

DEPARTMENT OF PHYSICS
UNIVERSITY OF CAMBRIDGE

PART II PHYSICS

Advanced Quantum Physics

Handout 3

Contents

| | |
|---|------------|
| 8 Identical Particles | 70 |
| 8.1 Quantum statistics | 70 |
| 8.2 Space and spin wavefunctions | 72 |
| 8.3 Physical consequences of particle statistics | 73 |
| 8.4 Many-body systems | 76 |
| 8.4.1 Non-interacting Fermi gas | 76 |
| 8.4.2 Non-interacting Bose gas | 77 |
| 9 Atomic structure | 81 |
| 9.1 The non-relativistic hydrogen atom | 81 |
| 9.1.1 The Zeeman effect | 82 |
| 9.2 The “real” hydrogen atom | 83 |
| 9.2.1 Relativistic correction to the kinetic energy | 84 |
| 9.2.2 Spin-orbit coupling | 84 |
| 9.2.3 Darwin term | 87 |
| 9.2.4 INFO: Lamb shift | 88 |
| 9.2.5 Hyperfine structure | 89 |
| 9.3 Multi-electron atoms | 90 |
| 9.3.1 Central field approximation | 91 |
| 9.3.2 Spin-orbit coupling | 96 |
| 9.3.3 INFO: jj coupling scheme | 99 |
| 9.3.4 Zeeman effect | 101 |
| 10 From molecules to solids | 103 |
| 10.1 The H_2^+ ion | 104 |
| 10.2 The H_2 molecule | 106 |
| 10.3 From molecules to solids | 109 |
| 11 Time-dependent perturbation theory | 114 |
| 11.1 Time-dependent potentials: general formalism | 114 |
| 11.1.1 Dynamics of a driven two-level system | 115 |
| 11.1.2 Paramagnetic resonance | 116 |
| 11.2 Time-dependent perturbation theory | 117 |
| 11.3 “Sudden” perturbation | 119 |
| 11.3.1 Harmonic perturbations: Fermi’s Golden Rule | 119 |
| 11.3.2 INFO: Harmonic perturbations: second-order transitions | 121 |
| 12 Scattering theory | 122 |
| 12.1 Basics | 123 |
| 12.2 The Born approximation | 124 |

| | |
|--|------------|
| 13 Radiative transitions | 126 |
| 13.1 Coupling of matter to the electromagnetic field | 126 |
| 13.1.1 Quantum fields | 127 |
| 13.1.2 Spontaneous emission | 127 |
| 13.1.3 Absorption and stimulated emission | 128 |
| 13.1.4 Einstein's A and B coefficients | 129 |
| 13.2 Selection rules | 130 |
| 13.3 Lasers | 132 |
| 13.3.1 Operating principles of a laser | 134 |
| 13.3.2 Gain mechanism | 134 |
| 14 Field theory: from phonons to photons | 136 |
| 14.1 Quantization of the classical atomic chain | 136 |
| 14.1.1 INFO: Classical chain | 136 |
| 14.1.2 INFO: Quantum chain | 139 |
| 14.2 Quantum electrodynamics | 142 |
| 14.2.1 INFO: Classical theory of the electromagnetic field | 143 |
| 14.2.2 Quantum field theory of the electromagnetic field | 145 |
| 14.2.3 Fock states | 146 |
| 14.2.4 Coherent states | 147 |
| 14.2.5 Non-classical light | 148 |

Chapter 8

Identical Particles

Until now, most of our focus has been on the quantum mechanical behaviour of *individual* particles, or problems which can be “factorized” into independent single-particle systems.¹ However, most physical systems of interest involve the interaction of large numbers of particles; electrons in a solid, atoms in a gas, etc. In classical mechanics, particles are always distinguishable in the sense that, at least formally, their “trajectories” through phase space can be traced and their identity disclosed. However, in quantum mechanics, the intrinsic uncertainty in position, embodied in Heisenberg’s principle, demands a careful and separate consideration of distinguishable and indistinguishable particles. In the present section, we will consider how to formulate the wavefunction of many-particle systems, and address some of the (sometimes striking and often counter-intuitive) implications of particle indistinguishability.

8.1 Quantum statistics

Consider then two indistinguishable particles confined to a box in one-dimension. Here, by indistinguishable, we mean that the particles can not be discriminated by some internal quantum number. For example, we might have two electrons of the same spin. The normalized two-particle wavefunction $\psi(x_1, x_2)$, which gives the probability $|\psi(x_1, x_2)|^2 dx_1 dx_2$ of finding simultaneously one particle in the interval x_1 to $x_1 + dx_1$ and another between x_2 to $x_2 + dx_2$, only makes sense if $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$, since we can’t know which of the two indistinguishable particles we are finding where. It follows from this that the wavefunction can exhibit two (and, generically, only two) possible symmetries under exchange: $\psi(x_1, x_2) = \psi(x_2, x_1)$ or $\psi(x_1, x_2) = -\psi(x_2, x_1)$.² If two identical particles have a symmetric wavefunction in some state, particles of that type always have symmetric wavefunctions, and are called **bosons**. (If in some other state they had an antisymmetric wavefunction, then a linear

¹For example, the treatment of the hydrogen atom involves the separation of the system into centre of mass and relative motion. Each could be referred to an effective single-particle dynamics.

²We could in principle have $\psi(x_1, x_2) = e^{i\alpha}\psi(x_2, x_1)$, with α a constant phase. However, in this case we would not recover the original wavefunction on exchanging the particles twice. Curiously, in some two-dimensional theories used to describe the fractional quantum Hall effect, there exist collective excitations of the electron system — called **anyons** — that do have this kind of property. For a discussion of this point, one may refer to the seminal paper of J. M. Leinaas and J. Myrheim, *On the theory of identical particles*. Il Nuovo Cimento **B37**, 1-23 (1977). Such anyonic systems have been proposed as a strong candidate for the realization of quantum computation. For a pedagogical discussion, we refer to an entertaining discussion by C. Nayak, S. H. Simon, A. Stern, M. Freedman, S. Das Sarma, *Non-Abelian Anyons and Topological Quantum Computation*, Rev. Mod. Phys. **80**, 1083 (2008). However, all ordinary “fundamental” particles are either bosons or fermions.

superposition of those states would be neither symmetric nor antisymmetric, and so could not satisfy the relation $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$.) Similarly, particles having antisymmetric wavefunctions are called **fermions**.

To construct wavefunctions for three or more fermions, let first suppose that the particles do not interact with each other, and are confined by a spin-independent potential, such as the Coulomb field of a nucleus. In this case, the Hamiltonian will be symmetric in the fermion degrees of freedom,

$$\hat{H} = \frac{\hat{\mathbf{p}}_1^2}{2m} + \frac{\hat{\mathbf{p}}_2^2}{2m} + \frac{\hat{\mathbf{p}}_3^2}{2m} \dots + V(\mathbf{r}_1) + V(\mathbf{r}_2) + V(\mathbf{r}_3) + \dots ,$$

and the solutions of the Schrödinger equation will be products of eigenfunctions of the single-particle Hamiltonian $\hat{H}_s = \hat{\mathbf{p}}^2/2m + V(\mathbf{r})$. However, single products such as $\psi_a(1)\psi_b(2)\psi_c(3)$ do not have the required antisymmetry property under the exchange of any two particles. (Here a, b, c, \dots label the single-particle eigenstates of \hat{H}_s , and 1, 2, 3,... denote *both* space and spin coordinates of single particles, i.e. 1 stands for (\mathbf{r}_1, s_1) , etc.)

We could achieve the necessary antisymmetrization for particles 1 and 2 by subtracting the same product wavefunction with the particles 1 and 2 interchanged, i.e. $\psi_a(1)\psi_b(2)\psi_c(3) \mapsto (\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1))\psi_c(3)$, ignoring the overall normalization for now. However, the wavefunction needs to be antisymmetrized with respect to *all* possible particle exchanges. So, for 3 particles, we must add together all $3!$ permutations of 1, 2, 3 in the state a, b, c , with a factor -1 for each particle exchange necessary to get to a particular ordering from the original ordering of 1 in a , 2 in b , and 3 in c . In fact, such a sum over permutations is precisely the *definition* of the determinant. So, with the appropriate normalization factor:

$$\psi_{abc}(1, 2, 3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} .$$

The determinantal form makes clear the antisymmetry of the wavefunction with respect to exchanging any two of the particles, since exchanging two rows of a determinant multiplies it by -1 . We also see from the determinantal form that the three states a, b, c *must* all be different, for otherwise two columns would be identical, and the determinant would be zero. This is just the manifestation of Pauli's exclusion principle: no two fermions can be in the same state.

Although these determinantal wavefunctions (known as **Slater determinants**), involving superpositions of single-particle states, are only strictly correct for non-interacting fermions, they provide a useful platform to describe electrons in atoms (or in a metal), with the electron-electron repulsion approximated by a single-particle potential. For example, the Coulomb field in an atom, as seen by the outer electrons, is partially shielded by the inner electrons, and a suitable $V(\mathbf{r})$ can be constructed self-consistently, by computing the single-particle eigenstates and finding their associated charge densities. (We shall discuss this "Hartree" approximation later in the course, when we study multielectron atoms.)

In the bosonic system, the corresponding many-particle wavefunction must be symmetric under particle exchange. We can obtain such a state by expanding all of the contributing terms from the Slater determinant and setting all of the signs to be positive. In other words, the bosonic wave function describes the uniform (equal phase) superposition of all possible permutations of product states. In this case, there is no requirement that the orbitals be different.

(There is no exclusion principle for bosons.) One important example is the groundstate wavefunction of a collection of non-interacting gas of bosons, in which all bosons occupy the lowest energy single particle wavefunction, $\psi_0(\mathbf{r})$. The many-particle state (for N indistinguishable bosons) is

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i=1}^N \psi_0(\mathbf{r}_i) \quad (8.1)$$

This is the wavefunction for a **Bose-Einstein condensate**, with all particles condensed in the state $\psi_0(\mathbf{r})$. It describes a non-interacting gas of (spinless) bosons at zero temperature: *i.e.* the many-particle groundstate.

▷ INFO. As you will learn in the Part II Statistical Physics course, Bose-Einstein condensation is defined by the condition that a finite fraction of the particles occupy the same quantum state. For a non-interacting gas of bosons in thermal equilibrium, this occurs at temperatures below a critical temperature

$$k_B T_c = \alpha \frac{\hbar^2}{m} \bar{n}^{2/3}, \quad \alpha = \frac{2\pi}{\zeta^{2/3}(3/2)},$$

where \bar{n} is the mean particle density of the 3D gas. Roughly speaking, T_c is the temperature below which the thermal de Broglie wavelength λ_T is larger than the mean interparticle spacing $\lambda_T \equiv \frac{\hbar}{\sqrt{2\pi m k_B T}} \gtrsim \frac{1}{\bar{n}^{1/3}}$.

8.2 Space and spin wavefunctions

Although the methodology for constructing a basis of many-particle states outlined above is generic, it is not particularly convenient when the Hamiltonian is spin-independent. In this case we can simplify the structure of the wavefunction by factorizing the spin and spatial components. Suppose we have two electrons (i.e. fermions) in some spin-independent potential $V(\mathbf{r})$. We know that the two-electron wavefunction must be antisymmetric under particle exchange. Since the Hamiltonian has no spin-dependence, we must be able to construct a set of common eigenstates of the Hamiltonian, the total spin, and the z -component of the total spin. For two electrons, there are four basis states in the spin space, the $S = 0$ spin singlet state, $|\chi_{S=0, S_z=0}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1 \downarrow_2\rangle - |\downarrow_1 \uparrow_2\rangle)$, and the three $S = 1$ spin triplet states,

$$|\chi_{11}\rangle = |\uparrow_1 \uparrow_2\rangle, \quad |\chi_{10}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1 \downarrow_2\rangle + |\downarrow_1 \uparrow_2\rangle), \quad |\chi_{1,-1}\rangle = |\downarrow_1 \downarrow_2\rangle.$$

Here the first arrow in the ket refers to the spin of particle 1, the second to particle 2.

▷ EXERCISE. By way of revision, it is helpful to recapitulate the discussion of the addition of spin $s = 1/2$ angular momenta. By setting $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$,³ where \mathbf{S}_1 and \mathbf{S}_2 are two spin 1/2 degrees of freedom, construct the matrix elements of the total spin operator \mathbf{S}^2 for the four basis states, $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$, and $|\downarrow\downarrow\rangle$. From the matrix representation of \mathbf{S}^2 , determine the four eigenstates. Show that one corresponds to a total spin zero state and three correspond to spin 1.

Hint: begin by proving that, for two spin $s = 1/2$ degree of freedom, $\mathbf{S}^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 = 2 \times s(s+1)\hbar^2 + 2S_1^z S_2^z + S_1^+ S_2^- + S_1^- S_2^+$.

³Here, for simplicity, we have chosen not to include hats on the spin angular momentum operators.

It is evident that the spin singlet wavefunction is antisymmetric under the exchange of two particles, while the spin triplet wavefunction is symmetric. For a general state, the total wavefunction for the two electrons in a common eigenstate of S^2 , S_z and the Hamiltonian \hat{H} then has the form:

$$\Psi(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2) = \psi(\mathbf{r}_1, \mathbf{r}_2)\chi(s_1, s_2),$$

where $\chi(s_1, s_2) = \langle s_1, s_2 | \chi \rangle$. For two electron degrees of freedom, the total wavefunction, Ψ , must be antisymmetric under exchange. It follows that a pair of electrons in the spin *singlet* state must have a *symmetric* spatial wavefunction, $\psi(\mathbf{r}_1, \mathbf{r}_2)$, whereas electrons in the spin *triplet* states must have an *antisymmetric* spatial wavefunction. Before discussing the physical consequences of this symmetry, let us mention how this scheme generalizes to more particles.

▷ INFO. **Symmetry of three-electron wavefunctions:** Unfortunately, in seeking a factorization of the Slater determinant into a product of spin and spatial components for three electrons, things become more challenging. There are now $2^3 = 8$ basis states in the spin space. Four of these are accounted for by the spin $3/2$ state with $S_z = 3/2, 1/2, -1/2, -3/2$. Since all spins are aligned, this is evidently a symmetric state, so must be multiplied by an antisymmetric spatial wavefunction, itself a determinant. So far so good. But the other four states involve two pairs of total spin $1/2$ states built up of a singlet and an unpaired spin. They are orthogonal to the symmetric spin $3/2$ state, so they can't be symmetric. But they can't be antisymmetric either, since in each such state, two of the spins must be pointing in the same direction! An example of such a state is presented by $|\chi\rangle = |\uparrow_1\rangle \otimes \frac{1}{\sqrt{2}}(|\uparrow_2\downarrow_3\rangle - |\downarrow_2\uparrow_3\rangle)$. Evidently, this must be multiplied by a spatial wavefunction symmetric in 2 and 3. But to recover a total wave function with *overall* antisymmetry it is necessary to add more terms:

$$\Psi(1, 2, 3) = \chi(s_1, s_2, s_3)\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \chi(s_2, s_3, s_1)\psi(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) + \chi(s_3, s_1, s_2)\psi(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2).$$

Requiring the spatial wavefunction $\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ to be symmetric in 2, 3 is sufficient to guarantee the overall antisymmetry of the total wavefunction Ψ .⁴ For more than three electrons, similar considerations hold. The mixed symmetries of the spatial wavefunctions and the spin wavefunctions which together make a totally antisymmetric wavefunction are quite complex, and are described by **Young diagrams (or tableaux)**.⁵ A discussion of this scheme reaches beyond the scope of these lectures.

▷ EXERCISE. A hydrogen atom consists of two fermions: the proton and the electron. By considering the wavefunction of two non-interacting hydrogen atoms under exchange, show that the atom transforms as a boson. In general, if the number of fermions in a composite particle is odd, then it is a fermion, while if even it is a boson. Quarks are fermions: baryons consist of three quarks and so translate to fermions while mesons consist of two quarks and translate to bosons.

8.3 Physical consequences of particle statistics

The overall antisymmetry demanded by the many-fermion wavefunction has important physical implications. In particular, it determines the magnetic

⁴Particle physics enthusiasts might be interested to note that functions exactly like this arise in constructing the spin/flavour wavefunction for the proton in the quark model (Griffiths, Introduction to Elementary Particles, page 179).

⁵For a simple introduction, see Sakurai's textbook (section 6.5) or chapter 63 of the text on quantum mechanics by Landau and Lifshitz.

properties of atoms. The magnetic moment of the electron is aligned with its spin, and even though the spin variables do not appear in the Hamiltonian, the energy of the eigenstates depends on the relative spin orientation. This arises from the electrostatic repulsion between electrons. In the spatially antisymmetric state, the probability of electrons coinciding at the same position necessarily vanishes. Moreover, the nodal structure demanded by the antisymmetry places the electrons further apart on average than in the spatially symmetric state. Therefore, the electrostatic repulsion raises the energy of the spatially symmetric state above that of the spatially antisymmetric state. It therefore follows that the lower energy state has the electron spins pointing in the same direction. This argument is still valid for more than two electrons, and leads to **Hund's rule** for the magnetization of incompletely filled inner shells of electrons in transition metal and rare earths atoms. (We shall discuss this later, when we consider multielectron atoms.) This is the first step in understanding ferromagnetism (see 8.4.2).

A gas of hydrogen molecules provides another manifestation of wavefunction antisymmetry. In particular, the specific heat depends sensitively on whether the two protons (spin 1/2) in H_2 have their spins parallel or antiparallel, even though that alignment involves only a very tiny interaction energy. If the proton spins occupy a spin singlet configuration, the molecule is called **parahydrogen** while the triplet states are called **orthoohydrogen**. These two distinct gases are remarkably stable - in the absence of magnetic impurities, para-ortho transitions take weeks.

The actual energy of interaction of the proton spins is of course completely negligible in the specific heat. The important contributions to the specific heat are the usual kinetic energy term, and the rotational energy of the molecule. This is where the overall (space \times spin) antisymmetric wavefunction for the protons plays a role. Recall that the parity of a state with rotational angular momentum ℓ is $(-1)^\ell$. Therefore, parahydrogen, with an antisymmetric proton spin wavefunction, must have a symmetric proton *spatial* wavefunction, and so can only have even values of the rotational angular momentum. Orthohydrogen can only have odd values. The energy of the rotational level with angular momentum ℓ is $E_\ell^{\text{tot}} = \hbar^2 \ell(\ell+1)/I$, where I denotes the moment of inertia of the molecule. So the two kinds of hydrogen gas have different sets of rotational energy levels, and consequently different specific heats.

▷ EXERCISE. Determine the degeneracy of ortho- and parahydrogen. By expressing the state occupancy of the rotational states through the Boltzmann factor, determine the low temperature variation of the specific heat for the two species.

▷ EXAMPLE: As a final example, and one that will feed into our discussion of multielectron atoms in the next chapter, let us consider the implications of particle statistics for the **excited state spectrum of Helium**. After Hydrogen, Helium is the simplest atom having two protons and two neutrons in the nucleus ($Z = 2$), and two bound electrons. As a complex many-body system, we have seen already that the Schrödinger equation is analytically intractable and must be treated perturbatively. Previously, in chapter 7, we have used the ground state properties of Helium as a vehicle to practice perturbation theory. In the absence of direct electron-electron interaction, the Hamiltonian

$$\hat{H}_0 = \sum_{n=1}^2 \left[\frac{\hat{\mathbf{p}}_n^2}{2m} + V(\mathbf{r}_n) \right], \quad V(\mathbf{r}) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r},$$

is separable and the wavefunction can be expressed through the states of the hydrogen atom, ψ_{nlm} . In this approximation, the ground state wavefunction involves both

electrons occupying the $1s$ state leading to an antisymmetric spin singlet wavefunction for the spin degrees of freedom, $|\Psi_{\text{g.s.}}\rangle = (|100\rangle \otimes |100\rangle) \otimes |\chi_{00}\rangle$. In chapter 7, we made use of both the perturbative series expansion and the variational method to determine how the ground state energy is perturbed by the repulsive electron-electron interaction,

$$\hat{H}_1 = \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

Now let us consider the implications of particle statistics on the spectrum of the lowest excited states.

From the symmetry perspective, the ground state wavefunction belongs to the class of states with symmetric spatial wavefunctions, and antisymmetric spin (singlet) wavefunctions. These states are known as **parahelium**. In the absence of electron-electron interaction, the first excited states are degenerate and have the form,

$$|\psi_p\rangle = \frac{1}{\sqrt{2}} (|100\rangle \otimes |2\ell m\rangle + |2\ell m\rangle \otimes |100\rangle) \otimes |\chi_{00}\rangle.$$

The second class of states involve an antisymmetric spatial wavefunction, and symmetric (triplet) spin wavefunction. These states are known as **orthohelium**. Once again, in the absence of electron-electron interaction, the first excited states are degenerate and have the form,

$$|\psi_o\rangle = \frac{1}{\sqrt{2}} (|100\rangle \otimes |2\ell m\rangle - |2\ell m\rangle \otimes |100\rangle) \otimes |\chi_{1S_z}\rangle.$$

The perturbative shift in the ground state energy has already been calculated within the framework of first order perturbation theory. Let us now consider the shift in the excited states. Despite the degeneracy, since the off-diagonal matrix elements vanish, we can make use of the first order of perturbation theory to compute the shift. In doing so, we obtain

$$\Delta E_{n\ell}^{\text{p,o}} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 |\psi_{100}(\mathbf{r}_1)\psi_{n\ell 0}(\mathbf{r}_2) \pm \psi_{n\ell 0}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|},$$

with the plus sign refers to parahelium and the minus to orthohelium. Since the matrix element is independent of m , the $m = 0$ value considered here applies to all values of m . Rearranging this equation, we thus obtain $\Delta E_{n\ell}^{\text{p,o}} = J_{n\ell} \pm K_{n\ell}$ where

$$J_{n\ell} = \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \frac{|\psi_{100}(\mathbf{r}_1)|^2 |\psi_{n\ell 0}(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

$$K_{n\ell} = \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \frac{\psi_{100}^*(\mathbf{r}_1)\psi_{n\ell 0}^*(\mathbf{r}_2)\psi_{100}(\mathbf{r}_2)\psi_{n\ell 0}(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

Physically, the term $J_{n\ell}$ represents the electrostatic interaction energy associated with the two charge distributions $|\psi_{100}(\mathbf{r}_1)|^2$ and $|\psi_{n\ell 0}(\mathbf{r}_2)|^2$, and it is clearly positive. By contrast, the **exchange term**, which derives from the antisymmetry of the wavefunction, leads to a shift with opposite signs for ortho and para states. In fact, one may show that, in the present case, $K_{n\ell}$ is positive leading to a positive energy shift for parahelium and a negative shift for orthohelium. Moreover, noting that

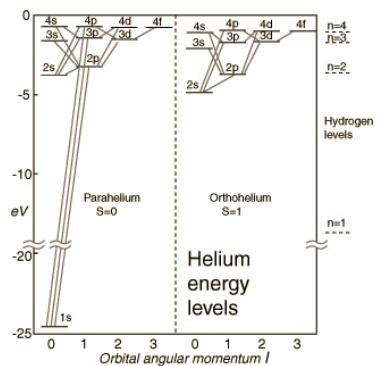
$$2\mathbf{S}_1 \cdot \mathbf{S}_2 = (\mathbf{S}_1 + \mathbf{S}_2)^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2 = \hbar^2 \left(S(S+1) - 2 \times \frac{3}{4} \right) = \hbar^2 \begin{cases} 1/2 & \text{triplet} \\ -3/2 & \text{singlet} \end{cases}$$

the energy shift can be written as

$$\Delta E_{n\ell}^{\text{p,o}} = J_{n\ell} - \frac{1}{2} \left(1 + \frac{4}{\hbar^2} \mathbf{S}_1 \cdot \mathbf{S}_2 \right) K_{n\ell}.$$

This result shows that the electron-electron interaction leads to an effective **ferromagnetic** interaction between spins – i.e. the spins want to be aligned.

In addition to the large energy shift between the singlet and triplet states, electric dipole decay selection rules $\Delta\ell = \pm 1$, $\Delta s = 0$ (whose origin is discussed later in the



Energy level diagram for ortho- and parahelium showing the first order shift in energies due to the Coulomb repulsion of electrons. Here we assume that one of the electrons stays close to the ground state of the unperturbed Hamiltonian.

course) cause decays from triplet to singlet states (or vice-versa) to be suppressed by a large factor (compared to decays from singlet to singlet or from triplet to triplet). This caused early researchers to think that there were two separate kinds of Helium. The diagrams (right) shows the levels for parahelium (singlet) and for orthohelium (triplet) along with the dominant decay modes.

8.4 Many-body systems

An important and recurring example of a many-body system is provided by the problem of free (i.e. non-interacting) non-relativistic quantum particles in a closed system – a box. The many-body Hamiltonian is then given simply by the kinetic energy

$$\hat{H}_0 = \sum_{i=1}^N \frac{\hat{\mathbf{p}}_i^2}{2m}, \quad (8.2)$$

where $\hat{\mathbf{p}}_i = -i\hbar\nabla_i$ and m denotes the particle mass. If we take the dimensions of the box to be L^d , and the boundary conditions to be periodic⁶ the normalized eigenstates of the single-particle Hamiltonian are simply given by plane waves, $\phi_{\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{L^{d/2}} e^{i\mathbf{k}\cdot\mathbf{r}}$, with wavevectors taking discrete values,⁷

$$\mathbf{k} = \frac{2\pi}{L}(n_1, n_2, \dots, n_d), \quad n_i \text{ integer.}$$

To address the quantum mechanics of the system, we start with fermions.

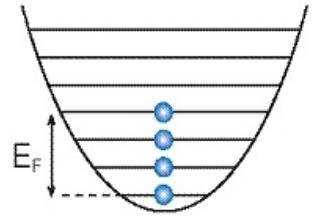
8.4.1 Non-interacting Fermi gas

In the (**spinless**) **Fermi system**, Pauli exclusion inhibits multiple occupancy of single-particle states. In this case, the many-body ground state wavefunction is obtained by filling states sequentially up to the **Fermi energy**, $E_F = \hbar^2 k_F^2 / 2m$, where the **Fermi wavevector**, k_F , is fixed by the number of particles. All the plane wave states $\phi_{\mathbf{k}}$ with energies lower than E_F are filled, while all states with energies larger than E_F remain empty. Since each state is associated with a k -space volume $(2\pi/L)^d$ (see Fig. 8.1), in the three-dimensional system, the total number of occupied states is given by $N = (\frac{L}{2\pi})^3 \frac{4}{3} \pi k_F^3$, i.e. defining the mean particle density $\bar{n} \equiv N/L^3 = k_F^3/6\pi^2$,

$$E_F = \frac{\hbar^2}{2m} (6\pi^2 \bar{n})^{\frac{2}{3}}.$$

The **density of states** per unit volume,

$$g(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{1}{6\pi^2} \frac{d}{dE} \left(\frac{2mE}{\hbar^2} \right)^{3/2} = \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2}.$$



Note that the volume of a d -dimensional sphere is given by $S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$.

⁶It may seem odd to consider such an unphysical geometry – in reality, we are invariably dealing with a closed system in which the boundary conditions translate to “hard walls” – think of electrons in a metallic sample. Here, we have taken the boundary conditions to be periodic since it leads to a slightly more simple mathematical formulation. We could equally well consider closed boundary conditions, but we would have to separately discriminate between “even and odd” states and sum them accordingly. Ultimately, we would arrive to the same conclusions!

⁷The quantization condition follows from the periodic boundary condition, $\phi[\mathbf{r} + L(m_x \hat{\mathbf{x}} + m_y \hat{\mathbf{y}} + m_z \hat{\mathbf{z}})] = \phi(\mathbf{r})$, where $\mathbf{m} = (m_x, m_y, m_z)$ denotes an arbitrary vector of integers.

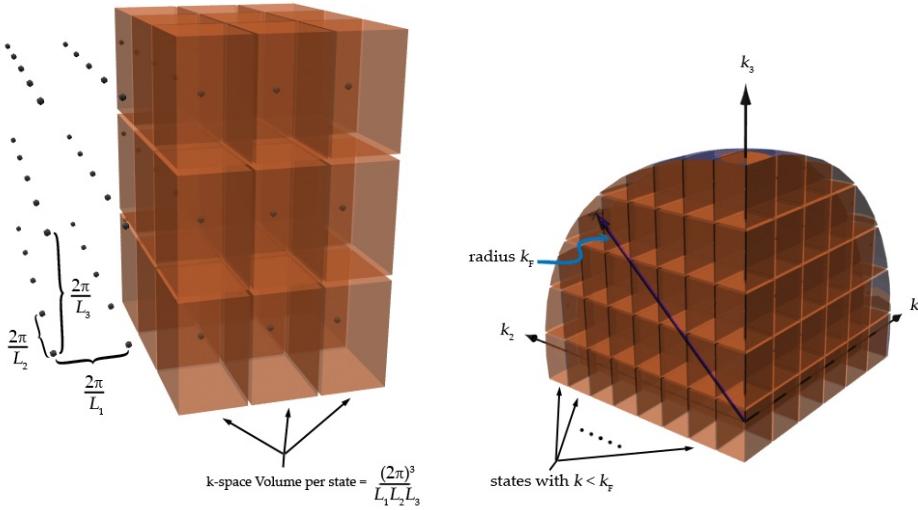


Figure 8.1: (Left) Schematic showing the phase space volume associated with each plane wave state in a Fermi gas. (Right) Schematic showing the state occupancy of a filled Fermi sea.

▷ EXERCISE. Obtain an expression for the density of states, $g(E)$ in dimension d . In particular, show that the density of states varies as $g(E) \sim E^{(d-2)/2}$.

We can also integrate to obtain the total energy density of all the fermions,

$$\frac{E_{\text{tot}}}{L^3} = \frac{1}{L^3} \int_0^{k_F} \frac{4\pi k^2 dk}{(2\pi/L)^3} \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{20\pi^2 m} (6\pi^2 \bar{n})^{5/3} = \frac{3}{5} \bar{n} E_F .$$

Therefore the energy per particle is

$$\frac{E_{\text{tot}}}{N} = \frac{E_{\text{tot}}}{\bar{n} L^3} = \frac{3}{5} E_F .$$

The groundstate wavefunction is the Slater determinant

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{\mathbf{k}_1}(\mathbf{r}_1) & \phi_{\mathbf{k}_2}(\mathbf{r}_1) & \dots & \phi_{\mathbf{k}_N}(\mathbf{r}_1) \\ \phi_{\mathbf{k}_1}(\mathbf{r}_2) & \phi_{\mathbf{k}_2}(\mathbf{r}_2) & \dots & \phi_{\mathbf{k}_N}(\mathbf{r}_2) \\ \vdots & \vdots & & \vdots \\ \phi_{\mathbf{k}_1}(\mathbf{r}_N) & \phi_{\mathbf{k}_2}(\mathbf{r}_N) & \dots & \phi_{\mathbf{k}_N}(\mathbf{r}_N) \end{vmatrix}$$

formed using the plane wave states with wavevectors $|\mathbf{k}_i| \leq k_F$ (i.e. those of the filled Fermi sea).

8.4.2 Non-interacting Bose gas

For the case of non-interacting bosons, confined in the same box, the nature of the groundstate is even simpler. There is no exclusion principle, so the groundstate is obtained by putting all bosons in the lowest energy single particle state, $\phi_{\mathbf{k}=0}(\mathbf{r})$. This single particle state has vanishing kinetic energy. Thus, the many-particle groundstate has energy

$$E_{\text{tot}} = 0 .$$

The groundstate wavefunction is

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i=1}^N \phi_{\mathbf{k}=0}(\mathbf{r}_i)$$

▷ INFO. **Interacting Fermi gas:**

For realistic metals or (doped) semiconductors, it is unreasonable to treat the electrons as non-interacting fermions. After all, they have electric charge e , so interact via the Coulomb interactions. As a result, in addition to the kinetic energy (8.2) in the Hamiltonian one should add an interaction⁸

$$\hat{V}_I = \frac{1}{2} \sum_{i \neq j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V_{\text{bg}}(\mathbf{r}_i), \quad (8.3)$$

The first term describes the electron-electron repulsion. In the second term, $V_{\text{bg}}(\mathbf{r})$ is the average (attractive) background potential due to the charged ions; this is required in order to make the system overall charge neutral.

Let us estimate how large this interaction energy is. For a gas with density \bar{n} , the typical interparticle spacing is $\bar{a} \equiv \bar{n}^{-1/3}$. Therefore, by dimensional analysis, the mean interaction energy per particle must be of order

$$\frac{\langle \hat{V}_I \rangle}{N} \sim \frac{e^2}{4\pi\epsilon_0 \bar{a}}$$

(We shall discuss the size of the prefactor below.) On the other hand, we derived the kinetic energy per particle for the groundstate of the spinless Fermi gas,

$$\frac{\langle \hat{H}_0 \rangle}{N} = \frac{3}{5} E_F \sim \frac{\hbar^2}{m\bar{a}^2}$$

where we have used $E_F \sim \hbar^2/(m\bar{a}^2)$.

The ratio of the mean interaction energy to the mean kinetic energy is

$$\frac{\langle \hat{V}_I \rangle}{\langle \hat{H}_0 \rangle} \sim \frac{e^2/(4\pi\epsilon_0 \bar{a})}{\hbar^2/m\bar{a}^2} = \frac{\bar{a}}{a_0}$$

where $a_0 \equiv \frac{4\pi\hbar^2\epsilon_0}{me^2}$ is the Bohr radius. We can understand the properties of the interacting Fermi gas in two simple limiting cases.

(1) Weak interactions: For $\bar{a} \ll a_0$ the typical interaction energy is small compared to the kinetic energy. In this case, one can treat the interactions as a perturbative correction to the groundstate of the kinetic energy. As discussed above, this groundstate is the filled Fermi sea. For weak interactions, the perturbative corrections are small and the interacting electron gas has the same qualitative properties as the non-interacting (ideal) Fermi gas. This case is typical for “good” metals such as copper.

(2) Strong interactions: For $\bar{a} \gg a_0$ the typical interaction energy is large compared to the kinetic energy. In this case, one should treat the kinetic energy as a perturbative correction to the groundstate of the interaction potential (8.3). The interaction potential is minimized by keeping the electrons as far apart from each other as possible. At a fixed average density \bar{n} , this is achieved by placing the electrons at points located at the sites of a regular lattice. This unperturbed wavefunction is therefore a *crystal of electrons*. Provided the kinetic energy is small ($\bar{a} \gg a_0$), the perturbation by \hat{H}_0 acts only to move the electrons a small distance from the sites of the crystalline lattice. The integrity of the lattice remains, as a so-called **Wigner crystal**. This limit of strong interactions is *not* achieved in typical metallic systems for which the density is too high. However, in two-dimensional electron gases (formed both in doped semiconductor devices and on the surface of liquid helium) there is evidence for the appearance of a 2D Wigner crystal in experiment.

The two limiting cases described above are simple to understand, since we can use perturbation theory. However, many metallic systems are in an intermediate regime,

⁸The electrons move in the background potential of the ionic cores of the atoms. This is a periodic potential acting on each electron, and leads to a more complicated band structure: with a new “effective mass” m^* and with band gaps at the Brillouin zone boundaries, *etc.* This can have important quantitative effects on the properties of the metal. Here we focus on the *qualitative* effects of electron-electron interactions. Therefore, for simplicity, we neglect this periodic potential.

with $\bar{a} \simeq a_0$, where perturbation theory in either \hat{V}_I or \hat{H}_0 is inappropriate. This case of intermediate coupling poses a very difficult challenge to theory, and few results are known with certainty. One technique that one can always fall back on is the **variational** approach. One can “guess” what possible phases of the electron system might appear, and construct trial wavefunctions for each of these phases. By comparing the expectation values of the total energy for these different trial wavefunctions, one obtains an indication of which of them best describes the groundstate and therefore what phase the electron system may be in. For example, this technique has been used to determine at what (numerical) value of \bar{a}/a_0 the transition between the Wigner crystal and Fermi-gas states occurs.

Another important example of the application of variational approach is in understanding **ferromagnetism** of an itinerant spin-1/2 Fermi gas. In this case, the two representative states to consider are: (a) a filled Fermi sea of spin-polarized electrons, say with all spins “up”; and (b) a filled Fermi sea of both spin components. By fixing the occupation numbers to be filled Fermi seas, these states have no remaining variational parameters. A more general class of wavefunctions is provided by “Hartree-Fock” theory, in which the orbitals in the Slater determinants are viewed as variational parameters.

