

Spring semester, 2009 (revised)
Oberlin College Physics 412

Notes for

Applied Quantum Mechanics

Dan Styer, Schiffer Professor of Physics

Contents

1	Introduction	1
1.1	Problems	1
2	Atoms	6
2.1	Many-particle systems in quantum mechanics	6
2.2	An antisymmetric basis for the helium problem	16
2.3	Ground state energy of helium	20
3	Molecules	24
3.1	The hydrogen molecule ion	24
3.2	Problems	27
4	Perturbation Theory for the Time Development Problem	29
4.1	Setup	29
4.2	Fermi's golden rule	33
4.3	Problems	35
5	Quantization of the Electromagnetic Field	36
5.1	Overview	36
5.2	Classical electromagnetism	37
5.3	Quantal electromagnetism — What is a photon?	42
A	The Spherical Harmonics	44

B Radial Wavefunctions for the Coulomb Problem	46
C Cheat Sheet	47

Chapter 1

Introduction

1.1 Problems

1.1 Quantal recurrence in the infinite square well

- Find the period as a function of energy for a classical particle of mass m in an infinite square well of width L .
- Show that any wavefunction, regardless of energy, in the same infinite square well is periodic in time with a period

$$\frac{4mL^2}{\hbar\pi}.$$

(This part can be solved knowing only the energy eigenvalues.)

- What happens after one-half of this time has passed? (This part requires some knowledge of the energy eigenfunctions.)

[*Note:* This problem raises deep and as yet unresolved questions about the classical limit of quantum mechanics. See D.F. Styer, “Quantum revivals versus classical periodicity in the infinite square well,” *American Journal of Physics* **69** (January 2001) 56–62.]

1.2 Quantal recurrence in the Coulomb problem

Show that in the Coulomb problem, any quantal state consisting of a superposition of two or more bound energy eigenstates with principal quantum numbers n_1, n_2, \dots, n_r evolves in time with a period of

$$\frac{h}{\text{Ry}}N^2,$$

where Ry is the Rydberg constant and the integer N is the least common multiple of n_1, n_2, \dots, n_r .

1.3 Atomic units

The Schrödinger equation for the Coulomb problem is

$$i\hbar \frac{\partial \Psi(x, y, z, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi(x, y, z, t) - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \Psi(x, y, z, t).$$

It is clear that the answer to any physical problem can depend only on the three parameters \hbar , m , and $e^2/4\pi\epsilon_0$. In class, we used these ideas to show that any problem that asked for a length had to have an answer which was a dimensionless number times the characteristic length, the so-called Bohr radius

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{e^2 m}.$$

- Show that there is only one characteristic energy, i.e. only one way to combine the three parameters to produce a quantity with the dimensions of energy. (In class, we found *one* way to perform this combination, but I want you to prove that this is the *only* way. *Hint*: Instead of the conventional basic dimensions of length, mass, and time, use the unconventional basic dimensions of length, mass, and energy.)
- Find the characteristic time τ_0 . What is its numerical value in terms of femtoseconds?
- Bonus*: Show that, in the Bohr model, the period of the innermost orbit is $2\pi\tau_0$. What is the period of the n th orbit?
- Estimate the number of heartbeats made in a lifetime by a typical person. If each Bohr model orbit corresponds to a heartbeat, how many “lifetimes of hydrogen” pass in a second?
- Write the time-dependent Schrödinger equation in terms of the scaled variables

$$\tilde{r} = \frac{r}{a_0} \quad \text{“lengths measured in atomic units”}$$

and

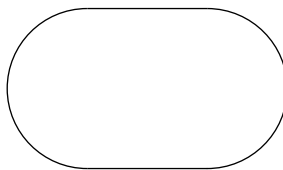
$$\tilde{t} = \frac{t}{\tau_0} \quad \text{“time measured in atomic units”}.$$

Be sure to use the dimensionless wavefunction

$$\tilde{\Psi}(\tilde{x}, \tilde{y}, \tilde{z}, \tilde{t}) = (a_0)^{3/2} \Psi(x, y, z, t).$$

1.4 Scaling in the stadium problem

The “stadium” problem is often used as a model chaotic system, in both classical and quantum mechanics. [See E.J. Heller, “Bound-State Eigenfunctions of Classically Chaotic Hamiltonian Systems: Scars of Periodic Orbits” *Phys. Rev. Lett.*, **53**, 1515 (1984); S. Tomsovic and E.J. Heller, “Long-Time Semiclassical Dynamics of Chaos: The Stadium Billiard” *Phys. Rev. E*, **47**, 282 (1993); E.J. Heller and S. Tomsovic, “Postmodern Quantum Mechanics” *Physics Today*, **46**, 38 (1993).] This is a two-dimensional infinite well shaped as a rectangle with semi-circular caps on opposite ends. Suppose one stadium has the same shape but is exactly three times as large as another. Show that in the larger stadium, wavepackets move just as they do in the smaller stadium, but nine times more slowly. (The initial wavepacket is of course also scaled up three times.) And show that the energy eigenvalues of the larger stadium are one-ninth the energy eigenvalues of the smaller stadium.



1.5 Solving the Coulomb problem through operator factorization

Griffiths (section 4.2) finds the bound state energy eigenvalues for the Coulomb problem using power series solutions of the Schrödinger equation. Here is another way, based on operator factorization (ladder operators). In atomic units, the radial wave equation is

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{2r^2} - \frac{1}{r} \right] u_{n,\ell}(r) \equiv h_\ell u_{n,\ell}(r) = \epsilon_{n,\ell} u_{n,\ell}(r)$$

where $u_{n,\ell}(r)$ is r times the radial wavefunction. Introduce the operators

$$D_\pm^{(\ell)} \equiv \frac{d}{dr} \mp \frac{\ell}{r} \pm \frac{1}{\ell}.$$

a. Show that

$$D_+^{(\ell)} D_-^{(\ell)} = -2h_\ell - \frac{1}{\ell^2}.$$

and that

$$D_-^{(\ell+1)} D_+^{(\ell+1)} = -2h_\ell - \frac{1}{(\ell+1)^2}$$

b. Conclude that

$$h_{\ell+1} D_+^{(\ell+1)} = D_+^{(\ell+1)} h_\ell,$$

and apply this operator equation to $u_{n,\ell}(r)$ to show that

$$D_+^{(\ell+1)} u_{n,\ell}(r) \propto u_{n,\ell+1}(r)$$

and that $\epsilon_{n,\ell}$ is independent of ℓ .

c. Argue that for every $\epsilon_{n,\ell} < 0$ there is a maximum ℓ . (Hint: Examine the effective potential for radial motion.) Call this ℓ value ℓ_n .

d. Define $n = \ell_n + 1$ and show that

$$\epsilon_{n,\ell} = -\frac{1}{2n^2} \quad \text{where} \quad \ell = 0, \dots, n-1.$$

(One can also continue this game to find the energy eigenfunctions.)

1.6 Kramers' relation

Kramers' relation states that for any energy eigenstate $\eta_{n\ell m}(\vec{r})$ of the Coulomb problem, the expected values of r^s , r^{s-1} , and r^{s-2} are related through

$$\frac{s+1}{n^2} \langle r^s \rangle - (2s+1) a_0 \langle r^{s-1} \rangle + \frac{s}{4} [(2\ell+1)^2 - s^2] a_0^2 \langle r^{s-2} \rangle = 0.$$

- a. Prove Kramers' relation. *Hints:* Use atomic units. Start with the radial equation in form

$$u''(r) = \left[\frac{\ell(\ell+1)}{r^2} - \frac{2}{r} + \frac{1}{n^2} \right] u(r),$$

and use it to express $\int(u(r)r^s u''(r)) dr$ in terms of $\langle r^s \rangle$, $\langle r^{s-1} \rangle$, and $\langle r^{s-2} \rangle$. Then integrate $\int(u(r)r^s u''(r)) dr$ by parts to find an integral involving $u'(r)$ as the highest derivative. Show that $\int(u(r)r^s u'(r)) dr = -[s/2]\langle r^{s-1} \rangle$ and that $\int(u'(r)r^s u'(r)) dr = -[2/(s+1)] \int(u''(r)r^{s+1} u'(r)) dr$.

- b. Use Kramers' relation with $s = 0$, $s = 1$, $s = 2$, and $s = 3$ to find formulas for $\langle r^{-1} \rangle$, $\langle r \rangle$, $\langle r^2 \rangle$, and $\langle r^3 \rangle$. Note that you could continue indefinitely to find $\langle r^s \rangle$ for any positive power.
- c. However, you can't use this chain to work downward. Try it for $s = -1$, and show that you get a relation between $\langle r^{-2} \rangle$ and $\langle r^{-3} \rangle$, but not either quantity by itself.

1.7 The Stark effect

This problem is based on Griffiths problems 6.36 and 6.37 and uses the notation and situation established there.

- a. In class we applied degenerate perturbation theory to the four $n = 2$ states, and found that all but two elements of the relevant 4×4 matrix were zero, and that those two remaining elements were equal. Find the value of these two elements.
- b. The "good" energy eigenstates for the $n = 2$ Stark effect are $|2, 1, +1\rangle$, $|2, 1, -1\rangle$, $(|2, 0, 0\rangle + |2, 1, 0\rangle)/\sqrt{2}$, and $(|2, 0, 0\rangle - |2, 1, 0\rangle)/\sqrt{2}$. Find the expectation value of the electric dipole moment ($\mathbf{p} = -e\mathbf{r}$) in each of these states.
- c. The expectation value of the dipole moment is zero in state $|2, 0, 0\rangle$ and zero in state $|2, 1, 0\rangle$, yet it is non-zero in state $(|2, 0, 0\rangle + |2, 1, 0\rangle)/\sqrt{2}$. At first glance, this might seem like a contradiction. What great principle of quantum mechanics allows this fact to escape the curse of contradiction? (Answer in one sentence.)
- d. (**Bonus.**) Describe these four states qualitatively and explain why they are the "good" states for use in the Stark effect.
- e. Consider the Stark effect for the $n = 3$ states of hydrogen. There are initially nine degenerate states. Construct a 9×9 matrix representing the perturbing Hamiltonian. (*Hint:* Use a selection rule to determine the best order of basis elements before actually working any integrals.)
- f. Find the eigenvalues and degeneracies.

1.8 Bonus

In the previous problem, on the Stark effect, we had to calculate a lot of matrix elements of the form

$$\int_0^\infty r^2 R_{n,\ell}(r) r R_{n',\ell'}(r) dr.$$

This was possible but (to put it mildly) tedious. Can you think of some easy way to do integrals of this form? Could the operator factorization technique (problem 1.4) give us any assistance? Can you derive any inspiration from our proof of Kramers' relation (problem 1.5)?

1.9 Variational principle for the harmonic oscillator

Find the best bound on the ground state energy of the one-dimensional harmonic oscillator using a trial wavefunction of form

$$\psi(x) = \frac{A}{x^2 + b^2},$$

where A is determined through normalization and b is an adjustable parameter. *Hint:* Put the integrals within $\langle H \rangle$ into dimensionless form so that they are independent of A and b , and are “just numbers”: call them C_K and C_P . Solve the problem in terms of these numbers, then evaluate the integrals only at the end.

Chapter 2

Atoms

2.1 Many-particle systems in quantum mechanics

One particle moves in one dimension. (Ignore spin.) How can we represent this system's state?

There are several ways: The ordinary wavefunction $\psi(x)$ represents the state in terms of the position basis. The momentum wavefunction

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx e^{-i(p/\hbar)x} \psi(x)$$

represents the state in terms of the momentum basis. The energy expansion coefficients

$$c_n = \int_{-\infty}^{+\infty} \eta_n^*(x) \psi(x) dx$$

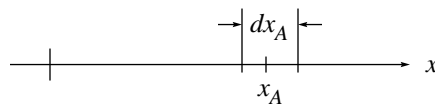
represent the state in terms of the energy basis. [Meaning that

$$\psi(x) = \sum_n c_n \eta_n(x). \quad]$$

Or we can represent the state in terms of the expectation of position $\langle x \rangle$, the expectation of momentum $\langle p \rangle$, the indeterminacy in position (which involves $\langle x^2 \rangle$), the indeterminacy in momentum (which involves $\langle p^2 \rangle$), the moments $\langle x^3 \rangle$ and $\langle p^3 \rangle$ and so forth, the correlation functions $\langle xp \rangle$ and $\langle xp^2 x^3 p \rangle$ and so forth. You can prove (it's not easy!) that if all these mean values are known then one can reconstruct the wavefunction.

Suppose we know the ordinary position wavefunction. Then if you measure the particle's location, the probability of finding it in a window of width dx_A about position x_A is $|\psi(x_A)|^2 dx_A$. This is *not* sufficient information to specify the particle's state: it tells you everything there is to know about position, but nothing about momentum or about position-momentum correlation functions.

One-particle window



Meanwhile the *amplitude* of finding the particle in this window is $\psi(x_A) \sqrt{dx_A}$. If you know the amplitude at every point x_A , then you *do* have full information about the state.

The normalization is of course

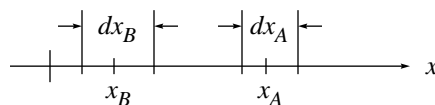
$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1.$$

Variations: If a spin-zero particle moves in three dimensions, the wavefunction is $\psi(x, y, z)$. If a spin-half particle ($s_z = \pm\frac{1}{2}$) moves in three dimensions, the wavefunction is $\psi(x, y, z, s_z)$, or $\tilde{\psi}(p_x, p_y, p_z, s_x)$. In general, when I say things like “the variable x ”, you will have to generalize in different circumstances to, for example, “the variables p_x, p_y, p_z, s_x ”.

Two particles, say an electron and a neutron, move in one dimension. (Ignore spin.) How can we represent this system’s state?

