

Q&A Lecture 10

If my microscopic operator had a dependence on the momentum, I'm assuming we could redo the analysis in the momentum basis, with appropriate fourier transform, but the resulting expression would not resemble classical dynamics anymore since the momentum terms cancel out in the classical case - however, can classical integrators still be reused in this case?

Actually, for momentum-dependent observables, the momentum basis is used only for the matrix element of the observable, or - oftentimes - estimators based on the coordinate representation of the momentum (i.e. a constant times the derivative with respect to the position) are taken as the starting point of the manipulations. The matrix element of the exponential of the Hamiltonian is still left in the coordinate basis. This is because a full momentum representation for the matrix elements involving the potential is typically impossible (how would you write, for example, something like a Morse potential in the momentum representation?). I'll say a bit more about this in the next class.

Out of curiosity, how would we tackle systems with temperature gradients in it while exchanging with the outside? Maybe these systems aren't of importance at the length scales that we are now interested in, that is also a possibility.

One way would be to create two slots of material (finite length) on the two sides of the simulation box that you wish to keep at different temperatures and thermostat the atoms in those slots at different temperatures. You can allow transfer of particles between these slots and the inside of the box or not. In any case, the asymmetry will generate a thermal gradient in the system (i.e. the interior of the original box) that would not be thermostatted - so simulated at NVE.

Why only talk about the quantum in the case of NVT, also get the density operator based on NVT, so we do not need to consider the case in NVE?

The probability density in the quantum microcanonical ensemble is defined in analogy with the classical one as the uniform probability defined by the delta of the Hamiltonian operator minus the "target" energy. However, this ensemble is used very rarely (meaning that I have never seen a calculation in realistic models molecular or solid state physics, it's different for Ising models). The canonical is more interesting from an experimental point of view - thermal equilibrium vs isolated system - and practical - easier to compute via path integrals.

In the last equation that we derived from the path integral formalism, by rearranging constants we end up with the ring polymer Hamiltonian. Then we can compute quantities such as $\langle O(r) \rangle$ for observables diagonal in the computational basis and use the empirical mean over our samples to compute this quantity. My problem is when we are to sample with NH or Langevin, how are the samples over the beads generated? Do we solve equations of motion under the Ring Polymer Hamiltonian, then we use Langevin or Nosé Hoover with the EOM to generate samples, and this for all beads?

The procedure you describe is the one adopted: thermostatted evolution equations for all beads. What is important to remember is that the path integral expression of the probability density DOES NOT involve momenta (the quantum kinetic energy is represented by the "springs") and so the additional momenta in the RPMD are added by multiplying and dividing by the classical distribution of momenta. So, the momenta distribution is - in general -

meaningless from a quantum point of view but it is part of our “extended” system to generate trajectories...another example of fictitious dynamics....

I don't remember to which part of the Langevin method was the fluctuation-dissipation theorem related to.

Without this relationship between the friction and the variance of the Gaussian random number that determines the intensity of the kick, the Langevin equation does not sample the canonical probability (this can be seen in the proof of this sampling via the Fokker-Planck equation).

Could you also briefly re-explain what the fluctuation-dissipation theorem is.

It is a statement on the condition necessary to have compensation between normal (equilibrium) fluctuations and response to an external perturbation so as to keep a non-isolated system in statistical equilibrium (by having a stationary probability distribution and therefore averages of static observables that do not depend on time).

When we say we are in equilibrium, do we assume a state of dynamic equilibrium, where the particles could escape and re-enter the system volume, such that the total number remains constant?

No, in NVT simulations, the number of particles in the system is strictly constant. There is a simulation set up called periodic boundary conditions in which one minimises boundary effects on the finite volume of the simulation box by allowing particles to escape from the box and re-enter it, but this is not like allowing the number of particles to change dynamically.

I assume that similarly to the discussion about the choice of the timestep for having a valid Trotter decomposition at minimal cost, we can also tune the number of auxiliary positions we add to compute the probabilistic weight. For the timestep we could look at the conserved energy but here which indicator can we use to know that our decomposition is "sufficient" enough?

Convergence tests on the average values computed with increasing number of auxiliary positions.

*And would the computational costs scale by $3N * M$ with M is the decomposition integer (correlated to the number of auxiliary positions)?*

The direct scaling is indeed the one you indicate. There's a lot of research that goes into finding algorithms that scale better, a family is known as contraction techniques in which some of the beads are “coarse grained” to reduce the number of additional degrees of freedom. We shall discuss a (drastic) alternative in the next class.

At the end of the class, it was said that what was done in the analysis of the quantum NVT ensemble was path-integral formalism. I understand the mathematical derivations that were performed; however, I don't understand why do we call this path-integral formalism, or what part of what we have done in the derivations was path-integral formalism?

As we shall see in the next class (or the one after that) the name “path integrals” was originally decided for the expression of the real time propagator in the coordinate basis “massaged” via exactly the same Trotter operations that we have seen. You can see the analogy by making the change $\beta \rightarrow -it/\hbar$ and think about the thermal density as a propagator in imaginary time.

Then - again more on this in the next class - you can see the sequence of integrals that we have introduced as a “sum over all possible paths in imaginary times” that the quantum system moves on to go from x_0 to x_0 .

Was the goal of path integral formalism to show that the expectation value in both quantum and classical NVT ensemble has the same mathematical structure?

Yes

And, is this the general goal of path integral formalism, to show that there are concepts that both in quantum and classical mechanics have the same mathematical structure?

It is certainly one of the advantages: by finding classical analogies of quantum properties, we gain some intuition. This needs, however, care as we could be misled in over-interpreting. The fact that the mathematical structure of a quantum average looks like the classical average of an extended system also leads to computational advantages since we can use classical algorithms to compute quantum averages.

