LECTURE 2
## BASIC CONCEPTS
States of a System

Let's consider how we specify the state of a system with N particles in both classical mechanics and quantum mechanics.

In classical mechanics if we have a single particle in one dimension then we can describe the system completely by specifying the position coordinate q and the momentum coordinate p. We can represent this graphically by labeling one axis with q and one axis with p:
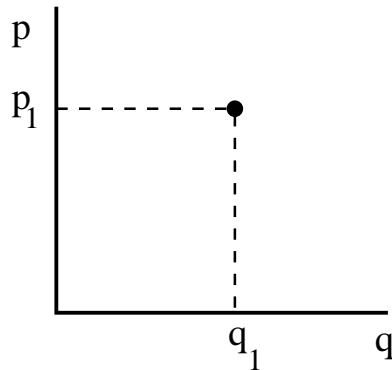
We call the space spanned by p and q "phase space." It's the space in which the point (q_1, p_1) exists. If our system has N particles and exists in 3D, then we must provide $\vec{p}_i$ and $\vec{q}_i$ for all N particles. Since $\vec{p}$ and $\vec{q}$ are each 3 dimensional vectors, if we want to represent the system as a point in phase space, we need 6N axes. In other words phase space is 6N dimensional. The coordinates of the point representing the system in phase space are $(q_{x1}, q_{x2}, ..., q_{zN}, p_{x1}, ...p_{zN})$.

Spatial coordinates and momenta are continuous variables. To obtain a countable number of states, we divide phase space into little boxes or cells. For our one particle in 1D example, the volume of one of these cells is

$$\delta q \, \delta p = h_o \tag{1}$$

where h_o is some small constant having the dimensions of angular momentum. The state of the system can then be specified by stating that its coordinate lies in some interval between q and $q + \delta q$ and between p and $p + \delta p$. For our N particle system, $f = 3N$ spatial coordinates and f momentum coordinates are required to specify the system. So the volume of a cell in phase space is

$$\delta q_1 ... \delta q_f \, \delta p_1 ... \, \delta p_f = h_o^f \tag{2}$$

Each cell in phase space corresponding to a state of the system can be labeled with some number. The state of a system is provided by specifying the number of the cell in phase space within which the system is located.

One microscopic state or *microstate* of the system of N particles is defined by specifying all the coordinates and momenta of each particle. An N particle system at any instant of time is specified by only one point in a 6N dimensional phase space and the corresponding microstate by the numerical label of the cell in which this point is located. As the system evolves in time, the coordinates of the particles change, and the N particle system follows some trajectory in this 6N dimensional phase space.

A macroscopic state or *macrostate* of the system is determined by only a few macroscopic parameters such as temperature, energy, pressure, magnetization, etc. Note that a macrostate contains much less information about a system than a microstate. So a given macrostate can correspond to any one of a large number of microstates.

How would quantum mechanics be used to describe a system? Any system of N interacting particles can be described by a wavefunction

$$\psi_{\{n\}}(q_1, ..., q_f) \tag{3}$$

where the q_i are the appropriate "coordinates" for the N particles. The coordinates include both spin and space coordinates for each particle in the system. A particular state (or a particular wavefunction) is then specified by providing the values of a set of quantum numbers $\{n\}$. This set of quantum numbers can be regarded as labelling this state. Different values of the quantum numbers correspond to different states. For simplicity let's just label the states by some index r, where $r = 1, 2, 3, ...$ The index r then labels the different microstates. In quantum mechanics, h_o is replaced by Planck's constant h.

Thus both classical and quantum mechanics lead to a countable number of microstates for an N particle system.

Ensemble

Statistical mechanics is based on probability considerations, and averages over appropriate quantities. Thus one approach to statistical mechanics is a consideration of a large number of identically prepared systems, all subject to the same initial conditions and the same set $\{X\}$ of external parameters such as the total energy, particle number, and volume. This hypothetical collection of identical systems is called an *ensemble*. The systems in the ensemble will, in general, be in different states and will, therefore, also be characterized by different macroscopic parameters (e.g., by different values of pressure or magnetic moment). (These macroscopic parameters are not the ones in set $\{X\}$, but they may be conjugate to them. For example pressure p is conjugate to volume V because $p\,dV$ is the work done by pressure p in changing the volume by dV.) We can calculate the probability of the occurrence of a particular value of such an external parameter, i.e., we can determine the fraction of cases in the ensemble when the parameter assumes this particular value. For example we can calculate the average pressure. Another way to say this: any variable or property that we are attempting to calculate will then be obtained from an averging procedure over all members of the ensemble. The aim of theory will be to predict the probability of occurrence in the ensemble of various values of such a parameter on the basis of some basic postulates.

