Q&A Lecture 8

Could you clarify the physical meaning of working with ρ(G) instead of ρ(r) in the minimization of E[ρ]?

Reciprocal space is often the preferred environment for discussing systems in DFT-related calculations. This is because (a) solids, in particular crystals, are periodic structures well suited for a description in Fourier space (b) other types of systems still interact predominantly via the Coulomb term that has a very simple (1/G^2) representation in Fourier space.

Since SHAKE and RATTLE remove high–frequency bond vibrations to allow larger MD timesteps, in what situations is it appropriate to use these constraints, and when should a fully unconstrained many–body simulation be preferred so that vibrational degrees of freedom are physically retained?

The computational model adopted to simulate any system depends on the system (obviously) and on the specific properties that one wishes to study. If rototraslations are the focus, constraining vibrations is fine (provided that any biasing of the statistical ensemble due to the constraints is accounted for); if one is interested in vibrational properties (e.g. IR or Raman spectra) then vibrations should be explicitly accounted for.

In Car-Parrinello dynamics, the approximation works best when the rate of change of the electronic degrees of freedom is small. Is this assumption generally valid in practice? Although the fictitious electronic mass (μ) is chosen to be very small, wouldn't a rapid acceleration of the electronic density make the product $\mu \cdot (d^2\rho/dt^2)$ non-negligible, thus challenging the validity of the approximation?

Why do you say that a small rate of change for the electronic dofs is necessary? What we need is that they move (also via rapid oscillations) around and close to the value that minimises the energy functional.

My question is about the idea of "inventing" dynamics whose trajectories have the same ensemble probability - in Car-Parrinello and MaZe, we saw that these trajectories are pretty close to (or equal) to the original trajectories. Are there any other formulations of the dynamics where the trajectories are wildly different but the probability in the ensemble average remains the same?

Yes, a good examples are different dynamics used to sample the canonical ensemble (so constant number of particles, volume, and temperature — until now we discussed constant energy, number of particles and volume). In that case, we can use trajectories as different as stochastic dynamics and extended systems. I'll say more about this in the next class.

MaZe seems like a very efficient method in terms of the computational cost. It also seems like we are, in a way, doing "classical mechanics" with the electrons as well and just adding the minimum energy, like the zero-point energy of the system for the dynamics. I may also be completely wrong, so, what are some key differences between MaZe and full classical treatment of the system? Also, it seems like MaZe has the same computational cost as classical dynamics.

The electronic energy and density are computed quantum mechanically via a minimisation mechanism that employs classical molecular MD methods. However, both quantities are quantum and, in particular, the electrons are not described as classical particles - e.g. they are delocalised and described via the density which is a quantum probability concept.

How can we intuitively understand that in the limit of zero mass, the Car-Parinello dynamics gives the Born-Oppenheimer dynamics?

Zero inertia means finding the minimum infinitely fast?

Also, do we encounter non-holonomic constraints in physical systems of interest, and in this case how does the SHAKE algorithm change?

Yes, e.g. if we want to constrain the motion of particles in a magnetic field. For non-holonomic constraints that depend linearly on the velocity, there is an adapted algorithm that works in a very similar way to SHAKE.

Are there limitations to the types of densities or potentials for which the Fourier-transformed conjugate gradient method is effective?

Not really though in practice, depending on the specific form of different terms in the energy functional - and in particular on the choice for the exchange correlation term - it is often convenient to switch between Fourier and real space representation for the density. This does not change much conceptually and is done on the fly.

Are there systematic ways to estimate the error introduced by finite grid spacing?

Yes, the usual convergence type of tests. A particularly useful quantity here is the convergence of the total energy of the system - evaluated for a small set of ionic positions - with the number of (Fourier) grid points.

Due to the same probability density, would the expectation value of any observable of the system be the same if one performed MaZe or if one performed BO-MD?

Yes, they are the same.

Are the expectation values given by MaZe closer to the experimental results compared to the ones from Car-Parrinello trajectories?

it depends on the observables, as Car-Parrinello remains a good approximant of the exact dynamics. however, it is known that there is a bias (that goes to zero for zero mass of the auxiliary variables) in the kinetic energy of the ions. this can manifest itself, for example, in shifts in the ionic vibrational spectra when compared to experiments or BO (or MaZe) ones.

If MaZe is capable of sampling from the probability density of BO-MD and Car-Parrinello is not, why would one use Car-Parrinello in practice instead of MaZe?

