Cambridge University Press
978-0-521-82575-7 - Statistical Mechanics: From First Principles to Macroscopic Phenomena
J. Woods Halley
Excerpt
More information

Introduction

The problems of statistical mechanics are those which involve systems with a larger number of degrees of freedom than we can conveniently follow explicitly in experiment, theory or simulation. The number of degrees of freedom which can be followed explicitly in simulations has been changing very rapidly as computers and algorithms improve. However, it is important to note that, even if computers continue to improve at their present rate, characterized by Moore's "law," scientists will not be able to use them for a very long time to predict many properties of nature by direct simulation of the fundamental microscopic laws of physics. This point is important enough to emphasize.

Suppose that, T years from the present, a calculation requiring computation time t_0 at present will require computation time $t(T) = t_0 2^{-T/2}$ (Moore's "law,"[1] see Figure 1). Currently, state of the art numerical solutions of the Schrödinger equation for a few hundred atoms can be carried out fast enough so that the motion of these atoms can be followed long enough to obtain thermodynamic properties. This is adequate if one wishes to predict properties of simple homogeneous gases, liquids or solids from first principles (as we will be discussing later). However, for many problems of current interest, one is interested in entities in which many more atoms need to be studied in order to obtain predictions of properties at the macroscopic level of a centimeter or more. These include polymers, biomolecules and nanocrystalline materials for example. In such problems, one easily finds situations in which a first principles prediction requires following 10^6 atoms dynamically. The first principles methods for calculating the properties increase in computational cost as the number of atoms to a power between 2 and 3. Suppose they scale as the second power so the computational time must be reduced by a factor 10^8 in order to handle 10^6 atoms. Using Moore's law we then predict that the calculation will be possible T years from the present where $T = 16/\log_{10}2 = 53$ years. In fact, this may be optimistic because Moore's "law" may not continue to be valid for that long and also because 10^6 atoms will not be enough in many cases. What this means is that,

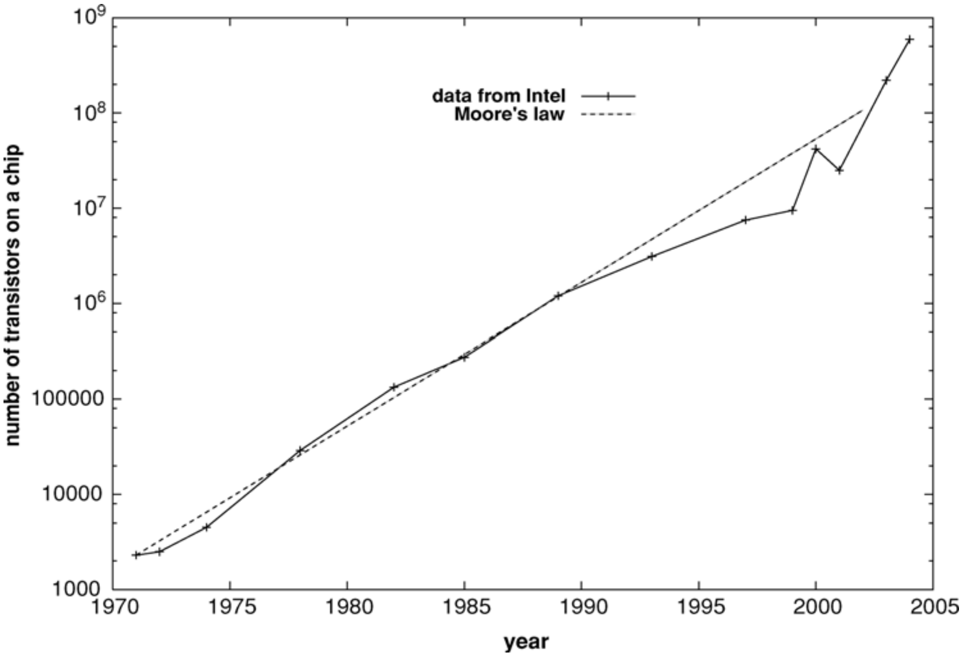

Figure 1 One version of Moore's "law."

for a long time, we will need means beyond brute force computation for relating the properties of macroscopic matter to the fundamental microscopic laws of physics.

Statistical mechanics provides the essential organizing principles needed for connecting the description of matter at large scales to the fundamental underlying physical laws (Figure 2). Whether we are dealing with an experimental system with intractably huge numbers of degrees of freedom or with a mass of data from a simulation, the essential goal is to describe the behavior of the many degrees of freedom in terms of a few "macroscopic" degrees of freedom. This turns out to be possible in a number of cases, though not always. Here, we will first describe how this connection is made in the case of equilibrium systems, whose average properties do not change in time. Having established (Part I) some principles of equilibrium statistical mechanics, we then provide (Part II) a discussion of how they are applied in the three most common phases of matter (gases, liquids and solids) and the treatment of phase transitions. Part III concerns dynamical and nonequilibrium methods.

