CHAPTER I

INTRODUCTION

Classical Statistical Mechanics[1]

Phase space

Consider a system of N molecules each with s degrees of freedom. The motion of these molecules is described by Hamilton's equations of motion,

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad i = 1,\ldots, Ns \quad \text{----}(1.1)$$

which constitute 2 Ns first order differential equations in 2 Ns unknowns. Once all the values of the p_i and q_i are known at any one time the motion of these molecules is completely known for all time. Thus, in principle, the macroscopic properties of any system can be determined by a detailed investigation of the microscopic motion of the molecules constituting the system. However, if the system consists of a large number of molecules such a procedure quickly becomes highly complecated and loses its value. For this reason, in order to obtain a description of the macroscopic properties of such a system, we shall adopt a quite different procedure.

---

[1] Henry Eyring, Douglas Henderson, Betsy Jones Stover, and Edward M. Eyring, Statistical Mechanics and Dynamics (New York: John Wiley and sons, Inc, 1964), pp. 80 - 110

The behavior, in time, of this system can be graphically represented by means of a single trajectory in a 2 Ns dimensional space which is determined by the 2 Ns coordinates p_1, ..., p_{Ns}, q_1,... q_{Ns}. At any time the instantaneous state of the system is specified by the position of a representative point in this space. This space is called the phase space or Γ space of the system. Sometimes a phase space is used to represent the motion of a single molecule contained in the system. Such a phase space is called μ space.

The macroscopic properties of the system are thus time averages over a segment of the trajectory in phase space. Instead of following this procedure, it is much more advantageous to follow the suggestion of Gibbs. He suggested that instead of taking time averages we consider a large number of similar systems, each consisting of N molecules and Ns degrees of freedom, but whose representative points in phase space are suitably random so that every state accessible to the actual system in the course of time is represented by at least one system at any instant of time. Such a collection of similar systems is called an ensemble. We assume that there are so many systems in the ensemble that we can speak of the density $D(p_i, q_i, t)$ with which the representative points are distributed in phase space. Gibbs' suggestion was that we replace time averages over a single system by ensemble averages over the ensemble at a fixed time. Thus the average value of a function A(p,q) is,

$$\overline{A(p_i,q_i)} = \frac{\int A(p_i,q_i)\, D(p_i,q_i)\, dq_1 \ldots dq_s}{\int D(p_i,q_i)\, dq_1 \ldots dp_s} \qquad \text{-----(1.2)}$$

Liouville's theorem

We shall now consider how the density of representative points changes with time. In considering how $D(p_i, q_i, t)$ changes with time we must distinguish between its rate of change with respect to a fixed point in phase space, which is $\partial D/\partial t$ since p_i and q_i will not vary, and its rate of change in a coordinate system moving with a particle through phase space, dD/dt. These two rates of change are related by

$$\frac{dD}{dt} = \frac{\partial D}{\partial t} + \sum_i \left(\frac{\partial D}{\partial p_i}\dot{p}_i + \frac{\partial D}{\partial q_i}\dot{q}_i\right) \qquad \text{-----(1.3)}$$

we shall prove that $dD/dt = 0$. This is "Liouville's Theorem". Thus the density in phase space remains constant in the neighborhood of a particle which is moving in phase space. An equivalent statement of this theorem is that if $\delta\omega'$ is the volume occupied at time $t + dt$ by the representative points which occupied a volume $\delta\omega$ at time t then

$$\delta\omega = \delta\omega' \qquad \text{-----(1.4)}$$

It is this property of conservation of volume in phase space during the motion that distinguishes phase space from any other space in which the motion of the system could be represented and which makes phase space important in the study of statistical mechanics. As we shall see, the fact that Hamilton's equations hold for the coordinates of phase space gives it this unique property.

