Chapter 7
The Canonical Commutation Relation

Learning goals:

(1) Explain why in quantum mechanics, the way we represent position and the way we represent momentum do not commute with each other.
(2) Describe how transitions between energy eigenstates yield spectral line energies.
(3) Explain how the Rydberg-Ritz combination principle requires quantum mechanics to be described with matrices (or operators).

Material to review:

(a) Matrix multiplication,
(b) Euler formula.
(c) Planck-Einstein relation.
(d) Bohr model.

In quantum mechanics, all of the nonclassical behavior arises from the canonical commutation relation, which summarizes how the degree of freedom of position and the degree of freedom of momentum do not commute. This result was first discovered by Heisenberg, and later was improved upon by Born and Jordan. We will investigate this derivation in detail later in the chapter. As we go through our arguments, it may not seem like what we are doing is something that is so important. But it is. And we will be exploring the consequences of this throughout the remainder of the book.

7.1 Rydberg-Ritz Combination Principle

In the late 1800s, the field of spectroscopy was an exceedingly important scientific enterprise. It was noticed that different elements emit light with very specific frequencies. This allowed us to identify elements in stars by just looking at the colors

of the light. It also allowed us to discover new materials that had not yet been seen on earth. It was a very powerful tool.

In spite of this importance, no one understood why the elements emitted these characteristic light frequencies only, and this was viewed as a major problem that needed to be solved. In many respects, as our knowledge of spectra improved, it allowed us to ultimately develop quantum mechanics as a formal theory that could describe how atoms emit light at characteristic frequencies. In fact, after World War II, precision spectroscopy provided critical tests for quantum field theory as well.

How are spectra measured? The experimental device is simple. First we need an atomic source that is excited with high energy. This is often done by using a vacuum tube, inserting nearly pure elemental gas into it, and then adding energy into the system (usually with high-voltage electrodes), similar to the cathode-ray tube. The atoms emit light in their characteristic frequencies, which can be measured if the light can be separated into its constituent colors.

There is a device that separates light into its characteristic colors that was known for a long time—it is a prism. But this was not accurate enough for spectroscopy measurements, so it was replaced by a diffraction grating. The better you could make a diffraction grating, which is done by drawing very fine nonreflective lines on glass, the better you could measure spectra. Note that the measurement converts the wavelength, or equivalently the energy of light into a measurement of position—many quantum measurements act in a similar way by relying on a position measurement because the detectors are always located somewhere. By 1900, the spectra of many atoms were tabulated, and a huge amount of data existed. Much of this work extended the visible light spectra into the ultraviolet and the infrared.

When Bohr developed his model for the atom, he postulated that certain orbits only were allowed and they would be stable even though classically they should emit light as the electron spiraled into the nucleus. So, from Bohr's perspective we only needed to determine the energy levels, and the spectra would result from the energy differences.

Bohr did not make up this idea of energy differences determining the spectra on his own. It was known well before his model, and describing how this works is our first job in this chapter. Note that the Planck-Einstein relation tells us that we can equally well think of frequencies to determine the allowed orbital energies.

Sommerfeld had extended the Bohr model from circular orbits to elliptical ones, and generalized it as well to better describe definite angular momentum states [15, 16] (the second citation is an English translation of the original). If an electron orbited in an ellipse, it would have a periodic orbit, repeating every period. Mathematically, such behavior is described by a Fourier series. This is a weighted sum of complex exponentials $e^{i\omega t}$, where the frequencies that appeared in the sum were all integer multiples of a fundamental frequency ω_0. Why you might ask? Well, if we consider the sum

$$\sum_{n=1}^{\infty} \left(a_n e^{in\omega_0 t} + b_n e^{-in\omega_0 t}\right) + c, \tag{7.1}$$

where a_n, b_n, and c are all constants, then one can easily see that this series is periodic with period $T = 2\pi/\omega_0$, because $e^{2\pi i n} = 1$. Furthermore, if we added any term with a different frequency in the exponent, it would not have this period. We will examine this concept in more detail a bit later in the chapter, but I hope it is clear to you that this series is periodic with period T.

What this told physicists at that time is that we should expect the frequencies used in our description to be of the form $n\omega_0$. This then tells us that the spectra should yield frequencies given by the fundamental and its harmonics—the intensities would be proportional to the modulus squared of the coefficients.

But, this was not seen in the data. Indeed, the data was found to be a bit more complicated, and this complication is what we want to explore next.

Energy (ev)	1.89	2.55	2.86	3.03
Wavelength (Å)	6562.79	4861.35	4340.472	4101.734

Table 7.1: Visible spectral lines of hydrogen.

Hydrogen was a favorite to examine its spectroscopy. Table 7.1 shows the visible spectra of hydrogen, which was the easiest one to measure. Johann Balmer, a Swiss Ph. D. who taught in a girls high school, enjoyed trying to make sense of the data for hydrogen (see Fig. 7.1). He found a remarkable formula that described all of the experimental results.

Fig. 7.1: Johann Balmer in the 1880s, around the time he discovered the Balmer formula. [Public domain. Converted to look like a stencil by the authors.]

Balmer determined that the spectral lines satisfied

$$\lambda = 3645.068\ \text{Å}\left(\frac{m^2}{m^2-4}\right) \quad \text{and} \quad E = 2\pi\hbar c R_H\left(\frac{1}{2^2} - \frac{1}{m^2}\right), \tag{7.2}$$

where $R_H = 10\,967\,800\ \mathrm{m}^{-1}$ and $m = 3,\ 4,\ 5,\ 6$ [1]. (Note that Balmer did not know about $\hbar$, so he used the inverse wavelength, but we use a more modern form with the energy.) The constant $2\pi\hbar c$ is given by 12 398.4 eV-Å. The Balmer formula was quite accurate. For example, we have for $m = 3$ that the Balmer transition has $\lambda = 6561.12$ Å as compared to the measured value of 6562.79 Å. If we focus on the energy form of the formula, we can see that it looks like the difference of two energies. This observation will turn out to be correct. It is consistent with the Bohr postulate.

Note further, that Balmer's formula predicts more transitions for $m > 6$. These transitions are in the ultraviolet, so they need more sophisticated detectors to measure them. Balmer's predictions for these higher-energy states were also confirmed. In honor of his work, these transitions in hydrogen are called the Balmer series.

Fig. 7.2: Johannes Rydberg, a Swedish physicist, who developed the Rydberg formula. [Photo by Per Bagge and courtesy of the University of Lund library (public domain). Converted to look like a stencil by the authors.]

Johannes Rydberg was a Swedish physicist who took the Balmer approach further [13, 14] (the second citation is a shortened English version) and found similar formulas could be approximately applied to alkali metals in addition to hydrogen (see Fig. 7.2). He also found that the Balmer formula was just the first of a series of formulas one could find, when he generalized the approach in 1888. His more general formula took the form

$$\lambda = 3645.068\ \text{Å}\left(\frac{m^2}{m^2-n^2}\right) \quad \text{and} \quad E = 2\pi\hbar c R_H\left(\frac{1}{n^2}-\frac{1}{m^2}\right), \tag{7.3}$$

where $m > n$. The Balmer formula corresponds to $n = 2$. This formula predicted a series of spectral lines in the ultraviolet, called the Lyman series, with $n = 1$, an infrared series with $n = 3$, called the Paschen series, and even more—the Brackett series with $n = 4$, the Pfund series with $n = 5$ and the Humpreys series with $n = 6$.

Niels Bohr was able to exactly produce these results with his model for hydrogen [2], which we will examine in more detail later in the chapter.

Fig. 7.3: Walther Ritz in 1911, around when he discovered the Ritz combination principle. [AIP Emilio Segrè Visual Archives, W. F. Meggers Collection. Public domain. Converted to look like a stencil by the authors.]

But, before Bohr came on the scene, a mathematician, Walther Ritz generalized the ideas even further (see Fig. 7.3). His result became the Rydberg-Ritz combination principle [12]. We will describe the basic idea first, and then discuss the additional subtleties associated with it. When applied to hydrogen, it was consistent with the Rydberg formula, and worked flawlessly. When applied to other atoms, it usually worked fairly well, but could predict transitions that do not exist, or could predict transitions with more sizable errors than one might have expected. We will explain both of these subtleties as we examine the principle in more detail.

