

1. The Molecular Hamiltonian

I. The Molecular Hamiltonian

This course will deal with the electronic structure and spectra of molecules. By electronic structure, we mean the information about where the electrons are and how they behave, which is deduced from the wave function for the electrons,

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \Psi(1, 2, \dots, N), \quad (1.1)$$

where \mathbf{x}_i (often abbreviated as i) is the coordinate specifying the state (position and spin) of the electron. We haven't learned about spin yet, but no matter. Just remember that sometimes electrons are "paired" (as in a chemical bond, lone pair, the singlet state of a carbene, or a "low-spin" metal complex), sometimes an electron is unpaired (as in a free radical or the hydrogen atom), and sometimes electrons have parallel (= the same) spin (as in a triplet carbene or a high-spin metal complex).

The electronic structure of a molecule is associated with the probability of observing various electrons in various states, which gives

$$\begin{aligned} \rho_N(\mathbf{x}_1, \dots, \mathbf{x}_N) &= |\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N)|^2 \\ &\equiv \Psi^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N). \end{aligned} \quad (1.2)$$

Recall that this is the Born postulate: the square of the wave function gives the probability of observing a certain state. In this case, the state is that electron 1 has position \mathbf{x}_1 , electron 2 has position \mathbf{x}_2 , etc.. Consequently, to determine the electronic structure of a molecule, we need to find the wave function. The problem of finding the wave function (or, more precisely, $|\Psi|^2$, the complex square of the wave function) is called the electronic structure problem. Electronic structure theory is the primary topic of this course.

Finding the molecular wave function, of course, is done by solving the Schrödinger equation. So let's suppose we have a molecule with N electrons and P nuclei. We'll say the nuclei have atomic numbers $\{Z_\alpha\}_{\alpha=1}^P$ and are located at the points $\{\mathbf{R}_\alpha\}_{\alpha=1}^P$. Now, the nuclei interact with one another according to Coulomb's law. That is,

$$V_{\alpha\beta} \equiv \frac{(\text{charge on nucleus } \alpha)(\text{charge on nucleus } \beta)}{4\pi\epsilon_0 (\text{distance between nucleus } \alpha \text{ and nucleus } \beta)} \quad (1.3)$$

Here, ε_0 is the permittivity of the vacuum.

We denote the distance between the nuclei with

$$|\mathbf{R}_\alpha - \mathbf{R}_\beta| = \sqrt{(x_\alpha - x_\beta)^2 + (y_\alpha - y_\beta)^2 + (z_\alpha - z_\beta)^2} \quad (1.4)$$

where $\mathbf{R}_\alpha = (x_\alpha, y_\alpha, z_\alpha)$ is the position of nucleus α . The charge on nucleus α is just the atomic number times the elementary charge, $Z_\alpha e$. (The charge of the proton is e ; the charge of the electron is $-e$.) So

$$V_{\alpha\beta} \equiv \frac{e^2 Z_\alpha Z_\beta}{4\pi\varepsilon_0 |\mathbf{R}_\alpha - \mathbf{R}_\beta|}. \quad (1.5)$$

The mass of the nucleus is denoted M_α . The mass could be looked up in the CRC handbook—be sure to use the appropriate nuclide masses, not the “isotope-weighted” ones in the periodic table! As a rough guess (loosely based on Weisacker mass formula),

$$M = (2Z + .0288 \cdot Z^{5/3}) m_H \quad (1.6)$$

The interaction between the electrons and nuclei has a similar form to the nuclear-nuclear repulsion potential (cf. (1.5)),

$$V_{\alpha i} \equiv \frac{e Z_\alpha (-e)}{4\pi\varepsilon_0 |\mathbf{R}_\alpha - \mathbf{r}_i|} = -\frac{e^2 Z_\alpha}{4\pi\varepsilon_0 |\mathbf{R}_\alpha - \mathbf{r}_i|}, \quad (1.7)$$

as does the electron-electron potential,

$$V_{ij} \equiv \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.8)$$

Now, we can write the Hamiltonian as,

$$\begin{aligned} \hat{H} = & \sum_{\alpha=1}^P \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 + \frac{1}{2} \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^P \frac{e^2}{4\pi\varepsilon_0} \left(\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \right) \right) \\ & + \sum_{i=1}^N \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^P \frac{e^2}{4\pi\varepsilon_0} \left(\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \right) \\ & + \sum_{\alpha=1}^P \sum_{i=1}^N -\frac{Z_\alpha e^2}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \end{aligned} \quad (1.9)$$

