The variational method in quantum mechanics

Bento Montenegro, PhD student • Gleb Wataghin Institute of Physics, State University of Campinas, Brazil

March 1, 2024

Here I present a simple introduction to the variational theory, which is a well-known method for estimating ground-state energies using trial wavefunctions. We begin by presenting and proving the variational theorem, following with two examples for illustration.

Introduction

In textbook approaches to quantum mechanics we usually learn that the only atomic system for which the Schrödinger equation is exactly-solvable is the Hydrogen (1 proton + 1 electron). For the case of multi-electron atoms one must, therefore, rely on other techniques, such as perturbation theory, numerical methods or the so-called variational theorem, which I now introduce.

Although very well established as a central technique for quantum chemistry and physical applications, particular attention has been raised today to this topic due to the rapid-growing development of quantum computing. For example, the Variational Quantum Eigensolver (VQE), one of the famous prominent quantum algorithms for Near-term Intermediate Scale Quanum (NISQ) devices, has been explored extensively by both academy and industry in practical problems involving molecular dynamics. As the name suggests, it is grounded upon the variational theory.

The discussion presented here is not intended to be original or to bring some new remarkable analysis. On the contrary, it is explicitly based on textbooks and other sources I used to learn this topic, together with my personal experience (see references at the end). My main intention is to provide a somewhat fast and concise material for those who are interested in learning the variational theorem for practical applications as quick as possible.

The variational theorem

Suppose we are given an atomic system whose Hamiltonian is H. Then suppose that we are asked to determine its ground-state (GS) energy, E_0. We know that if the system is not a Hydrogen atom, trying to do this analytically would be perhaps the 13-th task of Hercules. However, as good physicists we are, we can safely say that there *exists* a GS, $|\psi_0\rangle$, and that any other (excited) state $|\psi\rangle$ has an energy greater than E_0. Mathematically stated, we can write

$$\langle\psi| H |\psi\rangle \geq E_0 \tag{1}$$

where the equality holds only for $|\psi\rangle = |\psi_0\rangle$.

Now, the magic of the variational theorem is this: eq.(1) holds for *any* wavefunction, whatever it is, provided only that it satisfies the same boundary conditions as the original wavefunctions of the system (which we assume, of course, to be unknown).

Proof. Denote by $\{|\phi_i\rangle\}$ the eigenstates of H (which we do not know, but we can always assume that exist). As H is Hermitian, they form an orthonormal basis, so that any vector $|\psi\rangle$ belonging to the Hilbert space of the system can be expanded in terms of them:

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle. \tag{2}$$

Note that the possibility of this expansion requires that $|\psi\rangle$ satisfies the same boundary conditions as

the $|\phi_i\rangle$'s. Therefore, using that $H|\phi_i\rangle = E_i|\phi_i\rangle$, $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ and that $E_i > E_0$, $\forall i \neq 0$, the proof follows straightfowardly:

$$\langle\psi|(H - E_0)|\psi\rangle = \sum_{ij} c_i^* c_j \langle\phi_i|(H - E_0)|\phi_j\rangle$$
$$= \sum_{ij} c_i^* c_j (E_j - E_0)\delta_{ij}$$
$$= \sum_i |c_i|^2 (E_i - E_j) \geq 0 \blacksquare$$

In other words, if we do not know what $|\psi_0\rangle$ really is, we may try *guess* it in order to, at least, determine an upper bound to E_0. The smarter the guess, the closer we will be to E_0.

In practice, we usually suppose that $|\psi\rangle$ is a function of a set of parameters $|\psi\rangle = |\psi(\theta)\rangle$, $\theta = \{\theta_1, ..., \theta_N\}$. Then we start to *vary* these parameters, testing each new $|\psi(\theta)\rangle$ we get on eq.(1). The one that produces the lowest result will be the *optimal* approximation we can get to $|\psi_0\rangle$.