This concept, and the term enemble, were introduced by J. W. Gibbs, an American physicist around the turn of the 20th century.

Another basic approach to statistical mechanics, proposed by Boltzmann and Maxwell, is known as the *ergodic hypothesis*. According to this view, the macroscopic properties of a system represent averages taken over the microstates traversed by a single system in the course of time. It is supposed that systems traverse *all* the possible microstates fast enough that the time averages are identical with the averages taken over a large collection of identical and independent systems, i.e., an ensemble. This is the idea behind Monte Carlo simulations.

Basic Postulates of Statistical Mechanics

To make any progress, we need some basic postulates about the relative probability of finding a system in any of its accessible states. Usually one only has partial information about a system. We don't know every single thing about every particle. The states which are compatible with the information that we have about the system are called "accessible states." The accessible states don't violate or contradict any information that we have about the system. Now consider a thermally isolated system. It cannot exchange energy with the system so its total energy is fixed or conserved. We would like to make some statements about the system in equilibrium. When the system is in equilibrium, things are not changing in time. The macroscopic parameters are time–independent. The probability of finding the system in any one state is independent of time, i.e., the representative ensemble is the same irrespective of time. This leads to the fundamental postulate of statistical mechanics:

An isolated system in equilibrium is equally likely to be in any of its accessible states. In other words if phase space is subdivided into small cells of equal size, then an isolated system in equilibrium is equally likely to be in any of its accessible cells.

Certainly this seems reasonable. There is no reason for one microstate to be preferred over another, as long as each microstate is consistent with the macroscopic parameters. There are also more rigorous reasons to accept this postulate. It's a consequence of Liouville's theorem (see Appendix 13 of Reif) that if a representative ensemble of such isolated systems are distributed uniformly over their accessible states at any one time, then they will remain uniformly distributed over these states forever.

One can think of the accessible states as the "options" that a system has available to it. Lots of accessible states means lots of possible microstates that the system can be in.

What if an isolated system is not equally likely to be found in any of the states accessible to it? Then it is not in equilibrium. But it will approach equilibrium. The system will make transitions between all its various accessible states as a result of interactions between its constituent particles. Once the system is equally likely to be in any of its accessible states, it will be in equilibrium, and it will stay that way forever (at least as long as it is isolated). The idea that a nonequilibrium system will approach equilibrium is a consequence of the H theorem (Appendix 12 in Reif). If we think in terms of an ensemble of systems distributed over the points in phase space in some arbitrary way, then the ensemble will evolve slowly in time until phase space is uniformly occupied.

The characteristic time associated with attaining equilibrium is called the "relaxation time." The magnitude of the relaxation time depends on the details of the system. The relaxation time can range from less than a microsecond to longer than the age of the universe (e.g., glass). Indeed the glass transition is a good example of a system falling out of equilibrium because the experimenter cannot wait long enough for the system to equilibrate. Calculating the rate of relaxation toward equilibrium is quite difficult, but once equilibrium is reached and things become time–independent, the calculations become quite straightforward. For example, many of the properties of the early universe have been calculated using the assumption that things were in equilibrium.

Probability calculations

From this basic postulate, how do we calculate various quantities of interest? Let us consider a system of total energy between E and $E + \delta E$. Let $\Omega(E)$ be the total number of microstates that satisfy this condition. Suppose that $\Omega(E; y_k)$ is the number of states contained within $\Omega(E)$ that are characterized by the parameter y having the value y_k. For example, y might be the magnetic moment of the system or the pressure exerted by the system. Since all states are equally likely, we have for the probability $P(y_k)$ that the parameter y of the system assumes the value y_k

$$P(y_k) = \frac{\Omega(E; y_k)}{\Omega(E)} \tag{4}$$

To calculate the mean value of the parameter y of the system, we simply take the average over the systems in the ensemble; i.e.,

$$
\begin{aligned}
\overline{y} &= \sum_k P(y_k) \, y_k \\
&= \sum_k \frac{\Omega(E; y_k)}{\Omega(E)} \, y_k \\
&= \frac{\sum_k \Omega(E; y_k) \, y_k}{\Omega(E)}
\end{aligned}
\tag{5}
$$

Here the sum over k denotes a sum over all possible values which the parameter y can assume. Note that to calculate the average value of the parameter y, we simply need to count states. However, this may be highly nontrivial.

