I assume the reader has taken a standard course in nonrelativistic quantum mechanics, at the advanced undergraduate or basic graduate level. For convenience I briefly review the important points here. Ideally you should find the concepts familiar; if not, I recommend going back to your quantum textbook before going further.

Quantum mechanics is one of the great triumphs of twentieth-century physics. While it certainly seems odd, arbitrary, or even bizarre when first encountered, with sufficient practice one can develop a strong intuition for much of quantum mechanics.

Part of the problem with quantum mechanics is its stitched-up nature. The other great triumph of twentieth-century physics, relativity, has a simpler origin. Relativity sprang mostly from one person, Albert Einstein (though I do not mean to denigrate the important contributions by Poincare, Lorentz, Hilbert, and others who foresaw many elements or who contributed further developments), who laid the foundations for both the special and general theories with elegant, concrete statements about observations: for special relativity, that one cannot make an observation that determines absolute motion, and in particular that the speed of light is the same in all inertial frames; and for general relativity, that one cannot make an observation that distinguishes between acceleration and gravity. If you takes these axioms seriously and rigorously, and add in a lot of math, one can deduce length contraction and time dilation, the equivalence of matter and energy, gravitational redshift, the bending of the fabric of space-time, and so on.

There are no analogous starting axioms for quantum mechanics, or least none with the same simple clarity and comprehensibility. And unlike relativ-
ity, quantum mechanics has multiple fathers: Planck and Bohr and Schrödinger and Heisenberg and de Broglie and Born and many others, including Einstein himself. Even to this day quantum mechanics bears the scars of this chimeric heritage: under the standard classroom pedagogy, the Copenhagen interpretation, a quantum system can change in time in two very distinct ways, either evolving under the time-dependent Schrödinger equation, or through the so-called collapse of the wavefunction at the moment a measurement is made. While there are alternate interpretations available, they are not used on a day-to-day basis because they typically make calculations, the bread and butter of a theoretical physicist, more difficult without adding accuracy.

I will not devote time to foundational issues, but instead will focus on technical points. And here, assuming one does not obsess too much on trying to cram our macroscopic intuition into the microscopic realm (the assumption that the universe at all length scales must be beholden to our paleolithic hominid brains is the height of arrogance), the math issues are straightforward, nearly all linear algebra. This review will not attempt to be comprehensive, but instead aimed at reminding the reader of the points most relevant to the subject at hand.

So let us begin.

### 1.1 Quantum mechanics and linear algebra

Quantum mechanics is a linear theory. The time-dependent Schrödinger equation for the time evolution of a system,

\[
\hat{H}\Psi(t) = i\hbar \frac{\partial}{\partial t}\Psi(t),
\]

is a linear differential equation. Therefore many of the fundamental tools of quantum mechanics are basic concepts in linear algebra: eigenvalue problems, vectors and matrices, and linear transformations.

Linear algebra occurs even in the uttermost bedrock of physics: measurement. Suppose you measure something in your laboratory: position, velocity, angular momentum, you name it. In quantum mechanics the measurement is represented by a linear operator \(\hat{O}\), and the only values that you can read on your laboratory equipment are the eigenvalues of that operator:

\[
\hat{O}\Phi_\lambda = \lambda\Phi_\lambda.
\]

We assume the operator \(\hat{O}\) is Hermitian, because then the eigenvalues \(\lambda\), and thus what is measured in the laboratory, is real-valued. For some operators \(\lambda\) is continuous-valued, but for others \(\lambda\) has discrete values, and it is this that makes quantum mechanics 'quantum.'

Further, we assume that any system is described by a wavefunction \(\Psi\). Technically \(\Psi\) and the \(\Phi_\lambda\) are members of a vector space, but for the moment I'll be cagey in describing that vector space, because the entire point of this book is to develop the tools to handle wavefunctions appropriate to the many-body
1.1. QUANTUM MECHANICS AND LINEAR ALGEBRA

problem! But since you have seen such wavefunctions before, you should not be stumped by this statement.

The Hermiticity of $\hat{O}$ also means the eigenfunctions (or eigenvectors, I’m still being cagey) $\Phi_\lambda$ must span the space and are orthogonal, which means they form a convenient basis for the expansion of any wavefunction in:

$$\Psi = \sum_\lambda c_\lambda \Phi_\lambda.$$  \hfill (1.3)

As mentioned earlier, for any single experiment the only allowed values you can measure in the laboratory are the eigenvalues $\lambda$. But quantum mechanics is random, and so even with the same initial conditions, with the same starting wavefunction $\Psi$, one can have different outcomes. (The idea that somewhere there is unnoticed clockwork that deterministically determines the particular outcome is known as hidden variables theory. Bell’s inequality has squashed the simplest and most appealing, at least, hidden variables theories. But no more on foundational issues.) The probability of measuring any particular value is given by $|c_\lambda|^2$.

In order to find the coefficient $c_\lambda$, called amplitudes in quantum mechanics, one needs to take a dot product, also known as an inner or scalar product. Novices get the idea that a dot product is between two vectors, but it isn’t that simple. If you have two real-valued vectors, $\vec{v} = (v_1, v_2, v_3)$ and $\vec{w} = (w_1, w_2, w_3)$, one might write the dot product $\vec{v} \cdot \vec{w} = v_1w_1 + v_2w_2 + v_3w_3$. But this definition is unsatisfactory for complex-value vectors, especially since we want the inner product to be positive definite (the inner product of any vector with itself to be a real, nonnegative number and is zero if and only if the vector itself is zero).

You’re likely aware that the solution is to modify the dot product to be $\vec{v}^\dagger \cdot \vec{w} = v_1^\ast w_1 + v_2^\ast w_2 + v_3^\ast w_3$. This seems ad hoc at first, but mathematicians formalized the solution through the concept of adjoint vectors. The purpose of an adjoint vector is to map a vector onto a scalar (that is, a single real or complex number). A common way to denote the adjoint operation is through a raised ‘dagger’ or $\dagger$. Thus if $\vec{v}$ is a vector, then $\vec{v}^\dagger$ is the adjoint.

Visually, it is helpful to represent vectors as column ($N \times 1$) arrays, and adjoint vectors as row ($1 \times N$) arrays:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad \vec{v}^\dagger = \begin{pmatrix} v_1^\ast & v_2^\ast & v_3^\ast \end{pmatrix},$$  \hfill (1.4)

where the complex conjugation in the adjoint guarantees that one has positive-definite magnitudes, that is, $\vec{v}^\dagger \cdot \vec{v}$ is real and positive if $\vec{v}$ is nonzero.

This convention for vectors and adjoint vectors is useful because it calls upon our notions of matrix multiplication: if $\mathbf{C} = \mathbf{A}\mathbf{B}$, to get the $i, j$th element of $\mathbf{C}$ we multiply the $i$th row of $\mathbf{A}$ by the $j$ column of $\mathbf{B}$. It also helps us because we can see immediately that

$$\vec{w}^\dagger \cdot \vec{v} = (\vec{v}^\dagger \cdot \vec{w})^*,$$  \hfill (1.5)
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

an important technical requirement for inner products (consult a convenient linear algebra text, or a convenient mathematician, if you have questions about this).

