Q&A Lecture 9

We discussed time-dependent averages, such as the diffusion coefficient, for which we cannot use the ergodic hypothesis. (Maybe trajectory-dependent is the correct term?) For time-dependent averages, we can't "invent" dynamics anymore, can we? And what are some other examples of time-dependent properties?

We can - and we are - using the ergodic hypothesis. In fact, the "invented" dynamical system generates a trajectory - so a time sequence for the averages - and we are guaranteed that this trajectory samples the correct probability via that hypothesis.
No, we cannot. We can, however, systematically investigate the effect of the presence of the thermostat on the time-dependent averages and assess if it is behaving as a small enough perturbation. E.g. we can change the value of Q and see if, for small enough values, time-averages converge to a stable result.
It's a long list…some examples are vibrational spectra, conductivities (electrical and thermal), currents…I'll say a bit more about this in the next class.

From what I gather, there are simpler schemes for simulating NVT ensembles that just keep all the momenta constant, and therefore the temperature, too. Their main drawback is that they don't sample the correct distribution as they don't allow for proper fluctuations. How is this not a problem for NosÅLe-Hoover dynamics, which also modifies the momenta to keep their fluctuations low?

Methods that keep momenta constant (? not sure I know what you refer to) will, by construction, have no fluctuations. Consequently, they can keep the average temperature constant (by construction, given that no momenta change) but will miss things like the variance of the momenta intrinsic in their "Gaussian" distribution that appears in the Boltzmann probability. Nose-Hoover allows for fluctuation so it has a mechanism to estimate also the fluctuations. In fact it can be proven that the dynamical system that we have introduced has a well-defined probability in the extended space. When restricted to the space (r.p) this probability becomes exactly the canonical. I may say a few more words on this in the next class.

The Nose-Hoover thermostat is much more elegant than what I have previously learned in MD courses. Is it also numerically more accurate than velocity rescaling, Langevin or the Berendsen thermostat?

Glad you liked it ;-) It is as accurate as Langevin - I'll say something about this in the next class. Velocity rescaling and Berendsen are in fact not good thermostat in that they only get the average temperature right but do not really sample the correct Boltzmann distribution for the NVT ensemble. In particular, the fluctuations of the temperature are incorrect.

Where T is the target temperature. My question is, what is that c parameter there? What does it do? Why we need it there?

The microscopic estimator for the temperature is derived from the equipartition theorem that says - as far as we are interested here - that at thermal equilibrium any quadratic degree of freedom in the Hamiltonian has an average equal to 1/2 k_bT (where k_b is the Boltzmann constant and T the target temperature). Hence, each Cartesian component of the momenta of each particle contributes this value to the average and the average of the total kinetic energy (that involves a sum over components and particles of the square of the momentum) is equal to 3/2 N k_b T. Setting c=3 N k_b T ensures that the total microscopic kinetic energy (that we have indicated for notational simplicity without the sum) oscillates around the correct value.

Why do we calculate diffusion? for what?

Diffusion coefficients measure how quickly a species moves in the medium it is embedded in. This information is extremely important in different domains, ranging from materials science (where it gives for example the permeation properties of a container, or something like colloidal motion in suspensions) to pharma (diffusion of a drug across a membrane) or chemistry (motion towards reaction centers)…and more.

If I understood correctly, the fact that the Nosé-Hoover dynamics add a friction term to the derivative of the momentum, the trajectories described by it are no longer 'natural' for a system of molecules, and therefore time-dependent properties obtained from Nosé-Hoover dynamics are no longer reliable. Are there other dynamics that make time-dependent properties reliable from simulations of the NVT ensemble?

Correct. No, in general, sampling dynamics introduce biases in the time-dependent properties. The practice is to consider the behaviour of the properties with respect to changes in the quantity that determines the strength of the coupling between the physical and "fake' degrees of freedom (for Nose Hoover the Q parameter) and find regimes in which they become "small perturbations" to NVE references. Alternatively, one can use the extended system to equilibrate the system to the desired temperature and then switch to NVE dynamics. It is known that in the thermodynamic limit (so N and V going to infinity, keeping the density constant) fluctuations of thermodynamic variables (e.g. T in the NVE ensemble) tend to zero as the inverse of the square root of the number of particles. This means that, for large enough systems, all ensembles are in fact equivalent...

In the class we introduced both the microcanonical and canonical probability distributions, but we didn't explicitly use them to derive any equations. Could you clarify why it was important to define both distributions in this context?

Not sure where the question comes from…here is an attempt to reply. Defining the two different distributions is important to enable to simulate experiments in different conditions (isolate systems or systems in thermal equilibrium), We did not use the

explicit form of the probabilities to derive equations, but we did use the microcanonical form to understand the structure of the probability — counting states compatible with the thermodynamical condition (i.e. with the microscopic estimator of the energy H(r,p) equal to the thermodynamic target value E). After defining the canonical probability, we introduced the concept that this cannot be sampled in MD via an algorithm that leads to constant energy and provide an example of an extended dynamical system suitable for this sampling. So, both probabilities played an important conceptual role in our progress...

