

Lecture 0.1: A Review of Basic Principles

Boyi Zhang

January 8, 2020

1 The principles of quantum mechanics are derived from experimental observation

In this class, we will be learning a lot of abstract math. It is important to remember that all the mathematics of quantum mechanics is motivated by the desire to form a consistent worldview from experimental observations. When you get confused as to the meaning of all the mathematics, it is always possible to follow a path back to a physical grounding. Let's consider a classic thought experiment (Figure 1). Electrons are

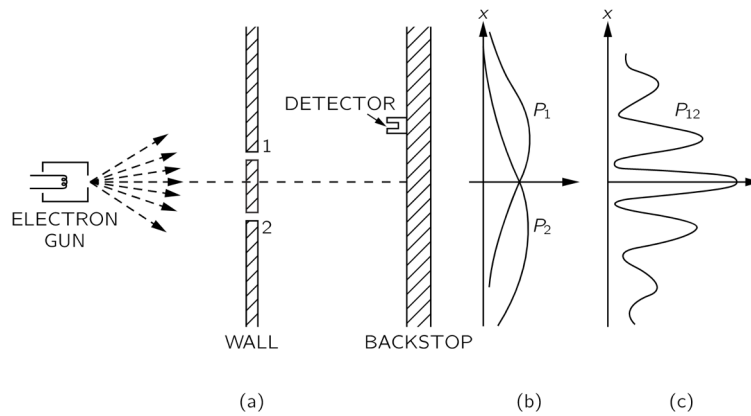


Figure 1: The Double Slit Experiment

shot out of an electron gun and arrive at the first wall. The electrons are only able to move through the two holes on the wall to the detector. The first question we can ask is: in what form do the electrons arrive at the detector? The answer is that the electrons always arrive at the detector in identical lumps, that is, they arrive as particles. The second question we can ask is: through what hole do the electrons come through? If we block off hole 2, the electrons can only come through hole 1, and arrives at the detector with probability distribution P_1 . If we block off hole 1, the electrons can only come through hole 2, and arrives at the detector with probability distribution P_2 . However, if we leave both holes open, the probability distribution observed is not the same as $P_1 + P_2$. It is instead something different, we call it P_{12} . There is interference. It turns out, the mathematics of these observations are simple, and can be described by 2 complex numbers, ϕ_1 and ϕ_2 .

$$P_1 = |\phi_1|^2$$

$$P_2 = |\phi_2|^2$$

$$\phi = \phi_1 + \phi_2$$

$$P_{12} = |\phi_1 + \phi_2|^2$$

We notice that this set of mathematics is the same as obtained for water waves. We can make a observation: electrons arrive in lumps, like particles, but the probability of arrival through both holes is distributed like the intensity of waves.

2 The wavefunction is a concept used to reflect experimental observation

There are three key conclusions we can make from this experiment, which in turn guide the general principles of quantum mechanics:

1. The probability of an event in an ideal experiment is given by the square of the absolute value of a complex number ϕ which is called the probability amplitude.
2. When an event can occur in several alternative ways, the probability amplitude for the even is the sum of the probability amplitudes for each way considered separately. There is interference.

$$P_{12} = |\phi_1 + \phi_2|^2$$

3. If an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost.

$$P = P_1 + P_2$$

Thus, the framework given by quantum mechanics is to give an amplitude for every event that can occur. If the event involves the reception of a particle, we can associate a probability amplitude to finding that particle at different place and times, which generally will vary with position and time. This function of probability amplitude at different positions and times is what we call a “wavefunction”.

3 Postulates of quantum mechanics

This leads us directly to an important postulate in quantum mechanics:

- *The state of the system is described by a wavefunction ψ that depends of the positions of the particles and time. ψ is also called the state function and it turns out to contain all knowable information about the system.*

I want to emphasize that the idea of a wavefunction was motivated from a desire to explain experimental observation. Our primary goal is to find the correct wavefunction for our system of interest, and use it to obtain our desired observable. To do this, we need to review the other fundamental postulates of quantum mechanics:

- *Every physical observable has a corresponding linear operator in quantum mechanics. (A algebraic operator is something that maps a function into another function) These operators can be obtained from classical counterparts in a Cartesian coordinate representation via replacements*

$$\begin{aligned} x &\rightarrow \hat{x} \\ p &\rightarrow -i\hbar \frac{\partial}{\partial x} \end{aligned}$$

- *The possible values that can result from the measurement of a physical observable are the eigenvalues of the corresponding operator:*

$$\hat{B}\chi_i = b_i\chi_i$$

Measurement of a property causes a sudden change of the state function, “collapse of psi” If the measurement yields b_i , then ψ becomes χ_i

- The average value of a physical observable at time t is given by

$$\langle \hat{B} \rangle = \int \Psi^*(\mathbf{x}, t) \hat{B} \Psi(\mathbf{x}, t) d\mathbf{x}$$

where $\Psi(\mathbf{x}, t)$ is the normalized state function of the system. $|\Psi(\mathbf{x}, t)|^2$ acts like a probability density function.