If the system to be represented by an ensemble is in thermal equilibrium, then the ensemble averages must be independent of time, since the macroscopic properties of a system in thermal equilibrium do not change with time. Therefore, we may reasonably require that the composition of the ensemble be independent of time. In other words,

$$\frac{\partial D}{\partial t} = 0 \qquad \text{-----(1.5)}$$

An ensemble satisfying Eq.(1.5) is called a <u>stationary ensemble</u> and we shall in general be concerned with this type of ensemble. It can readily be shown that if D is a function of the energy ϵ alone the ensemble is stationary. We shall find it convenient to work with a normalized density of representative points

$$\rho\,(p_i, q_i) = \frac{D(p_i, q_i)}{\int D(p_i, q_i)\,dq_i \ldots dp_s} \qquad \text{-----(1.6)}$$

The Maxwell-Boltzmann distribution

We consider a system consisting of N molecules, each molecule having s degrees of freedom. Each molecule will then have a representative point in the 2s dimensional μ space. We divide μ space into a number of cells each of whose volume is extremely small so that all molecules in the ith cell have the same energy ϵ_i. Let the number of allowed states associated with the energy ϵ_i (i.e. the degeneracy of ϵ_i) be g_i. Further we suppose that the state of the system is such that there are n_i molecules whose representative points lie in the ith cell.

We now calculate the number of ways of realizing this situation. In order to make this calculation, let us first calculate the number of ways of putting n_1 objects of N objects in one box, then n_2 objects out of $N-n_1$ in a second box, and so on until we have exhausted all of the objects. The number of ways of choosing n_1 objects out of N objects is given by

$$\frac{N\,!}{(N-n_1)!\,n_1!} \qquad \text{----(1.7)}$$

and the number of ways of choosing n_2 objects out of $N-n_1$ objects is

$$\frac{(N-n_1)!}{(N-n_1-n_2)!\,n_2!} \qquad \text{----(1.8)}$$

And so the number of ways of achieving this arrangement is

$$\frac{N!}{(N-n_1)!n_1!} \cdot \frac{(N-n_1)!}{(N-n_1-n_2)!n_2!} \cdots = \frac{N!}{n_1!n_2!\ldots} \quad \text{---(1.9)}$$

If all the $g_i = 1$ this would be the number of ways that the system can be in a macroscopic state such that there are n_i molecules whose representative points are in the ith cell. But if g_i is different from unity then we must multiply Eq. (1.9) by $g_i^{n_i}$. Hence the number of ways of achieving a macroscopic state of the system such that n_i molecules have representative points in the ith cell is

$$W = N! \prod_i \frac{g_i^{n_i}}{n_i!} \quad \text{----(1.10)}$$

We shall assume that the total number of molecules in the system is extremely large so that the number of molecules n_i whose representative points are in each cell is also large. Under these conditions we can regard W as a continuous function of the n_i. We seek the most probable distribution of the n_i which we assume to be overwhelmingly more probable than any other distribution. To do this we must determine the values of n_i which make Eq.(1.10) a maximum subject to the constraining conditions:

$$\sum_i n_i = N \ , \ \sum_i n_i \epsilon_i = E \quad \text{-----(1.11)}$$

By calculation, we obtain the most probable distribution.

$$n_i = \frac{N g_i e^{-\beta\epsilon_i}}{\sum_i g_i e^{-\beta\epsilon_i}} \qquad \text{-----(1.12)}$$

Since we have divided μ space into regions of extremely small volume to insure that every molecule in each cell has the same energy ϵ_i, it is more convenient to express this result in the differential form. We do this by making the transitions

$$\begin{aligned} \epsilon_i &\longrightarrow \epsilon\,(p,q) \\ g_i &\longrightarrow dq_1 \ldots d\,p_s = d\,\omega \end{aligned} \qquad \text{-----(1.13)}$$

Since g_i is proportional to the volume of the ith cell. Hence, we obtain for the number of molecules, dn, whose representative points lie in the region of μ space defined by the interval $(q_i, q_i + dq_i)$, $(p_i, p_i + dp_i)$ the following,

$$dn = \frac{N e^{-\beta\epsilon(p,q)} dq_1 \ldots dp_s}{\int e^{-\beta\epsilon(p,q)} dq_1 \ldots dp_s} \qquad \text{-----(1.14)}$$

We can calculate the value of β, and obtain that

$$\beta = 1/kT \qquad \text{-----(1.15)}$$

where k is Boltzmann's constant.