What are the algorithms used for sampling trajectories in NVT systems, if not velocity verlet? We discussed it for a bit but didnt touch on specific algorithms, so was curious to look into them.

There are several different options. You can have a look at the chapter on Nose Hoover chains (setting the chain length to one) on the book by Tuckerman. An overview of different algorithms - not just nose hoover - for MD is given in the attached paper. Happy to discuss this further.

Did we define the temperature the same way as last week class ? By this I mean, is it the same reasoning as we did to obtain T = sum π_s p^2_s = ∫dR_o dP π(R,P) p^2

Not sure I understand. The microscopic temperature estimator that we used in class this time is obtained from the equipartition theorem. In what context did you see the T definition that you propose? If you \pi(r,p) there is the canonical probability density then this is a way to obtain the estimator…but I don't know/remember how it emerged in the previous class. On a very cautious note: careful about not mixing this concept with the notion of quantum kinetic energy operator (sometimes also indicated with T) in dft. Of course, in that context, the integral you write does not mean much…in particular for the integral over P…

CPMD is based on this same idea of "inventing" a new trajectory with the same nuclear phase space even though the time sequence dynamics might be different. Would I then be correct in saying that CPMD should not be used to recover observables like diffusion and that only MaZe or BOMD should be used for such observables? (also here by BOMD I mean velocity-verlet + exact minimization)

When using extended ("invented") dynamical systems it is important to assess the amount of perturbation introduced by the ad hoc aspects of the dynamics on the statistical properties of the physical variables and - in particular - on time-dependent properties. In the case of CPMD, it can be shown - mathematically - that the perturbation in the dynamics goes to zero with the mass of the auxiliary variables. Therefore one is relatively safe if a convergence test of observables like the diffusion is performed by looking at values obtained with smaller and smaller mass for the auxiliary variables.

- All the above equations of motion become very symmetric by just renaming variables such that,

$$\dot{\eta} = P_\eta / Q, \tag{11}$$

which was done in class by Prof. Bonella. This gives the notion that η can be thought of as position and P_η as its conjugate momentum. My question is: Can this enable us to write an extended space Hamiltonian, $H(r, p, \eta, P_\eta)$, maybe given by,

$$H(r, p, \eta, P_\eta) = \frac{p^2}{2me^{2\eta}} + V(r) + \frac{P_\eta^2}{2Q} + cT\eta, \tag{12}$$

where the masses are scaled by $e^{2\eta}$ and Q is the thermostat mass? This lets us write the equations of motion as,

$$\dot{r} = \frac{\partial H}{\partial p} = \frac{p}{me^{2\eta}}, \quad \dot{p} = -\frac{\partial H}{\partial r} = -\frac{dV(r)}{dr}, \tag{13}$$

$$\dot{\eta} = \frac{\partial H}{\partial P_\eta} = P_\eta / Q, \quad \dot{P}_\eta = -\frac{\partial H}{\partial \eta} = \frac{p^2}{me^{2\eta}} - cT. \tag{14}$$

But this doesn't give the frictional force as in Eqn. (9). If I add a term like $\dot{\eta}pr$ to make the term pop up, then I lose the Hamiltonian structure. What am I doing wrong? Is my extended phase space Hamiltonian wrong? Or can this not be thought of as an extended phase space?

Very nice try! The dynamical system is non-Hamiltonian…it admits a constant of the motion, but the evolution equations cannot be derived in the usual way.

I do not understand well why Nosé-Hoover fails when we want to study time-dependant averages. Is it because when inventing a new trajectory we "loose" the notion of time?

The Nose-Hoover dynamics is guaranteed to sample the canonical (Boltzmann) probability density for the physical degrees of freedom. This means that it will pass in given regions of phase space with a frequency that is proportional to the Boltzmann probability times the volume of that region of phase space. The Nose-Hoover trajectory however, does not visit these regions in the time-order that would correspond to the "natural" dynamics of the system. Consequently, anything that does require knowledge of the exact time sequence of the visits in space (like when we ask what is the distance travelled in a given amount of time) is not automatically correct.

Why is Nose Hoover dynamics not good for identifying the time difference between positions, but good for positions?