In Fig. 7.4, we plot the energy levels and the possible photon energies given by the difference in energies for a jump between two levels. The different series are labeled by the final (lower) energy state. The Lyman series is primarily ultraviolet, the Balmer series primarily visible, and all other series are primarily infrared. The Ritz combination principle was simple: since energies that appear in spectra are energy differences, if we add or subtract the energies of two related spectral lines, it will be equal to the energy of another spectral line. In equations, we have

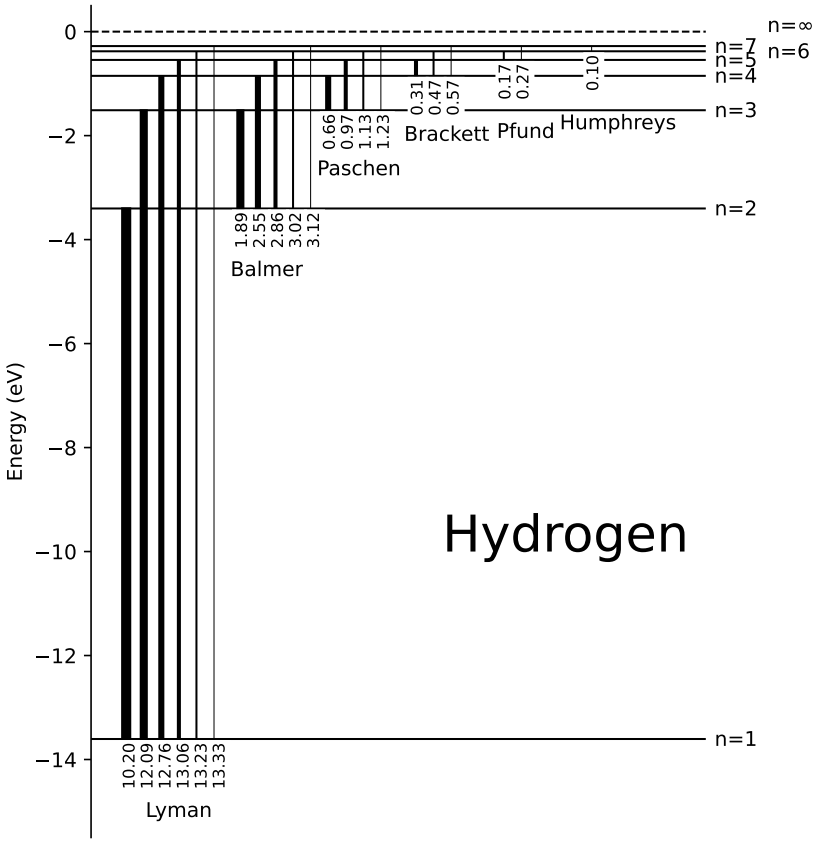

Fig. 7.4: Grotrian diagram for hydrogen. The horizontal lines are the different energy levels for hydrogen (in eV) and the vertical lines are transitions. Because of degeneracy in hydrogen, we do not need to worry about selection rules. This style for depicting spectra was developed by Grotrian [8].

$$E_{12} + E_{23} = (E_1 - E_2) + (E_2 - E_3) = E_1 - E_3 = E_{13}, \tag{7.4}$$

implying that if we see a transition from 1 to 2 and one from 2 to 3, then we should see one from 1 to 3. Indeed, the critical property of this result is that the energy of the light in an atomic spectra line is given by the difference in energies between two atomic energy levels—an idea that Bohr used heavily in the development of his theory.

But, the Ritz combination principle had some issues. Sometimes it predicted there to be a spectral line, but it was not seen. Why? The answer comes from the fact that spectra have selection rules. In Bohr's theory, he classified energy levels according to their angular momentum. The most common selection rule (for electric

dipole transitions) is that the angular momentum quantum number must increase or decrease by exactly one unit. The angular momentum states were labeled as you know from chemistry: $s = 0$, $p = 1$, $d = 2$, $f = 3$ and so on. Hence if we have a $s \rightarrow p$ transition and a $p \rightarrow d$ transition, then for the Ritz combination principle, we anticipate one at the sum of the two energies, but this is an $s \rightarrow d$ transition, which is not allowed for the electric dipole, so either it is very weak (coming from a magnetic dipole or electric quadrupole transition), or it is not seen at all. If, however, we have near degeneracy between energy levels of different angular momentum, then the Ritz combination principle works. Since hydrogen always has such near degeneracy, it satisfies the combination principle perfectly. But, other atoms don't always do so well. We give an example of how this works for helium with its partial Grotrian diagram in Fig. 7.5.

Let us look at three examples for helium in Fig. 7.5. The first is the transition from $1^1S \rightarrow 2^1P$ (21.22 eV) and the transition from $2^1P \rightarrow 3^1D$ (1.85 eV). Then Ritz says add these energies and there should be a transition from $1^1S \rightarrow 3^1D$. Adding them, we get 23.07 eV. Now, there is no $1^1S \rightarrow 3^1D$ transition due to selection rules, but there is a $1^1S \rightarrow 3^1P$ which is 23.09 eV. One can see it works! Another one is a similar pattern: $1^1S \rightarrow 3^1P$ (23.09 eV), $3^1P \rightarrow 4^1D$ (0.85 eV) and then looking at the sum, which is 23.74 eV, we see the transition from $1^1S \rightarrow 4^1P$ is 23.74 eV. It again works! The third is to compare the transition $2^1S \rightarrow 3^1P$ (2.48 eV) and $2^1S \rightarrow 4^1P$ (3.13 eV). Take the difference, which is (0.65 eV). There is a transition $3^1P \rightarrow 4^1D$ of 0.65 eV.

In all of these cases that worked, what we see is they occur because we have two states that are nearly degenerate, allowing the combination principle to approximately work. This is the general case one will see with most atoms other than hydrogen due to the selection rules. We also have an example where it does not work. Consider $2^1S \rightarrow 4^1P$ (3.13 eV) and $3^1S \rightarrow 4^1P$ (0.82 eV). Taking the difference, we find 2.31 eV. But, there is no transition for this—the closest is 2.48 eV for $2^1S \rightarrow 3^1P$. In this case the near degeneracy is not so near.

In summary, what was found in the early 1900s was that the spectra are described in terms of energy differences. At that time, there was no good theory to explain why this was so, and as we discuss in the next section, the Rydberg-Ritz combination principle illustrated how spectra could never be described classically. Then, Bohr developed a model that restricted the classical theories and was able to make his nonclassical theory work, but only for hydrogen. Relying on the same principle, Heisenberg finally discovered the proper way to incorporate it into a quantum theory. We will take you through this journey in the remainder of the chapter.

7.2 Atomic Spectra cannot be Described Classically

If we think of an atom as involving some kind of motion of its constituents, then that motion either is periodic, or is nonperiodic (creating a never repeating orbit). But, because the orbits of planets are periodic, and the Coulomb law is the same

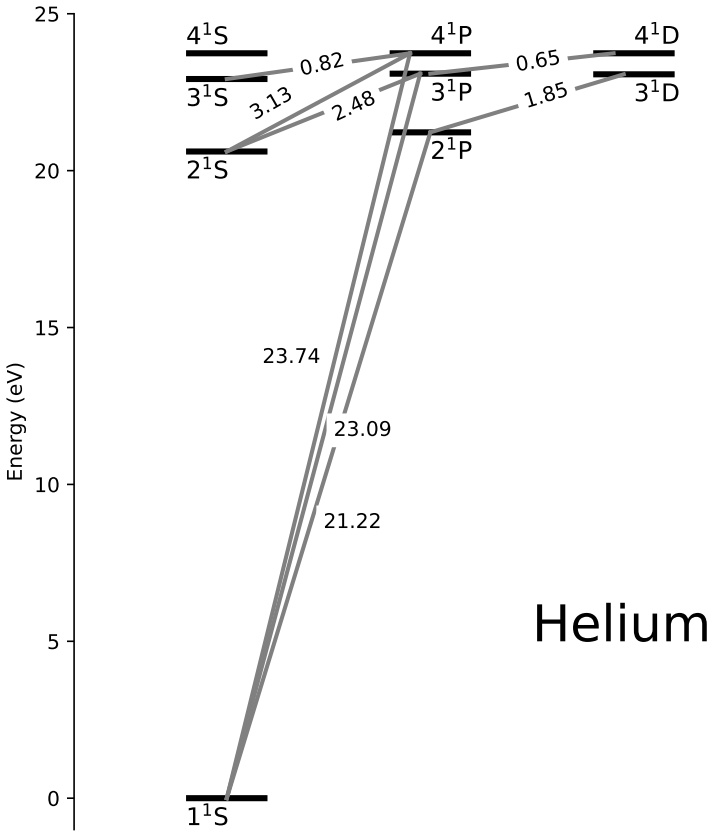

Fig. 7.5: Partial Grotrian diagram for the singlet states of helium. The horizontal lines are the different energy levels for hydrogen (in eV) and the angled lines are transitions, labeled by their energies in eV. Note that the energy levels are labeled according to the principle quantum number and the orbital angular momentum symmetry of the state. The 1 superscript indicates it is a singlet state. There are examples of lines that do satisfy the Rydberg-Ritz combination principle, but not all predictions lead to transitions due to selection rules. The lines are angled because one can only have transitions which change the angular momentum by ± 1.

functional form as the universal law of gravitation, we anticipate that atoms should be periodic in their motion as well.

