Q&A Lecture 10

With a matrix the trace is the sum of all the diagonal terms, however what does this represent physically (or more abstractly if it makes more sense to start there)?

Not sure I understand the question: the trace is a mathematical operation on matrices (or more generally operators). Formally, the trace of the product of two operators is one of the most common definitions of a “scalar product” between operators. The physical meaning, for me, is associated with the definition of the quantum average in which the trace is used to get the scalar product between the probability density and the observable. This can be read in analogy with the classical definition of the average which is also a scalar product (this time defined on the space of functions of phase space via the integral) between the (classical) probability density and the (classical) observable.

If we use the trace in calculating the quantum canonical average, does this imply that the system is discretized? And if so, how is this then analogous to a classical canonical average?

No, the trace is an abstract operation in which we sum over all diagonal elements of the matrix represenation of an operator (in our case, the operator resulting from the product of the probability operator and the operator representing the observable). This sum is discrete if we choose a discrete basis or continuous (and then becomes an integral) if we use a continuous basis. In the next lecture, we’ll play with the coordinate representation and see more about this.

The two approximations (The approximation that electrons go to the most stable site immediately following the movement of atoms in BO approximation and the approximation that electrons stay at the energy-minimum site in Mass Zero constraint) look very similar. Although I know that the equations of these methods are different, are there other differences in the sense of their concepts?

They are quite different conceptually: in BO the electronic variables are parameters to be minimized, in MaZe they are DYNAMICAL VARIABLES to be propagated.

The question that I had this week is about the physical interpretation of the terms of the added constraint force in mass-zero MD. I am having trouble understanding why the electronic density dynamics are in this form. P.ex. what does lambda exactly represent, and what does the derivative of sigma is ?

We are imposing the condition of minimum of the energy with respect to the electronic degrees of freedom as a constraint on their motion (sigma = derivative of the energy with respect to the electronic coefficients = 0 expresses the constraint) . To satisfy this constraint, we need to impose a force: the derivative of sigma times the lambda is this force.

Why is the kinetic energy average only an estimator and cannot be considered as intrinsically linked to the Temperature?

It is…as an estimator! The Temperature is the macroscopic quantity that we measure with a thermometer. This corresponds microscopically to the AVERAGE of the kinetic energy of the particles.

What is the inertia parameter Q in our expression for η dot ? Does it determine the efficiency of η at influencing our system and how do we determine or choose this value?

Q represents how “fast” the thermostat’s degrees of freedom adjust to changes in the kinetic energy of the system. So, in a sense, it does determine the efficiency of \eta as an influencer. It Q is very large, the system takes a long time to relax to the desired temperature but once it is there it’ll be difficult to “change the temperature”. If Q is small, the system will go quickly to the target temperature but can also leave it. In the first case we take long to relax but we have small fluctuations around the target temperature, in the second the opposite is true. As usual, the choice is a compromise between these case.

What would be the physical interpretation of the Nose variable η if there is any ?

I don’t think that there is a good physical interpretation. Also, note that the crucial fictitious variable is p_\eta that acts as a generalised (not constant, not a definite sign) friction coefficient on the physical variables.

To compute classical canonical average, do we need to solve equations for p, r, η and p_η until the system reaches the target temperature? When the dynamics is described by equations for p, r, η and p_η, is it possible that system's temperature fluctuates near target temperature?

Yes, this is called the equilibration phase of the run; Yes, there will be fluctuations. Fluctuations - however - exist also in the canonical ensemble where the distribution is not a delta in T. What is important is that it can be proven that the Nose - Hoover dynamics gives the correct (i.e. same as in the analytical ensemble) fluctuations. This is critical because the fluctuations are related to observables - e.g. specific heat…

Can you explain again why and how we introduce the Nosé variables ?

The idea is to extend the dynamical system to drive, via coupling with the nose variables, the momenta of the physical system to adapt and assume values such that the total physical kinetic (barring constants) is around the target temperature. This is done by (a) determining changes in the momenta of the fictitious variable (so p_\eta) that respond to variations in the instantaneous estimator of hte physical kinetic energy, and (b) by constructing an additional term in the changes of the physical momenta to make them slow down (cooling the system) or grow (heating the system) based on the changes of p_\eta. This is the intuition. It can be proved mathematically that the dynamical system that we constructed in class samples the classical canonical probability density.

How do we select the target temperature? Is it guided by some physical properties of the system we are studying? Also, due to the two-way viscous nature of our ficticious medium, will we not have similar problems with oscillations as before?

The target temperature is fixed by the experimental conditions. The running average of the microscopic temperature estimator will indeed fluctuate. The important thing, however, is that the total average will be equal to the target temperature and that the fluctuations of the average will be the same as those of the analytical canonical ensemble.

We wrote that the average of the ensemble was given by the trace of the probability operator times the observable operator, is that thing only valid when the state representation is diagonal to the whole operator ? Furthermore, couldn't we determine this average from the wavefunctions, as they are related to the probability of the states, at which we could apply the desired operator to measure any observable (assuming that we can solve the Schrödinger equation) ?