I am not sure I'd say that Nose Hoover is "good for positions", I would say that it gives averages of properties that depend on position but not on time (including the average position) consistently with the correct thermal probability for the system. The fact that properties that depend explicitly on time, and in particular on time

differences, are not necessarily accurate can be explained as follows. The Nose-Hoover dynamics is guaranteed to sample the canonical (Boltzmann) probability density for the physical degrees of freedom. This means that it will pass in given regions of phase space with a frequency that is proportional to the Boltzmann probability times the volume of that region of phase space. The Nose-Hoover trajectory however, does not visit these regions in the time-order that would correspond to the “natural” dynamics of the system. Consequently, anything that does require knowledge of the exact time sequence of the visits in space (like when we ask what is the distance travelled in a given amount of time) is not automatically correct.

Which system features determine whether a deterministic thermostat (Nosé–Hoover) or a stochastic one (Langevin/Andersen) is more appropriate?

Not all thermostats are equal…more specifically, some sample the correct probability density (ensuring, for example, that the correct average temperature is fixed by the target temperature and that fluctuations of the microscopic estimator of the temperature are correctly accounted for) others don’t (so they only keep the temperature close to the target one). Nose Hoover and Langevin fall in the first category and can be chosen interchangeably (I would not say that the choice depends on the system but mostly on wether it is more convenient to deal with a deterministic dynamics - for example when one is interested in “simple” time reversal properties - or a stochastic one - for example when a quick thermalisation or the possibility to define convenient likelihoods of the trajectory e.g. for free energy calculations are important), Andersen falls in the second and - in my opinion - should NOT be used.

In the Nosé-Hoover thermostat, how should one choose the parameter Q in practice, is it chosen empirically by testing different values? More specifically, how does the value of Q influence the time scale and amplitude of temperature fluctuations, and what are the consequences for the quality (I presume that ergodicity and decorrelation might be affected) of sampling the NVT ensemble ?

Q is indeed chosen empirically. I am not aware of exact results on ergodicity and decor relation, but the following observations typically hold. The smaller Q, the faster the system equilibrates to the target temperature but the larger the fluctuations around this target temperature. A reasonable value of Q is one that enables to relax to the target temperature in a few thousand time-steps and keeps fluctuations on the order of 5-10% of the target temperature. Of course, convergence tests of averages also as a function of Q should be done before production runs.

In real experiments, systems can lose or gain heat, leading to natural temperature fluctuations. So why do we deliberately constrain molecular dynamics simulations with a thermostat to keep the temperature constant, instead of allowing these fluctuations to occur naturally?

As we discussed in class, the microscopic temperature does fluctuate even in the presence of the thermostat. This is why the feedback mechanism to “add” or “reduce” friction on the momenta is important. It can be shown that the Nose Hoover dynamics preserves fluctuations of the temperature in line with the canonical distribution. Not sure, however, with what you mean with “occur naturally”. First of all, it is important to realise that here we discuss a system in equilibrium with its environment - so no heat sources or drains are present. If you mean that we could simulate the environment in addition to the system…we could but that would be very inefficient since we’d have to simulate the reservoir (as discussed in class) that for us has no other relevance than to keep the system in thermal equilibrium. So, integration of a very large number of evolution equations to focus on what only a few particles do.

– A follow-up question would be: if real systems exchange heat randomly with their surroundings, how was the deterministic thermostat mechanism used in simulations derived? Shouldn’t it instead involve a stochastic (Brownian) term with zero mean to reflect the random nature of heat exchange?

In modern simulations there are both deterministic and stochastic thermostats and it’s nice to have both kinds available so one can choose on the basis of the specific application. Nose Hoover was derived around the time of Car Parrinello MD when (deterministic) extended systems where - as they still are - popular. The ideas to construct the system are very much what was discussed in class. The additional ingredient is the proof that the dynamics does indeed sample the canonical ensemble for the physical degrees of freedom. You will find it in Tuckerman, I’ll say a bit more about this in class.

Given that the Nose-Hoover thermostat is not the only thermostat that exists, what are the main reasons to choose the Nose-Hoover thermostat instead of others such as the Andersen thermostat for example ?

Not all thermostats are equal…more specifically, some sample the correct probability density (ensuring, for example, that the correct average temperature is fixed by the target temperature and that fluctuations of the microscopic estimator of the temperature are correctly accounted for) others don’t (so they only keep the temperature close to the target one). Nose Hoover and Langevin fall in the first category and can be chosen interchangeably (I would not say that the choice depends on the system but mostly on wether it is more convenient to deal with a deterministic dynamics - for example when one is interested in “simple” time reversal properties - or a stochastic one - for example when a quick thermalisation or the possibility to define convenient likelihoods of the trajectory e.g. for free energy calculations are important), Andersen falls in the second and - in my opinion - should NOT be used.

The Nose-Hoover dynamics adds an extra term in the Hamiltonian equations of motion, which may corrupt the realistic motion of the system and introduce an error when computing the dynamic properties. So, which one (Hamiltonian or NosÅLe-Hoover) is the 'Real' dynamics of our system in NVT?