Case (a), the spin-polarized Fermi sea, was discussed in detail above. For a density \bar{n} the states are filled out to the Fermi wavevector $k_F = (6\pi^2\bar{n})^{1/3}$. A direct calculation of the mean energy per particle in this spin-polarized Fermi gas gives

$$\frac{E_{\text{tot}}^{\text{pol.}}}{N} \equiv \frac{\langle H_0 \rangle + \langle V_I \rangle}{N} = \frac{3}{5} \frac{\hbar^2}{2m\bar{a}^2} (6\pi^2)^{2/3} - \frac{3(6\pi^2)^{1/3}}{4\pi} \frac{e^2}{4\pi\epsilon_0\bar{a}}$$

The prefactor of the interaction term arises from a straightforward but lengthy calculation⁹ (not shown here!). The negative sign of the interaction term indicates that the energy is lowered by the electrons residing in the material, and thereby neutralising the background charge density of the ions. The overall size of the energy reduction is influenced by the extent to which the electrons repel each other.

Case (b), the spin-unpolarized Fermi gas, follows similar considerations. Note now that a total density \bar{n} will consist of density $\bar{n}/2$ of spin-up electrons and a density $\bar{n}/2$ of spin-down electrons. Since there is no Pauli exclusion between these two spin states, each spin-state can fill its own Fermi sea. The resulting Fermi wavevector (for both spin-up and spin-down) is therefore $k'_F = (6\pi^2\bar{n}/2)^{1/3} = k_F/2^{1/3}$. A direct calculation of the mean energy per particle in this spin-unpolarized Fermi gas leads to

$$\frac{E_{\text{tot}}^{\text{unpol.}}}{N} \equiv \frac{\langle H_0 \rangle + \langle V_I \rangle}{N} = \frac{3}{5} \frac{\hbar^2}{2m\bar{a}^2} (3\pi^2)^{2/3} - \frac{3(6\pi^2)^{1/3}}{4\pi 2^{1/3}} \frac{e^2}{4\pi\epsilon_0\bar{a}}$$

Compared to the polarized case, the kinetic energy is smaller. This is due to the fact that, since both spin states are occupied, the electrons are not forced up to such a high Fermi energy as in the polarized case. However, note also that the reduction in the energy due to the interactions is smaller. That is, the total interaction energy of the electrons is larger for the spin-unpolarized case than for the spin-polarized case. For the polarized Fermi gas, the wavefunction must vanish when any two electrons approach each other. This effectively keeps the electrons far apart from each other, and reduces the electron-electron repulsion. For the unpolarized Fermi gas, the wavefunction must vanish when any two particles of the same spin approach each other. However, particles of opposite spin can be located nearby in space, and can contribute a larger Coulomb energy.¹⁰

For high density (small \bar{a}), the kinetic energy dominates and favours the unpolarized Fermi gas. For low density (large \bar{a}), the interaction energy dominates and favours the polarized Fermi gas. A comparison of the total energies of the two variational states shows that the transition between these is at

$$\frac{\bar{a}}{a_0} = \frac{2\pi}{5} (3\pi^2)^{1/3} (1 + 2^{1/3}) \simeq 8.8$$

⁹The calculation of the interaction energy can be found in “Solid State Physics” by Ashcroft and Mermin.

¹⁰This is the same qualitative physics as in the excited states of helium in the example on page 74, and that we will find when we discuss Hund’s rules.

This provides a simple variational estimate for the condition **ferromagnetism** in the interacting Fermi gas. This competition is responsible for the ferromagnetism of iron and other metallic magnets. (For even larger values of \bar{a}/a_0 there should be a transition from the polarized Fermi gas into a Wigner crystal.)

▷ INFO. **Interacting Bose Gas:** For the BECs formed in cold atomic gases, there are residual interactions between the particles. To a good approximation, these can be represented by contact interactions, with

$$\hat{V}_I = \frac{1}{2} \sum_{i \neq j} g \delta(\mathbf{r}_i - \mathbf{r}_j).$$

A simple variational approach for the interacting Bose gas is provided by the **Gross-Pitaevksii** approximation. One assumes that the groundstate is a condensate, Eqn.(8.1), and treats the (normalized) condensate wavefunction $\psi(\mathbf{r})$ as the variational function. The mean energy per particle (for $N \gg 1$) is

$$\frac{\langle \hat{H}_0 + \hat{V}_I \rangle}{N} = \int \left[\frac{\hbar^2}{2m} |\nabla \psi|^2 + V^{\text{ext}}(\mathbf{r}) |\psi|^2 \right] d^3 r + \frac{1}{2} g N \int |\psi|^4 d^3 r$$

for the BEC in an external trapping potential $V^{\text{ext}}(\mathbf{r})$. The optimal condensate wavefunction $\psi(\mathbf{r})$ is found by minimizing the energy per particle, subject to the normalization $\langle \psi | \psi \rangle = 1$. This constraint can be imposed by introducing a Lagrange multiplier, $-E \int |\psi|^2 d^3 r$, with minimization over ψ then leading to

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + [V^{\text{ext}}(\mathbf{r}) + gN|\psi|^2] \psi(\mathbf{r}) = E\psi(\mathbf{r})$$

Since $\psi(\mathbf{r})$ is normalized, the mean particle density is $n(\mathbf{r}) = N|\psi(\mathbf{r})|^2$. The interaction term can be viewed as contributing to an effective potential to $V^{\text{total}}(\mathbf{r}) = V^{\text{ext}}(\mathbf{r}) + gn(\mathbf{r})$ in which $n(\mathbf{r})$ is a function of $\psi(\mathbf{r})$ and must therefore be self-consistently determined.

Chapter 9

Atomic structure

9.1 The non-relativistic hydrogen atom

The quantum mechanics of atomic hydrogen, and hydrogen-like atoms is characterized by a large degeneracy with eigenvalues separating into multiplets of n^2 -fold degeneracy, where n denotes the principal quantum number. However, although the idealized Schrödinger Hamiltonian,

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + V(r), \quad V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}, \quad (9.1)$$

provides a useful platform from which develop our intuition, there are several important effects which mean that the formulation is a little too naïve. These “corrections”, which derive from several sources, are important as they lead to physical ramifications which extend beyond the realm of atomic physics. Here we outline some of the effects which need to be taken into account even for atomic hydrogen, before moving on to discuss the quantum physics of multi-electron atoms. In broad terms, the effects to be considered can be grouped into those caused by the internal properties of the nucleus, and those which derive from relativistic corrections.

To orient our discussion, it will be helpful to summarize some of key aspects of the solutions of the non-relativistic Schrödinger equation, $\hat{H}_0\psi = E\psi$ on which we will draw:

Hydrogen atom revisited:

- ▷ As with any centrally symmetric potential, the solutions of the Schrödinger equation take the form $\psi_{\ell m_\ell}(\mathbf{r}) = R(r)Y_{\ell m_\ell}(\theta, \phi)$, where the spherical harmonic functions $Y_{\ell m_\ell}(\theta, \phi)$ depend only on spherical polar coordinates, and $R(r)$ represents the radial component of the wavefunction. Solving the radial wave equation introduces a radial quantum number, $n_r \geq 0$. In the case of a Coulomb potential, the energy depends on the principal quantum number $n = n_r + \ell + 1 \geq 1$, and not on n_r and ℓ separately.
- ▷ For atomic hydrogen ($Z = 1$), the energy levels of the Hamiltonian (9.1) are given by

$$E_n = -\frac{\text{Ry}}{n^2}, \quad \text{Ry} = \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m}{2\hbar^2} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{2a_0} = \frac{1}{2} mc^2 \alpha^2,$$

where $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{e^2 m}$ is the Bohr radius, $\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c}$ denotes the **fine structure constant**, and m represents the reduced mass of the electron

The fine structure constant is known to great accuracy and is given by,

$$\alpha = 7.297352570(5) \times 10^{-3} = \frac{1}{137.035999070(9)}.$$

and proton. Applied to single electron ions with higher atomic weight, such as He^+ , Li^{2+} , etc., the Bohr radius is reduced by a factor $1/Z$, where Z denotes the nuclear charge, and the energy is given by $E_n = -\frac{Z^2}{n^2}\text{Ry} = -\frac{1}{2n^2}mc^2(Z\alpha)^2$.

- ▷ Since $n \geq 1$ and $n_r \geq 0$, the allowed combinations of quantum numbers are shown on the right, where we have introduced the conventional notation whereby values of $\ell = 0, 1, 2, 3, 4 \dots$ are represented by letters $s, p, d, f, g \dots$ respectively.
- ▷ Since E_n depends only on n , this implies, for example, an exact degeneracy of the $2s$ and $2p$, and of the $3s$, $3p$ and $3d$ levels.

These results emerge from a treatment of the non-relativistic hydrogen atom which is inherently non-relativistic. In fact, the Hamiltonian (9.1) represents only the leading term in an expansion in $v^2/c^2 \simeq (Z\alpha)^2$ of the full relativistic Hamiltonian (see below). Higher order terms provide relativistic corrections, which impact significantly in atomic and condensed matter physics, and lead to a lifting of the degeneracy. In the following we will discuss and obtain the hierarchy of leading relativistic corrections.¹ This discussion will provide a platform to describe multi-electron atoms.

9.1.1 The Zeeman effect

We start by exploring how the spectrum of the hydrogen atom is influenced by a uniform magnetic field, \mathbf{B} . There are two main effects:

(1) The magnetic field couples to the magnetic moment associated with the *spin* of the electron.² This gives rise to an additional term in the Hamiltonian

$$\hat{H}_{\text{spin}} = -\boldsymbol{\mu}_e \cdot \mathbf{B} = g \frac{e}{2m} \hat{\mathbf{S}} \cdot \mathbf{B}$$

where we have used the spin magnetic dipole moment of the electron $\mu_e = -g\mu_B(\mathbf{S}/\hbar)$ where $\mu_B \equiv e\hbar/(2m)$ is the Bohr magneton, and g the electron *g*-factor which is close to $g = 2$.³

(2) The field couples to the orbital motion of the (charged) electron. This is accounted for by replacing the kinetic energy with

$$\frac{1}{2m}(\hat{\mathbf{p}})^2 \rightarrow \frac{1}{2m}(\hat{\mathbf{p}} + e\mathbf{A}(\hat{\mathbf{r}}))^2 \quad (9.2)$$

giving an additional change in the Hamiltonian

$$\hat{H}_{\text{orbital}} = \frac{1}{2m} \left[e(\hat{\mathbf{p}} \cdot \mathbf{A} + \mathbf{A} \cdot \hat{\mathbf{p}}) + \hat{e}^2 |\mathbf{A}|^2 \right]$$

Let us explore further the effects of these orbital terms. The term linear in \mathbf{A} is known as the **paramagnetic term**, and the quadratic term the **diamagnetic contribution**. For a constant magnetic field, the vector potential

¹It may seem odd to discuss relativistic corrections without introducing the Dirac equation and the relativistic formulation of quantum mechanics. However, such a discussion would present a lengthy and complex digression. We will therefore follow the normal practice of discussing relativistic corrections as perturbations to the familiar non-relativistic theory.

²There is also a coupling to the nuclear magnetic moment. Since $\mu_N/\mu_e \sim 10^{-3}$ this is a much weaker effect. However, it can be important in the hyperfine structure.

³The Dirac equation for the electron leads to the prediction that $g = 2$ exactly. Experimental measurements show $g = 2[1.0011596521884(\pm43)]$. The small correction is known as the anomalous magnetic dipole moment of the electron. It arises from the electron's interaction with virtual photons in quantum electrodynamics, and can be accurately computed within that theory.

| n | ℓ | Subshell(s) |
|-----|-----------------|-------------|
| 1 | 0 | 1s |
| 2 | 0, 1 | 2s 2p |
| 3 | 0, 1, 2 | 3s 3p 3d |
| 4 | 0, 1, 2, 3 | 4s 4p 4d 4f |
| n | $0 \dots (n-1)$ | $ns \dots$ |

To see that $v^2/c^2 \simeq (Z\alpha)^2$, we may invoke the **virial theorem**. The latter shows that the average kinetic energy is related to the potential energy as $\langle T \rangle = -\frac{1}{2}\langle V \rangle$. Therefore, the average energy is given by $\langle E \rangle = \langle T \rangle + \langle V \rangle = -\langle T \rangle \equiv -\frac{1}{2}mv^2$. We therefore have that $\frac{1}{2}mv^2 = \text{Ry} \equiv \frac{1}{2}mc^2(Z\alpha)^2$ from which follows the relation $v^2/c^2 \simeq (Z\alpha)^2$.

can be as $\mathbf{A} = -\hat{\mathbf{r}} \times \mathbf{B}/2$. In this case, the paramagnetic component takes the form

$$-\frac{ie\hbar}{m}\mathbf{A} \cdot \nabla = -\frac{ie\hbar}{2m}(\hat{\mathbf{r}} \times \nabla) \cdot \mathbf{B} = \frac{e}{2m}\hat{\mathbf{L}} \cdot \mathbf{B},$$

where $\hat{\mathbf{L}}$ denotes the angular momentum operator. Similarly, the diamagnetic term leads to

$$\frac{e^2}{2m}\mathbf{A}^2 = \frac{e^2}{8m}(\hat{\mathbf{r}}^2\mathbf{B}^2 - (\hat{\mathbf{r}} \cdot \mathbf{B})^2) = \frac{e^2B^2}{8m}(\hat{x}^2 + \hat{y}^2),$$

where, here, we have chosen the magnetic field to lie along the z -axis.

Before addressing the role of these separate contributions in atomic hydrogen, let us first estimate their relative magnitude. With $\langle \hat{x}^2 + \hat{y}^2 \rangle \simeq a_0^2$, where a_0 denotes the Bohr radius, and $\langle \hat{L}_z \rangle \simeq \hbar$, the ratio of the diamagnetic to paramagnetic terms is

$$\frac{(e^2/8m_e)\langle \hat{x}^2 + \hat{y}^2 \rangle B^2}{(e/2m_e)\langle \hat{L}_z \rangle B} = \frac{e}{4} \frac{a_0^2 B^2}{\hbar B} \simeq 10^{-6} B/\text{T}.$$

Therefore, provided the electron is bound in an orbital of size $\sim a_0$, for fields that can be achieved in the laboratory ($B \simeq 1 \text{ T}$) the diamagnetic term is negligible compared to the paramagnetic term.

Thus, for typical laboratory magnetic fields, the diamagnetic term can be neglected and the net change to the Hamiltonian is

$$\hat{H}_{\text{Zeeman}} \equiv \hat{H}_{\text{spin}} + \hat{H}_{\text{orbital}} \simeq \frac{e}{2m}(\hat{gS} + \hat{L}) \cdot \mathbf{B} \quad (9.3)$$

The overall Hamiltonian for an electron moving in a Coulomb potential in the presence of a constant magnetic field $\mathbf{B} = B\mathbf{z}$ becomes

$$\hat{H} \simeq \hat{H}_0 + \frac{eB}{2m}(g\hat{S}_z + \hat{L}_z),$$

where $\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r}$. Since $[\hat{H}_0, \hat{S}_z] = [\hat{H}_0, \hat{L}_z] = 0$, the eigenstates of the unperturbed Hamiltonian, $\psi_{n\ell m_\ell m_s}(\mathbf{r})$, remain eigenstates of \hat{H} and the corresponding energy levels are

$$E_{n\ell m_\ell m_s} \simeq -\frac{\text{Ry}}{n^2} + \hbar\omega_L(gm_s + m_\ell)$$

where $\omega_L = \frac{eB}{2m}$ denotes the **Larmor frequency**. From this result, we expect that a constant magnetic field will lead to a splitting of the $2 \times (2\ell + 1)$ -fold degeneracy of the energy levels. Before the knowledge of the existence of electron spin, one would have expected the level to split into $(2\ell + 1)$ levels, due to the orbital Zeeman effect. The fact that experiment showed an additional two-fold splitting, was one of the key insights that led to the identification of electron spin.

9.2 The “real” hydrogen atom

The relativistic corrections (sometimes known as the **fine-structure corrections**) to the spectrum of hydrogen-like atoms derive from three different sources:

- ▷ relativistic corrections to the kinetic energy;
- ▷ coupling between spin and orbital degrees of freedom;
- ▷ a contribution known as the Darwin term.

In the following, we will discuss each of these corrections in turn.

Sir Joseph Larmor 1857-1942
 A physicist and mathematician who made innovations in the understanding of electricity, dynamics, thermodynamics, and the electron theory of matter. His most influential work was *Aether and Matter*, a theoretical physics book published in 1900. In 1903 he was appointed Lucasian Professor of Mathematics at Cambridge, a post he retained until his retirement in 1932.



9.2.1 Relativistic correction to the kinetic energy

Previously, we have taken the kinetic energy to have the familiar non-relativistic form, $\frac{\hat{\mathbf{p}}^2}{2m}$. However, from the expression for the relativistic energy-momentum invariant, $p_\mu p^\mu = (mc)^2$, we can already anticipate that the leading correction to the non-relativistic Hamiltonian appears at order \mathbf{p}^4 ,

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} = mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{1}{8} \frac{(\mathbf{p}^2)^2}{m^3 c^2} + \dots$$

As a result, we can infer the following perturbation to the kinetic energy of the electron,

$$\boxed{\hat{H}_1 = -\frac{1}{8} \frac{(\hat{\mathbf{p}}^2)^2}{m^3 c^2}.}$$

When compared with the non-relativistic kinetic energy, $\mathbf{p}^2/2m$, one can see that the perturbation is smaller by a factor of $p^2/m^2 c^2 = v^2/c^2 \simeq (Z\alpha)^2$, i.e. \hat{H}_1 is only a small perturbation for small atomic number, $Z \ll 1/\alpha \simeq 137$. We can therefore make use of a perturbative analysis to estimate the scale of the correction.

In principle, the large-scale degeneracy of the hydrogen atom would demand an analysis based on the degenerate perturbation theory. However, fortunately, since the off-diagonal matrix elements vanish,⁴

$$\langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0 \quad \text{for } \ell \neq \ell' \quad \text{or} \quad m_\ell \neq m'_\ell,$$

degenerate states are uncoupled and such an approach is unnecessary. Then making use of the identity, $\hat{H}_1 = -\frac{1}{2mc^2} [\hat{H}_0 - V(r)]^2$, the scale of the resulting energy shift can be obtained from first order perturbation theory,

$$\langle \hat{H}_1 \rangle_{n\ell m_\ell} \equiv \langle n\ell m_\ell | \hat{H}_1 | n\ell m_\ell \rangle = -\frac{1}{2mc^2} (E_n^2 - 2E_n \langle V(r) \rangle_{n\ell} + \langle V^2(r) \rangle_{n\ell}).$$

Since the calculation of the resulting expectation values is not particularly illuminating, we refer to the literature for a detailed exposition⁵ and present here only the required identities (right). From these considerations, we obtain the following expression for the first order energy shift,

$$\langle \hat{H}_1 \rangle_{n\ell m_\ell} = -\frac{mc^2}{2} \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{n}{\ell+1/2} - \frac{3}{4} \right). \quad (9.4)$$

From this term alone, we expect the degeneracy between states with different values of total angular momentum ℓ to be lifted. However, as we will see, this conclusion is a little hasty. We need to gather all terms of the same order of perturbation theory before we can reach a definite conclusion. We can, however, confirm that (as expected) the scale of the correction is of order $\frac{\langle \hat{H}_1 \rangle_{n\ell m_\ell}}{E_n} \sim (\frac{Z\alpha}{n})^2$. We now turn to the second important class of corrections.

9.2.2 Spin-orbit coupling

As well as revealing the existence of an internal spin degree of freedom, Dirac’s relativistic formulation of quantum mechanics shows that there is a further relativistic correction to the Schrödinger operator which involves a coupling

⁴Proof: Since $[\hat{H}_1, \hat{\mathbf{L}}^2] = 0$, $\hbar^2 [\ell'(\ell'+1) - \ell(\ell+1)] \langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0$. Similarly, since $[\hat{H}_1, \hat{L}_z] = 0$, $\hbar(m'_\ell - m_\ell) \langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0$.

By making use of the form of the radial wavefunction for the hydrogen atom, one may obtain the identities,

$$\begin{aligned} \left\langle \frac{1}{r} \right\rangle_{n\ell} &= \frac{Z}{a_0 n^2} \\ \left\langle \frac{1}{r^2} \right\rangle_{n\ell} &= \frac{Z^2}{a_0^2 n^3 (\ell + 1/2)}. \end{aligned}$$

between the spin and orbital degrees of freedom. For a general potential $V(r)$, this spin-orbit coupling takes the form,

$$\hat{H}_2 = \frac{1}{2m^2c^2} \frac{1}{r} (\partial_r V(r)) \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}.$$

For a hydrogen-like atom, $V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}$, and

$$\hat{H}_2 = \frac{1}{2m^2c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r^3} \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}.$$

▷ INFO. Physically, the **origin of the spin-orbit interaction** can be understood from the following considerations. As the electron is moving through the electric field of the nucleus then, in its rest frame, it will experience this as a magnetic field. There will be an additional energy term in the Hamiltonian associated with the orientation of the spin magnetic moment with respect to this field. We can make an estimate of the spin-orbit interaction energy as follows: If we have a central field determined by an electrostatic potential $\phi(r)$, the corresponding electric field is given by $\mathbf{E} = -\nabla\phi(r) = -\mathbf{e}_r(\partial_r\phi) = +\mathbf{e}_r(\partial_rV)/e$ where in the last step we introduce the potential energy of the electron $V = -e\phi$. For an electron moving at velocity \mathbf{v} , this translates to an effective magnetic field $\mathbf{B} = -\frac{1}{c^2}\mathbf{v} \times \mathbf{E}$. The magnetic moment of the electron associated with its spin is equal to $\boldsymbol{\mu}_s = g_s \frac{-e}{2m} \mathbf{S} \equiv -\frac{e}{m} \mathbf{S}$, and thus the interaction energy is given by

$$-\boldsymbol{\mu}_s \cdot \mathbf{B} = -\frac{e}{mc^2} \mathbf{S} \cdot (\mathbf{v} \times \mathbf{E}) = -\frac{1}{(mc)^2} \mathbf{S} \cdot (\mathbf{p} \times \mathbf{e}_r(\partial_rV)) = +\frac{1}{(mc)^2} \frac{1}{r} (\partial_rV) \mathbf{L} \cdot \mathbf{S},$$

where we have used the relation $\mathbf{p} \times \mathbf{e}_r = \mathbf{p} \times \frac{\mathbf{r}}{r} = -\frac{\mathbf{L}}{r}$. In fact this isn’t quite correct; there is a relativistic effect connected with the precession of axes under rotation, called **Thomas precession** which multiplies the formula by a further factor of $\frac{1}{2}$.

Once again, we can estimate the effect of spin-orbit coupling by treating \hat{H}_2 as a perturbation. In the absence of spin-orbit interaction, one may express the eigenstates of hydrogen-like atoms in the basis states of the mutually commuting operators, \hat{H}_0 , $\hat{\mathbf{L}}^2$, \hat{L}_z , $\hat{\mathbf{S}}^2$, and \hat{S}_z . However, in the presence of spin-orbit coupling, the total Hamiltonian no longer commutes with \hat{L}_z or \hat{S}_z (exercise). It is therefore helpful to make use of the degeneracy of the unperturbed Hamiltonian to switch to a new basis in which the angular momentum components of the perturbed system are diagonal. This can be achieved by turning to the basis of eigenstates of the operators, \hat{H}_0 , $\hat{\mathbf{J}}^2$, \hat{J}_z , $\hat{\mathbf{L}}^2$, and $\hat{\mathbf{S}}^2$, where $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ denotes the total angular momentum. (For a discussion of the form of these basis states, we refer back to Chapters 4 and 5.)

Making use of the relation, $\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$, in this basis, it follows that,

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2}(\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2).$$

Combining the spin and angular momentum, the total angular momentum takes values $j = \ell \pm 1/2$. The corresponding basis states $|j = \ell \pm 1/2, m_j, \ell\rangle$ (with $s = 1/2$ implicit) therefore diagonalize the operator,

$$\hat{\mathbf{S}} \cdot \hat{\mathbf{L}} |j = \ell \pm 1/2, m_j, \ell\rangle = \frac{\hbar^2}{2} \begin{Bmatrix} \ell \\ -\ell - 1 \end{Bmatrix} |\ell \pm 1/2, m_j, \ell\rangle,$$

where the curly brackets index $j = \ell + 1/2$ (top) and $j = \ell - 1/2$ (bottom). As for the radial dependence of the perturbation, once again, the off-diagonal matrix elements vanish circumventing the need to invoke degenerate perturbation theory. As a result, at first order in perturbation theory, one obtains

$$\langle H_2 \rangle_{n,j=\ell\pm 1/2,m_j,\ell} = \frac{1}{2m^2c^2} \frac{\hbar^2}{2} \left\{ \begin{array}{c} \ell \\ -\ell - 1 \end{array} \right\} \frac{Ze^2}{4\pi\epsilon_0} \left\langle \frac{1}{r^3} \right\rangle_{n\ell}.$$

Then making use of the identity (right),⁶ one obtains

$$\langle \hat{H}_2 \rangle_{n,j=\ell\pm 1/2,m_j,\ell} = \frac{1}{4} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \frac{n}{j+1/2} \left\{ \begin{array}{c} \frac{1}{j} \\ -\frac{1}{j+1} \end{array} \right\}.$$

For $\ell > 0$,

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \left(\frac{mc\alpha Z}{\hbar n} \right)^3 \frac{1}{\ell(\ell + \frac{1}{2})(\ell + 1)}.$$

Note that, for $\ell = 0$, there is no orbital angular momentum with which to couple! Then, if we rewrite the expression for $\langle \hat{H}_1 \rangle$ (9.4) in the new basis,

$$\langle \hat{H}_1 \rangle_{n,j=\ell\pm 1/2,m_j,\ell} = -\frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left\{ \begin{array}{c} \frac{n}{j} - \frac{3}{4} \\ \frac{n}{j+1} - \frac{3}{4} \end{array} \right\},$$

and combining both of these expressions, for $\ell > 0$, we obtain

$$\langle \hat{H}_1 + \hat{H}_2 \rangle_{n,j=\ell\pm 1/2,m_j,\ell} = \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{j+1/2} \right),$$

while for $\ell = 0$, we retain just the kinetic energy term (9.4).

The Zeeman effect revisited

Our theory of the Zeeman effect §9.1.1 was based on the effect of a uniform magnetic field as a perturbation to the *non-relativistic* Hamiltonian \hat{H}_0 . How are these conclusions affected by the additional physics from spin-orbit coupling?

We showed that the effect of a magnetic field $\mathbf{B} = B\mathbf{z}$ on a one-electron atom is to produce an additional term

$$\hat{H}_{\text{Zeeman}} = \frac{eB}{2m} (\hat{L}_z + g\hat{S}_z).$$

This denotes the Zeeman energy associated with the coupling of the spin and orbital angular momentum degrees of freedom to the magnetic field. Depending on the scale of the magnetic field, the Zeeman term \hat{H}_{Zeeman} or the spin-orbit term \hat{H}_2 may dominate the spectrum of the atom.

We have just seen that, to leading order, the relativistic corrections lead to a fine-structure energy shift of

$$\Delta E_{n,j}^{\text{rel}} = \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{j+1/2} \right),$$

for states $|n, j = \ell \pm 1/2, m_j, \ell\rangle$. For weak magnetic fields, we can treat the Zeeman energy as a perturbation to this spectrum. Here, although states with common j values are degenerate, the two spatial wavefunctions have different parities, since they have orbital angular momenta ($\ell = j - 1/2$ or $\ell = j + 1/2$) that differ by 1. (For example, for a given shell, with $n \geq 1$, a $j = 1/2$ state can be formed either from the $\ell = 0$ orbital or from the $\ell = 1$ orbital.) As a consequence of the differing parities of the spatial wavefunctions, the

⁶For details see, e.g., Ref. [1].

off-diagonal matrix element of \hat{H}_{Zeeman} coupling these states vanishes. We may therefore avoid using degenerate perturbation theory. Making use of the relation (exercise – refer back to the discussion of the addition of angular momenta and spin in section 5.4.2),

$$\langle n, j = \ell \pm 1/2, m_j, \ell | S_z | n, j = \ell \pm 1/2, m_j, \ell \rangle = \pm \frac{\hbar m_j}{2\ell + 1},$$

we obtain the following expression for the first order energy shift,

$$\Delta E_{j=\ell \pm 1, m_j, \ell}^{\text{Zeeman}} = \ell \pm 1/2, m_j, \ell \rangle = \mu_B B m_j \left(1 \pm \frac{1}{2\ell + 1} \right),$$

where μ_B denotes the Bohr magneton. Therefore, we see that all degenerate levels are split due to the magnetic field. In contrast to the “normal” Zeeman effect, the magnitude of the splitting depends on ℓ .

▷ INFO. If the **field is strong**, the Zeeman energy becomes large in comparison with the spin-orbit contribution. In this case, we must work with the basis states $|n, \ell, m_\ell, m_s\rangle = |n, \ell, m_\ell\rangle \otimes |m_s\rangle$ in which both \hat{H}_0 and \hat{H}_{Zeeman} are diagonal. Within first order of perturbation theory, one then finds that (exercise)

$$\Delta E_{n, \ell, m_\ell, m_s} = \mu_B (m_\ell + m_s) + \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{\ell + 1/2} - \frac{nm_\ell m_s}{\ell(\ell + 1/2)(\ell + 1)} \right),$$

the first term arising from the Zeeman energy and the remaining terms from $\hat{H}_{\text{rel.}}$. At intermediate values of the field, we have to apply degenerate perturbation theory to the states involving the linear combination of $|n, j = \ell \pm 1/2, m_j, \ell\rangle$. Such a calculation reaches beyond the scope of these lectures and, for details, we refer to the literature (see, e.g., Ref. [1]).

9.2.3 Darwin term

The final contribution to the Hamiltonian from relativistic effects is known as the Darwin term and arises from the “**Zitterbewegung**” of the electron – trembling motion – which smears the effective potential felt by the electron. Such effects lead to a perturbation of the form,

$$\hat{H}_3 = \frac{\hbar^2}{8m^2c^2} \nabla^2 V = \frac{\hbar^2}{8m^2c^2} \left(\frac{e}{\epsilon_0} Q_{\text{nuclear}}(\mathbf{r}) \right) = \frac{Ze^2}{4\pi\epsilon_0} \frac{\hbar^2}{8(mc)^2} 4\pi\delta^{(3)}(\mathbf{r}),$$

where $Q_{\text{nuclear}}(\mathbf{r}) = Ze\delta^{(3)}(\mathbf{r})$ denotes the nuclear charge density. Since the perturbation acts only at the origin, it affects only states with $\ell = 0$. As a result, one finds that

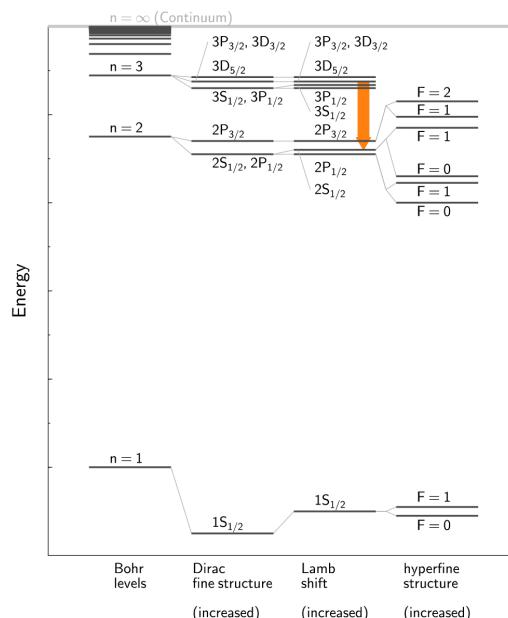
$$\langle \hat{H}_3 \rangle_{njm_j\ell} = \frac{Ze^2}{4\pi\epsilon_0} \frac{\hbar^2}{8(mc)^2} 4\pi |\psi_{\ell n}(0)|^2 = \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 n \delta_{\ell,0}.$$

Intriguingly, this term is formally identical to that which would be obtained from $\langle \hat{H}_2 \rangle$ at $\ell = 0$. As a result, combining all three contributions, the total energy shift is given simply by

$$\boxed{\Delta E_{n, j = \ell \pm 1/2, m_j, \ell} = \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{j + 1/2} \right)}, \quad (9.5)$$

a result that is independent of ℓ and m_j .

Figure 9.1: Figure showing the hierarchy of energy shifts of the spectra of hydrogen-like atoms as a result of relativistic corrections. The first column shows the energy spectrum predicted by the (non-relativistic) Bohr theory. The second column shows the predicted energy shifts from relativistic corrections arising from the Dirac theory. The third column includes corrections due quantum electrodynamics and the fourth column includes terms for coupling to the nuclear spin degrees of freedom. The H- α line, particularly important in the astronomy, corresponds to the transition between the levels with $n = 2$ and $n = 3$.



To discuss the predicted energy shifts for particular states, it is helpful to introduce some nomenclature from atomic physics. For a state with principal quantum number n , total spin s , orbital angular momentum ℓ , and total angular momentum j , one may use **spectroscopic notation** $n^{2s+1}L_j$ to define the state. For a hydrogen-like atom, with just a single electron, $2s + 1 = 2$. In this case, the factor $2s + 1$ is often just dropped for brevity.

If we apply our perturbative expression for the relativistic corrections (9.5), how do we expect the levels to shift for hydrogen-like atoms? As we have seen, for the non-relativistic Hamiltonian, each state of given n exhibits a $2n^2$ -fold degeneracy. For a given multiplet specified by n , the relativistic corrections depend only on j and n . For $n = 1$, we have $\ell = 0$ and $j = 1/2$: Both $1S_{1/2}$ states, with $m_j = 1/2$ and $-1/2$, experience a negative energy shift by an amount $\Delta E_{1,1/2,m_j,0} = -\frac{1}{4}Z^4\alpha^2$ Ry. For $n = 2$, ℓ can take the values of 0 or 1. With $j = 1/2$, both the former $2S_{1/2}$ state, and the latter $2P_{1/2}$ states share the same negative shift in energy, $\Delta E_{2,1/2,m_j,0} = \Delta E_{2,1/2,m_j,1} = -\frac{5}{64}Z^4\alpha^2$ Ry, while the $2P_{3/2}$ experiences a shift of $\Delta E_{2,3/2,m_j,1} = -\frac{1}{64}Z^4\alpha^2$ Ry. Finally, for $n = 3$, ℓ can take values of 0, 1 or 2. Here, the pairs of states $3S_{1/2}$ and $3P_{1/2}$, and $3P_{3/2}$ and $3D_{3/2}$ each remain degenerate while the state $3D_{5/2}$ is unique. These predicted shifts are summarized in Figure 9.1.

This completes our discussion of the relativistic corrections which develop from the treatment of the Dirac theory for the hydrogen atom. However, this does not complete our description of the “real” hydrogen atom. Indeed, there are further corrections which derive from quantum electrodynamics and nuclear effects which we now turn to address.

9.2.4 INFO: Lamb shift

According to the perturbation theory above, the relativistic corrections which follow from the Dirac theory for hydrogen leave the $2S_{1/2}$ and $2P_{1/2}$ states degenerate. However, in 1947, a careful experimental study by Willis Lamb and Robert Rutherford discovered that this was not in fact the case:⁷ $2P_{1/2}$ state is slightly lower in energy than the $2S_{1/2}$ state resulting in a small shift

⁷W. E. Lamb and R. C. Rutherford, *Fine Structure of the Hydrogen Atom by a Microwave Method*, Phys. Rev. **72**, 241 (1947).

Willis Eugene Lamb, 1913-2008
A physicist who won the Nobel Prize in Physics in 1955 “for his discoveries concerning the fine structure of the hydrogen spectrum”. Lamb and Polykarp Kusch were able to precisely determine certain electromagnetic properties of the electron.



of the corresponding spectral line – the Lamb shift. It might seem that such a tiny effect would be deemed insignificant, but in this case, the observed shift (which was explained by Hans Bethe in the same year) provided considerable insight into quantum electrodynamics.

In quantum electrodynamics, a quantized radiation field has a zero-point energy equivalent to the mean-square electric field so that even in a vacuum there are fluctuations. These fluctuations cause an electron to execute an oscillatory motion and its charge is therefore smeared. If the electron is bound by a non-uniform electric field (as in hydrogen), it experiences a different potential from that appropriate to its mean position. Hence the atomic levels are shifted. In hydrogen-like atoms, the smearing occurs over a length scale,

$$\langle(\delta\mathbf{r})^2\rangle \simeq \frac{2\alpha}{\pi} \left(\frac{\hbar}{mc}\right)^2 \ln \frac{1}{\alpha Z},$$

some five orders of magnitude smaller than the Bohr radius. This causes the electron spin *g*-factor to be slightly different from 2,

$$g_s = 2 \left(1 + \frac{\alpha}{2\pi} - 0.328 \frac{\alpha^2}{\pi^2} + \dots\right).$$

There is also a slight weakening of the force on the electron when it is very close to the nucleus, causing the $2S_{1/2}$ electron (which has penetrated all the way to the nucleus) to be slightly higher in energy than the $2P_{1/2}$ electron. Taking into account these corrections, one obtains a positive energy shift

$$\Delta E_{\text{Lamb}} \simeq \left(\frac{Z}{n}\right)^4 n\alpha^2 \text{Ry} \times \left(\frac{8}{3\pi}\alpha \ln \frac{1}{\alpha Z}\right) \delta_{\ell,0},$$

for states with $\ell = 0$.