There is now a wavefunction $\psi(x_A, x_B)$ with the interpretation that if you measure the location of both particles, then the probability of finding the electron in a window of width dx_A about position x_A , and the neutron in a window of width dx_B about position x_B , is $|\psi(x_A, x_B)|^2 dx_A dx_B$.

Two-particle windows



Note that the letters A and B refer to two different positions, not two different particles. The particles are represented by the sequence of arguments: the first argument pertains to the electron, the second argument pertains to the neutron.

I particularly emphasize that the wavefunction applying to the system “electron plus neutron” is *one* function of *two* variables, not *two* functions each of *one* variable:

$$\psi_{\text{electron}}(x_A)\psi_{\text{neutron}}(x_B) \quad \mathbf{NO!}$$

The difference feeds directly into this question: How many (real) numbers does it take to specify a state? In classical mechanics, the answer is straightforward. The state of any single particle is specified through two numbers: the position and momentum of that particle. The state of a collection of several particles is specified through the state of each particle. In summary

particles	real numbers needed to specify classical state
1	2
2	4
3	6
\vdots	\vdots
N	$2N$

In quantum mechanics, the answer is more subtle. To specify the state of even a single particle, one must give the wavefunction $\psi(x)$... an infinite number of complex numbers! For concreteness suppose we approximate this function on a computer, using a grid of 100 points. Then we need 100 complex numbers, that is $2(100)$ real numbers. But one of these numbers is fixed through the normalization condition, and one is an overall phase that can be set arbitrarily. The end result is that to specify a single-particle wavefunction to this degree of accuracy requires $2(100) - 2 = 198$ real numbers.

What about two particles? Now we have a wavefunction on a grid of 100×100 points, so specifying a two-particle wavefunction to this degree of accuracy requires $2(100)^2 - 2$ real numbers. This number (19998) is *much larger* than twice 198. To specify the two-particle states, we cannot get away with just specifying two one-particle states. Just as a particle might not have a location, so in a two-particle system an individual particle might not have a state.

In summary

particles	real numbers needed to specify quantal state
1	$2(100) - 2 = 198$
2	$2(100)^2 - 2 = 19998$
3	$2(100)^3 - 2 = 1999998$
\vdots	\vdots
N	$2(100)^N - 2$

Much of the spectacular richness and complexity of the quantum world arises from this rapid increase of information with particle number. (Design of quantum computer.)

Two identical particles, say two neutrons, move in one dimension. (Ignore spin.) How can we represent this system's state?

Of course, there is a wavefunction $\psi(x_A, x_B)$, but the interpretation is somewhat different. The question is *not* "What is the probability of finding neutron α within window A and neutron β within window B ?" These neutrons are identical, so there is no such thing as "neutron α " or "neutron β ." The question instead is "What is the probability of finding a neutron within window A and a neutron within window B ?" The answer to this question is

$$\begin{cases} 2|\psi(x_A, x_B)|^2 dx_A dx_B & \text{if the windows don't overlap} \\ |\psi(x_A, x_B)|^2 dx_A dx_B & \text{if } x_B = x_A \end{cases}$$

The normalization condition is

$$2 \int_{-\infty}^{+\infty} dx_A \int_{-\infty}^{x_A} dx_B |\psi(x_A, x_B)|^2 = 1 \quad \text{or} \quad \int_{-\infty}^{+\infty} dx_A \int_{-\infty}^{\infty} dx_B |\psi(x_A, x_B)|^2 = 1.$$

If the two particles are identical, then it's certainly true that

$$|\psi(x_A, x_B)|^2 = |\psi(x_B, x_A)|^2.$$

But this condition insures only that the position probabilities are unaffected if you swap the windows. If the two particles are identical, then the same holds for momentum probabilities. In other words, the wavefunctions

$$\psi(x_A, x_B) \text{ and } \psi(x_B, x_A)$$

represent the same state, so

$$\psi(x_A, x_B) = e^{i\delta} \psi(x_B, x_A),$$

where δ is a number, not a function of x_A or x_B . Thus, for example,

$$\psi(5, 7) = e^{i\delta} \psi(7, 5).$$

But

$$\psi(7, 5) = e^{i\delta} \psi(5, 7),$$

so

$$\psi(5, 7) = (e^{i\delta})^2 \psi(5, 7).$$

We conclude that

$$e^{i\delta} = \pm 1.$$

In other words, when the wavefunction swaps arguments, it either remains the same or changes sign. In the first case, the wavefunction is called “symmetric under exchange,” (or “under swapping,” or “under interchange”)¹ in the second, “antisymmetric.”

What if there are **three identical particles**? The wavefunction is $\psi(x_A, x_B, x_C)$ and you can swap either the first and second arguments, or the second and third arguments, or the first and third arguments. The arguments of the next three paragraphs will show that the wavefunction must be either symmetric under each of these three interchanges or else antisymmetric under each of these three interchanges.

After any swapping, you must produce a wavefunction representing the same state, so any swapping can introduce at most a constant phase factor. Thus

$$\begin{aligned} \psi(x_A, x_B, x_C) &= e^{i\alpha} \psi(x_B, x_A, x_C) \\ &= e^{i\beta} \psi(x_A, x_C, x_B) \\ &= e^{i\gamma} \psi(x_C, x_B, x_A) \end{aligned}$$

¹I prefer “swap” to emphasize that we’re swapping mathematical windows, not exchanging physical particles, but the most commonly used term is “exchange.”

The “double swap” argument above shows that $e^{i\alpha}$ is either +1 or -1 , that $e^{i\beta}$ is either +1 or -1 , and that $e^{i\gamma}$ is either +1 or -1 . We can gain more information through repeated swappings that return ultimately to the initial sequence. For example

$$\begin{aligned}\psi(x_A, x_B, x_C) &= e^{i\alpha}\psi(x_B, x_A, x_C) && \text{[[swapping first and second arguments]]} \\ &= e^{i\alpha}e^{i\beta}\psi(x_B, x_C, x_A) && \text{[[swapping second and third arguments]]} \\ &= e^{i\alpha}e^{i\beta}e^{i\gamma}\psi(x_A, x_C, x_B) && \text{[[swapping first and third arguments]]} \\ &= e^{i\alpha}e^{i\beta}e^{i\gamma}e^{i\beta}\psi(x_A, x_B, x_C) && \text{[[swapping second and third arguments]]}\end{aligned}$$

We already know that $(e^{i\beta})^2 = 1$, so this argument reveals that $e^{i\alpha}e^{i\gamma} = 1$, i.e., these two phase factors are either both +1 or both -1 .

Further arguments of this type will convince you that the three phase factors must either be all +1 or else all -1 . For suppose that

$$\begin{aligned}\psi(x_A, x_B, x_C) &= -\psi(x_B, x_A, x_C) \\ &= +\psi(x_A, x_C, x_B) \\ &= -\psi(x_C, x_B, x_A)\end{aligned}$$

(That is, antisymmetric under swaps of the first and second arguments or the first and third arguments, symmetric under swaps of the second and third arguments.) Then we can go from $\psi(x_A, x_B, x_C)$ to $\psi(x_B, x_C, x_A)$ via two different swapping routes:

$$\begin{aligned}\psi(x_A, x_B, x_C) &= (-1)\psi(x_B, x_A, x_C) && \text{[[swapping first and second arguments]]} \\ &= (-1)(+1)\psi(x_B, x_C, x_A) && \text{[[swapping second and third arguments]]}\end{aligned}$$

or

$$\begin{aligned}\psi(x_A, x_B, x_C) &= (-1)\psi(x_C, x_B, x_A) && \text{[[swapping first and third arguments]]} \\ &= (-1)(-1)\psi(x_B, x_C, x_A) && \text{[[swapping first and second arguments]]}\end{aligned}$$

The only function that satisfies both of conditions is $\psi(x_A, x_B, x_C) = 0$.

The other possible “mixed symmetric and antisymmetric” possibility is

$$\begin{aligned}\psi(x_A, x_B, x_C) &= +\psi(x_B, x_A, x_C) \\ &= -\psi(x_A, x_C, x_B) \\ &= +\psi(x_C, x_B, x_A)\end{aligned}$$

but this can be shown impossible by the “two route” argument of the previous paragraph.

- 2.1 **Problem:** Show that the same result applies for functions of four or more arguments by considering first swaps among the first, second, and third arguments; then swaps among the first, second, and fourth arguments; then swaps among the first, second, and fifth arguments; etc.

In conclusion, a wavefunction for any number of identical particles must be either “completely symmetric” (every swap introduces a phase factor of $+1$) or else “completely antisymmetric” (every swap introduces a phase factor of -1). This is called the “exchange symmetry” of the wavefunction.

- 2.2 **Problem:** If there are two particles, there is one possible swap. If there are three particles, there are three possible swaps. Show that for four particles there are six possible swaps and that for N particles there are $N(N - 1)/2$ possible swaps.
- 2.3 **Problem:** Show that the momentum wavefunction has the same interchange symmetry as the position wavefunction (i.e., symmetric or antisymmetric). How about the energy coefficients? (Exactly what does that last question mean?)
- 2.4 **Problem:** Show that exchange symmetry is conserved: If the system starts out in a symmetric state it will remain symmetric at all times in the future, and similarly for antisymmetric.

Given what we’ve said so far, I would guess that a collection of neutrons could start out in a symmetric state (in which case they would be in a symmetric state for all time) or else they could start out in an antisymmetric state (in which case they would be in an antisymmetric state for all time). In fact, however, this is not the case. A collection of neutrons is *always* in an antisymmetric state. Furthermore, if two neutrons that have never encountered each other are brought close together, then the system goes into an antisymmetric state. The exchange symmetry turns out to have nothing to do with history, and depends only on the type of particle. Neutrons, protons, electrons, carbon-13 nuclei, and sigma baryons are always in antisymmetric states — they are called *fermions*. Photons, alpha particles, carbon-12 nuclei, and pi mesons are always in symmetric states — they are called *bosons*.

Furthermore, all bosons have integral spin and all fermions have half-integral spin. There is a mathematical result in relativistic quantum field theory called “the spin-statistics theorem” that sheds some light on this astounding fact. (See *Pauli and the Spin-Statistics Theorem* by Ian Duck and E.C.G. Sudarshan, and the review of this book by A.S. Wightman in *Am. J. Phys.* **67** (August 1999) 742–746.)

Given their obvious importance, it makes sense to spend some time on the **mathematics of completely symmetric and completely antisymmetric functions**. Given a garden-variety two-variable “seed” function $f(x_A, x_B)$, we can build a symmetric function

$$f_S(x_A, x_B) = f(x_A, x_B) + f(x_B, x_A),$$

and an antisymmetric function

$$f_A(x_A, x_B) = f(x_A, x_B) - f(x_B, x_A).$$

Note that these built functions are not necessarily normalized.

A generalized process works for three-variable functions: The built functions are sums over all $3!$ permutations of arguments. The function

$$f_S(x_A, x_B, x_C) = f(x_A, x_B, x_C) + f(x_A, x_C, x_B) + f(x_C, x_A, x_B) + f(x_C, x_B, x_A) + f(x_B, x_C, x_A) + f(x_B, x_A, x_C)$$

is completely symmetric while the function

$$f_S(x_A, x_B, x_C) = f(x_A, x_B, x_C) - f(x_A, x_C, x_B) + f(x_C, x_A, x_B) - f(x_C, x_B, x_A) + f(x_B, x_C, x_A) - f(x_B, x_A, x_C)$$

is completely antisymmetric. This process of building an symmetric function f_S from arbitrary seed function f is called “symmetrization”. Similarly for “antisymmetrization”. If the function is a wavefunction, the (anti)symmetrization process is usually understood to include also normalizing the resulting wavefunction.

2.5 Problem: If the seed $f(x_A, x_B, x_C)$ happens to be symmetric to begin with, what are the symmetrized and antisymmetrized functions? What if the seed happens to be antisymmetric to begin with?

2.6 Problem: Show that any two-variable function can be represented as a sum of a symmetric and an antisymmetric function. Can any three-variable function be represented as a sum of a completely symmetric and a completely antisymmetric function?

Symmetric and antisymmetric bases

If we want to study identical particles, we’ll need to build a basis of symmetric states and a basis of antisymmetric states. Here’s how.

Start with a single particle subject to a potential, and solve the energy eigenproblem. Suppose the results are an energy basis of

$$\begin{array}{ll} \eta_1(x) & \epsilon_1 \\ \eta_2(x) & \epsilon_2 \\ \vdots & \vdots \\ \eta_M(x) & \epsilon_M \end{array}$$

In most cases the number of one-particle energy eigenstates M is infinite, but it’s useful to keep that number as a variable anyway. There might or might not be some degeneracies in the system. . . it doesn’t matter.

Three non-identical particles. Now suppose there is not one, but three particles moving in this potential, and that they’re not identical. We can build a basis of product wavefunctions.

