Guiding Principles of Statistical Mechanics

May 8, 2011

1 Kinetic Theory

We will concern ourselves only with classical statistical mechanics, i. e. the Hamiltonian is a scalar operator of q, and p. Kinetic theory concerns itself with the 3 following questions:

- What is statistical equilibrium?

- Why/When do things go to equilibrium?

- How fast?

Definition of Statistical Equilibrium The meaning of statistical equilibrium is ambiguous. Here's why: on a qualitative meaning, it means "macroscopic observables" don't change at equilibrium. But it depends on what you mean by macroscopic observables. For example, glass is at thermal equilibrium on time scales of less than seconds, but is NOT in mechanical equilibrium over thousands of years because it can be then considered like fluid that flows. We can, however give a rigorous definition of equilibrium as follows. Suppose we have a ensemble of microstates of our system of interest. Each microstate is a point in the phase space of possible configurations, i. e $\in \{\vec{q}, \vec{p}\}$. Then those ensembles will evolve over time according to classical mechanics. the probability of finding a certain ensemble in a element of volume dV of phase space is:

$$dP = \rho(p, q, t)dV$$

where ρ is a probability distribution over phase space describing all the ensembles.

Now. equilibrium, means that $\rho(p, q, t) = \rho(p, q)$, meaning that the probability distribution of interest does not change over time, which means, by Liouville's theorem:

$$[H, \rho(p, q)] = 0$$

THis also allows justification for the probability of a microstate of given energy to be the same, the microcanonical assumption, since, if $\rho(p, q) = \rho(H(p, q))$, then:

$$[\rho, H] = 0$$

follows trivially

Why do they go to statistical Equilibrium? Boltzmann H theorem. See Kardar $\rightarrow$ leads to entropy increase $\rightarrow$ leads to maximization of entropy/shannon entropy picture in stat mech.

How fast? Depends on the system in question.

2 The technology of Statistical Mechanics

Introduction The reason why the technology of stat mech – the Partition function, the Microcanonical Ensemble, the link to The link to thermodynamics – might not seem intuitive is because we have to make a clear divide when we make the following approximations:

- When do we replace averages by their most probable values and finally by their thermodynamic macrostate variables?

- How does that create Thermodynamic limit of Stat Mech?

Here are **Some misconceptions** I think I might have induced in the last time we met. I'd like to clarify them.

- A given microcanonical ensemble of a system is DISTINCTLY DIFFER-ENT from its canonical ensemble. The former considers a bunch of different copies of the system in the FIXED energy E, while the latter considers a bunch of copies of the system with fixed T, or other intensive parameters like Pressure, Electric field... In the thermodynamic limit, all copies have energies very very very closed to E. So they become asymptotically similar in the thermodynamic limit, but are otherwise different ensembles.

Because of the divergent techniques of stat mech, I'd like to focus on the following concept: given a macrosopic average value for an ensemble of copies of a system, what are the microscopic probabilities of the microstates? That question is the crucial link between thermo and stat mech: macroscopic averages over ensembles give you what you thermodynamically measure (by ergodicity), and microscopic probabilities is what you deal with using the technology of stat mech.

Equilibrium We defined equilibrium before as probabilities of microstates over the ensemble are time independent, but that also means ensemble averages are time independent, which means when you specify the macrostate of a system at equilibrium, what you are REALLY specifying are ensemble averages.

3 The Shannon Picture

If you believe that nature ensures maximal entropy via interactions (Boltzmann H), and that shannon entropy is a good measure of entropy, you believe in the Shannon picture. Basically, it says for a given ensemble of copies of a system, and you let them evolve according to the laws of classical mechanics, the probability distribution $\rho(q, p)$ will maximize entropy. defined by:

$$S = \sum_\mu p \ln(p)$$

Suppose $f_i(p, q)$ are functions of the microstate in question, i.e. can be total energy. The, maximizing entropy with the following constraints:

$$\langle f_i \rangle = F_i$$

leads directly to the bolztmann factor, if we are considering distinguishable particles, Fermi-dirac Factor if we are considering Fermions, and Bose-Einstein factors if we are considering Bosons.

For Boltzmann, we have the following probability maximizing distribution:

$$p(\alpha, \lambda_1, \lambda_2, ...) = \frac{1}{Z} e^{-\alpha - \sum_i \lambda_i F_i}$$

where Z is a normalizing factor for probabilities, α is the lagrange multiplier associated with conservation of probability (or particles, same thing), λ_i is the lagrange multiplier associated with the constraint F_i.