To conclude this section, three important details: (i) For some reason $|\psi(\theta)\rangle$ is usually called an *Ansatz*, which is a German word for something we put "by hand"; (ii) If we choose to use an Ansatz that is not normalized, we must replace the LHS of eq.(1) by $\langle\psi|H|\psi\rangle / \langle\psi|\psi\rangle$. This latter quantity is known as the *Rayleigh ratio*; and (iii) Eq.(1) is usually applied in the *position representation*, the good old way of doing quantum mechanics. Casting it in 1-D we have

$$\int_{-\infty}^{\infty} dx\,\psi^*(x)H\psi(x) \geq E_0. \tag{3}$$

As you probably know, you can show the equivalence between these two equations by simply inserting $\int_{\infty}^{\infty} dx\,|x\rangle\langle x| = 1$ between H and each wavefunction in eq.(1), with $\psi(x) = \langle x|\psi\rangle$. The advantage of eq.(1) is, of course, that it is representation-independent (i.e. it says that we could work in with momentum, spin or whatever eigenstates we would like to). However, eq.(3) will be more useful for us in the following examples.

Examples

Although the most celebrated applications of the variational theorem are for complicated systems, it is instructive to test its validity with two systems that we are already familiar with.

Infinite square well

We know from the standard textbook discussion that the orthonormal set of position wavefunctions for a particle trapped in an infinite square well with sides ate $x = 0$ and $x = L$ is given by $\psi_n(x) = \sqrt{2/L}\sin(kx/L)$, where the boundary conditions $\psi(0) = \psi(L) = 0$ impose $k = k_n = n\pi/L$. Each of these are, of course, associated with an energy $E_n = \hbar^2 k_n^2 / 2m$.

However, suppose we didn't know the answer, but only the boundary conditions to this problem. A perfectly acceptable trial wavefunction could be, for ex.,

$$\psi_t(x) = Ax(L - x). \tag{4}$$

where t refers to "trial" and A is a complex constant. But first, in order to apply eq.(3), we need to normalize it:

$$|A|^2 \int_0^L dx\,\psi_t^*(x)\psi_t(x) = |A|^2 \int_0^L dx\,x^2(L-x)^2$$
$$= |A|^2 \left\{ \frac{L^5}{3} + \frac{L^5}{5} + \frac{L^5}{2} \right\}$$
$$= 1 \Rightarrow A = \sqrt{\frac{30}{L^5}}$$

where I chose A real by convention.

Now we use that inside the well the potential energy vanishes, so that $H = -(\hbar^2/2m)\partial^2/\partial x^2$. Then, using that $\partial^2 \psi_t(x)/\partial x^2 = -2$ in eq.(3), we get

$$\int_0^L \psi_t^*(x)H\psi_t(x) = \frac{30\hbar^2}{mL^5} \int_0^L x(L-x)$$
$$= \frac{5\hbar^2}{mL^2}$$
$$= E_t > E_1 = \frac{\hbar^2\pi^2}{2mL^2}.$$

Hence, we can see that $E_t/E_1 = 10/\pi^2 \approx 1.013$, meaning that our result has an error of only 1.3% when compared to the actual GS-energy. Of course, in this case we were able to directly evaluate if our result was satisfactory, as the infinite square well is a well-known system. When one comes to analyze more complex systems, one must find other ways to test the validity of his results.

Harmonic oscillator

Now, consider a quantum harmonic oscillator (HO) whose Hamiltonian is $H = -(\hbar^2/2m)\partial^2/\partial x^2 + m\omega\omega^2 x^2/2$. Let us apply the variational theorem with the trial wavefunction

$$\psi_t(x) = Ae^{-bx^2}, \quad b \text{ real and } > 0. \tag{5}$$

Of course, the boundary conditions that the oscillator cannot oscillate to infinity are translated as $\psi(x = \pm\infty) = 0$, which is satisfied by this choice. The normalization constant A can be determined immediately by using the Gaussian integral result, $\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\pi/a}$. We have that [1]

$$\int dx \psi_t^*(x)\psi_t(x) = |A|^2 \sqrt{\pi/2b}$$
$$= 1 \Rightarrow A = (2b/\pi)^{1/4}. \tag{6}$$

