

Models and Methods for Statistical Mechanics

from the lectures of Prof. Lorenzani

Andrea Bonifacio

October 16, 2023

1 Introduction

Statistical mechanics may be naturally divided into two branches:

- equilibrium systems: independent of time,
- non-equilibrium systems: evolve in time.

Non-equilibrium phenomena are less understood at the moment. A notable exception is offered by the case of diluted gases.

A basic equation was established by Ludwig Boltzmann in 1872, which appears as a prototype of a reduced description taking into account only partial information about the underlying microscopic state (fully described by the coordinate and momenta of all the molecules), but nevertheless undergoing an autonomous time evolution. Thus, the problem of the rigorous derivation of the Boltzmann equation from the microscopic description has attracted a certain amount of interest among physicists and mathematicians. This, in turn, has revived the interest in the existence and uniqueness of the solution of the Boltzmann equation, since this problem has proved to be intimately tied with the previous one. According to the molecular theory of matter, a macroscopic volume of gas (e.g. 1 cm^3) is a system if a very large number of molecules ($N \approx 10^{20}$) moving in a rather irregular way.

In principle, we may assume that the laws of interaction between the molecules are perfectly known so that, idealistically, the evolution of the system is computable provided suitable initial data are given. If the molecules are, for example, mass points, the equation of motion are

$$\begin{cases} \dot{\vec{\xi}}_i = \vec{X}_i, \\ \dot{\vec{x}}_i = \vec{\xi}_i, \end{cases} \quad (1.1a)$$

or:

$$\ddot{\vec{x}}_i = \vec{X}_i, \quad (1.1b)$$

where \vec{x}_i is the position vector of the i^{th} particle of a system of $i = 1, \dots, N$ particles and $\vec{\xi}_i$ is the velocity vector. The term \vec{X}_i is the force acting upon the particle divided by its mass. Such a force will be usually the sum of the external forces (e.g. gravity) and the forces describing the action of the other particles on the one under observation. The expression of such forces must be given as a part of the description of the mechanical system.

In order to compute the time evolution of the system, one would need to solve the $6N$ first order differential equations (1.1a) in the $6N$ unknowns constituting the $2N$ vector $(\vec{x}_i, \vec{\xi}_i)$. A prerequisite for this is the knowledge of the $6N$ initial conditions

$$\begin{aligned} \vec{x}_i(0) &= \vec{x}_i^0, \\ \dot{\vec{x}}_i(0) &= \vec{\xi}_i(0) = \vec{\xi}_i^0, \end{aligned} \quad (1.2)$$

where the component of \vec{x}_i^0 and $\vec{\xi}_i^0$ are $6N$ given constants which describe the initial state of the system. However, solving the above initial value problem is an impossible and useless task for a realistic number of molecules. As a consequence, one should consider the evolution, not of a system, but of an ensemble of identical systems whose initial data differ from each other by quantities of the order of the accepted ones.

The conclusion is that the only significant and useful results are those about the behavior of many systems in the form of statistics, that is information about probable distributions. As a result, only overlaps can be computed, provided they are related to macroscopic quantities as pressure, temperature, stresses, heat flow, etc. This is the basic idea of statistical mechanics. When we deal with statistical mechanics, therefore, we talk about probabilities instead of certainties: that is, in our description, a given particle will not have a definite position and velocity, but only different probabilities and velocities. It is to be stressed that the particles which appear in statistical mechanics range through a continuous set of values.

Thus, in the case of n continuous variables $\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n$, that is a vector variable $\vec{Z} = (\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n)$, we have to introduce a probability density $P(\vec{z})$ such that $P(\vec{z})d^n\vec{z}$ is the probability that \vec{z} lies between

\vec{z} and $\vec{z} + d\vec{z}$ with $d^n \vec{z}$ denoting the volume of an infinitesimal cell, also denoted as $d\vec{z}_1, d\vec{z}_2, \dots, d\vec{z}_n$. In this case the property of the “sum” is one becomes

$$\int_Z P(\vec{z}) d^n \vec{z} = 1 \quad (1.3)$$

where Z is the region of the n -dimensional space in which \vec{z} varies. What is the use of a probability density? A probability density is needed to compute averages: if we know the probability density $P(\vec{z})$ we can compute the average value of any given functions. As a matter of fact, we can define averages as follows

$$\langle q(\vec{z}) \rangle = \overline{q(\vec{z})} = \int_Z P(\vec{z}) q(\vec{z}) d\vec{z}. \quad (1.4)$$

In other words, in order to compute the average value of a function $\varphi(\vec{z})$ we integrate it over all values of \vec{z} weighting each $d\vec{z}$ with the probability density for a value in $d\vec{z}$ to be realized.