While the concept of the adjoint clears up technical problems, notationally writing \( \vec{w} \) and \( \overline{\vec{v}} \) is difficult to distinguish, and writing out column and row vectors with their components is clunky. Therefore in the next section we bring up Dirac’s clever notation.

1.2 Dirac’s bra-ket notation

An alternate notational convention used by physicists (but not mathematicians) is Dirac’s bra-ket notation. Rather than drawing an arrow over a vector \( \vec{v} \), Dirac wrote it as a ‘ket’ \( |v\rangle \). The adjoint vector \( \overline{\vec{v}} \) he wrote as a ‘bra’ \( \langle v| \). You should still think of kets as column vectors and bras as adjoint row vectors. But we now write the dot or inner product as \( \langle v|w\rangle \) which, to my mind, is cleaner to see than \( \overline{\vec{v}} \cdot \vec{w} \). For example, we can see that

\[
\overline{\vec{v}} \cdot \vec{w} = \langle v|w\rangle = (v_1^* \ v_2^* \ v_3^*) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = v_1^* w_1 + v_2^* w_2 + v_3^* w_3
\]

correctly gives us the inner product, while

\[
\langle w|\overline{\vec{v}} \rangle = |w\rangle \langle v| = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} (v_1^* \ v_2^* \ v_3^*)
\]

\[
= (v_1^* w_1 \ v_2^* w_1 \ v_3^* w_1) \\
= (v_1^* w_2 \ v_2^* w_2 \ v_3^* w_2) \\
= (v_1^* w_3 \ v_2^* w_3 \ v_3^* w_3)
\]

(1.6)
does not.

We also write (1.5) as

\[
\langle w|v\rangle = \langle v|w\rangle^*,
\]

(1.7)

which we’ll have frequent occasion to use.

So now given a set of orthonormal basis vector \( \{|e_i\}\) , we can easily write the orthonormality condition,

\[
\langle e_i|e_j\rangle = \delta_{ij}
\]

(1.8)

and the completeness identity

\[
\sum_i |e_i\rangle\langle e_i| = 1,
\]

(1.9)

where \( 1 \) is the unit matrix.
1.2. DIRAC’S BRA-KET NOTATION

The Dirac notation allows us to easily distinguish between a representation-free vector, \( |v_i \rangle \), and its components in some basis \( \{ |e_i \rangle \} \) given by \( v_i = \langle e_i | v \rangle \). Indeed, the notation allows us to visually distinguish between a vector ‘|something\rangle’ and a scalar ‘\( c = \langle something|something\ else \rangle \)’, and draws out attention that the way we get a scalar from a vector is to take the inner product with an adjoint vector.

When I first aspired to be a physicist, but had no physics or mathematics, I saw the time-independent Schrödinger equation written as

\[
H\Psi = E\Psi
\]  

(1.10)

which of course shed no light on the topic whatsoever. Nowadays I write the Schrödinger equation instead as

\[
\hat{H}|\Psi\rangle = E|\Psi\rangle
\]  

(1.11)

where the notation clearly distinguishes between an operator \( \hat{H} \), a vector \( |\Psi\rangle \), and a scalar \( E \).

Along with the Schrödinger equation we can rewrite our eigenvalue equation:

\[
\hat{O}|\Phi_\lambda\rangle = \lambda|\Phi_\lambda\rangle.
\]  

(1.12)

If the \( \{ |\Phi_\lambda\rangle \} \) form an orthonormal set, we write

\[
\langle \Phi_\lambda | \Phi_{\lambda'} \rangle = \delta_{\lambda\lambda'}
\]  

(1.13)

and we can expand any vector in this orthonormal basis set

\[
|\Psi\rangle = \sum_\lambda c_\lambda |\Phi_\lambda\rangle,
\]  

(1.14)

where the amplitudes \( c_\lambda \) are found through the inner product

\[
c_\lambda = \langle \Phi_\lambda | \Psi \rangle.
\]  

(1.15)

As the probability to measure the value \( \lambda \) is \( |c_\lambda|^2 \) then after many measurements the average value of \( \lambda \) is

\[
\sum_\lambda \lambda |c_\lambda|^2 = \sum_\lambda \lambda c_\lambda^* c_\lambda = \sum_\lambda \langle \Psi | \Phi_\lambda \rangle \langle \Phi_\lambda | \Psi \rangle
\]

\[
= \sum_\lambda \langle \Psi | \hat{O} |\Phi_\lambda\rangle \langle \Phi_\lambda | \Psi \rangle = \langle \Psi | \hat{O} |\Psi\rangle
\]  

(1.16)

where I’ve used \( \langle \alpha | \beta \rangle^* = \langle \beta | \alpha \rangle \), the eigenvalue equation (1.12) and the completeness relation (1.9). The final result \( \langle \Psi | \hat{O} | \Psi \rangle \) is called the expectation value, a misleading name—it is it is not the most expected value, but simply the average. Hence it is sometimes written simply as \( \langle \hat{O} \rangle \), which indicates the average. It is this bracket notation for the average which inspired Dirac to invent his bra-ket, or bracket, notation for vectors.
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

Although the time-evolution of a quantum system will not be a major concern for us, it is an instructive topic to review. Starting from the time-independent Schrödinger equation,

\[ \hat{H}|n\rangle = E_n|n\rangle, \quad (1.17) \]

one writes the general solution of the time-dependent equation

\[ i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \]

by expanding, as usual, in the eigenstates:

\[ |\Psi(t)\rangle = \sum_n c_n e^{iE_n t/\hbar} |n\rangle, \quad (1.19) \]

where one finds the coefficient \( c_n \) from the initial condition:

\[ c_n = \langle \psi_n | \Psi(0) \rangle. \quad (1.20) \]

When students first learn about vectors, they commonly think of them as simple lists (or a one-dimensional array) of numbers \( \vec{v} = (v_1, v_2, v_3) \). But this is not the vector itself; this is a representation of a vector through its components \( v_i \), which are complex numbers or scalars, in a particular basis. If we choose a different basis, the components change, though the vector itself does not.

The Dirac notation helps us to remember this distinction. \( |v\rangle \) is the physical, representation-free vector. It is only when we choose a basis \( \{ |e_i\rangle \} \) that we can find the components \( v_i = \langle e_i | v \rangle \).

Now the components of a vector are a necessarily evil, because we humans and our numerical servants computers can really only compute with scalars. But by making clear the distinction between a vector and its components in a particular basis, it helps to free us to choose a different basis, and to transform one from basis to another.

So: suppose we start with the components of a vector \( |v\rangle \) in some basis \( \{ |e_i\rangle \} \), that is \( v_i = \langle e_i | v \rangle \). Further suppose we have a second basis, \( \{ |f_\alpha\rangle \} \), which yields different components \( v'_\alpha = \langle f_\alpha | v \rangle \). (Note: one can and should employ notation to emphasize and clarify ideas; for example, as here, using lower-case latin letters \( i, j, \ldots \) for the first basis and lower-case greek letters \( \alpha, \beta, \ldots \) for the second. You’ll see this strategy several more times in this book.) We assume all bases are orthonormal (see the exercises for hints on how to deal with non-orthonormal bases), so that the standard completeness relation (1.9) also holds. We easily derive the unitary transformation from one basis to another by inserting the completeness relation:

\[ v'_\alpha = \langle f_\alpha | v \rangle = \langle f_\alpha | \left( \sum_i |e_i\rangle \langle e_i | \right) \rangle |v\rangle = \sum_i \langle f_\alpha | e_i \rangle \langle e_i | v \rangle = \sum_i U_{\alpha i} v_i, \quad (1.21) \]

where

\[ U_{\alpha i} = \langle f_\alpha | e_i \rangle \quad (1.22) \]
is the unitary transformation that takes us from one basis to another. You may remember that the columns, and separately the rows, of a unitary matrix form an orthonormal set of vectors. A useful interpretation of the elements of a unitary matrix are the components of a set of (basis) vectors \{ |e_i\}\ in another basis \{ |f_{\alpha}\}\; this is exactly what (1.22) means.