The first term in Eq. (1.9) depends only on the nuclear positions,

$$\hat{T}_n + V_{nn} \equiv \sum_{\alpha=1}^M \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 + \sum_{\beta=\alpha+1}^P \frac{e^2}{4\pi\varepsilon_0} \left(\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \right) \right), \quad (1.10)$$

the second term depends only on the electronic positions,

$$\hat{T}_e + V_{ee} \equiv \sum_{i=1}^N \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N \frac{e^2}{4\pi\varepsilon_0} \left(\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \right) \quad (1.11)$$

and the last term couples the electrons to the nuclei,

$$V_{ne} = \sum_{\alpha=1}^P \sum_{i=1}^N -\frac{Z_{\alpha}e^2}{4\pi\epsilon_0|\mathbf{r}_i - \mathbf{R}_{\alpha}|}. \quad (1.12)$$

In Eqs. (1.10) and (1.11), the first terms are the nuclear and electronic kinetic energies, denoted \hat{T}_n and \hat{T}_e , respectively. The second terms are the nuclear-nuclear and electron-electron repulsion energies, which are denoted V_{nn} and V_{ee} respectively.¹

¹ We can write the interparticle potential in several ways. The fundamental idea is that

$$\text{interparticle potential} = \sum_{\substack{\text{all pairs} \\ \text{of particles}}} (\text{repulsive potential for the pair})$$

For example, for three particles we have

$$\begin{aligned} V_{rep} &= V(|\mathbf{r}_1 - \mathbf{r}_2|) + V(|\mathbf{r}_1 - \mathbf{r}_3|) + V(|\mathbf{r}_1 - \mathbf{r}_4|) + V(|\mathbf{r}_2 - \mathbf{r}_3|) + V(|\mathbf{r}_2 - \mathbf{r}_4|) + V(|\mathbf{r}_3 - \mathbf{r}_4|) \\ &= \sum_{i=1}^3 \sum_{j=i+1}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) = \sum_{i=1}^4 \sum_{\substack{j=1 \\ j>i}}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) \end{aligned}$$

In the last equality, the $i = 4$ term does not contribute because there is no value of j between one and four for which $j > i$. For this reason, we sometimes use shorthand notation like

$$V_{rep} = \sum_{i=1}^N \sum_{j>i}^N V(|\mathbf{r}_i - \mathbf{r}_j|)$$

or even just

$$V_{rep} = \sum_{j>i} V(|\mathbf{r}_i - \mathbf{r}_j|).$$

Due to the symmetry of the interaction between the particles, there are many equivalent ways to write the interparticle interaction potential. Among the more common formulations and notations are

$$\begin{aligned} V_{rep} &= V(|\mathbf{r}_2 - \mathbf{r}_1|) + V(|\mathbf{r}_3 - \mathbf{r}_1|) + V(|\mathbf{r}_3 - \mathbf{r}_2|) + V(|\mathbf{r}_4 - \mathbf{r}_1|) + V(|\mathbf{r}_4 - \mathbf{r}_2|) + V(|\mathbf{r}_4 - \mathbf{r}_3|) \\ &= \sum_{i=2}^4 \sum_{j=1}^{i-1} V(|\mathbf{r}_i - \mathbf{r}_j|) = \sum_{i=1}^4 \sum_{\substack{j=1 \\ j<i}}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) \\ &= \sum_{i=1}^N \sum_{j<i}^N V(|\mathbf{r}_i - \mathbf{r}_j|) = \sum_{j<i} V(|\mathbf{r}_i - \mathbf{r}_j|) \end{aligned}$$

