

Lecture 1: Introduction to Statistical Physics and Course Description

1 Basic concepts

The great success of physics in the past four centuries has been the formulation of equations which describe the motion of individual particles or objects. This is the basis of the study of dynamics whether it be classical, quantum or relativistic. In practical situations however one is frequently confronted with a multitude of particles all interacting with each other. In principle, the complete dynamical solution for this complex system can be worked out from first principles and then solved using a computer. However when the number of particles (or degrees of freedom more generally) exceeds a certain level this approach becomes completely impractical. Moreover one is often not at all interested in the location of an individual particle from this assembly of objects. Instead one is usually interested in only the “total” behavior of the assembly. For example, the total energy of the system or other gross properties. A concrete example of this situation is provided by a gas volume. For a unit volume ($1m^3$) this consists of around 6×10^{23} molecules colliding with both the container walls and themselves. We are not interested in the position of an individual molecule instead we are interested in things like the “temperature” which is proportional to its total energy.

We need therefore a method for passing from the individual description of the system to the large scale and useful description of the entire system. This is the central practical problem of statistical physics. Rather interestingly it is often the case that the laws governing the large scale are rather universal and not particularly dependent on the specific dynamics controlling individual particles. Such laws are known as the system thermodynamics. Clearly however we need to relate those laws in some way to what is happening to the individual particles. Given that the large scale quantities of interest are typically averages of basic dynamical variables for the individual particles this process is fundamentally statistical in nature. Statistical physics is the method for making this transition and thereby justifying the empirical laws of thermodynamics. The subject had two great pioneers in the late 19th and early 20th century namely Ludwig Boltzmann in Austria and Josiah Willard Gibbs in the US. Their conceptual framework still stands today as a great (and unfinished) achievement in mathematical physics.

2 Microstates

Consider a large number of particles and assume initially that their dynamics are controlled by the classical equations of Newton, Lagrange and Hamilton. The state of such a system can be specified precisely at a particular time by specifying each particle’s position and momentum. From these the future entire

state of the system can be determined precisely¹. The set of coordinates involved for a system of N particles may be written as

$$(\mathbf{q}, \mathbf{p}) \equiv (q_1, q_2, \dots, q_N, p_1, p_2, \dots, p_N)$$

This space of dimension $2N$ is known as the phase space of the system and the coordinates specify the system microstate. Note that when one makes the transition to describing the system using quantum mechanics this precise description is no longer possible since the uncertainty principle there asserts that

$$\Delta q_i \Delta p_i \geq \hbar/2$$

which implies that the volume elements of the phase space have an effective lower limit of

$$V_{min} = (\hbar/2)^{2N}$$

This effectively fixes a lower limit of resolution for microstates. It is also, as we shall see below, involved in the definition of entropy, a central concept in statistical physics. In this course we shall only deal in passing with quantum mechanics in order to keep the presentation clear. Quantum mechanics can become important under certain circumstances which we do not consider for pedagogical reasons.

The above microstate description of a system could refer to an entire closed system or to a (open) subsystem of a much larger closed system. The latter subsystem approach proves more useful ultimately as we shall see later in the course.

Now consider a very long time realization of this system. In general, given sufficient time T we expect the system will pass through any specified range of coordinates

$$\begin{aligned} &(q_i, q_i + \Delta q) \\ &(p_i, p_i + \Delta p) \end{aligned}$$

a certain number of times. Suppose the total amount of time spent in this subregion is Δt then we can consider the limit

$$w \equiv \lim_{T \rightarrow \infty} \frac{\Delta t}{T}$$

We assume that such a limit exists in general and regard it as the probability the system will be in the specified phase space volume at a particular time. By refining the phase space volume we can define a limiting phase space probability density function ρ :

$$dw = \rho(\mathbf{q}, \mathbf{p}) dq dp$$

where clearly we require

$$\int \rho(\mathbf{q}, \mathbf{p}) dq dp = 1$$

¹Notice that we need more than just the particle positions since Newton's law of motion is a second order differential equation.

