

Foundations of Statistical Mechanics

Paolo Bedaque

April 10, 2016

The object of Statistical Mechanics is to study systems with a large number of degrees of freedom by focusing solely on the macroscopic properties of the system.

Why, you may ask, are we only focusing on these macroscopic properties? Surely, given enough information we could model any system by following its individual members! This is not necessarily the case because the sheer volume of information necessary to do such a calculation is in fact an impediment to performing such a calculation. Let's illustrate this using a simple example of a system consisting of 20 particles in the ground state. Suppose, through divine intervention or otherwise, we have obtained the complete wave equation for each particle in the system. Before we begin to calculate any properties of this system we must first store this information.

As shown in figure 2, every particle wavefunction will be reasonably approximated by 10 points in each dimension. Taking the 3-dimensional case, we find that the amount of memory we would need would be $(10)^{3N} * 8 \simeq 10^{61}$ bytes. We include the factor of 8 since each real number will be represented by 8 bytes. Assuming we store all of this data on terabyte drives that have a storage capacity of $\frac{10^{12} \text{ bytes}}{0.4 \text{ kgs}}$, we find that the total mass of all of the drives required to store the information describing this 20 particle system to be on the order of 10^{49} kilograms, which is much much more than the mass of the Milky Way galaxy! Clearly, it is impossible to store this much information, much less manipulate it. Therefore, we employ statistics to deal with the system at a macroscopic level.

1 Hamiltonian Dynamics

Consider a hypothetical system consisting of some number of particles which can be described by their individual positions, q_i and their momentums, p_i . We can indicate the paths these particles travel on a **Phase Space** diagram.

This diagram displays the evolution of the systems constituent particles as their positions and momentums change. The evolution of this system is governed by the Hamilton Equations: $\frac{dq_i}{dt} = \frac{\delta \mathcal{H}}{\delta p_i}$ and $\frac{dp_i}{dt} = -\frac{\delta \mathcal{H}}{\delta q_i}$. The evolution of p and q are clearly deterministic, therefore we can express these values as time-dependant functions. $q_i(t), p_i(t) \rightarrow q_i(t=0) = q_i^0$ and $p_i(t=0) = p_i^0$

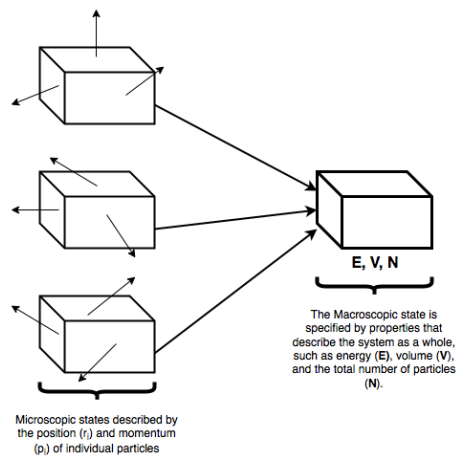


Figure 1: Multiple microscopic states make up a single macroscopic state of a system.

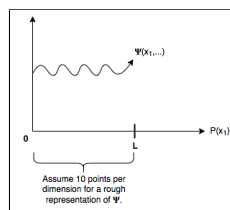


Figure 2: Wavefunction of our system in dimension x_1

There are a couple of tricks we can employ to idealize the measurement process and simplify manipulations of these functions.

2 Idealization of the Measurement Process

Let's start by defining some value \bar{f} , which is the result of measuring function $f(p_i, q_i)$.

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{\delta}^{\delta+T} dt f(q(t), p(t)) = \tilde{f}(q_i^0, p_i^0)$$

T = duration of the measurement, δ = time the measurement starts, \tilde{f} = some function of the initial conditions.

This expression is an idealization of an arbitrary measurement. In reality, the time T is much larger than the microscopic scales in which the functions of $q(t)$ and $p(t)$ vary. For example, suppose the function f is the pressure on the wall of a gas containing vessel. The quantity $f(q(t), p(t))$ fluctuates like the diagram below.

However, except in specialized cases, the measuring apparatus is not accurate enough to capture every tiny fluctuation of the force on the wall of the vessel. Instead, it measures an average of $f(q(t), p(t))$ over a finite period of time:

Taking the $\lim_{T \rightarrow \infty}$ is a mathematical idealization of this averaging out of function fluctuations. Returning to our analysis of the hypothetical measurement \bar{f} , we see that the function $\tilde{f}(q_i^0, p_i^0)$ is not really a function of the initial condition. Rather, it is a function of the *trajectory* of the system evolution. This may not be immediately apparent, but can be illustrated by choosing another initial point on the trajectory of our function $\tilde{f} = (q_i^0, p_i^0)$. As $T \rightarrow \infty$, we see that our new initial points $q_i^{0'}$ and $p_i^{0'}$ yield the same value of \tilde{f} as q_i^0 and p_i^0 . i.e., $\tilde{f}(q_i^0, p_i^0) = \tilde{f}(q_i^{0'}, p_i^{0'})$. This is because the difference between the two trajectories is negligible compared to the infinite extended trajectories as $T \rightarrow \infty$.

Now, it could be possible that two different trajectories give different values for \bar{f} .

However, it turns out that, for systems we are interested in, that it does not happen. These systems have the property that any individual trajectory will pass through almost **every** point in phase space. This is called the **Ergodic Property**.