With the Dirac notation in hand, we can easily accomplish many important tasks.

Example. Given an orthonormal basis that satisfies the completeness relation, the unitarity of (1.22) becomes nearly tautological:

$$
(UU^\dagger)^{\alpha\beta} = \sum_i U_{\alpha i} U_{\beta i}^* = \sum_i \langle f_{\alpha} | e_i \rangle \langle e_i | f_{\beta} \rangle^* = \sum_i \langle f_{\alpha} | f_{\beta} \rangle = \delta_{\alpha\beta}.
$$

Example. We can derive a lemma: Inner products are independent of basis. Although we write \( \langle v|w \rangle \) for an inner product, in practice we need to carry it out in a basis, which we becomes obvious by inserting a completeness relation:

$$
\langle v|w \rangle = \sum_i \langle v|e_i \rangle \langle e_i|w \rangle = \sum_i v_i^* w_i.
$$

But now let’s insert two more completeness relations:

$$
\langle v|w \rangle = \sum_\alpha \langle v|f_{\alpha} \rangle \sum_\beta \langle f_{\alpha}|e_i \rangle \langle e_i|f_{\beta} \rangle \langle f_{\beta}|w \rangle = \sum_\alpha v_{\alpha}^* U_{\alpha\beta} U_{\beta i}^* w_i^* = \sum_\alpha v_{\alpha}^* \delta_{\alpha\beta} w_i = \sum_\alpha v_{\alpha}^* w_i
$$

where I used the unitaruitarity of \( U \). Hence we see we can take the inner product in any basis and get the same answer. A fancy phrase is the inner product is invariant under a change of basis.

### 1.3 The matrix formulation of quantum mechanics

By realizing that the ‘real’ vector is \( |v\rangle \) and not its components helps us to an important insight in quantum mechanics. In introductory quantum mechanics we learn the Born probabilistic interpretation of the wavefunction, that \( |\Psi(x)|^2 \) is the probability to find a particle at point \( x \) (more technically, it is the probability density, and \( |\Psi(x)|^2 dx \) is the probability to find a particle in a range \( dx \) about \( x \)).

But the wavefunction is \( |\Psi\rangle \) and its representation in coordinate space is \( \Psi(x) = \langle x | \Psi \rangle \). This is a difficult point, because it is difficult to grasp what
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

the basis vectors $|x\rangle$ are. (In fact, we can’t really grasp them, because they are vectors, and to represent them means choosing a basis...) The most accurate description is that the $\{|x\rangle\}$ are the eigenfunctions of the position operator $\hat{x}$, that is, $\hat{x}|x\rangle = x|x\rangle$. Because they are continuous-valued states, their orthonormality relation is given not by the Kronecker $\delta$ but by the Dirac $\delta$-distribution:

$$\langle x'|x\rangle = \delta(x-x').$$

There are two very important ideas that come out this, so pay attention. The first is critical to a broader understanding of quantum mechanics (although not so much to the topic of this book); the second is absolutely critical to understanding the occupation representation.

The first idea is the generalized statistical interpretation. Under the Born interpretation, $|\Psi(x)|^2 = |\langle x|\Psi\rangle|^2$ is the probability to find the particle with wavefunction $|\Psi\rangle$ in the state $|x\rangle$, which we colloquially state as ‘the probability to find the particle at position $x$.’ But this is true for any basis state $|\phi\rangle$, so that $|\langle \phi|\Psi\rangle|^2$ is the probability to find the particle with wavefunction $|\Psi\rangle$ in the state $|\phi\rangle$. In particular, if $|\phi\rangle$ is an eigenstate of a operator representing a measurement, then this is also the probability to obtain the eigenvalue (measurement) corresponding to $|\phi\rangle$.

The second idea is to emphasize, once again, that $\Psi(x)$ is not the real wavefunction, but that $|\phi\rangle$ is, and that we can write the components of $|\Psi\rangle$ in any useful basis set, $\{|\phi_\lambda\rangle\}$, and that the representation of the wavefunction in this basis,

$$|\Psi\rangle = \sum_\lambda |\phi_\lambda\rangle \langle \phi_\lambda|\Psi\rangle = \sum_\lambda c_\lambda |\phi_\lambda\rangle$$

is just as real as the coordinate space representation.

So, so example, the Fourier transform of a wavefunction,

$$\tilde{\Psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \Psi(x)dx$$

is really just a unitary transformation from coordinate space to wavenumber space. We can write this as

$$\tilde{\Psi}(k) = \langle k|\Psi\rangle = \sum_x \langle k|x\rangle \langle x|\Psi\rangle$$

where I treat the integral over $x$ the same as a sum “$\sum_x$.” (In an abstract sense, one can treat states with continuous labels, such as $|x\rangle$, and states with discrete labels, $|\phi_n\rangle$, in nearly the same fashion. There are of course important technical differences; you sum over the latter and integrate over the former; the latter’s orthonormality is given by a Kronecker-$\delta$ and the former by a Dirac-$\delta$. Consult an advanced quantum textbook for more details.)

You can see that the element of the unitary transformation is $\langle k|x\rangle = \exp(ikx)/\sqrt{2\pi}$. More importantly, however, the Fourier transformed wavefunction $\tilde{\Psi}(k)$ is just as ‘true’ and no less original than the coordinate space wavefunction $\Psi(x)$. 

I emphasize this point because this text revolves around an even more abstract representations of wavefunctions, one that nonetheless is just as ‘true’ as the coordinate-space representation.

We can now turn to the idea of a function space. We can treat functions as vectors in an infinite-dimensional vector space. The concept of an infinite-dimensional space, to say nothing of functions as vectors is nonintuitive and even overwhelming at first, but if one pays attention to the rigorous mathematical definition of an abstract vector space, then functions can indeed act as vectors. I won’t go into rigorous details, but the important idea is that a linear combination of vectors $\alpha \tilde{v} + \beta \tilde{w}$ is also a vector—and a linear combination of functions $\alpha f(x) + \beta g(x)$ is also a function.

We also can introduce an inner product for functions:

$$\langle f|g \rangle = \int f^*(x)g(x)dx.$$  \hspace{1cm} (1.29)

An inner product must have certain properties (linearity, positive-definitiveness, and so on) which this definition satisfies. But it also makes intuitive sense; in particular, if you have any experience with numerical quadrature (calculation of definite integrals), then you know that one discretizes a function $f(x)$ as an array $f(x_i)$ and the integral (1.29) is approximately

$$\int f^*(x)g(x)dx \approx \sum_i f^*(x_i)g(x_i)\Delta x.$$  \hspace{1cm} (1.30)

Aside from the lattice spacing $\Delta x$, this is very reminiscent of the standard dot product for finite vectors. We can also arrive at this definition by simply using the completeness relation

$$\langle f|g \rangle = \sum_x \langle f|x \rangle \langle x|g \rangle$$  \hspace{1cm} (1.31)

and noting that for a continuous variable like $x$ we replace the sum by an integral, and remembering that $\langle f|x \rangle = \langle x|f^* \rangle = f^*(x)$, we arrive at (1.29).