As we described earlier the most general form of a periodic function is a Fourier series. It is easy to see that the Fourier series is always periodic with the period T. It is harder to prove that any periodic function with period T can be represented by such a series; we will not provide that proof here. Working with the Fourier series, we further use the Planck-Einstein relation to relate angular frequencies to energies (via $E = \hbar\omega$), and then we expect the spectra to correspond to the energies that appear in the Fourier series (that is, from the frequencies in the expansion). As you recall, they are all harmonics, so we would expect them all to be multiples of the harmonics (assuming those harmonic appear in the expansion with nonzero coefficients).

So, how would this work for some classical problems?

The simplest problem is the harmonic oscillator with frequency ω_0. We know that the periodic motion is given by $x(t) = x_0 \cos(\omega_0 t) + \frac{p_0}{\omega_0} \sin(\omega_0 t)$. You can see only one frequency appears in this expansion. Hence, there would be no transition if there is just one energy state, so perhaps the harmonic oscillator is not a good approximation for an atom, even though we tried to use it earlier when describing metals back in Chap. 1.

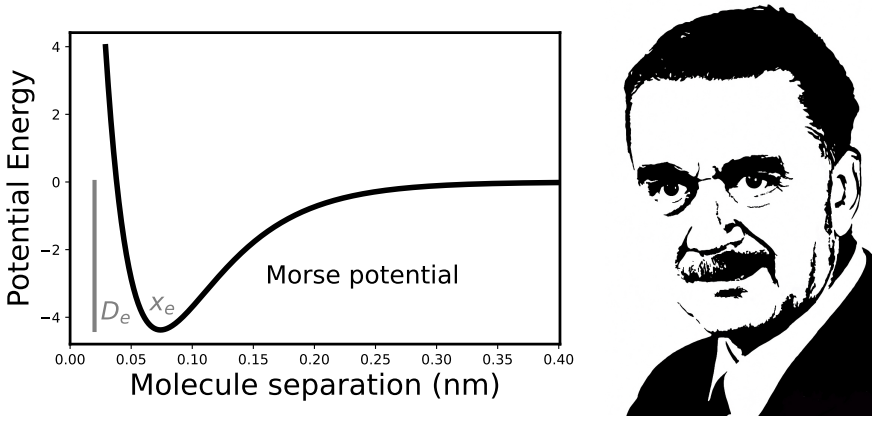

Fig. 7.6: Morse and his potential. The Morse potential goes to infinity on the left, and to zero on the right. It has a depth of D_e at $x = x_e$. Morse image taken when he was in his 50s. [Morse image from AIP Emilio Segrè Visual Archives, Physics Today Collection. Converted to look like a stencil by the authors.]

Another interesting potential is the Morse potential, which was originally created by Philip Morse at the dawn of quantum mechanics [10]. It is an asymmetric potential that was used to model the potential energy of a diatomic molecule of equilibrium extension x_e. You can see an image of Morse and his potential in Fig. 7.6. The potential is

$$V_{\text{Morse}}(x) = D_e\left(e^{-2a(x-x_e)} - 2e^{-a(x-x_e)}\right). \tag{7.5}$$

This looks like a complicated potential, and working out the force, by taking its negative spatial derivative, makes it seem impossible to find the orbits. But we can, as we show next. Just note, this is a French-cooking style of problem—many steps, but none too difficult. Just find the stamina to get through it all.

We solve it by looking at the conservation of energy. This is a standard approach in classical mechanics, but not always discussed in an introductory class. Let us see how it goes.

Energy conservation says $\frac{1}{2}m\dot{x}^2 + V(x) = E$ and requires $E < 0$ for a bound state with a periodic orbit from the shape of the potential. The velocity can be solved for giving us

$$\frac{dx}{dt} = \pm\sqrt{\tfrac{2}{m}(E - V(x))}, \tag{7.6}$$

and then allowing us to have a differential equation for the position, which is solved by integration, yielding

$$\sqrt{\tfrac{2}{m}}\int^t d\bar{t} = \pm\int^x \frac{d\bar{x}}{\sqrt{(E - V(\bar{x}))}}. \tag{7.7}$$

Next, we change variables, letting $y = e^{a(x-x_e)}$ and $dy/(ay) = d\bar{x}$. The integral becomes

$$\begin{aligned}\sqrt{\tfrac{2}{m}}t + c &= \pm\int^{e^{a(x-x_e)}} \frac{dy}{ay\sqrt{E - \frac{D_e}{y^2} + \frac{2D_e}{y}}} \\ &= \pm\int^{e^{a(x-x_e)}} \frac{dy}{a\sqrt{Ey^2 + 2D_e y - D_e}},\end{aligned} \tag{7.8}$$

with c the integration constant. The right hand side can be exactly integrated, to yield

$$\pm\sqrt{\tfrac{2a^2D_e}{m}}t + c = i\sqrt{\tfrac{D_e}{|E|}}\ln\left(2\sqrt{\tfrac{|E|}{D_e}y^2 + 2y - 1} + \frac{2i(|E|y + D_e)}{\sqrt{D_e|E|}}\right). \tag{7.9}$$

Define $\omega = \sqrt{\frac{2a^2|E|}{m}}$. Then, we have

$$\tfrac{c'}{2}e^{\mp i\omega t} - \frac{i(|E|y + D_e)}{\sqrt{D_e|E|}} = \sqrt{\tfrac{|E|}{D_e}y^2 + 2y - 1}, \tag{7.10}$$

with $c' = e^c$. Squaring both sides and recalling $E < 0$, then gives us

$$\frac{c'^2}{4}e^{\mp 2i\omega t} - \frac{ic'(|E|y + D_e)}{\sqrt{D_e|E|}}e^{\mp i\omega t} = -\frac{E + D_e}{E}. \tag{7.11}$$

Next, we solve for y and find that

$$y = i\sqrt{\frac{D_e}{|E|}}\frac{E+1}{Ec'}e^{\mp i\omega t} - \frac{D_e}{E} + i\frac{c'}{4}\sqrt{\frac{D_e}{|E|}}e^{\pm i\omega t}. \tag{7.12}$$

The term c' is determined by the initial conditions of the system. We pick $c' = 2i\sqrt{\frac{E+D_e}{|E|}}$, which corresponds to starting at a turning point of the potential (for a chosen E value). Then, we find

$$y = \frac{\sqrt{D_e^2 + D_e E}}{E}\cos(\omega t) - \frac{D_e}{E}. \tag{7.13}$$

Finally, going from y to x, we find that

$$x(t) = x_e + \frac{1}{a}\ln\left(\frac{\sqrt{D_e^2 + D_e E}}{E}\cos(\omega t) - \frac{D_e}{E}\right). \tag{7.14}$$

This has many harmonics in it, as can be seen by using the Taylor series expansion for $\ln(1-z)$, which gives

$$x(t) = x_e + \frac{1}{a}\left(\sum_{n=1}^{\infty}\left(\sqrt{\frac{D_e+E}{D_e}}\cos\omega t\right)^n + \ln\frac{D_e}{|E|}\right). \tag{7.15}$$

By using multiple angle formulas, to convert powers of cosine into cosines of $n\omega$, one can determine the full Fourier series. We do not need this though. One can clearly see that the frequencies would include all of the harmonics of ω. Interestingly, since ω is a function of E, the fundamental changes with E. So, if we had a Morse system with many energies, we would end up with spectra at every possible frequency. Only if the energies were restricted in some way to specific allowed values, would we have discrete spectra.

One can immediately see that if we took a classical point of view, insisted on periodic motion, and relied on the Planck-Einstein relation to convert frequencies to energies, then we are required to have all spectra to be harmonics of one or more fundamentals.