Because probabilities have to sum to 1,

$$Z(\alpha, \lambda_i) = \sum_\mu e^{-\alpha - \sum_i \lambda_i F_i}$$

Several Remarks

- For a given constraint set ($\{\alpha, \lambda_i\}$), the probabilities among microstates satisfying this set are equal. This is the principle of equal apriori probability, trivially satisfied by Shannon Entropy.

- A special case of this point is when F is average energy. In that case, a microcanonical ensemble which assumes equal probability density for an iso-energy hypersurface in phase space is directly derived from this fact.

- Different sets of $\{F_i\}$ generate different canonical ensembles: Gibbs canonical etc...

- We can derive relationships between the λ_i and F_i via pure dimensional analysis arguments.. $\rightarrow$ I'll let you think about that, too tired right now.

- Equilibrium becomes a trivial fact of mathematics. Suppose I bring in a different system and let them equilibrate. Then in the end, the λ_i's for both system must be equal by maximization. Hence, the λ_i's or some function thereof will be the parameters that determine thermodynamic equilibrium. Since equilibrium is intensive, obviously, the λ_i's are functions of intensive parameters... hint for last question raised up.

- If you go through the math, you'll see all the information thermodynamic information we actually measure ($\{F_i\}$)is actually encoded in Z. That's why it's called the "generating function". Specifically,

$$\partial_{\lambda_i} \ln(Z) = F_i$$

and

$$\partial_{\lambda_i}^2 \ln(Z) = \left\langle (f_i - F_i)^2 \right\rangle$$

This, in fact, is more easily seen via the Laplace Transform picture which is next.

4 The Laplace Transform Picture

This picture is the mathematically simplest, but the physically most unsatisfying. We will generalize the argument in the Legendre Transform paper.

Suppose we have a function of density of states for a fixed constraint set, F_i. Then,

$$\Omega_i(F_i) = \int dq dp \delta(f_i(q,p) - F_i)$$

We note that this assumption of counting with equal weight over iso-f_i surfaces is based on a priori probabilities, which in turn comes from maximization conditions of entropy. The reason why this is confusing is because the microcanonical ensemble considers a fixed $f_i = F_i$. The probabilistic nature of $< f_i >$ is hidden under the fact that different values of f_i leads to different volumes/hypersurfaces of phase space, + the fact of equal apriori probability. This is NOT the canonical ensemble.

This constrained integral is terrible to evaluate, so we take its laplace transform with respect to the parameter conjugate to F_i, call it β_i. This reason is purely dimensional-analytic for me, but if you have a better motivation, please point it out.

$$Z(\beta_i) = \int \Omega_i(F_i) e^{-\beta_i F_i} dF_i$$

This is the same partition function as obtained above when there is one constraint, specifically:

$$\langle f_i(q,p) \rangle = F_i$$

4

What is cute, is that the cumulants of f_i are generated directly by the expansion:

$$\ln(Z) = \sum_n \langle (f_i - F_i)^n \rangle \frac{\beta_i^n}{n!}$$

$$\rightarrow \partial_{\beta_i}^n \ln Z = \langle (f_i - F_i)^n \rangle$$

This fact might be easier to digest if we consider a laplace transform to be a fourier transform with the contour of integration rotated.

Remarks

- This picture lends itself nicely to the idea of smoothening out" information when you compress information via cumulants or averages. In signal processing, you encode a picture by its fourier transform, throwing out extra information. Similarly, the function Z encodes for the ρ, except here, we are throwing out the extra cumulants in the thermodynamic limit. In other words, when N is large, by CLT, we can just need the 1st cumulant of a distribution (the constraint average) and arguments about equal a priori probability to reconstruct ρ

- Z can be evaluated via the method of steepest descent. In the thermodynamic, it just picks out the most probable value of $\Omega(E_{mp})$, however, because of CLT, the distribution of $\Omega(E_{mp})$ spikes up like a gaussian goes to a delta function. The distinction between most probable and mean becomes irrelevant, hence, we don't need to distinguish between canonical and microcanonical ensemble. This is the crucial distinction: mean vs. most probable.

- Converting an integral by method of steepest descent is the same as converting a fourier to a legendre transform in the large N limit. Hence, fourier transforms in stat mech links to legendre transform in Thermodynamics.