Now let’s go back and look again at expanding a wavefunction $|\Psi \rangle$ in a set of discrete basis functions $\phi_\lambda$. We wrote this down formally before, but now we can see how to evaluate the inner products:

$$\langle \phi_\lambda|\Psi \rangle = \sum_x \langle \phi|x \rangle \langle x|\Psi \rangle = \int \phi^*_\lambda(x)\Psi(x)dx.$$  \hspace{1cm} (1.32)

Of course, we could evaluate this equally well in Fourier space

$$\langle \phi_\lambda|\Psi \rangle = \sum_k \langle \phi|k \rangle \langle k|\Psi \rangle = \int \tilde{\phi}^*_\lambda(k)\tilde{\Psi}(k)dk$$  \hspace{1cm} (1.33)

which might be more convenient in some cases.
With this experience under our belts, let us also return to the time-independent Schrödinger equation:

$$\hat{H}|i\rangle = E|i\rangle. \quad (1.34)$$

While this now makes mathematical sense, as written it does not illuminate any method of solution. Again, this is because as humans we are not very good with vectors and operators; we are better with scalars (numbers); computers are even more dependent on scalars. To convert to scalars, we take the inner product with an convenient basis vector $|i\rangle$ (which might even be the coordinate basis vector $|x\rangle$):

$$\langle \phi_\lambda | \hat{H} | \Psi \rangle = E \langle \phi_\lambda | \Psi \rangle. \quad (1.35)$$

To go further, we strategically insert a completeness relation:

$$\sum_\mu \langle \phi_\lambda | \hat{H} | \phi_\mu \rangle \langle \phi_\mu | \Psi \rangle = E \langle \phi_\lambda | \Psi \rangle. \quad (1.36)$$

Now everything here has been converted to scalars. We can define

$$H_{\lambda\mu} = \langle \phi_\lambda | \hat{H} | \phi_\mu \rangle, \quad (1.37)$$

which is the matrix element of the operator $\hat{H}$ in the basis $\{|\phi_\lambda\rangle\}$. Matrix elements are very important in quantum mechanics—among other things they allow us to carry out actual and not just formal calculations—and computing them will be a central concern to us.

Why do we call $H_{\lambda\mu}$ a matrix elements? Recalling that we expand

$$|\Psi\rangle = \sum_\lambda c_\lambda |\phi_\lambda\rangle, \quad (1.38)$$

with $c_\lambda = \langle \phi_\lambda | \Psi \rangle$, we now have turned the abstract Schrödinger equation into a matrix equation:

$$\sum_\mu H_{\lambda\mu} c_\mu = E c_\lambda, \quad (1.39)$$

a formulation suitable for numerical solution by a computer.

Back at the dawn of quantum mechanics, Schrödinger developed his wave mechanics and Heisenberg his matrix mechanics. It was our friend Dirac who showed the two were equivalent. If $\hat{H}$ started life as a differential operator, it can be converted to matrix form by (1.37). Specifically, we can even write this in integral form:

$$H_{\lambda\mu} = \langle \phi_\lambda | \hat{H} | \phi_\mu \rangle = \sum_x \sum_{x'} \langle \phi_\lambda | x \rangle \langle x | \hat{H} | x' \rangle \langle x' | \phi_\mu \rangle$$

$$= \int \int \phi^*_\lambda(x) \hat{H}(x, x') \phi_\mu(x') dx \, dx'. \quad (1.40)$$

If, as so often is the case, $\hat{H}$ is a local operator (just functions of $x$ and low derivatives of $x$), we just write this in familiar form

$$H_{\lambda\mu} = \int \phi^*_\lambda(x) \hat{H}(x) \phi_\mu(x) dx, \quad (1.41)$$
1.3. THE MATRIX FORMULATION OF QUANTUM MECHANICS

with the implicit assumption $\hat{H}$ is a (local) operator in coordinate space.

I cannot stress enough the importance to our topic the concept of expanding a wavefunction in basis functions, and rewriting any differential equation as a matrix equation.

In the same way we could explicitly expand any vector in a basis, we can also expand any operator in a basis:

$$\hat{A} = \sum_{i,j} |e_i\rangle \langle e_i| \hat{A} |e_j\rangle \langle e_j| = \sum_{i,j} |e_i\rangle A_{ij} \langle e_j|.$$  \hfill (1.42)

The orientation of the bra and ket vectors help us see this is an operator.

Example. Given the unitary transformation (1.22) of a vector from one basis to another, we can easily derive the equivalent transformation of the matrix elements of an operator, that is, the unitary transformation of a matrix. For some operator $\hat{A}$, the matrix elements in the final basis are $A'_{\alpha\beta} = \langle f_\alpha| \hat{A} |f_\beta\rangle$ (here is another case where notation is important: I use primed quantities and Greek indices for objects in the final basis, and unprimed quantities and lowercase Latin indices for the initial basis) and inserting the completeness relation twice

$$A'_{\alpha\beta} = \langle f_\alpha| \sum_j |e_i\rangle \langle e_i| \hat{A} \sum_j |e_j\rangle \langle e_j| f_\beta\rangle = \sum_{ij} U_{\alpha i} A_{ij} U^\dagger_{\beta j}$$  \hfill (1.43)

or in more compact matrix notation

$$A' = UAU^\dagger. \quad (1.44)$$

Example. We can prove that the trace of an operator is independent of the (orthonormal) basis, that is, the trace is invariant under a unitary transformation. Given an operator $\hat{A}$, for some orthonormal basis $\{|e_i\}\}$ the matrix elements are $A_{ij} = \langle e_i| \hat{A} |e_j\rangle$, and the trace is

$$\text{tr} A = \sum_i A_{ii} = \sum_i \langle e_i| \hat{A} |e_i\rangle.$$  \hfill (1.45)

Now suppose we are in a different basis, $\{|f_\alpha\}\}$. We can insert twice the completeness relation:

$$\text{tr} A = \sum_i \langle e_i| \sum_\alpha |f_\alpha\rangle \langle f_\alpha| \hat{A} \sum_\beta |f_\beta\rangle \langle f_\beta| e_i\rangle$$

$$= \sum_\alpha \sum_{i,\beta} \langle e_i| f_\alpha\rangle A'_{\alpha\beta} \langle f_\beta| e_i\rangle = \sum_\alpha \sum_{i,\beta} \langle f_\beta| e_i\rangle \langle e_i| f_\alpha\rangle A'_{\alpha\beta}$$

$$= \sum_{\alpha,\beta} \langle f_\beta| f_\alpha\rangle A'_{\alpha\beta} = \sum_{\alpha,\beta} \delta_{\alpha\beta} A'_{\alpha\beta} = \sum_\alpha A'_{\alpha\alpha} = \text{tr} A'.$$  \hfill (1.46)

You will have a chance to practice these kinds of derivations in the exercises at the end of the chapter.
1.4 Example: Diagonalization in a basis

The idea that solving a differential eigenvalue equation is equivalent to diagonalizing a matrix is a key concept. Therefore I’ll work out an example in detail. Actually, I’ll work out examples throughout the text, as we develop more and more technology.