As we already know, the spectra do not take this form, so they cannot be described with a classical theory. They appear instead in the form of energy differences. There are two ways to obtain energy differences and neither is classical. The first way, due to Bohr, is to only allow specific orbits, with a discrete set of energies determined by the Bohr-Sommerfeld quantization rules, and to only allow transitions between different energy levels. Then, we obtain spectra that are energy differences. The other way, is to see that the fundamental objects to work with are matrices of a very specific form. This is what Heisenberg did, and we will examine his approach in detail later in the chapter.

7.3 Bohr Model

As we discussed before in Chapter 2, once Rutherford discovered the nucleus of the atom, Bohr came up with his theory for the atom [2]—it worked beautifully for hydrogen, but not so well for other atoms. It was based on the Bohr-Sommerfeld quantization condition, where the integral $J = \int p dx$ over a period of motion was quantized to be $2\pi n\hbar$; the integral is called the action. This quantization principle appeared to work in some cases, but not in others. Then, for the allowed orbits and their allowed energies, the observed spectra arose from the transition from state n with E_n to state m with E_m, with the photon of angular frequency that satisfies $\hbar\omega = E_n - E_m$, from the Planck-Einstein relation. Bohr insisted that the orbits were stable, even though classically they should emit radiation and decay.

Fig. 7.7: Paul Ehrenfest in the 1920s. Ehrenfest was an Austrian physicist who succeeded Lorentz in Leiden. He was quite ahead of his times, supporting many women to earn Ph.D.s during his tenure. Unfortunately he committed suicide in 1933. [Frenkel, Leningrad Physics-Technical Institute, Courtesy of AIP Emilio Segrè Visual Archives. Converted to look like a stencil by the authors.]

So, how could one reason why such a quantization condition exists? Paul Ehrenfest (see Fig. 7.7), an Austrian physicist, came up with a clever argument to justify the procedure based on a principle called adiabatic invariants [5]. We now briefly summarize how these ideas worked. An adiabatic invariant is something that remains constant when a parameter is slowly changed. If we consider the Coulomb problem, where the energy is given by

$$E = \frac{\vec{p}^2}{2m} - \frac{k}{r} = \frac{p_r^2}{2m} + \frac{\vec{L}^2}{2mr^2} - \frac{k}{r}, \tag{7.16}$$

where p_r is the radial momentum, $\vec{L}$ is the angular momentum, and $k = e^2/4\pi\epsilon_0$. In an adiabatic invariant calculation, we assume k varies slowly in time and wish to find something that remains constant as k changes. We assume further that the energy is essentially constant over an individual period of motion, but changes over many periods (similar to how we calculated the decay of the classical electron orbiting in an atom). Then we have

$$\dot{E} = -\frac{\dot{k}}{r} = -\frac{k}{r}\frac{\dot{k}}{k} = PE\,\frac{\dot{k}}{k}, \tag{7.17}$$

where PE denotes the potential energy. If we average over a period of motion and use an overbar to indicate the averaging, we have

$$\frac{d}{dt}\bar{E} = \frac{d}{dt}\frac{1}{T}\int_0^T E\,dt = \overline{PE}\,\frac{\dot{k}}{k}. \tag{7.18}$$

Here T is the period of the motion. Let us take a moment to understand this equation. It says, if I average the energy over one period of the motion, then the rate that that averaged energy changes is given by the rate that the constant k changes relative to its size multiplied by the rate at which the averaged potential energy changes. So, if we can relate the average potential energy to the average energy, we can find a differential equation we can solve. This is our next task.

The way we relate the time-averaged potential energy to the total energy is via the so-called viral theorem. The idea is that we examine the time derivative of the virial, here given by rp_r, to find such a relation. We have

$$\frac{d}{dt}(rp_r) = \dot{r}p_r + r\dot{p}_r = \frac{p_r}{m}p_r + r\left(\frac{\vec{L}^2}{mr^3} - \frac{k}{r^2}\right) = \underbrace{\frac{p_r^2}{m} + \frac{\vec{L}^2}{mr^2}}_{2\times KE}\ \underbrace{-\frac{k}{r}}_{PE}. \tag{7.19}$$

We used the fact that $p_r = m\dot{r}$ and $\dot{p}_r$ is the radial force. Now, we integrate both sides over a period and divide by the period. The left hand side will vanish, because both r and p_r are periodic in the motion. Then, we find

$$0 = 2\overline{KE} + \overline{PE} \ \text{ or } \ \overline{KE} = -\tfrac{1}{2}\overline{PE}. \tag{7.20}$$

Since $\bar{E} = \overline{KE} + \overline{PE}$, we find that $\bar{E} = \frac{1}{2}\overline{PE}$. This then allows us to transform Eq. (7.18) into

$$\frac{\frac{d}{dt}\bar{E}}{\bar{E}} = \frac{2\dot{k}}{k}, \ \text{ or } \ \tfrac{d}{dt}\ln(\bar{E}) = 2\tfrac{d}{dt}\ln(k). \tag{7.21}$$

This is solved by the adiabatic invariant result given by

$$\frac{k^2}{\bar{E}} = \text{constant}. \tag{7.22}$$

So, as k and $\bar{E}$ vary, the ratio $k^2/\bar{E}$ remains constant.

How did Ehrenfest use this to relate to the Bohr-Sommerfeld quantization condition? The answer is that one re-organizes the constant object so that it has dimensions of the action, and then sets it equal to $2\pi n\hbar$. You might note that because the differential equation has an arbitrary constant associated with it, we do need more information to determine the precise result. This is examined further in Prob. 7.5.

To get the right dimensions, we use dimensional analysis. It is essentially a matching game. For example, the action has dimensions of angular momentum, or position times momentum or energy times time. Let's keep the last one in mind. Now, since k has the dimensions of energy times length, then the adiabatic invariant has dimensions of energy times length squared. The only constant I can use is mass. How can I get it to have the dimensions of the action. It seems hard to sort out, but after a little trial and error, one will discover that mass times length squared is equal to energy times time squared, so then we see the right combination to quantize is

$$\sqrt{\frac{mk^2}{-E}} \propto 2\pi n\hbar. \tag{7.23}$$

Here, we recall that the energy is negative. We use the proportional sign, because the way this analysis worked, we did not get the numerical constants correct, just the dimensions. Nevertheless, it says the energies must behave like $E_n \propto 1/n^2$, which is exactly what Rydberg discovered by looking at the spectra.

Even after all of that analysis, all we really have is a handwaving argument. While it is true the approach appears to be working, why is it one would quantize the adiabatic invariants of the motion? No one was able to answer that question. But, it seemed comforting that quantizing the adiabatic invariant meant that there was a unique way to apply the Bohr-Sommerfeld quantization condition.

The Bohr model had a few assumptions. It said atoms have stable orbits chosen from all possible classical orbits by a procedure where we use the Bohr-Sommerfeld quantization principle, or we use the adiabatic invariant argument. It says the transitions between the atoms correspond to changes from one stable orbit to another. The emitted or absorbed photon had an energy given by the energy difference, so $\hbar\omega = E_n - E_m$.

When applied to hydrogen, it properly predicted the Rydberg constant in terms of other fundamental constants. It did not do as well for other atoms. Yet, even though hydrogen worked so well, there was still a puzzle. It had long been known that spectra similar to the Rydberg expressions we discussed earlier could be found for something that was observed in certain massive stars. They were discovered by the astronomer Edward Charles Pickering (see Fig. 7.8), who discovered these spectral lines in the star Zeta-puppis and identified three of the lines as belonging to hydrogen, but with half-odd integer values used in the Rydberg formula [11]. Alfred Fowler (see Fig. 7.8), a British astronomer, tried to create the lines on earth, and found he could only get them if he used mixtures of hydrogen with helium [6]. The

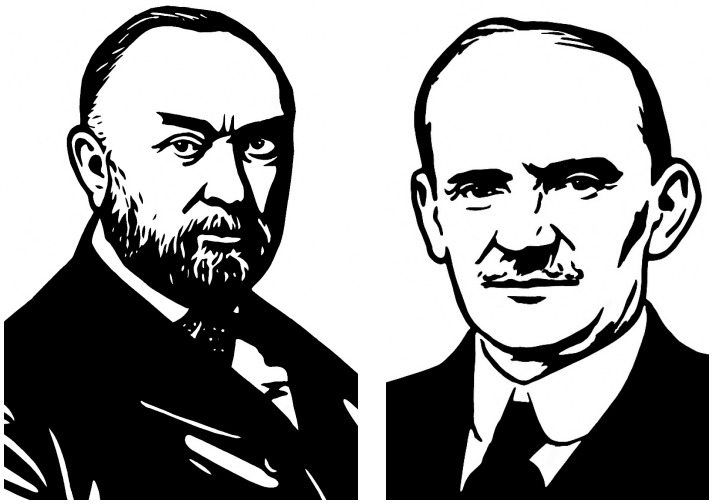

Fig. 7.8: (Left) Edward Charles Pickering from the 1880s. (Right) Alfred Fowler from 1900. Pickering was an American astronomer at Harvard, who discovered the extra spectral lines from the spectra in the star Zeta-puppis. Fowler was an English astronomer who performed experiments on hydrogen-helium mixtures and discovered he could produce the Pickering lines on earth. [Pickering image: public domain from wikimedia commons. Fowler image: Public domain from Library of Congress, George Bain Collection, LC-B2- 6573-10. Both converted to look like stencils by the authors.]

