical ensemble.

In fact, the connection between the two ensembles is made even clearer in the continuous formalism. Comparing the continuous canonical partition function with the microcanonical density of states from Eq. 5.11, we note an interesting connection:

$$Z(\beta, V, N) = \int_0^\infty dE e^{-\beta E} g(E, V, N) \tag{5.34}$$

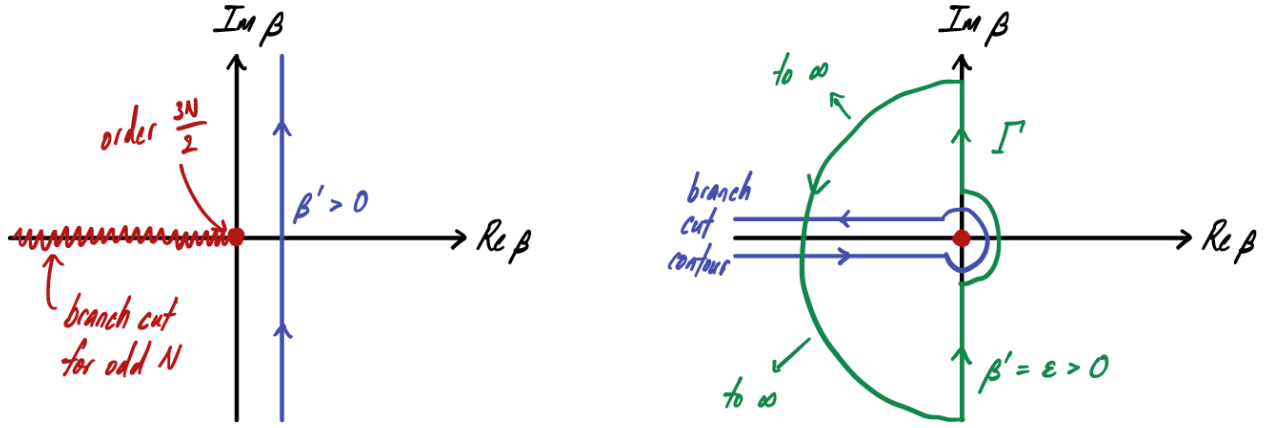


FIGURE 10: Contours for the Bromwich Integral to Obtain the Density of States. (Left) We start by placing a vertical contour on some $\beta' \in \mathbb{R}$ which lies to the right of all non-holomorphic domains of $Z(\beta, V, N)$. In this case, we have a pole of order $3N/2$ at the origin, which is also a branch point that necessitates a branch cut to $\beta \rightarrow -\infty$. (Right) If a branch cut is required, we choose the blue contour to enclose the pole at the origin whilst running along both separate holomorphic domains on either side of the branch cut. If the branch cut can be written off (as we have done by assuming the thermodynamic limit), we choose the green contour Γ which encloses the pole and is closed in the left half-plane. This is displaced from the pole by some small $\beta' = \epsilon$.

where we explicitly change the temperature dependency to β , and the integral is carried out over the domain $[0, \infty)$ since the energy is assumed to be bounded below by zero (otherwise the Laplace kernel $e^{-\beta E}$ would not converge). This means that the classical canonical partition function is simply the Laplace transform of the microcanonical density of states (with β serving as the complex frequency)! This also means the process can be inverted (in theory) to obtain the microcanonical density of states from the classical canonical partition function, though this will have to be computed using the Bromwich integral:

$$g(E, V, N) = \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta e^{\beta E} Z(\beta, V, N) \quad (5.35)$$

where $\beta' \in \mathbb{R}$ sits to the right of all poles of $Z(\beta, V, N)$. Of course, this is not the ideal contour to perform the integral along, and it should be deformed so the residue theorem can be exploited to obtain quick results. We will now attempt to compute this Bromwich integral for the classical ideal gas to obtain the microcanonical density of states (Eq. 5.17) from the classical canonical partition function in Eq. 5.29.