Note that the definition for ϱ above does not depend on the particular initial conditions chosen since it refers to a limiting infinite time interval T in order to describe the density. A vital problem we shall face is working out what the functional form of ϱ is precisely. Once we have it however we will then have a statistical as opposed to deterministic description of the system and this will constitute what is called equilibrium statistical mechanics. A central contribution of Gibbs was in deducing that the form of ϱ was as follows

$$\varrho = C \exp(-\alpha_i I_i(\mathbf{q}, \mathbf{p})) \quad (1)$$

where I_i are the time invariant quantities for the dynamical system and the α_i real numbers. Such a deduction will be looked at more closely in Lecture 2. It has been verified empirically in a large range of different systems and the distribution is commonly called the Gibbs distribution or measure. In many circumstances the only invariant of importance turns out to be the system energy.

3 Macrostates

Now the large scale variables of interest f for our subsystem, for example the subsystem energy, will be functions of the microstate coordinates:

$$f = f(\mathbf{q}, \mathbf{p})$$

and these will fluctuate as time progresses and we can regard this variable then as a random variable with statistics controlled by the density ϱ . Thus the long time average of this random variable will for instance be given by

$$\bar{f} = \int f(\mathbf{q}, \mathbf{p}) \varrho(\mathbf{q}, \mathbf{p}) d\mathbf{q} d\mathbf{p}$$

If the actual value of f is very close to \bar{f} then we say that the system is in equilibrium. If the system is externally perturbed in such a way as to modify f then empirically if we leave the system alone for a sufficient time then $f \rightarrow \bar{f}$ and the time involved is termed the non-equilibrium relaxation time. Note also that because the system is open the system energy can fluctuate however the mean value is fixed. For “large” systems in equilibrium, as we shall discuss in Lecture 2, the fluctuations in energy relative to this mean value are very small i.e. the distribution for energy is very sharply peaked at the mean value. Such mean values of large scale quantities specify the so-called macrostates of the system. They are also commonly called thermodynamical quantities and we shall discuss them further in Lectures 3 and 4. The fluctuations about the mean values are also of interest and these will be discussed further in Lectures 7 and 10.

4 Entropy

The microstates are related to the macrostates in a fundamental way through the concept of entropy. We follow a heuristic and intuitive definition of this

concept due to Landau and Lifshitz: Consider the case that ϱ depends only on the total energy E of the system and thus macrostates are defined by specifying \bar{E} . As the equilibrium subsystem evolves in time the total energy will execute a tight random path very close to the mean value \bar{E} . In general the microstates (\mathbf{q}, \mathbf{p}) will be confined to a very particular sub-volume ΔV of the total $2N$ dimensional space in order that this be the case.

We can derive an expression for ΔV by considering the one dimensional density $W(E)^2$ which gives the probability that the energy of the system lies in some range of values. Since the full system density ϱ is constant if the energy is held at a fixed value then we can write

$$W(E_0)\delta E = \varrho(E(\mathbf{p}, \mathbf{q}))V_{2N-1}(E(\mathbf{p}, \mathbf{q}) = E_0)\delta E \quad (2)$$

where V_{2N-1} is the $2N - 1$ dimensional volume of the usually compact sub-manifold of phase space given by the constraint $E(\mathbf{p}, \mathbf{q}) = E_0$ and δE is some small range of energy about E_0 . The volume mentioned usually increases with the energy defining it. Now the energy density $W(E)$ is very sharply peaked meaning that energy values are typically very close to \bar{E} and the typical range of such values ΔE about the peak value can be defined by

$$W(\bar{E})\Delta E = 1$$

For a Gaussian³ W it is trivial to show that $\Delta E = \sqrt{2\pi}sd(E)$. Using (2) we can define a sub-volume of phase space where the system typically resides by