9.2.5 Hyperfine structure

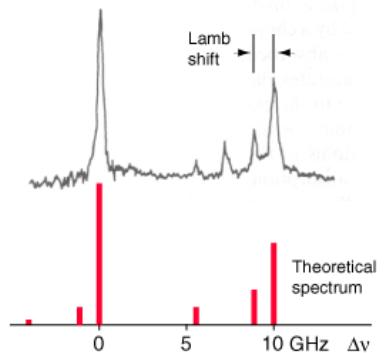
So far, we have considered the nucleus as simply a massive point charge responsible for the large electrostatic interaction with the charged electrons which surround it. However, the nucleus has a spin angular momentum which is associated with a further set of **hyperfine corrections** to the atomic spectra of atoms. As with electrons, the protons and neutrons that make up a nucleus are fermions, each with intrinsic spin 1/2. This means that a nucleus will have some total nuclear spin which is labelled by the quantum number, I . The latter leads to a nuclear magnetic moment,

$$\boldsymbol{\mu}_N = g_N \frac{Ze}{2M_N} \boldsymbol{I}.$$

where M_N denotes the mass of the nucleus, and g_N denotes the nuclear *g*-factor. Since the nucleus has internal structure, the nuclear *g*-factor is not simply 2 as it (nearly) is for the electron. For the proton, the sole nuclear constituent of atomic hydrogen, $g_P \approx 5.56$. Even though the neutron is charge neutral, its *g*-factor is about -3.83 . (The constituent quarks have *g*-factors of 2 (plus corrections) like the electron but the problem is complicated by the strong interactions which make it hard to define a quark’s mass.) We can compute (to some accuracy) the *g*-factor of nuclei from that of protons and neutrons as we can compute the proton’s *g*-factor from its quark constituents. Since the nuclear mass is several orders of magnitude higher than that of the electron, the nuclear magnetic moment provides only a small perturbation.

Hans Albrecht Bethe 1906-2005

A German-American physicist, and Nobel laureate in physics “for his work on the theory of stellar nucleosynthesis.” A versatile theoretical physicist, Bethe also made important contributions to quantum electrodynamics, nuclear physics, solid-state physics and particle astrophysics. During World War II, he was head of the Theoretical Division at the secret Los Alamos laboratory developing the first atomic bombs. There he played a key role in calculating the critical mass of the weapons, and did theoretical work on the implosion method used in both the Trinity test and the “Fat Man” weapon dropped on Nagasaki.



Hydrogen fine structure and hyperfine structure for the $n = 3$ to $n = 2$ transition (see Fig. 9.1).

According to classical electromagnetism, the magnetic moment generates a magnetic field

$$\mathbf{B} = \frac{\mu_0}{4\pi r^3} (3(\boldsymbol{\mu}_N \cdot \mathbf{e}_r) \mathbf{e}_r - \boldsymbol{\mu}_N) + \frac{2\mu_0}{3} \boldsymbol{\mu}_N \delta^{(3)}(\mathbf{r}).$$

To explore the effect of this field, let us consider just the s -electrons, i.e. $\ell = 0$, for simplicity.⁸ In this case, the interaction of the magnetic moment of the electrons with the field generated by the nucleus, gives rise to the hyperfine interaction,

$$\hat{H}_{\text{hyp}} = -\boldsymbol{\mu}_e \cdot \mathbf{B} = \frac{e}{m} \hat{\mathbf{S}} \cdot \mathbf{B}.$$

For the $\ell = 0$ state, the first contribution to \mathbf{B} vanishes while second leads to the first order correction,

$$\langle H_{\text{hyp}} \rangle_{n,1/2,0} = \left(\frac{Z}{n}\right)^4 n \alpha^2 \text{Ry} \times \frac{8}{3} g_N \frac{m}{M_N} \frac{1}{\hbar^2} \mathbf{S} \cdot \mathbf{I}.$$

Once again, to evaluate the expectation values on the spin degrees of freedom, it is convenient to define the total spin $\mathbf{F} = \mathbf{I} + \mathbf{S}$. We then have

$$\begin{aligned} \frac{1}{\hbar^2} \mathbf{S} \cdot \mathbf{I} &= \frac{1}{2\hbar^2} (\mathbf{F}^2 - \mathbf{S}^2 - \mathbf{I}^2) = \frac{1}{2} (F(F+1) - 3/4 - I(I+1)) \\ &= \frac{1}{2} \begin{cases} I & F = I + 1/2 \\ -I - 1 & F = I - 1/2 \end{cases} \end{aligned}$$

Therefore, the $1s$ state of Hydrogen is split into two, corresponding to the two possible values $F = 0$ and 1 . The transition between these two levels has frequency 1420 MHz, or wavelength 21 cm, so lies in the radio waveband. It is an important transition for radio astronomy. A further contribution to the hyperfine structure arises if the **nuclear shape** is not spherical thus distorting the Coulomb potential; this occurs for deuterium and for many other nuclei.

Finally, before leaving this section, we should note that the nucleus is not point-like but has a small size. The effect of **finite nuclear size** can be estimated perturbatively. In doing so, one finds that the s ($\ell = 0$) levels are those most affected, because these have the largest probability of finding the electron close to the nucleus; but the effect is still very small in hydrogen. It can be significant, however, in atoms of high nuclear charge Z , or for **muonic** atoms.

This completes our discussion of the “one-electron” theory. We now turn to consider the properties of multi-electron atoms.

A **muon** is a particle somewhat like an electron, but about 200 times heavier. If a muon is captured by an atom, the corresponding Bohr radius is 200 times smaller, thus enhancing the nuclear size effect.

9.3 Multi-electron atoms

To address the electronic structure of a multi-electron atom, we might begin with the hydrogenic energy levels for an atom of nuclear charge Z , and start filling the lowest levels with electrons, accounting for the exclusion principle. The degeneracy for quantum numbers (n, ℓ) is $2 \times (2\ell + 1)$, where $(2\ell + 1)$ is the number of available m_ℓ values, and the factor of 2 accounts for the spin degeneracy. Hence, the number of electrons accommodated in shell, n , would be $2 \times n^2$,

⁸For a full discussion of the influence of the orbital angular momentum, we refer to [6].

| n | ℓ | Degeneracy in shell | Cumulative total |
|-----|------------|---------------------------------|------------------|
| 1 | 0 | 2 | 2 |
| 2 | 0, 1 | $(1 + 3) \times 2 = 8$ | 10 |
| 3 | 0, 1, 2 | $(1 + 3 + 5) \times 2 = 18$ | 28 |
| 4 | 0, 1, 2, 3 | $(1 + 3 + 5 + 7) \times 2 = 32$ | 60 |

We would therefore expect that atoms containing 2, 10, 28 or 60 electrons would be especially stable, and that in atoms containing one more electron than this, the outermost electron would be less tightly bound. In fact, if we look at data (Fig. 9.2) recording the first ionization energy of atoms, i.e. the minimum energy needed to remove one electron, we find that the noble gases, having $Z = 2, 10, 18, 36 \dots$ are especially tightly bound, and the elements containing one more electron, the alkali metals, are significantly less tightly bound.

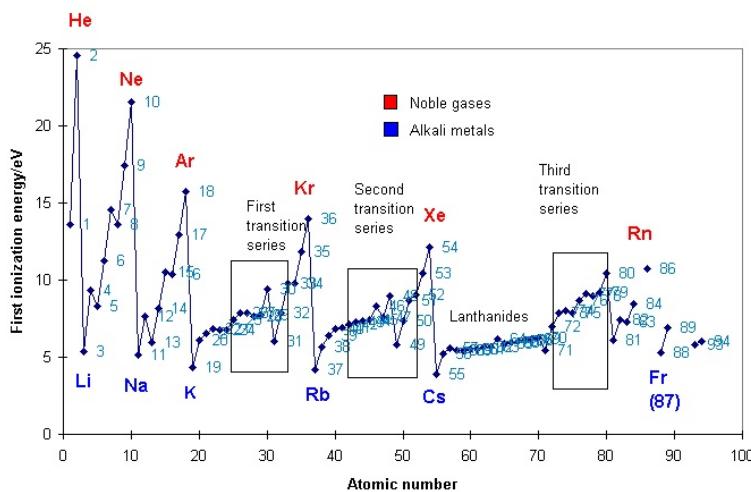


Figure 9.2: Ionization energies of the elements.

The reason for the failure of this simple-minded approach is that we have neglected the repulsion between electrons. In fact, the first ionization energies of atoms show a relatively weak dependence on Z ; this tells us that the outermost electrons are almost completely shielded from the nuclear charge.⁹ Indeed, when we treated the Helium atom as an example of the variational method in chapter 7 we found that the effect of electron-electron repulsion was sizeable, and really too large to be treated accurately by perturbation theory.

9.3.1 Central field approximation

Leaving aside for now the influence of spin or relativistic effects, the Hamiltonian for a multi-electron atom can be written as

$$\hat{H} = \sum_i \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_i} \right] + \sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}},$$

where $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$. The first term represents the “single-particle” contribution to the Hamiltonian arising from interaction of each electron with the nucleus, while the last term represents the mutual Coulomb interaction between the constituent electrons. It is this latter term that makes the generic

⁹In fact, the shielding is not completely perfect. For a given energy shell, the effective nuclear charge varies for an atomic number Z as $Z_{\text{eff}} \sim (1 + \alpha)^Z$ where $\alpha > 0$ characterizes the ineffectiveness of screening. This implies that the ionization energy $I_Z = -E_Z \sim Z_{\text{eff}}^2 \sim (1 + 2\alpha Z)$. The near-linear dependence of I_Z on Z is reflected in Fig. 9.2.

problem “many-body” in character and therefore very complicated. Yet, as we have already seen in the perturbative analysis of the excited states of atomic Helium, this term can have important physical consequences both on the overall energy of the problem and on the associated spin structure of the states.

The **central field approximation** is based upon the observation that the electron interaction term contains a large central (spherically symmetric) component arising from the “core electrons”. From the following relation,

$$\sum_{m_\ell=-\ell}^{\ell} |Y_{\ell m_\ell}(\theta, \phi)|^2 = \text{const.}$$

it is apparent that a **closed shell** has an electron density distribution which is isotropic (independent of θ and ϕ). We can therefore develop a perturbative scheme by setting $\hat{H} = \hat{H}_0 + \hat{H}_1$, where

$$\hat{H}_0 = \sum_i \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_i} + U_i(r_i) \right], \quad \hat{H}_1 = \sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}} - \sum_i U_i(r_i).$$

Here the one-electron potential, $U_i(r)$, which is assumed central (see below), incorporates the “average” effect of the other electrons. Before discussing how to choose the potentials $U_i(r)$, let us note that \hat{H}_0 is separable into a sum of terms for each electron, so that the total wavefunction can be factorized into components for each electron. The basic idea is first to solve the Schrödinger equation using \hat{H}_0 , and then to treat \hat{H}_1 as a small perturbation.

On general grounds, since the Hamiltonian \hat{H}_0 continues to commute with the angular momentum operator, $[\hat{H}_0, \hat{\mathbf{L}}] = 0$, we can see that the eigenfunctions of \hat{H}_0 will be characterized by quantum numbers (n, ℓ, m_ℓ, m_s) . However, since the effective potential is no longer Coulomb-like, the ℓ values for a given n need not be degenerate. Of course, the difficult part of this procedure is to estimate $U_i(r)$; the potential energy experienced by each electron depends on the wavefunction of all the other electrons, which is only known after the Schrödinger equation has been solved. This suggests that an iterative approach to solving the problem will be required.

To understand how the potentials $U_i(r)$ can be estimated – the **self-consistent field method** – it is instructive to consider a variational approach due originally to Hartree. If electrons are considered independent, the wavefunction can be factorized into the product state,

$$\Psi(\{\mathbf{r}_i\}) = \psi_{i_1}(\mathbf{r}_1)\psi_{i_2}(\mathbf{r}_2) \cdots \psi_{i_N}(\mathbf{r}_N),$$

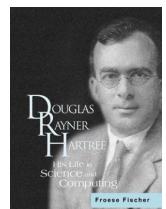
where the quantum numbers, $i_k \equiv (nlm_\ell m_s)_k$, indicate the individual state occupancies. Note that this product state is not a properly antisymmetrized Slater determinant – the exclusion principle is taken into account only in so far as the energy of the ground state is taken to be the lowest that is consistent with the assignment of different quantum numbers, $nlm_\ell m_s$ to each electron. Nevertheless, using this wavefunction as a trial state, the variational energy is then given by

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \sum_i \int d^3r \psi_i^* \left(-\frac{\hbar^2 \nabla^2}{2m} - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i + \frac{1}{4\pi\epsilon_0} \sum_{i < j} \int d^3r \int d^3r' \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}') \psi_i(\mathbf{r}).$$

Now, according to the variational principle, we must minimize the energy functional by varying $E[\{\psi_i\}]$ with respect to the complex wavefunction, ψ_i ,

Douglas Rayner Hartree FRS 1897-1958

An English mathematician and physicist most famous for the development of numerical analysis and its application to atomic physics. He entered St John's College Cambridge in 1915 but World War I interrupted his studies and he joined a team studying anti-aircraft gunnery. He returned to Cambridge after the war and graduated in 1921 but, perhaps because of his interrupted studies, he only obtained a second class degree in Natural Sciences. In 1921, a visit by Niels Bohr to Cambridge inspired him to apply his knowledge of numerical analysis to the solution of differential equations for the calculation of atomic wavefunctions.



subject to the normalization condition, $\langle \psi_i | \psi_i \rangle = 1$. The latter can be imposed using a set of Lagrange multipliers, ε_i , i.e.

$$\frac{\delta}{\delta \psi_i^*} \left[E - \varepsilon_i \left(\int d^3r |\psi_i(\mathbf{r})|^2 - 1 \right) \right] = 0.$$

Following the variation,¹⁰ one obtains the **Hartree equations**,

$$\begin{aligned} \left(-\frac{\hbar^2 \nabla^2}{2m} - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}) \\ = \varepsilon_i \psi_i(\mathbf{r}). \end{aligned} \quad (9.6)$$

Then according to the variational principle, amongst all possible trial functions ψ_i , the set that minimizes the energy are determined by the effective potential,

$$U_i(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}.$$

Equation (9.6) has a simple interpretation: The first two terms relate to the nuclear potential experienced by the individual electrons, while the third term represents the electrostatic potential due to the other electrons. However, to simplify the procedure, it is useful to engineer the radial symmetry of the potential by replacing $U_i(\mathbf{r})$ by its spherical average,

$$U_i(\mathbf{r}) \mapsto U_i(r) = \int \frac{d\Omega}{4\pi} U_i(\mathbf{r}).$$

Finally, to relate the Lagrange multipliers, ε_i (which have the appearance of one-electron energies), to the total energy, we can multiply Eq. (9.6) by $\psi_i^*(\mathbf{r})$ and integrate,

$$\begin{aligned} \varepsilon_i = \int d^3r \psi_i^* \left(-\frac{\hbar^2 \nabla^2}{2m} - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i \\ + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' d^3r |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} |\psi_i(\mathbf{r})|^2. \end{aligned}$$

If we compare this expression with the variational state energy, we find that

$$E = \sum_i \varepsilon_i - \frac{1}{4\pi\epsilon_0} \sum_{i < j} \int d^3r' d^3r |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} |\psi_i(\mathbf{r})|^2. \quad (9.7)$$

To summarize, if we wish to implement the central field approximation to determine the states of a multi-electron atom, we must follow the algorithm:

- Firstly, one makes an initial “guess” for a (common) central potential, $U(r)$. As $r \rightarrow 0$, screening becomes increasingly ineffective and we expect $U(r) \rightarrow 0$. As $r \rightarrow \infty$, we anticipate that $U(r) \rightarrow \frac{1}{4\pi\epsilon_0} \frac{(Z-1)e^2}{r}$, corresponding to perfect screening. So, as a starting point, we make take some smooth function $U(r)$ interpolating between these limits. For this trial potential, we can solve (numerically) for the eigenstates of the single-particle Hamiltonian. We can then use these states as a platform to build the product wavefunction and in turn determine the self-consistent potentials, $U_i(r)$.

¹⁰Note that, in applying the variation, the wavefunction ψ_i^* can be considered independent of ψ_i – you might like to think why.

2. With these potentials, $U_i(r)$, we can determine a new set of eigenstates for the set of Schrödinger equations,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} + U_i(r) \right] \psi_i = \varepsilon_i \psi_i .$$

3. An estimate for the ground state energy of an atom can be found by filling up the energy levels, starting from the lowest, and taking account of the exclusion principle.
4. Using these wavefunctions, one can then make an improved estimate of the potentials $U_i(r_i)$ and return to step 2 iterating until convergence.

Since the practical implementation of such an algorithm demands a large degree of computational flair, if you remain curious, you may find it useful to refer to the Mathematica code prepared by Ref.[4] where both the Hartree and the Hartree-Fock procedures (described below) are illustrated.

▷ INFO. An improvement to this procedure, known the **Hartree-Fock** method, takes account of exchange interactions. In order to do this, it is necessary to ensure that the wavefunction, including spin, is antisymmetric under interchange of any pair of electrons. This is achieved by introducing the **Slater determinant**. Writing the individual electron wavefunction for the i^{th} electron as $\psi_k(\mathbf{r}_i)$, where $i = 1, 2 \dots N$ and k is shorthand for the set of quantum numbers ($n\ell m_\ell m_s$), the overall wavefunction is given by

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \psi_1(\mathbf{r}_3) & \dots \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \psi_2(\mathbf{r}_3) & \dots \\ \psi_3(\mathbf{r}_1) & \psi_3(\mathbf{r}_2) & \psi_3(\mathbf{r}_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} .$$

Note that each of the $N!$ terms in Ψ is a product of wavefunctions for each individual electron. The $1/\sqrt{N!}$ factor ensures the wavefunction is normalized. A determinant changes sign if any two columns are exchanged, corresponding to $\mathbf{r}_i \leftrightarrow \mathbf{r}_j$ (say); this ensures that the wavefunction is antisymmetric under exchange of electrons i and j . Likewise, a determinant is zero if any two rows are identical; hence all the ψ_k s must be different and the Pauli exclusion principle is satisfied.¹¹ In this approximation, a variational analysis leads to the Hartree-Fock equations (exercise),

$$\begin{aligned} \varepsilon_i \psi_i(\mathbf{r}) &= \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right] \psi_i(\mathbf{r}) \\ &+ \sum_{j \neq i} \int d^3 r_j \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_j^*(\mathbf{r}') \left[\psi_j(\mathbf{r}') \psi_i(\mathbf{r}) - \psi_j(\mathbf{r}) \psi_i(\mathbf{r}') \delta_{m_{s_i}, m_{s_j}} \right] . \end{aligned}$$

The first term in the last set of brackets translates to the ordinary Hartree contribution above and describes the influence of the charge density of the other electrons, while the second term describes the non-local **exchange** contribution, a manifestation of particle statistics.

The outcome of such calculations is that the eigenfunctions are, as for hydrogen, characterized by quantum numbers n, ℓ, m_ℓ , with $\ell < n$, but that the states with different ℓ for a given n are not degenerate, with the lower values of ℓ lying lower. This is because, for the higher ℓ values, the electrons tend to lie further from the nucleus on average, and are therefore more effectively screened. The states corresponding to a particular value of n are generally

Vladimir Aleksandrovich Fock
1898-1974
A Soviet physicist, who did foundational work on quantum mechanics and quantum electrodynamics. His primary scientific contribution lies in the development of quantum physics, although he also contributed significantly to the fields of mechanics, theoretical optics, theory of gravitation, physics of continuous medium. In 1926 he derived the Klein-Gordon equation. He gave his name to Fock space, the Fock representation and Fock state, and developed the Hartree-Fock method in 1930. Fock made significant contributions to general relativity theory, specifically for the many body problems.



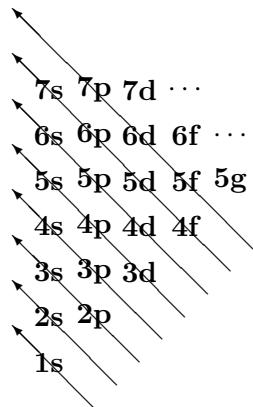
¹¹Note that for $N = 2$, the determinant reduces to the familiar antisymmetric wavefunction, $\frac{1}{\sqrt{2}}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) - \psi_2(\mathbf{r}_1)\psi_1(\mathbf{r}_2)]$.

referred to as a **shell**, and those belonging to a particular pair of values of n, ℓ are usually referred to as a **subshell**. The energy levels are ordered as below (with the lowest lying on the left):

| Subshell name | 1s | 2s | 2p | 3s | 3p | 4s | 3d | 4p | 5s | 4d | ... |
|---------------|----|----|----|----|----|----|----|----|----|----|-----|
| $n =$ | 1 | 2 | 2 | 3 | 3 | 4 | 3 | 4 | 5 | 4 | ... |
| $\ell =$ | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 1 | 0 | 2 | ... |
| Degeneracy | 2 | 2 | 6 | 2 | 6 | 2 | 10 | 6 | 2 | 10 | ... |
| Cumulative | 2 | 4 | 10 | 12 | 18 | 20 | 30 | 36 | 38 | 48 | ... |

Note that the values of Z corresponding to the noble gases, 2, 10, 18, 36, at which the ionization energy is unusually high, now emerge naturally from this filling order, corresponding to the numbers of electrons just before a new shell (n) is entered. There is a handy mnemonic to remember this filling order. By writing the subshells down as shown right, the order of states can be read off along diagonals from lower right to upper left, starting at the bottom.

We can use this sequence of energy levels to predict the ground state electron configuration of atoms. We simply fill up the levels starting from the lowest, accounting for the exclusion principle, until the electrons are all accommodated (the **aufbau principle**). Here are a few examples:



| Z | Element | Configuration | $^{2S+1}L_J$ | Ioniz. Pot. (eV) |
|----|---------|--|--------------|------------------|
| 1 | H | (1s) | $^2S_{1/2}$ | 13.6 |
| 2 | He | (1s) ² | 1S_0 | 24.6 |
| 3 | Li | He (2s) | $^2S_{1/2}$ | 5.4 |
| 4 | Be | He (2s) ² | 1S_0 | 9.3 |
| 5 | B | He (2s) ² (2p) | $^2P_{1/2}$ | 8.3 |
| 6 | C | He (2s) ² (2p) ² | 3P_0 | 11.3 |
| 7 | N | He (2s) ² (2p) ³ | $^4S_{3/2}$ | 14.5 |
| 8 | O | He (2s) ² (2p) ⁴ | 3P_2 | 13.6 |
| 9 | F | He (2s) ² (2p) ⁵ | $^2P_{3/2}$ | 17.4 |
| 10 | Ne | He (2s) ² (2p) ⁶ | 1S_0 | 21.6 |
| 11 | Na | Ne (3s) | $^2S_{1/2}$ | 5.1 |

Since it is generally the outermost electrons which are of most interest, contributing to chemical activity or optical spectra, one often omits the inner closed shells, and just writes O as (2p)⁴, for example. However, the configuration is not always correctly predicted, especially in the heavier elements, where levels may be close together. It may be favourable to promote one or even two electrons one level above that expected in this simple picture, in order to achieve a filled shell. For example, Cu ($Z = 29$) would be expected to have configuration $\dots (4s)^2 (3d)^9$, and actually has configuration $\dots (4s)^1 (3d)^{10}$. There are several similar examples in the transition elements where the d subshells are being filled, and many among the lanthanides (rare earths) and actinides where f subshells are being filled.

▷ INFO. Since the assignment of an electron configuration requires only the enumeration of the values of n and ℓ for all electrons, but not those of m_ℓ and m_s , each configuration will be accompanied by a **degeneracy** g . If $\nu_{n\ell}$ denotes the number of electrons occupying a given level $E_{n,\ell}$, and $\delta_\ell = 2 \times (2\ell + 1)$ is the degeneracy of that level, there are

$$d_{n\ell} = \frac{\delta_\ell!}{\nu_{n\ell}!(\delta_\ell - \nu_{n\ell})!} \quad (9.8)$$

ways of distributing the $\nu_{n\ell}$ electrons among the δ_ℓ individual states. The total degeneracy, g , is then obtained from the product.

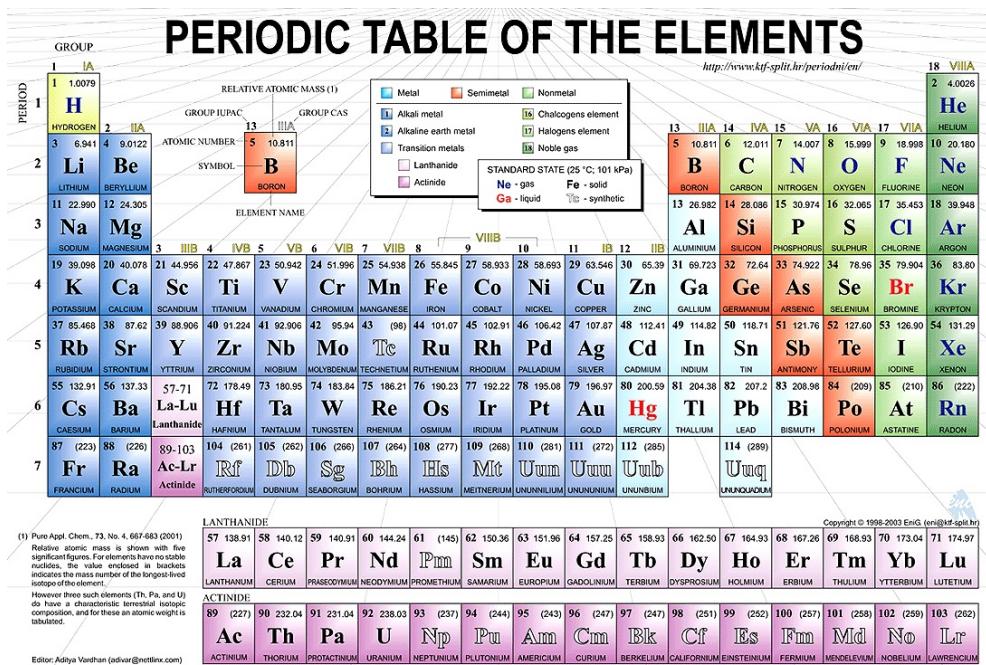


Figure 9.3: Periodic table of elements.

This scheme provides a basis to understand the **periodic table of elements** (see Fig. 9.3). We would expect that elements which have similar electron configurations in their outermost shells (such as Li, Na, K, Rb, Cs, Fr which all have $(ns)^1$ or F, Cl, Br, I, which all have $(np)^5$) would have similar chemical properties, such as valency, since it is the unpaired outer electrons which especially participate in chemical bonding. Therefore, if one arranges the atoms in order of increasing atomic number Z (which equals the number of electrons in the atom), periodic behaviour is seen whenever a new subshell of a given ℓ is filled.

9.3.2 Spin-orbit coupling

The procedure outlined above allows us to predict the occupation of subshells in an atomic ground state. This is not in general sufficient to specify the ground state fully. If there are several electrons in a partially filled subshell, then their spins and orbital angular momenta can combine in several different ways, to give different values of total angular momentum, with different energies. In order to deal with this problem, it is necessary to consider the spin-orbit interaction as well as the residual Coulomb interaction between the outer electrons.

Schematically we can write the Hamiltonian for this system as follows:

$$\hat{H} \approx \hat{H}_0 + \underbrace{\sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}} - \sum_i U_i(r_i)}_{\hat{H}_1} + \underbrace{\sum_i \xi_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}_i}_{\hat{H}_2},$$

where \hat{H}_0 includes the kinetic energy and central field terms, \hat{H}_1 is the residual Coulomb interaction, and (with $\xi(r) = \frac{1}{2m^2c^2} \frac{1}{r} (\partial_r V(r))$) \hat{H}_2 is the spin-orbit interaction. We can then consider two possible scenarios:

$\hat{H}_2 \ll \hat{H}_1$: This tends to apply in the case of light atoms. In this situation, one considers first the eigenstates of $\hat{H}_0 + \hat{H}_1$, and then treats \hat{H}_2 as a

perturbation. This leads to a scheme called **LS** (or **Russell-Saunders**) **coupling**.

$\hat{H}_2 \gg \hat{H}_1$: This can apply in very heavy atoms, or in heavily ionized light atoms, in which the electrons are moving at higher velocities and relativistic effects such as the spin-orbit interaction are more important. In this case, a scheme called **jj coupling** applies.

It is important to emphasise that both of these scenarios represent approximations; real atoms do not always conform to the comparatively simple picture which emerges from these schemes.

LS coupling scheme

In this approximation, we start by considering the eigenstates of $\hat{H}_0 + \hat{H}_1$. We note that this Hamiltonian must commute with the total angular momentum $\hat{\mathbf{J}}^2$ (because of invariance under rotations in space), and also clearly commutes with the total spin $\hat{\mathbf{S}}^2$. It also commutes with the total orbital angular momentum $\hat{\mathbf{L}}^2$, since \hat{H}_1 only involves internal interactions, and must therefore be invariant under global rotation of all the electrons. Therefore the energy levels can be characterised by the corresponding total angular momentum quantum numbers L, S, J . Their ordering in energy is given by **Hund's rules**:

1. Combine the spins of the electrons to obtain possible values of total spin S . The largest permitted value of S lies lowest in energy.
2. For this value of S , find the possible values of total angular momentum L . The largest value of L lies lowest in energy.
3. Couple the values of L and S to obtain the values of J (hence the name of the scheme). If the subshell is less than half full, the smallest value of J lies lowest; otherwise, the largest value of J lies lowest.

In deciding on the permitted values of L and S , in addition to applying the usual rules for adding angular momenta, one also has to ensure that the exclusion principle is respected, as we will see later when considering some examples.

These rules are empirical; there are exceptions, especially to the L and J rules (2 and 3). Nevertheless, Hund's rules are a useful guide, and we should try to understand their physical origin.

1. Maximising S makes the spin wavefunction as symmetric as possible. This tends to make the spatial wavefunction antisymmetric, and hence reduces the Coulomb repulsion, as we saw when discussing the exchange interactions in Helium and the origin of itinerant ferromagnetism.
2. Maximising L also tends to keep the electrons apart. This is less obvious, though a simple classical picture of electrons rotating round the nucleus in the same or different senses makes it at least plausible.
3. The separation of energies for states of different J arises from treating the spin-orbit term \hat{H}_2 as a perturbation (fine structure). It can be shown (using the Wigner-Eckart theorem – beyond the scope of these lectures) that

$$\begin{aligned} & \langle J, m_J, L, S | \sum_i \xi_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}_i | J, m_J, L, S \rangle \\ &= \zeta(L, S) \langle J, m_J, L, S | \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} | J, m_J, L, S \rangle \\ &= \frac{\zeta(L, S)}{2} [J(J+1) - L(L+1) - S(S+1)], \end{aligned} \quad (9.9)$$

where the matrix element $\zeta(L, S)$ depends on the total L and S values. Since one may show that the sign of $\zeta(L, S)$ changes according to whether the subshell is more or less than half-filled, the third Hund's rule is established.

To understand the application of LS coupling, it is best to work through some examples. Starting with the simplest multi-electron atom, **helium**, the ground state has an electron configuration $(1s)^2$, and must therefore have $L = S = J = 0$. In fact, for any completely filled subshell, we have $L = S = 0$ and hence $J = 0$, since the total m_L and m_S must equal zero if all substates are occupied. Consider now an excited state of helium, e.g. $(1s)^1(2p)^1$, in which one electron has been excited to the $2p$ level. We can now have $S = 1$ or $S = 0$, with the $S = 1$ state lying lower in energy according to Hund's rules. Combining the orbital angular momenta of the electrons yields $L = 1$ and thus, with $S = 0, J = 1$, while with $S = 1, J = 0, 1, 2$ with $J = 0$ lying lowest in energy.

Once again, as with the hydrogen-like states, we may index the states of multi-electron atoms by spectroscopic term notation, $^{2S+1}L_J$. The superscript $2S + 1$ gives the multiplicity of J values into which the level is split by the spin-orbit interaction; the L value is represented by a capital letter, S, P, D , etc., and J is represented by its numerical value. Thus, for the $(1s)^1(2p)^1$ state of helium, there are four possible states, with terms:

$$^3P_0 \quad ^3P_1 \quad ^3P_2 \quad ^1P_1,$$

where the three 3P states are separated by the spin-orbit interaction, and the singlet 1P state lies much higher in energy owing to the Coulomb interaction. The separations between the 3P_2 and 3P_1 and the 3P_1 and 3P_0 should be in the ratio 2:1. This is an example of the **Landé interval rule**, which states that the separation between a pair of adjacent levels in a fine structure multiplet is proportional to the larger of the two J values involved. This is easily shown using Eq. (9.9) – the separation in energy between states J and $J - 1$ is

$$\propto J(J + 1) - (J - 1)J = 2J.$$

Actually in the case of helium the situation is a bit more complicated, because it turns out that the spin-orbit interaction between *different* electrons makes a non-negligible additional contribution to the fine structure. Other excited states of helium, of the form $(1s)^1(n\ell)^1$, can be handled similarly, and again separate into singlet and triplet states.

▷ EXERCISE. For the case of **boron**, with the electron configuration $(1s)^2(2s)^2(2p)$, use Hund's rules to show that the ground state is $^2P_{1/2}$.

We next consider the case of **carbon**, which has ground state electron configuration $(1s)^2(2s)^2(2p)^2$. This introduces a further complication; we now have two *identical* electrons in the same *unfilled* subshell, and we need to ensure that their wavefunction is antisymmetric with respect to electron exchange. The total spin can either be the singlet $S = 0$ state, which has an antisymmetric wavefunction $\frac{1}{\sqrt{2}}[|\uparrow_1\rangle \otimes |\downarrow_2\rangle - |\downarrow_1\rangle \otimes |\uparrow_2\rangle]$, or one of the triplet $S = 1$ states, which are symmetric, $\frac{1}{\sqrt{2}}[|\uparrow_1\rangle \otimes |\downarrow_2\rangle + |\downarrow_1\rangle \otimes |\uparrow_2\rangle]$, $|\uparrow_1\rangle \otimes |\uparrow_2\rangle$ or $|\downarrow_1\rangle \otimes |\downarrow_2\rangle$. We must therefore choose values of L with the appropriate symmetry to partner each value of S . To form an antisymmetric state, the two electrons must have different values of m_ℓ , so the possibilities are as shown

| $m_\ell^{(1)}$ | $m_\ell^{(2)}$ | m_L |
|----------------|----------------|-------|
| 1 | 0 | 1 |
| 1 | -1 | 0 |
| 0 | -1 | -1 |

right. Inspecting the values of m_L we can deduce that $L = 1$.¹² By contrast, to form a symmetric total angular momentum state, the two electrons may have any values of m_ℓ , leading to the possibilities shown right. Inspecting the values of m_L we infer that $L = 2$ or 0.

We must therefore take $S = 1$ with $L = 1$ and $S = 0$ with $L = 2$ or 0. Finally, to account for the fine structure, we note that the states with $S = 1$ and $L = 1$ can be combined into a single $J = 0$ state, three $J = 1$ states, and five $J = 2$ states leading to the terms 3P_0 , 3P_1 , and 3P_2 respectively. Similarly the $S = 0$, $L = 2$ state can be combined to give five $J = 2$ states, 1D_2 , while the $S = 0$, $L = 0$ state gives the single $J = 0$ state, 1S_0 . Altogether we recover the $1 + 3 + 5 + 5 + 1 = 15$ possible states (cf. Eq. (9.8) with the ordering in energy given by Hund's rules (shown to the right). The experimental energy values are given using the conventional spectroscopic units of inverse wavelength. Note that the Landé interval rule is approximately obeyed by the fine structure triplet, and that the separation between L and S values caused by the electron-electron repulsion is *much* greater than the spin-orbit effect.

In an excited state of carbon, e.g. $(2p)^1(3p)^1$, the electrons are no longer equivalent, because they have different radial wavefunctions. So now one can combine any of $S = 0, 1$ with any of $L = 0, 1, 2$, yielding the following terms (in order of increasing energy, according to Hund's rules):

$$^3D_{1,2,3} \quad ^3P_{0,1,2} \quad ^3S_1 \quad ^1D_2 \quad ^1P_1 \quad ^1S_0.$$

For **nitrogen**, the electron configuration is given by $(1s)^2(2s)^2(2p)^3$. The maximal value of spin is $S = 3/2$ while L can take values 3, 2, 1 and 0. Since the spin wavefunction (being maximal) is symmetric, the spatial wavefunction must be completely antisymmetric. This demands that all three states with $m_\ell = 1, 0, -1$ must be involved. We must therefore have $L = 0$, leading to $J = 3/2$ and the term, $^4S_{3/2}$.