and

$$\begin{aligned} V_{rep} &= \frac{1}{2} \left(V(|\mathbf{r}_1 - \mathbf{r}_2|) + V(|\mathbf{r}_1 - \mathbf{r}_3|) + V(|\mathbf{r}_1 - \mathbf{r}_4|) + V(|\mathbf{r}_2 - \mathbf{r}_3|) + V(|\mathbf{r}_2 - \mathbf{r}_4|) + V(|\mathbf{r}_3 - \mathbf{r}_4|) \right. \\ &\quad \left. + V(|\mathbf{r}_2 - \mathbf{r}_1|) + V(|\mathbf{r}_3 - \mathbf{r}_1|) + V(|\mathbf{r}_3 - \mathbf{r}_2|) + V(|\mathbf{r}_4 - \mathbf{r}_1|) + V(|\mathbf{r}_4 - \mathbf{r}_2|) + V(|\mathbf{r}_4 - \mathbf{r}_3|) \right) \\ &= \frac{1}{2} \left(\sum_{i=1}^4 \sum_{\substack{j=1 \\ j>i}}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) + \sum_{i=1}^4 \sum_{\substack{j=1 \\ j<i}}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) \right) = \frac{1}{2} \left(\sum_{i=1}^4 \sum_{\substack{j=1 \\ j \neq i}}^4 V(|\mathbf{r}_i - \mathbf{r}_j|) \right) \\ &= \frac{1}{2} \sum_{j \neq i} V(|\mathbf{r}_i - \mathbf{r}_j|). \end{aligned}$$

There are several assumptions inherent in Eq. (1.9). First and foremost is the neglect of relativistic corrections for the motion of the nuclei and electrons. For light atoms and molecules composed of light atoms, this rarely causes qualitative errors. As a general rule, relativistic effects can be neglected for atoms in the first four rows of the periodic table (H-Kr), though relativistic corrections can be important for the weightier atoms in this sequence.

Second, we have assumed that the nucleus is a point-charge—that the nucleus is of zero size. This approximation makes very little impact on our study because the actual size of the nucleus with mass M is approximately

$$\begin{aligned} \text{radius} &\approx \left(1.2 \cdot 10^{-15} \text{ m}\right) \sqrt[3]{\frac{M}{m_H}} \\ &\approx (2.27 \cdot 10^{-5} \text{ Bohr}) \sqrt[3]{2Z}. \end{aligned} \tag{1.13}$$

Except for *very* heavy elements, the effect of finite nuclear size can be ignored. The effect of the finite nuclear size on the valence electrons in an atom or molecule is negligible. The fact that atomic nuclei are not simple point charges might be important in NMR, where the “Fermi-contact” interaction plays a role, though even in this context the model of the nucleus as a “point charge with spin” often seems appropriate.)

Third, we have neglected the presence of non-electromagnetic forces. Nuclear forces are short-ranged (the typical range of the strong and weak nuclear forces is 10^{-15} m), which means that they only act in a very small portion of the volume explored by the electrons. In general, in chemistry we are not interested in radioactivity. However, if one wishes to model electron capture by proton-rich nuclei, then one must include the nuclear forces in the Hamiltonian. (For example, ${}^7_4\text{Be}$ captures an electron to become ${}^7_3\text{Li}$. The primary force involved in electron capture is the weak force, which is ten million times weaker than the electromagnetic force.) In general, introduction of the nuclear forces should be done in the context of a quantum field theoretic and relativistic treatment.

We have also neglected the force of gravity, which is *not* short-ranged. We might plausibly add terms to the molecular Hamiltonian for the electron-electron, nuclear-nuclear, and electron-nuclear gravitational attraction. However, the gravitational interaction between particles is 10^{36} times weaker than the electromagnetic force, so neglecting gravitational effects is an excellent approximation.)

There has been much written on how to correct for the effects of the small terms in the molecular Hamiltonian that Eq. (1.9) omits. This course, however, will focus on molecules in the nonrelativistic and point-charge nucleus approximations.

II. Atomic Units

We can simplify Eq. (1.9) in several ways. First of all, we introduce atomic units, where

$$\hbar = m_e = \frac{e^2}{4\pi\epsilon_0} = 1 \quad (1.14)$$

In atomic units, the unit of length is the Bohr, and

$$1 \text{ Bohr} = .529177 \text{ \AA} \equiv .529177 \cdot 10^{-10} \text{ m} \quad (1.15)$$

A typical carbon-carbon single bond is about 3 Bohr (1.5 Å) long.

