

Q&A Lecture 12

I see how the stationary phase approximation can help us calculate each matrix element of the time-evolution operator as one evaluation of $e^{[it(\text{action}/\hbar)]}$ at the classical path instead of the integral over all paths (if we neglect the quantum weights from the second order term). Once we calculate all matrix elements, we can get $\phi(t)$ from $\phi(0)$ with a simple matrix-vector product. However, I don't see how one would use RPMD to get to $\phi(t)$ as described in the last part of class. How does one turn a classical trajectory of a ring polymer calculated with Nose-Hoover dynamics into the object $\phi(t)$, which gives the probability of existence at each point in space?

Excellent point. We do not use RPMD to obtain the wave function but to obtain quantum averages that depend on time, e.g. the mean square displacement of the quantum particle as a function of time. If we want to reconstruct the wave function, we have to go via stationary phase approximation. However, essentially all relevant observables for the molecular systems that we are typically interested in come in the form of averages. I'll comment more on this in class.

In practical simulations, we know that the time step must be chosen smaller than the fastest physical timescale in the system. But how should we choose the spatial discretization? Specifically, how fine should the coordinate grid be, and what determines the appropriate grid spacing when discretizing the position axis for path-integral or semiclassical calculations?

In the full path integral time-propagator, one should discretize each integral over x_k with a grid fine enough to represent the smaller space variations of the wavefunction at all times. This problem is reminiscent of the problem of finding a good grid spacing in the SOFT algorithm and should be addressed with similar convergence tests. Note that, as discussed in class, the full space-discretise calculation would scale exponentially with the number of degrees of freedom (like SOFT) and with the number of steps since at each step one has to "move" to all possible discrete coordinates at the next slice. In the stationary phase approximation scenario, only two grids are left: the one for x_0 and the one for x_N . The grid spacing on x_0 depends on the characteristics of the wave function only at time $t=0$ (that we assume to know). The grid spacing on x_N is not known a priori and must be treated as in SOFT.

The path integral treats all trajectories democratically, but the classical limit picks a unique one. Is there a notion of entropy or information loss associated with collapsing an infinite path into one trajectory?

I am not aware of a specific notion of entropy related to this process. I guess - but it's more a hunch than an idea - that one could try to compare the entropic information associated to the classical and quantum propagators (as one would do via likelihood in stochastic processes) and start from there. Interesting.

Also, if the semiclassical limit picks out stationary phase, can two different quantum systems share exactly the same set of stationary paths but still have different quantum dynamics?

I don't think so. The Hamiltonian defines (univocally) the quantum system and the quantum propagator. If two systems have the same set of weights associated to paths they are the same system.

I got a little confused with the semi-classical and 'ad hoc' methods. So semi-classical is where I assume a classical trajectory and compute quantum weights that cause fluctuations while ad-hoc methods generate intuition based trajectories and the fluctuations come from nose-hoover?

Fluctuation are a general concept indicating differences of a given (relevant) quantity with respect to the average value. We have used this concept in different ways in class. In semiclassical methods, we prove that classical trajectories provide the most important contributions to the quantum propagation and we can account for the difference between classical and quantum dynamics - to a first order correction, not all the differences - by including the paths that are "close to the classical trajectory" and that we described as the fluctuations of the paths. This is all analytical and rigorous. In ad hoc methods, the trajectories are purely classical, but of a non-standard system (the ring polymer that represents a single quantum particles). The fluctuations induced by the thermostat are not there to include corrections to the classical dynamics but to ensure that the quantum thermal (NVT) equilibrium is sampled.

Also, how can these fluctuations be useful in my simulation they seem to me like they can be largely regarded as noise? Do they contribute to the observables or just provide approximations to actual dynamics?

They are not noise - in fact they contain physical information also in classical dynamics (e.g. fluctuations of the energy in NVT are a measure of the specific heat of the system) - and they can indeed contribute to the observables since they give both amplitude and phase to the approximate time-evolved wave function. Remember that the averages in the semiclassical language that we discussed so far are NVE and so observables are averages over the wave function.

Why we want semiclassical case, why not just left only the classical path in the integral? are not it more simple?

If you mean why I introduced the quadratic fluctuations, the answer is that including them means that we are using a higher order (and therefore more accurate) Taylor expansion of the action. You are right in saying that looking only at the classical path would be simpler, but it's a worst approximation to the quantum dynamics.

Also I do not undersand why during the lesson, you specifcilly pointed that $M/(2\delta t)[(x_{k+1}-x_k)^2 - V(x_k)]$ could be negative, so not a probability. what does this mean? are not $M/(2\delta t)[(x_{k+1}-x_k)^2 - V(x_k)]$ just represent $L(\text{lagrange})$?