I'll take a drastic point of view and say (this would have to be coloured) that the full NVE description of the system + reservoir is the "real" dynamics.

In actual simulations, how can we ensure that NVT conservation is guaranteed while avoiding simulational errors?

Nose Hoover admits a conserved quantity of the motion (I'll say a bit more about this in class) that can be monitored just like you'd monitor total energy in NVE simulations. Monitoring the (microscopic estimator of the) temperature and ensuring that it fluctuates around the target temperature is also a valid test.

In Nose-Hoover, is there a system in which the temperature oscillations could get smaller initially (damping) and then larger at later time steps?

Not that I can think of. Do you have something in mind?

We concluded that this method is time independent, and for phenomenon like diffusion we need time dependence which means you can only use verlet with very large Q

The method is time-dependent, in the sense that gives a trajectory parametrised by "time", but it cannot be used without caution to compute time-dependent averages.

I'd imagine it would be easier experimentally to enforce a constant pressure rather than volume and the application are also wider – I think. So I was wondering what major changes that would need to solve such a problem?

Excellent point. There is an extended dynamical system also for NPT dynamics in which the volume (or related quantities) becomes a dynamical variable that fluctuates so as to keep the average of the pressure at the target value. This is called the Parrinello-Rahman dynamics.

If we run Nose-Hoover with $Q \to \infty$. The thermostat freezes. What ensemble do I get? NVE or something weird in between? And if $Q \to 0$, does the system explode?

NVE; yes. this can be controlled by reducing the time step, but at some point one reaches numerical precision limits

I don't understand if the thermostat parameter should be defined by optimization or if there are some typical value depending on if we want an NVE or NVT like system.

The NVE would only be obtained with infinite Q, all other values lead to NVT. As a rule of thumb, Q is chosen - by trial - to ensure that the microscopic average of the temperature goes to the target value in a few thousand steps and that fluctuations around the value are of the order 5-10%. However. the choice of Q depends on the specific calculation one is performing. For example, if the plan is to estimate time-dependent averages (e.g. diffusion) using Nose Hoover dynamics, then Q must be large enough to make the thermostat a small perturbation with respect to NVE dynamics for the physical degrees of freedom. Also this is checked via convergence tests (and benchmarks agains an NVE run).

In practice, when dealing with a molecular system, is one algorithm generally preferred over the other? Are there situations where it is better to use the Ritz–Rayleigh approach instead of Nosé–Hoover, or vice versa?

Careful: I have the impression that there is some confusion on the different physics and algorithms that we have discussed.
Ritz-Rayleigh and associated dynamics pertain to minimization of the electronic part of the energy functional when high-level methods are used to compute interactions. This leads to Born-Oppenheimer, Car-Parrinello or Maze extended systems (where the additional degrees of freedom are associated to the representation of the electrons) but it does not “care” if we then discuss the system as isolated (i.e. integrating with velocity Verlet the physical degrees of freedom) or in thermal equilibrium (i.e. integrating with Nose Hoover the physical degrees of freedom). In fact, you can have all combinations, e.g. Nose Hoover plus Car-Parrinello.
Nose Hoover is used both with Ritz-Rayleigh, empirical and machine learned potentials to evolve the physical degrees of freedom in the presence of a thermostat that keeps the microscopic estimator of the temperature around the target temperature. Here the additional degrees of freedom are associated with the representation of the reservoir (i.e. the large system with which energy is exchanged to maintain temperature).

Is it ever practical to equilibrate under a thermostat and then run production sampling NVE to model dynamical phenomena such as diffusion?

Practical, it depends. Correct, yes. It’s true that thermostatted dynamics can be used also to compute time-dependent averages but great care must be taken to ensure that the thermostat is not biasing the result. This can be done via convergence tests towards benchmarks obtained via a (fast) NVE run.

Do thermostats improve ergodicity by promoting energy exchange between different subsystems?

Not necessarily and the speed at which different thermostats converge to the asymptotic probability density depends on their nature and on the system. E.g. Langevin is “faster” than Nose Hoover. Also, the Nose Hoover that I discussed in class is not always ergodic (the dynamics admits the total momentum as a conserved quantity even when this should not be the case and this leads to spurious confinement on the constant momentum hypersurface). This can be corrected with something called Nose Hoover chains — see Tuckerman.

Is the invented dynamics for the use of Nosé-Hoover thermostats related to the generalized Langevin equation? Or not necessarily?

Only in the sense that both dynamics are ergodic with respect to the canonical probability. I’ll discuss this a bit more in the next class.

If Q is chosen extremely large, does the system still manage to sample the NVT probability, or does it remain effectively stuck in NVE like behaviour?

The larger is Q, the longer it takes to sample NVT. So…if Q is too big, the system remains effectively stuck somewhere - not necessarily NVE because there will always be some perturbation to the dynamics that affects long time scale phenomena.