In the Langevin evolution equations, we sample the random kick randomly from a Gaussian distribution. How does this random sampling allow for an accurate description of the dynamics? For example, if the random kick is sampled at a higher end of the distribution, would it not disturb the trajectory of the particle?

The kick and the friction (both) are indeed perturbations to the dynamics, but they are constructed so as to keep the average kinetic energy around the target value. As discussed for ALL the invented dynamics that we have seen until now, the trajectory itself does not need to be the “true” physical trajectory. All that we require is that it visits different regions of phase space with a frequency proportional to the target probability. In the case of Langevin, it can be proven that this is the case.

It is not very clear for me why when you derived the expression for quantum averages you said that the $R_y = R_o$ as if we were doing a loop?

The definition of the quantum average is given by the trace of the product of the probability times the operator corresponding to the observable. The trace is, by definition, the sum over the diagonal elements of the representation of the product in an arbitrary basis (the value of the trace does not depend on the choice of the basis). In the coordinate basis, this means that we consider the diagonal element $\langle x_0 | () | x_0 \rangle$ (where $()$ stands for the product of the probability times the operator). The following manipulations do not change the fact that we “start” at x_0 and end at x_0 so...a loop. I'll say a bit more about this in the next class.

In Langevin dynamics, a constant frictional force and a random positive force is used to have the system sample the same classical NVT distribution. Does there exist a hybrid where the force is proportional to the velocity but with a factor that is sampled from a random gaussian distribution centered at 0 so that it can be both positive and negative?

You mean a positive or negative friction? Not that I know of.

How to physically interpret the auxiliary particles introduced when we study the quantum system?

There is an interpretation of the classical system and there are quantum properties related to averages over the auxiliary particles. I'll say more about these in the next class.

Now since we turned one quantum particle into N classical particles, and the $W[\{x_j\}]$ represents the effective potential, we can evolve the system with Nose-Hoover. However, does the time evolution of the coordinates $x_0(t), x_1(t) \dots x_{N-1}(t)$ generated by the integrator has any physical meaning or it just samples probability $\propto \exp(-\beta W[\{x_j\}])$?

Rigorously, there is no meaning associated to the dynamics except for harmonic systems (or free particles) or in the very high temperature limit. We shall, however, see that this dynamics is used to approximate also time-dependent quantities...more on this in future lectures.

Is placing the auxiliary variables just a mathematical trick, or does it say something physical about the quantum system we are studying? Because I'm thinking we integrate over all variables so that means it goes from a mathematical trick to something that contributes in some way to the observable I'm measuring? I think.

There is indeed physical information in the (average) properties of the auxiliary particles. For example, the average distance squared of particles at "opposite" places in the chain of auxiliary particles provides information on how delocalised the quantum particle is. I'll say a bit more on this in the next class.

For a generic molecular system containing common atoms (H, C, O, N, ...), how many beads do we usually need for each atom in order for the ensemble average to converge?

The number of beads is determined numerically via convergence tests on the averages of the observables computed. In general, the "heavier" the atom, the less beads are needed. Typical calculations on water at room temperature are done, for example, with 32 to 64 beads per particle.

In the Langevin equation, how come the random kick is independent of the state of the system? I'd expect that, in order to sample the true distribution, the random kick would depend on the position and velocity of the particle.

I think that the correct physical picture is that the reservoir is so large that the state of the system is a small - negligible - perturbation on its state. The kick and the friction are, however, connected so as to keep the average kinetic energy of the system around (at within statistical fluctuations) the target temperature. They act on the acceleration so on how the velocity changes.

When passing onto the quantum picture, we use the same distribution function for the $\hat{\pi}$ operator as we would in the classical case. Why can we just "steal" the classical distribution function for the quantum case?

The fact that the thermal probability density remains invariant in form (of course with operators) where we had functions can be viewed (1) as a consequence of the classical correspondence principle - we have done this for all observables up until now; (2) proven via the quantum Fokker-Planck equation; (3) derived via a thermodynamic principle known as the maximisation of the entropy....

Finally, in the result of the last part of the lecture, did I understand correctly that we can model a single quantum particle with a ring polymer of classical particles? The operator \hat{A} corresponds to a single particle?

Yes, the operator \hat{A} is - in what we did in class - associated to a single position. However, using the invariance under permutations of the particles labels in the Hamiltonian - and therefore in the quantum probability - it can often be rewritten as the average of the observations over all the ring polymers particles. I'll say a bit more about this in the next class.

When using Trotter splitting for the quantum case, why is it still valid even though the operators inside H do not commute?

I find this question a bit surprising. Trotter is used PRECISELY to enable splitting of the exponential of the sum into the product of the exponential FOR NON COMMUTING OPERATORS in the sum, in our case "the operators inside H ". We have used this a few times now and it was also discussed in the last lecture, so I would recommend that you review the corresponding lectures, in particular the one on the SOFT algorithm.

We assumed that the observable A is a function of only position, but what about observables that are not, such as the kinetic energy or pressure?

Observables like the pressure and the kinetic energy (via the Hamiltonian) can be obtained via thermodynamic relations as derivatives of the partition function, i.e. the logarithm of the normalization of the probability. If one deals with other functions of momenta, a possibility is to express the matrix element of the observables via a basis change that goes to momentum representation and then comes back to coordinates - a bit like we did for SOFT. The only caveat with this is that the fact that the path starts and ends at the same point x_0 (we discussed this in connection with the fact that it comes from a trace - sum over diagonal elements) may no longer be true. That, however, is not a critical problem.