Is there a way to ensure that our trial space contains the ground state, i.e. that our trial wavefunction isn't orthogonal to the ground state?

Symmetries of the Hamiltonian can help to identify symmetries of the ground state. One can also use trial functions with different symmetries and check which has the lowest expectation of the Hamiltonian.

How do people find proper trial wave functions in practice? Are there some research that using the variational principle to calculate excited state energy? If so, what are the main methods they utilize?

There are well established strategies in quantum chemistry, using functions constructed from basis that account, e.g., for known symmetries of the system or from functions that work for "fragments", e.g. linear combination of atomic orbitals to build molecular states. In physics, one mostly uses density functional theory (more on this in future classes). There is indeed research on the excited states. Main methods used are direct diagonalization of the Hamiltonian, e.g. in a quantum chemistry basis, or time-dependent density functional theory. Modelling empirically excited states is also quite common.

Is it possible to have multiple different trial wavefunctions which give the same calculated energy? If so, can these represent different physical states?

If the ground state is not degenerate, I don't think that this is possible.

If the foundation of our simulations and also predictions in quantum chemistry relies on guesses and approximations, and also knowing that we may never truly find the exact ground or excited states, how can we be sure of the correctness of our systems and use them as benchmark for further simulations, and where do we draw the line between useful science and wishful thinking?

Internal consistency and experimental evidence.

Gradient based methods (as shown in class) only find the global minimum if the energy over the search space is convex. Why should we expect convexity over such a high dimensional space with highly non-linear functions?

Gradient based methods have proven reliable over a large class of Hamiltonians and can be complemented with annealing and/or stochasticity to verify the "quality" of the minimum that has been found. Convexity of the functional has been justified/proved, e.g. in the context of density functional theory - more on DFT in future classes - for rather general forms of the Hamiltonian.

I don’t understand the efficiency/motivation behind the 1st way to find the ground state energy. If we know our desired Hamiltonian is “approximately” similar to a harmonic oscillator or even any other model, we know the ground state energy of that model right?Why do we need to differentiate the energy functional with respect to the coefficients to give us that when we already know the answer is c0 and energy E0?

We used the harmonic oscillator as a simple test case to get familiar with the variational method and check the validity of the result that we obtain from it. Since the exact solutions are known, we could easily compare with them.

The steps we followed for the harmonic oscillator will also serve as a roadmap for deriving the Hartree–Fock equations in the next tutorial.

In the simple case of the harmonic oscillator, using a set of trial functions that does not adopt the known "e^(x^2)" form already yielded a result that is 40% off. It almost seems like we can only use this principle if we already know the analytical form of this solution. My question is then: Is this principle actually used for real-world examples? Do you have one in mind?

Essentially all calculations in quantum chemistry and in solid state physics are based on the variational principle. Methods like Hartree-Fock, or density functional theory are variational approaches to the ground state. The accuracy of the example is not representative of what can be achieved with good trial functions, where errors (compared to experimental values of the ground state) can be as small as 1%.

How to choose the family of trial wavefunctions that is used for calculation? Is the estimated ground energy value very sensitive to this choice?

There are well established strategies in quantum chemistry, using functions constructed from basis that account, e.g., for known symmetries of the system or from functions that work for “fragments”, e.g. linear combination of atomic orbitals to build molecular states. In physics, one mostly uses density functional theory (more on this in future classes). The estimated energy is indeed sensitive to the choice of the trial state. This is why several states are typically tried and why - e.g. when a basis expansion is employed - convergence tests - e.g. with the number of basis elements - are performed.

What is the strategy or thought process employed while selecting the family of trial wavefunctions, especially for an arbitrary state? Also, if we get a high error rate with a particular wavefunction, how do we tune our trial wavefunction to reduce the error? Is it by observing some behaviour of the trial function, or by random trial and error?

There are well established strategies in quantum chemistry, using functions constructed from basis that account, e.g., for known symmetries of the system or from functions that work for “fragments”, e.g. linear combination of atomic orbitals to build molecular states. In physics, one mostly uses density functional theory (more on this in future classes). There is indeed research on the excited states. Main methods used are direct diagonalization of the Hamiltonian, e.g. in a quantum chemistry basis, or time-dependent density functional theory. Modelling empirically excited states is also quite common.