A couple of important things to keep in mind. The assumption that the trajectory passes through every point in the phase space is contingent on the assumption that these points have the same value of energy (and any other conserved quantities) as the initial point. After all, these values are conserved by the Hamiltonian Flow. The phrase "almost every point" also has a specific meaning, in that it is "almost every point" in the measure theory sense.

It is really difficult to prove that a realistic system has the ergodic property. Only a few systems have actually been proven rigorously to be ergodic (for example, the Sinai Stadium). Moving forward, we will operate on the assumption

that the systems we are interested are, in fact, ergodic.

To expand on the concept of ergodicity, it is very easy to find systems that are not ergodic. Any conserved value restricts the trajectories to lie on a submanifold of the phase space. If the only conserved quantities are the ones resulting from the standard symmetries (i.e. energy, momentum, etc.) we can question the validity of ergodicity within the restricted subspace with constant values of these quantities. There are systems with so many other conserved quantities that the subspace is one dimensional (integrable systems), thereby invalidating the ergodicity of the system. The Integrable/Ergodic classification is also non-binary, there are intermediate types of systems that exist between the two extremes, such as mixing systems.

3 Introduction to Thermodynamics

At its core level, many thermodynamic problems can be phrased as a comparison between two states of a system. The pre-evolutionary equilibrium state and the post-evolutionary equilibrium state.

Since every microstate of the system is equally probable, the final macroscopic state will be the one which corresponds to the largest number of microstates:

$$\Gamma(E'_1, V'_1, N'_1, E'_2, V'_2, N'_2) = \Gamma_1(E'_1, V'_1, N'_1) \Gamma_2(E'_2, V'_2, N'_2)$$

Where $\Gamma(E'_1, V'_1, N'_1, E'_2, V'_2, N'_2)$ is the number of states of the composite system, $\Gamma_1(E'_1, V'_1, N'_1)$ is the number of microstates of subsystem 1, and $\Gamma_2(E'_2, V'_2, N'_2)$ is the number of microstates of subsystem 2.

Let's look at the case where the wall between the two subsystems allows the exchange of energy but not of volume or particle number. In this case we find:

$$\begin{aligned} E_1 + E_2 &= E'_1 + E'_2 \\ V_1 &= V'_1, V_2 = V'_2 \\ N_1 &= N'_1, N_2 = N'_2 \\ \Gamma(E'_1, E'_2) &= \Gamma_1(E'_1) \Gamma_2(E'_2) \quad (1) \end{aligned}$$

This implies that

$$\Gamma_{Max, E_{total} fixed} \Leftrightarrow \frac{1}{\Gamma_1} \frac{d\Gamma_1(E'_1)}{dE'_1} - \frac{1}{\Gamma_2(E'_2)} \frac{d\Gamma_2(E'_2)}{dE'_2} = 0$$

So, knowing the total number of initial microstates of the two subsystems allows us to predict the final state of the overall system. This is summed up in a very clean and useful fashion in the definition of **entropy**:

$$S = k_B \ln(\Gamma)$$

Notice that entropy is extensive as long as the subsystems are large enough that the boundary effects are negligible.

How can Entropy grow?

Within the constraints of the problem, there is an apparent paradox in staying that the entropy is maximized. That's because, using Hamiltonian equations, we can show S is a constant.

Define $\rho(q, p, t = 0)$ to be an initial distribution of identical systems (this is called an **ensemble**.) As the Hamiltonian Flow evolves in time every element of the ensemble is carried with it and the distribution evolves to $\rho(q, p, t) = \rho(q(t), p(t))$.

Because the number of systems in the ensemble is fixed ρ obeys the continuity equation:

$$\frac{d\rho}{dt} + \nabla \cdot (\rho \nabla) = 0$$

or

$$\begin{aligned} 0 &= \frac{\delta \rho}{\delta t} + \sum_{i=1}^{3N} \left[\frac{\delta(\rho \dot{q}_i)}{\delta q_i} + \frac{\delta \rho \dot{p}_i}{\delta p_i} \right] \\ &= \frac{\delta \rho}{\delta t} + \end{aligned} \tag{2}$$

This result is known as the **Liouville Theorem**. It implies that the density of elements on the ensemble does not change as it evolves along the Hamiltonian Flow (of course, it does change at any fixed position in phase space). As the number of systems in the ensemble is fixed, the Liouville Theorem implies that the volume of the phase space occupied by the ensemble is fixed.

This means that the entropy, being the log of the volume, cannot change either. So how can the entropy grow if the microscopic equations of motion show that it does not? A resolution to this paradox is to notice that, for systems in which statistical mechanics holds, the Hamiltonian Flow takes nice looking, civilized ensembles and turns them into convoluted shapes:

The average of smooth observables (i.e. the value of $f(q, p)$) does not distinguish between the average performed with the true $\rho(t)$ or the "coarse grained" $\tilde{\rho}(t)$, as shown below.

$$\bar{f} = \int dq dp f(q, p) \rho(q, p, t) \simeq \int dq dp f(q, p) \tilde{\rho}(q, p, t)$$

For all practical purposes, the entropy grows so long as we only look at observables that are very smooth. Macroscopic observables don't care about the precise position of the particles and tend to be smooth in phase space. Systems whose Hamiltonian Flow have this property are called "**Mixing**." It is easy to show that ergodic systems are mixing, but keep in mind that mixing systems are not necessarily ergodic.