The atomic unit of energy is the Hartree, and one Hartree is

$$\begin{aligned} 1 \text{ Hartree} &= 27.2114 \text{ eV} \\ &= 2.195 \cdot 10^5 \text{ cm}^{-1} \\ &= 2625 \text{ kJ/mol} \\ &= 627 \text{ kcal/mol} \\ &= 6.58 \cdot 10^{15} \text{ Hz} \end{aligned} \quad (1.16)$$

At room temperature $kT \approx 2.5 \text{ kJ/mol}$. Typical activation energies for chemical reactions are about 10 kcal/mol. Typical bond strengths are a few hundred kJ/mol. Accurately describing chemical thermodynamics and kinetics requires determining to within a factor of kT , and so we require accuracy of about 1 kcal/mol.² Consequently, modern computational techniques strive for errors of .002 Hartree. Methods this accurate are said to have “chemical accuracy.” Modern spectroscopic techniques have sub-wavenumber accuracy. “Spectroscopic accuracy” refers to computational methods which have errors less than 10^{-6} Hartree.

² Actually, computing chemical properties usually only requires that certain energy differences—e.g. the difference of energy between the reactants and products of a chemical reaction, or between the reactants and the transition state—be accurately known. Due to cancellation of errors, we can often achieve “chemical accuracy” for thermodynamic and kinetic properties of a reaction even when the errors in the raw energies (not the energy differences) of the individual molecular species are much larger.

The reason we use atomic units is best described by analogy. Suppose that the primary distance I traverse is the commute between my house and the university, the primary activity of my day is drinking canned sodas, and my most important tool is a pen. Then, I might define a “Paul-centered unit system” as

$$\begin{aligned} 1 \text{ Paulcommute} &= 1.4 \text{ km} \\ 1 \text{ sodadrinkingtime} &= 30 \text{ min} \\ 1 \text{ penweight} &= 30 \text{ g.} \end{aligned} \tag{1.17}$$

Now, this means that the primary activities of my day are all “1” unit: it is one unit of distance to get to work, 1 unit of time to drink a soda, and my pens have one unit of weight. I can define my other activities in terms of these; to drive to my parents’ home in North Carolina is about 1100 Paulcommutes and will take about 32 sodadrinkingtimes. My suitcase will probably weigh about 600 penweights. By comparing to objects I am familiar with, it is easier to understand what the trip entails: I will have to buy gas three times; I’d better pack a lot of sodas; I should be careful not to wrench my back when I pick up my suitcase.

The idea behind atomic units is similar. We are going to be talking about electronic properties of atoms and molecules, and it is more useful to describe these properties in terms of Hartree and Bohr than it is to describe them in terms of calories (how much energy does it take to raise the temperature of a gram of water by 1 °C) or meters (related to the wave length of a certain light source). Thus, when we say that something has an energy of 24 Hartree we’ll know its energy is 48 times that of the Hydrogen atom and when we say that the bond length in H_2 is 1.4011 Bohr we know that this is 1.4011 times the average distance of the electron from the proton in the hydrogen atom. Since the primary objects we shall be considering have energies at the Hartree scale and spatial extents at the Bohr scale, it is useful to measure energies and distances in terms of these “characteristic” atomic quantities; this is just like me wanting to measure the time it takes to drive home in terms of sodadrinkingtimes instead of seconds. In fact, the system of atomic units can be defined by

$$\begin{aligned}
1 \text{ Hartree} &= 2 \times \text{Energy of the Hydrogen atom} \\
1 \text{ Bohr} &= \text{average distance of the electron from the proton} \\
&\quad \text{in the hydrogen atom} \\
1 \text{ (atomic unit of time)} &= \text{"time" it requires an electron to circle the Hydrogen atom} \\
&\quad \text{(in the Bohr model)} \\
&= 2.418884 \cdot 10^{-17} \text{ s}
\end{aligned} \tag{1.18}$$

III. The Born-Oppenheimer Approximation

The nucleus is very heavy compared to the electron,

$$m_H \approx 1832 \cdot m_e \tag{1.19}$$

It is thus expected to move very slowly compared to the electron. For example, in the immediate vicinity of the hydrogen atom (which has the lightest = fastest nucleus), the kinetic energy of the electron is about $\frac{1}{\pi}$ Hartree. A typical vibration frequency for a C-H stretch is about 2900 cm^{-1} . So the zero-point energy of this motion is $\frac{2900}{2} \text{ cm}^{-1}$, the kinetic energy of the proton is approximately $\frac{2900}{4} \text{ cm}^{-1} = .0033$ Hartree. The kinetic energy is $\frac{1}{2}mv^2$, so $v_{electron} \approx \sqrt{\frac{2}{\pi}} \approx .8$ while $v_H = \sqrt{\frac{2.033}{1832}} = .006$. The nucleus moves about 200 times slower than the electron. We expect, then, that so far as the electrons are concerned, the nuclei are “still”. That is, we expect that we would make only a very small error if we assumed that the nuclei were infinitely heavy,³ in which case the nuclei don’t move at all (they are fixed or, to use the technical term, “clamped” in space) so that the electrons effectively move with respect to the Hamiltonian (in atomic units):

