

Chapter 1

Introduction

Statistical mechanics is the science that deals with macroscopic behavior of bodies—solids, gases, liquids, . . . —that consist of a great many components, of a great many atoms or molecules, for example. A detailed mechanical description of such a body would be impractical, even impossible, quantum mechanically. Moreover, such a detailed approach would yield the appropriate macroscopic description in terms of temperature, pressure, . . . , for example, only at the end of a laborious calculation. We will see that thermodynamics follows readily from an equilibrium statistical description of a system.

Many physicists contributed to the development of statistical mechanics, most notably Maxwell and Boltzmann. But the greatest synthesis was that of an American, J. Willard Gibbs, who in 1902 published a slim treatise, *Elementary Principles in Statistical Mechanics*, putting the subject on a firm and conceptually clear foundation that does not have to be modified even when quantum mechanical systems are discussed.

For definiteness, let us consider a classical system with s degrees of freedom. That is, the system is described by s generalized coordinates,

$$q_i = q_1, q_2, q_3, \dots, q_s, \quad (1.1)$$

and s conjugate momenta,

$$p_i = p_1, p_2, p_3, \dots, p_s. \quad (1.2)$$

We will regard the system as described by a Hamiltonian $H = H(q_i, p_i)$, where the Hamilton's equations of motion are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \quad (1.3)$$

For one degree of freedom, consider the example of a particle of mass m moving in a potential V . The Hamiltonian is

$$H = \frac{p^2}{2m} + V(q), \quad (1.4)$$

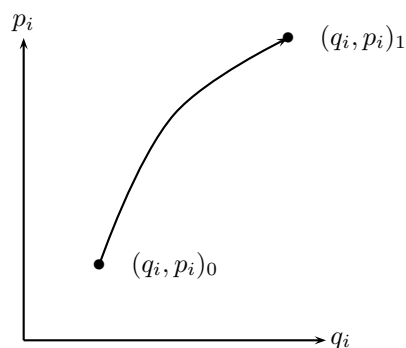


Figure 1.1: Trajectory of a system starting from a point in phase space $(q_i, p_i)_0$ at time t_0 and evolving in time to another point in phase space $(q_i, p_i)_1$ at time t_1 .

and the resulting Hamilton's equations are

$$\dot{q} = \frac{p}{m}, \quad \dot{p} = -V'(q), \quad (1.5)$$

simply Newton's law.

To specify the state of this system, one must give the value of each coordinate q_i and of each momentum p_i . More precisely, because these are continuous variables, we would say q_i lies between q_i and $q_i + dq_i$, while p_i lies between p_i and $p_i + dp_i$, or we say that the system is found in the infinitesimal *phase-space* volume

$$dq dp \equiv dq_1 dq_2 \dots dq_s dp_1 dp_2 \dots dp_s. \quad (1.6)$$

The system, in the course of time, follows a trajectory through the $2s$ -dimensional volume of phase space. Given initial conditions, $(q_i)_0, (p_i)_0$, Hamilton's equations uniquely determine the trajectory. See Fig. 1.1.

As s gets large, following the trajectory becomes increasingly difficult, even if we could somehow determine the initial conditions. What we will do instead is discuss the "average" behavior of the system in phase space. Boltzmann discussed the most probable distribution of the system in time. The *ergodic hypothesis* suggested that one trajectory could fill all of phase space (consistent with a given total energy). This is false; what is more plausible is Ehrenfest's *quasi-ergodicity*: a single trajectory is dense on the energy surface—that is, it passes arbitrarily close to any given point. But this is also not true in general.

Instead of time averages, Gibbs tells us to think of *ensemble* averages. That is, we consider N identical replicas of the same system, each of which is *weakly interacting* with the others, in a heat bath of the others. As $N \rightarrow \infty$, we can ask what is the probability dw of finding one of the systems in the phase-space volume $dq dp$,

$$dw = \rho(q_1, \dots, q_s; p_1, \dots, p_s) dq_1 \dots dq_s dp_1 \dots dp_s \equiv \rho(q_i, p_i) dq dp. \quad (1.7)$$

The quantity ρ is called the density of the probability distribution. Because dw is simply the fraction of systems in this state, it must be true that integrating