Thus, Eq.(1.14) becomes

$$dn = \frac{N e^{-\epsilon(p,q)/kT} d\omega}{\int e^{-\epsilon(p,q)/kT} d\omega} \qquad \text{----(1.16)}$$

Entropy in statistical mechanics

As a preliminary to this discussion let us calculate the entropy of one mole of a perfect gas from the combined first and second laws of thermodynamics

$$T\,dS = dE + p\,dV \qquad \text{----(1.17)}$$

But for one mole of a perfect gas

$$E = C_V T \qquad pV = RT \qquad \text{----(1.18)}$$

Hence,

$$S = C_V \ln T + R \ln V + C \qquad \text{----(1.19)}$$

where C is a constant. If we have one mole each of two different perfect gases contained in vessels of equal volume at the same temperature and pressure, but separated by a partition, then the entropies of the two gases are

$$S_A = C_V \ln T + R \ln V + C$$

$$S_B = C_V \ln T + R \ln V + C \qquad \text{-----(1.20)}$$

If the partition is removed and the gases allowed to diffuse into one another, then the entropy of the mixture is

$$S_{A+B} = 2\,C_V \ln T + 2\,R \ln 2\,V + 2\,C$$

$$= 2C_V \ln T + 2R \ln V + 2R \ln 2 + 2C \qquad \text{--(2.21)}$$

$$S_{A+B} = S_A + S_B + 2R \ln 2 \qquad \text{-----(2.22)}$$

Thus, from purely thermodynamic reasoning we see that the entropy of the mixture is greater than the combined entropy of the separated gases.

As a result of the diffusion of the different perfect gases we have lost a certain degree of knowledge about the gases, since before the partition was removed, knowledge of the position of molecule also gave knowledge of its species whereas this is no longer true after the partition is removed. In other words, an element of randomness has been introduced into the system. Because of this it is resonable to postulate a relationship between the entropy of a system and the randomness or degree of disorder of the system in the given state. Thus,

$$S = f(W) \qquad \text{-----}(1.23)$$

where W is the number of a priori equally probable states accessible to the system, which is proportional to the total volume of accessible phase space.

If we consider two separate systems with entropies S_1 and S_2 then,

$$S_1 = f(W_1)$$

$$S_2 = f(W_2) \qquad \text{-----}(1.24)$$

Thermodynamics tells us that the entropy of the combined system is

$$S = S_1 + S_2 \qquad \text{----}(1.25)$$

But, since the systems are independent, the number of a priori equally probable states is

$$W_1W_2 \qquad \text{-----}(1.26)$$

Thus

$$S_1+S_2 = f(W_1W_2) \qquad \text{-----}(1.27)$$

The only functional relationship that can exist between S and W satisfying Eq.(1.27) is

$$S = k' \ln W + C \qquad \text{-----}(1.28)$$

where k′ and C are some constants. But we assume that W is one at absolute zero, then the third law of thermodynamics allows us to set C equal to zero. Let us now evaluate k′. From Eq.(1.10) we have

$$W = \sum_{(n_i)} W(n_i) = \sum_{(n_i)} N!\prod_i \frac{g_i^{n_i}}{n_i!} \qquad \text{-----}(1.29)$$

the summation being over all possible sets of the n_i. But since the most probable distribution of the n_i is overwhelmingly more probable than any other, only the most probable distribution of the n_i will contribute to W so that

$$W = N!\prod_i \frac{g_i^{n_i}}{n_i!} \qquad \text{-----}(1.30)$$

where

$$n_i = \frac{Ng_i\, e^{-\epsilon_i/kT}}{\sum_i g_i\, e^{-\epsilon_i/kT}}$$

Thus

$$S = k' \ln W = k' \left[\ln N! + \sum_i (n_i \ln g_i - \ln n_i!)\right] \quad \text{---(1.31)}$$

Applying Stirling's approximation, we obtain

$$S = k' \left[N \ln \sum_i g_i e^{-\epsilon_i/kT} + \frac{E}{kT}\right] \quad \text{-----(1.32)}$$

Thus

$$\left(\frac{\partial S}{\partial E}\right)_V = \frac{k'}{kT} \quad \text{-----(1.33)}$$