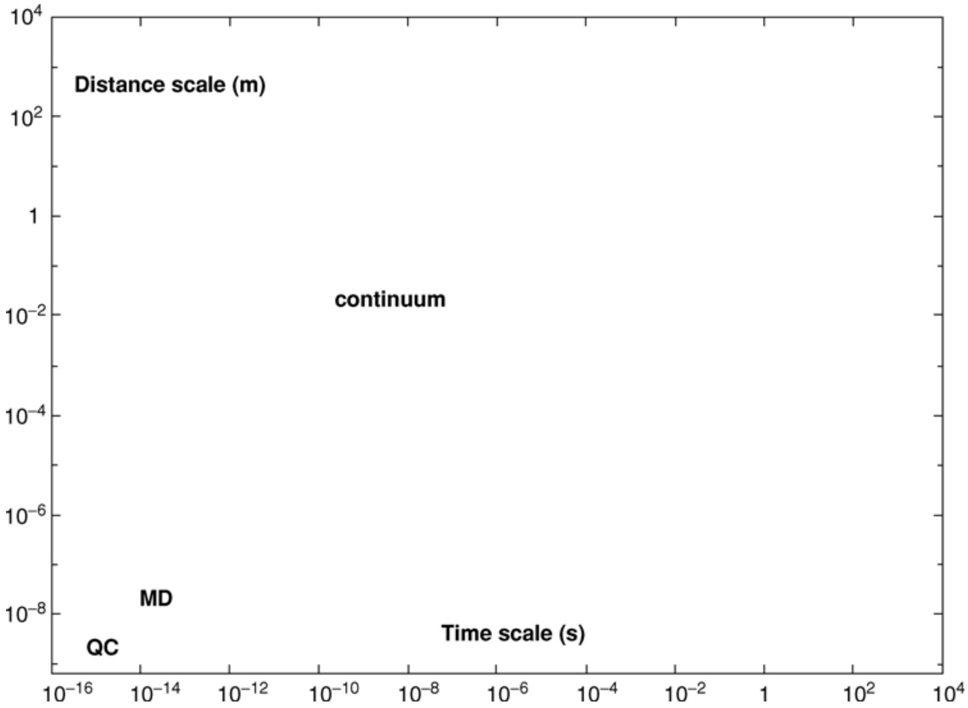

Figure 2 Computational length and time scales. QC stands for quantum chemistry methods in which the Schrödinger equation is solved. MD stands for molecular dynamics in which classical equations of motion for atomic motion are solved. Continuum includes thermodynamics, hydrodynamics, continuum mechanics, micromagnetism in which macroscopic variables describe the system. Statistical mechanics supplies the principles by which computations at these different scales are connected.

Reference

1. C. E. Moore, *Electronics*, April 19 (1965).

Cambridge University Press
978-0-521-82575-7 - Statistical Mechanics: From First Principles to Macroscopic Phenomena
J. Woods Halley
Excerpt
More information

Part I

Foundations of equilibrium statistical mechanics

1

The classical distribution function

Historically, the first and most successful case in which statistical mechanics has made the connection between microscopic and macroscopic description is that in which the system can be said to be in equilibrium. We define this carefully later but, to proceed, may think of the equilibrium state as the one in which the values of the macroscopic variables do not drift in time. The macroscopic variables may have an obvious relation to the underlying microscopic description (as for example in the case of the volume of the system) or a more subtle relationship (as for temperature and entropy). The macroscopic variables of a system in equilibrium are found experimentally (and in simulations) to obey historically empirical laws of thermodynamics and equations of state which relate them to one another. For systems at or near equilibrium, statistical mechanics provides the means of relating these relationships to the underlying microscopic physical description.

We begin by discussing the details of this relation between the microscopic and macroscopic physical description in the case in which the system may be described classically. Later we run over the same ground in the quantum mechanical case. Finally we discuss how thermodynamics emerges from the description and how the classical description emerges from the quantum mechanical one in the appropriate limit.