$$\begin{aligned}
\hat{H}_{el} &= (\hat{T}_e + V_{en} + V_{ee}) + V_{nn} \\
&\equiv \sum_{i=1}^N \left(-\frac{1}{2} \nabla_i^2 + \sum_{\alpha=1}^M -\frac{Z_{\alpha}}{|\mathbf{r}_i - \mathbf{R}_{\alpha}|} + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^P \left(\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \right) + V_{nn}
\end{aligned} \tag{1.20}$$

This is called the Born-Oppenheimer approximation, though it was already employed by several workers (notably Heitler, London, and Slater) before the Born and Oppenheimer derived it. The approximation dates to 1927. Some workers omit the nuclear-nuclear repulsion term in Eq. (1.20), but it is convenient to include it here.

³ Clearly then, the Born-Oppenheimer approximation is most accurate for heavy atoms and least accurate for hydrogen-containing molecules.

The Born-Oppenheimer form, Eq. (1.20), *is* an approximation, and it is useful to “derive it” so that we can see when it works and when it fails. We start by writing the Schrödinger equation for the electrons and the nuclei as,

$$\left(\hat{T}_n + \hat{T}_e + V_{nn} + V_{ee} + V_{en}\right)\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) = E\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right). \quad (1.21)$$

Next, we write the exact wave function obtained by solving the Schrödinger equation as a product of two terms,

$$\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) = \psi_e\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\chi_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right). \quad (1.22)$$

The interpretation of Eq. (1.22) is that $\chi_n(\{\mathbf{R}_\alpha\})$ is the wave function for the nuclei in the molecule, while $\psi_e(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ is the wave function for the electrons in the molecule when the nuclei are described by the state function, $\chi_n(\{\mathbf{R}_\alpha\})$.⁴ Equation (1.22) is

⁴ The motivation for this construction comes from Bayes’ theorem. Bayes’ theorem states that the probability of observing events A and B together, $P(A, B)$ is the probability of observing event B , $P(B)$, times the probability of observing event A given that B occurs, $P(A|B)$. For example, suppose I want to calculate the likelihood that I both (a) meet a beautiful woman tomorrow afternoon and (b) go out on a date tomorrow night. First, I can calculate the likelihood that I will meet a lovely lass tomorrow afternoon. Let’s say that there is a 95% chance of that. Next, I can calculate the probability that I will go out on a date tomorrow night (given that I met a pretty girl earlier in the day, which *slightly* boosts the odds of that). Just to stave off public humiliation, allow me to overestimate the probability of having a hot date tomorrow night as, say, $\frac{1}{365}$. Then, we conclude that the probability that I will meet a voluptuous vixen (or a ravishing waif—I’m not picky) tomorrow afternoon and have a hot date (or even a not-so-hot date) tomorrow night is

$$\begin{aligned} \left(\begin{array}{c} \text{probability of Paul having a date} \\ \text{AND} \\ \text{meeting a fetching female} \end{array} \right) &= \left(\begin{array}{c} \text{probability of Paul meeting} \\ \text{a fetching female} \end{array} \right) \left(\begin{array}{c} \text{probability of Paul} \\ \text{having a date if he} \\ \text{meets a fetching female} \end{array} \right) \\ &= .95 \cdot \frac{1}{365} = .0026. \end{aligned}$$

Based on this analysis, I can conclude that it is safe to leave the house tomorrow without programming my VCR to tape tomorrow night’s basketball games; the likelihood of me missing the games is roughly equivalent to the probability that the devil is in the market for a zamboni.