The definition via the trace is general. In fact, it can be shown that the trace (which is a scalar) is independent of the choice of the basis; Great question. Probability enters in quantum mechanics in two ways. The first is at the level of the wavefunction in the interpretation of measurements. This refers to the probability to get as the result of a measure one of the eigenvalues of the operator being measured when the system is in a given state. Classically (stretching things a bit) this would be equivalent to asking the value of a microscopic observable when the system is in a single state or point in phase space (this value is unique classically). The second way that probabilty enters in quantum mechanics is when the system can be found in a set of states, each with a given probability given - for example in the case discussed in class given by the weight of the canonical probability. Classically (stretching things a bit less) this corresponds to asking the average value of a microscopic observable when the system is in a given ensemble. In the first case, in quantum, we say that the system is in a pure state. In the second, we say that the system is in a statistical mixture.

I don't clearly see the link between the classical canonical statistical mean, where we sum (with the integral) over all possible R and P and the quantum canonical mean which sums only over the diagonal elements (trace) of the pi * A operator?

The trace operation is one of the definitions of a scalar product in the space of operators. One can see the classical average as the scalar product of the probability density and the function representing the microscopic estimator. So, this is a full analogy. You can also look (stretching it a bit) at the integral in the classical average as a sum over states. Because we have only one integral(sum) this is a "diagonal" set of products of the two functions in the integrand. I may say a bit more on this in class.

Car-Parrinello vs. Mass-Zero Constraint Dynamics for Studying Vibrational Spectra: What Are the Implications of Freezing the Electron Density in the Latter?

In Car-Parrinello, vibrational spectra are (sligthly) shifted because the finite mass of the electronic degrees of freedom induces a (small) bias on the distribution of the momenta of the ions with respect to the exact Born-Oppenheimer probability. This bias is removed by the zero mass limit in MaZe.

I'm not sure to have understood why we want to solve the evolution equation for a fixed temperature. Is there a difference in the way we compute the average of an observable ?

The expression to compute the averages, both classically and quantum mechanically, in all ensembles has the same structure: sum over the states of the product of the probability density times the observables. What changes depending on the ensemble is the expression of the probability density. We may want to have constant temperature to mimic experimental conditions.

Is there an analytical expression for the Nozet variable, or is it guessed? And are there algorithms for QM dynamics in the Grand canonical ensemble with varying particle number, or is it too complex?

Analytical means what here? I would say that the form of the Nose dynamics is constructed to mimic the behaviour of a thermal reservoir. The evolution equations of the Nose variables then create a feedback mechanism for the physical degrees of freedom that keeps the microscopic temperature around the target one. There is no analytical solution for the evolution of these variables, but we can obtain their trajectory numerically via a generalisation of the Verlet algorithm. I must say that I don't think I have ever seen a grand canonical quantum calculation…not sure why though.

In the canonical, we look at the kinetic energy which allows us to create something, but the hamiltonian to compute the canonical average still depends on the potential energy, wrt to what system (we should have some kind of finite space, otherwise the potential energy could we any value if we play with the referencial) do we compute it?

The answer to this question is a bit different in classical and quantum. In classical, the estimator of the microscopic estimator of temperature is (usually, it is not the only possibility) related only to the kinetic energy. Since the classical canonical probability factors into the exponential of the kinetic energy times the exponential of the potential, we can control the temperature looking only at the kinetic energy for ANY potential, i.e. no need of a reference what we discussed works for any potential. In quantum, the uncertainty principle couples the kinetic and potential parts in the probability (this we saw yesterday when we noticed that the integral of the exponential of the kinetic energy in path integrals results in a product of coordinate dependent Gaussians). As a consequence, the estimator of the temperature - or of the kinetic energy - involves both the kinetic and the potential terms. In path integrals, we can however use the classical isomorphism discussed yesterday to still use a Nose-Hoover like dynamics for any W (not V in the notation of the lecture).

*When we implement the mass zero constrained dynamics, why the dynamics of the iron is changed? Is it because we change the total energy E[rho_G, R] into E[rho_G, R] -\lambda * \sigma?*

The dynamic of the ions is changed ONLY IF we don't take the zero limit for mass of the electronic degrees of freedom. Before the limit, the change comes from the constraint force which is indeed due to the term in the "energy" proportional to \lambda. After the limit is taken the dynamics of the ions is unchanged compared to the original problem.

Still the question about solving the variational principle of min_{\rho_G} E[\rho_G, R]. If I understand correctly, the formula of mass zero constrained dynamics is similar to the idea of Nesterov accelerated gradient descent, which consider the dissipation via the use of momentum. There are many well established opitimization methods that can solve this minimization problem. Why don't we use them directly? In this way, we just focus on the optimization problem of the E[rho_G, R] with a fixed R, and the dynamics *of the iron won't change. Is it because solving the mass zero constrained dynamics is more efficient? (I am not sure which method is more efficient.)*

I dont know Nesterov but, based on what you write, I think that MaZe has a few advantages: the most important is the promotion of the electronic degrees of freedom to dynamical variables that obey an inertial (not dissipative) dynamics. This means that we have a formal framwork to discuss the formal (e.g. statistical ensemble) and algorithmic (e.g. time-reversible) properties of the system. Of course, it is possible that alternative optimization method might be more efficient…so far we did not find one that combined efficiency and interpretation…