If $P(\vec{z}) = \delta(\vec{z} - \vec{z}_0)$, then $\overline{q(\vec{z})} = q(\vec{z}_0)$, that is, in the case of a probability density equal to the delta function, we have a probability density meant to represent certainty.

2 Phase space and Liouville's theorem

In order to discuss the behavior of a system of N mass points satisfying (1.1a), it is very convenient to introduce the so-called **phase space**; that is, a $6N$ -dimensional space where the Cartesian coordinates are the $3N$ components of the N position vectors \vec{x}_i and the $3N$ components of the N position $\vec{\xi}_i$. Let us introduce the $6N$ -dimensional vector \vec{z} which gives the position of the representative point in phase space

$$\vec{z} = (\vec{x}, \vec{\xi}).$$

The evolution equation for \vec{z} is from (1.1a)

$$\dot{\vec{z}} = \frac{d\vec{z}}{dt} = \vec{Z}, \quad (2.1)$$

where \vec{Z} is a $6N$ -dimensional vector, whose components are respectively given by the $3N$ components of the N three-dimensional vectors $\vec{\xi}_i$ and the $3N$ components of the N three-dimensional vectors \vec{X}_i . Given the initial state, that is a point \vec{z}_0 in the phase space, (2.1) determines \vec{z} at subsequent times. If the initial data are not known with absolute accuracy, we must introduce a probability density $P_0(\vec{z})$ which gives us the distribution of probability for the initial data, and we can try to set up the problem of computing the probability density at subsequent times $P(\vec{z}, t)$. This can be done provided the forces are known; that is if the only uncertainty is in the initial data.

The equation satisfied by $P(\vec{z}, t)$ can be defined in the following way: the evolution equation (2.1) defines, at each instant, a mapping of the phase space into itself; in this mapping, to each point \vec{z}_0 there corresponds the point $\vec{z} = \vec{z}(\vec{z}_0, t)$ reached at time t by a point which was at \vec{z}_0 at $t = 0$. The mapping is 1 – 1 if the equations have, as we shall assume, a unique solution to given initial data, for both $t > 0$ and $t < 0$. The probability for the point representing the system to be found in a region R of the phase space at time t is

$$Prob(\vec{z} \in R) = \int_R P(\vec{z}, t) d\vec{z} \quad (2.2)$$

The above-mentioned probability will be equal to the probability that the representative point was, at $t = 0$, in the region R_0 consisting of the points \vec{z}_0 which are the inverse images of the points $\vec{z} \in R$ in the mapping considered above. In fact a point can be in R at time t if and only if it was in R_0 at $t = 0$. Hence:

$$\int_R P(\vec{z}, t) d\vec{z} = \int_{R_0} P_0(\vec{z}_0) d\vec{z}_0, \quad (2.3)$$

where the set of values of $\vec{z} \in R$ coincides with the set of points $\vec{z} = \vec{z}(\vec{z}_0, t)$ with $\vec{z}_0 \in R_0$. We can exploit this fact to change the integration variables from the components of \vec{z} to those of \vec{z}_0 in the first integral obtaining:

$$\int_R P(\vec{z}, t) d\vec{z} = \int_{R_0} P_0(\vec{z}(\vec{z}_0, t), t) J(\vec{z}/\vec{z}_0) d\vec{z}_0, \quad (2.4)$$

where $J(\vec{z}/\vec{z}_0)$ is the Jacobian determinant of the old variables with respect to the new ones. Comparison of (2.4) and (2.3) gives, due to the arbitrariness of R_0 ,

$$P(\vec{z}(\vec{z}_0, t), t) J(\vec{z}/\vec{z}_0) = P_0(\vec{z}_0). \quad (2.5)$$

Since the right-hand side is time independent, the total derivative of the left-hand side with respect to time must vanish identically

$$J \left(\frac{\partial P}{\partial t} + \frac{\partial \vec{z}}{\partial t} \cdot \frac{\partial P}{\partial \vec{z}} \right) + P \cdot \frac{\partial J}{\partial t} = 0, \quad (2.6)$$

where the argument of P are \vec{z}, t and those of J are \vec{z}_0, t . Since $J \neq 0$, we can use (2.1) to obtain

$$\frac{\partial P}{\partial t} + \vec{Z} \cdot \frac{\partial P}{\partial \vec{z}} + \frac{P}{J} \frac{\partial J}{\partial t} = 0. \quad (2.7)$$