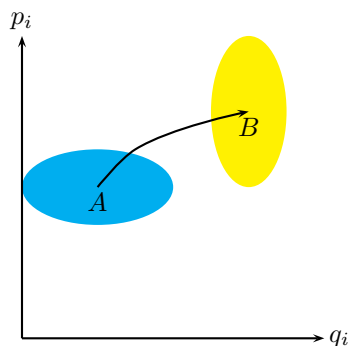


Figure 1.2: Evolving in time according to Hamilton's equations, a phase-space volume A at time t_0 evolves into another phase-space volume B at time t_1 . The volumes of these two regions in phase space are the same.

over all states gives

$$\int dw = \int \rho(q_i, p_i) dq dp = 1, \quad (1.8)$$

where the integrals extend over all of phase space. Given ρ we can calculate the mean value of some physical quantity $f(q_i, p_i)$:

$$\langle f \rangle = \int f(q_i, p_i) \rho(q_i, p_i) dq dp. \quad (1.9)$$

Boltzmann would say this mean value could be obtained as a time average,

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt f(t), \quad (1.10)$$

but this time average may not exist. It would, if the ergodic hypothesis were true, but it is not. In fact, it is ensemble averages, not time averages, that are measured—one repeats the experiment over and over again, and quotes an average value; one does not watch one system forever.

1.1 Liouville's theorem

Liouville's theorem is a statement of the conservation of phase-space volume, as illustrated in Fig. 1.2.

To prove this theorem, consider a small interval in one of the momenta, say,

$$dp_1 = p_1'' - p_1', \quad (1.11)$$

and consider a system entering this interval from the left, that is, with a smaller value of p_1 during a very small time interval dt . To accomplish this, the value of p_1 must be between p_1' and $p_1' - \dot{p}_1' dt$. Likewise any systems in that dp_1 interval lying between p_1'' and $p_1'' - \dot{p}_1'' dt$ will leave the element. (This assumes that both

\dot{p}'_1 and \dot{p}''_1 are positive; if not, change the word enter to leave and vice versa.) We assume that dt is so small that

$$\dot{p}'_1 dt, \dot{p}''_1 dt \ll dp_1. \quad (1.12)$$

Therefore the fraction of systems entering dp_1 is

$$\rho(p'_1, \dots) \dot{p}'_1 dt dp_2 \dots dq_1 \dots, \quad (1.13)$$

while the fraction of systems leaving dp_1 is

$$\rho(p''_1, \dots) \dot{p}''_1 dt dp_2 \dots dq_1 \dots \quad (1.14)$$

Consequently, the decrease in the fraction of systems between p'_1 and $p''_1 = p'_1 + dp_1$ is

$$\frac{d(\rho \dot{p}_1)}{dp_1} dp_1 dp_2 \dots dq_1 dq_2 \dots dt. \quad (1.15)$$

We argue in just the same way for all the other coordinates and momenta. Therefore the decrease in the fraction of systems in the phase-space element $dq dp$ is

$$\sum_{k=1}^s \left(\frac{\partial(\rho \dot{p}_i)}{\partial p_i} + \frac{\partial(\rho \dot{q}_i)}{\partial q_i} \right) dq dp dt. \quad (1.16)$$

Now, from Hamilton's equations (1.3)

$$\frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial}{\partial p_i} \left(-\frac{\partial H}{\partial q_i} \right) = -\frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} = -\frac{\partial \dot{q}_i}{\partial q_i}, \quad (1.17)$$

so the decrease in the fraction of systems in the $dq dp$ phase-space element is

$$-\frac{\partial \rho}{\partial t} dt dq dp = \sum_{i=1}^s \left(\frac{\partial \rho}{\partial p_i} \dot{p}_i + \frac{\partial \rho}{\partial q_i} \dot{q}_i \right) dq dp dt, \quad (1.18)$$

where $\partial \rho / \partial t$ means that p_i and q_i are held fixed. Therefore, we conclude that the total time derivative of the probability density vanishes:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^s \left(\frac{\partial \rho}{\partial p_i} \dot{p}_i + \frac{\partial \rho}{\partial q_i} \dot{q}_i \right) = 0. \quad (1.19)$$

1.1.1 Statistical Equilibrium

If the probability density does not explicitly depend on time,

$$\frac{\partial \rho}{\partial t} = 0, \quad (1.20)$$

that is, ρ is only a function of q_i, p_i , we say the system is in statistical equilibrium. Then

$$\frac{d\rho}{dt} = \sum_{i=1}^s \left(\frac{\partial \rho}{\partial p_i} \dot{p}_i + \frac{\partial \rho}{\partial q_i} \dot{q}_i \right) = 0. \quad (1.21)$$