- The state function $\Psi(\mathbf{x}, t)$ of the system obeys the time-dependent Schrödinger equation

$$\hat{H}\Psi(\mathbf{x}, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t)$$

where \hat{H} is the Hamiltonian, or energy operator.

With these postulates, we see that the wavefunction contains everything we can know about a system, and that the way to solve for it is to use the time-dependent Schrödinger equation.

4 The time-independent Schrödinger wave equation

In an introductory quantum mechanics course, you were most likely introduced to the time-independent Schrödinger wave equation as a fundamental equation of quantum mechanics:

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r})$$

which was derived from the time-dependent form with some assumptions and mathematical manipulations. Here, I've switched from writing \mathbf{x} to \mathbf{r} to represent the space coordinates. The Hamiltonian \hat{H} is made up of a kinetic energy term $(-\frac{\hbar^2}{2m} \nabla^2)$ and a potential energy term $(V(\mathbf{r}))$. We can rewrite the equation as:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

In this form, it is clear that the time-independent Schrödinger wave equation is a differential equation whose solution, $\psi(\mathbf{r})$, describes a particle of mass m moving in a potential field described by $V(\mathbf{r})$.

It is through the Schrödinger wave equation that we can obtain solve for our wavefunction, which we have stated is the primary goal of all this. The steps to do this are: 1. Write down $V(\mathbf{r})$ for your system of interest. 2. Use the Schrödinger equation to solve for the wavefunction. Everyone has probably had practice with this (recall finding solutions for particle in a box, harmonic oscillator, and the hydrogen atom).

For systems of chemical interest, we are not working with single particle systems, but multiple particles. Following the steps described above, we first need to find an appropriate Hamiltonian for many-electron systems. Let's consider a diatomic system with 2 nuclei and 2 electrons. There are 4 kinetic energy terms, one for each nucleus and electron. The potential energy between two charges i and j is give as $\frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}$ in atomic units. There are three types of pairwise interactions in our diatomic system: electron–electron, electron–nucleus, and nucleus–nucleus. We can now write our Hamiltonian compactly, in atomic units, as:

$$\hat{H} = -\sum_i^N \frac{1}{2} \nabla_i^2 - \sum_A^M \frac{1}{2M_A} \nabla_A^2 - \sum_i^N \sum_A^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_i^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}$$

The first two terms are the kinetic energy terms for the electrons and nuclei, respectively, while the last three terms are the electron–nucleus, electron–electron, and nucleus–nucleus pairwise potential energy interactions. We can invoke the Born-Oppenheimer approximation, and use the idea that the nucleus moves much slower

than the electrons. We can thus “fix” the nucleus and solve for the movement of the electrons only (“clamped-nuclei” approximation). The kinetic energy term for the nucleus becomes zero, and our electron-nucleus potential energy terms now become a constant which is added onto the final energy. We thus obtain our electronic Hamiltonian:

$$\hat{H}_e = -\sum_i^N \frac{1}{2} \nabla_i^2 - \sum_i^N \sum_A^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_i^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

We can use the electronic Schrödinger equation to solve for the electronic wavefunction:

$$\hat{H}_e \Phi_e = E_e \Phi_e$$

Φ_{elec} describes the motion of electrons, explicitly depend on electronic coordinates but parametrically on nuclear coordinates. (Dr. Schaefer’s class uses the electronic Hamiltonian to derive wavefunctions for simple chemical systems, namely, multi-electron atoms and diatomics) Once we have solved for the electronic wavefunction, we can then solve for the positions of the nucleus in the average field of the electrons. Our nuclear Hamiltonian is given by:

$$\hat{H}_{nucl} = -\sum_A^M \frac{1}{2M_A} \nabla_A^2 + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} + E_e$$

And the nuclear Schrödinger equation is:

$$\hat{H}_{nucl} \Phi_{nucl} = E_{nucl} \Phi_{nucl}$$

The nuclear Schrodinger equation describes the vibration, rotation, and translation of a molecule. (Dr. Allen’s class finds various solutions for the nuclear Schrödinger equation.)

In this class, we will focus on using the electronic Hamiltonian to solve for general chemical systems. The electronic structure methods covered in the class are in their essence different approaches for finding the appropriate wavefunction. We will be learning tools of abstraction to make the math doable, and then applying these tools of abstraction to derive commonly used methods in electronic structure. The goal is to equip you to understand the mathematics used to derive theories in quantum chemistry.

5 A different notation for the electronic Hamiltonian

The Hamiltonian used in this class will always stay the same. It is the electronic Hamiltonian written down in the previous section. Instead of writing the terms explicitly, we choose to write it in terms of two operators: the one-electron operator called \hat{h} and the two-electron operator called \hat{g} , where

$$\begin{aligned} \hat{h}(i) &= -\frac{1}{2} \nabla_i^2 - \sum_A^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \\ \hat{g}(i, j) &= \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ \hat{H}_e &= \sum_i^N \hat{h}(i) + \sum_{i<j}^N \hat{g}(i, j) \end{aligned}$$

It’s easy to see that these operators are named based on the number of electrons involved in their terms. From here on out we will not write the terms explicitly, but you can remember that the symbols we use are just representations of concrete algebraic terms.