In this way, our normalized trial wavefunction depends on a free parameter b, i.e. $\psi_t \equiv \psi_t(x, b)$. From the discussion above, we could use several methods to vary b as wisely as possible in order to make E_0 a minimum. However, because our trial wavefunction is simple enough, we can first integrate the Rayleigh ratio and use the condition that if E_0 is a mininum with respect to b, then $\partial E_0/\partial b = 0$. Using that $\partial^2 \psi_t/\partial x^2 = -2b(1 - 2bx^2)Ae^{-bx^2}$ and that $\int dx x^2 e^{-ax^2} = -(\partial/\partial a)\int dx e^{-ax^2} = (1/2)\sqrt{\pi/a^3}$, we have

$$\int dx \psi_t^*(x)H\psi_t(x) = |A|^2 \left\{ \frac{\hbar^2 b}{m} \int dx e^{-2bx^2} + \right.$$
$$\left. + \left(\frac{m\omega^2}{2} - \frac{2\hbar^2 b^2}{m} \right) \int dx x^2 e^{-2bx^2} \right\}$$
$$= \frac{\hbar^2 b}{2m} + \frac{m\omega^2}{8b} \equiv E_0. \tag{7}$$

Thus,

$$\frac{\partial E_0}{\partial b} = 0 \Leftrightarrow \frac{m\omega^2}{8b^2} = \frac{\hbar^2}{2m}$$
$$\therefore b = \frac{m\omega}{2\hbar}. \tag{8}$$

This is the optimum value for b. Substituting it into the last equation we get

$$E_0 = \frac{\hbar\omega}{2}, \tag{9}$$

which is the *exact* GS-energy of the HO. This was expected, as the trial wavefunction we have chosen is, in fact, the exact wavefunction of this system. We did so in order to make clear where the parameters (to be optimized) may appear in applications of the variational theorem. We also confirm that it yields the right solution as we make an Ansatz that is closer to the exact solution.

---

[1]To avoid writting these "$\pm\infty$" so many times, let me fix that whenever integrals appear without their limits of integration, these must be understood.

Concluding remarks

In this article I explained and proved the variational theorem, following with two practical examples of how to use it. I would like to end by noting three useful facts, which you might explore from yourself:

1. There are several ways of justifying a good Ansatz, such as, for example, imposing that the gs-wavefunction should have no nodes (as can be rigorously shown for 1-D systems), or even using alternative formulations of quantum mechanics, like *second quantization*, for many-particle systems (See, for ex., [3], [2], [4]).

2. In general, even if our estimate of the gs-energy is pretty close to the actual value for the system we are considering, this does not mean that our Ansatz was equally close to the actual gs-wavefunction (See, for ex., [1]). For example, in the case of the infinite square well we can observe the two plots below:
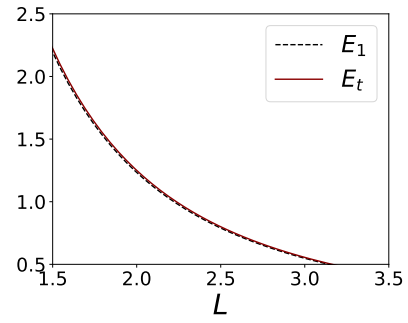
(a)
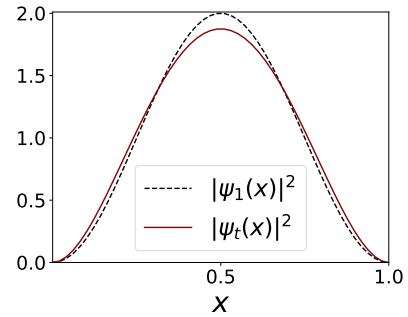
(b)

Figure 1: Comparison between actual and trial (a) GS-energies; and (b) GS-wavefunction for the infinite square well. We can see that the estimate of the GS-energy as a function of L fits very closely to the actual values, whereas the estimate for the GS-probability density does not match the same level of precision for all values of x.

3. It is possible to extend the variational method to estimate the energy of, not only the GS, but also the excited states of a system. See [REF]

References

References

[1] Peter Atkins and Ronald Friedman. *Molecular Quantum Mechanics*. Oxford University Press, 2010. ISBN: 0199541426.

[2] Gabriel T. Landi. *Second Quantization (Lecture notes)*. URL: `http://www.fmt.if.usp.br/~gtlandi/courses/second-quantization.pdf`. (accessed: 03.01.2024).

[3] Ira N. Levine. *Quantum Chemistry*. Pearson Education, 2013. ISBN: 0321803450.

[4] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, 2020. ISBN: 1108473229.