The basic setup reads:

$$\begin{aligned} g(E, V, N) &= \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta e^{\beta E} \frac{V^N}{N!} \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3N/2} \\ &= \frac{V^N}{N!} \left(\frac{m}{2\pi\hbar^2} \right)^{3N/2} \frac{1}{2\pi i} \int_{\beta' - i\infty}^{\beta' + i\infty} d\beta \frac{e^{\beta E}}{\beta^{3N/2}} \end{aligned}$$

As is the case in any complex integral, we need to analyse the structure of the integrand first to determine the appropriate contour. Firstly, the integrand has a pole of (thermodynamically large) order $3N/2$ at $\beta = 0$. Next, we also see that the integrand demands a branch cut from $\beta = 0$ to $\beta \rightarrow -\infty$ if N is odd. However, we can treat N as being even since it is thermodynamically large, and so we don't expect the result to change whether it is of the value 10^{23} or $10^{23} + 1$. We thus assume N is even and close the contour on the left half-plane. This allows us to invoke the residue theorem on the pole at the origin, which we evaluate as:

$$\begin{aligned}\int_{\Gamma} d\beta \frac{e^{\beta E}}{\beta^{3N/2}} &= 2\pi i \operatorname{Res} \left(\frac{e^{\beta E}}{\beta^{3N/2}} \right) \\ &= 2\pi i \frac{E^{3N/2-1}}{(3N/2-1)!}\end{aligned}\quad (5.36)$$

which is easily obtained using the limit formula for residues of higher order poles (or using a Laurent series expansion). We thus obtain the density of states as:

$$g(E, V, N) = \frac{V^N}{N!} \left(\frac{m}{2\pi\hbar^2} \right)^{3N/2} \frac{E^{3N/2-1}}{(3N/2-1)!} \quad (5.37)$$

which agrees with the microcanonical result (up to the Boltzmann counting error, which was fixed later)! While we had the slightly sketchy step of assuming N was even, it should be noted that expressing the factorial of $3N/2 - 1$ in terms of the gamma function restores meaning for odd-valued N , and this will return the correct original partition function when subject to the Laplace transform of Eq. 5.34.

5.4.1 Revisiting Fluctuations in the Canonical Ensemble

Earlier when we were studying the canonical ensemble, we found in Eq. 3.16 that the absolute size of energy fluctuations scaled with the square root of the heat capacity (the relative fluctuations scaled as $1/\sqrt{N}$). We now wish to revisit this idea in the continuous classical ensemble to further cement the notion that the canonical and microcanonical ensembles are, in practice, identical.

We first start with the fact that the canonical probability distribution is constructed as:

$$P(E) dE \propto e^{-\beta E} g(E) dE \quad (5.38)$$

This is the product of the Boltzmann factor $e^{-\beta E}$ which monotonically decreases with E , and the density of states $g(E)$ which monotonically increases with E . The product must therefore have an *extremum*²⁹ at some value of $E = E^*$. The value of E^* is easily determined by solving:

$$\left. \frac{\partial}{\partial E} [e^{-\beta E} g(E)] \right|_{E=E^*} = 0 \quad (5.39)$$

$$\implies \left. \frac{\partial \log g(E)}{\partial E} \right|_{E=E^*} = \beta \quad (5.40)$$

From our thermodynamic definition of temperature, we had:

$$\left. \frac{\partial S(E)}{\partial E} \right|_{E=\langle E \rangle} = \frac{1}{T} = k_B \beta \quad (5.41)$$

which immediately leads to the condition that $\langle E \rangle = E^*$. This is neat because it directly tells us that the most likely energy value E^* is equal to its mean value $\langle E \rangle$ that corresponds to the measurable thermodynamic internal energy. This justifies our usual trickery of interchangeably using both of these quantities. To get a better understanding of the probability distribution itself, let's expand this distribution around the maximum³⁰:

²⁹At this stage, we technically have no information on whether this extremal point is a minimum or maximum. However, knowing that this is a probability distribution which will have to be normalised, we can deduce that this must be a maximal point, otherwise there is no way to normalise the distribution (assuming it is well-behaved).

³⁰I have dropped the energy shell thickness dE here since it is physically obvious as to what we're doing. We just have to keep in mind that the units of $g(E)$ are that of $[E]^{-1}$.

$$\begin{aligned}\log [e^{-\beta E} g(E)] &= \left(-\beta \langle E \rangle + \frac{S}{k_B} \right) + \frac{1}{2} \frac{\partial^2}{\partial E^2} \log [e^{-\beta E} g(E)] \Big|_{E=\langle E \rangle} (E - \langle E \rangle)^2 + \mathcal{O}(E^3) \\ &\approx -\beta(\langle E \rangle - TS) - \frac{1}{2k_B T^2 C_V} (E - \langle E \rangle)^2\end{aligned}\quad (5.42)$$

where I have chosen to expand the logarithm of the distribution instead of the distribution itself, since this splits up the multiplicative factors into additive terms for easy expansion. We also directly use the result from Eq. 3.16 in the second derivative term. This gives us the distribution as:

$$P(E) dE \sim e^{-\beta(\langle E \rangle - TS)} \exp \left[-\frac{(E - \langle E \rangle)^2}{2k_B T^2 C_V} \right] dE \quad (5.43)$$

where the prefactor of $e^{-\beta(\langle E \rangle - TS)} = e^{-\beta F}$ is just our old friend – the canonical partition function Z ! This tells us that the normalised probability distribution $P(E)/Z$ is an overall Gaussian distribution in E with mean $\langle E \rangle$ and variance $k_B T^2 C_V$. In terms of the nondimensionalised energy $E/\langle E \rangle$, we see that the distribution has its mean at unity, and a variance which now scales as $1/\sqrt{N}$ (as we saw before in Eq. 3.16). In the thermodynamic limit $N \gg 1$, this rapidly approaches a Dirac delta distribution!

We can look at this for the specific situation of the classical ideal gas, to get a grasp of the numbers. We have already seen from Eq. 5.17 (up to the incorrect Boltzmann counting) that $g(E)$ scales as $E^{3N/2-1}$. The maximal probability thus occurs at:

$$E^* = \frac{1}{\beta} \left(\frac{3N}{2} - 1 \right) \quad (5.44)$$

On the other hand, we also know from the temperature equation of state that the internal energy is given by $\langle E \rangle = 3N/2\beta$, which is incredibly close to E^* for thermodynamically large N . An example of this situation is shown in Fig. 11 for $N = 10$.

5.5 The Classical Ideal Gas in the GCE

Finally, we turn to the grand canonical ensemble to analyse the classical ideal gas. We expect that this should return us the same results as before, only now we have the fluctuating particle number which could make for some interesting discussion. In a similar manner to the canonical ensemble, we construct the grand canonical probability density and grand partition function as:

$$\rho_{\text{gc}}(\mathbf{q}, \mathbf{p}) = \frac{1}{(2\pi\hbar)^{3N}} \frac{e^{\beta\mu N}}{\mathcal{Z}(T, V, \mu)} \exp[-\beta H(\mathbf{q}, \mathbf{p})] \quad (5.45)$$

$$\text{where } \mathcal{Z}(T, V, \mu) = \frac{1}{(2\pi\hbar)^{3N}} \sum_{N=0}^{\infty} \int d\mathbf{q} d\mathbf{p} e^{\beta\mu N} \exp[-\beta H(\mathbf{q}, \mathbf{p})] \quad (5.46)$$

which now includes the fugacity factor z^N and a sum over the particle numbers N , in contrast to the classical canonical partition function which just integrates over phase space. With the same Hamiltonian from Eq. 5.12, we now construct the grand partition function starting from the previously obtained canonical partition function in Eq. 5.29:

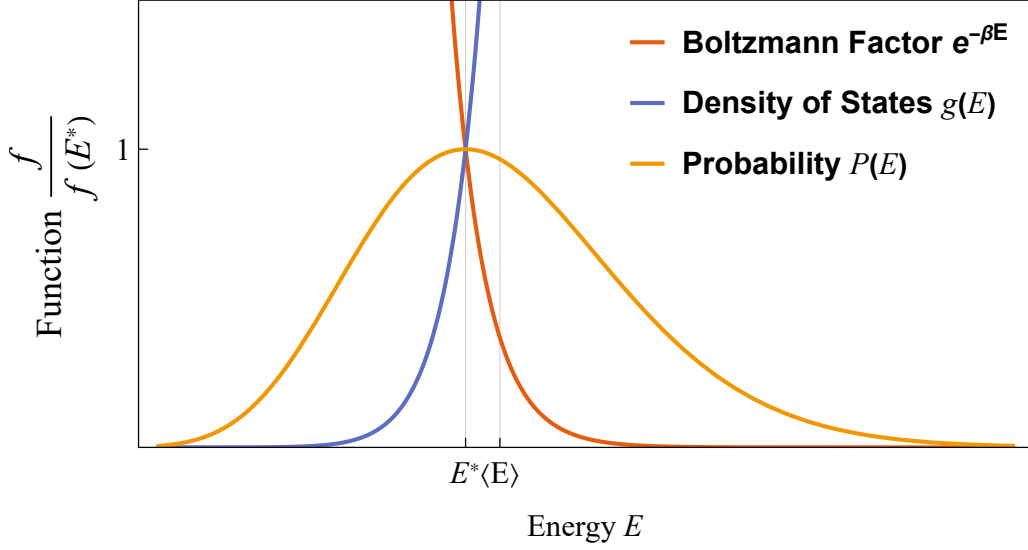


FIGURE 11: Behaviour of the Boltzmann factor, density of states and their product. In this plot, the two factors and their product (probability) is shown normalised to their respective values at $E = E^*$. This plot shows the situation the classical ideal gas with $N = 10$, so there is a noticeable deviation between E^* and $\langle E \rangle$, but this rapidly shrinks in the limit of larger N . The skewness of the distribution is also due to the small value of N , and will appear as a $O(E^3)$ correction. Adapted from [5].