Density of States

Density of states is a useful concept. A macroscopic system, like a cup of coffee or a block of copper, has a great many degrees of freedom. Let E be the energy of the system. Suppose we divide up the energy scale into small regions, each of size δE, where δE is much larger than the spacing between energy levels but macroscopically small. Let $\Omega(E)$ be the number of states whose energy lies between E and $E + \delta E$. Then $\Omega(E)$ must be proportional to δE and we can write

$$\Omega(E) = \rho(E)\delta E \tag{6}$$

where $\rho(E)$ is the "density of states". (Your book writes it as $w(E)$.) The density of states is a characteristic property of the system which measures the number of states per unit energy range. For example one could have the number of states per eV. The density of states is an important concept in systems with many particles. For example in a simple metal where electrons conduct electric current, the density of electron states near the Fermi energy determines how good a conductor the metal is. If the density of states is high, then the metal is a good conductor because the electrons near the Fermi energy will have a lot of empty states to choose from when they hop. If $\rho(E)$ is small, then the metal is a poor conductor because the electrons will not have many empty states to hop to. The density of states is often useful in converting sums into integrals over energy:

$$\sum_i f_i \rightarrow \int dE \rho(E) f(E) \tag{7}$$

Interaction Between Macroscopic Systems

Macroscopic systems are described by specifying some macroscopically measurable independent parameters like the volume V or the applied external electric and magnetic field. Now consider two macroscopic systems A and A' which can interact with one another so that they can exchange energy. Their total energy $E + E'$ remains constant since the combined system A^o consisting of A and A' is isolated. They can interact in two ways: mechanically and/or thermally. If they interact thermally, they exchange heat but the energy levels of the systems do not change though their occupation might. If they interact mechanically, the external parameters (like volume) are changed and some of the energy levels are shifted.

Thermal Interaction

Let's consider the thermal interaction. In a purely thermal interaction, energy is transferred from one system to the other. If we have an ensemble of interacting systems (A+A'), the amount of energy transferred to each system A is not exactly the same for the different members of the ensemble. We can however talk in terms of the change in the *mean* energy of each of the systems. This is called "heat." More precisely, the change $\Delta \overline{E}$ of the mean energy of system A is called the "heat Q absorbed" by this system; i.e.,

$$Q = \Delta \overline{E} \tag{8}$$

Heat is energy transfer. The heat can be positive or negative. $-Q$ is the heat given off by a system; Q is the heat absorbed by the system. Since the total energy is unchanged

$$\Delta \overline{E} + \Delta \overline{E}' = 0 \tag{9}$$

or

$$Q + Q' = 0 \tag{10}$$

or

$$Q = -Q' \tag{11}$$

This is just conservation of energy.

Mechanical Interaction

Now suppose that the systems A and A′ cannot interact thermally, i.e., they are thermally isolated. However they can interact mechanically. For example A′ could do work on A. Consider a cylinder separated into two parts by a movable piston. Let A be the gas in one part and A′ be the gas in the other part. Suppose A′ expands, moves the piston and compresses the gas in A. This changes the energy of A by heating up the gas. As before we think of an ensemble of identical systems and speak in terms of the change in the mean energy. If the change in the mean energy due to the change of the external parameters is denoted by $\Delta_x \overline{E}$, then the macroscopic work done *on* the system is defined as

$$\mathcal{W} = \Delta_x \overline{E} \tag{12}$$

The macroscopic work W done *by* the system is the negative of $\mathcal{W}$:

$$W = -\mathcal{W} = -\Delta_x \overline{E} \tag{13}$$

Conservation of energy dictates that

$$W + W' = 0 \tag{14}$$

or

$$W = -W' \tag{15}$$

Doing work on a system changes the positions of the energy levels and the occupation of different states.