$\eta_1(x_A)\eta_1(x_B)\eta_1(x_C)$	$ 1, 1, 1\rangle$	$\epsilon_1 + \epsilon_1 + \epsilon_1$
$\eta_1(x_A)\eta_1(x_B)\eta_2(x_C)$	$ 1, 1, 2\rangle$	$\epsilon_1 + \epsilon_1 + \epsilon_2$
$\eta_1(x_A)\eta_2(x_B)\eta_1(x_C)$	$ 1, 2, 1\rangle$	$\epsilon_1 + \epsilon_2 + \epsilon_1$
$\eta_2(x_A)\eta_1(x_B)\eta_1(x_C)$	$ 2, 1, 1\rangle$	$\epsilon_2 + \epsilon_1 + \epsilon_1$
\vdots	\vdots	\vdots
$\eta_1(x_A)\eta_7(x_B)\eta_3(x_C)$	$ 1, 7, 3\rangle$	$\epsilon_1 + \epsilon_7 + \epsilon_3$
$\eta_7(x_A)\eta_3(x_B)\eta_1(x_C)$	$ 7, 3, 1\rangle$	$\epsilon_7 + \epsilon_3 + \epsilon_1$
\vdots	\vdots	\vdots
$\eta_M(x_A)\eta_M(x_B)\eta_M(x_C)$	$ M, M, M\rangle$	$\epsilon_M + \epsilon_M + \epsilon_M$

A few remarks: (1) There are M^3 elements in the basis. (2) We have a basis of product wavefunctions, but that doesn't mean that every state is a product state, because an arbitrary state is a sum of basis elements. (3) It's tiring to write always the form in the left column so we abbreviate it through the form in the center column. (4) If the three particles don't interact, then this is an energy basis with the eigenvalues shown. But even if they do interact, it's a basis. (5) If the particles don't interact, then there is necessarily degeneracy in this basis. (6) To keep in mind the distinction between this basis for the three-particle system and the basis for the one-particle system from which it is built, we often call the three-particle basis elements "states" and the one-particle basis elements "levels". The levels are the building blocks out of which states are constructed.

Building an antisymmetric basis. We need a basis of states each of which is antisymmetric. Now any wavefunction can be expressed as a sum over the above basis,

$$\psi(x_A, x_B, x_C) = \sum_{r=1}^M \sum_{s=1}^M \sum_{t=1}^M c_{r,s,t} \eta_r(x_A) \eta_s(x_B) \eta_t(x_C) = \sum_{r,s,t} c_{r,s,t} |r, s, t\rangle,$$

but we're not interested in any wavefunction, we're interested in antisymmetric wavefunctions. To build an antisymmetric wavefunction, we execute the antisymmetrization process on $\psi(x_A, x_B, x_C)$. Doing so, we conclude that this antisymmetric wavefunction can be expressed as a sum over the antisymmetrization of each basis element.

Let's think a bit about the antisymmetrization of

$$\eta_r(x_A)\eta_s(x_B)\eta_t(x_C) \quad \text{also known as} \quad |r, s, t\rangle.$$

When we introduced the antisymmetrization process, we permuted the variables (representing position windows) x_A , x_B , and x_C . But if the seed function is a product like this, it's obviously the same thing to permute the level indices r , s , and t . We represent the antisymmetrization of $|r, s, t\rangle$ as

$$\hat{A}|r, s, t\rangle = \text{const} (|r, s, t\rangle - |r, t, s\rangle + |t, r, s\rangle - |t, s, r\rangle + |s, t, r\rangle - |s, r, t\rangle)$$

where "const" is a normalization constant.

If we go through and antisymmetrize each element of the basis for three non-identical particles, we will find a basis for antisymmetric states. Let's start with $|1, 1, 1\rangle$. This antisymmetrizes to zero. Same with $|1, 1, 2\rangle$:

$$\hat{\mathcal{A}}|1, 1, 2\rangle = \text{const} (|1, 1, 2\rangle - |1, 2, 1\rangle + |2, 1, 1\rangle - |2, 1, 1\rangle + |1, 2, 1\rangle - |1, 1, 2\rangle) = 0.$$

It's clear, in fact, that *any* basis element with two indices the same will antisymmetrize to zero. The only way to avoid antisymmetrization to zero is for all the level indices to be zero. Furthermore

$$\hat{\mathcal{A}}|r, s, t\rangle = -\hat{\mathcal{A}}|r, t, s\rangle = \hat{\mathcal{A}}|t, r, s\rangle = -\hat{\mathcal{A}}|t, s, r\rangle = \hat{\mathcal{A}}|s, t, r\rangle = -\hat{\mathcal{A}}|s, r, t\rangle$$

so the six distinct basis elements $|1, 7, 3\rangle$, $|7, 3, 1\rangle$, $|3, 7, 1\rangle$, etc. all antisymmetrize to the same thing.

We are left with a basis of

$$\frac{M(M-1)(M-2)}{3!}$$

antisymmetric elements. One of the neat things about these elements is that they're long ... for example one of them is

$$\frac{1}{\sqrt{3!}} [\eta_1(x_A)\eta_7(x_B)\eta_3(x_C) - \eta_1(x_A)\eta_3(x_B)\eta_7(x_C) + \eta_3(x_A)\eta_1(x_B)\eta_7(x_C) - \eta_3(x_A)\eta_7(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_3(x_B)\eta_1(x_C) - \eta_7(x_A)\eta_1(x_B)\eta_3(x_C)]$$

but to specify them we need only state the three levels that go into building it (the three "building blocks" that go into making it). [This was *not* the case for three non-identical particles.] Consequently one often speaks of this state as "a particle in level 1, a particle in level 7, and a particle in level 3". This phrase is not absolutely correct: If a particle *were* in level 7, then it could be distinguished as "the particle in level 7" and hence would not be identical to the other two particles. The correct statement is that the *system* is in the antisymmetric state given above, and that the *individual particles* do not have states. On the other hand, the correct statement is a mouthful and you may use the "balls in buckets" picture as shorthand — as long as you say it but don't believe it.

Building a symmetric basis. We can build a basis of states, each of which is symmetric, in a parallel manner by symmetrizing each element of the basis for non-identical particles and discarding duplicates.

The symmetrization of

$$\eta_r(x_A)\eta_s(x_B)\eta_t(x_C) \quad \text{also known as} \quad |r, s, t\rangle$$

results in

$$\hat{\mathcal{S}}|r, s, t\rangle = \text{const} (|r, s, t\rangle + |r, t, s\rangle + |t, r, s\rangle + |t, s, r\rangle + |s, t, r\rangle + |s, r, t\rangle)$$

where "const" is again a normalization constant.

Let's start with $|1, 1, 1\rangle$. This symmetrizes to itself:

$$\hat{\mathcal{S}}|1, 1, 1\rangle = |1, 1, 1\rangle.$$

Next comes $|1, 1, 2\rangle$:

$$\hat{\mathcal{S}}|1, 1, 2\rangle = \text{const} (|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle + |2, 1, 1\rangle + |1, 2, 1\rangle + |1, 1, 2\rangle) = 2 \text{const} (|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle).$$

It's clear, then, that

$$\hat{S}|1, 1, 2\rangle = \hat{S}|1, 2, 1\rangle = \hat{S}|2, 1, 1\rangle.$$

and in fact that all states built out of any three given levels are the same state. The discarded duplicates result in a basis of

$$\frac{M(M+1)(M+2)}{3!}$$

symmetric elements.

Once again these states have long expressions like

$$\frac{1}{\sqrt{3!}} [\eta_1(x_A)\eta_7(x_B)\eta_3(x_C) + \eta_1(x_A)\eta_3(x_B)\eta_7(x_C) + \eta_3(x_A)\eta_1(x_B)\eta_7(x_C) \\ + \eta_3(x_A)\eta_7(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_3(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_1(x_B)\eta_3(x_C)]$$

but to specify the three-particle state we need only list the one-particle building blocks (“levels”) used in its construction. This results in almost the same “balls in buckets” picture that we drew for antisymmetric wavefunctions, except that now multiple balls can go into the same bucket. Once again you may use the “balls in buckets” picture as a shorthand, as long as you keep in mind that it conceals a considerably more intricate process of building and symmetrizing.

2.7 Problem: Count the elements in the antisymmetric and antisymmetric bases for N particles rather than three. (Continue to use M levels.) Does your expression have the proper limits when $N = 1$ and when $N = M$?

2.8 Problem: Find the normalization constant for $\hat{S}|7, 3, 7\rangle$.

2.9 Problem: Any two-variable function may be written as a sum of a symmetric and an antisymmetric function. Consequently the union of the symmetric basis and the antisymmetric basis is a basis for the set of all two-variable functions. Show that neither of these statements is true for functions of three variables.

2.10 Mean separation

(Be sure to read Griffiths section 5.1.2, “Exchange forces,” before attempting this problem.) Two noninteracting particles, each of mass m , are in an infinite square well of width L . The associated one-body energy eigenstates are $\eta_n(x)$ and $\eta_m(x)$, where

$$\eta_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right).$$

Calculate the root-mean-square separation

$$\sqrt{\langle (x_A - x_B)^2 \rangle}$$

if these are

- a. two non-identical particles, one in state $\eta_n(x_A)$ and the other in state $\eta_m(x_B)$

- b. two identical bosons, in state

$$\frac{1}{\sqrt{2}} [\eta_n(x_A)\eta_m(x_B) + \eta_m(x_A)\eta_n(x_B)]$$

- c. two identical fermions, in state

$$\frac{1}{\sqrt{2}} [\eta_n(x_A)\eta_m(x_B) - \eta_m(x_A)\eta_n(x_B)]$$

2.11 Building basis states

Suppose you had three particles and three “building block” levels (say the orthonormal levels $\eta_1(x)$, $\eta_3(x)$, and $\eta_7(x)$). Construct and count all the possible three-particle states representing

- three non-identical particles
- three identical bosons
- three identical fermions

Repeat for three particles with four one-particle levels, but in this case simply count and don’t write down all the three-particle states.

2.2 An antisymmetric basis for the helium problem

If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, nuclear motion, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for two electrons and one nucleus is

$$\hat{H} = \underbrace{\widehat{\text{KE}}_A + \hat{U}_{nA}}_{\equiv \hat{H}_A} + \underbrace{\widehat{\text{KE}}_B + \hat{U}_{nB}}_{\equiv \hat{H}_B} + \underbrace{\hat{U}_{AB}}_{\equiv \hat{H}'}$$

Recall that in using the subscripts “A” and “B” we are not labeling the *electrons* as “electron A” and “electron B”: the electrons are identical and can’t be labeled. Instead we are labeling the *points in space* where an electron might appear as “point A” and “point B”.

We will look for eigenstates of the partial Hamiltonian $\hat{H}_A + \hat{H}_B$. These are not eigenstates of the full Hamiltonian, but they *are* a basis, and they can be used as a place to begin perturbation theory.

One-particle levels

We begin by finding the one-particle levels (or “orbitals”) for just the Hamiltonian \hat{H}_A . We combine these with levels for \hat{H}_B and antisymmetrize the result.

The problem \hat{H}_A is just the Hydrogen atom Coulomb problem with two changes:

nuclear mass is $4m_p \implies$ very small effect (“ignore nuclear motion”)

nuclear charge is $2e \implies$ the Rydberg is $\text{Ry} = \frac{m_e}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2$, so $E_{nA}^{(A)} = -\frac{4\text{Ry}}{n_A^2}$

Similarly, the energy eigenstates for \hat{H}_A are represented by familiar functions like

$$\eta_{n\ell m}(\mathbf{r})|\uparrow\rangle \quad \text{or} \quad \eta_{n\ell m}(\mathbf{r})\chi_+.$$

Soon we will need to keep track of \hat{H}_A versus \hat{H}_B . A notation like $\eta_{n\ell m}(\mathbf{r}_A)|\uparrow\rangle$ is fine for the space part of the eigenstate, but leaves the spin part ambiguous. We will hence use notation like

$$\eta_{n\ell m}(A)\chi_+(A)$$

to mean the same thing.

[[Notice that the eigenstates don't *have* to take on the factorized form of "space part" \times "spin part" — for example

$$\frac{1}{\sqrt{2}}[\eta_{200}(\mathbf{r})\chi_+ + \eta_{210}(\mathbf{r})\chi_-]$$

is a perfectly good eigenstate — but that the factorized form is particularly convenient for working with. (If we were to consider spin-orbit coupling, then the eigenstates could *not* take the factorized form.)]]

Antisymmetrization

Recall how we build an antisymmetrized wavefunction from a product of two one-particle levels, $\eta_n(A)$ and $\eta_m(B)$:

$$\hat{A}\eta_n(A)\eta_m(B) = \frac{1}{\sqrt{2}}[\eta_n(A)\eta_m(B) - \eta_m(A)\eta_n(B)]$$

(The normalization factor $1/\sqrt{2}$ holds when $\eta_n(A)$ and $\eta_m(A)$ are orthogonal.)