Pickering-Fowler spectra lines made no sense. While they fit the form of the Rydberg formula, they required half-odd integers, rather than integers. In addition, they were the only spectral lines thought to belong to hydrogen that Bohr had no theory for.

Bohr set out to solve this problem. He thought of an out-of-the-box solution based one the tidbit of information from Fowler's experiments—one could not get the lines unless there were both hydrogen and helium present. He worked his theory out for singly ionized helium. This ion looks just like hydrogen except the nuclear charge is $+2e$ instead of $+e$. You will get a chance to examine the details for how this works in the problems. It requires a small refinement of the Bohr model, which we discuss next, called the reduced mass.

One important refinement of the Bohr model is to take into account the fact that the nucleus is not infinitely massive. This means the atomic problem is to describe the joint behavior of the nucleus and the electron—called the two-body problem. We will illustrate it with a simplification of a one-dimensional system. The equations of motion of the two particles are

$$\dot{x}_1 = \frac{p_1}{m_1}, \quad \dot{p}_1 = -\frac{dV(x_1 - x_2)}{dx_1}, \quad \dot{x}_2 = \frac{p_2}{m_2}, \quad \text{and} \quad \dot{p}_2 = -\frac{dV(x_1 - x_2)}{dx_2}. \tag{7.24}$$

Note how we assume the potential depends on the distance between the two particles $x_1 - x_2$, which means it automatically satisfies Newton's third law—the force on particle one from particle two is equal and opposite to the force on particle two from particle one.

This two-body problem is reduced to an effective single-particle problem by going to average and relative coordinates. We work out these details now. The average coordinates are the center-of-mass position and the total momentum, or

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} = \frac{m_1 x_1 + m_2 x_2}{M} \quad \text{and} \quad P = p_1 + p_2, \tag{7.25}$$

with $M = m_1 + m_2$ the total mass. The relative position and momentum are

$$x_{\text{rel}} = x_1 - x_2 \quad \text{and} \quad p_{\text{rel}} = \mu \left(\frac{p_1}{m_1} - \frac{p_2}{m_2} \right), \tag{7.26}$$

where $\mu = m_1 m_2 / M$ is called the reduced mass.

To see that the system decouples into two independent systems, we need to compute the equations of motion for the average and relative coordinates. We find

$$\dot{X} = \frac{p_1 + p_2}{m_1 + m_2} = \frac{P}{M} \quad \text{and} \quad \dot{P} = 0, \tag{7.27}$$

because the two forces are equal and opposite and so they cancel. For the relative coordinates, we have

$$\dot{x}_{\text{rel}} = \dot{x}_1 - \dot{x}_2 = \frac{p_1}{m_1} - \frac{p_2}{m_2} = \frac{p_{\text{rel}}}{M} \quad \text{and} \quad \dot{p}_{\text{rel}} = -\frac{dV(x_{\text{rel}})}{dx_{\text{rel}}}. \tag{7.28}$$

This shows that the average degrees of freedom are those of a free particle with effective mass M, while the relative degrees of freedom are those of a particle with effective mass μ and having the same force act on it. For the Bohr model, it says the entire analysis is just like before with $m \to \mu$, and for hydrogen, we have

$$\mu = \frac{m_e m_p}{m_e + m_p} \approx \frac{1836}{1837} m_e \approx 0.9995 m_e. \tag{7.29}$$

This may seem like a small effect, and it is, but it can be seen. In fact, it can be used to "weigh" the mass of the nucleus by measuring spectroscopy. This is how deuterium was first discovered, which is investigated further in Prob. 7.10.

Now, we can get back to the Pickering-Fowler lines story. As we mentioned above, the lines were seen from stars, but not in conventional experiments on earth until Fowler did them, and he found they required hydrogen and helium in the tube to be observed. Bohr's bold idea was that the lines came from the transitions of singly ionized helium (He^+) [3]. How can one tell the difference—is it a half-odd integer quantum number for hydrogen, or an integer transition for singly ionized helium? If we calculate the two transitions, there is a slight difference because the nuclear mass of hydrogen and helium are different. This is what Bohr did and he found that

the data fit He^+ better than hydrogen. It showed how high precision spectroscopy could resolve the source for spectral lines that had been improperly assigned before. Einstein was so impressed with this result that he said it must be true that the Bohr theory is right.

Hence, the Bohr theory had significant success, even if it also had issues with it. The most important aspects of the theory for our next task are that it showed (i) that atoms have definite energy states and (ii) the emitted light has its frequency determined by the Planck-Einstein relation and energy conservation via $\hbar\omega = E_i - E_f$ for a transition from an initial to final state. These two facts are critical for determining the canonical commutation relation.

7.4 First Derivation of the Canonical Commutation Relation

As we saw when we were investigating spin, the commutation relation between operators is important in determining the quantum properties of spin. Now, as we move toward examining the quantum properties that depend on position and momentum, we need to find the commutation relation between the position and momentum operators or $[\hat{x}, \hat{p}]$. How do we do this?

This was one of the hallmark discoveries of Heisenberg, Born, and Jordan, and we will see how they did it in the next section. Their approach was based on matrix mechanics and provides a rigorous derivation. But, it relies on properties of matrices that one only studies in a linear algebra class. So, it is useful to examine an alternative way. This is what we do here.

Herbert Green (see Fig. 7.9) was a New Zealand physicist who wrote a remarkable quantum mechanics textbook in the sixties called *Matrix Mechanics* [7]. In it, he found a simpler path toward the canonical commutation relation than what Heisenberg, Born, and Jordan worked out. The derivation is not completely rigorous, as you will see, but it does provide a simple and very strong argument in favor of the result.

We use all of the ideas developed so far. To start, we consider the atom in an initial state $|i\rangle$ that transitions to a final state $|f\rangle$. We make a simplification to consider a one-dimensional system only, and we choose the Hamiltonian to be general and of the form

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}). \tag{7.30}$$

The atomic states are definite energy states, so $\hat{H}\,|i\rangle = E_i\,|i\rangle$ and $\hat{H}\,|f\rangle = E_f\,|f\rangle$.

Our first step is to consider the expression

$$\langle f|[\hat{x}, \hat{H}]|i\rangle\,, \tag{7.31}$$

and evaluate it two different ways. First, we expand the commutator to $\hat{x}\hat{H} - \hat{H}\hat{x}$ and then let the $\hat{H}$ operator act to the left or to the right depending on which state it is adjacent to. Using the fact that the states have definite energy yields

Fig. 7.9: Herbert Green was a New Zealander, who developed a simple argument to determine the canonical commutation relation. [Image credit: Australian Academy of Science. Converted to look like a stencil by the authors.]

$$(E_i - E_f)\,\langle f|\hat{x}|i\rangle = \hbar\omega_{if}\,\langle f|\hat{x}|i\rangle\,, \tag{7.32}$$

after using the Planck-Einstein relation.

The next step uses Heisenberg's idea that

$$\langle f|\hat{x}|i\rangle = e^{-i\omega_{if}t}\,\mathbf{X}_{fi}, \tag{7.33}$$

where $\mathbf{X}_{fi}$ is a matrix-element that does not depend on time. This specific variation in time is called a harmonic variation in time. Using the Heisenberg postulate, leads us to

$$\langle f|[\hat{x},\hat{H}]|i\rangle = i\hbar\frac{d}{dt}\,\langle f|\hat{x}|i\rangle\,. \tag{7.34}$$

We next make the assumption that all of the time dependence is assigned to the operator, not the states, which is called the Heisenberg representation. Then because $\dot{\hat{x}} = \hat{p}/m$, we finally arrive at

$$\langle f|[\hat{x},\hat{H}]|i\rangle = i\hbar\,\langle f|\frac{\hat{p}}{m}|i\rangle\,. \tag{7.35}$$

This concludes our first evaluation.