▷ EXERCISE. Construct the $L = 0$ state involving the addition of three $\ell = 1$ angular momentum states. Hint: make use of the total antisymmetry condition.

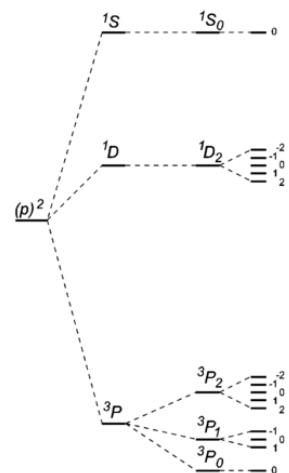
As a final example, let us consider the ground state of **oxygen**, which has electron configuration $(2p)^4$. Although there are four electrons in the $(2p)$ subshell, the maximum value of $S = 1$. This is because there are only three available values of $m_\ell = \pm 1, 0$, and therefore one of these must contain two electrons with opposite spins. Therefore, the maximum value of $m_S = 1$, achieved by having electrons with $m_s = +\frac{1}{2}$ in both the other m_ℓ states. By pursuing this argument, it is quite easy to see that the allowed values of L, S and J are the same as for carbon $(2p)^2$. This is in fact a general result – the allowed quantum numbers for a subshell with n electrons are the same as for that of a subshell with n “holes”. Therefore, the energy levels for the oxygen ground state configuration are the same as for carbon, except that the fine structure multiplet is inverted, in accordance with Hund's third rule.

9.3.3 INFO: jj coupling scheme

When relativistic effects take precedence over electron interaction effects, we must start by considering the eigenstates of $\hat{H}_0 + \hat{H}_2 = \hat{H}_0 + \sum_i \xi_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}_i$.

¹²This result would also be apparent if we recall that angular momentum states are eigenstates of the parity operator with eigenvalue $(-1)^L$. Since there are just two electrons, this result shows that both the $L = 0$ and $L = 2$ wavefunction must be symmetric under exchange.

| $m_\ell^{(1)}$ | $m_\ell^{(2)}$ | m_L |
|----------------|----------------|----------------------|
| 1 | 1 | 2 |
| 1 | 0 | 1 |
| 1 | -1 | 0 |
| 0 | 0 | 0 |
| 0 | -1 | -1 |
| -1 | -1 | -2 |
| | | E / cm^{-1} |
| | | 20649 |
| | | 10195 |
| | | 43 |
| | | 16 |
| | | 0 |



Level scheme of the carbon atom $(1s)^2(2s)^2(2p)^2$. Drawing is not to scale. On the left the energy is shown without any two-particle interaction. The electron-electron interaction leads to a three-fold energy splitting with L and S remaining good quantum numbers. Spin-orbit coupling leads to a further splitting of the states with J remaining a good quantum number. Finally on the right, the levels show Zeeman splittings in an external magnetic field. In this case, the full set of 15 levels become non-degenerate.

These must be eigenstates of $\hat{\mathbf{J}}^2$ as before, because of the overall rotational invariance, and also of $\hat{\mathbf{J}}_i^2$ for each electron. Therefore, in this case, the coupling procedure is to find the allowed j values of individual electrons, whose energies will be separated by the spin-orbit interaction. Then these individual j values are combined to find the allowed values of total J . The effect of the residual Coulomb interaction will be to split the J values for a given set of js . Sadly, in this case, there are no simple rules to parallel those of Hund.

As an example, consider a configuration $(np)^2$ in the jj coupling scheme, to be compared with the example of carbon which we studied in the LS scheme. Combining $s = 1/2$ with $\ell = 1$, each electron can have $j = 1/2$ or $3/2$. If the electrons have the same j value, they are equivalent, so we have to take care of the symmetry of the wavefunction. We therefore have the following possibilities:

- ▷ $j_1 = j_2 = 3/2 \Rightarrow J = 3, 2, 1, 0$, of which $J = 2, 0$ are antisymmetric.
- ▷ $j_1 = j_2 = 1/2 \Rightarrow J = 1, 0$, of which $J = 0$ is antisymmetric.
- ▷ $j_1 = 1/2, j_2 = 3/2 \Rightarrow J = 2, 1$.

In jj coupling, the term is written $(j_1, j_2)_J$, so we have the following terms in our example:

$$(1/2, 1/2)_0 \quad (3/2, 1/2)_1 \quad (3/2, 1/2)_2 \quad (3/2, 3/2)_2 \quad (3/2, 3/2)_0$$

in order of increasing energy. Note that both LS and jj coupling give the same values of J (in this case, two states with $J = 0$, two with $J = 2$ and one with $J = 1$) and in the same order. However, the pattern of levels is different; in LS coupling we found a triplet and two singlets, while in this ideal jj scenario, we have two doublets and a singlet. The sets of states in the two coupling schemes must be expressible as linear combinations of one another, and the physical states for a real atom are likely to differ from either approximation.

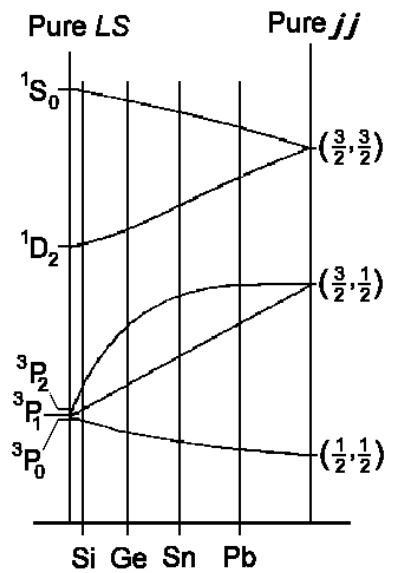
In fact, this idealized form of jj coupling is not seen in the heaviest such atom in the periodic table, lead $(6p)^2$. However, it is seen in some highly ionized states, for example in Cr^{18+} , which has the same electron configuration $(2p)^2$ as carbon, but where, because of the larger unscreened charge on the nucleus, the electrons are moving more relativistically, enhancing the spin-orbit effect. However, a classic example of the transition from LS to jj coupling is seen in the series C-Si-Ge-Sn-Pb in the excited states $(2p)(3s)$, $(3p)(4s)$, \dots $(6p)(7s)$ (see figure right). Here, the electrons are not in the same subshell, so their wavefunctions overlap less, and the Coulomb repulsion is reduced compared to the spin-orbit interaction. Analysing this situation in the LS coupling approximation, one expects a triplet and a singlet:

$$^3\text{P}_{0,1,2} \quad ^1\text{P}_1,$$

while in the jj scheme one expects two doublets:

$$(1/2, 1/2)_{0,1} \quad (1/2, 3/2)_{2,1}.$$

Experimentally, C and Si conform to the LS expectation and Pb to the jj scheme, while Ge and Sn are intermediate.



9.3.4 Zeeman effect

For a multi-electron atom in a weak magnetic field, the appropriate unperturbed states are given by $|J, M_J, L, S\rangle$, where J , L , S refer to the total angular momenta. To determine the Zeeman energy shift, we need to determine the matrix element of \hat{S}_z . To do so, we can make use of the following argument. Since $2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2$, this operator is diagonal in the basis of states, $|J, M_J, L, S\rangle$. Therefore, the matrix element of the operator (exercise, hint: recall that $[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k$ and $[\hat{L}_i, \hat{S}_k] = 0$),

$$-i\hbar\hat{\mathbf{S}} \times \hat{\mathbf{L}} \equiv \hat{\mathbf{S}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) - (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}})\hat{\mathbf{S}}$$

must vanish. Moreover, from the identity $[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{\mathbf{J}}] = 0$, it follows that the matrix element of the vector product,

$$-i\hbar(\hat{\mathbf{S}} \times \hat{\mathbf{L}}) \times \hat{\mathbf{J}} = \hat{\mathbf{S}} \times \hat{\mathbf{J}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) - (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}})\hat{\mathbf{S}} \times \hat{\mathbf{J}},$$

must also vanish. If we expand the left hand side, we thus find that the matrix element of

$$(\hat{\mathbf{S}} \times \hat{\mathbf{L}}) \times \hat{\mathbf{J}} = \hat{\mathbf{L}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) - \hat{\mathbf{S}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{J}}) \stackrel{\hat{\mathbf{L}}=\hat{\mathbf{J}}-\hat{\mathbf{S}}}{=} \hat{\mathbf{J}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) - \hat{\mathbf{S}}\hat{\mathbf{J}}^2,$$

also vanishes. Therefore, it follows that $\langle \hat{\mathbf{S}}\hat{\mathbf{J}}^2 \rangle = \langle \hat{\mathbf{J}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) \rangle$, where the expectation value is taken over the basis states. Then, with $\hat{\mathbf{S}} \cdot \hat{\mathbf{J}} = \frac{1}{2}(\hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 - \hat{\mathbf{L}}^2)$, we have that

$$\langle \hat{S}_z \rangle = \langle \hat{J}_z \rangle \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}.$$

As a result, we can deduce that, at first order in perturbation theory, the energy shift arising from the Zeeman term is given by

$$\Delta E_{J,M_J,L,S} = \mu_B g_J M_J B,$$

where

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)},$$

denotes the effective **Landé g-factor**, which lies between 1 and 2. Note that, in the special case of hydrogen, where $S = 1/2$ and $J = L \pm 1/2$, we recover our previous result. The predicted Zeeman splitting for sodium is shown in figure 9.5.

> INFO. In the **strong field limit**, where the influence of Zeeman term dominates, the appropriate basis states are set by $|L, M_L, S, M_S\rangle$, in which the operators \hat{L}^2 , \hat{L}_z , \hat{S}^2 , \hat{S}_z , and \hat{H}_{Zeeman} are diagonal. In this case, the energy splitting takes the form

$$\Delta E_{L,M_L,S,M_S} = \mu_B B(M_L + 2M_S) + \frac{1}{2}mc^2 \left(\frac{Z\alpha}{n} \right)^4 \frac{nM_L M_S}{\ell(\ell+1/2)(\ell+1)},$$

where the second term arises from the spin-orbit interaction.

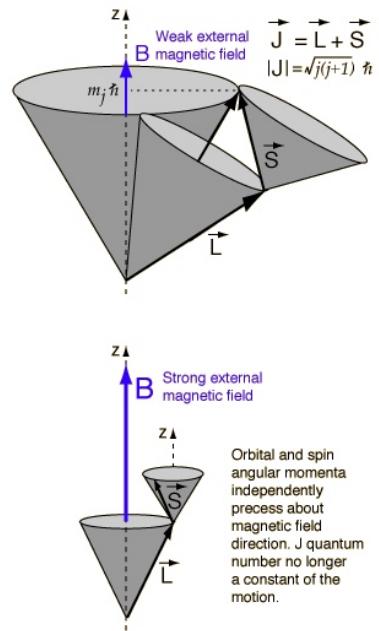


Figure 9.4: In the weak field case, the vector model (top) implies that the coupling of the orbital angular momentum \mathbf{L} to the spin angular momentum \mathbf{S} is stronger than their coupling to the external field. In this case where spin-orbit coupling is dominant, they can be visualized as combining to form a total angular momentum \mathbf{J} which then precesses about the magnetic field direction. In the strong field case, \mathbf{S} and \mathbf{L} couple more strongly to the external magnetic field than to each other, and can be visualized as independently precessing about the external field direction.

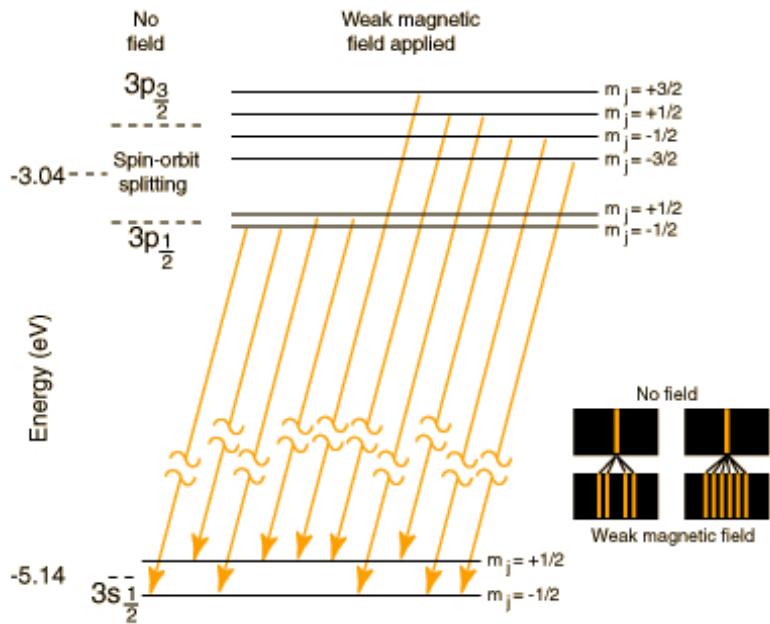


Figure 9.5: The well known doublet which is responsible for the bright yellow light from a sodium lamp may be used to demonstrate several of the influences which cause splitting of the emission lines of atomic spectra. The transition which gives rise to the doublet is from the $3p$ to the $3s$ level. The fact that the $3s$ state is lower than the $3p$ state is a good example of the dependence of atomic energy levels on orbital angular momentum. The $3s$ electron penetrates the $1s$ shell more and is less effectively shielded than the $3p$ electron, so the $3s$ level is lower. The fact that there is a doublet shows the smaller dependence of the atomic energy levels on the total angular momentum. The $3p$ level is split into states with total angular momentum $J = 3/2$ and $J = 1/2$ by the spin-orbit interaction. In the presence of an external magnetic field, these levels are further split by the magnetic dipole energy, showing dependence of the energies on the z -component of the total angular momentum.

Chapter 10

From molecules to solids

In the previous section, we studied the quantum mechanics of multi-electron atoms – the subject of atomic physics. In this section, we will begin to explore how these concepts get translated into systems with many atoms, from simple molecular structures to the solid state. As with atomic physics, the subjects of molecular and solid state physics represent fields in their own right and it would be fanciful to imagine that we could do more than touch on the basic principles. Nevertheless, in establishing the foundations of these subjects, we will see two things: firstly, by organising the hierarchy of approximation schemes, much can be understood about the seemingly complex quantum dynamics of many-particle systems. Secondly, we will find that constraints imposed by symmetries (such as translation) allow a simple phenomenology to emerge from complex solid state systems. We begin our discussion with the molecular system.

A molecule consists of electrons moving in the complicated potential set up by all the constituent electric charges. Even in classical mechanics, it would be extremely difficult to solve the equations of motion of the internal molecular degrees of freedom. Fortunately, for most purposes, we can treat the motion of the electrons and nuclei separately, due to their very different masses. As the forces on a nucleus are similar in magnitude to those that act on an electron, so the electrons and nuclei must have comparable momenta. Therefore the nuclei are typically moving much more slowly than the electrons. In studying the motion of the electrons, we can therefore treat the nuclei as being “nailed down” in fixed positions. Conversely, in studying the nuclear motion (vibrations and rotations of the molecule) we can assume, as a first approximation, that the electrons adjust instantly to changes in the *molecular conformation* defined by the positions of the nuclei. This picture forms the basis of the **Born-Oppenheimer approximation**.

In quantum mechanics, the wavefunction of a molecule, $\Psi(\{\mathbf{r}_n\}, \{\mathbf{R}_N\})$, is a function of the positions of all the electrons and nuclei, and the Hamiltonian has the form

$$\hat{H} = \sum_n \frac{\hat{\mathbf{p}}_n^2}{2m_e} + \sum_N \frac{\hat{\mathbf{p}}_N^2}{2m_N} + V(\{\hat{\mathbf{r}}_n\}, \{\hat{\mathbf{R}}_N\}),$$

where, as usual, the momentum operators, $\hat{\mathbf{p}}_n$ and $\hat{\mathbf{p}}_N$, act only on the corresponding coordinates. Here we have labelled electrons by lower case letters $n = 1, 2, \dots$ to distinguish them from nuclei which are denoted by capitals, $N = a, b, \dots$. The statement that the electrons and nuclei have comparable momenta translates into the fact that $\nabla_n^2 \Psi$ and $\nabla_N^2 \Psi$ will be comparable. Therefore the second term in the Hamiltonian above, the sum over nuclear

J. Robert Oppenheimer 1904-1967

An American theoretical physicist best known for his role as the scientific director of the Manhattan Project: the World War II effort to develop the first nuclear weapons at Los Alamos National Laboratory. In reference to the Trinity test in New Mexico, where his Los Alamos team first tested the bomb, Oppenheimer famously recalled the Bhagavad Gita: “If the radiance of a thousand suns were to burst at once into the sky, that would be like the splendor of the mighty one.” and “Now I am become Death, the destroyer of worlds.” As a scientist, Oppenheimer is remembered most for being the chief founder of the American school of theoretical physics while at the University of California, Berkeley. As director of the Institute for Advanced Study he held Einstein’s old position of Senior Professor of Theoretical Physics. Oppenheimer’s notable achievements in physics include the Born-Oppenheimer approximation, work on electron-positron theory, the Oppenheimer-Phillips process, quantum tunneling, special relativity, quantum mechanics, quantum field theory, black holes, and cosmic rays.



kinetic energies, can be neglected as a first approximation when we solve for the dependence of Ψ on the electron position vectors $\{\mathbf{r}_n\}$.

In the Born-Oppenheimer approximation, the time-independent Schrödinger equation for the electronic motion is therefore given by

$$\left[-\sum_n \frac{\hbar^2 \nabla_n^2}{2m_e} + V(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}) \right] \psi_k(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}) = E_k(\{\mathbf{R}_N\}) \psi_k(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}),$$

where the eigenfunctions ψ_k , with $k = 0, 1, 2, \dots$, describe the electronic ground state and excited states with the nuclei nailed down at positions $\mathbf{R}_a, \mathbf{R}_b, \dots$; E_k are the corresponding energy levels. Notice that the nuclear positions appear as parameters in ψ_k and E_k . As the molecular conformation is varied by changing $\mathbf{R}_a, \mathbf{R}_b, \dots$, the ground state energy E_0 follows a curve called the molecular potential energy curve and the minimum of this curve defines the equilibrium conformation of the molecule.

10.1 The H₂⁺ ion

The simplest system that exhibits molecular properties is the hydrogen ion H₂⁺, which consists of two protons with positions $\mathbf{R}_a, \mathbf{R}_b$ and one electron at \mathbf{r} . With the potential energy

$$V(\mathbf{r}, \mathbf{R}_a, \mathbf{R}_b) = \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{R}_a - \mathbf{R}_b|} - \frac{1}{|\mathbf{r} - \mathbf{R}_a|} - \frac{1}{|\mathbf{r} - \mathbf{R}_b|} \right),$$

the Schrödinger equation takes the form,

$$\left[-\frac{\hbar^2 \nabla_r^2}{2m_e} + V(\mathbf{r}, \mathbf{R}_a, \mathbf{R}_b) \right] \psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b) = E\psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b).$$

Although this equation can be solved exactly using elliptical polar coordinates, it will be more instructive for our purposes to seek an approximate method of solution. Since there is no obvious parameter in which to develop a perturbative expansion, we will instead follow a variational route to explore the low energy states of system.

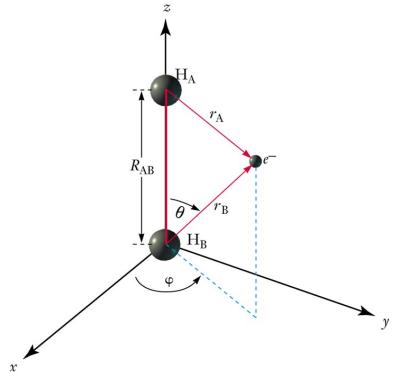
If the electron is close to one of the protons, one would expect the other proton to have a small influence on its dynamics, and that the wavefunction in this region would be close to that of a hydrogen atomic orbital. Therefore, in seeking the ground state of the H₂⁺ ion, we may take a trial wavefunction that is a linear combination of the ground state (1s) wavefunctions centred on the two protons,

$$\psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b) = \alpha\psi_a(\mathbf{r}; \mathbf{R}_a) + \beta\psi_b(\mathbf{r}; \mathbf{R}_b),$$

where $\psi_{a,b} = (\pi a_0^3)^{-1/2} \exp[-|\mathbf{r} - \mathbf{R}_{a,b}|/a_0]$, represents the corresponding hydrogenic wavefunction with a_0 the atomic Bohr radius. In this case, the coefficients α and β can be taken as real. The variational expression to be minimized in order to estimate the ground state energy is given by

$$\langle E \rangle = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\alpha^2 H_{aa} + \beta^2 H_{bb} + 2\alpha\beta H_{ab}}{\alpha^2 + \beta^2 + 2\alpha\beta S}$$

where $H_{aa} = \langle \psi_a | \hat{H} | \psi_a \rangle = H_{bb} = \langle \psi_b | \hat{H} | \psi_b \rangle$, and $H_{ab} = \langle \psi_a | \hat{H} | \psi_b \rangle = \langle \psi_b | \hat{H} | \psi_a \rangle$. Note that the matrix elements $H_{aa} = H_{bb}$ because \hat{H} is symmetric with respect to \mathbf{R}_a and \mathbf{R}_b . Moreover, since ψ_a and ψ_b are not orthogonal we have to introduce the **overlap integral**, $S = \langle \psi_a | \psi_b \rangle$, which measures the overlap



between the two atomic wavefunctions. In fact we can simplify this expression further, because the potential is symmetric about the mid-point between the two protons. The wavefunction must therefore be either symmetric or antisymmetric, $\alpha = \pm\beta$, and hence,

$$E_0 \leq \langle E \rangle = \frac{H_{aa} \pm H_{ab}}{1 \pm S}.$$

The matrix elements in this expression can be evaluated in closed form, though the calculation is rather tedious.¹

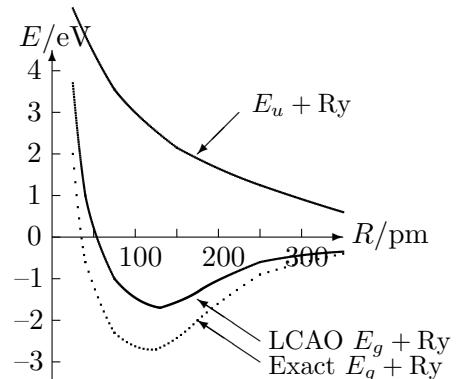
We have, therefore, found two possible wavefunctions for the H₂⁺ ion,

$$\psi_g = \frac{\psi_a + \psi_b}{\sqrt{2(1+S)}}, \quad \psi_u = \frac{\psi_a - \psi_b}{\sqrt{2(1-S)}},$$

with energies $E_g = \frac{H_{aa}+H_{ab}}{1+S}$, $E_u = \frac{H_{aa}-H_{ab}}{1-S}$. The subscript *g* refers to the term *gerade* (German for even) used in molecular physics to denote a state that is even under the operation of inverting the electronic wavefunction through the centre of symmetry of the molecule, without changing the positions of the nuclei. Such an inversion changes $\mathbf{r} \rightarrow \mathbf{R}_a + \mathbf{R}_b - \mathbf{r}$, which interchanges ψ_a and ψ_b . Note that this is not the same as parity inversion, which would also affect the nuclear coordinates. The *ungerade* (odd) state is denoted by subscript *u*. Note that ψ_g and ψ_u are orthogonal, even though ψ_a and ψ_b are not. In fact, ψ_g and ψ_u are just the orthonormal states that diagonalize the Hamiltonian, if we limit ourselves to linear combinations of ψ_a and ψ_b . In chemistry, ψ_g and ψ_u are called **molecular orbitals** and the assumption that they are linear combinations of atomic stationary states is called the **linear combination of atomic orbitals** (LCAO) approximation.

The state ψ_g has the lower energy, while ψ_u represents an excited state of the molecular ion. Physically, the reason for this is that, in the ψ_g state, the two atomic wavefunctions interfere constructively in the region between the protons, giving an enhanced electron density in the region where the electron is attracted strongly by both protons, which serves to screen the two protons from each other. Conversely, in ψ_u we have destructive interference in the region between the protons. If we plot E_g and E_u as functions of the nuclear separation $R = |\mathbf{R}_a - \mathbf{R}_b|$, the results are as shown in the figure right. For both curves, we have plotted $E + Ry$ since $-Ry$ is the ground state energy of the hydrogen atom. The curve of $E_g + Ry$ in the LCAO approximation has a minimum value of -1.8eV at $R = R_0 = 130\text{pm}$, which is the predicted equilibrium nuclear separation. The predicted energy required to dissociate the molecular ion into a hydrogen atom and a free proton is thus 1.8eV . The curve of $E_u + Ry$ does not have a minimum, suggesting that the odd wavefunction ψ_u does not correspond to a bound molecular state.

As the variational method provides an upper limit on the ground state energy, it is no surprise that the true molecular potential energy curve (shown dotted) lies below the LCAO one. The true values of the equilibrium separation and dissociation energy are 106pm and 2.8eV respectively. Clearly



Molecular potential energy curves for H₂⁺ ion.

¹We simply quote the results here:

$$S = \left(1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2}\right) e^{-R/a_0}, \quad H_{aa} = -Ry + \frac{e^2}{4\pi\epsilon_0 R} \left(1 + \frac{R}{a_0}\right) e^{-2R/a_0}$$

$$H_{ab} = \frac{e^2}{4\pi\epsilon_0 a_0} \left(1 + \frac{R}{a_0}\right) e^{-R/a_0} + S \left(-Ry + \frac{e^2}{4\pi\epsilon_0 R}\right),$$

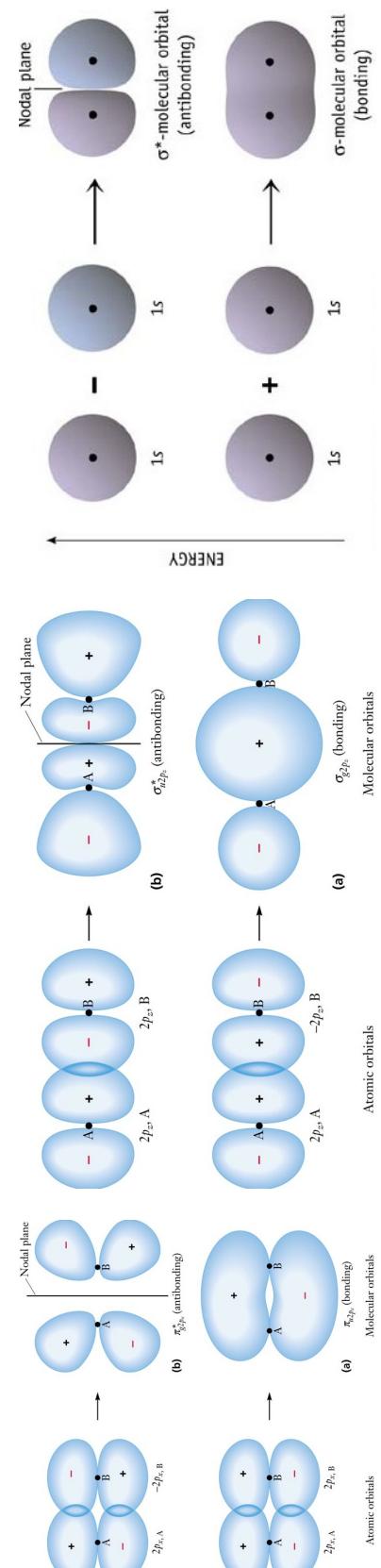
where $R = |\mathbf{R}_a - \mathbf{R}_b|$ and Ry is the Rydberg constant, the binding energy of the Hydrogen atom in its ground state.

the LCAO wavefunction ψ_g is not a very accurate approximation to the true ground state. We could improve it by introducing further variational parameters or additional atomic orbitals. For example, when R becomes very small the true wavefunction should approach that of a He⁺ ion; we could include such a term in the trial function. We could also include an **effective charge parameter** in ψ_a and ψ_b , which is equivalent to replacing the Bohr radius a_0 by a free parameter.

Although not very reliable quantitatively, the LCAO wavefunction ψ_g does however exhibit a number of important features of the true ground state: (i) it is even (g) with respect to inversion of the electron wavefunction; (ii) there is constructive interference which leads to an enhanced probability of finding the electron in the region between the two nuclei. These features will be important when we come to discuss **bonding**. Since the odd u -states are orthogonal to the even g -states, and the true ground state is a g -state, the curve of E_u actually represents an upper limit on the energy of the lowest u -state. Thus the fact that the curve has no minimum does not really prove that there are no bound u -states; but this does turn out to be the case. The LCAO wavefunction ψ_u shows the characteristic feature of an **anti-bonding** state: there is destructive interference in the region between the two nuclei, so that the electron is actually forced out of the region of overlap of the two atomic wavefunctions.

▷ INFO. At this stage, it is helpful to introduce some notation to **label the molecular orbitals**. Although the wavefunctions $\psi_{g,u}$ that we have been discussing are formed in the LCAO approximation from linear combinations of atomic $1s$ ($n = 1, \ell = 0$) states, with no orbital angular momentum, they do not themselves necessarily have zero orbital angular momentum. An $\ell = 0$ state must be proportional to Y_{00} , i.e., it must have no dependence on θ and ϕ , giving an isotropic probability distribution. But these states are certainly not isotropic: they have a ‘dumbbell’ shape, concentrated around the two protons. They do not have unique electronic orbital angular momentum because the operator $\hat{\mathbf{L}}^2$ for the electron does not commute with the Hamiltonian on account of the non-central terms $1/|\mathbf{r} - \mathbf{R}_a|$ and $1/|\mathbf{r} - \mathbf{R}_b|$ in the potential.

The only component of $\hat{\mathbf{L}}$ that does commute with these terms is \hat{L}_z , provided we choose the z -axis parallel to the internuclear axis $\mathbf{R}_a - \mathbf{R}_b$. Therefore instead of classifying the states as s, p, d, \dots orbitals according to whether $\ell = 0, 1, 2, \dots$, we call them $\sigma, \pi, \delta, \dots$ orbitals according to whether $\Lambda = 0, 1, 2, \dots$, where $\Lambda \equiv |m_\ell|$. A subscript u or g denotes whether the state is even or odd under inversion; this notation can be applied to all *homonuclear* diatomic molecules, in which the potential is symmetric about the median plane of the molecule. Thus the ground state of the hydrogen molecular ion is σ_g and the corresponding odd state is σ_u^* , where the star signifies an antibonding orbital. In the LCAO approximation used above, these molecular orbitals are linear combinations of $1s$ atomic orbitals and so they can be written as $1s\sigma_g$ and $1s\sigma_u^*$. To get some insight into how this notation applies, it is helpful to refer to the figures on the right.



10.2 The H₂ molecule

We now turn to consider the hydrogen molecule. One might imagine that this would be a trivial extension of the H₂⁺ ion, but in fact several new features arise when we consider this simple molecule. For two electrons with positions $\mathbf{r}_{1,2}$ and two protons at $\mathbf{R}_{a,b}$, in the Born-Oppenheimer approximation, the

| Approximation | B.E./eV | R_0/pm |
|-------------------------------|---------|-----------------|
| Experiment | 4.75 | 74.1 |
| MO fixed charge | 2.70 | 85.2 |
| MO variable charge | 3.49 | 73.0 |
| VB variable charge | 3.78 | 74.6 |
| Variable λ and charge | 4.03 | 75.7 |
| 13 variable parameters | 4.72 | 74.1 |

Table 10.1: Binding energy and equilibrium nuclear separation of the hydrogen molecule in various approximations.

Hamiltonian is given by

$$\hat{H} = -\frac{\hbar^2}{2m_e} (\nabla_{r_1}^2 + \nabla_{r_2}^2) + \frac{e^2}{4\pi\epsilon_0} \left[\frac{1}{r_{ab}} + \frac{1}{r_{12}} - \frac{1}{r_{1a}} - \frac{1}{r_{1b}} - \frac{1}{r_{2a}} - \frac{1}{r_{2b}} \right],$$

where $r_{1b} = |\mathbf{r}_1 - \mathbf{R}_b|$, etc. This is just the sum of two Hamiltonians for H₂⁺ ions, plus the additional term $\frac{e^2}{4\pi\epsilon_0}(\frac{1}{r_{12}} - \frac{1}{r_{ab}})$. It is plausible that the expectation values of $1/r_{12}$ and $1/r_{ab}$ will be comparable and therefore the extra term can be treated as a perturbation. Thus, as a first approximation we neglect it and assign each electron to one of the H₂⁺ molecular orbitals defined above. There are four ways of filling the two orbitals σ_g and σ_u^* , which we can represent by

$$\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2), \quad \psi_g(\mathbf{r}_1)\psi_{u^*}(\mathbf{r}_2), \quad \psi_{u^*}(\mathbf{r}_1)\psi_g(\mathbf{r}_2), \quad \psi_{u^*}(\mathbf{r}_1)\psi_{u^*}(\mathbf{r}_2).$$

Of these, we would expect the first to be the ground state.

However, at this stage, we have given no consideration to the constraints imposed by particle statistics. In fact, since the electrons are identical fermions, the wavefunction must be antisymmetric with respect to their interchange. Taking into account the spin degree of freedom, for both electrons to occupy the bonding σ_g molecular orbital, they must have opposite spins and occupy the singlet spin state, $X_{0,0} = \frac{1}{\sqrt{2}}(\chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2))$.

If we calculate the energy of the state $\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2)X_{0,0}$ as a function of the nuclear separation R , the minimum ground state energy occurs at $R_0 = 85\text{pm}$ and corresponds to a binding energy of 2.7eV. The true molecule is smaller and more tightly bound. Allowing for more variation in the atomic orbitals, in the form of a variable effective charge, gives an equilibrium value of R_0 much closer to experiment, but a binding energy that is still not high enough (see Table 10.1). The reason is that the σ_g^2 configuration alone is not a very good representation of the ground state. To understand why, consider the following.

If we multiply out the spatial part of the σ_g^2 wavefunction in the LCAO approximation, we see that it has a rather strange form,

$$\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2) \propto [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] + [\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_b(\mathbf{r}_2)].$$

The terms in the first square bracket correspond to the two electrons being shared between the two hydrogen atoms. This is the **covalent bonding** picture of the bound state. In the other square bracket, however, both electrons are assigned to the same atom, corresponding to **ionic bonding**. Since all the terms have equal coefficients, the ionic and covalent contributions are equal, which seems rather constraining, if not implausible. For example, it means that, when the two protons are pulled apart, the system is just as likely to be found to consist of an H⁺ and an H⁻ ion as two neutral atoms.

If we go to the pure **valence bonding** (VB) approximation and drop the ionic part of the wavefunction altogether, we find that the predicted binding energy and nuclear separation are both improved (see Table 10.1). Including a small parameter λ for the amplitude of the ionic component, i.e., taking

$$\psi^{VB} \propto [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] + \lambda[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_b(\mathbf{r}_2)],$$

we find that the variational method gives an optimal value of λ of about 1/6, meaning only about a 3% probability of finding the ionic configuration. Still, even with this refinement, the agreement with experiment is far from perfect. However, this doesn't mean quantum mechanics is failing; by taking enough free parameters in the trial function an excellent result can be obtained, as shown in the Table.

▷ INFO. As mentioned above, there are four possible ways of putting two electrons into the $1s\sigma_g$ and $1s\sigma_u^*$ molecular orbitals. From these four two-electron states we can make three states that are symmetric under interchange of the positions of the electrons, all of which need to be combined with the antisymmetric spin state $X_{0,0}$,

$$\begin{aligned} ^1\Sigma_g &: X_{0,0}\sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2) \\ ^1\Sigma_u &: X_{0,0}[\sigma_g(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2) + \sigma_u^*(\mathbf{r}_1)\sigma_g(\mathbf{r}_2)]/\sqrt{2} \\ ^1\Sigma_g &: X_{0,0}\sigma_u^*(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2). \end{aligned}$$

In addition we can make a triplet of states from an antisymmetric spatial wavefunction and the symmetric triplet spin states X_{1,m_S} ($m_S = 0, \pm 1$),

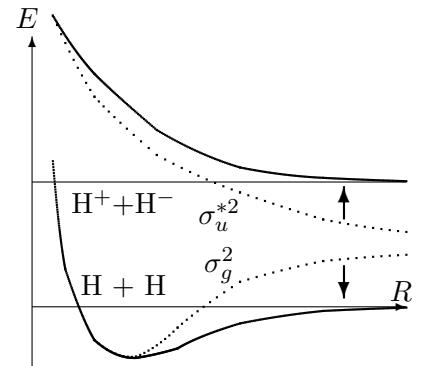
$$^3\Sigma_u : X_{1,m_S}[\sigma_g(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2) - \sigma_u^*(\mathbf{r}_1)\sigma_g(\mathbf{r}_2)]/\sqrt{2}.$$

We have introduced to the left of these equations a new notation, called the **molecular term**, to describe the overall quantum numbers of the molecule, all of which must be good quantum numbers because they correspond to operators which commute with the molecular Hamiltonian. It derives from a historic spectroscopic notation used in atomic physics. The term is written ${}^{2S+1}\Lambda_{u/g}$. The prefix $2S+1$ denotes the multiplicity of the total spin state of the electrons, hence 1 for a singlet and 3 for a triplet state. The central greek capital letter represents the magnitude of the total \hat{L}_z quantum number, $\Lambda = 0, 1, 2 \dots$ being represented by $\Sigma, \Pi, \Delta \dots$. In the case of the molecular orbitals above based on the $1s$ atomic orbitals, all the wavefunctions clearly have zero orbital angular momentum about the internuclear axis and hence they are all Σ states. The g or u suffix means even or odd under inversion, and is only meaningful for homonuclear molecules. Notice that since $(\pm 1)^2 = +1$ we get a g -state by combining two g or two u -states, and a u -state by combining a g and a u -state.