We compute now $\frac{\partial J}{\partial t}$. Let J_{rs} be the cofactor of $\frac{\partial \vec{z}_r}{\partial \vec{z}_s^o}$ in the Jacobian determinant, then try the rule for differentiating determinants, keeping in mind that

$$J_{rs} = J \frac{\partial \vec{z}_r}{\partial \vec{z}_s^o}, \quad \text{Laplace's theorem}$$

we have

$$\begin{aligned} \frac{\partial J}{\partial t} &= \sum_{r,s=1}^{6N} \frac{\partial}{\partial t} \left(\frac{\partial \vec{z}_r}{\partial \vec{z}_s^o} \right) J_{rs} = \sum_{r,s=1}^{6N} \frac{\partial}{\partial \vec{z}_s^o} \left(\frac{\partial \vec{z}_r}{\partial t} \right) J_{rs} = \sum_{r,s=1}^{6N} \frac{\partial \vec{z}_r}{\partial \vec{z}_s^o} J_{rs} = \\ &= J \sum_{r,s=1}^{6N} \frac{\partial \vec{Z}_r}{\partial \vec{z}_s^o} \frac{\partial \vec{z}_s^o}{\partial \vec{z}_r} = J \sum_{r=1}^{6N} \frac{\partial \vec{Z}_r}{\partial \vec{z}_r} = J \operatorname{div} \vec{Z}. \end{aligned} \quad (2.8)$$

Inserting (2.8) in (2.7) gives

$$\frac{\partial P}{\partial t} + \vec{Z} \cdot \frac{\partial P}{\partial \vec{z}} + \frac{P}{J} J \operatorname{div} \vec{Z} = 0,$$

which is written as

$$\frac{\partial P}{\partial t} + \vec{Z} \cdot \nabla P + P \operatorname{div} \vec{Z} = 0. \quad \text{Liouville equation}$$

Normally, $\operatorname{div} \vec{Z} = 0$. In fact, since \vec{x}_i and $\vec{\xi}_i$ are independent variables.

$$\operatorname{div} \vec{Z} = \sum_{i=1}^N \left(\frac{\partial}{\partial \vec{x}_i} \cdot \vec{\xi}_i + \frac{\partial}{\partial \vec{\xi}_i} \cdot \vec{X}_i \right) = \sum_{i=1}^N \frac{\partial}{\partial \vec{\xi}_i} \cdot \vec{X}_i. \quad (2.9)$$

If the force per unit mass is velocity-independent, then also $\frac{\partial}{\partial \vec{\xi}_i} \cdot \vec{X}_i = 0$, and $\operatorname{div} \vec{Z} = 0$, as stated. We shall always consider forces such that $\operatorname{div} \vec{Z} = 0$.

Hence, we write the Liouville equation in the following form

$$\frac{\partial P}{\partial t} + \vec{Z} \cdot \frac{\partial P}{\partial \vec{z}} = 0. \quad (2.10)$$

(2.10) can be, of course, rewritten in terms of the variables $\vec{x}_i, \vec{\xi}_i$

$$\frac{\partial P}{\partial t} + \sum_{i=1}^N \vec{\xi}_i \cdot \frac{\partial P}{\partial \vec{\xi}_i} + \sum_{i=1}^N \vec{X}_i \cdot \frac{\partial P}{\partial \vec{\xi}_i} = 0. \quad (2.11)$$

We note, incidentally, that (2.8) shows that $J = \text{const.}$ for $\text{div } \vec{Z} = 0$ and, since $J = 1$ for $t = 0$, then $J = 1$ for any t . As a consequence, the volume of a region in phase space is invariant with respect to the mapping induced by the motion of the systems, provided $\text{div } \vec{Z} = 0$ (Liouville's theorem). In fact, we can write (2.4) for the case $P = 1$

$$\int_R d\vec{z} = \int_{R_0} J(\vec{z}/\vec{z}_0) d\vec{z}_0 = \int_{R_0} d\vec{z}_0. \quad (2.12)$$

3 Lecture 2

In the previous sections, we considered the case of mass points which continuously interact with each other according to the equations of motion (1.1a). It is frequently convenient to consider limiting cases in which the system has only discrete interactions with finite impulses (hard collisions); in such a case the forces are not describable by means of ordinary functions and the Liouville equation must be handled differently. The limiting case of hard collision is useful because it gives a more intuitive idea of the evolution of the system, and it is a good approximation to the strong repulsive forces which actual molecules mutually exert when they are close to each other.

These considerations lead to the concept of a gas of **hard spheres**, that is a system of many “billiard balls” which do not interact at a distance and collide according to the law of elastic impact. The diameter σ of the spheres is equivalent to the range of force through which actual molecules interact; as a matter of fact a gas of rigid spheres can be pictured as a system of mass points which do not interact when their mutual distance is larger than σ , but interact with a formally infinite repulsive central force when the distance is exactly σ .