Two theorems:

- If you antisymmetrize a product of the same two levels, you end up with zero:

$$\hat{A}\eta_n(A)\eta_n(B) = 0.$$

- If you antisymmetrize two levels in the opposite sequence, you end up with the same state:

$$\hat{A}\eta_n(A)\eta_m(B) = -\hat{A}\eta_m(A)\eta_n(B).$$

The ground state

The ground levels of \hat{H}_A and of \hat{H}_B are both doubly degenerate due to spin. So if you had distinguishable particles, the ground state of $\hat{H}_A + \hat{H}_B$ would be four-fold degenerate:

distinguishable

$$\begin{aligned} &\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_-(B) \end{aligned}$$

But if you have identical fermions, two of these basis states antisymmetrize to zero, and the other two antisymmetrize to the same state:

distinguishable	antisymmetrized
$\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)$	0
$\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B) - \eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)]$
$\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)$	$-\frac{1}{\sqrt{2}}[\text{above}]$
$\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)$	0

Hence the Hamiltonian $\hat{H}_A + \hat{H}_B$ has a non-degenerate ground state, namely

$$\eta_{100}(A)\eta_{100}(B)\frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)].$$

States built from one ground level

Now build a state by combining the ground level of one Hamiltonian with $|n\ell m\rangle$ from the other. If you had distinguishable particles, this “combination” means a simple multiplication, and there would be eight states (all with the same energy):

distinguishable
$\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B)$
$\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B)$
$\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B)$
$\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B)$
$\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)$
$\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)$
$\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)$
$\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)$

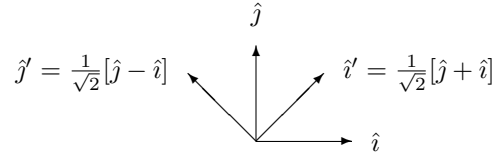
But if you have identical fermions, the “combination” means a multiplication followed by an antisymmetrization. Because of the second theorem concerning antisymmetrization, each of the last four products above antisymmetrize to the same state as the product four lines above it. The first four products result in antisymmetrized states as follows:

distinguishable	antisymmetrized
(a) $\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B) - \eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)]$
(b) $\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B) - \eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)]$
(c) $\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B) - \eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)]$
(d) $\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B) - \eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)]$

Antisymmetrized expressions (a) and (d) readily factor into a space part times a spin part:

$$\begin{aligned} \text{(a)} &\implies \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_+(B) \\ \text{(d)} &\implies \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_-(B) \end{aligned}$$

But expressions (b) and (c) do *not* factor. One thing to do about this is nothing — after all, there’s no *requirement* that the wavefunctions factorize. But another approach is to look for a simple change of basis (remember, these four states all have the same energy of $\hat{H}_A + \hat{H}_B$). Someone (I don’t know who) thought about the favorite change of basis in planar geometry — a rotation of the axes by 45° :



Applying this transformation to the basis elements (“unit vectors”) given through (b) and (c) results in the new basis elements

$$\begin{aligned} \frac{1}{\sqrt{2}}[(b) + (c)] &= \frac{1}{2} \{ [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_-(B) \\ &\quad + [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_+(B) \} \\ &= \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)] \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\sqrt{2}}[(b) - (c)] &= \frac{1}{2} \{ [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_-(B) \\ &\quad - [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_+(B) \} \\ &= \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)]. \end{aligned}$$

This process results in an antisymmetric basis of

$$\begin{aligned} &\frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_+(B) \\ &\frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)] \\ &\frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_-(B) \\ &\frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)]. \end{aligned}$$

The first three elements are called a “triplet” (with “space antisymmetric, spin symmetric”). The last element is called a “singlet” (with “space symmetric, spin antisymmetric”). This particular basis has three nice properties: (1) Every basis element factorizes into a spatial part times a spin part. (2) Every basis element factorizes into a symmetric part times an antisymmetric part. (3) All three elements in the triplet have identical spatial parts.

The third point means that when we take account of electron-electron repulsion through perturbation theory, we will necessarily find that all three elements of any triplet remain degenerate even when the effects of the sub-Hamiltonian \hat{U}_{AB} are considered.

[[This process works for combining an arbitrary level $|n\ell m\rangle$ with a ground state level $|100\rangle$. Since this should work for any $|n\ell m\rangle$, what happens if we take $|n\ell m\rangle = |100\rangle$, the situation we first considered? In particular, what's up with the normalization?]]

States built from two excited levels

What happens if we carry out the above process but combining an excited level of one sub-Hamiltonian (say $\eta_{200}(A)$) with an arbitrary level of the other sub-Hamiltonian (say $\eta_{\ell m}(B)$)?

The process goes on in a straightforward way, but it turns out that the resulting eigenenergies are always so high that the atom is unstable: it decays rapidly to a positive helium atom plus an ejected electron. Such electrons are called “Auger electrons” (pronounced “oh-jey” because Pierre Victor Auger was French) and Auger electron spectroscopy is an important analytical technique in surface and materials science.

2.12 Electron-electron repulsion

In class we wrote the (approximate) Hamiltonian for helium as $\hat{H}_A + \hat{H}_B + \hat{U}_{AB}$ and found antisymmetric energy eigenstates for $\hat{H}_A + \hat{H}_B$ that we called 1^1S , 2^3S , 2^1P , and so forth. Then we qualitatively discussed how the energy associated with such a state would change, under perturbation theory, through the electron-electron repulsion term \hat{U}_{AB} . Write down expressions for the first-order energy shifts due to \hat{U}_{AB} for 2^1S , 2^3S , and 2^1P . (That is, set up the integrals in terms of the one-particle eigenstates $\eta_{\ell m}(\mathbf{r})$. Do *not* evaluate the integrals.) *Bonus*: Argue that the energy shift for 2^1P is greater than the shift for 2^1S .

2.13 Two-electron ions

Apply the techniques of Griffiths, section 7.2, “Ground State of Helium,” to the H^- and Li^+ ions. Each of these ions has two electrons, like helium, but nuclear charges $Z = 1$ and $Z = 3$, respectively. For each ion find the effective (partially shielded) nuclear charge and determine the best upper bound on the ground state energy.

2.3 Ground state energy of helium

The role of theory

Jacov Ilich Frenkel (also Yakov Ilich Frenkel or Iakov Ilich Frenkel; 1894–1952) was a prolific physicist. Among other things he coined the term “phonon”. In a review article on the theory of metals (quoted by M.E. Fisher in “The Nature of Critical Points”, Boulder lectures, 1965) he said:

The more complicated the system considered, the more simplified must its theoretical description be. One cannot demand that a theoretical description of a complicated atom, and all the more of a molecule or a crystal, have the same degree of accuracy as of the theory of the simplest hydrogen atom. Incidentally, such a requirement is not only impossible to fulfill but also essentially useless. . . . An exact calculation of the constants characterizing the simplest physical system has essential significance as a test on the correctness of the basic principles of the theory. However, once it passes this test brilliantly there is no sense in subjecting it to further tests as applied to more complicated systems. The most ideal theory cannot pass such tests, owing to the practically unsurmountable mathematical difficulties unavoidably encountered in applications to complicated systems. In this case all that is demanded of the theory is a correct interpretation of the general character of the quantities and laws pertaining to such a system. The theoretical physicist is in this respect like a cartoonist, who must depict the original, not in all details like a photographic camera, but simplify and schematize it in a way as to disclose and emphasize the most characteristic features. Photographic accuracy can and should be required only of the description of the simplest system. A good theory of complicated systems should represent only a good “caricature” of these systems, exaggerating the properties that are most difficult, and purposely ignoring all the remaining inessential properties.

Which case is the ground state of He?

- 1) Fundamental test of symmetrization postulate.
- 2) Test to see whether QM breaks down for complex systems (Anthony J. Leggett).
- 3) Refinements can involve new physical ideas.
- 4) Physical effects other than ground state energy.

Experiment

$$E_g = -78.975 \text{ eV.}$$

Theory

(Summarizing Griffiths 5.2.1 and 7.2.) If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, nuclear motion, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for two electrons and one nucleus is

$$\hat{H} = \hat{H}_A + \hat{H}_B + \hat{U}_{AB} \quad (2.1)$$

where

$$\hat{U}_{AB} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_A - \mathbf{r}_B|}. \quad (2.2)$$

The ground state wavefunction for H is

$$\eta_{100}(\mathbf{r}) = \frac{1}{\sqrt{\pi}a_0^{3/2}}e^{-r/a_0}. \quad (2.3)$$

But if the nucleus had charge $+Ze$, this would be

$$\eta_{100}(\mathbf{r}) = \frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}}e^{-Zr/a_0}. \quad (2.4)$$

So the $\hat{U}_{AB} = 0$ ground state is

$$\eta_{100}(\mathbf{r}_A)\eta_{100}(\mathbf{r}_B) = \frac{Z^3}{\pi a_0^3}e^{-Z(r_A+r_B)/a_0} \quad \text{with} \quad Z = 2. \quad (2.5)$$

This state gives a ground state energy of $E_g = -8(\text{Ry}) = -109 \text{ eV}$.

Turning on the electron-electron repulsion, perturbation theory finds $\langle \hat{U}_{AB} \rangle$ and jacks up E_g to -75 eV .

The variational method uses the same wavefunction as above, but considers Z not as 2 but as an adjustable parameter. Interpretation: “shielding” — expect $1 < Z_{\min} < 2$. And in fact minimizing $\langle H \rangle$ with over this class of trial wavefunctions gives $Z_{\min} = 1.69$ and $E_g = -77.5 \text{ eV}$. (Sure enough, an overestimate.) Griffiths stops here and suggests that the rest of the work is humdrum.

Further theory

Review: A. Hibbert, *Rept. Prog. Phys.* **38** (1975) 1222–1225.

Hylleraas (1929): Trial wavefunction of form (atomic units)

$$\psi(\mathbf{r}_A, \mathbf{r}_B) = e^{-Z(r_A+r_B)} \sum c_{nlm} (Z(r_A+r_B))^n (Z(r_B-r_A))^{2l} (Z|\mathbf{r}_A-\mathbf{r}_B|)^m.$$

[I won’t go into all the reasons why he picked this trial wavefunction, but...ask why only even powers $2l$.] Using Z and six terms in sum as variational parameters, he got an energy good to 2 parts in 10,000.

This is a good energy. Is there any point in doing better? Yes. Although it gives you a good energy, it gives you a poor wavefunction: Think of a $d = 2$ landscape with a hidden valley — e.g. a crater, an absolute minimum. The $d = 2$ landscape represents two variational parameters — by coincidence, the exact wavefunction has the form that you guessed. If you tried just one variational parameter, you’d be walking a line in this landscape. The line could be quite far from the valley bottom while giving very good elevation estimates for the valley bottom, because the valley is flat at the bottom. [Sketch.]

In fact, you can show that *no* wavefunction of this form, no matter how many terms you pick, can satisfy the Schrödinger Equation — even if you picked an infinite number of terms, you’d never hit the wavefunction right on!

Is there any reason to get the wavefunction right? Yes! For example if you wanted to calculate Stark or Zeeman effect, or spin-orbit, or whatever, you’d need those wavefunctions for doing perturbation theory!

Kinoshita (1959): One of the “great fiddlers of physics”. Trial wavefunction of form (atomic units)

$$\psi(\mathbf{r}_A, \mathbf{r}_B) = e^{-Z(r_A+r_B)} \sum c_{nlm} (Z(r_A+r_B))^n \left(\frac{r_B-r_A}{|\mathbf{r}_A-\mathbf{r}_B|} \right)^{2l} \left(\frac{|\mathbf{r}_A-\mathbf{r}_B|}{r_A+r_B} \right)^m .$$

He showed that this could satisfy the Schrödinger Equation exactly if sum were infinite. Used 80 terms for accuracy 1 part in 100,000.

Pekeris (1962): A different trial wavefunction guaranteed to get the correct form when both electrons are far from nucleus. Used 1078 terms, added fine structure and hyperfine structure, got accuracy 1 part in 10^9 .

Schwartz (1962): Added terms like $[Z(r_A+r_B)]^{n/2}$... not smooth. Got better energies with 189 terms!

Frankowski and Pekeris (1966): Introduced terms like $\ln^k(Z(r_A+r_B))$... not smooth. 246 terms, accuracy 1 part in 10^{12} .

Kato: (See Drake, page 155.) Looked at condition for two electrons close, both far from nucleus. In this case it's like H atom, wavefunction must have cusp. Allow electrons to show this cusp.

State of art: Gordon W.F. Drake, ed. *Atomic, Molecular, and Optical Physics Handbook* page 163. [Reference QC173.A827 1996]

New frontiers: experiment. S.D. Bergeson, *et al.*, “Measurement of the He ground state Lamb shift”, *Phys. Rev. Lett.* **80** (1998) 3475–3478.

New frontiers: theory. S.P. Goldman, “Uncoupling correlated calculations in atomic physics: Very high accuracy and ease,” *Phys. Rev. A* **57** (1998) 677–680. 8066 terms, 1 part in 10^{18} .

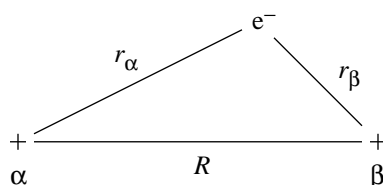
New frontiers: Lithium, metallic Hydrogen.

Chapter 3

Molecules

3.1 The hydrogen molecule ion

The hydrogen molecule ion is two protons and a single electron... H_2^+ . If we had managed to successfully solve the helium atom problem we would also have solved this one, because it's just three particles interacting through $1/r^2$ forces. However, you know that this problem has not been exactly solved even in the classical limit. Thus we don't even look for an exact solution: we look for the approximation most applicable to the case of two particles much more massive than the third.



If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for one electron and two protons (α and β) is

$$\hat{H} = \widehat{\text{KE}}_\alpha + \widehat{\text{KE}}_\beta + \widehat{\text{KE}}_e + \hat{U}_{\alpha\beta} + \hat{U}_{\alpha e} + \hat{U}_{\beta e} \quad (3.1)$$

This is, of course, also the Hamiltonian for the helium atom, or for any three-body problem with pair interactions. Now comes the approximation suitable for the hydrogen molecule ion (but *not* appropriate for the helium atom): Assume that the two protons are so massive that they are fixed, and the interaction between them is treated classically. In equations, this approximation demands

$$\widehat{\text{KE}}_\alpha = 0; \quad \widehat{\text{KE}}_\beta = 0; \quad \hat{U}_{\alpha\beta} = U_{\alpha\beta} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{R}. \quad (3.2)$$

The remaining, quantum mechanical, piece of the full Hamiltonian is the electronic Hamiltonian

$$\hat{H}_e = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{r_\alpha} + \frac{1}{r_\beta}\right). \quad (3.3)$$

This approximation is called the “Born-Oppenheimer” approximation.

What shall we do with the electronic Hamiltonian? It would be nice to have an analytic solution of the energy eigenproblem. Then we could do precise comparisons between these results and the experimental spectrum of the hydrogen molecule ion, and build on them to study the hydrogen molecule, in exactly the same way that we built on our exact solution for He^+ to get an approximate solution for He. This goal is hopelessly beyond our reach. [Check out Gordon W.F. Drake, editor, *Atomic, Molecular, and Optical Physics Handbook* (AIP Press, Woodbury, NY, 1996) Reference QC173.A827 1996. There’s a chapter on high-precision calculations for helium, but no chapter on high-precision calculations for the hydrogen molecule ion.] Instead of giving up, we might instead look for an exact solution to the ground state problem. This goal is also beyond our reach. Instead of giving up, we use the variational method to look for an approximate ground state.