L can indeed be positive and negative (think even to a classical system: the kinetic energy is not always larger than the potential energy). However, this is not the main problem: the problem is the i in the exponent that makes the $e^{i/\hbar S}$ an oscillating function that can have positive and negative values— it can be expressed as $\cos(i/\hbar S) + i \sin(i/\hbar S)$ via a basic identity known as Euler formula. This means, in particular, that it cannot be interpreted as a probability (a probability must be equal or greater than zero). As we discussed in previous lectures, the advantage of having a probability is that it can be sampled via trajectories (as in the path integral expression of quantum thermal averages). Because real time path integrals do not have a positive weight we can't do that.

When we talked about the fluctuations around the classical path in the first part of the class. It was mentioned that we need to approximate up to the quadratic term. However, I did not understand what exactly we approximate up to that term

In a stationary phase approximation, we consider the function at the exponent (phase). We perform a Taylor series expansion of this function around stationary points (so where the first derivative is null). The expansion then leaves us with the function at the stationary point and terms of second and higher order of the expansion. Typically, the expansion is truncated at second order, and it is this second-order term that is associated to the fluctuations. For the path integral, this means that we consider the second derivatives of the action and multiply by the square of the difference between any given path and the classical path (these are the quadratic fluctuations).

In ad-hoc methods such as PIMD, is there a recommended iterative strategy to choose numerically converged parameters (number of beads, time step, thermostat parameters, total simulation time)? For example, should we first fix P , the thermostat and Δt and increase the trajectory length until statistical errors are under control, then, with that trajectory length, refine Δt and thermostat parameters, and only at the end test convergence with respect to P , or would you suggest a different sequence of convergence tests in practice?

For each P : time-step always comes first - this one typically tests with NVE runs (so for PIMD you don't add a thermostat in the evolution equations), then thermostat parameters, then length of trajectory that gives converged averages. Then change P and repeat (usually the first two points are ok unless the number of beads changes dramatically) until converged in P . This is a pretty general rule (so, not just for PIMD and not just for ad-hoc dynamics). When looking at time-dependent quantum averages via RPMD or CMD, these tests will ensure that the run is correct - in the sense that there are no errors or statistical problems - but they will not say much about the accuracy of the results obtained with respect to the "true" quantum time-dependent averages.

I understood that the Nosé-Hoover and Langevin dynamics were introduced such that the trajectories that they sample follow the probability distribution of the classical NVT ensemble. When introducing the ad-hoc methods during this lecture, it was stated that we can use the trajectories given by Nosé-Hoover and Langevin dynamics to approximate the evolution of a quantum system, but only when we are sure that Nosé-Hoover and Langevin trajectories correspond to the true trajectories that a classical system would have. I don't understand how can we make sure that the Nosé-Hoover and Langevin trajectories correspond to the true trajectories of a classical system, since the goal of introducing them was not that they correspond to the true trajectory, but that they sample the good probability distribution?

Your understanding of Nose-Hoover and Langevin dynamics is correct. What I said is that when the thermostat acts as a small perturbation with respect to the NVE dynamics of the system, we can use these sampling dynamics also to provide time-dependent averages. The relevance of the perturbation can be assessed, for example for Langevin dynamics, by computing the averages for different values of diminishing friction. At convergence, these can be compared with NVE results and will not be very different. The good sampling properties remain because the dynamics is still ergodic with respect to the target probability. Let me know if you wish to discuss this further.

We established that the dynamics we run on the ring polymer (using the 'fake' momenta) are just a tool to sample the probability distribution. Does this mean the time-dependent

properties of this specific trajectory have no physical meaning, and we can only trust the static averages that we calculate from it?

It means that the static properties can be fully trusted (once all convergence parameters have been established, e.g. number of beads in the ring polymer), while, for non-harmonic systems, the time-dependent ones have to be considered with care. The reliability increases with “temperature” and “mass” of the particles - i.e. when the system is almost classical - and has been shown empirically for several properties. Furthermore, there is a growing literature addressing the known shortcomings of these methods that can also assist.

You mentioned the tube of quadratic fluctuations around the classical path. Is the width of this tube a fixed physical quantity (or a hyperparameter) or does the width change dynamically along the trajectory?

The width, measured in the full semiclassical expression by the square root of a determinant of matrices that measure how the classical trajectories change when small changes in initial conditions are made (these are called monodromy matrices), is a dynamical quantity that changes along the trajectory.

When is it more useful to use the ad-hoc approximate trajectory over the stationary phase approach?

In general, semiclassical methods (i.e. methods based on the stationary phase approximation) are computationally more expensive than the alternatives. So, typically ad-hoc approaches are used for relatively large systems, e.g. hundreds of degrees of freedom or more. Also, there are some subtleties in the nature of the sampling coming from stationary phase that I'll mention in the next class and that sometimes produce pathologies that - paradoxically - methods based on ring polymer dynamics do not have.

Is there a way to estimate the error introduced by the approximation in the ad-hoc method?

There are specific caveats for different observables (e.g. vibrational spectra) but not a fully general and “clean”/systematic correction scheme...as far as I am aware. Research is ongoing here.