The statement of Liouville's theorem given in Fig. 1.2 then follows from the equilibrium statement that the fraction of systems in a phase-space element, $\rho dq dp$ is constant, because systems are neither created nor destroyed, and since $d\rho/dt = 0$, we must have

$$\frac{d}{dt}(dq dp) = 0. \quad (1.22)$$

This also can be seen explicitly from the Jacobian of the transformation from the coordinates and momenta at one time, q_i, p_i , and the coordinates and momenta at an infinitesimally displaced time,

$$q_i \rightarrow \bar{q}_i = q_i + \dot{q}_i dt, \quad p_i \rightarrow \bar{p}_i = p_i + \dot{p}_i dt, \quad (1.23)$$

so

$$d\bar{q} d\bar{p} = \frac{\partial(\bar{q}, \bar{p})}{\partial(q, p)} dq dp, \quad (1.24)$$

where

$$\begin{aligned} \frac{\partial(\bar{q}, \bar{p})}{\partial(q, p)} &= \begin{vmatrix} 1 + \frac{\partial \dot{q}_1}{\partial q_1} dt & \frac{\partial \dot{p}_1}{\partial q_1} dt & \dots \\ \frac{\partial \dot{q}_1}{\partial p_1} dt & 1 + \frac{\partial \dot{p}_1}{\partial p_1} dt & \dots \\ \dots & \dots & \dots \end{vmatrix} \\ &= 1 + \sum_{i=1}^s \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) dt + O(dt^2) \\ &= 1 + O(dt^2), \end{aligned} \quad (1.25)$$

where the last step follows from the identity (1.17).

Here are some alternate ways of writing the Liouville equation. Recall from mechanics that the Poisson bracket is defined by

$$\{A, B\} = \sum_{i=1}^s \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right). \quad (1.26)$$

Therefore Eq. (1.19) reads

$$\frac{\partial \rho}{\partial t} = -\{\rho, H\}. \quad (1.27)$$

Or, if we define $2s$ -dimensional vectors by

$$\mathbf{v} = (\dot{q}_i, \dot{p}_i), \quad \nabla = \left(\frac{\partial}{\partial q_i}, \frac{\partial}{\partial p_i} \right), \quad (1.28)$$

that equation can be written as

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho = 0, \quad (1.29)$$

from which we see that d/dt has the interpretation as a substantial or co-moving derivative. Because $\nabla \cdot \mathbf{v} = 0$, we see that the Liouville equation is simply the equation of continuity,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (1.30)$$

The question now arises, what is ρ ? It is not measurable experimentally, and must be guessed. In fact, physical consequences are insensitive to the form of ρ . In equilibrium, $\partial\rho/\partial t = 0$, $\rho = \rho(q_i, p_i)$ must be a function of constants of the motion. The most obvious possibility is the Hamiltonian H , which we also assume to be independent of time,

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = 0. \quad (1.31)$$

But ρ could also be a function of the linear momentum \mathbf{P} or the angular momentum \mathbf{J} . But usually we think of systems in “boxes,” so space-translation invariance (which implies $d\mathbf{P}/dt = 0$) and rotational invariance (which implies $d\mathbf{J}/dt = 0$) are not present. But in any case, there are always $2s - 1$ constants of the motion. Why single out the energy? The other constants of the motion are generally discontinuous along the path and are multiple-valued. Only surfaces of constant energy can be used to describe motion globally.

The simplest possibility would seem to be to confine the system to a surface of definite energy E :

$$\rho_{MC}(q_i, p_i) = \frac{1}{\Omega(E)} \delta(E - H). \quad (1.32)$$

This is called the *microcanonical* distribution. Here $\Omega(E)$ is a normalization constant, determined by Eq. (1.8):

$$\Omega(E) = \int \delta(E - H) dq dp. \quad (1.33)$$

Another important distribution is the *canonical* one:

$$\rho_C(q_i, p_i) = \frac{1}{Z} e^{-\beta H}, \quad Z = \int e^{-\beta H} dq dp. \quad (1.34)$$

(It will turn out that $\beta = 1/kT$, where k is Boltzmann’s constant, and T is the absolute temperature. Z is called the partition function.) The canonical distribution has the important property that if $H = H_1 + H_2$,

$$\rho = \frac{1}{Z} e^{-\beta H} = \frac{1}{Z_1} e^{-\beta H_1} \frac{1}{Z_2} e^{-\beta H_2}; \quad (1.35)$$

the probability distributions multiply, as we might expect for two independent systems.