When we set the gradient of the energy with respect to a coefficient to zero to minimise the energy, how do we ensure that the optimisation procedure does ont get stuck in a local minima? What would be some strategies that avoid this in practice? I wonder if optimisation techniques used in neural network optimisers are relevant here (stochasticity in gradients, momentum, etc)

Convexity of the functional has been proven for a relatively large set of trial states. Annealing and stochastic approaches can be used to verify the stability of the minimum identified.

BO says that because nuclei are heavy, R changes slowly. So the electronic wavefunction $\phi\alpha(r;R)$ can “adjust adiabatically”. does not this mean that $\partial\phi\alpha(r;R)/\partial R$ is very big? then why on the contrary ,we think the coupling term $D12(R)=\langle\phi 1|\partial/\partial R|\phi 2\rangle\approx 0$?

No, it means that the derivative is small (slow changes of a function with respect to a variable) so things are consistent.

Can a neural network be used to define an Ansatz for the wavefunction ? If so, is it used in practice ?

Yes, typically in association with approaches like variational Monte Carlo. But I am not an expert here.

By using the variational principle for a parametrized wavefunction, or for a wavefunction that is expanded into a truncated basis set, it is possible to find as solution multiple parameters or combination of coefficients such that the equations obtained while applying the variational principle are verified. Therefore, multiple energies can be obtained when applying the variational principle. The lowest of the energies obtained is the best approximation for the ground state energy. Are the other energies good approximations for the excited states, for example, is the second lowest energy obtained a good approximation for the first excited state?

No, in fact typically the same “type of” trial state (e.g. same symmetry) cannot describe both the ground and first excited state - that are orthogonal. Also, I am not sure that the scenario in which there are multiple solutions is very common - unless we are trapped in local minima, which in any case is not our best estimate of the ground state energy.

Is it possible to converge towards a wrong expectation value for the energy and a wrong state approximation given an unlucky parametrization of the trial wave functions?

Yes. This is one of the reasons why typically more than one trial state is considered. Stability of the result for sufficiently different trial states is “reassuring"

Is the harmonic oscillator always used to do this method computationally ? Is it possible to know by how far we are from the exact energy ? Does it always involve comparison with laboratory-obtained data? Does the Born-Oppenheimer approx. imply that all the electrons are in the ground state ? If yes, why ?

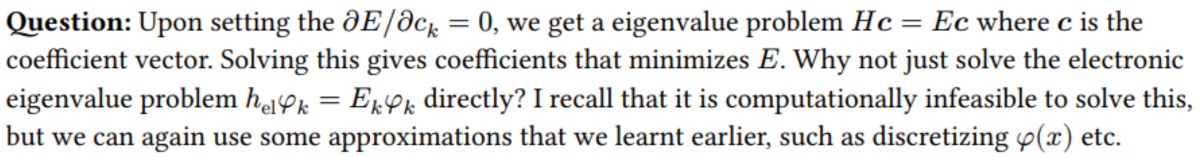
No, it was used as a “simple” example to clarify the concepts and math introduced in class; Experimental data is the ultimate test. Convergence, for example with respect to basis size in the trial state, can also provide some indication; Yes, the states that we are considering are all electron states. Not sure I understand the why question. The argument on the fast degrees of freedom (electrons) relaxing to the ground state is general.

In practice, how do we guess the energy function? Do we generate many random functions, or is there a more systematic approach?

There are well established strategies in quantum chemistry, using functions constructed from basis that account, e.g., for known symmetries of the system or from functions that work for “fragments”, e.g. linear combination of atomic orbitals to build molecular states. In physics, one mostly uses density functional theory (more on this in future classes). There is indeed research on the excited states. Main methods used are direct diagonalization of the Hamiltonian, e.g. in a quantum chemistry basis, or time-dependent density functional theory. Modelling empirically excited states is also quite common.

In practice, what is the limit of BO approximation,are there sone case where the approximation results in significant errors ?

The B-O approximation works quite well for essentially all molecular and solid systems if they are not subject to specific photo-excitations. As discussed in class, there are however experiments, e.g. ultrafast spectroscopic, where it would misrepresent the situation and lead to significant errors.



Question: Upon setting the $\partial E/\partial c_k = 0$, we get a eigenvalue problem $H\boldsymbol{c} = E\boldsymbol{c}$ where $\boldsymbol{c}$ is the coefficient vector. Solving this gives coefficients that minimizes E. Why not just solve the electronic eigenvalue problem $h_{\text{el}}\varphi_k = E_k\varphi_k$ directly? I recall that it is computationally infeasible to solve this, but we can again use some approximations that we learnt earlier, such as discretizing $\varphi(x)$ etc.