For the second one, we directly evaluate the commutator in the bra-ket:

$$\left[\hat{x},\frac{\hat{p}^2}{2m}+V(\hat{x})\right] = \left[\hat{x},\frac{\hat{p}^2}{2m}\right] = \frac{1}{2m}\,(\hat{p}[\hat{x},\hat{p}] + [\hat{x},\hat{p}]\hat{p})\,. \tag{7.36}$$

The first equality occurs because an operator commutes with a function of the operator and the second follows by applying the Leibniz rule for commutators of products of operators.

Our next step involves the assumption we will use. We assume that the commutator $[\hat{x}, \hat{p}]$ commutes with $\hat{p}$, or $[\hat{p}, [\hat{x}, \hat{p}]] = 0$. This holds if the commutator is a number, and not an operator (more generally if it is a function of $\hat{p}$). We assume, however, that it is a number, which means we can factor it out of the matrix element and we obtain

$$\langle f|[\hat{x}, \hat{H}]|i\rangle = [\hat{x}, \hat{p}]\,\langle f|\frac{\hat{p}}{m}|i\rangle\,. \tag{7.37}$$

Comparing Eq. (7.35) to Eq. (7.37), we see that we must have

$$[\hat{x}, \hat{p}] = i\hbar. \tag{7.38}$$

This relation is called the canonical commutation relation. In most quantum mechanics textbooks it is postulated, but you see here that we can derive it, with one small assumption, which is true *a posteriori*.

What observations went into this derivation? We needed the Planck-Einstein relation between the frequency and energy of light, the Bohr idea that atoms have static states with definite energy and light is emitted when the electron transitions from an initial state to a final state, the Heisenberg result that the matrix element of an operator between two states oscillates harmonically according to their energy difference, and that the total energy is given by the sum of the kinetic and potential energies. But, really, these ideas come primarily from Einstein's concept of a photon and the experimentally observed spectra of atoms.

Note that we did not need to say anything about the form of the potential—indeed, it does not enter into the analysis, which holds for all potentials.

We have achieved a remarkable result. By carefully following the ideas that evolved from the observed spectral lines of atoms, we have found a way to motivate the canonical commutation relation. We will soon discover that essentially all of the quantum properties follow from this canonical commutation relation. Yes, it is that important.

7.5 Second Derivation of the Canonical Commutation Relation

Our next enterprise is to provide a fully rigorous derivation of the canonical commutation relation as carried out by Heisenberg, Born, and Jordan. We will start with a bit of the history before we get into the math. Note, however, that our derivation does not follow that of the pioneers. It instead follows a more recent study [17]. We will point out where we deviate from the older approach and why, when we get to that place in the derivation.

Legend has it, that Werner Heisenberg suffered from hay fever, so he traveled to a desolate rock of an island called Helgoland where hay fever symptoms would be

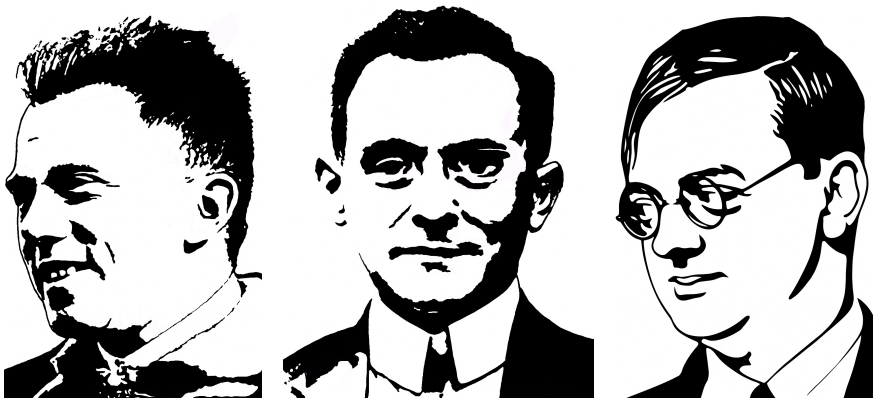

Fig. 7.10: (Left) Werner Heisenberg, (center) Max Born, and (right) Pascaul Jordan in the mid 1920s. These three developed the theory of matrix mechanics, and provided a path towards a rigorous derivation of the canonical commutation relation, although they did not provide such a derivation themselves. While Heisenberg and Born both won Nobel prizes, Jordan sympathized with the Nazis and did not. Heisenberg also supported the Nazis by heading their atomic bomb project. [Heisenberg image: Friedrich Hund, AIP Emilio Segrè Visual Archives, Physics Today Collection. Born image: AIP Emilio Segrè Visual Archives, Lande Collection. Jordan image: AIP Emilio Segrè Visual Archives, Segrè Collection. All figures converted to stencils by the authors.]

minimized. During this time he had been hard at work trying to figure out how to work with quantum mechanics in a more reasonable way than the Bohr-Sommerfeld quantization. In a moment of inspiration, in the very early hours of the morning, he saw how all the pieces fall together and created matrix mechanics.

The reality is, of course, a bit different. He had been working on this problem for quite some time and indeed much work was completed in Helgoland, but the process was perhaps a bit more gradual than sudden as is often described. See Fig. 7.10 for images of the pioneers who develop quantum mechanics through the properties of matrices.

Heisenberg's main train of thought was that he did not want to work with position and momentum as functions of time, as in a classical orbit, because, as he would say, one cannot observe such orbits. Instead, he wanted to focus on objects that describe the Rydberg-Ritz principle properly, and here, we land on having to consider matrices, just as we discussed previously. Because it is uncommon for most to have studied linear algebra at this time, we take a small detour to discuss matrices and matrix multiplication before heading to the main event.

Matrices are arrays of numbers, with two indices. The first is the row index and the second is the column index. Matrices have many different properties, and arise in many different contexts in physics, and you have probably worked with 2×2 and

3×3 versions previously. As an example, we show a 3×3 matrix, first abstractly and then concretely for a specific realization:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \mathbf{M}_{13} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} \end{pmatrix} = \begin{pmatrix} 3 & 1+i & 0 \\ 1-i & 2 & 4 \\ 0 & 4 & -1 \end{pmatrix}. \tag{7.39}$$

This specific matrix is also Hermitian, because $\mathbf{M}_{ji} = \mathbf{M}^*_{ij}$. Please check it carefully to verify it is Hermitian.

We want to remind you of the concept of matrix multiplication, which we discussed earlier in Chap. 3. The way it works for the product of two matrices **A** and **B** is that we take each row of **A** and each column of **B** and we pair them and thinking of each of these as vectors, we form the inner product between them. That entry goes into the respective entry of the $\mathbf{C} = \mathbf{AB}$ matrix.

Let us take the product of **M** with itself to illustrate.

$$\begin{pmatrix} 3 & 1+i & 0 \\ 1-i & 2 & 4 \\ 0 & 4 & -1 \end{pmatrix} \begin{pmatrix} 3 & 1+i & 0 \\ 1-i & 2 & 4 \\ 0 & 4 & -1 \end{pmatrix} = \begin{pmatrix} 11 & 5+5i & 4+4i \\ 5-5i & 22 & 4 \\ 4-4i & 4 & 17 \end{pmatrix}. \tag{7.40}$$

For example, let us look at the 13 element of the product. The first row vector is $(3, 1+i, 0)$ and the third column vector is $(0, 4, -1)$, so the product is

$$M^2_{13} = 3 \times 0 + (1+i) \times 4 + 0 \times (-1) = 4 + 4i. \tag{7.41}$$

The other cases are worked out similarly. Please carefully check that you can multiply these two matrices together.

More abstractly, we can write

$$\mathbf{C}_{mn} = \sum_{l=0}^{N} \mathbf{A}_{ml} \mathbf{B}_{ln}, \tag{7.42}$$

which summarizes the procedure we just described. This is the main result we need to discuss Heisenberg's matrix mechanics.

Heisenberg assumed the matrices in his theory took the form [9]

$$\mathbf{M}_{mn} = M_{mn} e^{\frac{i(E_m - E_n)t}{\hbar}}, \tag{7.43}$$

where M_{mn} were numbers independent of time, and all of the time dependence was given by the harmonic time variation. Heuristically, one can think of the mth row as corresponding to the state with energy E_m and the nth column as corresponding to the state with energy E_n, but Heisenberg knew nothing about states at that time, and the concept of a state is not needed for this analysis.