But thermodynamics, informs us that

$$\left(\frac{\partial S}{\partial E}\right)_V = \frac{1}{T} \quad \text{-----(1.34)}$$

and hence we have that k' is Boltzmann's constant, and this gives the famous Boltzmann relation

$$S = k \ln W \quad \text{-----(1.35)}$$

The microcanonical ensemble

Let us consider an isolated system composed of large number, N, of molecules having total energy ϵ_0. Since the system is isolated the energy ϵ_0 is a constant, independent of time. In order to obtain the macroscopic properties of this system, we must find an appropriate ensemble of similar systems. The representative points of such an ensemble must be uniformly distributed in the accessible regions of phase space (to give a priori probabilities to equal volumes of phase space). The ensemble described by the density

$$\rho(\epsilon) = \text{constant} \quad \epsilon_0 \leq \epsilon \leq \epsilon_0 + \delta\epsilon \qquad \text{-----}(1.36)$$

$$\rho(\epsilon) = 0 \qquad \text{elsewhere}$$

satisfies these requirements. Such an ensemble is called a microcanonical ensemble. Since ρ is a function of ϵ alone the microcanonical is a stationary ensemble. Thus the microcanonical ensemble is appropriate to represent a system of known energy in a marcoscopically steady state.

Since the meaning of Eq.(1.36) is that the energy of the system is constant and that equal volumes of accessible phase space have equal a priori probabilities, the microcanical ensemble yields no more information than we obtained by more elementary means in previous section.Nevertheless Eq. (1.36) will prove useful in deriving the density for canonical ensemble which allows for a relaxation of the condition of ϵ being constant.

The canonical ensemble

Following Gibbs, the canonical ensemble is defined by the density,

$$\rho(\epsilon) = C\, e^{-\epsilon/kT} \qquad \text{-----}(1.37)$$

ϵ being the energy of the entire system. As we shall see the canonical ensemble can be used to represent a system in the thermal contact with a heat reservoir.

In order to prove this point, we consider a microcanonical ensemble representing a very large isolated system. We wish to consider the behavior of a constituent subsystem of this isolated system which is in thermal equilibrium with the rest of the system. We shall denote the properties of the subsystem by the subscript s, properties of the remainder of the total system (which acts as a heat reservoir) by the subscript r, and the properties of the total system by the subscript t. Since the total system is isolated its energy, $\epsilon_i = \epsilon_s + \epsilon_r$, is constant. As the total system is part of a microcanonical ensemble, the probability $d\, p_t$ that the representative point of the total system is in an element $d\, \Omega_t$ of phase space is

$$d\, p_t = C'\, d\, \Omega_t \qquad \text{-----}(1.38)$$

for $\epsilon_t \leqslant \epsilon \leqslant \epsilon_t + d\, \epsilon_t$ and

$$d\, p_t = 0 \qquad \text{-----}(1.39)$$

otherwise, C' being a constant. We may write

$$d\, \Omega_t = d\, \Omega_s\, d\, \Omega_r \qquad \text{-----}(1.40)$$

where $d\, \Omega_s$ contains only the coordinates and momenta belonging to the subsystem and $d\Omega_r$ contains only the coordinates and momenta of the reservoir. Hence

$$d\, p_t = C'\, d\, \Omega_s\, d\, \Omega_r \qquad \text{-----}(1.41)$$

for $\epsilon_t \leqslant \epsilon \leqslant \epsilon_t + d\, \epsilon_t$ and

$$d\, p_t = 0 \qquad \text{-----}(1.42)$$

otherwise.

We seek the probability $d\,p_s$ that the representative point of the subsystem is in the volume element $d\,\Omega_s$, without specifying the condition of the reservoir but still requiring that ϵ_t be a constant. Thus,

$$d\,p_s = C'\,d\,\Omega_s\,\Delta\,\Omega_r \qquad \text{----(1.43)}$$

where $\Delta\,\Omega_r$ is the volume of phase space accessible to the representative point of the reservoir if the representative point of the subsystem is in $d\,\Omega_s$. Let us evaluate $\Delta\,\Omega_r$. Since each accessible state of the reservoir is associated with the volume h^m, the entropy of the reservoir is

$$S_r = k \ln \frac{\Delta\,\Omega_r}{h^m} \qquad \text{----(1.44)}$$

where m is the number of degrees of freedom of the reservoir, and thus,

$$\Delta\,\Omega_r = h^m e^{S_r/k} \qquad \text{----(1.45)}$$

Since the subsystem is small in comparison with the total system we have that $\epsilon_r << \epsilon_t$ and so, using a Taylor's expansion, we have

$$S_r(\epsilon_r) = S_r(\epsilon_t - \epsilon_s) = S_r(\epsilon_t) - \left(\frac{\partial S_r}{\partial \epsilon_r}\right)_{\epsilon_r = \epsilon_t} \epsilon_s + \ldots\ldots \qquad \text{----(1.46)}$$