In terms of the molecular orbital approach, the VB approximation implies that the ground state is not simply the σ_g^2 configuration but rather a mixture of the two ${}^1\Sigma_g$ states,

$$\psi^{VB} \propto (1 + \lambda)(1 + S)\sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2) - (1 - \lambda)(1 - S)\sigma_u^*(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2).$$

Thus there is **configuration mixing** in the hydrogen molecule. The two states can mix because they have the same overall quantum numbers. At large nuclear separations the energy eigenstates are clearly separated into those that are almost a pair of neutral atoms and those consisting of an H^+ and an H^- ion. In the case of the u -states, we can see by expanding the spatial wave functions that ${}^3\Sigma_u$ is of the former type and ${}^1\Sigma_u$ of the latter. Since these configurations have different resultant electron spins they do not mix significantly (just like singlet and triplet states of atomic helium). Also the u - and g -states are prevented from mixing by their different inversion symmetry. However, the two ${}^1\Sigma_g$ configurations have the same electron spin and symmetry and can mix to give the above covalent-bonded ground state and an orthogonal excited state that is more ionic.



Molecular potential energy and configuration mixing in H_2 .

10.3 From molecules to solids

With these basic principles in hand, we could go on to discuss the orbital and electronic structure of more complicated molecules. In doing so, we would sink deeper into the realm of quantum chemistry. Instead, we will use these simple ideas of molecular bonding to develop a surprisingly versatile and faithful description of the ordered solid state. With *ca.* 10^{23} nuclei and electrons now involved such an enterprise seems foolhardy. However, we will see that, by exploiting symmetries, we can capture many of the basic principles of the solid state.

Let us then consider the electronic structure of an ordered crystalline array of equivalent atoms. (The consideration of more complicated periodic structures would not present conceptual challenges, but would bring unnecessary complications to the discussion.) Once again, we can draw upon the Born-Oppenheimer approximation and focus solely on the motion of electrons around an otherwise ordered array of nuclei. Even then, we are confronted with a many-particle Hamiltonian of apparently great complexity,

$$\hat{H} = - \sum_n \frac{\hbar^2 \nabla_n^2}{2m_e} - \sum_{n,N} \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_{nN}} + \sum_{m < n} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{mn}}.$$

Here the second term represents the interaction of the electrons with the constituent nuclei, and the third term involves the electron-electron interaction. In a physical system, we would have to take into account the influence of further relativistic corrections which would introduce additional spin-orbit couplings.

To address the properties of such a complex interacting system, we will have to draw upon many of the insights developed previously. To begin, it is helpful to partition the electrons into those which are bound to the core and those which are able to escape the potential of the individual atomic nuclei and propagate “freely” through the lattice. The electrons which are tightly bound to the nuclei screen the nuclear charge leading to a modified nuclear potential, $V_{\text{eff}}(\mathbf{r})$. Focussing on those electrons which are free, the effective Hamiltonian can be written as, $\hat{H} \simeq \sum_n \hat{H}_n + \sum_{m,n} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{mn}}$, where

$$\hat{H}_n = - \frac{\hbar^2 \nabla_n^2}{2m_e} + V_{\text{eff}}(\mathbf{r}_n)$$

represents the single-particle Hamiltonian experienced by each of the electrons – i.e. \hat{H}_n describes the motion of an electron moving in a periodic lattice potential, $V_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}(\mathbf{r} + \mathbf{R})$ with \mathbf{R} belonging to the set of periodic lattice vectors.

Despite engineering this approximation, we are still confronted by a challenging many-particle problem. Firstly, the problem remains coupled through the electron-electron Coulomb interaction. Secondly, the electrons move in a periodic potential. However, if we assume that the electrons remain mobile – the jargon is **itinerant** – and free to propagate through the lattice, they screen each other and diminish the effect of the Coulomb interaction.² Therefore, in these circumstances, we can proceed by neglecting the Coulomb interaction altogether. Of course, we still have to contend with the constraints placed by Pauli exclusion and wavefunction antisymmetry. But we are now in the realm of the molecular orbital theory, and can proceed analogously using the variational LCAO approach.

²This screening is effective for $\bar{a} \ll a_0$, in the notation of §8.4.2.

In particular, we can build a trial wavefunction by combining orbital states of the single ion, $V_{\text{ion}}(\mathbf{r})$, where $V_{\text{eff}}(\mathbf{r}) = \sum_{\mathbf{R}} V_{\text{ion}}(\mathbf{r} - \mathbf{R})$. As with the hydrogen molecule, the Hamiltonian for the individual nuclei, $\hat{H}_0 = -\frac{\hbar^2 \nabla^2}{2m_e} + V_{\text{ion}}(\mathbf{r})$ are associated with a set of atomic orbitals, ψ_q , characterized by a set of quantum numbers, q . In the **atomic limit**, when the atoms are far-separated, these will mirror the simple hydrogenic states. To find the variational ground state of the system, we can then build a trial state from a linear combination of these atomic orbitals. Taking only the lowest orbital, $q = 0$, into account we have,

$$\Psi(\mathbf{r}) = \sum_{\mathbf{R}} \alpha_{\mathbf{R}} \psi(\mathbf{r} - \mathbf{R}),$$

where, as before, $\alpha_{\mathbf{R}}$ represent the set of variational coefficients, one for each site (and, in principle, each atomic orbital if we had taken more than one!).

Once again, we can construct the variational state energy,

$$E = \frac{\langle \Psi | \hat{H}_{1e} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}}^* H_{\mathbf{R}\mathbf{R}'} \alpha_{\mathbf{R}'}}{\sum_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}}^* S_{\mathbf{R}\mathbf{R}'} \alpha_{\mathbf{R}'}}$$

where, as before, $H_{\mathbf{R}\mathbf{R}'} = \int d^d r \psi^*(\mathbf{r} - \mathbf{R}) \hat{H}_{1e} \psi(\mathbf{r} - \mathbf{R}')$ denote the matrix elements of the orbital wavefunction on the Hamiltonian $\hat{H}_{1e} = -\frac{\hbar^2 \nabla^2}{2m_e} + V_{\text{eff}}(\mathbf{r})$ and $S_{\mathbf{R}\mathbf{R}'} = \int d^d r \psi^*(\mathbf{r} - \mathbf{R}) \psi(\mathbf{r} - \mathbf{R}')$ represent the overlap integrals. Then, varying the energy with respect to $\alpha_{\mathbf{R}}^*$, we find that the coefficients obey the **secular equation** (exercise)

$$\sum_{\mathbf{R}'} [H_{\mathbf{R}\mathbf{R}'} - E S_{\mathbf{R}\mathbf{R}'}] \alpha_{\mathbf{R}'} = 0.$$

Note that, if the basis functions were orthogonal, this would just be an eigenvalue equation.

▷ EXERCISE. To develop this idea, let us first see how the method relates back to the problem of H_2^+ : In this case, the secular equation translates to the 2×2 matrix equation,

$$\begin{pmatrix} H_{aa} - E & H_{ab} - ES \\ H_{ab} - ES & H_{aa} - E \end{pmatrix} \boldsymbol{\alpha} = 0,$$

where the notation and symmetries of matrix elements follow from section 10.1. As a result, we find that the states $\boldsymbol{\alpha}$ divide into even and odd states as expected.

Now let us consider a one-dimensional periodic lattice system. If we assume that the atoms are well-separated, it is evident that both the overlap integrals and matrix elements will decay exponentially fast with orbital separation. The dominant contribution to the energy will then derive from matrix elements coupling only neighbouring states. In this case, the secular equation separates into the coupled sequence of equations,

$$(\varepsilon - E) \alpha_n - (t + ES)(\alpha_{n+1} + \alpha_{n-1}) = 0,$$

for each n , where $H_{nn} = \varepsilon$ denotes the atomic orbital energy, $H_{n,n+1} = H_{n+1,n} = -t$ is the matrix element between neighbouring states, $S_{n,n} = 1$ and $S_{n,n+1} = S_{n+1,n} = S$. All other matrix elements are exponentially small. Here we consider a system with N lattice sites and impose the periodic boundary condition, $\alpha_{n+N} = \alpha_n$. The periodicity of the secular equation suggests a

solution of the form $\alpha_n = \frac{1}{\sqrt{N}} e^{ikna}$, where $k = 2\pi m/Na$ denote a discrete set of N wavevectors with m integer lying in the range $-N/2 < m \leq N/2$. (The size of this range, $k_{\max} - k_{\min}$, is fixed by the magnitude of the **reciprocal lattice vector** for the one-dimensional lattice, $2\pi/a$.) Substitution confirms that this is a solution with energies (exercise),

$$E_k = \frac{\varepsilon - 2t \cos(ka)}{1 + 2S \cos(ka)}.$$

The wavevector k parameterizes a **band** of states.

Then, according to the LCAO approximation, for a single electron, distributed throughout the system, the lowest energy state is predicted to be the uniform state $\alpha_n = \frac{1}{\sqrt{N}}$ with the energy $E_0 = \frac{\varepsilon - 2t}{1 + 2S}$. (Here we have assumed that the matrix element, t , is positive – in the atomic limit, consider why this is a sound assumption.) However, if we suppose that each atom contributes a non-zero fraction of electrons to the system, we must consider the influence of Pauli exclusion and particle statistics.

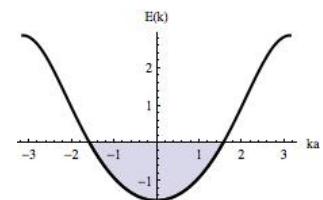
Since the electrons are fermions, each state k can host two electrons in a total spin singlet configuration. The lowest energy state is then obtained by adding electron pairs sequentially into states of increasing k until all electrons are accounted for. If the maximum k value, known as the Fermi wavevector, k_F , lies within the band of states, elementary excitations of the electrons cost vanishingly small energy leading to **metallic** behaviour. On the other hand, if the system is stoichiometric, with each atom contributing an integer number of electrons, the Fermi wavevector may lie at a **band gap** between two different band of states. In this case, the system has a gap to excitations and the material forms a **band insulator**.

▷ EXERCISE. Consider how the calculation above can be extended to a two-dimensional square lattice system. What would be the corresponding band dispersion?

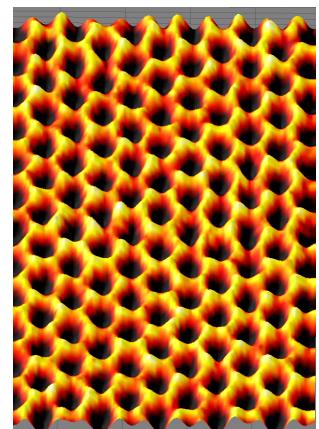
▷ EXERCISE. In the one-dimensional system, extend the calculation to compute the electron band structure for a periodic lattice with two atomic orbitals on each site, one with s -wave symmetry, and another with p -wave. In particular, consider how the dispersion would change with the orientation of the p state relative to the chain direction.

▷ INFO. To add flesh to these ideas, let us then consider a simple, but prominent problem from the realm of quantum condensed matter physics. In recent years, there has been great interest in the properties of **graphene**, a single layer of graphite. (These developments were recognized by the award of the 2010 Nobel prize in physics to Geim and Novoselov.) Remarkably, high quality single crystals of graphene can be obtained by running graphite – a pencil! – over an adhesive layer. The resulting electron states of the single layer compound have been of enormous interest to physicists. To understand why, let us implement the LCAO technology to explore the valence electron structure of graphene.

Graphene forms a periodic two-dimensional honeycomb lattice structure with two atoms per each unit cell. With an electron configuration $1s^2 2s^2 2p^2$, the two $1s$ electrons are bound tightly to the nucleus. The remaining $2s$ electrons **hybridize** with one of the p orbitals to form three sp^2 hybridized orbitals. These three orbitals form the basis of a strong covalent bond of σ orbitals that constitute the honeycomb lattice. The remaining electron, which occupies the out-of-plane p_z orbital, is then capable of forming an itinerant band of electron states. It is this band which we now address.



One dimensional band structure for E_k for $S = 0.3$, $t = 1$ and $\varepsilon = 0$. Here we have shown schematically a band-filling with $k_F a = \pi/2$.



Atomically resolved STM image of graphene sheet at 77K.

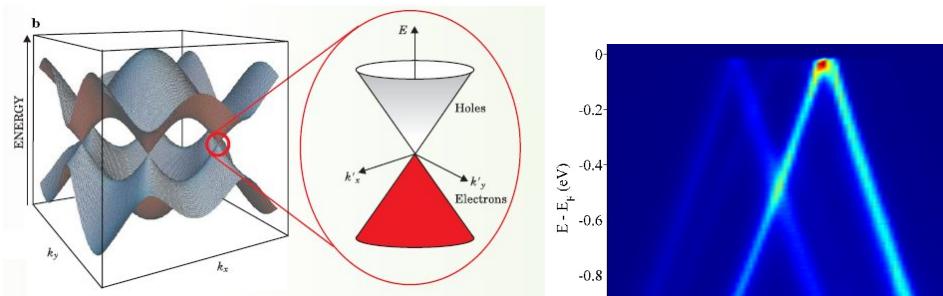


Figure 10.1: Left: dispersion relation of graphene $E_{\mathbf{k}}$ obtained with the LCAO approximation. Notice that near the centre of the band, the dispersion becomes point-like. Around these points, the dispersion takes form of a linear (Dirac) cone. Right: band structure of a sample of multilayer epitaxial graphene. Linear bands emerge from the K-points in the band structure. Three “cones” can be seen from three layers of graphene in the MEG sample.

Once again, let us suppose that the wavefunction of this band involves the basis of single p_z orbital states localized to each lattice site,

$$\psi(\mathbf{r}) = \sum_{\mathbf{R}} [\alpha_{\mathbf{R}} \psi_1(\mathbf{r} - \mathbf{R}) + \beta_{\mathbf{R}} \psi_2(\mathbf{r} - \mathbf{R})].$$

Here the p_z orbital wavefunction ψ_1 is centred on one of the atoms in the unit cell, and ψ_2 is centred on the other. Once again, taking into account matrix elements involving only nearest neighbours, the trial wavefunction translates to the secular equation,

$$\begin{aligned} (\varepsilon - E)\alpha_{\mathbf{R}} - (t + ES)(\beta_{\mathbf{R}} + \beta_{\mathbf{R}-\mathbf{a}_1} + \beta_{\mathbf{R}-\mathbf{a}_2}) &= 0 \\ (\varepsilon - E)\beta_{\mathbf{R}} - (t + ES)(\alpha_{\mathbf{R}} + \alpha_{\mathbf{R}+\mathbf{a}_1} + \alpha_{\mathbf{R}+\mathbf{a}_2}) &= 0, \end{aligned}$$

where the lattice vectors $\mathbf{a}_1 = (\sqrt{3}/2, 1/2)a$ and $\mathbf{a}_2 = (\sqrt{3}/2, -1/2)a$, with a the lattice spacing, are shown in the figure to the right. Note that the off-diagonal matrix elements involve only couplings between atoms on different sublattices. Once again, we can make the ansatz that the solutions are of the form of plane waves with $\alpha_{\mathbf{R}} = \frac{\alpha_{\mathbf{k}}}{\sqrt{N}} e^{i\mathbf{k}\cdot\mathbf{R}}$ and $\beta_{\mathbf{R}} = \frac{\beta_{\mathbf{k}}}{\sqrt{N}} e^{i\mathbf{k}\cdot\mathbf{R}}$. Notice that, in this case, we must allow for different relative weights, $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$. Substituting, we find that this ansatz is consistent if

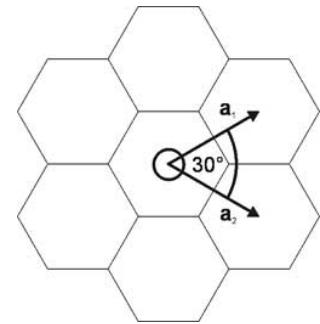
$$\begin{aligned} (\varepsilon - E)\alpha_{\mathbf{k}} - (t + ES)f_{\mathbf{k}}\beta_{\mathbf{k}} &= 0 \\ (\varepsilon - E)\beta_{\mathbf{k}} - (t + ES)f_{\mathbf{k}}^*\alpha_{\mathbf{k}} &= 0 \end{aligned}$$

where $f_{\mathbf{k}} = 1 + 2e^{-i\sqrt{3}k_x a/2} \cos(k_y a/2)$. Although this equation can be solved straightforwardly, it takes a particularly simple form when the overlap integral between neighbouring sites, S , is neglected. In this case, one obtains,

$$E_{\mathbf{k}} = \varepsilon \pm |f_{\mathbf{k}}|t.$$

The corresponding band structure is shown right as a function of \mathbf{k} . In particular, one may note that, at the band centre, the dispersion relation for the electrons becomes point-like (see Fig. 10.3).

In the half-filled system, where each carbon atom contributes a single electron to the band, the Fermi level lies precisely at the centre of the band where the dispersion, $E_{\mathbf{k}}$ is point like. Doping electrons into (or removing electrons from) the system results in (two copies) of a linear dispersion, $E_{\mathbf{k}} \simeq c|\mathbf{k}|$, where c is a constant (velocity). Such a linear dispersion relation is the hallmark of a relativistic particle (cf. a photon). Of course, in this case, the electrons are not moving at relativistic velocities. But their properties will mirror the behaviour of relativistic particles. And herein lies the interest that these materials have drawn.



Finally, let us comment on the influence of the electron-electron interaction effects that were neglected in our treatment above. In principle, we could adopt a Hartree or Hartree-Fock scheme to address the effects of electron interaction in a perturbative manner. Indeed, such a programme would lead to interaction corrections which would modify the electronic band structure derived above. It is a surprising yet robust feature of Fermi systems that the properties of the non-interacting ground state wavefunction remain qualitatively correct over an unreasonably wide range of interaction strengths.³ This *rigidity* can be ascribed to the constraints on the nodal structure of the wavefunction imposed by particle statistics. However, in some cases, the manifestations of electron interactions translate to striking modifications in the observable properties. We have already discussed the possibility of electron-electron interactions leading to Wigner crystallization or to itinerant ferromagnetism; other effects include interaction-driven electron localization (the “Mott transition”), local moment formation, quantum Hall fluids, and superconductivity. Such phases, which by their nature lie outside any perturbative scheme built around the non-interacting ground state, form the field of quantum condensed matter and solid state physics.

³This feature is embodied in the theory of the electron liquid known as **Landau’s Fermi liquid theory**.

Chapter 11

Time-dependent perturbation theory

So far, we have focused largely on the quantum mechanics of systems in which the Hamiltonian is time-independent. In such cases, the time dependence of a wavepacket can be developed through the time-evolution operator, $\hat{U} = e^{-i\hat{H}t/\hbar}$, or, when cast in terms of the eigenstates of the Hamiltonian, $\hat{H}|n\rangle = E_n|n\rangle$, as $|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle = \sum_n e^{-iE_nt/\hbar}c_n(0)|n\rangle$. Although this framework provides access to any closed quantum mechanical system, it does not describe interaction with an external environment such as that imposed by an external electromagnetic field. In such cases, it is more convenient to describe the induced interactions of a small isolated system, \hat{H}_0 , through a time-dependent interaction $V(t)$. Examples include the problem of magnetic resonance describing the interaction of a quantum mechanical spin with an external time-dependent magnetic field, or the response of an atom to an external electromagnetic field. In the following, we will develop a formalism to treat time-dependent perturbations.

11.1 Time-dependent potentials: general formalism

Consider then the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}(t)$, where all time-dependence enters through the potential $\hat{V}(t)$. In the **Schrödinger representation**, the dynamics of the system are specified by the time-dependent wavefunction, $|\psi(t)\rangle_S$ through the Schrödinger equation $i\hbar\partial_t|\psi(t)\rangle_S = \hat{H}|\psi(t)\rangle_S$. However, in many cases, and in particular with the current application, it is convenient to work in the **Interaction representation**,¹ defined by

$$|\psi(t)\rangle_I = e^{i\hat{H}_0 t/\hbar}|\psi(t)\rangle_S,$$

where $|\psi(0)\rangle_I = |\psi(0)\rangle_S$. With this definition, one may show that the wavefunction obeys the equation of motion (exercise)

$$i\hbar\partial_t|\psi(t)\rangle_I = \hat{V}_I(t)|\psi(t)\rangle_I, \quad (11.1)$$

where $\hat{V}_I(t) = e^{i\hat{H}_0 t/\hbar}\hat{V}e^{-i\hat{H}_0 t/\hbar}$. Then, if we form the eigenfunction expansion, $|\psi(t)\rangle_I = \sum_n c_n(t)|n\rangle$, and contract the equation of motion with a general state, $\langle n|$, we obtain

$$i\hbar\dot{c}_m(t) = \sum_n V_{mn}(t)e^{i\omega_{mn}t}c_n(t), \quad (11.2)$$

¹Note how this definition differs from that of the **Heisenberg representation**, $|\psi\rangle_H = e^{i\hat{H}t/\hbar}|\psi(t)\rangle_S$ in which all time-dependence is transferred into the operators.

where the matrix elements $V_{mn}(t) = \langle m | \hat{V}(t) | n \rangle$, and $\omega_{mn} = (E_m - E_n)/\hbar = -\omega_{nm}$. To develop some intuition for the action of a time-dependent potential, it is useful to consider first a periodically-driven two-level system where the dynamical equations can be solved exactly.

▷ INFO. The **two-level system** plays a special place in the modern development of quantum theory. In particular, it provides a platform to encode the simplest quantum logic gate, the **qubit**. A classical computer has a memory made up of bits, where each bit holds either a one or a zero. A quantum computer maintains a sequence of qubits. A single qubit can hold a one, a zero, or, crucially, any quantum superposition of these. Moreover, a pair of qubits can be in any quantum superposition of four states, and three qubits in any superposition of eight. In general a quantum computer with n qubits can be in an arbitrary superposition of up to 2^n different states simultaneously (this compares to a normal computer that can only be in one of these 2^n states at any one time). A quantum computer operates by manipulating those qubits with a fixed sequence of quantum logic gates. The sequence of gates to be applied is called a quantum algorithm.

An example of an implementation of qubits for a quantum computer could start with the use of particles with two spin states: $|\downarrow\rangle$ and $|\uparrow\rangle$, or $|0\rangle$ and $|1\rangle$). In fact any system possessing an observable quantity A which is conserved under time evolution and such that A has at least two discrete and sufficiently spaced consecutive eigenvalues is a suitable candidate for implementing a qubit. This is true because any such system can be mapped onto an effective spin-1/2 system.

11.1.1 Dynamics of a driven two-level system

Let us consider a two-state system with

$$\hat{H}_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad \hat{V}(t) = \begin{pmatrix} 0 & \delta e^{i\omega t} \\ \delta e^{-i\omega t} & 0 \end{pmatrix}.$$

Specifying the wavefunction by the two-component vector, $\mathbf{c}(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}$,

Eq. (11.2) translates to the equation of motion (exercise)

$$i\hbar\partial_t \mathbf{c} = \delta \begin{pmatrix} 0 & e^{i(\omega-\omega_{21})t} \\ e^{-i(\omega-\omega_{21})t} & 0 \end{pmatrix} \mathbf{c}(t), \quad (11.3)$$

where $\omega_{21} = (E_2 - E_1)/\hbar$. With the initial condition $c_1(0) = 1$, and $c_2(0) = 0$, this equation has the solution,

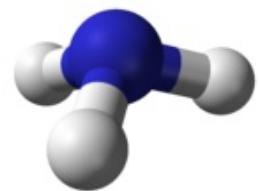
$$|c_2(t)|^2 = \frac{\delta^2}{\delta^2 + \hbar^2(\omega - \omega_{21})^2/4} \sin^2 \Omega t, \quad |c_1(t)|^2 = 1 - |c_2(t)|^2,$$

where $\Omega = ((\delta/\hbar)^2 + (\omega - \omega_{21})^2/4)^{1/2}$ is known as the **Rabi frequency**. The solution, which varies periodically in time, describes the transfer of probability from state 1 to state 2 and back. The maximum probability of occupying state 2 is a Lorentzian function of frequency ω , with

$$|c_2(t)|_{\max}^2 = \frac{\delta^2}{\delta^2 + \hbar^2(\omega - \omega_{21})^2/4},$$

taking the value of unity at resonance, $\omega = \omega_{21}$.

▷ EXERCISE. Derive the solution from the equations of motion for $\mathbf{c}(t)$. Hint: eliminate c_1 from the equations to obtain a second order differential equation for c_2 .



▷ INFO. The dynamics of the driven two-level system finds practical application in the **Ammonia maser**: The ammonia molecule NH₃ has a pyramidal structure with an orientation characterised by the position of the “lone-pair” of electrons sited on the nitrogen atom. At low temperature, the molecule can occupy two possible states, |A⟩ and |S⟩, involving symmetric (S) or antisymmetric (A) atomic configurations, separated by a small energy splitting, ΔE . (More precisely, along the axis of three-fold rotational symmetry, the effective potential energy of the nitrogen atom takes the form of a double-well. The tunneling of the nitrogen atom through the double well leads to the symmetric and antisymmetric combination of states.) In a time-dependent uniform electric field the molecules experience a potential $V = -\boldsymbol{\mu}_d \cdot \mathbf{E}$, where $\mathbf{E} = E\hat{\mathbf{e}}_z \cos \omega t$, and $\boldsymbol{\mu}_d$ denotes the electric dipole moment. Since $\boldsymbol{\mu}_d$ is odd under parity transformation, $P\boldsymbol{\mu}_d P = -\boldsymbol{\mu}_d$, and as $P|A\rangle = -|A\rangle$ and $P|S\rangle = |S\rangle$, the matrix elements of the electric dipole moment are off-diagonal: $\langle S|\boldsymbol{\mu}_d|S\rangle = \langle A|\boldsymbol{\mu}_d|A\rangle = 0$ and $\langle S|\boldsymbol{\mu}_d|A\rangle = \langle S|\boldsymbol{\mu}_d|A\rangle \neq 0$.

If we start with all of the molecules in the symmetric ground state, we have shown above that the action of an oscillating field for a particular time can drive a collection of molecules from their ground state into the antisymmetric first excited state. The ammonia maser works by sending a stream of ammonia molecules, traveling at known velocity, down a tube having an oscillating field for a definite length, so the molecules emerging at the other end are all (or almost all, depending on the precision of ingoing velocity, etc.) in the first excited state. Application of a small amount of electromagnetic radiation of the same frequency to the outgoing molecules will cause some to decay, generating intense radiation and therefore a much shorter period for all to decay, emitting coherent radiation.

Charles Hard Townes 1915-(left)

is

an American Nobel prize-winning physicist and educator. Townes is known for his work on the theory and application of the maser – microwave amplification by stimulated emission of radiation, on which he got the fundamental patent, and other work in quantum electronics connected with both maser and laser devices. He received the Nobel Prize in Physics in 1964.



11.1.2 Paramagnetic resonance

An important example of a dynamical two-level system is provided by a spin-1/2 nucleus under the influence of external magnetic fields. In § 5.3 we analysed the effect of a uniform magnetic field, and showed that this leads to spin precession. We found that the spin precession frequency is independent of the angle of the spin with respect to the field direction. Consider then how this looks in a frame of reference which is itself rotating with angular velocity ω about the z -axis. Let us specify the magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, since we'll soon be adding another component. In the rotating frame, the observed precession frequency is $\omega_r = -\gamma(\mathbf{B}_0 + \boldsymbol{\omega}/\gamma)$, so there is a different effective field $(\mathbf{B}_0 + \boldsymbol{\omega}/\gamma)$ in the rotating frame. If the frame rotates exactly at the precession frequency, $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = -\gamma\mathbf{B}_0$, spins pointing in any direction will remain at rest in that frame – there is no effective field at all.

Suppose we now add a small rotating magnetic field with angular frequency ω in the xy plane, so the total magnetic field is

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_1(\hat{\mathbf{e}}_x \cos(\omega t) - \hat{\mathbf{e}}_y \sin(\omega t)).$$

The effective magnetic field in the frame rotating with the same frequency ω as the small added field is then given by

$$\mathbf{B}_r = (B_0 + \omega/\gamma)\hat{\mathbf{z}} + B_1\hat{\mathbf{e}}_x.$$

Now, if we tune the angular frequency of the small rotating field so that it exactly matches the precession frequency in the original static magnetic field, $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = -\gamma\mathbf{B}_0$, all the magnetic moment will see in the rotating frame is the small field in the x -direction! It will therefore precess about the x -direction at the slow angular speed γB_1 . This matching of the small field rotation frequency with the large field spin precession frequency is the “resonance”.

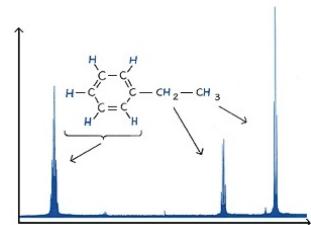
If the spins are lined up preferentially in the z -direction by the static field, and the small resonant oscillating field is switched on for a time such that $\gamma B_1 t = \pi/2$, the spins will be preferentially in the y -direction in the rotating frame, so in the lab they will be rotating in the xy plane, and a coil will pick up an a.c. signal from the induced e.m.f.

► INFO. **Nuclear magnetic resonance** is an important tool in chemical analysis. As the name implies, the method uses the spin magnetic moments of nuclei (particularly hydrogen) and resonant excitation. **Magnetic resonance imaging** uses the same basic principle to get an image (of the inside of a body for example). In basic NMR, a strong static B field is applied. A spin 1/2 proton in a hydrogen nucleus then has two energy eigenstates. After some time, most of the protons fall into the lower of the two states. We now use an electromagnetic wave (RF pulse) to excite some of the protons back into the higher energy state. The proton's magnetic moment interacts with the oscillating B field of the EM wave through the Hamiltonian

$$\hat{H} = -\boldsymbol{\mu} \cdot \mathbf{B} = \frac{g_p e}{2m_p c} \mathbf{S} \cdot \mathbf{B} = \frac{g_p e \hbar}{4m_p c} \boldsymbol{\sigma} \cdot \mathbf{B} = \frac{g_p}{2} \mu_N \boldsymbol{\sigma} \cdot \mathbf{B},$$

where the gyromagnetic ratio of the proton is about +5.6. The magnetic moment is $2.79\mu_N$ (nuclear magnetons). Different nuclei will have different gyromagnetic ratios giving more degrees of freedom with which to work. The strong static B field is chosen to lie in the z direction and an oscillating B -field (EM wave) is applied in the xy plane. For an oscillating field that is *circularly* polarized, with $\mathbf{B}(t) = B_+(\cos \omega t, \sin \omega t, 0)$, the dynamics of the two level system are identical to (11.3), with $\omega_1 - \omega_2 = g_p \mu_N B_z / (\hbar)$ (this is the Larmor frequency ω_0) and $\delta = g_p \mu_N B_+ / (2\hbar)$. A *linearly* polarized wave, with $\mathbf{B} = B_x(\cos \omega t, 0, 0)$, may be decomposed as two circularly polarized waves with $B_{\pm} = B_x/2$ (\pm refer to the handedness of the circular polarizations). Provided $\delta \ll \omega_0$, only the circular polarization that follows the precessing magnetic moment has any significant effect (the other circular polarization provides a force that oscillates as $e^{\pm i(\omega+\omega_0)t}$ and averages to zero). In the **rotating wave approximation**, we can again model the system according to (11.3), with $\omega_1 - \omega_2 = g_p \mu_N B_z / (\hbar) = \omega_0$ and $\delta = g_p \mu_N B_x / (4\hbar)$. Note that the resonance condition $\omega = \omega_0$ is exactly the condition that ensures that the energy of a photon of the EM field $E = \hbar\omega$ is equal to the energy difference between the two spin states $\Delta E = \hbar\omega_0$. The conservation of energy must be satisfied well enough to get a significant transition rate. In NMR, we observe the transitions back to the lower energy state. These emit EM radiation at the same frequency and we can detect it after the stronger input pulse ends (or by more complex methods).

NMR is a powerful tool in chemical analysis because the molecular field adds to the external B field so that the resonant frequency depends on the molecule as well as the nucleus. We can learn about molecular fields or just use NMR to see what molecules are present in a sample. In MRI, we typically concentrate on the one nucleus like hydrogen. We can put a gradient in B_z so that only a thin slice of the material has ω tuned to the resonant frequency. Therefore we can excite transitions to the higher energy state in only a slice of the sample. If we vary (in the orthogonal direction!) the B field during the decay, we can recover 3d images.



A proton NMR spectrum of a solution containing a simple organic compound, ethyl benzene. Each group of signals corresponds to protons in a different part of the molecule.



High resolution MRI scan of a brain!

11.2 Time-dependent perturbation theory

We now turn to consider a generic time-dependent Hamiltonian for which an analytical solution is unavailable – sadly the typical situation! In this case, we must turn to a perturbative analysis, looking for an expansion of the basis coefficients $c_n(t)$ in powers of the interaction,

$$c_n(t) = c_n^{(0)} + c_n^{(1)}(t) + c_n^{(2)}(t) + \dots,$$

where $c_n^{(m)} \sim O(V^m)$ and $c_n^{(0)}$ is some (time-independent) initial state. The programme to complete this series expansion is straightforward but technical.

▷ INFO. In the interaction representation, the state $|\psi(t)\rangle_I$ can be related to an initial state $|\psi(t_0)\rangle_I$ through the time-evolution operator, $U_I(t, t_0)$, i.e. $|\psi(t)\rangle_I = U_I(t, t_0)|\psi(t_0)\rangle_I$. Since this is true for any initial state $|\psi(t_0)\rangle_I$, from Eq. (11.1), we must have

$$i\hbar\partial_t U_I(t, t_0) = V_I(t)U_I(t, t_0),$$

with the boundary condition $U_I(t_0, t_0) = \hat{I}$. Integrating this equation from t_0 to t , formally we obtain,

$$U_I(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t') U_I(t', t_0).$$

This result provides a *self-consistent* equation for $U_I(t, t_0)$, i.e. if we take this expression and substitute $U_I(t', t_0)$ under the integrand, we obtain

$$U_I(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t') + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^t dt' V_I(t') \int_{t_0}^{t'} dt'' V_I(t'') U_I(t'', t_0).$$

Iterating this procedure, we thus obtain

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n V_I(t_1) V_I(t_2) \cdots V_I(t_n), \quad (11.4)$$

where the term $n = 0$ translates to \hat{I} . Note that the operators $V_I(t)$ are organised in a time-ordered sequence, with $t_0 \leq t_n \leq t_{n-1} \leq \cdots \leq t_1 \leq t$. With this understanding, we can write this expression more compactly as

$$U_I(t, t_0) = T \left[e^{-\frac{i}{\hbar} \int_{t_0}^t dt' V_I(t')} \right],$$

where “T” denotes the time-ordering operator and its action is understood by Eq. (11.4).

If a system is prepared in an initial state, $|i\rangle$ at time t_0 , at a subsequent time, t , the system will be in a final state,

$$|i, t_0, t\rangle = U_I(t, t_0)|i\rangle = \sum_n |n\rangle \overbrace{\langle n|U_I(t, t_0)|i\rangle}^{c_n(t)}.$$

Making use of Eq. (11.4), and the resolution of identity, $\sum_m |m\rangle\langle m| = \hat{I}$, we obtain

$$c_n(t) = \underbrace{c_n^{(0)}}_{\delta_{ni}} - \underbrace{\frac{i}{\hbar} \int_{t_0}^t dt' \langle n|V_I(t')|i\rangle}_{c_n^{(1)}} - \underbrace{\frac{1}{\hbar^2} \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \sum_m \langle n|V_I(t')|m\rangle \langle m|V_I(t'')|i\rangle}_{c_n^{(2)}} + \dots$$

Recalling that $V_I = e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 t/\hbar}$, we thus find that

$$\begin{aligned} c_n^{(1)}(t) &= -\frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{ni} t'} V_{ni}(t') \\ c_n^{(2)}(t) &= -\frac{1}{\hbar^2} \sum_m \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' e^{i\omega_{nm} t' + i\omega_{mi} t''} V_{nm}(t') V_{mi}(t''), \end{aligned} \quad (11.5)$$

where $V_{nm}(t) = \langle n|V(t)|m\rangle$ and $\omega_{nm} = (E_n - E_m)/\hbar$, etc. In particular, the probability of effecting a transition from state $|i\rangle$ to state $|n\rangle$ for $n \neq i$ is given by $P_{i \rightarrow n} = |c_n(t)|^2 = |c_n^{(1)}(t) + c_n^{(2)}(t) + \dots|^2$.