In order for the Heisenberg ansatz to be valid, when we construct composite matrices from some set of already defined matrices, they must continue to have the same functional form. Let us check by evaluating the square of **M**. We have

$$\mathbf{M}^2_{mn} = \sum_l \mathbf{M}_{ml}\mathbf{M}_{ln} = \sum_l \underbrace{M_{ml}M_{ln}}_{\text{indep. of } t} e^{\frac{i(E_m-E_l)t}{\hbar}+\frac{i(E_l-E_n)t}{\hbar}}$$
$$= \sum_l M_{ml}M_{ln}e^{\frac{i(E_m-E_n)t}{\hbar}}. \tag{7.44}$$

Hence, we immediately see that under matrix multiplication, the Heisenberg ansatz is preserved. This means it holds for any function that can be expressed in terms of a power series as well. It is said that Heisenberg rediscovered matrix multiplication to properly make his definitions robust and hold under matrix multiplication.

Okay. We have the ansatz for the form of matrices in Heisenberg's theory. Let us look at a few consequences of this. First, the Hamiltonian matrix $\mathbf{H}$, which is time independent, must be a diagonal matrix, with the E_n values along the diagonal. Second, these matrices can be infinite in size, as nothing precludes taking the limit $N \to \infty$. Third, we can compute time derivatives easily, as we did in the previous section, namely we have

$$\frac{d}{dt}\mathbf{M}_{mn} = \frac{i(E_m - E_n)}{\hbar}\mathbf{M}_{mn} = i\omega_{mn}\mathbf{M}_{mn}, \tag{7.45}$$

which defines ω_{mn} in the last term. Fourth, because the Hamiltonian is diagonal, we also have

$$\frac{d}{dt}\mathbf{M}_{mn} = \frac{i}{\hbar}[\mathbf{H}, \mathbf{M}]. \tag{7.46}$$

This last result is called the Heisenberg equation of motion.

Keeping these facts in mind, we are now ready to derive the canonical commutation relation. The approach we take follows closely the technical ideas from Born and Jordan [4].

We start with the correspondence principle, which relates quantum mechanics to classical mechanics and we have to decide how to implement it. The simplest thing to do is to insist that the position $\mathbf{X}$ and momentum $\mathbf{P}$ matrices satisfy the conventional equations of motion, namely that

$$\dot{\mathbf{X}} = \frac{1}{m}\mathbf{P} \text{ and } \dot{\mathbf{P}} = -V'(\mathbf{X}). \tag{7.47}$$

Note that we can think of the derivative in the second equation as being found by considering the scalar function $V(x)$, taking its derivative to form $V'(x)$ and then replacing $x \to \mathbf{X}$ in the power-series expression for V'. Making this requirement, ensures that we will eventually relate our quantum expressions to the well-known classical expressions, even if we are not completely sure what we do with these Heisenberg matrices.

The correspondence principle, then allows us to examine the time dependence of the commutator. We have

$$\frac{d}{dt}[\mathbf{X}, \mathbf{P}] = [\dot{\mathbf{X}}, \mathbf{P}] + [\mathbf{X}, \dot{\mathbf{P}}]. \tag{7.48}$$

Using the correspondence principle, the first term is proportional to $[\mathbf{P}, \mathbf{P}] = 0$, while the second term is proportional to $[\mathbf{X}, V'(\mathbf{X})] = 0$, because we can expand the force in a power series in $\mathbf{X}$ and $\mathbf{X}$ commutes with itself and any of its powers. This means

$$\frac{d}{dt}[\mathbf{X}, \mathbf{P}] = 0. \tag{7.49}$$

Because time-independent matrices must be diagonal, due to the Heisenberg form, we learn that $[\mathbf{X}, \mathbf{P}]$ is a diagonal matrix.

Next, we follow the argument Green used. We take the commutator of position with $\mathbf{H}$ and notice that it is proportional to a time derivative and hence proportional to the momentum. We have

$$[\mathbf{X}, \mathbf{H}] = i\hbar \frac{d}{dt}\mathbf{X} = \frac{i\hbar}{m}\mathbf{P}. \tag{7.50}$$

Next, we evaluate the commutator directly via

$$[\mathbf{X}, \mathbf{H}] = \frac{1}{2m}(\mathbf{P}[\mathbf{X}, \mathbf{P}] + [\mathbf{X}, \mathbf{P}]\mathbf{P}). \tag{7.51}$$

But, the commutator is a diagonal matrix, so by combining with the first result, we learn that

$$i\hbar \mathbf{P}_{mn} = \tfrac{1}{2}(c_m + c_n)\mathbf{P}_{mn}, \tag{7.52}$$

where we have that $[\mathbf{X}, \mathbf{P}]_{nn} = c_n$. This says the average of the diagonal matrix elements of the commutator is always equal to $i\hbar$, for all m and n that have $\mathbf{P}_{mn} \neq 0$. Now, you might think this means that we can immediately say that $c_n = i\hbar$, but there are other possible solutions. One is to have $P_{mn} \neq 0$ only if $n = m \pm 1$. Then we could have $c_{2n} = 2i\hbar$ and $c_{2n+1} = 0$, for example.

The correct argument to show that $c_n = i\hbar$ is to note that we want the canonical commutation relation to be independent of the potential. If we compare the basis used, which diagonalizes the Hamiltonian for one potential, and for another potential, they are different. The only matrix that is unchanged for any basis change is the identity matrix. Hence, the canonical commutation relation must satisfy $c_n = i\hbar$ for all n because it must have the same form for all potentials.

So how did this differ from the original work? There, they still believed the Bohr-Sommerfeld quantization condition was correct. It just needed to be refined. Their idea was to somehow differentiate the integral condition. They did this by taking finite differences of the quantization condition for two consecutive integers. This then immediately leads to the fact that the diagonal elements of the canonical commutation relation are all $i\hbar$. This is all Heisenberg could derive. Born and Jordan then showed that the off-diagonal elements are all zero. Born was so proud of this result, that he had it carved onto his gravestone.

7.6 Summary

In this chapter, we took you through a journey to examine the scientific thinking that went behind the derivation of the canonical commutation relation. This journey was replete with experimental data, and built off the work of others in the so-called old quantum theory. By the end, we were able to rigorously derive the canonical commutation relation. It might look a little odd to you. You almost certainly do not know what it means or how to use it yet. This is all fine. We will carefully show you the consequences of this as we move forward in our derivations.

Chapter 7 problems

Problem 7.1 Check Balmer's predictions. Check the accuracy of the other three visible spectral lines with the predictions of the Balmer formula. What is the average error over the four data points?

Problem 7.2 Selection rules and the Rydberg-Ritz combination principle. Write an essay that describes why the general Rydberg-Ritz principle should fail for dipole allowed transitions where the angular momentum changes by one unit (up or down) for each transition, unless two neighboring angular momentum states have nearly the same energy, allowing for a chain to occur. (For example $1s \rightarrow 2p$ followed by $2p \rightarrow 3s$ suggests a line with $3s \rightarrow 1s$, but this line is forbidden by the selection rules. However, if $3s$ and $3p$ are nearly degenerate than the $3p \rightarrow 1s$ transition would allow the principle to work.) But, if we consider three transitions, then we always will be able to make predictions that yield spectral lines that are not forbidden, without a degeneracy requirement. Interestingly, this idea was not used historically, although it seems like it could have been valuable. Explain how the three-difference rule would work.

Problem 7.3 Morse orbit derivation. In the derivation of the orbit of the Morse potential, we calculated an integral in Eq. (7.9) exactly. Rather than have you try to construct the integral, let us do the opposite. Differentiate the result we quoted and show that it is indeed equal to the integrand. You will need to do some creative algebra to carry this out.

Problem 7.4 There is no classical theory of spectroscopy possible. In the main text, we suggested that because periodic orbits are described by Fourier series, and if we assume the energies of the system are given by $\hbar$ multiplied by an angular frequency that appears in the Fourier series, then write an essay that explains why we cannot have a classical theory of spectroscopy using this approach after comparing with the experimental data.

Problem 7.5 Action-angle approach to quantization. One of the properties you will learn in advanced classical mechanics (but certainly not in an introductory course) is that the action is related to the period via

$$\frac{\partial J}{\partial E} = T(E), \tag{7.53}$$

where $T(E)$ is the period as a function of the energy. We will derive the formula now.

The action is given by $\oint p dx = J$, with the circle meaning the integral is performed over a period of the motion. Now, take the partial derivative with respect to E, using $p = m\sqrt{2m(E - V(x))}$. Verify that this implies that $\oint dt = T(E) = \partial J/\partial E$. This is why we have that formula.