From one perspective, it is easier to deal with the matrix-vector product version, since we can use tools from numerical linear algebra, and it is straight forward to transfer the problem to a computer. From another perspective, any time-independent Schrödinger equation can itself be seen as the result of a variational principle, doing the variation over the states instead of the coefficients. So you can introduce the expansion of the wavefunction in a basis either within the variational method, as described in your question, or directly in the time-independent Schrödinger equation, and in both cases, you obtain equivalent eigenvalue problems.Maybe the three first pages of the chapter 3 from the book Computational Physics by Thijssen could help to clarify it. If not, we can further discuss it on Wednesday.

The exact steps and approximations made to go from the coupled-channels equation to the BO approximation are still not clear to me. What are the exact steps and approximations that are performed?

Step 1: neglect coupling terms (the D) in the equations. This is an approximation, very good when there are no photo excitations and/or the thermal energy of the system is lower than typical separations between the electronic energies. As a result of this approximation, the coupled equations separate in uncoupled equations each corresponding to nuclear motion on a single electronic surface.

Step 2: restrict nuclear motion (still quantum) to the lowest energy state of the electrons, the ground state. This is justified on the basis of the different time scales of the motions: typically electrons have time to relax to their lowest energy state in the time that it takes the nuclei to move appreciably.

Step 3: change the evolution of the nuclei from quantum to classical. This is justified on the basis of the “large mass” of the nuclei that quenches quantum effects such as delocalisation.

In general, this may not be the case for some other trial wavefunction, which may require more sanity checks. Hence, is there any definitive form of observable, in relation to our Hamiltonian, whose expectation could be calculated or any other suitable check which could tell us if our trial function of choice was ’good’?

The situation in which errors are different for the state and different observables is very common in variational approaches. Trial states are often targeted to a specific observation - predominantly the energy. The mathematical properties of the square of the state provide a lot of information. E.g., the divergence that you are describing is easily "predictable" when you think about the integrals related to the average of x^4, not sure that you need more in terms of sanity check. In general, comparing results obtained with different states is a good test of the robustness of a trial.

For simple systems the variational principle make sense to me, but I really struggle on how could we find a good trial function for a more complicated system, without being to computationaly expensive. Is there any "methods" ?

The variational principle is the most widely used method to estimate the ground state energy and is applied also to very complex systems. There is a large body of work that develops trial states (in quantum chemistry based on expansions in basis sets that embody key symmetry and physical, e.g. correlation, aspects of the electronic system, in solid state physics and in some chemistry applications based on the density - density functional theory). We shall see a bit more about this in future lectures. There are alternatives, such as direct diagonalization of the Hamiltonian in a basis or variational Monte Carlo, but these are typically too expensive for large system sizes.

Why when we were discussing the main equation to solve, the coupled channels eq, we could neglect all the terms with the D constants?

This is based on the Born-Oppenheimer approximation, usually valid when there are no photo excitations. (See also previous lectures).

The systems we are now looking at are all nicely behaving diatomic systems. However, let's say we look at more exotic examples, e.g. nuclear fission, Born-Oppenheimer would start to fall apart since nuclei are now suddenly no longer part of the same system. Can we recover anything from the algorithmic approach or is this so far out that we have to devise whole new ways of finding solutions?

Although the example considered in class was a diatomic, our methods are valid for more complicated systems and also allow for dissociation as illustrated via the example (so I am not sure what you mean when you say nuclei are no longer part of the same system). Fission would be a very different experiment because you are actually breaking the nucleus and in that case you need more than the type of quantum mechanics we discussed - not just an algorithmic problem.

What other methods could we have used to estimate the electronic potential energy acting on the nuclei? What are the main advantages and disadvantages of these alternative approaches, and why did we ultimately choose the variational method?

Alternatives include - for example - direct diagonalizaiton of the electronic Hamiltonian in a convenient basis set, diffusion (quantum) Monte Carlo. The variational method, in its various incarnations that will be discussed also in future lectures, is the best compromise between accuracy and efficiency in the context of dynamical simulations.

Gradient based methods (as shown in class) only find the global minimum if the energy over the search space is convex. Why should we expect convexity over such a high dimensional space with highly non-linear functions?

Gradient based methods have proven reliable over a large class of Hamiltonians and can be complemented with annealing and/or stochasticity to verify the "quality" of the minimum that has been found. Convexity of the functional has been justified/proved, e.g. in the context of density functional theory - more on DFT in future classes - for rather general forms of the Hamiltonian.