▷ EXAMPLE: **The kicked oscillator:** Suppose a simple harmonic oscillator is prepared in its ground state $|0\rangle$ at time $t = -\infty$. If it is perturbed by a small time-dependent potential $V(t) = -eEx e^{-t^2/\tau^2}$, what is the probability of finding it in the first excited state, $|1\rangle$, at $t = +\infty$?

Working to first order of perturbation theory, the probability is given by $P_{0 \rightarrow 1} \simeq |c_1^{(1)}|^2$ where $c_1^{(1)}(t) = -\frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{10}t'} V_{10}(t')$, with $V_{10}(t') = -eE \langle 1|x|0\rangle e^{-t'^2/\tau^2}$ and $\omega_{10} = \omega$. Using the ladder operator formalism, with $|1\rangle = a^\dagger|0\rangle$ and $x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)$, we have $\langle 1|x|0\rangle = \sqrt{\frac{\hbar}{2m\omega}}$. Therefore, making use of the identity $\int_{-\infty}^{\infty} dt' \exp[i\omega t' - t'^2/\tau^2] = \sqrt{\pi}\tau \exp[-\omega^2\tau^2/4]$, we obtain the transition amplitude, $c_1^{(1)}(t \rightarrow \infty) = ieE\tau \sqrt{\frac{\pi}{2m\hbar\omega}} e^{-\omega^2\tau^2/4}$. As a result, we obtain the transition probability, $P_{0 \rightarrow 1} \simeq (eE\tau)^2 (\pi/2m\hbar\omega) e^{-\omega^2\tau^2/2}$. Note that the probability is maximized for $\tau \sim 1/\omega$.

▷ EXERCISE. Considering the same perturbation, calculate the corresponding transition probability from the ground state to the *second* excited state. Hint: note that this calculation demands consideration of the second order of perturbation theory.

11.3 “Sudden” perturbation

To further explore the time-dependent perturbation theory, we turn now to consider the action of fast or “sudden” perturbations. Here we define sudden as a perturbation in which the switch from one time-independent Hamiltonian \hat{H}_0 to another \hat{H}'_0 takes place over a time much shorter than any natural period of the system. If the system is initially in an eigenstate $|n\rangle$ of \hat{H}_0 , its time evolution following the switch will follow that of \hat{H}'_0 , i.e. one simply has to expand the initial state as a sum over the eigenstates of \hat{H}'_0 , $|n\rangle = \sum_{n'} |n'\rangle \langle n'|n\rangle$. The non-trivial part of the problem lies in establishing that the change is sudden enough. This is achieved by estimating the actual time taken for the Hamiltonian to change, and the periods of motion associated with the state $|n\rangle$ and with its transitions to neighboring states.

11.3.1 Harmonic perturbations: Fermi’s Golden Rule

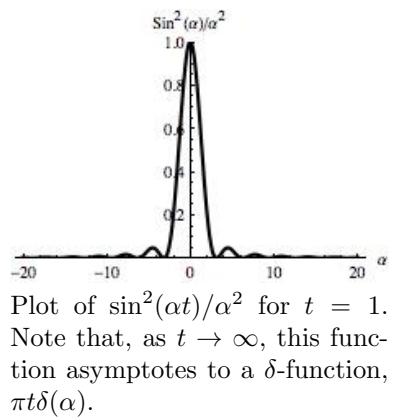
Let us then consider a system prepared in an initial state $|i\rangle$ and perturbed by a periodic harmonic potential $\hat{V}(t) = \hat{V} e^{-i\omega t}$ which is abruptly switched on at time $t = 0$. This could represent an atom perturbed by an external oscillating electric field, such as an incident light wave. What is the probability that, at some later time t , the system lies in state $|f\rangle$?

From Eq. (11.5), to first order in perturbation theory, we have

$$c_f^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle f|\hat{V}|i\rangle e^{i(\omega_f - \omega)t'} = -\frac{i}{\hbar} \langle f|\hat{V}|i\rangle \frac{e^{i(\omega_f - \omega)t} - 1}{i(\omega_f - \omega)}.$$

The probability of effecting the transition after a time t is therefore given by

$$P_{i \rightarrow f}(t) \simeq |c_f^{(1)}(t)|^2 = \frac{1}{\hbar^2} |\langle f|\hat{V}|i\rangle|^2 \left(\frac{\sin((\omega_f - \omega)t/2)}{(\omega_f - \omega)/2} \right)^2.$$



Setting $\alpha = (\omega_{\text{fi}} - \omega)/2$, the probability takes the form $\sin^2(\alpha t)/\alpha^2$ with a peak at $\alpha = 0$, with maximum value t^2 and width of order $1/t$ giving a total weight of order t . The function has more peaks positioned at $\alpha t = (n + 1/2)\pi$. These are bounded by the denominator at $1/\alpha^2$. For large t their contribution comes from a range of order $1/t$ also, and as $t \rightarrow \infty$ the function tends towards a δ -function centred at the origin, but multiplied by t , i.e. the likelihood of transition is proportional to time elapsed. We should therefore divide by t to get the transition rate.

Finally, with the normalisation, $\int_{-\infty}^{\infty} d\alpha (\frac{\sin(\alpha t)}{\alpha})^2 = \pi t$, we may effect the replacement, $\lim_{t \rightarrow \infty} \frac{1}{t} (\frac{\sin(\alpha t)}{\alpha})^2 = \pi \delta(\alpha) = 2\pi \delta(2\alpha)$ leading to the following expression for the transition rate,

$$\Gamma_{\text{i} \rightarrow \text{f}} = \lim_{t \rightarrow \infty} \frac{P_{\text{i} \rightarrow \text{f}}(t)}{t} = \frac{2\pi}{\hbar^2} |\langle \text{f} | \hat{V} | \text{i} \rangle|^2 \delta(\omega_{\text{fi}} - \omega). \quad (11.6)$$

This expression is known as **Fermi’s Golden Rule**.² It is often written with the delta-function applying to the energy, as

$$\boxed{\Gamma_{\text{i} \rightarrow \text{f}}(t) = \frac{2\pi}{\hbar} |\langle \text{f} | \hat{V} | \text{i} \rangle|^2 \delta(E_{\text{f}} - E_{\text{i}} - \hbar\omega).} \quad (11.7)$$

One might worry that, in the long time limit, we found that the probability of transition is in fact diverging — so how can we justify the use of perturbation theory? For a transition with $\omega_{\text{fi}} \neq \omega$, the “long time” limit is reached when $t \gg 1/(\omega_{\text{fi}} - \omega)$, a value that can still be very short compared with the mean transition time, which depends on the matrix element. In fact, Fermi’s Rule agrees extremely well with experiment when applied to atomic systems.

▷ INFO. **Alternative derivation of the Golden Rule:** When light falls on an atom, the full periodic potential is not suddenly imposed on an atomic time scale, but builds up over many cycles (of the atom and of the light). If we assume that $\hat{V}(t) = e^{i\varepsilon t} \hat{V} e^{-i\omega t}$, with ε very small, \hat{V} is switched on very gradually in the past, and we are looking at times much smaller than $1/\varepsilon$. We can then take the initial time to be $-\infty$, that is,

$$c_{\text{f}}^{(1)}(t) = -\frac{i}{\hbar} \int_{-\infty}^t \langle \text{f} | \hat{V} | \text{i} \rangle e^{i(\omega_{\text{fi}} - \omega - i\varepsilon)t'} dt' = -\frac{1}{\hbar} \frac{e^{i(\omega_{\text{fi}} - \omega - i\varepsilon)t}}{\omega_{\text{fi}} - \omega - i\varepsilon} \langle \text{f} | \hat{V} | \text{i} \rangle,$$

i.e. $|c_{\text{f}}(t)|^2 = \frac{1}{\hbar^2} \frac{e^{2\varepsilon t}}{(\omega_{\text{fi}} - \omega)^2 + \varepsilon^2} |\langle \text{f} | \hat{V} | \text{i} \rangle|^2$. Applied to the transition rate $\frac{d}{dt} |c_{\text{f}}^{(1)}(t)|^2$, the identity $\lim_{\varepsilon \rightarrow 0} \frac{2\varepsilon}{(\omega_{\text{fi}} - \omega)^2 + \varepsilon^2} \rightarrow 2\pi \delta(\omega_{\text{fi}} - \omega)$ leads to the Golden Rule.

From the expression for the Golden rule (11.7) we see that, for transitions to occur, and to satisfy energy conservation:

- (a) the final states must exist over a continuous energy range to match $\Delta E = \hbar\omega$ for fixed perturbation frequency ω , or
- (b) the perturbation must cover a sufficiently wide spectrum of frequency so that a discrete transition with a fixed $\Delta E = \hbar\omega$ is possible.

For two discrete states, since $|V_{\text{fi}}|^2 = |V_{\text{if}}|^2$, we have the semiclassical result $P_{\text{i} \rightarrow \text{f}} = P_{\text{f} \rightarrow \text{i}}$ — a statement of **detailed balance**.

²Curiously, although named after Fermi, most of the work leading to the Golden Rule was undertaken in an earlier work by Dirac (P. A. M. Dirac, *The quantum theory of emission and absorption of radiation*. Proc. Roy. Soc. (London) A **114**, 243265 (1927)) who formulated an almost identical equation, including the three components of a constant, the matrix element of the perturbation and an energy difference. It is given its name due to the fact that, being such a useful relation, Fermi himself called it “Golden Rule No. 2” (E. Fermi, *Nuclear Physics*, University of Chicago Press, 1950).

Enrico Fermi 1901-1954:

An Italian physicist most noted for his work on the development of the first nuclear reactor, and for his contributions to the development of quantum theory, nuclear and particle physics, and statistical mechanics. Fermi was awarded the Nobel Prize in Physics in 1938 for his work on induced radioactivity and is today regarded as one of the most influential scientists of the 20th century. He is acknowledged as a unique physicist who was highly accomplished in both theory and experiment. Fermium, a synthetic element created in 1952, is named after him.



11.3.2 INFO: Harmonic perturbations: second-order transitions

Although first order perturbation theory is often sufficient to describe transition probabilities, sometimes the first order matrix element, $\langle f|\hat{V}|i\rangle$ is identically zero due to symmetry (e.g. under parity, or through some selection rule, etc.), but other matrix elements are non-zero. In such cases, the transition may be accomplished by an indirect route. We can estimate the transition probabilities by turning to the second order of perturbation theory (11.5),

$$c_f^{(2)}(t) = -\frac{1}{\hbar^2} \sum_m \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' e^{i\omega_{fm}t'+i\omega_{mi}t''} V_{fm}(t') V_{mi}(t'').$$

If, as above, we suppose that a harmonic potential perturbation is gradually switched on, $\hat{V}(t) = e^{\varepsilon t} \hat{V} e^{-i\omega t}$, with the initial time $t_0 \rightarrow -\infty$, we have

$$c_f^{(2)}(t) = -\frac{1}{\hbar^2} \sum_m \langle f|\hat{V}|m\rangle \langle m|\hat{V}|i\rangle \int_{-\infty}^t dt' \int_{-\infty}^{t'} dt'' e^{i(\omega_{fm}-\omega-i\varepsilon)t'} e^{i(\omega_{mi}-\omega-i\varepsilon)t''}.$$

The integrals are straightforward, and yield

$$c_n^{(2)} = -\frac{1}{\hbar^2} e^{i(\omega_{fi}-2\omega)t} \frac{e^{2\varepsilon t}}{\omega_{fi} - 2\omega - 2i\varepsilon} \sum_m \frac{\langle f|\hat{V}|m\rangle \langle m|\hat{V}|i\rangle}{\omega_m - \omega_i - \omega - i\varepsilon}.$$

Then, following our discussion above, we obtain the transition rate:

$$\frac{d}{dt} |c_n^{(2)}|^2 = \frac{2\pi}{\hbar^4} \left| \sum_m \frac{\langle f|\hat{V}|m\rangle \langle m|\hat{V}|i\rangle}{\omega_m - \omega_i - \omega - i\varepsilon} \right|^2 \delta(\omega_{fi} - 2\omega).$$

This is a transition in which the system gains energy $2\hbar\omega$ from the harmonic perturbation, i.e. two “photons” are absorbed in the transition, the first taking the system to the intermediate energy ω_m , which is short-lived and therefore not well defined in energy – indeed there is no energy conservation requirement for the virtual transition into this state, only between initial and final states. Of course, if an atom in an arbitrary state is exposed to monochromatic light, other second order processes in which two photons are emitted, or one is absorbed and one emitted (in either order) are also possible.

Chapter 12

Scattering theory

Almost everything we know about nuclei and elementary particles has been discovered in scattering experiments, from Rutherford's surprise at finding that atoms have their mass and positive charge concentrated in almost point-like nuclei, to the more recent discoveries, on a far smaller length scale, that protons and neutrons are themselves made up of apparently point-like quarks. More generally, the methods that we have to probe the properties of condensed matter systems rely fundamentally on the notion of scattering. In this section, we will provide a brief introduction to the concepts and methodology of scattering theory.

As preparation for the quantum mechanical scattering problem, let us first consider the classical problem. This will allow us to develop some elementary concepts of scattering theory, and to introduce some notation. In a classical scattering experiment, one considers particles of energy $E = \frac{1}{2}mv_0^2$ (mass m and asymptotic speed v_0), incident upon a target with a central potential $V(r)$. For a repulsive potential, particles are scattered through an angle θ (see figure). The **scattering cross-section**, σ , can be inferred from the number of particles dn scattered into some element of solid angle, $d\Omega$, at angle (θ, ϕ) , i.e. for an incident flux j_i (number of particles per unit time per unit area), $dn = j_i \sigma d\Omega$. The total cross-section is then obtained as $\sigma_T = \int d\Omega \sigma(\theta, \phi) = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \sigma(\theta, \phi)$. The angle of deflection of the beam depends on the impact parameter, b (see figure right). We therefore have that $dn = j_i b db d\phi = j_i \sigma \sin \theta d\theta d\phi$ and

$$\sigma(\theta, \phi) = \frac{b}{\sin \theta} \frac{db}{d\theta}.$$

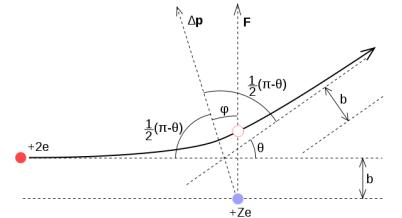
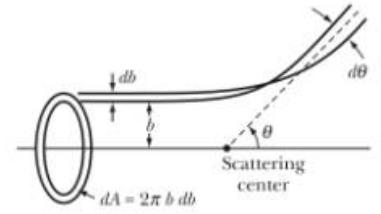
▷ **EXAMPLE:** Let us consider then the case of **classical Coulomb scattering** from a repulsive potential $V(r) = \frac{\kappa}{r}$ where $\kappa > 0$. From classical physics, we know that the particle will follow a hyperbolic trajectory with

$$r = \frac{L^2}{m\kappa(e \cos \varphi - 1)},$$

where $\mathbf{r} = (r, \varphi)$ parameterises the relative coordinates of the particle and target,¹ and $e = (1 + \frac{2EL^2}{\kappa^2 m})^{1/2} > 1$ denotes the eccentricity. Since the potential is central, the angular momentum L is conserved and can be fixed asymptotically by the condition $L = mv_0 b$.

To obtain the scattering angle, θ , we can use the relation above to find the limiting angle, $\cos \varphi_0 = 1/e$, where $\varphi_0 = (\pi - \theta)/2$. We therefore have $\tan(\theta/2) = \cot \varphi_0 =$

¹Note that the angle φ is distinct form the azimuthal angle ϕ associated with the axis of scattering.



$1/\sqrt{e^2 - 1} = (\frac{m\kappa^2}{2EL^2})^{1/2} = \frac{\kappa}{2Eb}$. Then, from this relation, we obtain the cross-section

$$\sigma = \frac{b}{\sin \theta} \frac{db}{d\theta} = \frac{\kappa^2}{16E^2} \frac{1}{\sin^4 \theta/2},$$

known as the **Rutherford formula**.

12.1 Basics

Let us now turn to the quantum mechanical problem of a beam of particles incident upon a target. The potential of the target, $V(r)$, might represent that experienced by a fast electron striking an atom, or an α particle colliding with a nucleus. As in the classical problem, the basic scenario involves directing a stream or flux of particles, all at the same energy, at a target and detect how many particles are deflected into a battery of detectors which measure angles of deflection. In principle, if we assume that all the in-going particles are represented by wavepackets of the same shape and size, our challenge is to solve the full time-dependent Schrödinger equation for such a wavepacket,

$$i\hbar\partial_t\Psi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Psi(\mathbf{r},t),$$

and find the probability amplitudes for out-going waves in different directions at some later time after scattering has taken place. However, if the incident beam of particles is switched on for times very long as compared with the time a particle would take to cross the interaction region, steady-state conditions apply. Moreover, if we assume that the wavepacket has a well-defined energy (and hence momentum), so it is many wavelengths long, and we may consider it a plane wave. Setting $\Psi(\mathbf{r},t) = \psi(\mathbf{r})e^{-iEt/\hbar}$, we may therefore look for solutions $\psi(\mathbf{r})$ of the time-*independent* Schrödinger equation,

$$E\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}),$$

subject to the boundary condition that the incoming component of the wavefunction is a plane wave, $e^{i\mathbf{k}\cdot\mathbf{x}}$. Here $E = \mathbf{p}^2/2m = \hbar^2\mathbf{k}^2/2m$ denotes the energy of the incoming particles while their flux is given by

$$\mathbf{j} = -i\frac{\hbar}{2m}(\psi^*\nabla\psi - \psi\nabla\psi^*) = \frac{\hbar\mathbf{k}}{m}.$$

In the one-dimensional geometry, the impact of a plane wave with the localized target resulted in a portion of the wave being reflected and a portion transmitted through the potential region. From energy conservation, we may deduce that both components of the outgoing scattered wave are plane waves with wavevector $\pm k$, while the influence of the potential is encoded in the amplitude of the reflected and transmitted beams, and a potential phase shift. Both amplitudes and phase shifts are then determined by solving the time-independent Schrödinger equation subject to the boundary conditions which ensure energy and flux conservation. In the three-dimensional system, the phenomenology is similar: In this case, the wavefunction well outside the localized target region will involve a superposition of the incident plane wave and the scattered wave

$$\psi(\mathbf{r}) \simeq e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta, \phi) \frac{e^{ikr}}{r},$$

where the function $f(\theta, \phi)$ records the relative amplitude and phase of the scattered components along the direction (θ, ϕ) relative to the incident beam.

The corresponding asymptotic flux is then given by

$$\mathbf{j} = \frac{\hbar}{m} \text{Im} \left\{ \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta) \frac{e^{ikr}}{r} \right]^* \nabla \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta) \frac{e^{ikr}}{r} \right] \right\}.$$

In general, an expansion then leads to a formidable collection of contributing terms. However, for most of these contributions, there remains an exponential factor, $e^{\pm ikr(1-\cos\theta)}$ where θ denotes the angle between \mathbf{k} and \mathbf{r} . For $r \rightarrow \infty$, the small angular integration implied by any physical measurement leads to a fast oscillation of this factor. As a result, such terms are strongly suppressed and can be neglected. Retaining only those terms where the phase cancellation is complete, we obtain

$$\mathbf{j} = \frac{\hbar\mathbf{k}}{m} + \frac{\hbar k}{m} \hat{\mathbf{e}}_r \frac{|f(\theta)|^2}{r^2} + O(1/r^3).$$

The first term represents the incident flux, while the remainder describes the radial flux of scattered particles. In particular, the number of particles crossing the area that subtends a solid angle $d\Omega$ at the origin (the target) is given by

$$\mathbf{j} \cdot \hat{\mathbf{e}}_r dA = \frac{\hbar k}{m} \frac{|f(\theta)|^2}{r^2} r^2 d\Omega + O(1/r).$$

The terms of order $1/r$ can be dropped, as they are negligible in the asymptotic limit.

The **differential cross-section** is defined as

$$\frac{d\sigma}{d\Omega} = \frac{\text{rate of scattering into } d\Omega}{\text{incident flux} \times d\Omega}$$

Thus, it is the ratio of the scattered flux $\mathbf{j} \cdot \hat{\mathbf{e}}_r$ times dA to the incident flux $(\hbar k/m)$ times $d\Omega$, giving

$$\boxed{\frac{d\sigma}{d\Omega} = |f(\theta)|^2.}$$

The total cross-section is then given by $\sigma_T = \int d\sigma = \int |f(\theta)|^2 d\Omega$.

12.2 The Born approximation

We shall consider only the case where the scattering is weak. We can then derive the scattering properties perturbatively.

We view the scattering as a time-dependent process, in which an incident plane wave with wavevector \mathbf{k} is scattered into an outgoing wave $\mathbf{k}' \neq \mathbf{k}$ by the potential $V(\mathbf{r})$. We calculate the rate of these transitions, employing Fermi's golden rule in the limit of zero frequency (since the potential is time-independent). We work with normalized plane wave states, defined on a system of volume L^3 with periodic boundary conditions.

Consider a particle initially in a plane wave state $|\mathbf{k}\rangle$, with position representation

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}}.$$

Note that the state is normalized to unity. Therefore a particle incident in state $|\mathbf{k}\rangle$ has flux $(\hbar k/m)(1/L^3)$.

By Fermi's golden rule, the rate at which the particle is scattered into states $|\mathbf{k}'\rangle$ within a small solid angle $d\Omega$ is

$$\Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega} = \frac{2\pi}{\hbar} \sum_{\mathbf{k}' \in d\Omega} |\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle|^2 \delta(E_{k'} - E_k)$$

where $E_k = \frac{\hbar^2 |\mathbf{k}|^2}{2m}$. The matrix element is

$$\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \frac{1}{L^3} \int d^3r V(\mathbf{r}) e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} \equiv \frac{1}{L^3} \tilde{V}(\mathbf{q})$$

which defines \tilde{V} and the **wavevector transfer** $\mathbf{q} \equiv \mathbf{k}' - \mathbf{k}$. The expression for the rate becomes

$$\begin{aligned} \Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega} &= \frac{2\pi}{\hbar} \frac{1}{L^6} \sum_{\mathbf{k}' \in d\Omega} |\tilde{V}(\mathbf{q})|^2 \delta(E_{k'} - E_k) \\ &= \frac{2\pi}{\hbar} \frac{1}{L^6} \frac{L^3}{(2\pi)^3} d\Omega \int_0^\infty k'^2 dk' |\tilde{V}(\mathbf{q})|^2 \delta\left[\frac{\hbar^2}{2m}(k'^2 - k^2)\right] \\ &= \frac{2\pi}{\hbar} \frac{1}{L^3} \frac{1}{(2\pi)^3} d\Omega k^2 \frac{m}{\hbar^2 k} |\tilde{V}(\mathbf{q})|^2 \end{aligned}$$

with the constraint from the delta-function that $|\mathbf{k}'| = |\mathbf{k}|$. As a result of this constraint, the momentum transfer \mathbf{q} can be written as $|\mathbf{q}| = 2|\mathbf{k}| \sin(\theta/2)$ where θ is the scattering angle.

The differential cross-section is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{\text{rate of scattering into } d\Omega}{\text{incident flux} \times d\Omega} \\ &= \frac{\Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega}}{(\hbar k/m)(1/L^3)d\Omega} \\ &= \frac{1}{(2\pi)^2} \frac{mk}{\hbar^3} \frac{1}{L^3} d\Omega |\tilde{V}(\mathbf{q})|^2 \times \frac{mL^3}{\hbar k d\Omega} \\ &= \left| \frac{m}{2\pi\hbar^2} \tilde{V}(\mathbf{q}) \right|^2 \end{aligned}$$

Thus, we obtain the **Born approximation** expression for the scattering cross section

$$\boxed{\frac{d\sigma}{d\Omega} = \left| \frac{m}{2\pi\hbar^2} \int V(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} d^3\mathbf{r} \right|^2}$$

The scattering cross-section is related to the Fourier transform of the scattering potential.

▷ INFO. Since the potential is time-independent one can also derive the scattering cross-section within the Born approximation using standard time-independent perturbation theory. One calculates corrections to the wavefunction of a particle in state $|\mathbf{k}\rangle$, that are first order in \hat{V} , and that involve the other states $|\mathbf{k}'\rangle$. However, note that the scattering involves coupling to degenerate states ($E_{k'} = E_k$). As a result, the perturbative correction has a singular amplitude. When integrating over \mathbf{k}' , careful handling of this singularity is required. The physically relevant solution is found by imposing the boundary condition that the correction to the wavefunction involves only *outward* propagating waves. This is a slightly technical point which is unsuitable for this core lecture course. It is covered in TP2, and related issues of causality appear in TP1.

Chapter 13

Radiative transitions

Previously, we have addressed the quantum theory of atoms coupled to a time-independent classical electromagnetic field, cf. our discussion of the Zeeman and Stark effects. However, to develop a complete quantum mechanical description of light-matter interaction, we have to address both the quantum theory of the electromagnetic field and the coupling of light to matter. In the following section, we will address both of these issues in turn. Our motivation for developing such a consistent theory is that it will (a) provide us with a platform to address the problem of radiative transitions in atoms and (b) it forms the basis of the field of **quantum optics**.

13.1 Coupling of matter to the electromagnetic field

Let us then consider the Hamiltonian of a single-electron atom subject to a time-dependent external electromagnetic field,

$$\hat{H}_{\text{atom}} = \frac{1}{2m} (\hat{\mathbf{p}} + e\mathbf{A}(\hat{\mathbf{r}}, t))^2 - e\phi(\hat{\mathbf{r}}, t) + V(\hat{\mathbf{r}}).$$

Here $V(\mathbf{r})$ denotes the binding potential associated with the atomic nucleus. To keep our discussion of a complex problem as simple as possible, we focus on the single electron system. However, a generalization of the methodology to multi-electron atoms would not present significant challenges. Expanding the kinetic energy, the atomic Hamiltonian can be recast as $\hat{H}_{\text{atom}} = \hat{H}_0 + \hat{H}_{\text{para}} + \hat{H}_{\text{dia.}}$, where

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}),$$

denotes the usual non-interacting Hamiltonian of the isolated atom,

$$\boxed{\hat{H}_{\text{para}}(t) = \frac{e}{m} \hat{\mathbf{A}}(t) \cdot \hat{\mathbf{p}},}$$

represents the time-dependent paramagnetic term arising from the coupling of the electron to the electromagnetic field, and $\hat{H}_{\text{dia.}} = (e\mathbf{A})^2/2m$ represents the diamagnetic term. Since we will be interested in the absorption and emission of single photons, we shall be able to neglect the influence of the diamagnetic term which presents only a small perturbation in the atomic system.

13.1.1 Quantum fields

As we shall see in chapter 14.2, the quantum Hamiltonian for the electromagnetic field can be expressed as,

$$\hat{H}_{\text{rad}} = \sum_{\mathbf{k}, \lambda=1,2} \hbar \omega_{\mathbf{k}} \left(a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + \frac{1}{2} \right),$$

where the operators $a_{\mathbf{k}\lambda}^\dagger$ and $a_{\mathbf{k}\lambda}$ create and annihilate photons with wavevector \mathbf{k} and polarization λ , and $\omega_{\mathbf{k}} = c|\mathbf{k}|$. These ladder operators obey the (bosonic) commutation relations, $[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}$, with $[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}] = [a_{\mathbf{k}\lambda}^\dagger, a_{\mathbf{k}'\lambda'}^\dagger] = 0$, and act on photon number states as

$$a_{\mathbf{k}\lambda}|n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda}}|n_{\mathbf{k}\lambda}-1\rangle$$

$$a_{\mathbf{k}\lambda}^\dagger|n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda}+1}|N_{\mathbf{k}\lambda}+1\rangle.$$

Here $|n_{\mathbf{k}\lambda}\rangle$ represents a photon number state with $n_{\mathbf{k},\lambda}$ photons in the mode $(\mathbf{k}\lambda)$. In terms of these field operators, the vector potential in the Heisenberg representation can be expanded as

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}, \lambda} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_{\mathbf{k}}t)} \right).$$

Taken together, $\hat{H} = \hat{H}_0 + \hat{H}_{\text{rad}} + \hat{H}_{\text{para}}(t)$ specify the full quantum mechanical Hamiltonian of the atom light system.

13.1.2 Spontaneous emission

With this background, let us now consider the probability for an atom, initially in a state $|i\rangle$ to make a transition to a state $|f\rangle$ leading to the emission of a photon of wavevector \mathbf{k} and polarization λ – a process of spontaneous emission. If we suppose that the radiation field is initially prepared in the vacuum state, $|0\rangle$, then the final state involves one photon, $|\mathbf{k}\lambda\rangle = a_{\mathbf{k}\lambda}^\dagger|0\rangle$. Therefore, making use of Fermi's Golden rule (11.7), with the perturbation $\hat{H}_{\text{para}} = \frac{e}{m}\hat{\mathbf{A}}(t)\cdot\hat{\mathbf{p}}$, we have the transition probability

$$\Gamma_{i\rightarrow f}(t) = \frac{2\pi}{\hbar^2} |\langle f | \otimes \langle \mathbf{k}\lambda | \hat{H}_{\text{para}} | i \rangle \otimes | 0 \rangle|^2 \delta(\omega_{if} - \omega_{\mathbf{k}})$$

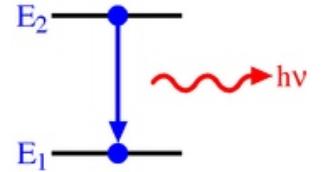
where $\omega_{if} = (E_i - E_f)/\hbar$. Then substituting the field operator expansion of $\hat{\mathbf{A}}$, we have

$$\Gamma_{i\rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \otimes \langle 0 | a_{\mathbf{k}\lambda} \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \otimes | 0 \rangle \right|^2 \delta(\omega_{if} - \omega_{\mathbf{k}})$$

As a result, we finally obtain

$$\Gamma_{i\rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_i - E_f - \hbar\omega_{\mathbf{k}})$$

To determine the transition rate, we have to analyse matrix elements of the form $\langle f | e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle$. Let us begin by estimating its magnitude. For a typical atomic state, $\langle \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} \rangle \simeq p \simeq Zmc\alpha$, where we have included a general nuclear



Schematic showing spontaneous emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$.

charge, Z . But what about the exponential factor? With $r \sim \hbar/p \simeq \hbar/mZc\alpha$, and $\omega_{\mathbf{k}} = c|\mathbf{k}| \simeq \frac{p^2}{2m}$ (for electronic transitions), we have

$$\mathbf{k} \cdot \mathbf{r} \simeq \frac{\omega_{\mathbf{k}} \hbar}{c p} \simeq \frac{\hbar p}{mc} \simeq Z\alpha.$$

This means that, for $Z\alpha \ll 1$, we can expand the exponential as a power series in $\mathbf{k} \cdot \mathbf{r}$ with the lowest terms being dominant. Taking the zeroth order term, and making use of the operator identity, $\frac{\hat{\mathbf{p}}}{m} = \frac{i}{\hbar}[\hat{H}_0, \mathbf{r}]$, the matrix element may be written as

$$\langle f | \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle = m \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \langle f | \frac{i}{\hbar} [\hat{H}_0, \mathbf{r}] | i \rangle = im \frac{E_f - E_i}{\hbar} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \langle f | \mathbf{r} | i \rangle = -im\omega_{\mathbf{k}} \langle f | \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \mathbf{r} | i \rangle.$$

This result, which emerges from the leading approximation in $Z\alpha$, is known as the **electric dipole approximation**. Effectively, we have set (exercise)

$$\frac{e}{m} \hat{\mathbf{A}}(\mathbf{r}, t) \cdot \hat{\mathbf{p}} \simeq e \hat{\mathbf{E}}(\mathbf{r}, t) \cdot \hat{\mathbf{r}},$$

translating to the potential energy of a dipole, with moment $\mathbf{d} = -e\mathbf{r}$, in an oscillating electric field.

13.1.3 Absorption and stimulated emission

Let us now consider the absorption of a photon with wave number \mathbf{k} , and polarization λ . If we assume that, in the initial state, there are $n_{\mathbf{k}\lambda}$ photons in state $(\mathbf{k}\lambda)$ then, after the transition, there will be $n_{\mathbf{k}\lambda} - 1$. Then, if the initial state of the atom is $|i\rangle$ and the final state is $|f\rangle$, the transition amplitude involves the matrix element,

$$\begin{aligned} \langle f | \otimes \langle (n_{\mathbf{k},\lambda} - 1) | \hat{H}_{\text{para}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle \\ = \langle f | \otimes \langle (n_{\mathbf{k},\lambda} - 1) | \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle \end{aligned}$$

Then, using the relation $a_{\mathbf{k}\lambda} | n_{\mathbf{k}\lambda} \rangle = \sqrt{n_{\mathbf{k}\lambda}} (n_{\mathbf{k}\lambda} - 1) \rangle$,

$$\langle f | \otimes \langle (n_{\mathbf{k},\lambda} - 1) | \hat{H}_{\text{para}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle = \langle f | \frac{e}{m} \sqrt{\frac{\hbar n_{\mathbf{k}\lambda}}{2\epsilon_0\omega_{\mathbf{k}}V}} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{p}} | i \rangle$$

As a result, using Fermi's Golden rule, we obtain the transition amplitude,

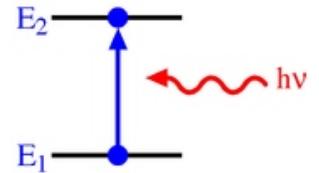
$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar n_{\mathbf{k}\lambda}}{2\epsilon_0\omega_{\mathbf{k}}V}} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega_{\mathbf{k}})$$

In particular, we find that the absorption rate *increases* linearly with photon number, $n_{\mathbf{k}\lambda}$.

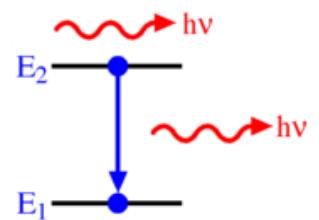
Similarly, if we now consider the emission process in which there are already $n_{\mathbf{k},\lambda}$ photons in the initial state, we have the revised transition rate,

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar(n_{\mathbf{k},\lambda} + 1)}{2\epsilon_0\omega_{\mathbf{k}}V}} e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega_{\mathbf{k}}).$$

This enhancement of the transition rate by the photon occupancy is known as **stimulated emission**.



Schematic showing absorption from an initial state at energy $E_i = E_1$ to a final state at energy $E_f = E_2$.



Schematic showing the stimulated emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$.

Altogether, in the dipole approximation, we have the transition rates,

$$\Gamma_{i \rightarrow f, k\lambda} = \frac{\pi \omega_k}{\epsilon_0 V} |\langle 2 | \hat{e}_{k\lambda} \cdot \mathbf{d} | 1 \rangle|^2 \begin{cases} n_{k\lambda} \delta(E_f - E_i - \hbar\omega_k) & \text{absorption} \\ (n_{k\lambda} + 1) \delta(E_i - E_f - \hbar\omega_k) & \text{emission} \end{cases}$$

where $|1\rangle$ and $|2\rangle$ are the states with energies E_1 and E_2 (with $E_2 > E_1$). If there are no photons present initially, this expression reduces to that obtained from spontaneous emission. The $n_{k\lambda}$ -independent component of the expression for absorption and emission coincide, an equality known as **detailed balance**.

If we are interested in the total rate, $d\Gamma_\lambda$ at which photons of polarization λ are scattered into the solid angle $d\Omega$, we must compute $dR_\lambda = \sum_{k \in d\Omega} \Gamma_{i \rightarrow f, k\lambda}$. Since, in the elemental volume $d^3 k = k^2 dk d\Omega$, there are $d^3 k V / (2\pi)^3$ states, we may set $\sum_k = \frac{V}{(2\pi)^3} \int k^2 dk d\Omega$. Finally, if we assume that the photon occupation of state $(k\lambda)$ is isotropic, dependent only on $|\mathbf{k}|$, we find that the integrated transition rate per unit solid angle is given by $\frac{dR_\lambda}{d\Omega} = V \int \frac{k^2 dk}{(2\pi)^3} \Gamma_{i \rightarrow f, k\lambda}$ from which we obtain

$$\frac{dR_\lambda}{d\Omega} = \frac{1}{4\pi\epsilon_0} \frac{\omega^3}{2\pi\hbar c^3} |\langle 2 | \hat{e}_{k\lambda} \cdot \mathbf{d} | 1 \rangle|^2 \begin{cases} n_\lambda(\omega) & \text{absorption} \\ n_\lambda(\omega) + 1 & \text{emission} \end{cases}$$

Here, in carrying out the integral, we have used the relation $\omega_k = c|\mathbf{k}|$ and $\hbar\omega = |E_f - E_i|$. For a thermal distribution of photons, with the energy density specified by the Planck formula,

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \bar{n}_\lambda(\omega), \quad \bar{n}_\lambda(\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1},$$

this equates to an absorption / stimulated emission rate,

$$\frac{dR_\lambda}{d\Omega} = \frac{1}{4\epsilon_0} \frac{1}{2\hbar^2} |\langle 2 | \hat{e}_{k\lambda} \cdot \mathbf{d} | 1 \rangle|^2 u(\omega)$$

From these expressions, we can obtain the power loss as $P_\lambda = \hbar\omega R_\lambda$. Before discussing the selection rules implied by the form of the dipolar coupling, it is first helpful to digress and discuss connections of this result to a famous result due to Einstein.