$$\begin{aligned}
 \mathcal{Z}(T, V, \mu) &= \sum_{N=0}^{\infty} e^{\beta\mu N} Z(T, V, N) \\
 &= \sum_{N=0}^{\infty} \frac{1}{N!} \left[V e^{\beta\mu} \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \right]^N \\
 &= \exp \left[V \exp \left(\frac{\mu}{k_B T} \right) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \right]
 \end{aligned} \tag{5.47}$$

From the grand partition function, we can then easily obtain the grand potential:

$$\Phi(T, V, \mu) = -k_B T \log \mathcal{Z}(T, V, \mu) = -k_B T V \exp \left(\frac{\mu}{k_B T} \right) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \tag{5.48}$$

This easily gives us the various equations of state through appropriate derivatives, which you can try on your own. I would like to focus specifically on the particle number equation of state:

$$\langle N \rangle (T, V, \mu) = - \left(\frac{\partial \Phi}{\partial \mu} \right)_{T, V} = V \exp \left(\frac{\mu}{k_B T} \right) \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \tag{5.49}$$

which tells us that we can write the grand partition function simply as:

$$\mathcal{Z} = \exp(\langle N \rangle) \tag{5.50}$$

Now, let's consider the probability for a system in the grand canonical ensemble to have N_s particles. This is written:

$$P(N_s) = \frac{1}{\mathcal{Z}} e^{\beta\mu N_s} Z(T, V, N_s) \tag{5.51}$$

which is essentially the grand canonical probability, after integrating out all other variables except the particle number. Comparing our expression for $\langle N \rangle$ against the classical canonical partition function in Eq. 5.29, we identify that the probability can be written as:

$$\begin{aligned} P(N_s) &= \exp(-\langle N \rangle) \frac{1}{N_s!} \langle N \rangle^{N_s} \\ &= \text{Pois}(k = N_s; \lambda = \langle N \rangle) \end{aligned} \quad (5.52)$$

which is a Poisson distribution! This also agrees with the fact that the number fluctuations evaluate to give:

$$\langle (\Delta N)^2 \rangle = \frac{1}{\beta} \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V} = \langle N \rangle \quad (5.53)$$

which is characteristic for a Poisson distribution. We thus see that the classical ideal gas in the grand canonical ensemble follows Poissonian statistics for its number distribution, in that the number of particles in a partitioned subsystem will demonstrate fluctuations dictated by the Poisson distribution³¹.

³¹For those of you who took PHY 304 and saw my review on quantum optics at the end of the course, you may recognise this Poissonian fluctuation as being characteristic of the photon field. This is not a coincidence, and it is for this same reason that the photon gas (blackbody radiation) is often called the *ideal quantum gas*. It is the only quantum gas that truly resides in the $\mu = 0$ regime.

References

1. Callen, H. B. **Thermodynamics and an Introduction to Thermostatistics** 493. ISBN: 0471862568 (Wiley, 1985).
2. Reif, F. **Fundamentals of Statistical and Thermal Physics** 651. ISBN: 9781577666127 (Waveland Press, 2009).
3. Gibbs, J. W. **Elementary Principles in Statistical Mechanics** 224. ISBN: 9780486789958 (Dover Publications, Incorporated, 2014).
4. Englert, B.-G. **Lectures on Statistical Mechanics** (World Scientific, 2020).
5. Pathria, R. K. & Beale, P. D. **Statistical Mechanics** ISBN: 9780081026922 (Elsevier Science & Technology, 2021).
6. Sekerka, R. F. **Thermal Physics: Thermodynamics and Statistical Mechanics for Scientists and Engineers** 610. ISBN: 9780128033043 (Elsevier Science & Technology Books, 2015).
7. Garanin, D. **Statistical Thermodynamics - Spring 2017 Lecture Notes** 2017. https://www.lehman.edu/faculty/dgaranin/teaching-Statistical_Thermodynamics-Spring-2021.php.
8. Dalvit, D. A. R., Frastai, J. & Lawrie, I. D. **Problems on Statistical Mechanics** 284. ISBN: 0750305215 (Institute of Physics Pub., 1999).