For now let us consider a problem in one-dimensional quantum mechanics, the delta-function potential:

\[ \hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} - \lambda \delta(x). \]  

(1.48)

This is a standard problem in introductory quantum mechanics and can be solved analytically; it has one bound state with an energy \( E = -m\lambda^2/2\hbar^2 \) and an value \( \langle x^2 \rangle = \hbar^2/m\lambda. \)

But we are going to find approximate solutions by diagonalizing. To do this we choose a basis \( \{ |n \rangle \} \) and compute matrix elements of the Hamiltonian, \( H_{mn} = \langle m | \hat{H} | n \rangle \). We then find the eigenvalues and eigenvectors of \( H \). Because a function space is infinite-dimensional, one must truncate the basis and thus the Hamiltonian matrix. From the variational theorem, see section 1.7, such a truncation is guaranteed to be an upper bound.

Almost always a basis is found by using the eigenstates of some other, simpler Hamiltonian. For this particular example we’ll use a particle in a box between \( x = -L/2 \) and \( x = L/2 \). The basis eigenstates are sine and cosine functions; the ones that interest us here are the even functions \( \phi_n(x) = \langle x | n \rangle = \sqrt{2/L} \cos((2n+1)\pi x/L), n = 0, 1, 2, \ldots, \) which vanish at the walls of the box. (For this box there are also sine eigenfunctions, but they vanish at the origin and so do not “see” the potential.) The matrix elements of the kinetic energy are diagonal in the basis of sines and cosines (which is why I choose it; for a general basis the kinetic energy matrix elements will not be diagonal):

\[ \langle m | -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} | n \rangle = \delta_{mn} \frac{\hbar^2 \pi^2 (2n + 1)^2}{2mL^2}. \]  

(1.49)

The matrix element of the delta-function potential between any two of these is simple:

\[ \langle m | -\lambda \delta(x) | n \rangle = -\lambda \frac{2}{L}. \]  

(1.50)

independent of \( m \) and \( n \).

(At the risk of repeating myself, let me emphasize that these matrix elements depend on the choice of basis; later on, in section 3.4 I’ll give the matrix elements of the kinetic energy and \( \delta \)-potential but in a basis of harmonic oscillator eigenfunctions, with very different individual values. Nonetheless, in the limit of large basis dimension, the ground state energy and other quantities converge to the same result.)
To further simplify things, dimensional analysis tells us that the basic unit of length is $\hbar/m\lambda$ and the basic unit of energy is $m\lambda^2/h^2$. Therefore let

$$L = L_0 \frac{h^2}{m\lambda},$$  \hspace{1cm} (1.51)

where $L_0$ is a dimensionless parameter.

Then the matrix element is

$$H_{mn} = \frac{m\lambda^2}{h^2} \times \left( -\frac{2}{L_0} + \delta_{mn} \frac{\pi^2}{2L_0^2} (2n+1)^2 \right).$$  \hspace{1cm} (1.52)

Now you can have fun: choose some value of $L_0$. For practical reasons, you have to truncate and choose a maximal value of $n$. With this truncated Hamiltonian in hand, you can run it through an eigensolver, for example in Mathematica, MatLab, Fortran or C.

I’ll choose a couple of cases. To begin with, I’ll let $L_0 = \pi^2/2$; as we’ll see later on, this is motivated by variational theory.

If we take only the lowest state, $n = 0$, then $E \approx H_{00} = -(2/\pi^2)m\lambda^2/h^2 = -0.2026m\lambda^2/h^2$, a mediocre approximation to the exact ground state of $-0.5m\lambda^2/h^2$.

Now let’s take $n = 0, 1$. The matrix is now

$$\frac{2}{\pi^2} \frac{m\lambda^2}{h^2} \begin{pmatrix} -1 & -2 \\ -2 & 7 \end{pmatrix},$$  \hspace{1cm} (1.53)

and the ground state energy is $(-6 + 4\sqrt{5})/\pi^2 \times (m\lambda^2/h^2) = -0.2983m\lambda^2/h^2$, not a vast improvement, but an improvement nonetheless.

We can continue numerically by including more states. Figure 1.1 show the results of finding the ground state energy as one increases the number of basis states from 1 to 500. In order to better examine the convergence, Figure 1.2 is a log-log plot of the difference $\Delta E$ between the numerical energy and the exact ground state energy as a function of the size of the basis. Also shown are different box sizes $L_0$. Notice that as the box size increases from 5($\approx \pi^2/2$) the energy starts out higher for one state, but then becomes lower. Thus Fig. 1.1 , 1.2 illustrate how complicated convergence can be. Depending on the basis chosen, one can start at a lower value, but may not ultimately converge to the correct result.

By diagonalizing we get not only information about the energies, but also the eigenstates themselves. By solving

$$\sum_n H_{mn} v_n = E v_n,$$

we can expand the wavefunction in coordinate space:

$$\psi(x) = \langle x | \psi \rangle = \sum_n v_n \langle x | n \rangle = \sum_{n=0}^{n_{\text{max}}} v_n \sqrt{\frac{2}{L}} \cos \left( \frac{(2n+1)\pi x}{L} \right).$$  \hspace{1cm} (1.54)
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

Figure 1.1: Convergence of the ground state energy for the δ-potential as a function of the number of basis states and of the size of the box \( L_0 \) (in units of \( \hbar^2 / m \)).

This I illustrate in Figure 1.3, where for a box of size \( L_0 = 7 \), I show the approximate ground state wavefunctions for 1, 10, and 100 basis states. For comparison I also show the exact ground state wavefunction, \( \psi(x) = (\sqrt{m\lambda}/\hbar) \exp(-m\lambda|x|^2/\hbar^2) \). By the time we get to 100 basis states, the cusp at \( x = 0 \) is well reproduced, and the only deviation from the exact answer is at the tails: the numerical wavefunction must vanish at the box boundaries.

1.5 Copenhagen, commutators, and conservation laws

You undoubtedly have heard the statement ‘it is impossible to know simultaneously the position and momentum of a particle’ as a fundamental principle of quantum mechanics. But what does that mean in terms of our mathematical framework?

Suppose we have a particle with a wavefunction \( |\Psi\rangle \). If we measure its position, that measurement is represented by the position operator \( \hat{x} \) with eigenvalue equation \( \hat{x} |x\rangle = x |x\rangle \), and the probability to find the particle at position \( x \) is given by \( |\langle x|\Psi\rangle|^2 \). Immediately thereafter the particle is now found in a new state \( |\Psi'\rangle = |x\rangle \). In the Copenhagen interpretation of quantum mechanics one says the wavefunction has ‘collapsed’ from \( |\Psi\rangle \) to \( |x\rangle \). If you don’t like this idea, and you aren’t alone in this, you can think of the Copenhagen interpretation as shorthand for some other, subtler process, such as decoherence.