We are doing something similar here. Write the probability of observing electrons at positions $\{\mathbf{r}_i\}_{i=1}^N$ and nuclei at positions $\{\mathbf{R}_\alpha\}_{\alpha=1}^P$ as $P\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$. Next, write the probability of observing the nuclei at positions $\{\mathbf{R}_\alpha\}_{\alpha=1}^P$ —irregardless of the position of the electrons—as $p_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$. Finally, write the probability of observing the electrons at $\{\mathbf{r}_i\}_{i=1}^N$ *given* the fact that the atomic nuclei are located at $\{\mathbf{R}_\alpha\}_{\alpha=1}^P$ as $p_e\left(\{\mathbf{r}_i\}_{i=1}^N \middle| \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$. Bayes’ theorem indicates that

merely defines a reasonable decomposition of the total wave function into a product of two terms; it is not an approximation.⁵

We substitute Eq. (1.22) into Eq. (1.21), and obtain

$$\left(\hat{T}_n + \hat{H}_{el}\right) \psi_e \left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \chi_n \left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) = E \psi_e \left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \chi_n \left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right). \quad (1.23)$$

Next, we use the expression for the nuclear kinetic energy (cf. Eq. (1.10))

$$\hat{T}_n \equiv \sum_{\alpha=1}^M \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2\right) \quad (1.24)$$

and the identity

$$\nabla^2 (f(\mathbf{r}) g(\mathbf{r})) = f(\mathbf{r}) \nabla^2 g(\mathbf{r}) + 2 \nabla f(\mathbf{r}) \cdot \nabla g(\mathbf{r}) + g(\mathbf{r}) \nabla^2 f(\mathbf{r}) \quad (1.25)$$

$$P\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) = p_e\left(\{\mathbf{r}_i\}_{i=1}^N \mid \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) p_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$$

Now, recall that the probability distribution function is the absolute square of the wave function. We use this to define wave function corresponding to the probability distribution function just introduced, namely,

$$\begin{aligned} \left|\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2 &= P\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \\ \left|\psi_e\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2 &= p_e\left(\{\mathbf{r}_i\}_{i=1}^N \mid \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \\ \left|\chi_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2 &= p_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \end{aligned}$$

Substituting into Bayes' theorem gives

$$\left|\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2 = \left|\psi_e\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2 \left|\chi_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)\right|^2$$

which motivates Eq. (1.22) and gives the aforementioned interpretation: $\chi_n(\{\mathbf{R}_\alpha\})$ is a nuclear wave function (corresponding to the probability of observing a given configuration of nuclei or, in chemical language, a given molecular structure); $\psi_e(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\})$ represents the positions of the electrons for given conformations of nuclei, including the coupling between electronic and nuclear degrees of freedom.

You can think of this as saying that the probability that the nuclei (which, being positively charged, are very attractive particles to electrons) are in a set location, times the probability that the electrons interact with (or “date,” if you insist on stretching the analogy) the nuclei in a specific way given that they are in that location, gives the total probability of a certain type of event (with the electrons and nuclei being in specified locations).

⁵ For example, given an exact wave function, $\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$, and a wave function for the nuclei, $\chi_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)$, Eq. (1.22) can be considered the *definition* of the “electronic” wave function,

$$\psi_e\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right) \equiv \frac{\Psi\left(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)}{\chi_n\left(\{\mathbf{R}_\alpha\}_{\alpha=1}^P\right)}.$$

to obtain

$$\begin{aligned}
\hat{T}_n \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) &= \sum_{\alpha=1}^M \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&= \sum_{\alpha=1}^M \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \frac{\hbar^2}{M_\alpha} \left(\nabla_\alpha \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \cdot \nabla_\alpha \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \right).
\end{aligned} \tag{1.26}$$

We substitute Eq. (1.26) into Eq. (1.23), obtaining,

$$\begin{aligned}
E \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) &= \sum_{\alpha=1}^M \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \frac{\hbar^2}{M_\alpha} \left(\nabla_\alpha \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \cdot \nabla_\alpha \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \right) \\
&\quad + \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(\hat{H}_{el} \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P).
\end{aligned} \tag{1.27}$$

We could solve Eq. (1.27) directly (if we were very smart and very, very patient) but it is better to write it as a set of coupled equations. That is, solve the equations⁶

$$\begin{aligned}
U(\{\mathbf{R}_i\}) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) &= \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(\hat{H}_{el} \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \frac{\hbar^2}{M_\alpha} \left(\nabla_\alpha \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \cdot \nabla_\alpha \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \right)
\end{aligned} \tag{1.28}$$

and

$$\begin{aligned}
E \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) &= \sum_{\alpha=1}^M \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(-\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \right) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&\quad + \sum_{\alpha=1}^M \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) U(\{\mathbf{R}_\alpha\}) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P) \\
&= \psi_e(\{\mathbf{r}_i\}_{i=1}^N, \{\mathbf{R}_\alpha\}_{\alpha=1}^P) \left(\hat{T}_n + U(\{\mathbf{R}_\alpha\}) \right) \chi_n(\{\mathbf{R}_\alpha\}_{\alpha=1}^P)
\end{aligned} \tag{1.29}$$