Before doing so, however, we notice one exact symmetry of the electronic Hamiltonian that will guide us in our search for approximate solutions. The Hamiltonian is symmetric under the interchange of symbols α and β or, what is the same thing, symmetric under inversion about the point midway between the two nuclei. Any discussion of parity (see, for example, Gordon Baym *Lectures on Quantum Mechanics* pages 99–101) shows that this means the energy eigenfunctions can always be chosen either odd or even under the interchange of α and β .

Where will we find a variational trial wavefunction? If nucleus β did not exist, the ground state wavefunction would be the hydrogen ground state wavefunction centered on nucleus α :

$$\eta_\alpha(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\alpha/a_0} \equiv |\alpha\rangle. \quad (3.4)$$

Similarly if nucleus α did not exist, the ground state wavefunction would be

$$\eta_\beta(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\beta/a_0} \equiv |\beta\rangle. \quad (3.5)$$

We take as our trial wavefunction a linear combination of these two wavefunctions. This trial wavefunction is called a “linear combination of atomic orbitals” or “LCAO”. So the trial wavefunction is

$$\psi(\vec{r}) = A\eta_\alpha(\vec{r}) + B\eta_\beta(\vec{r}). \quad (3.6)$$

At first glance, it seems that the variational parameters are the complex numbers A and B , for a total of four real parameters. However, one parameter is taken up through normalization, and one through overall phase. Furthermore, because of parity the swapping of α and β can result in at most a change in sign, whence $B = \pm A$. Thus our trial wavefunction is

$$\psi(\vec{r}) = A_\pm[\eta_\alpha(\vec{r}) \pm \eta_\beta(\vec{r})], \quad (3.7)$$

where A_{\pm} is the normalization constant, selected to be real and positive. (The notation A_{\pm} reflects the fact that depending on whether we take the + sign or the - sign, we will get a different normalization constant.)

This might seem like a letdown. We have discussed exquisitely precise variational wavefunction involving hundreds or even thousands of real parameters. Here the only variational parameter is the binary choice: + sign or - sign! Compute $\langle \hat{H}_e \rangle$ both ways and see which is lower! You don't even have to take a derivative at the end! Clearly this is a first attempt and more accurate calculations are possible. Rather than give in to despair, however, let's recognize the limitations and forge on to see what we can discover. At the very least what we learn here will guide us in selecting better trial wavefunctions for our next attempt.

There are only two steps: normalize the wavefunction and evaluate $\langle \hat{H}_e \rangle$. However, these steps can be done through a frontal assault (which is likely to get hopelessly bogged down in algebraic details) or through a more subtle approach recognizing that we already know quite a lot about the functions $\eta_{\alpha}(\vec{r})$ and $\eta_{\beta}(\vec{r})$, and using this knowledge to our advantage. Let's use the second approach.

Normalization demands that

$$\begin{aligned} 1 &= |A_{\pm}|^2 (\langle \alpha | \pm \langle \beta |) (| \alpha \rangle \pm | \beta \rangle) \\ &= |A_{\pm}|^2 (\langle \alpha | \alpha \rangle \pm \langle \alpha | \beta \rangle \pm \langle \beta | \alpha \rangle + \langle \beta | \beta \rangle) \\ &= 2|A_{\pm}|^2 (1 \pm \langle \alpha | \beta \rangle) \end{aligned}$$

where in the last step we have used the normalization of $|\alpha\rangle$ and $|\beta\rangle$. The integral $\langle \alpha | \beta \rangle$ is not easy to calculate, so we set it aside for later by naming it the *overlap integral*

$$I(R) \equiv \langle \alpha | \beta \rangle = \int \eta_{\alpha}(\vec{r}) \eta_{\beta}(\vec{r}) d^3r. \quad (3.8)$$

In terms of this integral, we can select the normalization to be

$$A_{\pm} = \frac{1}{\sqrt{2(1 \pm I(R))}}. \quad (3.9)$$

Evaluating the electronic Hamiltonian in the trial wavefunction gives

$$\begin{aligned} \langle \hat{H}_e \rangle &= \frac{(\langle \alpha | \pm \langle \beta |) \hat{H}_e (| \alpha \rangle \pm | \beta \rangle)}{2(1 \pm I(R))} \\ &= \frac{\langle \alpha | \hat{H}_e | \alpha \rangle \pm \langle \alpha | \hat{H}_e | \beta \rangle \pm \langle \beta | \hat{H}_e | \alpha \rangle + \langle \beta | \hat{H}_e | \beta \rangle}{2(1 \pm I(R))} \\ &= \frac{\langle \alpha | \hat{H}_e | \alpha \rangle \pm \langle \beta | \hat{H}_e | \alpha \rangle}{1 \pm I(R)} \end{aligned} \quad (3.10)$$

But we have already done large parts of these two integrals:

$$\begin{aligned} \hat{H}_e | \alpha \rangle &= \left[\widehat{\text{KE}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\alpha}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\beta}} \right] | \alpha \rangle \\ &= \left[\widehat{\text{KE}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\alpha}} \right] | \alpha \rangle - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\beta}} | \alpha \rangle \end{aligned}$$

$$\begin{aligned}
&= -\text{Ry} |\alpha\rangle - 2 \text{Ry} a_0 \frac{1}{r_\beta} |\alpha\rangle \\
&= -\text{Ry} \left[|\alpha\rangle + 2 \frac{a_0}{r_\beta} |\alpha\rangle \right]
\end{aligned} \tag{3.11}$$

whence

$$\langle \alpha | \hat{H}_e | \alpha \rangle = -\text{Ry} \left[1 + 2 \left\langle \alpha \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \right] \tag{3.12}$$

$$\langle \beta | \hat{H}_e | \alpha \rangle = -\text{Ry} \left[\langle \beta | \alpha \rangle + 2 \left\langle \beta \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \right]. \tag{3.13}$$

On the right-hand side we recognize the exchange integral, $I(R) = \langle \beta | \alpha \rangle$, and two new (dimensionless) integrals, which are called the *direct integral*

$$D(R) \equiv \left\langle \alpha \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \tag{3.14}$$

and the *exchange integral*

$$X(R) \equiv \left\langle \beta \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle. \tag{3.15}$$

These two integrals are not easy to work out (I will assign them as homework) but once we do them (plus the overlap integral) we can find the expectation value of the electronic Hamiltonian in the trial wavefunction. It is

$$\begin{aligned}
\langle \hat{H}_e \rangle &= -\text{Ry} \frac{1 + 2D(R) \pm I(R) \pm 2X(R)}{1 \pm I(R)} \\
&= -\text{Ry} \left[1 + 2 \frac{D(R) \pm X(R)}{1 \pm I(R)} \right].
\end{aligned} \tag{3.16}$$

This, remember, is only the electronic part of the Hamiltonian. In the Born-Oppenheimer approximation the nuclear part has no kinetic energy and Coulombic potential energy

$$\frac{e^2}{4\pi\epsilon_0} \frac{1}{R} = 2 \text{Ry} \frac{a_0}{R}, \tag{3.17}$$

so the upper bound on the total ground state energy is

$$\text{Ry} \left[2 \frac{a_0}{R} - 1 - 2 \frac{D(R) \pm X(R)}{1 \pm I(R)} \right]. \tag{3.18}$$

3.2 Problems

3.1 The hydrogen molecule ion: Evaluation of integrals

Evaluate the direct and exchange integrals $D(R)$ and $X(R)$. (*Hint:* Remember that $\sqrt{x^2} = |x|$.) Plot as a function of R the overlap integral, $I(R)$, as well as $D(R)$ and $X(R)$.

3.2 The hydrogen molecule ion: Thinking about integrals

For the hydrogen molecule ion, find and plot the expectation values of nuclear potential energy, total electronic energy, kinetic electronic energy, and potential electronic energy for the state $\psi_+(\vec{r})$, as functions of R . Do these plots shed any light on our initial question of “Why is stuff hard?” (We gave possible answers of “repulsion hardness,” “Heisenberg hardness,” and “Pauli hardness.”) *Bonus:* The hydrogen molecule ion cannot display Pauli hardness, because it has only one quantal particle. Can you generalize this discussion to the neutral hydrogen molecule?

3.3 Improved variational wavefunction

Everett Schlawin ('09) suggested using “shielded” subwavefunctions like equation (2.4) in place of the subwavefunctions (3.4) and (3.5) that go into making trial wavefunction (3.7). Then there would be a variational parameter Z in addition to the binary choice of $+$ or $-$. I haven't tried this, but through the usual variational argument, it can't be worse than what we've tried so far! (That is, the *results* can't be worse. The *amount of labor* involved can be far, far worse.) Execute this suggestion. Show that this trial wavefunction results in the exact helium ion ground state energy in the case $R = 0$.

Chapter 4

Perturbation Theory for the Time Development Problem

4.1 Setup

Here's our problem:

Solve the initial value problem for the Hamiltonian

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{H}'(t) \tag{4.1}$$

given the solution $\{|n\rangle\}$ of the unperturbed energy eigenproblem

$$\hat{H}^{(0)}|n\rangle = E_n|n\rangle. \tag{4.2}$$

Here we're thinking of $\hat{H}'(t)$ as being in some sense “small” compared to the unperturbed Hamiltonian $\hat{H}^{(0)}$. Note also that it doesn't make sense to solve the energy eigenproblem for $\hat{H}(t)$, because this Hamiltonian depends upon time, so it doesn't have stationary state solutions!

We solve this problem by expanding the solution $|\psi(t)\rangle$ in the basis $\{|n\rangle\}$:

$$|\psi(t)\rangle = \sum_n C_n(t)|n\rangle \quad \text{where} \quad C_n(t) = \langle n|\psi(t)\rangle. \tag{4.3}$$

Once we know the $C_n(t)$, we'll know the solution $|\psi(t)\rangle$. Now, the state vector evolves according to

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \tag{4.4}$$

so the expansion coefficients evolve according to

$$\begin{aligned}
\frac{dC_n(t)}{dt} &= -\frac{i}{\hbar} \langle n | \hat{H} | \psi(t) \rangle \\
&= -\frac{i}{\hbar} \sum_m \langle n | \hat{H} | m \rangle C_m(t) \\
&= -\frac{i}{\hbar} \sum_m \left[\langle n | \hat{H}^{(0)} | m \rangle + \langle n | \hat{H}' | m \rangle \right] C_m(t) \\
&= -\frac{i}{\hbar} \sum_m [E_m \delta_{m,n} + H'_{n,m}] C_m(t) \\
&= -\frac{i}{\hbar} \left[E_n C_n(t) + \sum_m H'_{n,m} C_m(t) \right]
\end{aligned} \tag{4.5}$$

This result is exact: we have yet to make any approximation.

Now, if $\hat{H}'(t)$ vanished, the solutions would be

$$C_n(t) = C_n(0) e^{-(i/\hbar)E_n t}, \tag{4.6}$$

which motivates us to define new variables $c_n(t)$ through

$$C_n(t) = c_n(t) e^{-(i/\hbar)E_n t}. \tag{4.7}$$

Because the “bulk of the time development” comes through the $e^{-(i/\hbar)E_n t}$ term, the $c_n(t)$ presumably have “less time dependence” than the $C_n(t)$. In other words, we expect the $c_n(t)$ to vary slowly with time.

Plugging this definition into the time development equation (4.5) gives

$$\frac{dc_n(t)}{dt} e^{-(i/\hbar)E_n t} + c_n(t) (-i/\hbar)E_n e^{-(i/\hbar)E_n t} = -\frac{i}{\hbar} \left[E_n c_n(t) e^{-(i/\hbar)E_n t} + \sum_m H'_{n,m} c_m(t) e^{-(i/\hbar)E_m t} \right] \tag{4.8}$$

or

$$\frac{dc_n(t)}{dt} = -\frac{i}{\hbar} \sum_m H'_{n,m} c_m(t) e^{+(i/\hbar)(E_n - E_m)t}. \tag{4.9}$$

Once again, this equation is exact. Its formal solution, given the initial values $c_n(0)$, is

$$c_n(t) = c_n(0) - \frac{i}{\hbar} \sum_m \int_0^t H'_{n,m}(t') c_m(t') e^{+(i/\hbar)(E_n - E_m)t'} dt'. \tag{4.10}$$

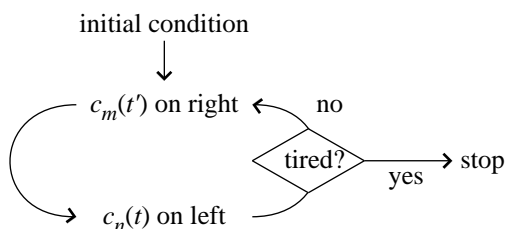
This set of equations (one for each basis element) is exact, but at first glance seems useless. The unknown quantities $c_n(t)$ are present on the left, but also the right-hand sides.

We can make progress using our idea that the coefficients $c_n(t)$ are changing *slowly*. In a very crude approximation, we can think that they’re not changing at all. So on the right-hand side of equation (4.10) we plug in not functions, but the constants $c_m(t') = c_m(0)$, namely the given initial conditions.

Having made that approximation, we can now perform the integrations and produce, on the left-hand side of equation (4.10), functions of time $c_n(t)$. These coefficients aren’t exact, because they were based

on the crude approximation that the coefficients were constant in time, but they're likely to be better approximations than we started off with.

Now, armed with these more accurate coefficients, we can plug these into the right-hand side of equation (4.10), perform the integration, and produce yet more accurate coefficients on the left-hand side. This process can be repeated over and over, for as long as our stamina lasts.