Foundations of equilibrium statistical mechanics

Here we will suppose that the systems with which we deal are nonrelativistic and can be described fundamentally by $3N$ time dependent coordinates labelled $q_i(t)$ and their time derivatives $\dot{q}_i(t)$ $(i = 1, \ldots, 3N)$. A model for the dynamics of the system is specified through a Lagrangian $L(\{q_i\}, \{\dot{q}_i\})$ (not explicitly time dependent) from which the dynamical behavior of the system is given by the principle of least action

$$\delta \int L \, \mathrm{d}t = 0 \tag{1.1}$$

or equivalently by the Lagrangian equations of motion

$$\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \tag{1.2}$$

Alternatively one may define momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{1.3}$$

and a Hamiltonian

$$H = \sum_{i=1}^{N} p_i \dot{q}_i - L \tag{1.4}$$

Expressing H as a function of the momenta p_i and the coordinates q_i one then has the equations of motion in the form

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \tag{1.5}$$

$$-\frac{\partial H}{\partial q_i} = \dot{p}_i \tag{1.6}$$

In examples, we will often be concerned with a system of identical particles with conservative pair interactions. Then it is convenient to use the various components of the positions of the particles $\vec{r}_1, \vec{r}_2, \ldots$ as the quantities q_i, and the Hamiltonian takes the form

$$H = \sum_k \vec{p}_k^{\,2}/2m + (1/2) \sum_{k \neq l} V(\vec{r}_k, \vec{r}_l) \tag{1.7}$$

where the sums run over particle labels and $\vec{p}_k = \nabla_{\dot{\vec{r}}_k} H$. Then the Hamiltonian equations reduce to simple forms of Newton's equation of motion. It turns out, however, that the more general formulation is quite useful at the fundamental level, particularly in understanding Liouville's theorem, which we will discuss later.

In keeping with the discussion in the Introduction, we wish to relate this microscopic description to quantities which are measured in experiment or which are conveniently used to analyze the results of simulations in a very similar way. Generically we denote these observable quantities as $\phi(q_i(t), p_i(t))$. It is also possible to consider properties which depend on the microscopic coordinates at more than one time. We will defer discussion of these until Part III. Generally, these quantities, for example the pressure on the wall of a vessel containing the system, are not constant

Cambridge University Press
978-0-521-82575-7 - Statistical Mechanics: From First Principles to Macroscopic Phenomena
J. Woods Halley
Excerpt
More information

in time and what is measured is a time average:

$$\bar{\phi}_t = \frac{1}{\tau}\int_{t-\tau/2}^{t+\tau/2} \phi(q_i(t'), p_i(t'))\,\mathrm{d}t' \tag{1.8}$$

τ is an averaging time determined by the apparatus and the measurement made (or chosen for analysis by the simulator). Experience has shown that for many systems, an experimental situation can be achieved in which measurements of $\bar{\phi}_t$ are independent of τ for all $\tau > \tau_0$ for some finite τ_0. It is easy to show that, in such a case, $\bar{\phi}_t$ is also independent of t. If this is observed to be the case for the macroscopic observables of interest, then the system is said to be in equilibrium. A similar operational definition of equilibrium is applied to simulations. In practice it is never possible to test this equilibrium condition for arbitrarily long times, in either experiment or simulation. Thus except in the rare cases in which mathematical proofs exist for relatively simple models, the existence and nature of equilibrium states are hypothesized on the basis of partial empirical evidence. Furthermore, in experimental situations, we do not expect any system to satisfy the equilibrium condition for arbitrarily long times, because interactions with the surroundings will inevitably change the values of macroscopic variables eventually. Making the system considered ever larger and the time scales longer and longer does not help here, because there is no empirical evidence that the universe itself is in equilibrium in this sense. Nevertheless, the concept of equilibrium turns out to be an extremely useful idealization because of the strong evidence that many systems do satisfy the relevant conditions over a very wide range of averaging times τ and that, under sufficiently isolated conditions, many systems spontaneously evolve rapidly toward an approximately equilibrium state whose characteristics are not sensitive to the details of the initial microscopic conditions. These empirical statements lack mathematical proofs for most systems of experimental or engineering interest, though mathematicians have made progress in proving them for simple models.

For systems in equilibrium defined in this way we are concerned with the calculation of averages of the type

$$\bar{\phi}_t = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^{\tau} \phi(\{q_i(t')\}, \{p_i(t')\})\,\mathrm{d}t' \tag{1.9}$$

We will show that it is always possible in principle to write this average in the form

$$\bar{\phi}_t = \int \rho(\{q_i\}, \{p_i\})\phi(\{q_i\}, \{p_i\})\,\mathrm{d}^{3N}q\,\mathrm{d}^{3N}p \tag{1.10}$$

in which $\rho(\{q_i\}, \{p_i\})$ is called the classical distribution function. The demonstration provides useful insight into the meaning of $\rho(\{q_i\}, \{p_i\})$. We consider the $6N$ dimensional space of the variables $\{q_i\}, \{p_i\}$, called phase space. In this space the time