Thus, from(1.45)

$$\Delta\,\Omega_r = h^m e^{S_r/k} \exp\left[-(\partial S_r/\partial \epsilon_r)_{\epsilon_r = \epsilon_t}\,\epsilon_s/k\right] \quad \text{----(1.47)}$$

But from the thermodynamics; we have, since $\epsilon_r \sim \epsilon_t$,

$$\frac{\partial S_r}{\partial \epsilon_r} = \frac{1}{T} \qquad \text{-----(1.48)}$$

where T is the absolute temperature of every part of the system. Then,

$$\Delta \Omega_r = h^m e^{S_r/k} (e^{-\epsilon_s/kT}) \qquad \text{-----(1.49)}$$

and hence the probability that the representative point of the subsystem be in the volume $d \Omega_s$ is

$$d p_s = C e^{-\epsilon_s/kT} d \Omega_s \qquad \text{-----(1.50)}$$

where C is a constant determined by the normalization condition

$$C \int e^{-\epsilon_s/kT} d \Omega_s = 1 \qquad \text{-----(1.51)}$$

Equation (1.50) justifies our statement that the canonical ensemble represents a system in thermal contact with a heat reservoir.

The averge value of property P(p,q) over a canonical distribution is then

$$\bar{P} = \frac{\int P(p,q) e^{-\epsilon(p,q)/kT} d\Omega}{\int e^{-\epsilon(p,q)/kT} d\Omega} \qquad \text{-----(1.52)}$$

The strong resemblance between the canonical ensemble and the Maxwell - Boltzmann distribution function Eq.(1.16)

Should be noted. Indeed, we could look upon the canonical ensemble as consisting of a system of 'molecules' which are weakly interacting.

In the case where the system of interest contains a large number of particles then nearly all the systems in the ensemble have energies which differ only slightly from the average of the ensemble. This makes it possible to use canonical ensembles to represent systems of interest which have defined energies.

One further property of the canonical ensemble is of interest. If two canonical ensembles, representing two systems in thermal contact, are coupled together, the resulting ensemble is again a canonical ensemble. Let the properties of the first system be denoted by the index 1 and the corresponding properties of the second system by the index 2 then,

$$\ln \rho_1 = \ln C_1 - (\epsilon_1/kT)$$

$$\ln \rho_2 = \ln C_2 - (\epsilon_2/kT) \qquad \text{-----}(1.53)$$

and adding

$$\ln \rho_1 \rho_2 = \ln C_1 C_2 - \left[(\epsilon_1 + \epsilon_2)/kT\right] \qquad \text{------}(1.54)$$

so that with $\rho = \rho_1 \rho_2$, $C = C_1 C_2$ and $\epsilon = \epsilon_1 + \epsilon_2$

we have

$$\ln \rho = \ln C - (\epsilon/kT) \qquad \text{-----}(1.55)$$

showing that the coupled ensemble is also canonical.

The grand canonical ensemble

We imagine the system under consideration to be embedded not only in a heat bath but also in a 'particle bath' and ridding ourselves of the restriction of a constant number of particles we shall obtain the grand canonical ensemble.