In the case of a collision between two identical rigid spheres, the equations which relate the velocities after impact $\vec{\xi}_1'$ and $\vec{\xi}_2'$ to those before impact, $\vec{\xi}_1$ and $\vec{\xi}_2$ are:

$$\begin{cases} \vec{\xi}_1' = \vec{\xi}_1 - \hat{n} \left[\hat{n} \cdot (\vec{\xi}_1 - \vec{\xi}_2) \right], \\ \vec{\xi}_2' = \vec{\xi}_2 - \hat{n} \left[\hat{n} \cdot (\vec{\xi}_2 - \vec{\xi}_1) \right], \end{cases} \quad (3.1)$$

where \hat{n} is the unitary vector directed along the line joining the two centers of the spheres.

Next, we will show that the Liouville theorem (conservation of volume in phase space) remains valid for the instantaneous interactions considered here. Instead, in the following we will consider the heuristic arguments proposed by Boltzmann to write an evolution equation for the one-particle probability density $P^{(1)}(\vec{x}, \vec{\xi}, t)$.

Let us first consider the meaning of $P^{(1)}(\vec{x}, \vec{\xi}, t)$; it gives the probability density of finding one fixed particle (say the one labeled with (1)) at a certain point $(\vec{x}, \vec{\xi})$ of the 6-dimensional reduced phase space associated with the position and velocity of the particle. We assume that the molecules are hard spheres, whose center has position \vec{x} . The velocities after impact $\vec{\xi}_1'$ and $\vec{\xi}_2'$ are related to those before the impact $\vec{\xi}_1$ and $\vec{\xi}_2$, by (3.1). In the absence of collisions, $P^{(1)}$ would remain unchanged along the trajectory of the particle.

Accordingly, we must evaluate the effects of collisions on the time evolution of $P^{(1)}$. Note that the probability of occurrence of a collision is related to the probability of finding another molecule with a center at exactly σ from the particle under observation.

Thus, in order to write the evolution equation for $P^{(1)}$ we shall need another equation $P^{(2)}$ which gives us the probability of finding one particle in $(\vec{x}_1, \vec{\xi}_1)$ and the other in $\vec{x}_2, \vec{\xi}_2$. Obviously: $P^{(2)} = P^{(2)}(\vec{x}_1, \vec{x}_2, \vec{\xi}_1, \vec{\xi}_2, t)$. Hence, $P^{(1)}$ satisfies an equation of the following form:

$$\frac{\partial P^{(1)}}{\partial t} + \vec{\xi}_1 \cdot \frac{\partial P^{(1)}}{\partial \vec{x}_1} = G - L. \quad (3.2)$$

$L d\vec{x}_1 d\vec{\xi}_1 dt$ gives the expected number of particles with position between \vec{x}_1 and $\vec{x}_1 + d\vec{x}_1$ and velocity between $\vec{\xi}_1$ and $\vec{\xi}_1 + d\vec{\xi}_1$ that disappear from those range of values because of a collision in the time interval between t and $t + dt$, and $G d\vec{x}_1 d\vec{\xi}_1 dt$ gives the number of particles that, instead, enter the same range.

Counting these numbers is easy, provided we use the trick of imagining particle 1 as a sphere at rest and endowed with twice the actual diameter σ and the other particles to be mass points with velocity $(\vec{\xi}_i - \vec{\xi}_1) = \vec{V}_i$ (relative velocity).

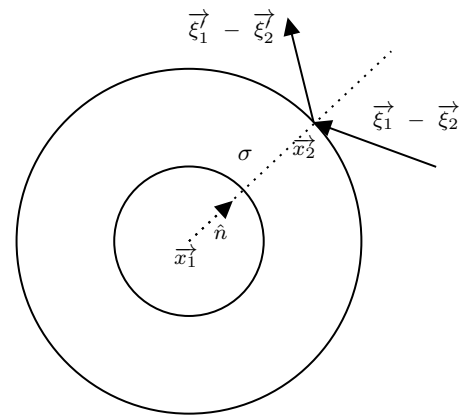


Figure 1: Example of a collision

In fact, each collision will send particle 1 out of the above range and the number of collisions of that particle will be the number of expected collisions of any other particle with that sphere. Since there are exactly $(N - 1)$ identical mass points and multiple collisions are disregarded, $G = (N - 1)g$ and $L = (N - 1)l$, where the lowercase letters indicate the contribution of a fixed particle, for example particle 2. We shall then compute the effect of collisions of particle 2 with particle 1.

Let \vec{x}_2 be a point of the sphere such that the vector joining the center of the sphere with \vec{x}_2 is $\sigma\hat{n}$, where \hat{n} is a unit vector.