13.1.4 Einstein's A and B coefficients

In fact, the frequency dependence of the spontaneous emission rate can be inferred without invoking quantum field theoretic methods by means of an ingenious argument due to Einstein which showed that the stimulated and spontaneous transitions must be related. Consider an ensemble of atoms exposed to a black-body radiation field at temperature T . Let us consider transitions between two states $|\psi_j\rangle$ and $|\psi_k\rangle$, with $E_k - E_j = \hbar\omega$. Suppose the numbers of atoms in the two states are n_j and n_k . The possible transitions and their rates per atom are given by:

$$\begin{array}{lll} \text{absorption} & j \rightarrow k & B_{j \rightarrow k} u(\omega) \\ \text{stimulated emission} & k \rightarrow j & B_{k \rightarrow j} u(\omega) \\ \text{spontaneous emission} & k \rightarrow j & A_{k \rightarrow j}(\omega) \end{array}$$

where $u(\omega)$, the energy density of radiation per unit ω . A and B are known as Einstein's A and B coefficients, and, as we have seen, are properties of the atomic states concerned.

Now, in thermodynamic equilibrium, the rates must balance, so that

$$n_k [A_{k \rightarrow j}(\omega) + B_{k \rightarrow j} u(\omega)] = n_j B_{j \rightarrow k} u(\omega).$$

At the same time, the relative populations of the two states (assumed non-degenerate for simplicity), are given by a Boltzmann factor

$$\frac{n_j}{n_k} = \frac{e^{-E_j/k_B T}}{e^{-E_k/k_B T}} = e^{\hbar\omega/k_B T}.$$

Thus we have:

$$A_{k \rightarrow j}(\omega) = \left[B_{j \rightarrow k} e^{\hbar\omega/k_B T} - B_{k \rightarrow j} \right] u(\omega). \quad (13.1)$$

For a black-body, the energy density $u(\omega)$ is just given by Planck's formula, $u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}$. The $A_{k \rightarrow j}$ coefficient in Eq. (13.1) certainly cannot depend on temperature, so T must cancel on the right hand side. Hence,

$$B_{k \rightarrow j} = B_{j \rightarrow k} \quad \text{and} \quad A_{k \rightarrow j}(\omega) = B_{k \rightarrow j} \frac{\hbar\omega^3}{\pi^2 c^3}.$$

So, the A and B coefficients are related, and if we can calculate the B coefficient for stimulated emission from Fermi's golden rule, we can infer A , and *vice versa*.

13.2 Selection rules

It is clear from the formulae for the transition rates that no transition, either spontaneous or stimulated, will occur between the states $|i\rangle$ and $|f\rangle$ unless at least one component of the dipole transition matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ is non-zero. It is often possible to show that the matrix elements are zero for certain pairs of states. If so, the transition is not allowed (at least in the electric dipole approximation), and the results can often be summarised in terms of simple **selection rules** governing the allowed changes in quantum numbers in transitions.

Since the dipole operator $\hat{\mathbf{d}} = q\hat{\mathbf{r}}$ for the atom¹ changes sign under parity ($\mathbf{r} \rightarrow -\mathbf{r}$), the matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ will trivially vanish if the states $|f\rangle$ and $|i\rangle$ have the same parity.² Therefore, **the parity of the wavefunction must change in an electric dipole transition**.

To determine further restrictions on the possible transitions induced by electric dipole coupling, we consider first the case in which spin-orbit interactions are absent. Then the energy eigenstates are also eigenstates of the electron spin $\hat{\mathbf{S}}^2$ and its z -component \hat{S}_z , with eigenvalues $s(s+1)\hbar^2$ and $m_s\hbar$.³ Since the dipole operator only acts on the spatial part of the wavefunction, it commutes with the spin: $[\hat{\mathbf{S}}^2, \hat{\mathbf{d}}] = [\hat{S}_z, \hat{\mathbf{d}}] = 0$. Taking the matrix element of the relation $\hat{S}_z \hat{\mathbf{d}} = \hat{\mathbf{d}} \hat{S}_z$ gives $\langle f|\hat{S}_z \hat{\mathbf{d}}|i\rangle = \langle f|\hat{\mathbf{d}} \hat{S}_z|i\rangle$ which leads to $\hbar(m_s^f - m_s^i)\langle f|\hat{\mathbf{d}}|i\rangle = 0$ where $\hbar m_s^{i/f}$ is the eigenvalue of \hat{S}_z for the initial/final state. Hence if the matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ is to be non-zero, we must have

¹More generally, for an atom with N electrons we should consider $\mathbf{d} = q \sum_{i=1}^N \hat{\mathbf{r}}_i$. The conclusions remain unchanged.

²This result may be found in an operator approach, by noting that $\hat{P}\hat{\mathbf{d}} = -\hat{\mathbf{d}}\hat{P}$ where \hat{P} is the parity operator. Hence $\langle f|\hat{P}\hat{\mathbf{d}}|i\rangle = -\langle f|\hat{\mathbf{d}}\hat{P}|i\rangle$ and so $(P_f + P_i)\langle f|\hat{\mathbf{d}}|i\rangle = 0$ where $P_{i,f}$ are the parities (± 1) of the initial/final states. A non-zero matrix element $\langle f|\hat{\mathbf{d}}|i\rangle \neq 0$ requires $P_f = -P_i$, *i.e.* that the parity changes.

³We consider the possibility of a multi-electron atom so the total spin s need not just be $s = 1/2$.

$\Delta m_s = (m_s^f - m_s^i) = 0$. A similar argument holds for the change in total spin, and we find the selection rule

$$\Delta s = 0, \quad \Delta m_s = 0 \quad (13.2)$$

The spin state is not altered in an electric dipole transition.

Let us now consider the selection rules for the orbital angular momenta. From the operator identity, $[\hat{L}_i, r_j] = i\hbar\epsilon_{ijk}r_k$ (exercise), it follows that

$$[\hat{L}_z, \hat{z}] = 0, \quad [\hat{L}_z, \hat{x} \pm i\hat{y}] = \pm(\hat{x} \pm i\hat{y})\hbar,$$

We therefore obtain the relation,

$$\langle \ell', m'_\ell | [\hat{L}_z, \hat{z}] | \ell, m_\ell \rangle = (m'_\ell - m_\ell)\hbar \langle \ell', m'_\ell | \hat{z} | \ell, m_\ell \rangle = 0.$$

Similarly, since $\langle \ell', m'_\ell | [\hat{L}_z, \hat{x} \pm i\hat{y}] | \ell, m_\ell \rangle = \pm \langle \ell', m'_\ell | \hat{x} \pm i\hat{y} | \ell, m_\ell \rangle$, it follows that

$$(m'_\ell - m_\ell \mp 1) \langle \ell', m'_\ell | \hat{x} \pm i\hat{y} | \ell, m_\ell \rangle = 0.$$

Therefore, to get non-zero component of the dipole matrix element, we require.

$$\Delta m_\ell = 0, \pm 1.$$

Similarly, using the operator identity $[\hat{\mathbf{L}}^2, [\hat{\mathbf{L}}^2, \hat{\mathbf{r}}]] | \ell, m_\ell \rangle = 2\hbar^2(\hat{\mathbf{r}}\hat{\mathbf{L}}^2 + \hat{\mathbf{L}}^2\hat{\mathbf{r}})$ (exercise), we have

$$\begin{aligned} \langle \ell', m'_\ell | [\hat{\mathbf{L}}^2, [\hat{\mathbf{L}}^2, \hat{\mathbf{r}}]] | \ell, m_\ell \rangle &= [\ell'(\ell' + 1) - \ell(\ell + 1)]^2 \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle \\ &= 2[\ell'(\ell' + 1) + \ell(\ell + 1)] \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle \end{aligned}$$

i.e. $(\ell + \ell')(\ell + \ell' + 2)[(\ell' - \ell)^2 - 1] \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle = 0$. Since $\ell, \ell' \geq 0$, we can conclude that, to effect an electric dipole transition, we must have

$$\Delta \ell = \pm 1.$$

One may summarize the selection rules for ℓ and m_ℓ by saying that the photon carries off (or brings in, in an absorption transition) one unit of angular momentum. It should be noted, however, that these rules were derived for the specific case of an electric dipole transition of the system. It is possible, though much less likely in the case of an atom, for the electromagnetic field to interact with some other observable such as the magnetic dipole moment or the electric quadrupole moment. In such transitions the selection rules are different. For example, the magnetic dipole operator is $\hat{\mu} = -\mu_B \hat{\mathbf{L}}/\hbar$ (or $-2\mu_B \hat{\mathbf{S}}/\hbar$ for the spin) and since the angular momentum does not change sign under the parity transformation, there is no change of parity in a magnetic dipole transition. To avoid confusion, we shall continue to confine the discussion to electric dipole transitions, which are responsible for the prominent lines in atomic spectra.

For transitions with $\Delta m_\ell = 0$, the dipole matrix element $\langle f | \hat{\mathbf{d}} | i \rangle \sim \hat{\mathbf{e}}_z$, so the only component of polarization is along the z -direction. Similarly, for electric dipole transitions with $m' = m \pm 1$, $\langle \ell', m' | x \mp iy | \ell, m \rangle \neq 0$, and $\langle f | \mathbf{d} | i \rangle \sim (1, \mp i, 0)$ lies in the $x - y$ plane. In this case, if the wavevector of photon lies along z , the emitted light is circularly polarized with a polarization which depends on helicity. Conversely, if the wavevector lies in xy plane, the emitted light is linearly polarized, while in general the polarization is elliptical.

We conclude by discussing how the selection rules are affected by spin-orbit coupling. In the most general case (i.e. of strong spin-orbit interaction), the

stationary states are labelled by their parity and by the quantum numbers j, m_j associated with the total angular momentum (squared), $\hat{\mathbf{J}}^2$, and its z -component \hat{J}_z . In this case, the selection rules can be inferred by looking for the conditions for non-zero matrix elements $\langle j', m'_j | \hat{\mathbf{r}} | j, m_j \rangle$. One finds that, for dipole transitions to take place, we require that

- ▷ Parity must change;
- ▷ $\Delta j = \pm 1, 0$ (but $0 \rightarrow 0$ is not allowed) and $\Delta m_j = \pm 1, 0$.

Atomic states are always eigenstates of parity and of total angular momentum so these selection rules can be regarded as absolutely valid in electric dipole transitions. It should be emphasized again, though, that the electric dipole approximation is an approximation, and higher order processes may occur, albeit at a slower rate, and have their own selection rules.

Even in the presence of spin-orbit coupling, if this coupling is sufficiently weak that the system is described by ideal LS coupling, then additional selection rules apply. Specifically, we also require:

- ▷ $\Delta s = 0$;
- ▷ $\Delta \ell = \pm 1, 0$ (but $0 \rightarrow 0$ is not allowed);
- ▷ and $\Delta \ell_i = \pm 1$ if only electron i is involved in the transition.

These follow from the fact that, in LS coupling, the states are eigenstates also of total spin $\hat{\mathbf{S}}^2$ and total orbital angular momentum $\hat{\mathbf{L}}^2$. Since the dipole operator does not operate on the spin part of the wavefunction, then the conditions (13.2) also apply. Combining this and the absolute rules relating to j , imply the rules for ℓ and m_ℓ . The rule for $\Delta \ell_i$ follows from the parity change rule, since the parity of the atom is the product of the parities of the separate electron wavefunctions, given by $(-1)^{\ell_i}$. However, since LS coupling is only an approximation, these rules should themselves be regarded as approximate.

▷ INFO. As another example of selection rules, consider a charged particle moving in a one-dimensional harmonic potential. The wavefunctions are characterised by the quantum number n , so the state $|n\rangle$ corresponds to energy $(n+1/2)\hbar\omega$. An oscillating electric field in the x -direction can induce transitions between states $|n\rangle$ and $|n'\rangle$, governed by matrix elements of the form $\langle n' | x | n \rangle$. These can be evaluated by use of the ladder operators \hat{a} and \hat{a}^\dagger . Since $x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)$, the matrix element becomes

$$\langle n' | x | n \rangle = \sqrt{\frac{\hbar}{2m\omega}} (\sqrt{n+1}\langle n' | n+1 \rangle + \sqrt{n}\langle n' | n-1 \rangle) ,$$

and therefore vanishes unless $n = n' \pm 1$. Hence the selection rule, in the electric dipole approximation, is $\Delta n = \pm 1$.

13.3 Lasers

Finally, to close this section, we will consider an important application of light matter interaction – the laser. The laser is the light source which enables modern spectroscopy. The term “laser” is an acronym for “light amplification by stimulated emission of radiation”. However, a laser not only amplifies light, but it acts as a special kind of light source which is characterised by a number of properties:

Monochromaticity: The emission of the laser generally corresponds to just one of the atomic transitions of the gain medium (in contrast to discharge lamps, which emit on all transitions). The spectral line width can be much smaller than that of the atomic transition. This is because the emission is affected by the optical cavity. In certain cases, the laser can be made to operate on just one of the modes of the cavity. Since the Q of the cavity⁴ is generally rather large, the mode is usually much narrower than the atomic transition, and the spectral line width is orders of magnitude smaller than the atomic transition.

Coherence: In discussing the coherence of an optical beam, we must distinguish between spatial and temporal coherence – laser beams have a high degree of both. **Spatial coherence** refers to whether there are irregularities in the optical phase in a cross-sectional slice of the beam. **Temporal coherence** refers to the time duration over which the phase of the beam is well defined. In general, the temporal coherence time, t_{coh} is given as the reciprocal of the spectral line width, $\Delta\nu$. Thus the coherence length ℓ_{coh} is given by,

$$\ell_{coh} = ct_{coh} = \frac{c}{\Delta\nu}.$$

Typical values of the coherence length for a number of light sources are given in the table below:

| Source | $\Delta\nu$ (Hz) | t_{coh} (s) | ℓ_{coh} (m) |
|--|--------------------|---------------------|--------------------|
| Na discharge lamp (D-lines at 589nm) | 5×10^{11} | 2×10^{-12} | 6×10^{-4} |
| Multi-mode HeNe laser (632.8nm line) | 1.5×10^9 | 6×10^{-10} | 0.2 |
| Single-mode HeNe laser (632.8nm line) | 1×10^6 | 1×10^{-6} | 300 |

These figures explain why it is much easier to conduct interference experiments with a laser than with a discharge lamp. If the path difference exceeds ℓ_{coh} you will not get interference fringes, because the light is incoherent.

Brightness: The brightness of lasers arises from two factors. First of all, the fact that the light is emitted in a well-defined beam means that the power per unit area is very high, even though the total amount of power can be rather low. Then we must consider that all the energy is concentrated within the narrow spectrum of the active atomic transition. This means that the spectral brightness (i.e. the intensity in the beam divided by the width of the emission line) is even higher in comparison with a white light source like a light bulb. For example, the spectral brightness of a 1 mW laser beam could easily be millions of times greater than that of a 100 W light bulb.

Ultra-short pulse generation: In some cases, lasers can be made to operate in pulses. The time duration of the pulses t_p is linked to the spectral band width of the laser light $\Delta\nu$ by the “uncertainty” product $\Delta t\Delta\nu \sim 1$. This follows from taking the Fourier transform of a pulse of duration t_p . As an example, the bandwidth of the 632.8nm line in the HeNe laser is 1.5 GHz (see above), so that the shortest pulses that a HeNe laser can produce would be 0.67 ns long. This is not particularly short by modern standards. Dye lasers typically have gain bandwidths greater than 10^{13} Hz, and can be used to generate pulses shorter than 100 fs. This is achieved by a technique called “mode-locking”. These short pulsed lasers are very useful for studying fast processes in physics, chemistry and biology.

⁴Recall that the Q -factor is approximately the number of oscillations required for a freely oscillating system’s energy to fall by a factor of $1/e^{2\pi}$ of its original energy.

Arthur Leonard Schawlow 1921-1999

American physicist and corecipient, with Nicolaas Bloembergen of the US and Kai Manne Siegbahn of Sweden, of the 1981 Nobel Prize for Physics for his work in developing the laser and in laser spectroscopy. In 1949 he went to Columbia University, where he began collaborating with Charles Townes on the development of masers, lasers, and laser spectroscopy. Schawlow worked on the project that led to the construction of the first working maser in 1953 (for which Townes received a share of the 1964 Nobel Prize for Physics). Schawlow was a research physicist at Bell Telephone Laboratories from 1951 to 1961. In 1958 he and Townes published a paper in which they outlined the working principles of the laser, though the first such working device was built by another American physicist, Theodore Maiman, in 1960. In 1961 Schawlow became a professor at Stanford University. He became a world authority on laser spectroscopy, and he and Bloembergen earned their share of the 1981 Nobel Prize by using lasers to study the interactions of electromagnetic radiation with matter.



13.3.1 Operating principles of a laser

Light amplification is achieved by stimulated emission. Ordinary optical materials do not amplify light. Instead, they tend to absorb or scatter the light, so that the light intensity out of the medium is less than the intensity that went in. To get amplification you have to drive the material into a non-equilibrium state by pumping energy into it. Positive optical feedback is achieved by inserting the amplifying medium inside a resonant cavity. Light in the cavity passes through the gain medium and is amplified. It then bounces off the end mirrors and passes through the gain medium again, getting amplified further. This process repeats itself until a stable equilibrium condition is achieved when the total round trip gain balances all the losses in the cavity.

The losses in the cavity fall into two categories: useful, and useless. The useful loss comes from the output coupling. One of the mirrors (called the “output coupler”) has reflectivity less than unity, and allows some of the light oscillating around the cavity to be transmitted as the output of the laser. The value of the transmission is chosen to maximise the output power. If the transmission is too low, very little of the light inside the cavity can escape, and thus we get very little output power. On the other hand, if the transmission is too high, there may not be enough gain to sustain oscillation, and there would be no output power. The optimum value is somewhere between these two extremes. Useless losses arise from absorption in the optical components (including the laser medium), scattering, and the imperfect reflectivity of the other mirror (the “high reflector”).

In general we expect the gain to increase as we pump more energy into the laser medium. At low pump powers, the gain will be small, and there will be insufficient gain to reach the oscillation condition. The laser will not start to oscillate until there is enough gain to overcome all the losses. This implies that the laser will have a threshold in terms of the pump power.

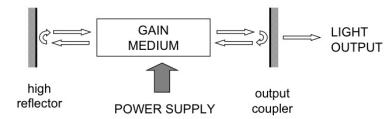
13.3.2 Gain mechanism

Laser operation relies upon the phenomenon of stimulated emission. In a gas of atoms in thermal equilibrium, the population of lower levels will always be greater than the population of upper levels. Therefore, if a light beam is incident on the medium, there will always be more upward transitions due to absorption than downward transitions due to stimulated emission. Hence there will be net absorption, and the intensity of the beam will diminish on progressing through the medium.

To amplify the beam, we require that the rate of stimulated emission exceeds the rate of absorption. If the light beam is sufficiently intense that we can ignore spontaneous emission, and the levels are non-degenerate, this implies that the number of atoms in some upper level, N_2 , must exceed that of the lower level N_1 . This is a highly non-equilibrium situation, and is called **population inversion**. Once we have population inversion, we have a mechanism for generating gain in the laser medium. The art of making a laser operate is to work out how to get population inversion for the relevant transition.

To develop a theory of the laser threshold, we can consider separately the rate equations for the photon and atomic excitation. Starting with photons, let us consider excitations created by the transitions between just two levels of the atom – a lower level 1, and an excited state 2. If the dipole matrix elements, W , between the two levels are independent of position and frequency, the net downwards transition rate is given by

$$W(N_2(n+1) - N_1n)$$



where n denotes the total number of photons in the cavity, and $N_{1,2}$ is the number of atoms in states 1,2. The first term represents the contribution from stimulated and spontaneous emission, while the latter is associated with absorption. Taking into account photon loss from the leaky cavity, the rate of change of photon number is therefore given by

$$\dot{n} = DWn + N_2W - \frac{n}{\tau_{\text{ph}}}, \quad (13.3)$$

where $D = N_2 - N_1$ represents the population imbalance and $1/\tau_{\text{ph}}$ is the photon loss rate. This equation shows that the gain in a laser medium is directly proportional to the degree of population inversion. Laser operation will occur when there is enough gain to overcome the losses in the cavity. This implies that a minimum amount of population inversion must be obtained before the laser will oscillate.

To achieve population inversion atoms must be “pumped” into the upper level 2. This can be achieved by a variety of techniques: Lasers are classified as being either three-level or four-level systems. In the following, we will consider the case of a three-level laser, although four-level lasers are more common. Examples of four-level lasers include Helium Neon or Nd:YAG. In a four-level laser, the levels comprise the ground state (0), the two lasing levels (1 and 2), and a fourth level (3) which is used as part of the pumping mechanism. In the three-level system, such as the first laser, ruby, level 1 is the ground state, and pumping is achieved by exciting atoms to level 3 with a bright flash lamp or by an electrical discharge, and then allowing them to decay rapidly to level 2. In this case, the corresponding rate equations for the populations of levels 1 and 2 can be written as

$$\dot{N}_2 \simeq -w_{21}N_2 + w_{12}N_1 - (N_2 - N_1)Wn \simeq -\dot{N}_1,$$

where w_{12}, w_{21} denote the “effective” transition rates between states 1 and 2 due to the pumping via the third state, and we have dropped the small contribution from spontaneous emission. From this equation, we can deduce that $N_1 + N_2 = N$, a constant, i.e. the decay from state 3 is so rapid that its population is always negligible. In this case, we obtain

$$\dot{D} = \frac{D_0 - D}{T} - 2DWn. \quad (13.4)$$

where $D_0 = N(w_{12} - w_{21})/(w_{12} + w_{21})$ denotes the **unsaturated inversion** (i.e. the degree of population inversion that would exist if there were no photons in the cavity, $n = 0$) and $1/T = w_{12} + w_{21}$.

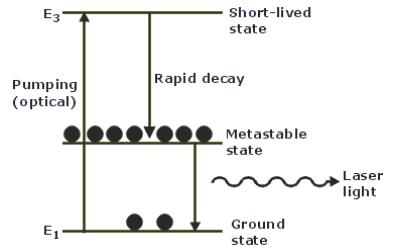
In steady-state, $\dot{n} = \dot{D} = 0$, and Eq. (13.4) translates to a population imbalance,

$$D \equiv N_2 - N_1 = \frac{D_0}{1 + 2TWn}.$$

From this result, we find the steady state photon number is given by

$$n = \frac{D_0W - 1/\tau_{\text{ph}}}{2TW/\tau_{\text{ph}}}.$$

When $D_0 > 1/W\tau_{\text{ph}}$, the **laser threshold condition**, there is a rapid increase in the number of photons in the cavity and the system starts lasing.



Although this analysis addresses the threshold conditions, it does not provide any insight into the coherence properties of the radiation field. In fact, one may show that the radiation field generated by the laser cavity forms a **coherent** state. The properties of this state will be discussed further when we discuss quantum states of light fields.

Chapter 14

Field theory: from phonons to photons

In our survey of quantum mechanics so far, it has been possible to work with a discrete representation in which we index individual constituent particles. However, when the “elementary excitations” of the system involve the coherent *collective* motion of many individual discrete particle degrees of freedom – such as the wave-like atomic vibrations of an ordered elastic solid, or where discrete underlying classical particles can not even be identified – such as the electromagnetic field, such a representation is inconvenient or even inaccessible. In such cases, it is useful to turn to a continuum formulation of quantum mechanics. In the following, we will develop these fundamental ideas on the background of the simplest continuum theory: lattice vibrations of the atomic chain. As we will see, this study will provide a platform to investigate the quantum mechanics of the electromagnetic field – the subject of quantum electrodynamics – which also paves the way to the development of quantum field theory of relativistic particles.

14.1 Quantization of the classical atomic chain

As a simplified model of an ordered (one-dimensional) crystal, let us consider a chain of point particles each of mass m (atoms) which are elastically connected by springs with spring constant k_s (chemical bonds) (see Fig. 14.1). Although our target will be to construct a *quantum* theory of the elementary vibrational excitations, it is helpful to begin our analysis by reviewing the classical properties of the system.

14.1.1 INFO: Classical chain

For reasons that will become clear, it is instructive to consider the **Lagrangian formulation** of the problem. For the N -atom chain, the classical Lagrangian is given by,

$$L = T - V = \sum_{n=1}^N \left[\frac{m}{2} \dot{x}_n^2 - \frac{k_s}{2} (x_{n+1} - x_n - a)^2 \right], \quad (14.1)$$

where the first term accounts for the kinetic energy of the particles whilst the second describes their coupling.¹ For convenience, we adopt periodic boundary

¹In real solids, the inter-atomic potential is, of course, more complex than our quadratic approximation. Yet, for “weak coupling”, the harmonic contribution dominates (cf. our discussion of molecular vibrations). For the sake of simplicity we, therefore, neglect the effects caused by higher order contributions.

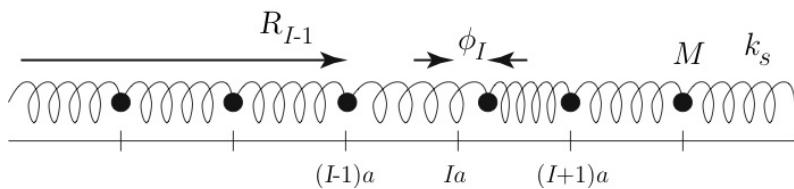


Figure 14.1: Toy model of a one-dimensional solid: a chain of point-like particles each of mass m coupled elastically by springs with spring constant k_s .

conditions such that $x_{N+1} \equiv Na + x_1$ where a denotes the ‘natural’ equilibrium lattice spacing. Anticipating that the effect of lattice vibrations on the solid is weak (i.e. long-range atomic order is maintained) we will assume that (a) the n -th atom has its equilibrium position at $\bar{x}_n \equiv na$, and (b) that the deviation from the equilibrium position remains small ($|x_n(t) - \bar{x}_n| \ll a$), i.e. the integrity of the solid is maintained. With $x_n(t) = \bar{x}_n + \phi_n(t)$ ($\phi_{N+1} = \phi_1$) the Lagrangian (14.1) then takes the form

$$L = \sum_{n=1}^N \left[\frac{m}{2} \dot{\phi}_n^2 - \frac{k_s}{2} (\phi_{n+1} - \phi_n)^2 \right].$$

Now, typically, we are not concerned with the behaviour of a given system on ‘atomic’ length scales. (For such purposes, our model is in any case much too primitive!) Rather, we are interested in **universal** features, i.e. experimentally observable behaviour, common to a wide range of physical systems, that manifests itself on macroscopic length scales where the detailed form of the model is inessential. For example, we might wish to study the specific heat of the solid in the limit of infinitely many atoms (or at least a macroscopically large number, $\mathcal{O}(10^{23})$). Under these conditions, microscopic models can usually be substantially simplified. In particular it is often permissible to subject a discrete lattice model to a **continuum approximation**, i.e. to neglect the discreteness of the microscopic entities of the system and to describe it in terms of effective continuum degrees of freedom. In the present case, taking a continuum limit amounts to describing the lattice displacements ϕ_n in terms of *smooth functions*, $\phi(x)$ of a continuous variable x (see figure). Clearly such a description makes sense only if relative fluctuations on atomic scales are weak (for otherwise the smoothness condition would be violated).

▷ EXERCISE. Starting with the discrete form of the Lagrangian, or otherwise, show the classical equations of motion take the form,

$$m\ddot{\phi}_n = k_s a^2 (\phi_{n+1} - 2\phi_n + \phi_{n-1}).$$

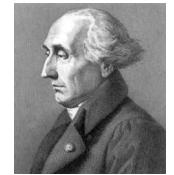
Remembering that the boundary conditions are periodic, obtain the normal modes. From this result, determine the condition under which the continuum approximation can be justified.

Introducing continuum degrees of freedom, $\phi(x)$, and applying a first order Taylor expansion,² we can define

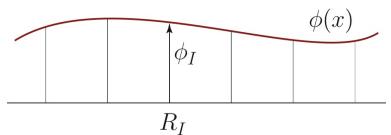
$$\phi_n \rightarrow \phi(x) \Big|_{x=na}, \quad \phi_{n+1} - \phi_n \rightarrow a \partial_x \phi(x) \Big|_{x=na}, \quad \sum_{n=1}^N \rightarrow \frac{1}{a} \int_0^L dx,$$

Joseph-Louis Lagrange 1736-1813:

A mathematician who excelled in all fields of analysis, number theory, and celestial mechanics.



In 1788 he published *Mécanique Analytique*, which summarised all of the work done in the field of mechanics since the time of Newton, and is notable for its use of the theory of differential equations. In it he transformed mechanics into a branch of mathematical analysis.



Hint: consider the ansatz, $\phi_n(t) = e^{i(kna - \omega t)}$.

²Indeed, for reasons that will become clear, higher order contributions to the Taylor expansion do not contribute to the low-energy properties of the system where the continuum approximation is valid.

where $L = Na$ (not to be confused with the Lagrangian itself!) denotes the total length of the chain. Expressed in terms of the new degrees of freedom, the continuum limit of the Lagrangian then reads $L[\phi] = \int_0^L dx \mathcal{L}(\dot{\phi}, \phi)$, where

$$\mathcal{L}(\dot{\phi}, \phi) = \frac{\rho}{2} \dot{\phi}^2 - \frac{\kappa_s a^2}{2} (\partial_x \phi)^2, \quad (14.2)$$

denotes the Lagrangian density, $\rho = m/a$ denotes the mass per unit length and $\kappa_s = k_s/a$. The corresponding **classical action** is given by

$$S[\phi] = \int dt L[\phi]. \quad (14.3)$$

Thus, we have succeeded in trading the N -point particle description in for one involving *continuous* degrees of freedom, $\phi(x)$, a (**classical**) **field**. The dynamics of the latter are specified by ‘functionals’ $L[\phi]$ and $S[\phi]$ which represent the continuum generalizations of the discrete classical Lagrangian and action, respectively.³ However, although we have achieved a continuum formulation, we have yet to extract concrete physical information from the action. To do so, we need to derive equations of motion. At first sight, it may not be entirely clear what is meant by ‘equations of motion’ in the context of an infinite dimensional model. The answer to this question lies in Hamilton’s extremal principle of classical mechanics:

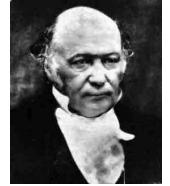
▷ INFO. **Hamilton’s extremal principle:** Suppose that the dynamics of a classical *point* particle with coordinate $x(t)$ is described by the classical Lagrangian $L(x, \dot{x})$, and action $S[x] = \int dt L(x, \dot{x})$. Hamilton’s extremal principle states that the configurations $x(t)$ that are actually realized are those that extremize the action. This means that, for any smooth curve $y(t)$,

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (S[x + \epsilon y] - S[x]) = 0, \quad (14.4)$$

i.e. to first order in ϵ , the action has to remain invariant. Applying this condition, one finds that it is fulfilled if and only if $x(t)$ obeys the **Euler-Lagrange equation** of motion (exercise),

$$\frac{d}{dt} (\partial_{\dot{x}} L) - \partial_x L = 0. \quad (14.5)$$

Sir William Rowan Hamilton
1805-1865:
A mathematician credited with the discovery of quaternions, the first non-commutative algebra to be studied. He also invented important new methods in Mechanics.



Now, in Eq. (14.3), we are dealing with a system of infinitely many degrees of freedom, $\phi(x, t)$. Yet Hamilton’s principle is general, and we may see what happens if (14.3) is subjected to an extremal principle analogous to Eq. (14.4). To do so, we must effect the substitution $\phi(x, t) \rightarrow \phi(x, t) + \epsilon \eta(x, t)$ into Eq. (14.3) and demand that the contribution first order in ϵ vanishes. When applied to the specific Lagrangian (14.2), a substitution of the ‘varied’ field leads to

$$S[\phi + \epsilon \eta] = S[\phi] + \epsilon \int dt \int_0^L dx \left(\rho \dot{\phi} \dot{\eta} - \kappa_s a^2 \partial_x \phi \partial_x \eta \right) + \mathcal{O}(\epsilon^2).$$

Integrating by parts (with respect to time for the first term under the integral, and space in the second) and demanding that the contribution linear

³In the mathematics and physics literature, mappings of functions into the real or complex numbers are generally called **functionals**. The argument of a functional is commonly indicated in rectangular brackets $[\cdot]$. For example, in this case, S maps the ‘functions’ $\partial_x \phi(x, t)$ and $\dot{\phi}(x, t)$ to the real number $S[\phi]$.

in ϵ vanishes, one obtains $\int dt \int_0^L dx (\rho \ddot{\phi} - \kappa_s a^2 \partial_x^2 \phi) \eta = 0$. (Notice that the boundary terms associated with both t and x vanish identically.⁴ Now, since η was defined to be any arbitrary smooth function, the integral above can only vanish if the term in parentheses is globally vanishing. Thus the equation of motion takes the form of a **wave equation**,

$$\rho \ddot{\phi} = \kappa_s a^2 \partial_x^2 \phi. \quad (14.6)$$

The solutions of Eq. (14.6) have the general form $\phi_+(x + vt) + \phi_-(x - vt)$ where $v = a\sqrt{\kappa_s/\rho}$, and ϕ_{\pm} are arbitrary smooth functions of their argument. From this we can deduce that the basic low energy **elementary excitations** of our model are lattice vibrations propagating as **sound waves** to the left or right at a constant velocity v (see figure). The trivial behaviour of our model is of course a direct consequence of its simplistic definition — no dissipation, dispersion or other non-trivial ingredients. Adding these refinements leads to the general classical theory of lattice vibrations. With this background, let us now turn to the consider the quantization of the quantum mechanical chain.



14.1.2 INFO: Quantum chain

In addressing the quantum description, the first question to ask is a conceptual one: is there a general methodology to quantize models of the form described by the atomic chain (14.2)? Indeed, there is a standard procedure to quantize continuum theories which closely resembles the quantization of point mechanics. The first step is to introduce canonical momenta conjugate to the continuum degrees of freedom (coordinates), ϕ , which will later be used to introduce canonical commutation relations. The natural generalization of the definition $p_n \equiv \partial_{\dot{x}_n} L$ of point mechanics to a continuum suggests setting

$$\pi = \partial_{\dot{\phi}} \mathcal{L}(\dot{\phi}, \phi). \quad (14.7)$$

In common with $\phi(x, t)$, the **canonical momentum**, $\pi(x, t)$, is a continuum degree of freedom. At each space point it may take an independent value. From the Lagrangian, we can define the Hamiltonian, $H[\phi, \pi] \equiv \int dx \mathcal{H}(\phi, \pi)$, where $\mathcal{H}(\phi, \pi) \equiv \pi \dot{\phi} - \mathcal{L}(\dot{\phi}, \phi)$. represents the **Hamiltonian density**. Applied to the atomic chain (14.2), the canonical momentum $\pi = \rho \dot{\phi}$ and $\mathcal{H}(\phi, \pi) = \frac{\pi^2}{2\rho} + \frac{\kappa_s a^2}{2} (\partial_x \phi)^2$.

In this form, the Hamiltonian can be quantized according to the following rules: (a) promote the fields $\phi(x)$ and $\pi(x)$ to operators: $\phi \mapsto \hat{\phi}$, $\pi \mapsto \hat{\pi}$, and (b) generalise the canonical commutation relations of one-particle quantum mechanics, $[\hat{p}_m, x_n] = -i\hbar \delta_{mn}$, according to the relation⁵

$$[\hat{\pi}(x), \hat{\phi}(x')] = -i\hbar \delta(x - x'). \quad (14.8)$$

Operator-valued functions like $\hat{\phi}$ and $\hat{\pi}$ are generally referred to as **quantum fields**. Employing these definitions, we obtain the quantum Hamiltonian density

$$\hat{\mathcal{H}}(\hat{\phi}, \hat{\pi}) = \frac{1}{2\rho} \hat{\pi}^2 + \frac{\kappa_s a^2}{2} (\partial_x \hat{\phi})^2.$$

⁴If we assume that the function ϕ already obeys the boundary conditions, we must have $\eta(0, t) = \eta(L, t) = \eta(x, 0) = \eta(x, T) = 0$.