What does it mean to say the particle is in the state \( |x\rangle \)? It means that
if you measure the position immediately again, you will find the particle still at position $x$ and not some other position $x'$; quantum mechanics is weird, but not that weird. A little more formally, the probability to find the particle at position $x'$ is $|\langle x'|\Phi\rangle|^2 = |\langle x'|x\rangle|^2$ which is zero unless $x' = x$.

(Even more formally: $|\langle x'|\Phi\rangle|^2 = |\delta(x' - x)|^2$ and one needs to integrate over $x$ and $x'$ to make everything finite. It’s unfortunate that in quantum mechanics some of the most intuitive situations, such as measuring position and momentum, have such mathematical complications. I take it as more evidence that our Olduvai-honed intuition is not trustworthy when it comes to the subatomic world.)

So, once you have made a measurement and collapsed a state into an eigenstate of a measurement operator, you can measure over and over and you will now get 100% of the time the same eigenvalue—which is another way of saying the uncertainty of the measurement is zero.

(Comment: one has to make the measurement immediately, which here means such a short period of time the wavefunction has not had time to evolve under the time-dependent Schrödinger equation.)

If you finally get tired of measuring the position over and over again, you might switch to measuring the momentum with operator $\hat{p}$ and eigenvalues $p$ and eigenstates $|p\rangle$. The same procedure applies: the probability to measure $p$ is $|\langle p|\Phi\rangle|^2$. But since we are in the eigenstate $|x\rangle$ from the previous measurement, this is $|\langle p|x\rangle|^2$, which because $\langle p|x\rangle \approx \exp(ipx/\hbar)$, is constant and independent of $p$. In other words, you can take on any momentum value, and the uncertainty is the largest possible.
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

You are now in the state $|\Psi''\rangle = |p\rangle$. You can measure momentum over and over and you will now always get the same value $p$, etc.

If you measure position, again, however, the probability to measure any $x'$ will be $|\langle x'\psi'' \rangle|^2 = |\langle x'|p\rangle|^2$, which means now that you could find yourself at any $x'$ independent of your original $x$. And now if you measure position over and over again you will get over and over again the same value $x'$.

And so on.

The reason all this happens is because it is impossible to have a state $|\Psi\rangle$ that is a simultaneous eigenstate of both the position operator $\hat{x}$ and the momentum operator $\hat{p}$. And it is impossible because $\hat{x}$ and $\hat{p}$ are not simultaneously diagonalizable: there is no basis in which both $\hat{x}$ and $\hat{p}$ are both diagonal.

How do we know this? Because the commutator of $\hat{x}$ and $\hat{p}$:

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar$$

is nonzero. (You prove: if two operators are both diagonal, then they commute.) Although the commutator appears to be an abstract and even spurious construct, we’ll later see that commutator play a central role in the occupation representation, so pay attention!

Theorem (which I won’t prove here): If two operators commute (i.e., their commutator vanishes), then it is possible to have simultaneous eigenfunctions.

This leads to Noether’s theorem. Classically, if the Hamiltonian of a system is invariant under some operation, such as translation in time or space, or rotation, then there is some conserved quantity. So: if the Hamiltonian is constant (invariant) in time, then energy is conserved; if it is translationally...
invariant, then linear momentum is conserved; if rotationally invariant, then angular momentum conserved; and so on.

For the quantum version of Noether’s theorem, consider an operator \( \hat{A} \) which commutes with the Hamiltonian \( \hat{H} \). In that case, it is possible to have a state \( |E, a\rangle \) that is a simultaneous eigenstate of both \( \hat{H} \) and \( \hat{A} \). Then, even as this wavefunction evolves under the time-dependent Schrödinger equation, it will always remain 100% with eigenvalue \( a \). Often we call a conserved quantity such as \( a \) a quantum number. This is especially true when it is indexed by a simple number such as integers. The examples that will concern us most are parity \( \pi \), total angular momentum \( J \), and the \( z \)-component of angular momentum \( J_z = M \).

Some quantum numbers are only partially conserved. For example, in nuclear physics, isospin is approximately conserved, but is broken at the level of a few percent. Parity is broken at only a part in a million. If a quantum number is only partially conserved, it is often still a useful concept.

1.6 Quantum mechanics in three dimensions

While one can easily write down a Hamiltonian for one particle in three dimensions,
\[
\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(x, y, z),
\]
in general the differential equations are coupled and intractable.

The only cases we can solve analytically are those where we can–you may have guessed it–reduce the problem to a set of three one-dimensional problems. These cases are of particular importance because they will provide important examples, or paradigms, for the occupation representation.

Let’s start with a special case: the spherically symmetric harmonic oscillator. In cartesian coordinates, the Hamiltonian is
\[
\hat{H} = -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \frac{1}{2} \omega^2 \left( x^2 + y^2 + z^2 \right)
\]
which we rearrange as
\[
\left( \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} \omega^2 x^2 \right) + \left( \frac{\hbar^2}{2m} \frac{\partial^2}{\partial y^2} + \frac{1}{2} \omega^2 y^2 \right) + \left( \frac{\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + \frac{1}{2} \omega^2 z^2 \right) = \hat{H}_x + \hat{H}_y + \hat{H}_z.
\]
Now we invoke separation of variables and assume a product wavefunction:
\[
\Psi(x, y, z) = \psi_1(x)\psi_2(y)\psi_3(z).
\]
Separation of variables means that we reduce the original problem in three variables
\[
\left( \hat{H}_x + \hat{H}_y + \hat{H}_z \right) \psi_1(x)\psi_2(y)\psi_3(z) = E\psi_1(x)\psi_2(y)\psi_3(z)
\]
to three independent problems each with one variable:

\[
\begin{align*}
\hat{H}_x \psi_1(x) &= E_x \psi_1(x); \\
\hat{H}_y \psi_2(y) &= E_y \psi_2(y); \\
\hat{H}_z \psi_3(z) &= E_z \psi_3(z).
\end{align*}
\tag{1.61}
\]

Each of \(\psi_1, \psi_2, \psi_3\) are the well-known 1-D harmonic oscillator eigensolutions, and the energies \(E_x = \hbar \omega (n_x + 1/2)\), etc, so that the total energy is \(E = \hbar \omega (N + 3/2)\), \(N = n_x + n_y + n_z = 0, 1, 2, \ldots\). What we want to take away from this is not the particular solution, however, but how the 3-D wavefunction is a product of one-variable functions.

Luckily for us, the same principle continues to hold for the general case when the potential is rotationally invariant, that is, when we go from cartesian \(x, y, z\) coordinates to spherical \(r, \theta, \phi\) coordinates and the potential \(V(r, \theta, \phi) = V(r)\) only. Although the details are slightly messier, and so I skip over them, we can apply separation of variables again.