MaZe is a relatively young method so adoption is lower than with alternatives. Also, as mentioned above the error in Car-Parrinello sampling may not be that big. Note that, whenever not too costly numerically, the most commonly used method today is BO-MD, not Car-Parrinello.

Why pointed out $\partial E[\{\rho\sim\_G, \rho\sim\_G^*\}; R] / \partial\rho\sim\_G = 0$ *is non-linear conjugate gradient during the lecture? does this imply something important?*

The fact that we must solve a non-linear system of equations (and therefore use non-linear conjugate gradient) augments the numerical complexity of the problem. In particular, it can be proven that for linear systems conjugate gradient solves the problem with the smallest possible number of steps. The same is not true for non-linear.

To compute the ensemble average of a variable, we need to know the density of the trajectory (or the function $\pi(R, p)$*) in the phase space. How can we know this without computing the trajectory of an actual system and counting the intersections of the trajectory and the cells?*

These probabilities are actually known analytically. The one corresponding to the constant energy (number of particles and volume) is a delta(H(r,p)-E), the one corresponding to the constant temperature (number of particles and volume) is the Boltzmann distribution, etc. The formal derivations of these densities is a cornerstone of statistical mechanics and you can find it also in the book by Tuckerman.

In the MaZe procedure you described, the constraint force keeps ideally the electronic coefficients on the exact BO minimum manifold at each step. But in practice, how sensitive is the ionic sampling to small numerical drift away from that manifold (due to imperfect constraint solution), and do we diagnose good sampling and convergence by checking that the total energy stays constant as we've done previously?

One of the advantages of using the constraints is that the error in the minimisation is not propagated from one time step to the other. That means that we cannot drift away from the constraints manifold and we'll stay as close to it as the convergence threshold for the numerical satisfaction of the constraint imposes. Energy conservation is a good check (if we propagate in the NVE ensemble as discussed in class). Convergence and stationarity of the averages is another.

You divided the space in n cells S, so you could approximate every point in those as having the same probability. I was wondering if it would be possible to remove all the cells where there is nothing (or where the path is not going), so when summing on all the cells, it would reduce the computational cost. How would you be able to predict which cells stay "empty"?

In practice, if you sample via the trajectory - so averages over time - this is automatically done because the trajectory does not go in that specific cell but for the integral, which is never solved by brute force discretisation - it would be complicated to see a priori. Also, keep in mind that the finite size cell is a device to construct the process but that we take the limit for the volume of the cell to go to zero to discuss ensembles. Zero - or infinitesimal - measure volumes are non-trivial to remove.

Are there some systems of practical interest where the ergodic hypothesis is not satisfied?

Not in “standard” states (e.g. solid, liquid, gas) typically simulated. However, there are systems that are locked in areas of the phase space for very long-times or whose dynamics is so slow that in practice it is not possible to verify ergodicity (e.g. glasses).

We should only have one macroscopic value for temperature/position but a molecule going from point A to B can take any path. So are the values still the same even if I run a simulation twice with 2 completely different paths?

What do you call a macroscopic position? For temperature - that has a clear macroscopic meaning - the answer to your question is that if the two paths are sufficiently long to converge the microscopic average then yes, the value of the temperature will be the same. Also, the notion of going to point A to B is not super relevant in this context…it would be if we were talking about, e.g. a chemical reaction in which point A is the reactant in phase space and point B the product. Is this what you had in mind?

I don't understand the meaning of ̃ρ_G ?

The density is a continuous function of space and, as such, not representable on the computer. We need to introduce a discrete representation for the function. The specific choice is in our hands: it can be a three d grid in real space or, as we did in class, a discrete Fourier representation. For the latter, the \tilde{\rho}_G are the expansion coefficients in the Fourier representation.

Why is it useful?

Going to Fourier space is convenient because Coulomb interactions (classical) are very easy to represent and are given simply by 1/G^2.

Also, in the Taylor expansion of ̃ρ_G(t+δ) (in the context of MAZE), did we add the correction term Fc or is it part of the expansion? Because in my notes I wrote ˜ρ_G(t+δ) = ˜ρ_G(t) + δ ̃ρ'_G(t) + δ^2 Fc, if it is part of the Taylor expansion, shouldn't it have a ½ coefficient?

The Taylor series here represents the “coordinate like” update in a velocity Verlet propagation. I don’t remember what I wrote on the board but yes, force terms come with a ½.

Given the equivalence between time and ensemble averages, does constructing a simplified trajectory require knowledge of the true microscopic path?