There is actually a theorem assuring us that this process will converge!

Theorem (Picard) If the matrix elements $H'_{n,m}(t)$ are continuous in time and bounded, and if the basis is finite, then this method converges to the correct solution.

The theorem does not tell us how many iterations will be needed to reach a desired accuracy. In practice, one usually stops upon reaching the first non-zero correction.

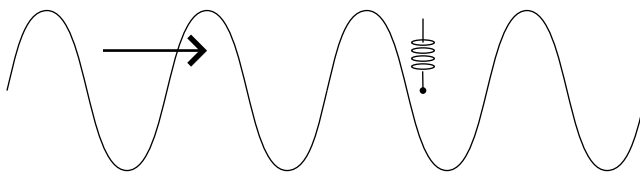
In particular, if the initial state is some eigenstate $|a\rangle$ of the unperturbed Hamiltonian $\hat{H}^{(0)}$, then to first order

$$\begin{aligned}
 c_n(t) &= -\frac{i}{\hbar} \int_0^t H'_{n,a}(t') e^{+(i/\hbar)(E_n - E_a)t'} dt' & \text{for } n \neq a \\
 c_a(t) &= 1 - \frac{i}{\hbar} \int_0^t H'_{a,a}(t') dt'
 \end{aligned}
 \tag{4.11}$$

If the system is in energy state $|a\rangle$ at time zero, then the probability of finding it in energy state $|b\rangle$ at time t , through the influence of perturbation $\hat{H}'(t)$, is called the *transition probability*

$$P_{a \rightarrow b}(t) = |C_b(t)|^2 = |c_b(t)|^2.
 \tag{4.12}$$

Example: An electron bound to an atom is approximated by a one-dimensional simple harmonic oscillator of natural frequency ω_0 . The oscillator is in its ground state $|0\rangle$ and then exposed to light of electric field amplitude E_0 and frequency ω for time t . (The light is polarized in the direction of the oscillations.) What is probability (in first-order perturbation theory) of ending up in state $|b\rangle$?



Solution part A — *What is the Hamiltonian?* If it were a classical particle of charge $-e$ exposed to electric field $E_0 \cos \omega t$, it would experience a force $-eE_0 \cos \omega t$ and hence have a potential energy of $eE_0 x \cos \omega t$. (We can ignore the spatial variation of electric field because the electron is constrained to move only up and down — that’s our “one dimensional” assumption. We can ignore magnetic field for the same reason.)

The quantal Hamiltonian is then

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2}{2}\hat{x}^2 + eE_0\hat{x} \cos \omega t. \quad (4.13)$$

We identify the first two terms as the time-independent Hamiltonian $\hat{H}^{(0)}$ and the last term as the perturbation $\hat{H}'(t)$.

Solution part B — *Apply perturbation theory.* The matrix element is

$$H_{n,0}(t) = \langle n | \hat{H}'(t) | 0 \rangle = eE_0 \cos \omega t \langle n | \hat{x} | 0 \rangle = eE_0 \cos \omega t \sqrt{\frac{\hbar}{2m\omega_0}} \delta_{n,1}. \quad (4.14)$$

(Remember your raising and lowering operators! See equation (C.30).) Invoking equations (4.11), we obtain

$$c_n(t) = 0 \quad \text{for } n \neq 0, 1 \quad (4.15)$$

$$c_1(t) = -\frac{i}{\hbar} eE_0 \sqrt{\frac{\hbar}{2m\omega_0}} \int_0^t \cos \omega t' e^{i\omega_0 t'} dt' \quad (4.16)$$

$$c_0(t) = 1 \quad (4.17)$$

We will eventually need to perform the time integral in equation (4.16), but even before doing so the main qualitative features are clear: First, probability is *not* conserved within first order perturbation theory. The probability of remaining in the ground state is 1, but the probability of transition to the first excited state is finite! Second, to first order transitions go only to the first excited state. This is an example of a *selection rule*.

The time integral in equation (4.16) will be evaluated at equation (4.29). For now, let’s just call it $I(t)$. In terms of this integral, the transition probabilities are

$$P_{0 \rightarrow b}(t) = 0 \quad \text{for } b \neq 0, 1 \quad (4.18)$$

$$P_{0 \rightarrow 1}(t) = \frac{e^2 E_0^2}{2m\hbar\omega_0} I(t) I^*(t) \quad (4.19)$$

$$P_{0 \rightarrow 0}(t) = 1 \quad (4.20)$$

4.2 Fermi's golden rule

How do atoms absorb light?

More specifically, if an electron in atomic energy eigenstate $|a\rangle$ (usually but not always the ground state) is exposed to a beam of monochromatic, polarized light for time t , what is the probability of it ending up in atomic energy eigenstate $|b\rangle$? We answer this question to first order in time-dependent perturbation theory.

First, we need to find the effect of light on the electron. We'll treat the light classically — that is, we'll ignore the quantization of the electromagnetic field (quantum electrodynamics) that gives rise to the concept of photons. Consider the light wave (polarized in the \hat{k} direction, with frequency ω) as an electric field

$$\vec{E}(\vec{r}, t) = E_0 \hat{k} \cos(\vec{k} \cdot \vec{r} - \omega t). \quad (4.21)$$

Presumably, the absorption of light by the atom will result in some sort of diminution of the light beam's electric field, but we'll ignore that. (A powerful beam from a laser will be somewhat diminished when some of the light is absorbed by a single atom, but not a great deal.) The light beam has a magnetic field as well as an electric field, but the magnetic field amplitude is $B_0 = E_0/c$, so the electric force is on the order of eE_0 while the magnetic force is on the order of $evB_0 = e(v/c)E_0$. Since the electron moves at non-relativistic speeds, $v/c \ll 1$ and we can ignore the magnetic effect. Finally, the electric field at one side of the atom differs from the electric field at the other side of the atom, but the atom is so small compared to the wavelength of light (atom: about 0.1 nm; wavelength of violet light: about 400 nm) that we can safely ignore this also.

Using these approximations, the force experienced by an electron due to the light beam is

$$\vec{F}(t) = -eE_0 \hat{k} \cos(\omega t), \quad (4.22)$$

so the associated potential energy is

$$U(t) = eE_0 z \cos(\omega t). \quad (4.23)$$

Turning this classical potential energy into a quantal operator gives

$$\hat{H}'(t) = eE_0 \hat{z} \cos(\omega t). \quad (4.24)$$

(Note that the hat \hat{k} in equation (4.22) signifies unit vector, whereas the hat \hat{z} in equation (4.24) signifies quantal operator. I'm sorry for any confusion... there just aren't enough symbols in the world to represent everything unambiguously!)

Now that we have the quantal operator for the perturbation, we can turn to the time-dependent perturbation theory result (4.11). (Is it legitimate to use perturbation theory in this case? See the problem.)

For all of the atomic energy states $|a\rangle$ we've considered in this book,

$$H'_{a,a}(t) = \langle a | H'(t) | a \rangle = eE_0 \langle a | \hat{z} | a \rangle \cos(\omega t) = 0, \quad (4.25)$$

whence $c_a(t) = 1$ and $P_{a \rightarrow a} = 1$. Most of the atoms *don't* make transitions.

But what about those that do? For these we need to find the matrix elements

$$H'_{b,a}(t) = \langle b|H'(t)|a\rangle = eE_0\langle b|\hat{z}|a\rangle \cos(\omega t). \quad (4.26)$$

These are just the $z_{b,a}$ matrix elements that we calculated for the Stark effect. (And after all, what we're considering here is just the Stark effect with an oscillating electric field.) The transition amplitudes are

$$c_b(t) = -\frac{i}{\hbar}eE_0\langle b|\hat{z}|a\rangle \int_0^t \cos(\omega t')e^{+(i/\hbar)(E_b-E_a)t'} dt'. \quad (4.27)$$

It is convenient (and conventional!) to follow the lead of Einstein's $\Delta E = \hbar\omega$ and define

$$E_b - E_a = \hbar\omega_0. \quad (4.28)$$

The time integral is then

$$\begin{aligned} \int_0^t \cos(\omega t')e^{i\omega_0 t'} dt' &= \int_0^t \frac{e^{+i\omega t'} + e^{-i\omega t'}}{2} e^{i\omega_0 t'} dt' \\ &= \frac{1}{2} \left[\int_0^t e^{i(\omega_0+\omega)t'} dt' + \int_0^t e^{i(\omega_0-\omega)t'} dt' \right] \\ &= \frac{1}{2} \left[\frac{e^{i(\omega_0+\omega)t'}}{i(\omega_0+\omega)} + \frac{e^{i(\omega_0-\omega)t'}}{i(\omega_0-\omega)} \right]_0^t \\ &= -\frac{i}{2} \left[\frac{e^{i(\omega_0+\omega)t} - 1}{\omega_0 + \omega} + \frac{e^{i(\omega_0-\omega)t} - 1}{\omega_0 - \omega} \right] \\ &= -\frac{i}{2} \left[e^{i(\omega_0+\omega)t/2} \frac{e^{i(\omega_0+\omega)t/2} - e^{-i(\omega_0+\omega)t/2}}{\omega_0 + \omega} + e^{i(\omega_0-\omega)t/2} \frac{e^{i(\omega_0-\omega)t/2} - e^{-i(\omega_0-\omega)t/2}}{\omega_0 - \omega} \right] \\ &= -\frac{i}{2} \left[e^{i(\omega_0+\omega)t/2} \frac{2i \sin((\omega_0 + \omega)t/2)}{\omega_0 + \omega} + e^{i(\omega_0-\omega)t/2} \frac{2i \sin((\omega_0 - \omega)t/2)}{\omega_0 - \omega} \right] \\ &= e^{i(\omega_0+\omega)t/2} \frac{\sin((\omega_0 + \omega)t/2)}{\omega_0 + \omega} + e^{i(\omega_0-\omega)t/2} \frac{\sin((\omega_0 - \omega)t/2)}{\omega_0 - \omega}. \end{aligned} \quad (4.29)$$

Enrico Fermi thought about this integral and realized that most of the time it would not be substantial. The only time it's big is when $\omega \approx \omega_0$, and when that's true only the right-hand part is big. So it's legitimate to ignore the left-hand part and write

$$c_b(t) = -\frac{i}{\hbar}eE_0\langle b|\hat{z}|a\rangle e^{i(\omega_0-\omega)t/2} \frac{\sin((\omega_0 - \omega)t/2)}{\omega_0 - \omega}. \quad (4.30)$$

The transition probability is then

$$P_{a \rightarrow b} = \frac{e^2 E_0^2 |\langle b|\hat{z}|a\rangle|^2 \sin^2((\omega_0 - \omega)t/2)}{\hbar^2 (\omega_0 - \omega)^2}. \quad (4.31)$$

This rule, like all rules, has limits on its applicability: we've already mentioned that it applies when the wavelength of light is much larger than an atom, when the light can be treated classically, when $\omega \approx \omega_0$, etc. Most importantly, it applies only when the transition probability is small, because when that probability is large the whole basis of perturbation theory breaks down. You might think that with all these restrictions, it's not a very important result. You'd be wrong. In fact Fermi used it so often that he called it "the golden rule."

4.3 Problems

4.1 Is light a perturbation?

Is it legitimate to use perturbation theory in the case of light absorbed by an atom? After all, we're used to thinking of the light from a powerful laser as a big effect, not a tiny perturbation. However, whether an effect is big or small depends on context. Estimate the maximum electric field due to a laser of XX watts, and the electric field at an electron due to its nearby nucleus. Conclude that while the laser is very powerful on a human scale (and you should *not* stick your eye into a laser beam), it is nevertheless very weak on an atomic scale.

Chapter 5

Quantization of the Electromagnetic Field

5.1 Overview

The classical electromagnetic field is a pair of vector functions, $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$, that evolve in time according to the Maxwell equations

$$\vec{\nabla} \cdot \vec{E}(\vec{r}, t) = \rho(\vec{r}, t)/\epsilon_0 \quad (5.1)$$

$$\vec{\nabla} \cdot \vec{B}(\vec{r}, t) = 0 \quad (5.2)$$

$$\vec{\nabla} \times \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t} \quad (5.3)$$

$$\vec{\nabla} \times \vec{B}(\vec{r}, t) = \mu_0 \vec{J}(\vec{r}, t) + \mu_0 \epsilon_0 \frac{\partial \vec{E}(\vec{r}, t)}{\partial t} \quad (5.4)$$

The energy of the classical EM field interacting with charged particles is

$$H(t) = \frac{1}{2} \int \left(\epsilon_0 \vec{E}^2(\vec{r}, t) + \frac{1}{\mu_0} \vec{B}^2(\vec{r}, t) \right) d^3r + \text{Hamiltonian of particles}. \quad (5.5)$$

This chapter considers the “free field” case, in which there are no particles, no charge, and no current, so the only source of $\vec{E}(\vec{r}, t)$ is the change in $\vec{B}(\vec{r}, t)$, and the only source of $\vec{B}(\vec{r}, t)$ is the change in $\vec{E}(\vec{r}, t)$.

In this case the (classical) EM field Hamiltonian already looks much like the (classical) simple harmonic oscillator Hamiltonian

$$H = \frac{1}{2m} p^2(t) + \frac{m\omega^2}{2} x^2(t)$$

in that it’s the sum of two quadratic functions. The EM field is more complicated in that there’s “an infinite number of oscillators”, one at each point in space, and you have to sum (integrate) over all of them to

find the total energy. This analogy is correct but not immediately useful, because the analogous oscillators adjacent in space are coupled: if there is EM field at one point only, it will spread out into adjacent points.

In this chapter we first look at the classical EM field to represent it in terms of oscillators that *aren't* coupled: The answer turns out to involve oscillators in \vec{k} -space rather than in \vec{r} -space. Once we've represented the EM field as a sum of oscillators it's easy to write a quantum field theory of the EM field, because we already know so much about the quantum properties of oscillators.