Because we know what the answer will be, we write a separable \textit{ansatz} for the wavefunction in spherical coordinates:

\[
\psi(r, \theta, \phi) = R(r) Y_{\ell m}(\theta, \phi)
\tag{1.62}
\]

where \(Y_{\ell m}\) is of course the spherical harmonic. (For a marginally less brief discussion, see the Appendix.) Then

\[
\nabla^2 \psi = \nabla^2 R(r) Y_{\ell m}(\theta, \phi) = \left( \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} R(r) \right) Y_{\ell m}(\theta, \phi) - R(r) \frac{\ell (\ell + 1)}{r^2} Y_{\ell m}(\theta, \phi).
\tag{1.63}
\]

With this in hand we can factor out the \(Y_{\ell m}\) leaving the Schrödinger equation for the radial part,

\[
-\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} R(r) + \frac{\hbar^2}{2m} \frac{\ell (\ell + 1)}{r^2} R(r) + V(r) R(r) = E R(r),
\tag{1.64}
\]

or, if as usual, we make the substitution \(R(r) = u(r)/r\), we get the standard radial Schrödinger equation,

\[
-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} u(r) + \frac{\hbar^2}{2m} \frac{\ell (\ell + 1)}{r^2} u(r) + V(r) u(r) = E u(r).
\tag{1.65}
\]

We’ve successfully reduced a three-dimensional differential equation to a one-dimensional differential equation. Almost. There is a price to be paid: we must solve this equation separately for each value of \(\ell = 0, 1, 2, \ldots\), so what we have done is to replace a single three-dimensional differential equation with an infinite number of different one-dimensional differential equations.

For some analytic cases such as the hydrogen atoms this poses no problem, but for a general potential (1.65) must be solved numerically, and solved numerically for each value of \(\ell\). Many times, however, one is only interested in a few low values of \(\ell\) and so the price is a reasonable one to pay.
Sadly, we cannot always count on separation of variables to save us. Or, rather, a *single* product of one-variable (or even one-body) functions will not always yield exactly decoupled equations. Nonetheless nearly all of our approaches in many-body physics flow from this idea. We will either expand our solutions as *sums* of product wavefunctions, as we’ll see in configuration-interaction calculations, or else choose a single product wavefunction as an approximate solution, which leads to mean-field approaches such as the Hartree-Fock and Gross-Pitaevskii approximations. These mean-field equations are effectively one-body problems but are nonlinear and must be solved iteratively, a different kind of price to pay for simplification. We will discuss such approaches in detail later on.

### 1.7 The variational theorem

Most quantum mechanical problems cannot be solved exactly, even with powerful computers. Therefore we must turn to approximate solutions. The two most powerful tools for dealing with approximations are perturbation theory (which we will not use here) and the variational theorem.

The Raleigh-Ritz variational theorem is: *Let $\hat{H}$ be some Hamiltonian with ground state energy $E_0$. Then for any wavefunction $\Psi$,*

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0.$$  \hspace{1cm} (1.66)

This theorem is quite easy to prove. (Do so.) It is also extremely useful when trying to find an approximate solution. Even if you don’t know the ground state energy— or especially when you don’t know the ground state energy—you can place an upper bound on it. If you have two wavefunctions, $|\Psi\rangle$ with energy $E$ from (1.66) and $|\Psi'\rangle$ with energy $E'$, whichever energy is lower is unequivocably better and whichever wavefunction gives the lower energy is better.

We can sharpen up this discussion. Suppose you have a trial wavefunction, often called an *ansatz*, which depends on a parameter, $|\Psi(\alpha)\rangle$. Then compute

$$E(\alpha) = \frac{\langle \Psi(\alpha) | \hat{H} | \Psi(\alpha) \rangle}{\langle \Psi(\alpha) | \Psi(\alpha) \rangle}.$$  \hspace{1cm} (1.67)

The best choice of $\alpha$ minimizes the energy, which in turn leads to the variational condition

$$\frac{\partial}{\partial \alpha} E(\alpha) = 0.$$  \hspace{1cm} (1.68)

One can of course generalize to multiple parameters and this will allow us to derive the Hartree-Fock approximation.

To illustrate, let me take a couple of examples in one dimension. The first relates to one we have already addressed: the $\delta$-potential (1.48). We choose a gaussian *ansatz*:

$$\psi(x, \alpha) = \left(\frac{\alpha}{2\pi}\right)^{1/4} \exp \left( -\frac{1}{2} \alpha x^2 \right)$$  \hspace{1cm} (1.69)
which is always normalized. Then the trial energy $E(\alpha) = \langle \hat{H} \rangle$ is
\begin{equation}
E(\alpha) = \langle \hat{H} \rangle = \frac{\hbar^2 \alpha}{2m} - \lambda \sqrt{\frac{\alpha}{\pi}}. \tag{1.70}
\end{equation}

Next, finding the extremum
\begin{equation}
\frac{\partial E(\alpha)}{\partial \alpha} = \frac{\hbar^2}{4m} - \frac{\lambda}{2} \frac{1}{\sqrt{\pi} \alpha} = 0 \tag{1.71}
\end{equation}
one gets
\begin{equation}
\alpha_{\text{min}} = \frac{4\lambda^2 m^2}{\hbar^4 \pi} \tag{1.72}
\end{equation}
and
\begin{equation}
E(\alpha_{\text{min}}) = -\frac{m \lambda^2}{\pi \hbar^2} \tag{1.73}
\end{equation}
which is slightly greater than the exact ground state energy $-\frac{m \lambda^2}{2\hbar^2}$. Furthermore, the exact variance (square of the width) of the ground state is $\langle x^2 \rangle_{\text{exact}} = \frac{\hbar^4}{2m^2 \lambda^2}$, but from our variational solution the variance is
\begin{equation}
\langle x^2 \rangle = \frac{1}{2\alpha_{\text{min}}} = \frac{\pi \hbar^2}{8m^2 \lambda^2} = \frac{\pi}{4} \langle x^2 \rangle_{\text{exact}}. \tag{1.74}
\end{equation}

For a second example, consider a particle in one dimension with a quartic potential:
\begin{equation}
\hat{H} = \frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{6} \lambda x^4. \tag{1.75}
\end{equation}
This equation can be easily solved in coordinate space numerically, but we will tackle it variationally.

We assume a gaussian ansatz (1.69) and compute
\begin{equation}
E(\alpha) = \frac{\hbar^2 \alpha}{4m} + \lambda \frac{1}{8 \alpha^2}. \tag{1.76}
\end{equation}
Then taking the first derivative
\begin{equation}
\frac{\partial}{\partial \alpha} E(\alpha) = \frac{\hbar^2}{4m} - \lambda \frac{1}{4\alpha^3} = 0 \tag{1.77}
\end{equation}
we find the extremum at a value
\begin{equation}
\alpha_{\text{min}} = \left( \frac{\lambda m}{\hbar^2} \right)^{1/3} \tag{1.78}
\end{equation}
and then
\begin{equation}
\frac{3}{8} \left( \frac{\hbar^4 \lambda}{m^2} \right)^{1/3}. \tag{1.79}
\end{equation}
1.8. SUMMARY

Numerically, the actually ground energy is $0.3676 \hbar^{4/3} \lambda^{1/3} / m^{2/3}$. Furthermore, we can compute the ground state variance as

$$\langle x^2 \rangle = 0.5 \frac{\hbar^{2/3}}{\lambda^{1/3} m^{1/3}},$$

while the numerically exact value is $0.522 \hbar^{2/3} \lambda^{1/3} m^{1/3}$.

We will return to both the $\delta$-potential and the quartic potential in Chapter 3.