We consider a large isolated system t, whose total number of particles and energy is constant, which is represented by a microcanonical ensemble. Let s be a subsystem of the large system which can exchange particles and energy with the remainder, r, the total system. We have,

$$\epsilon_t = \epsilon_s + \epsilon_r$$

$$N_t = N_s + N_r \qquad \text{-----(1.56)}$$

As in our treatment of the canonical ensemble we have for the probability, $d\,p_s$, that representative point of the subsystem is in a volume $d\,\Omega_s$ of its own phase space without specifying the condition of the reservoir

$$d\,p_s = C'\,d\,\Omega_s\,\Delta\,\Omega_r \qquad \text{-----(1.57)}$$

where $\Delta\,\Omega_r$ is again the volume of phase space accessible to the representative point of the reservoir when the representative point of the subsystem is in $d\,\Omega_s$. The entropy of the reservoir is

$$S_r = k \ln (\Delta\,\Omega_r/h^m) \qquad \text{-----(1.58)}$$

m being the number of degrees of freedom of the reservoir. We note that since $N_s << N_r$, m is well defined. Thus,

$$\Delta \Omega_r = h^m e^{S_r/k} \qquad \text{-----(1.59)}$$

But expanding in a Taylor's series using the fact that the subsystem is small compared to the total system, i.e. $N_s << N_t$ and $E_s << E_t$, we have

$$\begin{aligned} S_r(\epsilon_r, N_r) &= S_r(\epsilon_t - \epsilon_s,\ N_t - N_s) \\ &= S_r(\epsilon_t, N_t) - \epsilon_s \left(\frac{\partial S_r}{\partial \epsilon_r}\right)_{\epsilon_t, N_t} - N_s \left(\frac{\partial S_r}{\partial N_r}\right)_{\epsilon_t, N_t} \end{aligned} \qquad \text{-----(1.60)}$$

In eq. (1.60) ϵ_s and N_s behave as differential quantities because of the small size of the subsystem. But from thermodynamics we have,

$$\left(\frac{\partial S_r}{\partial \epsilon_r}\right)_V = \frac{1}{T} \qquad \text{-----(1.61)}$$

and we define the chemical potential μ by

$$T\, ds = d\,E + p\, dV - \mu\, d\,N \qquad \text{-----(1.62)}$$

Thus,

$$\frac{\partial S}{\partial N} = -\mu/T \qquad \text{-----(1.63)}$$

and we get for $d\,p_s$,

$$d\,p_s = C\, e^{(\mu N_s - \epsilon_s)/kT}\, d\,\Omega_s \qquad \text{-----(1.64)}$$

where C is a normalization constant. We are therefore led to the grand canonical ensemble, which is defined by

$$\rho\,(N, \epsilon) = C\, e^{(N\mu - \epsilon)/kT} \qquad \text{-----(1.65)}$$

The normalization is given by

$$\sum_{N=0}^{\infty} \int \rho\, d\Omega = 1 \qquad \text{-----(1.66)}$$

and the average value of quantity A(p,q) over a grand canonical ensemble is defined by

$$\bar{A} = \sum_{N=0}^{\infty} \int A(p,q)\, \rho\, d\Omega \qquad \text{-----(1.67)}$$

If more than one molecular species is present we replace

$N\,\mu$ by $\Sigma_i\, N_i\, \mu_i$.

Suppose two grand canonical ensembles, representing two systems engaged in thermal and particle exchange, are coupled together, then, if we represent the two systems by the indices 1 and 2 and assuming, for convenience, that only one molecular species is present in each system, we have,

$$\ln \rho_1 = \ln C_1 + \frac{N_1 \mu_1}{kT} - \frac{\epsilon_1}{kT} \qquad \text{---}$$

$$\ln \rho_2 = \ln C_2 + \frac{N_2 \mu_2}{kT} - \frac{\epsilon_2}{kT} \qquad \text{-----(1.68)}$$

Adding,

$$\ln \rho_1 \rho_2 = \ln C_1 C_2 + \frac{N_1 \mu_1 + N_2 \mu_2}{kT} - \frac{\epsilon_1 + \epsilon_2}{kT} \quad \text{----(1.69)}$$

So that with $\rho = \rho_1 \rho_2$, $C = C_1 C_2$ and $\epsilon = \epsilon_1 + \epsilon_2$ the resulting ensemble is a grand canonical ensemble. If both systems have the same species of molecule, then the resulting ensemble is a grand canonical ensemble only if $\mu_1 = \mu_2$, which is another way of arriving at the well known equilibrium condition.

From the above considerations we can conclude that the grand canonical ensemble is appropriate to represent an open system in equilibrium and at constant temperature.