5.2 Classical electromagnetism

Free field solutions

One solution of the free field Maxwell equations is the linearly polarized plane wave

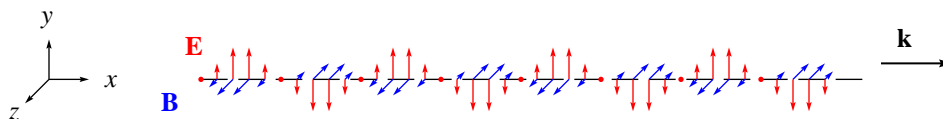
$$\vec{E}(\vec{r}, t) = E_0 \hat{y} \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (5.6)$$

$$\vec{B}(\vec{r}, t) = B_0 \hat{z} \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (5.7)$$

where

$$E_0 \text{ and } \mathbf{k} \text{ are arbitrary, but } \omega = kc \text{ and } B_0 = E_0/c.$$

This plane wave has the familiar snapshot visualization below. (You should be cautious of this visualization — and of all others as well! For example, the visualization suggests strongly that the wave is restricted to the x -axis, whereas in fact it spreads out over all space . . . but if I represented that, the picture would be completely covered with ink! You're supposed to multiply the image in your mind and realize that this is going on at all lines pointing to the right, not just the x -axis. You're also supposed to reject that strong suggestion that the light travels in wavy lines.)



Of course, any superposition of plane waves also solves Maxwell's equations, and these superpositions give rise to the extraordinary variety of free-field electromagnetic phenomena: circular and elliptical polarization; standing waves (optical cavities); maroon, white, and all other colors not part of rainbow; beams of finite width (such as the beams from a flashlight or a laser); pulses; circular waves.

It is not obvious, but nevertheless true, that in fact *every* free-field solution is a superposition of plane waves. This is a consequence of Fourier's theorem.

Fourier analysis

We'll use the conventions

$$F(\vec{r}) = \int f(\vec{k}) e^{+i\vec{k}\cdot\vec{r}} d^3k \quad (5.8)$$

$$f(\vec{k}) = \int F(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} \frac{d^3r}{(2\pi)^3} \quad (5.9)$$

where both \vec{r} and \vec{k} integrals range over all space. One consequence is that if $F(\vec{r})$ is real, then $f^*(\vec{k}) = f(-\vec{k})$. Another (Parseval's theorem) is that

$$\int F^2(\vec{r}) d^3r = \int f(\vec{k}) f(-\vec{k}) \frac{d^3k}{(2\pi)^3}. \quad (5.10)$$

Potentials

The fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ encode the electromagnetic information through six functions of space and time (three components of electric field, three components of magnetic field). The same information is encoded more compactly, through four functions of space and time, in the scalar potential $\phi(\vec{r}, t)$ and the vector potential $\vec{A}(\vec{r}, t)$. These functions satisfy

$$\vec{E}(\vec{r}, t) = -\vec{\nabla}\phi(\vec{r}, t) - \frac{\partial\vec{A}(\vec{r}, t)}{\partial t} \quad (5.11)$$

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t). \quad (5.12)$$

For given fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ there are many possible potentials $\phi(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ that satisfy these equations. For historical reasons, each choice is called a "gauge". In the free-field situation, the easiest choice is the "Coulomb gauge" in which

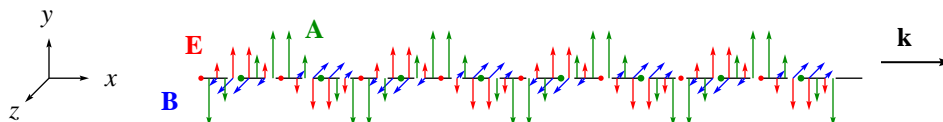
$$\vec{\nabla} \cdot \vec{A}(\vec{r}, t) = 0 \text{ and } \phi(\vec{r}, t) = 0. \quad (5.13)$$

The Maxwell equations for a free field in the Coulomb gauge are equivalent to

$$\vec{\nabla}^2 \vec{A}(\vec{r}, t) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2} = 0. \quad (5.14)$$

For the linearly polarized plane wave shown above, the vector potential in the Coulomb gauge is

$$\vec{A}(\vec{r}, t) = -\frac{E_0}{\omega} \hat{y} \cos(\vec{k} \cdot \vec{r} - \omega t). \quad (5.15)$$



Fourier transform of the vector potential

Let

$$\vec{A}(\vec{r}, t) = \int \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \quad (5.16)$$

What can we say about $\vec{a}(\vec{k}, t)$ in the Coulomb gauge?

1. We know

$$\begin{aligned} 0 &= \vec{\nabla} \cdot \vec{A}(\vec{r}, t) \\ &= \int \vec{\nabla} \cdot \left(\vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \right) \frac{d^3k}{(2\pi)^3} \\ &= \int i\vec{k} \cdot \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3} \end{aligned}$$

and so, for all \vec{k} ,

$$0 = \vec{k} \cdot \vec{a}(\vec{k}, t) \quad (5.17)$$

(because Fourier series are unique). That is, $\vec{a}(\vec{k}, t)$ is perpendicular to \vec{k} .

2. **Electric field.** The electric field is

$$\begin{aligned} \vec{E}(\vec{r}, t) &= -\frac{\partial \vec{A}(\vec{r}, t)}{\partial t} \\ &= -\int \dot{\vec{a}}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \end{aligned}$$

Thus the Fourier transform of the electric field is

$$\vec{e}(\vec{k}, t) = -\dot{\vec{a}}(\vec{k}, t). \quad (5.18)$$

3. **Magnetic field.** The magnetic field is

$$\begin{aligned} \vec{B}(\vec{r}, t) &= \vec{\nabla} \times \vec{A}(\vec{r}, t) \\ &= \int \vec{\nabla} \times \left(\vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \right) \frac{d^3k}{(2\pi)^3} \\ &= \int i\vec{k} \times \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \end{aligned}$$

Thus the Fourier transform of the magnetic field is

$$\vec{b}(\vec{k}, t) = i\vec{k} \times \vec{a}(\vec{k}, t). \quad (5.19)$$

4. **Maxwell equations.** The Maxwell equations, applied to $\vec{A}(\vec{r}, t)$ in the Coulomb gauge, result in

$$0 = \vec{\nabla}^2 \vec{A}(\vec{r}, t) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2}$$

$$\begin{aligned}
&= \nabla^2 \vec{A}(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2} \\
&= \int \left[\vec{a}(\vec{k}, t) \nabla^2 e^{+i\vec{k}\cdot\vec{r}} - \frac{1}{c^2} \frac{d^2 \vec{a}(\vec{k}, t)}{dt^2} e^{+i\vec{k}\cdot\vec{r}} \right] \frac{d^3 k}{(2\pi)^3} \\
&= \int \left[\vec{a}(\vec{k}, t) (-k^2) - \frac{1}{c^2} \ddot{\vec{a}}(\vec{k}, t) \right] e^{+i\vec{k}\cdot\vec{r}} \frac{d^3 k}{(2\pi)^3}.
\end{aligned}$$

Thus, for all \vec{k} ,

$$\ddot{\vec{a}}(\vec{k}, t) = -(kc)^2 \vec{a}(\vec{k}, t). \quad (5.20)$$

The general solution is

$$\vec{a}(\vec{k}, t) = \vec{a}(\vec{k}, 0) e^{-i\omega t}. \quad (5.21)$$

Combining this with equation (5.18) results in

$$\vec{e}(\vec{k}, t) = i\omega \vec{a}(\vec{k}, t). \quad (5.22)$$

Hamiltonian in terms of the Fourier transform of vector potential

Using the property (5.10), the Hamiltonian becomes

$$\begin{aligned}
H(t) &= \frac{1}{2} \int \left(\epsilon_0 \vec{E}^2(\vec{r}, t) + \frac{1}{\mu_0} \vec{B}^2(\vec{r}, t) \right) d^3 r \\
&= \frac{1}{2} \int \left(\epsilon_0 \vec{e}(\vec{k}, t) \cdot \vec{e}(-\vec{k}, t) + \frac{1}{\mu_0} \vec{b}(\vec{k}, t) \cdot \vec{b}(-\vec{k}, t) \right) \frac{d^3 k}{(2\pi)^3} \\
&= \frac{1}{2} \int \left(\epsilon_0 \vec{e}(\vec{k}, t) \cdot \vec{e}^*(\vec{k}, t) + \frac{1}{\mu_0} \vec{b}(\vec{k}, t) \cdot \vec{b}^*(\vec{k}, t) \right) \frac{d^3 k}{(2\pi)^3}
\end{aligned}$$

where in the last step we have used the fact that $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ are real.

Now, using first equation (5.22) and then equation (5.21),

$$\vec{e}(\vec{k}, t) \cdot \vec{e}^*(\vec{k}, t) = \omega^2 \vec{a}(\vec{k}, t) \cdot \vec{a}^*(\vec{k}, t) = \omega^2 \vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0).$$

Similarly, using both equation (5.19) and the transverse property (5.17),

$$\vec{b}(\vec{k}, t) \cdot \vec{b}^*(\vec{k}, t) = k^2 \vec{a}(\vec{k}, t) \cdot \vec{a}^*(\vec{k}, t) = k^2 \vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0).$$

Thus

$$H(t) = \frac{1}{2} \int \left(\epsilon_0 \omega^2 + \frac{k^2}{\mu_0} \right) \left(\vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0) \right) \frac{d^3 k}{(2\pi)^3} \quad (5.23)$$

and we find what we've suspected all along, namely that $H(t)$ is independent of time. We write $\vec{a}(\vec{k}, t) = \vec{a}_{\mathbf{k}}$ and allow the symbol $\vec{a}_{\mathbf{k}}$ to either be a function of time or a constant, depending on context. A little manipulation shows that $\epsilon_0 \omega^2 + k^2/\mu_0 = 2k^2/\mu_0$, whence

$$H = \frac{1}{\mu_0} \int k^2 \vec{a}_{\mathbf{k}} \cdot \vec{a}_{\mathbf{k}}^* \frac{d^3 k}{(2\pi)^3} = \frac{1}{\mu_0} \sum_{\alpha} \int k^2 a_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}^* \frac{d^3 k}{(2\pi)^3} \quad (5.24)$$

where the index α stands for one of the two polarization directions. This is our formula for the field Hamiltonian as an integral over all \vec{k} and sum over each of the two polarization directions at each \vec{k} . We call this a “sum over modes”.

Each mode as a simple harmonic oscillator

The Hamiltonian is the sum over all modes of the energy contribution from each mode, namely

$$\frac{1}{\mu_0} k^2 a_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}^*. \quad (5.25)$$

Our goal is to write this energy in the form of a simple harmonic oscillator. Since we’re dealing with only one particular mode, I’ll drop the elaborate subscript and write our desired form as

$$\frac{1}{\mu_0} k^2 a(t) a^*(t) = \frac{1}{2} [p^2(t) + \omega^2 x^2(t)] \quad (5.26)$$

where $x(t)$ and $p(t)$ are real and $\dot{x}(t) = p(t)$. This is the Hamiltonian for a simple harmonic oscillator of mass 1.

There is more than one way to define $x(t)$ and $p(t)$ that satisfy these conditions. My favorite choice starts with

$$x(t) = D (a(t) + a^*(t)),$$

where D is to be determined. Then (remembering that $\dot{a}(t) = -i\omega a(t)$)

$$p(t) = \dot{x}(t) = -i\omega D (a(t) - a^*(t))$$

so that

$$\begin{aligned} \frac{1}{\mu_0} k^2 a(t) a^*(t) &= \frac{1}{2} [-\omega^2 D^2 (a^2 - 2aa^* + a^{*2}) + \omega^2 D^2 (a^2 + 2aa^* + a^{*2})] \\ &= \frac{1}{2} [4\omega^2 D^2 aa^*] \\ &= 2k^2 c^2 D^2 aa^*. \end{aligned}$$

This choice works when $D = 1/\sqrt{2\mu_0 c^2}$.

We’ve done it. We’ve written the Maxwell equations for the free electromagnetic field in Hamiltonian form, and the Hamiltonian turns out to be nothing but a sum of simple harmonic oscillators, one for each mode (“mode” meaning a given one of the two polarizations at a given wavevector \vec{k}). The correspondence between the vector potential strength $a(t)$ of a given mode and the position and momentum of this analogous simple harmonic oscillator is given through

$$x(t) = \frac{1}{\sqrt{2\mu_0 c^2}} (a(t) + a^*(t)) \quad (5.27)$$

$$p(t) = -i \frac{\omega}{\sqrt{2\mu_0 c^2}} (a(t) - a^*(t)) \quad (5.28)$$

$$a(t) = \frac{1}{2} \sqrt{2\mu_0 c^2} \left(x(t) + \frac{i}{\omega} p(t) \right). \quad (5.29)$$

For completeness, we add that the field amplitudes in this mode are

$$e(t) = i\omega a(t) \quad (5.30)$$

$$b(t) = ika(t) = e(t)/c. \quad (5.31)$$

Everything we've done in this section has been classical. In fact, what we've done is just like solving a coupled oscillator problem in classical mechanics: We've found the independent normal modes of the coupled system, each one of which behaves like a single harmonic oscillator. *Any* time development of the whole system can be expressed as a sum over (superposition of) the normal modes. You will recall from your classical mechanics course that sometimes it's not obvious, looking at the motion of the whole system, that this motion can be expressed as nothing more than the sum of simple harmonic oscillators, but it's true whether obvious or not.