Generalized Force

In introductory physics we defined work as the force on an object times the distance it moves ("force times distance"). Now we have a system with 10^{23} particles. How do we define work? "Pressure times volume." We mentioned earlier that when something changes the volume by applying pressure, the mean energy changes and work has been done on the system. Pressure has units of force per unit area. The gas pushes on a wall and produces pressure which is the force per unit area on the wall. Notice that F/A has the same units as energy/volume. In fact the definition of pressure is

$$p = -\frac{\partial \overline{E}}{\partial V} \tag{16}$$

(We should keep the entropy fixed in this derivative.)

We can make this more formal. When we say "macroscopic work," we mean more than pdV or Fdx where F is a force. Let the energy of some microstate r depend on external parameters $x_1, ..., x_n$.

$$E_r(x_1, ..., x_n) \tag{17}$$

Then when the parameters are changed by infinitesimal amounts, the corresponding change in energy is

$$dE_r = \sum_{\alpha=1}^{n} \frac{\partial E_r}{\partial x_\alpha} \, dx_\alpha \tag{18}$$

The work dW done *by* the system when it remains in this particular state r is then defined as

$$dW_r \equiv -dE_r = \sum_{\alpha} X_{\alpha,r} \, dx_\alpha \tag{19}$$

where

$$X_{\alpha,r} \equiv -\frac{\partial E_r}{\partial x_\alpha} \tag{20}$$

This is called the "generalized force" conjugate to the external parameter x_α in the state r. Note that if x_α denotes a distance, then X_α simply is an ordinary force.

Consider an ensemble of similar systems. If the external parameters are changed quasi–statically so that the system remains in equilibrium at all times, then we can calculate the mean value averaged over all accessible states r

$$dW = \sum_{\alpha=1}^{n} \overline{X}_\alpha \, dx_\alpha \tag{21}$$

where

$$\overline{X}_\alpha \equiv -\overline{\frac{\partial E_r}{\partial x_\alpha}} \tag{22}$$

is the mean generalized force conjugate to x_α. Note that

$$
\begin{aligned}
\overline{X}_\alpha &\equiv -\overline{\frac{\partial E_r}{\partial x_\alpha}} \\
&= \sum_{r=1}^{N} P_r \left(-\frac{\partial E_r}{\partial x_\alpha} \right) \\
&= -\frac{\partial}{\partial x_\alpha} \sum_{r=1}^{N} P_r E_r \\
&= -\frac{\partial}{\partial x_\alpha} \overline{E} \tag{23}
\end{aligned}
$$

The macroscopic work W resulting from a finite quasi–static change of external parameters can then be obtained by integration.

Examples

1. Force times distance

$$dW = F_x \, dx \tag{24}$$

   where x is the linear dimension and F_x is the ordinary force in the x direction.

2. Pressure times volume

$$dW = \overline{p}\, dV \tag{25}$$

where $\overline{p}$ is the average pressure and V is volume. We wrote down the expression for pressure before, but now we can be more precise. The pressure is the generalized force associated with changes in volume.

$$\overline{p} = -\overline{\frac{\partial E_r}{\partial V}} = \sum_{r=1}^{N} P_r \left(-\frac{\partial E_r}{\partial V} \right) = -\frac{\partial}{\partial V} \sum_{r=1}^{N} P_r E_r = -\frac{\partial}{\partial V} \overline{E} \tag{26}$$

or

$$\overline{p} = -\frac{\partial \overline{E}}{\partial V} \tag{27}$$

where E is the macroscopic energy and V is the volume.

Your book talks about quasi–static processes in which the process occurs so slowly that the system can be regarded as being in equilibrium throughout. For example, the piston can be moved so slowly that the gas is always arbitrarily close to equilibrium as its volume is being changed. In this case the mean pressure has a well–defined meaning. If the volume is changed by an infinitesimal amount dV, then the work done is

$$dW = \overline{p}dV \tag{28}$$

If the volume is changed from an initial volume V_i to a final volume V_f, then the macroscopic amount of work done is given by

$$W_{if} = \int_{V_i}^{V_f} dW = \int_{V_i}^{V_f} \overline{p}dV \tag{29}$$

This integral depends on the path taken from the initial to the final volume. It is not path independent. So dW is not an exact differential. (Recall in electromagnetism, the potential difference is path independent.) dW is not the difference of 2 numbers referring to 2 neighboring macrostates; rather it is characteristic of the process of going from state i to state f. Similarly the infinitesimal amount of heat dQ absorbed by the system in some process is also not an exact differential and in general, will depend how the process occurs.