We assume the correct action J is of the form we derived in the main text plus a multiplicative constant α, so

$$J(E) = \alpha \frac{\sqrt{m}k}{\sqrt{-E}}. \tag{7.54}$$

Use the formula for the energy of the Kepler problem as a function of the semi-major axis a and Kepler's third law, which relates the period to the semi-major axis (it is okay to look up these formulas from reliable sources, if you do not recall them) to determine the constant α. Then quantize this formula for the action and compute the energy. It turns out this formula for the energy is exact. If you used the general Kepler's laws, the formula holds for elliptic as well as circular orbits too.

Problem 7.6 Circular orbit approach to quantization. We follow the same approach of the previous problem, where we assume the form of the action with the multiplicative constant α. But now, we compute the action for a circular orbit in terms of the angular momentum L. So, using the formula for the action with the constant α replace $-E$ by the angular momentum L and other constants to relate the angular momentum to the action. Next, use the quantization condition for the angular momentum used by Bohr, such that $L = n\hbar$ and $J = 2\pi n\hbar$. Compute the energy for these circular orbits by quantizing the action.

Problem 7.7 The adiabatic invariant for the harmonic oscillator. Follow the steps to derive the adiabatic invariant for the harmonic oscillator, assuming the slowly varying object is the oscillator frequency ω. Show that the adiabatic invariant is then E/ω, which does have the dimensions of an action. Use this to determine the the energy levels of the harmonic oscillator by quantizing the invariant.

Problem 7.8 Pickering-Fowler lines. Using $R_\infty = 10\ 973\ 731.568\ 16\ \text{m}^{-1}$, and $R_H = R_\infty \frac{\mu_H}{m_e}$, compute the transitions from $n = 3$, 4, 5, and 6 to $m = 2$, Use $m_e c^2 = 511000$ eV and $m_p c^2 = 938000000$ eV. Compute the wavelength of the transitions in Angstroms to an accuracy of 0.1 Å.

Next, try the Pickering hypothesis, using $n = 3.5$, 4.5, 5.5, and 6.5. Compute the wavelength in Angstroms to an accuracy of 0.1 Å.

Two of the experimental lines were observed at 5412 Å and 4542 Å. Compare these experimental results to the Pickering calculation and to the result for singly ionized helium as Bohr suggested. For Bohr's calculation, we have the Rydberg constant is $4R_\infty/(1 + \frac{m_e}{M_{He}})$. Using $1 + \frac{m_e}{m_{He}} = 1.000\ 137$. Work with the transitions from $n = 7$ and $n = 9$ to $m = 4$.

Which agrees with the data better?

Note how it is the small differences from a high-precision measurement that allows us to determine which theory is correct.

Problem 7.9 Reduced mass for central-force problems in three dimensions. Repeat the calculation for the reduced mass and the separation of variables for the case where there is a central force, where the potential $V(|\vec{r}_1 - \vec{r}_2|)$ depends only on the *distance* between the two particles. Verify that the reduced mass that you find is the same as what we found in the one-dimensional case.

Problem 7.10 Discovery of deuterium. We have discussed in the main text how one can use high-precision spectroscopy to measure the mass of the nucleus. This idea was used by Harold Urey to discover the deuterium isotope of hydrogen.

The most common way to measure the mass of quantum particles is to use a mass spectrometer. This is a device that deflects charged particles in a magnetic field, similar to what we did in the Stern-Gerlach experiment, but here the particles have charge, so they feel the full Lorentz force and then they are bent according to their mass. We give them charge by ionizing them in the beam, typically by firing an electron beam at them. This is a very effective way to measure masses.

But, it is useless to find the mass two isotope of hydrogen. Why? Because mass spectrometers always have ionized hydrogen molecules in them. These also have mass two, so they completely mask the signal you might find from deuterium. One had to find another way.

In 1931, the neutron was not known, so a potential mass 2 nucleus of hydrogen was thought to be a combination of two protons and one electron. They knew that the reduced mass entered into the spectra, so they looked for the shifts due to deuterium. But the hydrogen lines were too strong and the deuterium contribution too weak to see. So they needed to find an enriched deuterium source.

Ferdinand Brickwedde worked in low temperature physics. The idea was to liquify hydrogen, and then let it evaporate. The lighter hydrogen evaporates more easily than the heavier deuterium. So if we repeatedly do this, we enhance the concentration of deuterium to higher and higher levels. Would it be high enough to see?

By now you know the answer is yes. Write an essay that described how much you might need to enhance the percentage of deuterium in the hydrogen sample to see the new peaks. How accurately do you need to measure the wavelengths to see the two separate peaks?

Problem 7.11 Flaw in the Herbert Green argument. Explain why the derivation motivated by Herbert Green is not a rigorous derivation even though the assumption we made—that the commutator is a number, not an operator, turned out to be true.

References

1. Balmer, J.J.: Notiz über die Spectrallinien des Wasserstoffs. Ann. Phys. **261**, 80–87 (1885). DOI 10.1002/andp.18852610506. URL https://doi.org/10.1002/andp.18852610506

2. Bohr, N.: I. On the constitution of atoms and molecules. The London, Edinburgh, and Dublin Phil. Mag. and J. Sci. **26**, 1–25 (1913). DOI 10.1080/14786441308634955. URL http://dx.doi.org/10.1080/14786441308634955
3. Bohr, N.: The Spectra of Helium and Hydrogen. Nature **92**, 231–232 (1913). DOI 10.1038/092231d0. URL https://doi.org/10.1038/092231d0
4. Born, M., Jordan, P.: Zur Quantenmechanik. Z. Phys. **34**, 858–888 (1925). DOI 10.1007/BF01328531. URL https://doi.org/10.1007/BF01328531
5. Ehrenfest, P.: Xlviii. adiabatic invariants and the theory of quanta. Phil. Mag. **33**, 500–513 (1917). DOI 10.1080/14786440608635664. URL https://doi.org/10.1080/14786440608635664
6. Fowler, A.: Observations of the Principal and other Series of Lines in the Spectrum of Hydrogen. Monthly Notices Roy. Astro. Soc. **73**, 62–63 (1912). DOI 10.1093/mnras/73.2.62. URL https://doi.org/10.1093/mnras/73.2.62
7. Green, H.S.: Matrix Mechanics. P. Noordhoff Ltd. (1965)
8. Grotrian, W.: Graphische Darstellung der Spektren von Atomen und Ionen mit ein, zwei und drei Valenzelektronen, ed. by M. Born and J. Franck, *Struktur der Materie in Einzeldarstellungen*, vol. 7. Springer-Verlag (1928)
9. Heisenberg, W.: uber quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. Z. Phys. **33**, 879–893 (1925). DOI 10.1007/BF01328377. URL https://doi.org/10.1007/BF01328377
10. Morse, P.M.: Diatomic molecules according to the wave mechanics. ii. vibrational levels. Phys. Rev. **34**, 57–64 (1929). DOI 10.1103/PhysRev.34.57. URL https://doi.org/10.1103/PhysRev.34.57
11. Pickering, E.C., Fleming, W.P.: Stars having peculiar spectra. New variable stars in Crux and Cygnus. Astrophys. J. **4**, 369–370 (1896). DOI 10.1086/140291. URL https://doi.org/10.1086/140291
12. Ritz, W.: Magnetische Atomfelder und Serienspektren. Ann. Phys, **330** (1908). DOI 10.1002/andp.19083300403. URL https://doi.org/10.1002/andp.19083300403
13. Rydberg, J.R.: Researches sur la constitution des spectres d'émission des éléments chimiques. Kongliga Svenska Vetenskaps-Akademiens Handlingar, 2nd series **23**, 1–177 (1889)
14. Rydberg, J.R.: Xxxiv. on the structure of the line-spectra of the chemical elements. Phil. Mag. **29**, 331–337 (1890). DOI 10.1080/14786449008619945. URL https://doi.org/10.1080/14786449008619945
15. Sommerfeld, A.: Zur Quantentheorie der Spektrallinien. Ann. Phys. **51**, 1–94 (1916). DOI 10.1002/andp.19163561702. URL http://dx.doi.org/10.1002/andp.19163561702
16. Sommerfeld, A.: On the theory of the balmer series: Presented at the meeting on 6 december 1915. Eur. Phys. J. H **39**, 157–177 (2014). DOI 10.1140/epjh/e2013-40053-8. URL https://doi.org/10.1140/epjh/e2013-40053-8
17. Tran, J., Doughty, L., Freericks, J.K.: The right way to introduce complex numbers in damped harmonic oscillators. Am. J. Phys. **93**, 437–440 (2025). DOI 10.1119/5.0220797. URL https://doi.org/10.1119/5.0220797