⁵Note that the dimensionality of both the quantum and classical continuum fields is compatible with the dimensionality of the Dirac δ -function, $[\delta(x - x')] = [\text{Length}]^{-1}$.

▷ EXERCISE. To develop this field theoretical formulation of the Hamiltonian, we have pursued a Lagrangian formulation. If you feel uncertain about this methodology, you should explore the derivation of \hat{H} directly from the discrete atomic formulation. First, show that, for the discrete harmonic chain, the classical Hamiltonian is given by

$$\hat{H} = \sum_n \left[\frac{p_n^2}{2m} + \frac{k_s}{2} (\phi_{n+1} - \phi_n)^2 \right].$$

Promoting the displacements and momenta to operators, and applying the canonical quantization conditions, $[\hat{p}_n, \phi_{n'}] = -i\hbar\delta_{nn'}$, obtain the discrete form of the Hamiltonian. Taking the continuum limit, show that the Hamiltonian recovers the continuum form derived through the Lagrangian formulation.

The Hamiltonian represents a quantum field theoretical *formulation* of the problem but not yet a *solution*. To address the quantum properties of the system, it is helpful now to switch to a Fourier representation. As with any function, operator-valued functions can be represented in a variety of different ways. In particular they can be subjected to Fourier expansion,

$$\begin{cases} \hat{\phi}_k \equiv \frac{1}{L^{1/2}} \int_0^L dx e^{\{\mp ikx\}} \begin{cases} \hat{\phi}(x) \\ \hat{\pi}(x) \end{cases}, \\ \hat{\pi}_k \end{cases} \quad \begin{cases} \hat{\phi}(x) \\ \hat{\pi}(x) \end{cases} = \frac{1}{L^{1/2}} \sum_k e^{\{\pm ikx\}} \begin{cases} \hat{\phi}_k \\ \hat{\pi}_k \end{cases}. \quad (14.9)$$

where \sum_k represents the sum over all Fourier coefficients indexed by quantized wavevectors $k = 2\pi m/L$, m integer. Note that, since the classical field $\phi(x)$ is *real*, the quantum field $\hat{\phi}(x)$ is *Hermitian*, i.e. $\hat{\phi}_k = \hat{\phi}_{-k}^\dagger$ (and similarly for $\hat{\pi}_k$). In the Fourier representation, the transformed field operators obey the canonical commutation relations (exercise),

$$[\hat{\pi}_k, \hat{\phi}_{k'}] = -i\hbar\delta_{kk'}.$$

▷ EXERCISE. Making use of Eqs. (14.8) and (14.9) derive the canonical commutation relation above.

When expressed in the Fourier representation, making use of the identity

$$\int_0^L dx (\partial\hat{\phi})^2 = \sum_{k,k'} (ik\hat{\phi}_k)(ik'\hat{\phi}_{k'}) \underbrace{\frac{1}{L} \int_0^L dx e^{i(k+k')x}}_{\delta_{k+k',0}} = \sum_k k^2 \hat{\phi}_k \hat{\phi}_{-k},$$

together with the parallel relation for $\int_0^L dx \hat{\pi}^2$, the Hamiltonian assumes the “near diagonal” form,

$$\hat{H} = \sum_k \left[\frac{1}{2\rho} \hat{\pi}_k \hat{\pi}_{-k} + \frac{1}{2} \rho \omega_k^2 \hat{\phi}_k \hat{\phi}_{-k} \right], \quad (14.10)$$

where $\omega_k = v|k|$, and $v = a(\kappa_s/\rho)^{1/2}$ denotes the classical sound wave velocity. In this form, the Hamiltonian can be identified as nothing more than a superposition of independent quantum harmonic oscillators. The only difference between (14.10) and the canonical form of an oscillator Hamiltonian $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2$ is the presence of sub-indices k and $-k$ (a consequence of $\hat{\phi}_k^\dagger = \hat{\phi}_{-k}$). As we will show shortly, this difference is inessential. This

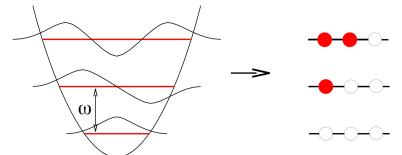
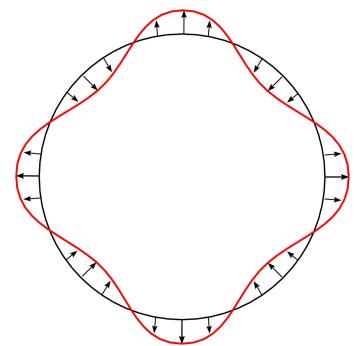
result is actually not difficult to understand (see figure): Classically, the system supports a discrete set of wave-like excitations, each indexed by a wave number $k = 2\pi m/L$. Within the quantum picture, each of these excitations is described by an oscillator Hamiltonian with a k -dependent frequency. However, it is important not to confuse the atomic constituents, also oscillators (albeit they coupled), with the independent *collective* oscillator modes described by \hat{H} .

The description above, albeit perfectly valid, still suffers from a deficiency: Our analysis amounts to explicitly describing the effective low energy excitations of the system (the waves) in terms of their microscopic constituents (the atoms). Indeed the different contributions to \hat{H} keep track of details of the microscopic oscillator dynamics of individual k -modes. However, it would be much more desirable to develop a picture where the relevant excitations of the system, the waves, appear as fundamental units, without explicit account of underlying microscopic details. (As with hydrodynamics, information is encoded in terms of collective density variables rather than through individual molecules.) To understand how this programme can be achieved let us recall the properties of the quantum harmonic oscillator.

▷ INFO. In quantum mechanics, the **harmonic oscillator** has the status of a single-particle problem. However, the fact that the energy levels, $\epsilon_n = \hbar\omega(n + 1/2)$, are *equidistant* suggests an alternative interpretation: One can think of a given energy state ϵ_n as an accumulation of n elementary entities, or **quasi-particles**, each having energy $\hbar\omega$. What can be said about the features of these new objects? First, they are structureless, i.e. the only ‘quantum number’ identifying the quasi-particles is their energy $\hbar\omega$ (otherwise n -particle states formed of the quasi-particles would not be equidistant). This implies that the quasi-particles must be *bosons*. (The same state $\hbar\omega$ can be occupied by more than one particle — see figure.) This idea can be formulated in quantitative terms by employing the formalism of ladder operators in which the operators \hat{p} and \hat{x} are traded for the pair of Hermitian adjoint operators $a \equiv \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} + \frac{i}{m\omega}\hat{p})$, $a^\dagger \equiv \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} - \frac{i}{m\omega}\hat{p})$. Up to a factor of i , the transformation $(\hat{x}, \hat{p}) \rightarrow (a, a^\dagger)$ is canonical, i.e. the new operators obey the canonical commutation relation, $[a, a^\dagger] = 1$. More importantly, in the a -representation, the Hamiltonian takes the simple form, $\hat{H} = \hbar\omega(a^\dagger a + 1/2)$, as can be checked by direct substitution. The complete hierarchy of higher energy states can be generated by setting $|n\rangle \equiv \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle$.

While the a -representation provides another way of constructing eigenstates of the quantum harmonic oscillator, its real advantage is that it naturally affords a many-particle interpretation. Temporarily forgetting about the original definition of the oscillator, we can *declare* $|0\rangle$ to be a ‘vacuum’ state, i.e. a state with no particles present. $a^\dagger|0\rangle$ then represents a state with a single featureless particle (the operator a^\dagger does not carry any quantum number labels) of energy $\hbar\omega$. Similarly, $(a^\dagger)^n|0\rangle$ is considered as a many-body state with n particles, i.e. within the new picture, a^\dagger is an operator that creates particles. The total energy of these states is given by $\hbar\omega \times$ (occupation number). Indeed, it is straightforward to verify that $a^\dagger a|n\rangle = n|n\rangle$, i.e. the Hamiltonian basically counts the number of particles. While, at first sight, this may look unfamiliar, the new interpretation is internally consistent. Moreover, it fulfills our objective: it allows an interpretation of the excited states of the harmonic oscillator as a superposition of independent structureless entities.

With this background, we may return to the harmonic atomic chain (14.10)



and, inspired by the ladder operator formalism, define⁶

$$a_k \equiv \sqrt{\frac{m\omega_k}{2\hbar}} \left(\hat{\phi}_k + \frac{i}{m\omega_k} \hat{\pi}_{-k} \right), \quad a_k^\dagger \equiv \sqrt{\frac{m\omega_k}{2\hbar}} \left(\hat{\phi}_{-k} - \frac{i}{m\omega_k} \hat{\pi}_k \right).$$

With this definition, one finds that the ladder operators obey the commutation relations (exercise)

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad [a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0, \quad (14.11)$$

and the Hamiltonian assumes the diagonal form

$$\hat{H} = \sum_k \hbar\omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right). \quad (14.12)$$

Eqs. (14.11) and (14.12) represent the final result of our analysis: The low-lying elementary excitations of the discrete atomic chain are described by oscillator wave-like modes – known as **phonons** – each characterised by a wavevector k and a linear dispersion, $\omega_k = v|k|$. A generic state of the system is given by

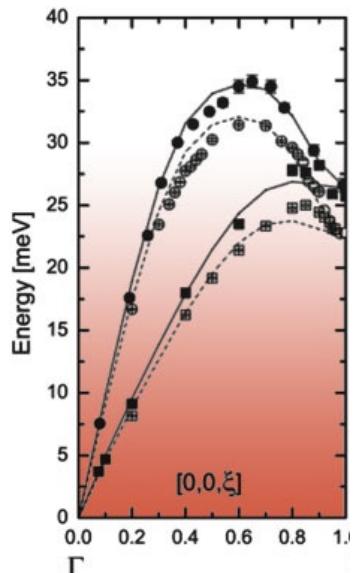
$$|\{n_k\} = (n_1, n_2, \dots)\rangle = \frac{1}{\sqrt{\prod_i n_i!}} (a_{k_1}^\dagger)^{n_1} (a_{k_2}^\dagger)^{n_2} \dots |0\rangle.$$

The representation derived above illustrates the capacity to think about quantum problems in different complementary “pictures”, a principle that finds innumerable applications. The existence of different interpretations of a given system is by no means heretic but, rather, is consistent with the spirit of quantum mechanics. Indeed, it is one of the prime principles of quantum theories that there is no such thing as ‘the real system’ which underpins the phenomenology. The only thing that matters is observable phenomena. For example, the ‘fictitious’ quasi-particle states of the harmonic chain, the phonons, *behave* as ‘real’ particles, i.e. they have dynamics, can interact, be detected experimentally, etc. From a quantum point of view there is actually no fundamental difference between these objects and ‘real’ particles.

This completes our discussion of the classical and quantum field theory of the harmonic atomic chain. In this example, we have seen how we can effect a quantum formulation of a continuum system. Using the insights obtained in this example, we now turn to consider the quantization of the electromagnetic field.

14.2 Quantum electrodynamics

In common with the continuous formulation of the atomic chain, in vacua, the electromagnetic (EM) field satisfies a wave equation. The generality of the procedure outlined above suggests that the quantization of the EM field might proceed in an entirely analogous manner. However, there are a number of practical differences that make quantization a slightly more difficult enterprise: Firstly, the vector character of the vector potential \mathbf{A} , alongside relativistic covariance, gives the problem a non-trivial internal geometry. Moreover, the gauge freedom of the vector potential introduces redundant degrees of freedom whose removal on the quantum level is not straightforward. To circumvent a



The figure shows a typical measured phonon dispersion of an ordered crystalline solid obtained by neutron scattering. The x-axis indexes wavenumbers along a lattice direction (specified in units of π/a). Three generic aspects are visible: (1) near $k = 0$, the dispersion is, as expected, linear. (2) The several branches are associated with different “polarizations” of the lattice fluctuations. (3) For wavelengths comparable to the lattice spacing, $k \sim \pi/a$, non-universal features specific to the particular material become visible.

⁶As for the consistency of these definitions, recall that $\hat{\phi}_k^\dagger = \hat{\phi}_{-k}$ and $\hat{\pi}_k^\dagger = \hat{\pi}_{-k}$. Under these conditions the second of the definitions below indeed follows from the first upon taking the Hermitian conjugate.

lengthy discussion of these issues, we will not address the problem of EM field quantization in all its detail.⁷ On the other hand, the photon field plays a much too important role in all branches of physics for us to drop the problem altogether. We will therefore aim at an intermediate exposition, largely insensitive to the problems outlined above but sufficiently general to illustrate the main principles. As with the harmonic chain, to prepare the way, we begin by developing the classical field theory of the EM field.

14.2.1 INFO: Classical theory of the electromagnetic field

In vacuum, the Lagrangian density of the EM field is given by $\mathcal{L} = -\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu}$ (summation convention implied) where $\mu_0 = 4\pi \times 10^{-7} \text{ Hm}^{-1}$ denotes the **vacuum permeability**,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix}_{\mu\nu}$$

denotes the EM field tensor, $\mathbf{E} = \dot{\mathbf{A}}$ is the electric field, and $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field. As a first step towards quantization, we must specify a gauge. In the absence of charge, a particularly convenient choice is the **Coulomb gauge**, $\nabla \cdot \mathbf{A} = 0$, with the scalar component $\phi = 0$.⁸ Using these gauge conditions, one may verify that the classical Lagrangian assumes the form,

$$L[\mathbf{A}(\mathbf{x}, t)] = \frac{1}{2\mu_0} \int d^3x \left[\frac{1}{c^2} \dot{\mathbf{A}}^2 - (\nabla \times \mathbf{A})^2 \right]. \quad (14.13)$$

The corresponding classical Euler-Lagrange equations of motion, $\partial_\mu F^{\mu\nu} = 0$, translate to the wave equation (exercise)

$$\frac{1}{c^2} \ddot{\mathbf{A}} = \nabla^2 \mathbf{A}.$$

The structural similarity between the EM field and the continuous formulation of the harmonic chain is clear. By analogy with our discussion above, we should now switch to the Fourier representation and quantize the classical field. However, in contrast to our analysis of the chain, we are now dealing (i) with the full three-dimensional Laplacian acting upon (ii) the vector field \mathbf{A} that is (iii) subject to the constraint $\nabla \cdot \mathbf{A} = 0$. It is these aspects which lead to the complications outlined above.

We can circumvent these difficulties by considering cases where the geometry of the system reduces the complexity of the eigenvalue problem while still retaining the key conceptual aspects of the problem. This restriction is less artificial than it might appear. For example, just as the field ϕ in the classical atomic chain can be expanded in Fourier harmonics, in long waveguides, the EM vector potential can be expanded in solutions of the eigenvalue equation⁹

$$-\nabla^2 \mathbf{u}_k(\mathbf{x}) = \lambda_k \mathbf{u}_k(\mathbf{x}), \quad (14.14)$$

⁷Readers interested in a more thorough and illuminating exposition are referred to the literature, e.g., L. H. Ryder, *Quantum Field Theory* (Cambridge University Press, 1996), or the excellent lecture notes of Eduardo Fradkin that have been made available online at <http://webusers.physics.illinois.edu/~efradkin/phys582/physics582.html>.

⁸Keep in mind that, once a gauge has been set, we cannot expect further results to display “gauge invariance.”

⁹More precisely, one should say that Eq. (14.14) defines the set of eigenfunctions relevant for the *low-energy* dynamics of the waveguide. More complex eigenfunctions of the Laplace operator exist but they involve much higher energy.

where k denotes a discrete *one-dimensional* index, and the vector-valued functions \mathbf{u}_k are real and orthonormalized, $\int d^3x \mathbf{u}_k \cdot \mathbf{u}_{k'} = \delta_{kk'}$. The dependence of the eigenvalues λ_k on k depends on details of the geometry and need not be specified for the moment.

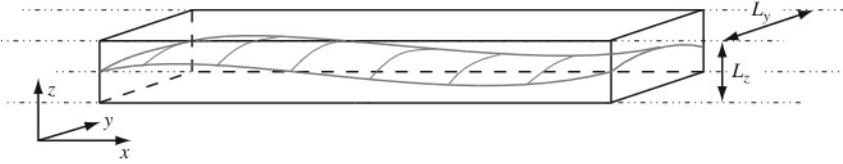


Figure 14.2: EM waveguide with rectangular cross-section. The structure of the eigenmodes of the EM field is determined by boundary conditions at the walls of the cavity.

▷ INFO. An **electrodynamic waveguide** is a quasi one-dimensional cavity with metallic boundaries (see Fig. 14.2). The practical advantage of waveguides is that they are good at confining EM waves. At large frequencies, where the wavelengths are of order meters or less, radiation loss in conventional conductors is high. In these frequency domains, hollow conductors provide the only practical way of transmitting radiation. EM field propagation inside a waveguide is constrained by boundary conditions. Assuming the walls of the system to be perfectly conducting,

$$\mathbf{E}_{\parallel}(\mathbf{x}_b) = 0, \quad \mathbf{B}_{\perp}(\mathbf{x}_b) = 0, \quad (14.15)$$

where \mathbf{x}_b parameterize points on the boundary of the system, and \mathbf{E}_{\parallel} (\mathbf{B}_{\perp}) is the parallel (perpendicular) component of the electric (magnetic) field. Applied to the problem at hand, let us consider a long cavity with uniform rectangular cross-section $L_y \times L_z$. To conveniently represent the Lagrangian of the system, we wish to express the vector potential in terms of eigenfunctions \mathbf{u}_k that are consistent with the boundary conditions (14.15). A complete set of functions fulfilling this condition is given by

$$\mathbf{u}_k = \mathcal{N}_k \begin{pmatrix} c_1 \cos(k_x x) \sin(k_y y) \sin(k_z z) \\ c_2 \sin(k_x x) \cos(k_y y) \sin(k_z z) \\ c_3 \sin(k_x x) \sin(k_y y) \cos(k_z z) \end{pmatrix}.$$

Here $k_i = n_i \pi / L_i$, with $i = x, y, z$ and n_i is integer, \mathcal{N}_k is a factor normalizing \mathbf{u}_k to unit modulus, and the coefficients c_i are subject to the condition $c_1 k_x + c_2 k_y + c_3 k_z = 0$ (reflecting the gauge choice $\nabla \cdot \mathbf{A} = 0$). Indeed, it is straightforward to verify that a general superposition of the type $\mathbf{A}(\mathbf{x}, t) \equiv \sum_k \alpha_k(t) \mathbf{u}_k(\mathbf{x})$, $\alpha_k(t) \in \mathbb{R}$, is divergenceless, and generates an EM field compatible with (14.15). Substitution of \mathbf{u}_k into (14.14) identifies the eigenvalues as $\lambda_k = k_x^2 + k_y^2 + k_z^2$.

In the physics and electronic engineering literature, eigenfunctions of the Laplace operator in a quasi-one-dimensional geometry are commonly described as **modes**. As we will see shortly, the energy of a mode (i.e. the Hamiltonian evaluated on a specific mode configuration) grows with λ_k . In cases where one is interested in the low-energy dynamics of the EM field, only configurations with small λ_k are relevant. For example, let us consider a massively anisotropic waveguide with $L_z < L_y \ll L_x$. In this case the modes with smallest λ_k are those with $k_z = 0$, $k_y = \pi / L_y$, and $k_x \equiv k \ll L_{z,y}^{-1}$. With this choice,

$$\mathbf{u}_k = \frac{2}{\sqrt{V}} \sin(\pi y / L_y) \sin(kx) \hat{\mathbf{e}}_z, \quad \lambda_k = k^2 + \left(\frac{\pi}{L_y} \right)^2, \quad (14.16)$$

and a scalar index k suffices to label both eigenvalues and eigenfunctions \mathbf{u}_k . A caricature of the spatial structure of the functions \mathbf{u}_k is shown in Fig. 14.2.

Returning to the problem posed by (14.13) and (14.14), one can expand the vector potential in terms of eigenfunctions \mathbf{u}_k as $\mathbf{A}(\mathbf{x}, t) = \sum_k \alpha_k(t) \mathbf{u}_k(\mathbf{x})$,

where the sum runs over all allowed values of the index parameter k . (In a waveguide of length L , $k = \pi n/L$ with n integer.) Substituting this expansion into (14.13), and using the normalization properties of \mathbf{u}_k , we obtain the Lagrangian,

$$L[\dot{\alpha}, \alpha] = \frac{1}{2\mu_0} \sum_k \left[\frac{1}{c^2} \dot{\alpha}_k^2 - \lambda_k \alpha_k^2 \right],$$

i.e. a decoupled representation where the system is described in terms of independent dynamical systems with coordinates α_k . From this point on, the quantization procedure mirrors that of the atomic chain.

14.2.2 Quantum field theory of the electromagnetic field

To achieve the electromagnetic field quantization, we first define the canonical momenta through the relation,

$$\pi_k = \partial_{\dot{\alpha}_k} \mathcal{L} = \epsilon_0 \dot{\alpha}_k,$$

where $\epsilon_0 = 1/\mu_0 c^2$ denotes the **vacuum permittivity**, which leads to the classical Hamiltonian $H = \sum_k (\frac{1}{2\epsilon_0} \pi_k^2 + \frac{1}{2} \epsilon_0 c^2 \lambda_k \alpha_k^2)$. Next we quantize the theory by promoting fields to operators $\alpha_k \rightarrow \hat{\alpha}_k$ and $\pi_k \rightarrow \hat{\pi}_k$, and declare the canonical commutation relations $[\hat{\pi}_k, \hat{\alpha}_{k'}] = -i\hbar \delta_{kk'}$. The quantum Hamiltonian operator, again of harmonic oscillator type, then reads

$$\hat{H} = \sum_k \left[\frac{\hat{\pi}_k^2}{2\epsilon_0} + \frac{1}{2} \epsilon_0 \omega_k^2 \hat{\alpha}_k^2 \right],$$

where $\omega_k^2 = c^2 \lambda_k$.

Then, guided by the analysis of the atomic chain, we now introduce the ladder operators,

$$a_k = \sqrt{\frac{\epsilon_0 \omega_k}{2\hbar}} \left(\hat{\alpha}_k + \frac{i}{\epsilon_0 \omega_k} \hat{\pi}_k \right), \quad a_k^\dagger = \sqrt{\frac{\epsilon_0 \omega_k}{2\hbar}} \left(\hat{\alpha}_k - \frac{i}{\epsilon_0 \omega_k} \hat{\pi}_k \right),$$

whereupon the Hamiltonian assumes the now familiar form

$$\hat{H} = \sum_k \hbar \omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right).$$

(14.17)

For the specific problem of the first excited mode in a waveguide of width L_y , $\hbar \omega_k = c[k^2 + (\pi/L_y)^2]^{1/2}$. Eq. (14.17) represents our final result for the quantum Hamiltonian of the EM waveguide. Physically, the quantum excitations described by this Hamiltonian are the **photons** of the EM field. The unfamiliar appearance of the dispersion ω_k is a peculiarity of the waveguide. In the limit of large longitudinal wave numbers $k \gg L_y^{-1}$, the dispersion approaches $\omega_k \simeq c|k|$, i.e. the familiar linear (relativistic) dispersion of the photon field.

With the analysis of the waveguide complete, let us go back and consider the quantization of the full three-dimensional system. For the waveguide, we have found that the vector potential can be expanded in modes of the cavity as $\hat{\mathbf{A}}(\mathbf{x}) = \sum_k \hat{\alpha}_k \mathbf{u}_k$ where, rearranging the expressions for the ladder operators, $\hat{\alpha}_k = \sqrt{\frac{\hbar}{2\epsilon_0 \omega_k}} (a_k + a_k^\dagger)$. More generally, in a fully three-dimensional cavity, one may show that¹⁰

$$\hat{\mathbf{A}}(\mathbf{x}) = \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar}{2\epsilon_0 \omega_{\mathbf{k}} V}} \left[\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}} \right],$$

¹⁰In the infinite system, the mode sum becomes replaced by an integral, $\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3 k$.

where V denotes the volume of the system, $\omega_{\mathbf{k}} = c|\mathbf{k}|$, and the two sets of polarization vectors, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$, are in general complex and normalized to unity, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda} = 1$. To ensure that the vector potential satisfies the Coulomb gauge condition, we require that $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{k} = \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \mathbf{k} = 0$, i.e. the two polarization vectors are orthogonal to the wave vector. Two real vectors, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$ correspond to two linear polarizations while, for circular polarization, the vectors are complex. It is also convenient to assume that the two polarization vectors are mutually orthogonal, $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}\mu} = \delta_{\mu\nu}$. The corresponding operators obey the commutation relations,

$$[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda\lambda'}, \quad [a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}] = 0 = [a_{\mathbf{k}\lambda}^\dagger, a_{\mathbf{k}'\lambda'}^\dagger].$$

With these definitions, the Hamiltonian then takes the familiar form

$$\hat{H} = \sum_{\mathbf{k}\lambda} \hbar\omega_{\mathbf{k}} \left[a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + 1/2 \right], \quad (14.18)$$

while, defining the vacuum, $|\Omega\rangle$, the eigenstates involve photon number states,

$$|\{n_{\mathbf{k}\lambda}\}\rangle \equiv |\dots, n_{\mathbf{k}\lambda}, \dots\rangle = \prod_{\mathbf{k}\lambda} \frac{(a_{\mathbf{k}\lambda}^\dagger)^{n_{\mathbf{k}\lambda}}}{\sqrt{n_{\mathbf{k}\lambda}!}} |\Omega\rangle. \quad (14.19)$$

Finally, in practical applications (including our study of radiative transitions in atoms), it is convenient to transfer the time-dependence to the operators by turning to the Heisenberg representation. In this representation, the field operators obey the Heisenberg equations of motion (exercise),

$$\dot{a}_{\mathbf{k}\lambda} = \frac{i}{\hbar} [\hat{H}, a_{\mathbf{k}\lambda}] = -i\omega_{\mathbf{k}} a_{\mathbf{k}\lambda}.$$

Integrating, we have $a_{\mathbf{k}\lambda}(t) = a_{\mathbf{k}\lambda}(0)e^{-i\omega_{\mathbf{k}}t}$, which translates to the relation

$$\hat{\mathbf{A}}(\mathbf{x}, t) = \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \left[\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} \right].$$

From the vector potential, we can obtain the electric field and magnetic fields, via $\mathbf{E} = -\dot{\mathbf{A}}$ and $\mathbf{B} = \nabla \times \mathbf{A}$, noting that we have chosen a gauge in which the scalar potential vanished, $\phi = 0$. In terms of the field operators, these are

$$\begin{aligned} \hat{\mathbf{E}}(\mathbf{x}, t) &= i \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 V}} \left[\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} - \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} \right], \\ \hat{\mathbf{B}}(\mathbf{x}, t) &= i \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar\omega_{\mathbf{k}}\mu_0}{2V}} \left[\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} - \hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} \right]. \end{aligned}$$

14.2.3 Fock states

The energy eigenstates of the EM field (14.19) are referred to as “number” or “Fock” states. For these states, the expectation value of the electric and magnetic fields both vanish.

$$\langle \hat{\mathbf{E}} \rangle = \langle \hat{\mathbf{B}} \rangle = 0.$$

These results follow from the facts that: (i) $\hat{\mathbf{A}}$ is linear in the operators \hat{a} and \hat{a}^\dagger , and therefore so too are $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$; and (ii) for any Fock state $|n\rangle$,

$$\langle n|\hat{a}|n\rangle = \langle n|\hat{a}^\dagger|n\rangle = 0.$$

The expectation values of the *squares* of the electric and magnetic fields do not vanish for Fock states. Direct calculations show that for the Fock state (14.19) the mean square electric and magnetic fields at a point \mathbf{r} are

$$\begin{aligned}\langle |\hat{\mathbf{E}}(\mathbf{r})|^2 \rangle &= \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_{\mathbf{k}}}{\epsilon_0 V} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right) \\ \langle |\hat{\mathbf{B}}(\mathbf{r})|^2 \rangle &= \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_{\mathbf{k}}\mu_0}{V} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right).\end{aligned}$$

These results are consistent with the expected total energy in the EM field of these Fock states

$$\langle U \rangle \equiv \int \left[\frac{\epsilon_0}{2} \langle |\mathbf{E}(\mathbf{r})|^2 \rangle + \frac{1}{2\mu_0} \langle |\mathbf{B}(\mathbf{r})|^2 \rangle \right] d^3\mathbf{r} = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right).$$

Note that the vacuum state, with all $\{n_{\mathbf{k},\lambda} = 0\}$, has a non-zero energy density. This is a manifestation of the zero-point motion of the EM fields. The vacuum energy density is a (formally divergent) background energy density predicted by the theory. These zero-point fluctuations have a number of observable physical consequences. Indeed, the spontaneous emission described in §13.1.2 can be viewed as arising from zero-point motion of the fields. Another dramatic consequence is given by the **Casimir effect**. In its simplest setting, this refers to the force between two parallel perfectly conducting planes that arises from the change in the zero-point energy due to the boundary conditions on the plates.

14.2.4 Coherent states

Coherent states – also called Glauber states – of light are relevant to much of modern optical physics. They are the “most classical” form of light field. They are generated from classical sources, and they also describe the state of light in a laser cavity. To discuss the properties of coherent states, we focus on a single mode with frequency ω and ladder operators $\hat{a}^{(\dagger)}$.

A coherent state $|\alpha\rangle$ is defined by the condition

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle. \quad (14.20)$$

That is, it is an eigenstate of the destruction operator, with eigenvalue α . Since \hat{a} is not Hermitian, the eigenvalue α is, in general, complex.

The coherent state may be written

$$|\alpha\rangle = \frac{1}{N_\alpha} \exp(\alpha\hat{a}^\dagger)|0\rangle = \frac{1}{N_\alpha} \sum_{n=0}^{\infty} \frac{(\alpha\hat{a}^\dagger)^n}{n!} |0\rangle \quad (14.21)$$

where $|0\rangle$ is the oscillator vacuum and N_α is a normalization factor.

▷ EXERCISE. Prove that (14.20) is satisfied by the state (14.21).

Roy Jay Glauber 1925-

An American theoretical physicist, he is the Mallinckrodt Professor of Physics at Harvard University and Adjunct Professor of Optical Sciences at the University of Arizona. Born in New York City, he was awarded one half of the 2005 Nobel Prize in Physics “for his contribution to the quantum theory of optical coherence”, with the other half shared by John L. Hall and Theodor W. Hänsch.



To find the normalization, consider

$$\begin{aligned}\langle \alpha | \alpha \rangle &= \frac{1}{N_\alpha^2} \sum_{n,n'} \langle 0 | \frac{(\alpha^* \hat{a})^{n'}}{n'!} \frac{(\alpha \hat{a}^\dagger)^n}{n!} | 0 \rangle \\ &= \frac{1}{N_\alpha^2} \sum_n \frac{(\alpha^* \alpha)^n}{n!} \\ &= \frac{1}{N_\alpha^2} e^{|\alpha|^2}.\end{aligned}$$

Thus, a normalized state is obtained for $N_\alpha = \exp(|\alpha|^2/2)$, and

$$|\alpha\rangle = \exp(-|\alpha|^2/2) \exp(\alpha \hat{a}^\dagger) |0\rangle$$

Since the coherent state is not an energy eigenstate, expectation values of observables are, in general, time-dependent. In the Heisenberg representation, where the time-dependence is in the operators, the electric field (for this mode) is

$$\hat{\mathcal{E}}(\mathbf{r}, t) = i \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \left[-\hat{a} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + \hat{a}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \right]$$

Writing α in polar form

$$\alpha = |\alpha| e^{i\theta}$$

one finds

$$\begin{aligned}\langle \alpha | \hat{\mathcal{E}}(\mathbf{r}, t) | \alpha \rangle &= i \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \left[-|\alpha| e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t + \theta)} + |\alpha| e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t + \theta)} \right] \\ &= 2|\alpha| \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \sin [\omega t - \mathbf{k} \cdot \mathbf{r} - \theta]\end{aligned}$$

This looks like a classical field, with amplitude $\propto |\alpha|$ and phase θ .

Furthermore, a calculation of the fluctuations of $\hat{\mathcal{E}}(\mathbf{r}, t)$ leads to

$$\langle (\Delta \mathcal{E})^2 \rangle = \frac{\hbar\omega}{2\epsilon_0 V}$$

which is identical to the fluctuations for the vacuum state.

The coherent-state is the “most classical” state of the oscillator in the sense that it exhibits the same time dependence as a classical field, while having only the noise of the vacuum.

14.2.5 Non-classical light

Since the EM field is a quantum system, all states of the EM field are in some sense non-classical. Of course, in practice it may be difficult to prepare or detect quantum effects in a light field. For example, classical sources typically generate coherent states (described above). If the amplitude of the field is large (or our sensitivity to measuring the field is low), then the quantum fluctuations of the field in the coherent state will be undetectable.

In modern optical physics – particularly in cavities with small mode volume V where even a single photon can generate a sizeable electric field – techniques are very advanced, and it is routine to generate and to detect non-classical effects of light. “Non-classical” in this context is used to denote a state of the light field that is *not* a coherent state. As a definition this is rather vague, qualifying all but this special coherent state as non-classical. There are, however, a few central examples of non-classical states. These are referred

to as **squeezed states** of light. “Squeezing” refers to the reduction in the uncertainty in one particular observable. (Owing to the uncertainty relations, this will lead to an increase in the uncertainty of the variable conjugate to this.)

▷ **EXAMPLE: Squeezed light:** We have already come across what is perhaps the most non-classical of all such states: the Fock state, for which the EM field is time-independent, with vanishing expectation value. This can be described as an amplitude, or “number”, squeezed state, having the minimum possible uncertainty in photon number. (By contrast there is a conjugate variable, which is loosely defined as the phase of the wave, which is completely uncertain.)

Another form of squeezed light is obtained by “quadrature” squeezing. The quadrature operators are defined by

$$\begin{aligned}\hat{X}_1 &= \frac{1}{2}(\hat{a} + \hat{a}^\dagger) \\ \hat{X}_2 &= \frac{1}{2i}(\hat{a} - \hat{a}^\dagger)\end{aligned}$$

For a simple harmonic oscillator, they can be viewed as (proportional to) the position and momentum operators.

For a single mode of the EM field, the electric field may be written

$$\hat{\mathcal{E}} = (\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{+i\omega t})$$

(For simplicity, we neglect spatial dependence and polarisation dependence, and have rescaled the electric field by the vacuum amplitude $\sqrt{\hbar\omega/(2\epsilon_0V)}$.) In terms of the quadrature operators, this becomes

$$\hat{\mathcal{E}} = 2 \left[\hat{X}_1 \cos(\omega t) + \hat{X}_2 \sin(\omega t) \right]$$

Thus, X_1 and X_2 are the field amplitudes for the components that oscillate $\pi/2$ out of phase with each other.¹¹

The quadrature operators satisfy the commutation relation

$$[\hat{X}_1, \hat{X}_2] = \frac{i}{2}$$

from which it follows that

$$\Delta X_1 \Delta X_2 \geq \frac{1}{4}.$$

For the number states $|n\rangle$

$$\Delta X_1 = \Delta X_2 = \frac{\sqrt{(2n+1)}}{2}.$$

Of these, only the vacuum $|0\rangle$ is a minimum uncertainty state.

The coherent states $|\alpha\rangle$ have the (remarkable!) property that

$$\Delta X_1 = \Delta X_2 = \frac{1}{2}$$

independent of α . Thus, they are minimum uncertainty states. Since $\Delta X_1 = \Delta X_2$, at all times during their time evolution, the uncertainty in the electric field

$$(\Delta \mathcal{E})^2 = (2\Delta X_1)^2 \cos^2(\omega t) + (2\Delta X_2)^2 \sin^2(\omega t)$$

is constant and equal to the uncertainty of the vacuum.

The “quadrature squeezed” states are minimum uncertainty states, $\Delta X_1 \Delta X_2 = 1/4$, for which $\Delta X_1 \neq \Delta X_2$. Thus, either ΔX_1 or ΔX_2 is less than $1/2$. Now, the uncertainty in the electric field¹²

$$(\Delta \mathcal{E})^2 = (2\Delta X_1)^2 \cos^2(\omega t) + (2\Delta X_2)^2 \sin^2(\omega t)$$

¹¹These two components are in quadrature, hence the name.

¹²We assume that $\langle \hat{X}_1 \hat{X}_2 + \hat{X}_2 \hat{X}_1 \rangle = 2\langle \hat{X}_1 \rangle \langle \hat{X}_2 \rangle$, which can always be achieved for quadrature squeezed light by introducing an offset in time.

varies in time. At certain times in the cycle, the uncertainty in the electric field is *smaller than the uncertainty in the vacuum*. (If $\Delta X_1 < 1/2$, then the uncertainty is below its vacuum value for $t = \pi/\omega \times \text{integer}$).

Quadrature squeezed light can be generated using nonlinear optical elements. That there is less “noise” in one of the quadratures than in a coherent state, or in the vacuum state, has possible technological applications, particularly in precision measurements and the detection of weak signals. One example is in the detection of gravity waves with large-scale interferometers as in LIGO and VIRGO.