Cambridge University Press
978-0-521-82575-7 - Statistical Mechanics: From First Principles to Macroscopic Phenomena
J. Woods Halley
Excerpt
More information

evolution of the system is described by the motion of a point. Take a small region of this space whose volume is denoted $\Delta^{3N}p\Delta^{3N}q$ centered at the point (p, q). (Henceforth we denote $(p, q) \equiv (\{q_i\}, \{p_i\})$ and similarly $(\Delta p, \Delta q) \equiv (\{\Delta q_i\}, \{\Delta p_i\})$.) Consider the interval of time Δt defined as

$$\Delta t(q_0, p_0, t_0; q, p, t; \Delta p, \Delta q) \tag{1.11}$$

equivalent to the time which the point describing the system spends in the region $\Delta^{3N}p\Delta^{3N}q$ around (q, p) between t_0 and t if it started at the point (q_0, p_0) at time t_0.

Now consider the fraction of time that the system point spends in $\Delta^{3N}p\Delta^{3N}q$, denoted Δw:

$$\Delta w(q_0, p_0; q, p; \Delta p, \Delta q) = \lim_{t\to\infty}\left(\frac{\Delta t}{t - t_0}\right) \tag{1.12}$$

which is the fraction of the total time between t_0 and $t \to \infty$ which the system spends in the region $\Delta^{3N}p\Delta^{3N}q$ around (q, p).

Now we express the time average $\bar{\phi}_t$ of equation (1.9) in terms of Δw by dividing the entire phase space into small regions labelled by an index k and each of volume $\Delta^{3N}p\Delta^{3N}q$:

$$\bar{\phi}_t = \sum_k \phi(q_0, p_0; q_k, p_k)\Delta w(q_0, p_0; q_k, p_k; \Delta p, \Delta q) \tag{1.13}$$

We then suppose that $\Delta w(q_0, p_0; q, p; \Delta p, \Delta q)$ is a well behaved function of the arguments $(\Delta p, \Delta q)$ and write

$$\Delta w = \left[\frac{\partial^{6N}\Delta w}{\partial^{3N}\Delta q\,\partial^{3N}\Delta p}\right]_{\Delta p=\Delta q=0} \Delta^{3N}q\,\Delta^{3N}p + \cdots \tag{1.14}$$

Defining

$$\rho(q_0, p_0; q, p) = \left[\frac{\partial^{6N}\Delta w}{\partial^{3N}\Delta q\,\partial^{3N}\Delta p}\right]_{\Delta p=\Delta q=0} \tag{1.15}$$

we then have in the limit $\Delta p\Delta q \to 0$ that

$$\bar{\phi}_t = \int \rho(q_0, p_0; q, p)\phi(q, p)\,\mathrm{d}^{3N}q\,\mathrm{d}^{3N}p \tag{1.16}$$

which is of the form (1.10). Several of the smoothness assumptions made in this discussion are open to question as we will discuss in more detail later.

Equation (1.16) is most useful if $\bar{\phi}_t$ depends only on a few of the $6N$ initial conditions q_0, p_0. Experimentally (and in simulations) it is found that the time averages of many macroscopic quantities measured in equilibrium systems are very insensitive to the way the system is prepared. We will demonstrate that under certain

Cambridge University Press
978-0-521-82575-7 - Statistical Mechanics: From First Principles to Macroscopic Phenomena
J. Woods Halley
Excerpt
More information

conditions, the only way in which these averages can depend on the initial conditions is through the values of the energy, linear momentum and angular momentum of the entire system. The general study of the dependence of averages of the form (1.16) on the initial conditions is part of ergodic theory. An ergodic system is (loosely from a mathematical point of view) defined as an energetically isolated system for which the phase point eventually passes through every point on the surface in phase space consistent with its energy. It is not hard to prove that the averages $\bar{\phi}_t$ in such an ergodic system depend only on the energy of the system. It is worth pointing out that the existence of ergodic systems in phase space of more than two dimensions is quite surprising. The trajectory of the system in phase space is a topologically one dimensional object (a path, parametrized by one variable, the time) yet we want this trajectory to *fill* the $6N - 1$ dimensional surface defined by the energy. The possibility of space filling curves is known mathematically (for a semipopular account see reference 1). However, for a large system, the requirement is extreme: the trajectory must fill an enormously open space of the order of 10^{23} dimensions! By contrast the path of a random walk has dimension 2 (in any embedding dimension)! (Very briefly, the (fractal or Hausdorff–Besicovitch) dimension of a random walk can be understood to be 2 as follows. The dimension of an object in this sense is determined as D_{H} defined so that when one covers the object in question with spheres of radius η a minimum of $N(\eta)$ spheres is required and