5.3 Quantal electromagnetism — What is a photon?

Once classical electrodynamics has been cast into this “sum over oscillators” form, it's easy to see how to invent quantum electrodynamics: Change the functions $x(t)$ and $p(t)$ into operators \hat{x} and \hat{p} , subject to the commutation rule $[\hat{x}, \hat{p}] = i\hbar$. Here are a few notable features of this quantization:

The traditional SHO lowering operator is

$$\bar{a} = \sqrt{\frac{\omega}{2\hbar}} \left(\hat{x} + \frac{i}{\omega} \hat{p} \right),$$

so this operator corresponds to the vector potential amplitude function $a(t)$ through

$$\bar{a} \iff \sqrt{\frac{\omega}{\hbar\mu_0 c^2}} a(t).$$

The classical mode Hamiltonian

$$\frac{1}{\mu_0} k^2 a(t) a^*(t)$$

thus corresponds to

$$\hbar\omega \bar{a} \bar{a}^\dagger.$$

(I'm not sure how the traditional $\frac{1}{2}\hbar\omega$ ground state energy fits into this prescription.) And the field amplitude operators are

$$\begin{aligned} \hat{e} &= i\sqrt{\hbar\omega} \sqrt{\mu_0 c^2} \bar{a} \\ \hat{b} &= \hat{e}/c \end{aligned}$$

The outline is clear: States with definite energy are eigenstates of $\bar{a} \bar{a}^\dagger$. States with definite field amplitude are eigenstates of \bar{a} . These operators don't commute:

$$[\bar{a}, \bar{a} \bar{a}^\dagger] = \bar{a}[\bar{a}, \bar{a}^\dagger] + [\bar{a}, \bar{a}] \bar{a}^\dagger = \bar{a},$$

so you can have an energy, or you can have a field strength, but you can't have both.

To say “the \mathbf{k}, α mode is in energy eigenstate n ” is pronounced “there are n photons of mode \mathbf{k}, α ”. Don't think that photons are hard round marbles. Don't think that photons are wavepackets of EM field.

Remember from classical mechanics that it can be hard to visualize the motion of a set of coupled oscillators as a superposition of normal modes. And remember from earlier in quantum mechanics that it can be hard to visualize the time development of a quantal state in a simple harmonic oscillator as a superposition of energy states. For the quantized EM field, we must do both at once, so it's no surprise that photons are hard to visualize.

There exist states of the EM field that have definite position (the “hard round marble” idea) — but these states do not have energy at all, much less energy $E = \hbar\omega$. There exist states of the EM field that have definite electric and magnetic field — but these states do not have either energy or position.

When he was my teacher, I heard Kurt Gottfried call this principle of “energy or field strength, but not both” the most astounding, “in your face”, instance of complementarity that he knew of. His written comment is a bit more reserved (pages 14–15): “It is impossible to determine the field strengths and the number of photons simultaneously. This shows the mutual exclusiveness of the photon and field-strength descriptions of electromagnetic phenomena and is a very illuminating example of Bohr's principle of complementarity.”

Appendix A

The Spherical Harmonics

A “function on the unit sphere” is a function $f(\theta, \phi)$. Another convenient variable is $\mu = \cos \theta$. “Integration over the unit sphere” means

$$\int d\Omega f(\theta, \phi) = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi f(\theta, \phi) = \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi f(\theta, \phi).$$

$$\nabla^2 Y_\ell^m(\theta, \phi) = -\frac{1}{r^2} \ell(\ell + 1) Y_\ell^m(\theta, \phi) \quad (\text{A.1})$$

$$\int Y_{\ell'}^{m'*}(\theta, \phi) Y_\ell^m(\theta, \phi) d\Omega = \delta_{\ell', \ell} \delta_{m', m} \quad (\text{A.2})$$

$$f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell, m} Y_\ell^m(\theta, \phi) \quad \text{where} \quad (\text{A.3})$$

$$f_{\ell, m} = \int Y_\ell^{m*}(\theta, \phi) f(\theta, \phi) d\Omega \quad (\text{A.4})$$

In the table, square roots are always taken to be positive.

$$\begin{aligned}
Y_0^0(\mu, \phi) &= \left(\frac{1}{2^2\pi}\right)^{1/2} \\
Y_1^0(\mu, \phi) &= \left(\frac{3}{2^2\pi}\right)^{1/2} \mu &= \left(\frac{3}{2^2\pi}\right)^{1/2} \frac{z}{r} \\
Y_1^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3}{2^3\pi}\right)^{1/2} \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3}{2^3\pi}\right)^{1/2} \frac{1}{r} (x \pm iy) \\
Y_2^0(\mu, \phi) &= \left(\frac{5}{2^4\pi}\right)^{1/2} (3\mu^2 - 1) &= \left(\frac{5}{2^4\pi}\right)^{1/2} \left(3\frac{z^2}{r^2} - 1\right) \\
Y_2^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3 \cdot 5}{2^3\pi}\right)^{1/2} \mu \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3 \cdot 5}{2^3\pi}\right)^{1/2} \frac{z}{r^2} (x \pm iy) \\
Y_2^{\pm 2}(\mu, \phi) &= \left(\frac{3 \cdot 5}{2^5\pi}\right)^{1/2} (1-\mu^2) e^{\pm 2i\phi} &= \left(\frac{3 \cdot 5}{2^5\pi}\right)^{1/2} \frac{1}{r^2} (x \pm iy)^2 \\
Y_3^0(\mu, \phi) &= \left(\frac{7}{2^4\pi}\right)^{1/2} (5\mu^3 - 3\mu) &= \left(\frac{7}{2^4\pi}\right)^{1/2} \left(5\frac{z^3}{r^3} - 3\frac{z}{r}\right) \\
Y_3^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3 \cdot 7}{2^6\pi}\right)^{1/2} (5\mu^2 - 1) \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3 \cdot 7}{2^6\pi}\right)^{1/2} \left(5\frac{z^2}{r^2} - 1\right) \frac{1}{r} (x \pm iy) \\
Y_3^{\pm 2}(\mu, \phi) &= \left(\frac{3 \cdot 5 \cdot 7}{2^5\pi}\right)^{1/2} \mu (1-\mu^2) e^{\pm 2i\phi} &= \left(\frac{3 \cdot 5 \cdot 7}{2^5\pi}\right)^{1/2} \frac{z}{r^3} (x \pm iy)^2 \\
Y_3^{\pm 3}(\mu, \phi) &= \mp \left(\frac{5 \cdot 7}{2^6\pi}\right)^{1/2} (1-\mu^2) \sqrt{1-\mu^2} e^{\pm 3i\phi} &= \mp \left(\frac{5 \cdot 7}{2^6\pi}\right)^{1/2} \frac{1}{r^3} (x \pm iy)^3
\end{aligned}$$

Appendix B

Radial Wavefunctions for the Coulomb Problem

Based on Griffiths, page 154, but with scaled variables and with integers factorized.

$$R_{10}(r) = 2e^{-r}$$

$$R_{20}(r) = \frac{1}{\sqrt{2}} \left(1 - \frac{1}{2}r\right) e^{-r/2}$$

$$R_{21}(r) = \frac{1}{\sqrt{2^3 \cdot 3}} r e^{-r/2}$$

$$R_{30}(r) = \frac{2}{\sqrt{3^3}} \left(1 - \frac{2}{3}r + \frac{2}{3^3}r^2\right) e^{-r/3}$$

$$R_{31}(r) = \frac{2^3}{3^3 \sqrt{2 \cdot 3}} \left(1 - \frac{1}{2 \cdot 3}r\right) r e^{-r/3}$$

$$R_{32}(r) = \frac{2^2}{3^4 \sqrt{2 \cdot 3 \cdot 5}} r^2 e^{-r/3}$$

$$R_{40}(r) = \frac{1}{2^2} \left(1 - \frac{3}{2^2}r + \frac{1}{2^3}r^2 - \frac{1}{2^6 \cdot 3}r^3\right) e^{-r/4}$$

$$R_{41}(r) = \frac{\sqrt{5}}{2^4 \sqrt{3}} \left(1 - \frac{1}{2^2}r + \frac{1}{2^4 \cdot 5}r^2\right) r e^{-r/4}$$

$$R_{42}(r) = \frac{1}{2^6 \sqrt{5}} \left(1 - \frac{1}{2^2 \cdot 3}r\right) r^2 e^{-r/4}$$

$$R_{43}(r) = \frac{1}{2^8 \cdot 3 \sqrt{5 \cdot 7}} r^3 e^{-r/4}$$

Appendix C

Cheat Sheet

Delta functions:

$$\int_{-\infty}^{+\infty} e^{ikx} dk = 2\pi\delta(x) \quad (\text{C.1})$$

$$\int_{-\infty}^{+\infty} e^{i(p/\hbar)x} dp = 2\pi\hbar\delta(x) \quad (\text{C.2})$$

$$\int_{-\infty}^{+\infty} e^{i\omega t} d\omega = 2\pi\delta(t) \quad (\text{C.3})$$

Fourier transforms:

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x) e^{-i(p/\hbar)x} dx \quad (\text{C.4})$$

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\psi}(p) e^{+i(p/\hbar)x} dp \quad (\text{C.5})$$

$$\tilde{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \quad (\text{C.6})$$

$$f(t) = \int_{-\infty}^{+\infty} \tilde{f}(\omega) e^{+i\omega t} \frac{d\omega}{2\pi} \quad (\text{C.7})$$

Gaussian integrals:

$$\int_{-\infty}^{+\infty} e^{ax^2+bx} dx = \sqrt{\frac{\pi}{-a}} e^{-b^2/4a} \quad \Re\{a\} \leq 0 \quad (\text{C.8})$$

$$\frac{\int_{-\infty}^{+\infty} x^2 e^{-x^2/2\sigma^2} dx}{\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2} dx} = \sigma^2 \quad (\text{C.9})$$

Time development:

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \quad (\text{C.10})$$

$$\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x}, t) \quad (\text{C.11})$$

$$|\psi(t)\rangle = \sum_n e^{-(i/\hbar)E_n t} c_n |\eta_n\rangle \quad (\text{C.12})$$

$$\frac{d\langle\hat{A}\rangle}{dt} = -\frac{i}{\hbar}\langle[\hat{A}, \hat{H}]\rangle \quad (\text{C.13})$$

Momentum:

$$\hat{p} \iff -i\hbar\frac{\partial}{\partial x} \quad (\text{C.14})$$

$$[\hat{x}, \hat{p}] = i\hbar \quad (\text{C.15})$$

Dimensions:

$$\psi(x) \text{ has dimensions } [\text{length}]^{-1/2} \quad (\text{C.16})$$

$$\psi(\mathbf{x}_1, \mathbf{x}_2) \text{ has dimensions } [\text{length}]^{-6/2} \quad (\text{C.17})$$

$$\tilde{\psi}(p) \text{ has dimensions } [\text{momentum}]^{-1/2} \quad (\text{C.18})$$

$$\hbar \text{ has dimensions } [\text{length} \times \text{momentum}] \text{ or } [\text{energy} \times \text{time}] \quad (\text{C.19})$$

Energy eigenfunction sketching: (one dimension)

$$n\text{th excited state has } n \text{ nodes} \quad (\text{C.20})$$

$$\text{if classically allowed: regions of high } V(x) \text{ have large amplitude and long wavelength} \quad (\text{C.21})$$

$$\text{if classically forbidden: regions of high } V(x) \text{ have faster cutoff} \quad (\text{C.22})$$

Infinite square well: (width L)

$$\eta_n(x) = \sqrt{2/L} \sin kx \quad k = n\pi/L \quad n = 1, 2, 3, \dots \quad (\text{C.23})$$

$$E_n = \frac{\hbar^2 k^2}{2m} = n^2 \frac{\pi^2 \hbar^2}{2mL^2} \quad (\text{C.24})$$

Simple harmonic oscillator: ($V(x) = \frac{1}{2}Kx^2$, $\omega = \sqrt{K/m}$)

$$E_n = (n + \frac{1}{2})\hbar\omega \quad n = 0, 1, 2, \dots \quad (\text{C.25})$$

$$[\hat{a}, \hat{a}^\dagger] = \hat{1} \quad (\text{C.26})$$

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}) \quad (\text{C.27})$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad (\text{C.28})$$

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (\text{C.29})$$

$$\hat{x} = \sqrt{\hbar/2m\omega}(\hat{a} + \hat{a}^\dagger) \quad (\text{C.30})$$

$$\hat{p} = -i\sqrt{m\hbar\omega/2}(\hat{a} - \hat{a}^\dagger) \quad (\text{C.31})$$

Coulomb problem:

$$E_n = -\frac{\text{Ry}}{n^2} \quad \text{Ry} = \frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 = 13.6 \text{ eV} \quad (\text{C.32})$$

$$a_0 = \frac{\hbar^2}{m} \left(\frac{4\pi\epsilon_0}{e^2} \right) = 0.0529 \text{ nm} \quad (\text{Bohr radius}) \quad (\text{C.33})$$

$$\tau_0 = \frac{\hbar}{2\text{Ry}} = 0.0242 \text{ fsec} \quad (\text{characteristic time}) \quad (\text{C.34})$$

Angular Momentum:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \text{ and cyclic permutations} \quad (\text{C.35})$$

The eigenvalues of \hat{J}^2 are

$$\hbar^2 j(j+1) \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \quad (\text{C.36})$$

For a given j , the eigenvalues of \hat{J}_z are

$$\hbar m \quad m = -j, -j+1, \dots, j-1, j. \quad (\text{C.37})$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y \quad (\text{C.38})$$

by

$$\hat{J}_+ |j, m\rangle = \hbar \sqrt{j(j+1) - m(m+1)} |j, m+1\rangle \quad (\text{C.39})$$

$$\hat{J}_- |j, m\rangle = \hbar \sqrt{j(j+1) - m(m-1)} |j, m-1\rangle. \quad (\text{C.40})$$