Not necessarily, it does require knowledge of the true microscopic probability density.

If I understood it right, BOMD is using KS-DFT, CPMD is using a similar approach, whereas MaZe is using orbital free DFT. Is that correct, and how different are KS-DFT and OF-DFT?

All these methods can be used in combination with KS-DFT or OF-DFT. It’s just a matter of what you identify as the auxiliary dynamical variables (so coefficients of the density in a basis or coefficient of the orbitals in a basis).

Is OF-DFT similar to extended Lagrangian (XL-BOMD)?

See above…

How different dynamical schemes (8-0, C-P) can lead to same statistical ensemble, even though they have different equations of motion?

the notion that different “lotteries” can produce results with the same probabilities is not unusual, is it? Think about tossing a coin 1000 times or having a 1000 coins is a bag, shaking the bag and then extracting the coins and looking at the results. do you expect a significantly different number of heads and tails to result from the two experiments? More rigorously, the probability sampled by a dynamical system can - in the cases that we are considering - derived mathematically by solving something called the Fokker-Plank equation so one has the information exactly. In molecular dynamics, there are also other analytical methods to know the ensemble.

How do the SHAKE algorithm and Car-Parrinello dynamics preserve the phase space distribution?

The phase space distribution sampled by SHAKE is known analytically and - while it in general is biased for the constrained degrees of freedom - the unbiasing procedures are well known and not difficult to implement (there is an additional volume element in the phase space measure that can be computed along the trajectory). The important thing for the application we considered in the lecture, however, is that - thanks to the zero mass limit - there is no bias on the ionic degrees of freedom so here we don’t have to correct anything.

The phase space distribution sampled by Car-Parrinello dynamics is also known analytically and it can be shown that there is a bias on the ionic momenta. (by the way, as far as I know, the exact form of this bias was not known until worked out via MaZe) In this case, however, due to the form of the bias it is more difficult to correct for it. That said, if the mass of the auxiliary variable is sufficiently small - something that can be tested via convergence analysis of the observables - this bias can be safely “neglected”. The problem here is that the smaller the mass of the auxiliary variables, the smaller the time step that must be used and therefore the higher the overall cost of the calculation.

If we were to compare the 'invented' trajectories to experimental measurements of the microscopic system would we be able to see great discrepancies?

That depends on what you consider the “ground truth” for the experimental measurement. If we go with Born-Oppenheimer, then MaZe would give the true trajectory for the ionic degrees of freedom. The trajectory of the coefficients of the density, on the other hand, has no experimental counterpart. CPMD would give a good set of coordinates, but there would be a bias in the momenta of the ions. In general, the “similarity” of the invented trajectories with respect to the chosen reference depends on the form of the dynamical system. We’ll see more about this in the next class.

Since the ensemble average we described does not care about the order in which the cells/volumes are entered, but only the frequency at which, couldn't this describe a completely different physical system than what is being observed?

Here we are in danger of bordering on philosophy - not a bad thing, but one that can lead to very long discussions. I would say that two systems are the same as long as we cannot distinguish them via a measurement. So, for macroscopic observations, the answer is no. As for the microscopic observations, there are controls - see the statements above about MaZe but also CPMD (since we know what the bias is we can remove it) - that give us an idea of how “different’ the systems are.

So these dynamics are only useful for considering macroscopic averages?

This is not a small thing...

I see μ has units of mass × time2, so it isn’t exactly mass. I remember you used the term “inertia” but the units still don’t match. I can interpret it as “mass” of electrons but still not quite convinced.

In extended dynamical systems, the term inertia is used to indicate the constant associated with the kinetic energy of the auxiliary variables introduced. The units are determined by this prescription. The name is used due to the physical meaning of the parameter, not necessarily because it has the units of the “standard” inertia.

I am having trouble understanding the difference between CPMD and MAZE conceptually. If its just the difference in lagrange multipliers, how come Roberto Car and Michele Parrinello missed it?

The two methods differ in the use of the inertia - finite in CPMD, zero in MaZe - and in the fact that the dynamical system not only incorporates constraints but, due to the fact that the minimum is imposed as a constraint, has a different structure. I am not sure I understand what you mean by “just the difference in Lagrance multipliers”. This is a substantial change in how the Born-Oppenheimer approximation is enforced exactly via the constraint condition (MaZe) or approximated via a physically fast motion induced by small inertia (CPMD). As for why Roberto and Michele missed this way to proceed…that’s a question for them!