Equation (1.29) can be simplified to take the form of a Schrödinger equation for the nuclear coordinates alone,

⁶ Note that the potential energy surface, $U(\{\mathbf{R}_\alpha\})$ in Eq. (1.28) only depends on the nuclear positions, and not the electronic positions. This occurs because the remaining terms in the Schrödinger equation (the first nuclear kinetic energy term, which is included in Eq. (1.27) but not in Eq. (1.28)) can be written as $\psi_e(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) \hat{T}_n \chi_n(\{\mathbf{R}_\alpha\})$, which shows that the operator only effects the nuclear coordinates. Thus the electronic wave function does not enter into Eq. (1.30), with the consequence that the potential energy surface must depend on the nuclear coordinates alone.

$$(\hat{T}_n + U(\{\mathbf{R}_\alpha\}))\chi_n(\{\mathbf{R}_\alpha\}) = E\chi_n(\{\mathbf{R}_\alpha\}). \quad (1.30)$$

The procedure, then, is to solve Eq. (1.30) for $\chi_n(\{\mathbf{R}_\alpha\})$, substitute the result into Eq. (1.28) and then, leaving $\chi_n(\{\mathbf{R}_\alpha\})$ unchanged, solve Eq. (1.28) for $U(\{\mathbf{R}_\alpha\})$. Then, substitute $U(\{\mathbf{R}_\alpha\})$ into Eq. (1.30) and solve for an improved approximation to $\chi_n(\{\mathbf{R}_\alpha\})$. One uses the improved approximation to the nuclear wave function, $\chi_n(\{\mathbf{R}_\alpha\})$, to obtain an improved approximation to the potential energy surface, $U(\{\mathbf{R}_\alpha\})$, and repeats the process over and over and over again, until (hopefully) the process converges.⁷

To establish the validity of the coupled equations, Eqs. (1.28) and (1.29), merely substitute the equation for $U(\{\mathbf{R}_\alpha\})$ into Eq. (1.29). The original Schrödinger equation, Eq. (1.27), is then obtained.

Now, look at Eqs. (1.20) and (1.28). Note that the last two terms in Eq. (1.28) have the nuclear mass in the denominator. Even the mass of the proton is much larger than the mass of the electron (by 1800 times!), so for hydrogen-containing molecules, the last two terms in Eq. (1.28) are expected to be about one thousand times smaller than the first term. (The discrepancy is even larger for molecules composed entirely of “heavy” atoms.⁸) This suggests that we might ignore the last two terms, which gives us the Born-Oppenheimer approximation,

$$\hat{H}_{el}\psi_e^{BO}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\}) = U^{BO}(\{\mathbf{R}_\alpha\})\psi_e^{BO}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\}). \quad (1.31)$$

The nuclear wave function can then be determined using the analogue to Eq. (1.30),

$$(\hat{T}_n + U^{BO}(\{\mathbf{R}_\alpha\}))\chi_n^{BO}(\{\mathbf{R}_\alpha\}) = E^{BO}\chi_n^{BO}(\{\mathbf{R}_\alpha\}). \quad (1.32)$$

Eqs. (1.31) and (1.32) are usually called the electronic Schrödinger equation and the nuclear Schrödinger equation, respectively. Note that in the Born-Oppenheimer approximation the Schrödinger equation for nuclear motion (cf. Eqs. (1.29) and (1.32)) and Schrödinger equation for electronic motion (compare Eqs. (1.28) and (1.31)) are decoupled. The interpretation is that the motion of the electrons no longer depends on the motion of the nuclei (though the electrons’ motion clearly depends on the position of

⁷ This is called the self-consistent solution to a set of coupled equations; this method is related to the mathematical technique of “fixed-point iteration.” The self-consistent solution method will recur throughout this course.

⁸ To a quantum chemist, any atom other than hydrogen or helium is “heavy.”

the nuclei). Similarly, the motion of the nuclei no longer depends on the detailed motion of the electrons, but only on an “average” property (the energy of the electrons for a given nuclear configuration).