$$\Delta V = V_{2N-1}(E(\mathbf{p}, \mathbf{q}) = E_0)\Delta E$$

and evidently now

$$\Delta V = \frac{1}{\varrho(\bar{E})} \quad (3)$$

Thus intuitively ΔV is a measure of the “number” of microstates which are consistent with a given macrostate. One problem with such a measure is that it is a dimensional quantity with the same units as action⁴ raised to the power N . Given that quantum mechanics sets a fundamental lower limit on allowable phase space volumes it is conventional to consider instead the dimensionless quantity:

$$\frac{\Delta V}{(2\pi\hbar)^N} \quad (4)$$

This statistical weight, is used to define the macrostate entropy via the equation

$$S = \log \frac{\Delta V}{(2\pi\hbar)^N}$$

²Note that this density is different from $\varrho(E(\mathbf{p}, \mathbf{q}))$ which is a density for the full phase space i.e. it is of dimension $2N$

³This is usually a very good approximation for the fluctuation distribution of energy due to the central limit theorem.

⁴Units of action are Joule-second.

Such a definition can be easily extended unchanged conceptually to the quantum rather classical case. Entropy can be thought of then intuitively as the logarithm of the number of microstates consistent with a given macrostate.

It is interesting to consider a different expression for it given the deduction of Gibbs that

$$\log \varrho(E) = a + bE$$

It follows from (3) and (4) that since the logarithm of ϱ is linear in E that⁵

$$\begin{aligned} S &= -\log \left[\overline{(2\pi\hbar)^N \varrho(E)} \right] \\ &= - \int dqdp \varrho \log \varrho - N \log (2\pi\hbar) \end{aligned}$$

the first term as we shall see in the next lecture is the information theoretic entropy of Shannon. This deep connection between information theory and statistical physics shall be emphasized in this course. For equilibrium systems this entropy is also referred to as the Gibbs entropy.

Suppose we have a series of subsystems making up a larger closed system. How should one define the statistical weight and hence entropy of the complete system? Given that in any subsystem any microstate consistent with the macrostate could be chosen a little thought shows that if molecules in different subsystems are approximately statistically independent, the total statistical weight should be simply the product of the individual weights since one is really considering the tensor product of phase spaces. The definition of entropy as the logarithm of the statistical weight then ensures that the total entropy of the entire system is the sum of the entropies of the individual subsystems. Such a property of a variable in a compound system is desirable and makes it additive. The same property also obviously holds for the conserved quantity of energy as well as the particle number.

Entropy was originally introduced in the empirical subject of thermodynamics by Clausius in 1865. It was a central concept in this discipline. Boltzmann showed how it related to a molecular view of matter via the idea of microstates⁶. He thus provided a physical justification for an earlier empirical concept.

5 Course overview

The laws governing macrostates are known collectively as thermodynamics. A comprehensive overview of this large scale description is provided in Lectures 3 and 4. The justification of these laws is provided by the subject of equilibrium statistical mechanics which we study in Lectures 5 and 6. As we saw in the previous section systems in equilibrium do not have absolutely steady large

⁵The overbar is the expectation with respect to the Gibbs density

⁶The definition of entropy as proportional to the logarithm of the number of microstates consistent with a given macrostate is inscribed on his tombstone in the Central Cemetery of Vienna.

scale variables. These fluctuations are studied in Lectures 7 and 10. Systems not in equilibrium are less well understood. Nevertheless some progress has been made on this rather important subject which we study in Lectures 8 through 11. The course has been prepared from a number of well known texts in the area which include:

1. Statistical Physics by Landau and Lifshitz
2. Equilibrium and Non-equilibrium Statistical Mechanics by Mazenko (2 volumes).
3. Foundations of Statistical Mechanics. A deductive treatment by Oliver Penrose.
4. Equilibrium and Non-equilibrium Statistical Mechanics by C.M. Van Vliet

The approach we follow is relatively non-rigorous as an intuitive introduction is the aim. There are however rigorous presentation of this field. Two examples are

1. Entropy, Large Deviations, and Statistical Mechanics by Ellis
2. Time's Arrow by Mackey

The latter reference focusses primarily on non-equilibrium behaviour and dynamical systems while the former is a treatment of equilibrium systems.