1.8 Summary

What you should get from this chapter:

- The theory of quantum mechanics is worked out in a linear vector space; what we call wavefunctions should be thought of as abstract vectors, $|\psi\rangle$, where Dirac’s bra-ket notation is a useful convention, particularly for inner products.
- The usual coordinate-space representation of a wavefunction is found by projecting onto a basis, $\psi(x) = \langle x|\psi\rangle$. But projection onto any basis is just as valid. This includes projection onto a linear function space spanned by (typically orthonormal) basis functions, $|\phi_n\rangle$, so that $\psi = \langle \phi_n | \psi \rangle$.
- Projection of vectors onto a discrete basis allows one to develop the matrix formulation of quantum mechanics. Solving the time-independent Schrödinger equation is now equivalent to diagonalizing (finding the eigenvalues and eigenvectors) of a matrix. Other important quantities, such as expectation values and transition amplitudes, can also be recast in terms of matrix elements. The computation of matrix elements is a fundamental topic in many-body theory.
- The variational principle is a useful tool for approximation in quantum mechanics.
- Another useful concept in quantum mechanics is, whenever possible, to break a multi-dimensional problem into separate one-dimensional problems, for example using a product ansatz. Sometimes, for example in three-dimensional quantum mechanics with a rotationally invariant potential, this can be done exactly. In future work we won’t be able to do this exactly, but we will still find product wavefunctions a powerful guiding principle.

1.9 Exercises

(1.1) Prove for any two operators $\hat{A}, \hat{B}$ that trace ($\hat{A}\hat{B}$) = trace ($\hat{B}\hat{A}$).

(1.2) Prove that the trace of an operator is the sum of its eigenvalues and the determinant of an operator is the product of its eigenvalues. (Assume the operator is, if not Hermitian, diagonalizable).

(1.3) Show that it is plausible, at least, that

$$\text{tr} \ln \hat{A} = \ln \det \hat{A}.$$  

(1.81)
CHAPTER 1. A REVIEW OF QUANTUM MECHANICS

Hint: assume \( \hat{A} \) is diagonalizable. In an advanced chapter we'll have use for this identity.

(1.4) Derive the matrix eigenvalue equation from a variational principle. Hint: For an operator \( \hat{A} \), we want to find a vector \( |\psi\rangle \) that is an extremum. To do so, we expand in an orthonormal basis \( \{|n\rangle\} \), that is, \( |\psi\rangle = \sum_n v_n |n\rangle \). Assume the \( v_n \) are real. Compute \( \langle \psi | \hat{H} | \psi \rangle \) in terms of the components \( v_n \) and the matrix elements \( H_{mn} = \langle m | H | n \rangle \). Note that even though the basis is orthonormal, we cannot assume that \( |\psi\rangle \) is normalized. Therefore we cannot just minimize \( \langle \psi | \hat{H} | \psi \rangle \); instead we have to add \( \langle \psi | \psi \rangle = 1 \) as a constraint. The way we do this is through Lagrange multipliers (consult a suitable mathematical methods text if you are unfamiliar with this concept) and so minimize \( \langle \psi | \hat{H} | \psi \rangle - E \langle \psi | \psi \rangle \) with respect to \( v_n \). You should regain the matrix eigenvalue equation with the Lagrange multiplier \( E \) now becoming the energy. We will use a similar method later on deriving the Hartree-Fock equations.

(1.5) Now consider an nonorthogonal basis, that is a basis \( \{|\mu\rangle\} \) such that \( \langle \mu | \nu \rangle = S_{\mu \nu} \) which is not the unit matrix. As long as the basis space spans the space \( S \) will be invertible. Now let's tackle some related problems:

(a) Prove that all the eigenvalues of \( S \) are real and positive, that is, \( S \) is positive definite. To do this you only have to show that for any set of components \( v_\mu \), that

\[
\sum_{\mu \nu} v_\mu^* S_{\mu \nu} v_\nu > 0.
\]

(b) Suppose we want to expand a vector in this basis, \( |\psi\rangle = \sum_\nu v_\nu |\nu\rangle \). Deduce how to find the coefficients \( v_\nu \) from the overlaps \( \langle \mu | \psi \rangle \).

(c) To find the eigenvalues and eigenstates in this nonorthogonal basis, derive the generalized eigenvalue equation

\[
\sum_\nu H_{\mu \nu} v_\nu = E \sum_\nu S_{\mu \nu} v_\nu \quad (1.82)
\]

(1.6) Consider the \( \delta \)-potential, with the Hamiltonian matrix elements given in (1.52). Show that for \( m = n = 0 \), this is minimized by \( L_0 = \pi^2/2 \).

(1.7) For the 3-D simple harmonic oscillator, the radial Schrödinger equation is

\[
\left( -\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2 \ell (\ell + 1)}{2mr^2} + \frac{1}{2}m\omega^2 r^2 \right) u(r) = Eu(r). \quad (1.83)
\]

It is well known that, for a given angular momentum \( \ell \) the lower energy is \( (\ell + 3/2)\hbar\omega \). Now make a variational estimate, using the exponential ansatz

\[
\psi(r) = \sqrt{\alpha^3/2r} \exp \left( -\frac{\alpha}{2r} \right). \quad (1.84)
\]

This are normalized by

\[
\int_0^\infty u^2(r) dr = 1, \quad (1.85)
\]
though of course you should check this yourself. Use this to find an upper bounds for the lowest energy for each $\ell$.

For further work, try

$$\psi(r) = C r^{\ell+1} \exp\left(-\frac{a}{2} r\right),$$

(1.86)

for angular momentum $\ell$; normalize and find the variational energy. Is it better than (1.84)?

(1.8) Diagonalize the simple harmonic oscillator Hamiltonian,

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2,$$

in the same basis we used for the $\delta$-potential in section 1.4, that is, cosines and sines between $x = -L/2$ and $x = L/2$. For even functions this is $|n\rangle = \sqrt{2/L} \cos((2n+1)\pi x/L), n = 0, 1, 2, \ldots$. The matrix elements for the kinetic energy are easy, while the matrix elements $\langle m | x^2 | n \rangle$ are tractable but more complicated integrals. As we did for the $\delta$-potential, I recommend you rewrite $L = L_0 b$, where $b = \sqrt{\hbar/m\omega}$ is the oscillator length and $L_0$ is a dimensionless parameter. What value of $L_0$ gives you the minimum for the ground state energy, using just the $n = 0$ basis state? Now diagonalize numerically, truncating at some $n_{\text{max}}$, and plot the converage of the ground state energy as a function of the truncation. Try this for different values of $L_0$. As we did for the $\delta$-potential, you should find that for larger $L_0$, the convergence is slower, but ultimately yields a better approximation to the ‘exact’ answer.

Look at the excited state energies as well. Try a scatter plot: the numerical energies versus the exact harmonic oscillator energies. You will get almost a straight line for a while, then a deviation. What happens at the deviation? Later in Chapter 3 we’ll find that the rms width of the $N$th excited state is $b\sqrt{N} + 1/2$, so it makes sense that when the state starts touching the wall, that is when $L_0/2 \sim \sqrt{N}$, you’ll get distortion, although keep in mind this is not an equality. Look at different values of $L_0$ and see if you have evidence for this behavior.

While you’re at it, take the eigenvectors from your diagonalization and compute $\langle x^2 \rangle$ for the ground state and for various excited states. The exact answer is $b^2 (N + 1/2)$.

Keep in mind the above was just for even functions; you could also do odd functions using $\sqrt{2/L} \sin(2n\pi x/L), n = 1, 2, \ldots$. 