The Born-Oppenheimer energy is always less than the true energy. However, in most cases it is very close to the true energy. Sometimes the assumption that the electronic coordinates are “immune” to the nuclear motion is a poor approximation, though. This occurs when the electrons and nuclei move at similar speeds (e.g. when one “fires” one atom at another atom at very high speed, as to perform reactive (or non-reactive) scattering measurements). Similarly, if the “gap” between the electronic state and other nearby electronic states is very small, the nuclei can “donate” the energy needed to excite the electrons to an electronic excited state. But because electronic states are generally well-spaced in energy,⁹ this effect is usually small.

A word or two should be said about the interpretation of Eq. (1.32). The interpretation is that the nuclei move on a “potential energy surface” determined by the electronic energy, $U^{BO}(\{\mathbf{R}_\alpha\})$. The reason for including the nuclear-nuclear repulsion energy in the electronic Hamiltonian is precisely so that Eq. (1.32) takes a simple form, with $U^{BO}(\{\mathbf{R}_\alpha\})$ resembling the potential energy surfaces you are used to. For example, the Lennard-Jones potential for a diatomic molecule is an approximate model for the way the potential energy of the molecule, $U^{BO}(\mathbf{R}_1, \mathbf{R}_2)$ depends on the bond length, R .

The physical interpretation of a potential energy surface is that because the electrons are moving much faster than the nuclei, one can effectively consider the nuclei “fixed” when one solves the electronic Schrödinger equation. This allows one to neglect the momentum of the nuclei, and thus $\hat{\mathbf{p}}_\alpha \psi_e(\{\mathbf{r}_i\}, \{\mathbf{R}_\alpha\}) = 0$, where $\hat{\mathbf{p}}_\alpha \equiv -i\hbar \nabla_\alpha$ is the momentum operator for the α^{th} atomic nucleus, which leads to Eq. (1.31). Similarly, the nuclei move on a “potential energy surface,” with no detailed dependence on the individual electronic coordinates.

By analogy, suppose you were going on a hike through one of Canada’s great boreal forests on a bright summer day. Invariably, a cloud of flies and mosquitoes follows your every move. However, your progress is determined primarily by the

⁹ This is most commonly true for ground states; this assumption is generally less appropriate for excited states, where corrections to the Born-Oppenheimer method are often important.

landscape (whether it is uphill or downhill, etc.; this is the “potential energy surface” you are walking on); your motion is imperceptibly changed by the individual insects that swarm around you. (To you, it is just a “cloud” of insects, and you are oblivious to the fact that the swarm is composed of hungry individuals until the end of the day, when you sit in your tent and calamine your insect bites.) Similarly, an atom in a molecule sees a “swarm” or “cloud” of electrons flitting around it, but (in the Born-Oppenheimer approximation) the motion of the nuclei is independent of the individual electronic motions.

From the alternative perspective of the hungry insects, you are a buffet lunch. Because they fly so much faster than you walk and are so much more maneuverable than you are, the motion of the insects is rigidly coupled to yours. The pesky buggers follow your every move; it makes little difference to them whether you stay still or not because they can “adapt” to changes in your position very quickly—there is little chance that you can force them to forego dinner by “getting ahead of them.”. The same is true for the electrons. Because they move so much faster than the nuclei, changes in the nuclear positions are adapted to extremely quickly, and so a change in nuclear position does not lead to the system being “out of equilibrium” for long. In the Born-Oppenheimer approximation, one assumes that the electrons and the nuclear motions are rigidly coupled in this sense: the nuclei move so slowly that the electrons are always “optimally arranged” around the nuclei, so that the electronic structure of the molecule is never distorted from its equilibrium configuration about the nuclei.

Finally, to see when the Born-Oppenheimer approximation breaks down, imagine what happens if you start running through the forest. If you are fleet afoot, you will slowly leave behind the insects as they struggle to keep up—there will be fewer insects in front of you and a long tail of insects behind you in hot pursuit. (Of course, the entire hoard, with their appetites whet by the chase, will catch up to you when you pause to catch your breath.) For atoms that are moving very fast, this phenomenon is observed—the electrons have problems “keeping up” with the nuclei, and so the distribution of the electrons about the nuclear positions is no longer optimal. The Born-Oppenheimer approximation is no longer accurate in this situation, which is common in atomic and molecular scattering experiments.