$$L_{\mathrm{H}} = \lim_{\eta \to 0} N(\eta)\eta^{D_{\mathrm{H}}}$$

is finite and nonzero. For a random walk of mean square radius $\langle R^2 \rangle$, $N(\eta) = \langle R^2 \rangle / \eta^2$ and $D_{\mathrm{H}} = 2$. See reference 1 for details.) Nevertheless something like ergodicity is required for statistical mechanics to work, and so the paths in phase space of large systems must in fact achieve this enormous convolution in order to account for the known facts from experiment and simulation. It is not true that every system consisting of small numbers of particles is ergodic. Some of the problems at the end of this section illustrate this point. For example, a one dimensional harmonic oscillator is ergodic, but a billiard ball on a two dimensional table is not (Figure 1.1). On the other hand, in the latter case, the set of initial conditions for which it is not ergodic is in some sense “small.” Another instructive example is a two dimensional harmonic oscillator (Problem 1.1).

There are several rationally equivalent ways of talking about equation (1.10). These occur in textbooks and other discussions and reflect the history of the subject as well as useful approaches to its extension to nonequilibrium systems. What we have discussed so far may be termed the Boltzmann interpretation of ρ (in which ρ is related to the time which the system phase point spends in each region of phase space). This is closely related to the *probability* interpretation of ρ because the

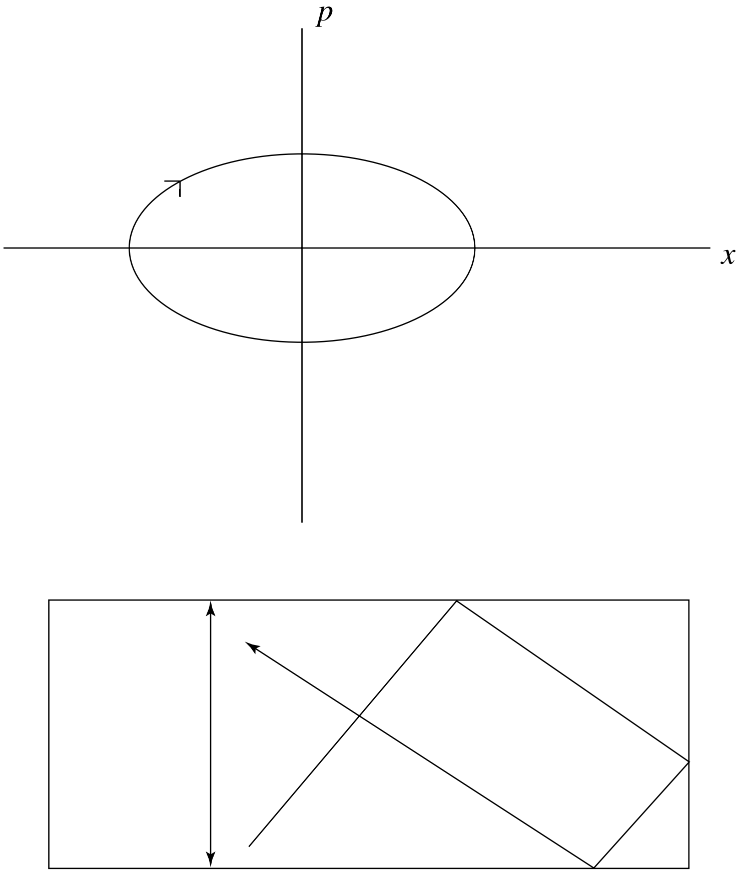

Figure 1.1 Phase space trajectory of a one dimensional oscillator fills the energy surface. For some initial conditions, a ball on a billiard table with elastic specularly reflecting walls is not ergodic.

probability that the system is found in $d^{3N}q\,d^{3N}p$ is just $\rho d^{3N}q\,d^{3N}p$ according to the standard observation frequency definition of probability. In such an interpretation, one takes no interest in the question of how the system got into each phase space region and could as well imagine that it hopped discontinuously from one to another for some purposes. Indeed such discontinuous hops (which we do not believe occur in real experimental systems obeying classical mechanics to a good approximation) do occur in certain numerical methods of computing the integrals (1.10) once the form of ρ is known. Regarding $\rho d^{3N}q\,d^{3N}p$ as a probability opens the way to the use of information theoretic methods for approximating its form under all sorts of conditions in which various constraints are applied. For mechanical systems in equilibrium this approach leads to the same forms which we will obtain and use here. The reader is referred to the book by Katz[2] and to many papers by Jaynes