General Interaction between 2 Systems

In general two systems interact both thermally and mechanically. Let Q be the heat absorbed *by* the system and let W be the work done *by* the system. Then the change in the mean energy $\Delta \overline{E}$ is given by

$$\Delta \overline{E} = Q - W \tag{30}$$

This is the first law of thermodynamics. If we write

$$Q = \Delta \overline{E} + W \tag{31}$$

then we can view the heat Q as the mean energy change not due to a change in the external parameters. For infinitesimal changes, we can write

$$dQ = d\overline{E} + dW \tag{32}$$

Note that $d\overline{E}$ is an exact differential. The change in the mean energy is independent of the path taken between the initial and final states. The energy is characteristic of the state, not of the process in getting to that state.

For example, suppose we push a cart over a bumpy road to the top of a hill. Let us suppose there are 2 roads to the top of the hill. How much work we do and how much is lost to friction and heat depends on which road we take and how long the road is. However, at the end of our journey at the top of the hill, the (potential) energy is independent of the road we chose. This is why dQ and dW are inexact differentials but dE is an exact differential.

Note that if $dQ = 0$, $d\overline{E} = -dW$ is an exact differential. So if $Q = 0$, then $\Delta\overline{E}_{if} = -W_{if}$. On the other hand, if $dW = 0$, $d\overline{E} = dQ$ and dQ is an exact differential. So if $W = 0$, $\Delta\overline{E}_{if} = Q_{if}$.

Exact and Inexact Differentials

We must now make a small digression to remind ourselves of the difference between exact and inexact differentials. Consider any function of two independent variables $F(x, y)$. Then the differential dF is defined by

$$
\begin{aligned}
dF &= F(x + dx, y + dy) - F(x, y) = \frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy \\
&= A(x, y)dx + B(x, y)dy \tag{33}
\end{aligned}
$$
$$\tag{34}$$

and

$$\Delta F = F_f - F_i = \int_i^f dF = \int_i^f (A\, dx + B\, dy) \tag{35}$$

Note that the integral of an exact differential depends only on the endpoints (initial and final points) and not on the path of integration.

However, not every function is an exact differential. Consider

$$dG = A'(x, y)\, dx + B'(x, y)\, dy \tag{36}$$

It is *not* guaranteed that there will exist a function $G(x, y)$ such that

$$dG = G(x + dx, y + dy) - G(x, y) \tag{37}$$

That is, it is not always true that

$$\int_i^f dG \tag{38}$$

is independent of the path between the endpoints. The integral may depend on the path of integration. As an example, consider
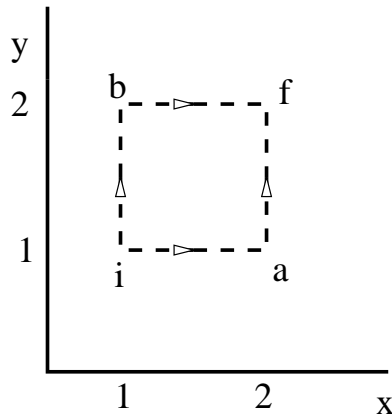
$$dG = \alpha dx + \beta \frac{x}{y} \, dy = \alpha dx + \beta x d(\ln y) \tag{39}$$

It is easy to show that

$$\int_{i \to a \to f} dG = \alpha + 2\beta \ln 2 \tag{40}$$

and

$$\int_{i \to b \to f} dG = \alpha + \beta \ln 2 \tag{41}$$

Note however that if

$$dF \equiv \frac{dG}{x} = \frac{\alpha}{x} dx + \frac{\beta}{y} dy \tag{42}$$

then dF is an exact differential with

$$F = \alpha \ln x + \beta \ln y \tag{43}$$

and

$$\int_i^f dF = \int_i^f \frac{dG}{x} = (\alpha + \beta) \ln 2 \tag{44}$$

independent of path. The factor $1/x$ is called an integrating factor for dG.

10