I. SYLLABUS AND INTRODUCTION

Let us start with a brief overview of the items that will (hopefully) be covered in this course and to give a guideline of what we are trying to learn.

Quantum mechanics 660 and 661 are advanced quantum mechanics courses designed for graduate students. The courses will be treated as a one-year course. It will be assumed that students have already some background in quantum mechanics (the concepts of waves, quantization, expectation values, etc.). An advised introductory textbook is *Introduction to quantum mechanics* by Griffiths. A number of things will, however, be repeated albeit in a more mathematical fashion. Also some basic knowledge in Fourier analysis, differential equations and linear algebra (calculation of eigenvalues and eigenstates) will be expected. All the information you need to know are in the lecture notes. However, the lecture notes are relatively concise and further reading is recommended. The book that was used last year was Liboff, *Introductory Quantum Mechanics*, fourth edition. However, a lot of other books treat the material of 660 in a similar fashion, so pick a book that you like best. Some of the material is not covered in Liboff (the older editions don't have relativistic quantum mechanics and quite a bit of the material of 661 is not in there).

The course is developed from a physicists point of view. Some concepts and mathematical techniques will be introduced along the way (such as matrix techniques) under the assumption that most of you are already somewhat familiar with them (after all, there are also 3 and 4 hundred level quantum mechanics courses). For a more comprehensive view of those things, I will refer to the textbook. The subjects that will be covered are

- **History**, Planck's quantum hypothesis, photoelectric effect,specific heat, atomic physics.
  The introduction is somewhat different from what is usually given in textbooks. However, since most of you have already followed a quantum mechanics course and since this is a graduate course, I think it is interesting to do it this way. Quite some time will be devoted to Planck's hypothesis. The physical language is quite different from that of the rest of the course (it is more thermodynamics than quantum mechanics). Some of these things are treated in the exercizes of Chapter 2 of Liboff. Other things will be mentioned more briefly.

- **Wave mechanics**, Schrödinger equation, probability density, free particles, Heisenberg's uncertainty principle.
  This section tries to show some of the thinking that led to the Schrödinger equation. As soon as you leave the idea of a particle as a billiard ball behind you, a number of other things follow directly, such as Heisenberg's uncertainty principle.

- **Harmonic oscillator**, Schrödinger equation approach, Solution using operators, Operators and wavefunctions, Correspondence principle.
  Although most of you have seen the harmonic oscillator already, it is an important example, since the number of exactly solvable models in quantum mechanics is limited. Try to pay special attention to the two approaches to attack this problem. The Schrödinger approach, solving the differential equation for the wavefunction, and the algebraic or operator approach due to Heisenberg, Jordan, and Born. Although the Schrödinger approach is more intuitive, the operator approach will become more important as we go on, since often we do not really care that much about the wavefunction, but we are more interested in the spectral properties (energies) Of the problem.

- **The Hydrogen atom**, Legendre polynomials, angular momentum, matrix notation.
  Again pay attention to the dual approach to this problem. The Schrödinger approach (solving the wavefunction) and the operator approach (angular momentum). This involves quite a bit of mathematics (solving differential equations). Here we will only deal with the potential from the nucleus, leading to the orbitals of the atom. Other aspects follow in the next section.

- **Relativistic quantum mechanics**, Klein-Gordan equation, Dirac equation, spin-orbit coupling.
  This section introduces the Dirac equation and its consequences for the problems that we are interested in: the introduction of spin, relativistic correction to the Hydrogen atom (spin-orbit coupling, lowest order kinetic corrections). We could introduce (which would have been simpler) these aspects in a phenomenological or semi quantum/classical approach, but this is more fun and avoids having to treat relativistic quantum mechanics as some appendix. Note that we are not really interested in things such as time dilation, length contraction, etc.

- **Perturbation theory and absorption and emission of photons**, time-independent perturbation theory, time-dependent perturbation theory, interaction of radiation and atom
  To be able to describe the emission and absortion of photons by an atom, we discuss perturbation theory. This will be done in some detail because of its relevance to later problems.

- **Many-electron atom** Pauli's exclusion principle, the Helium atom, periodic table, Coulomb interactions, addition of angular momenta.

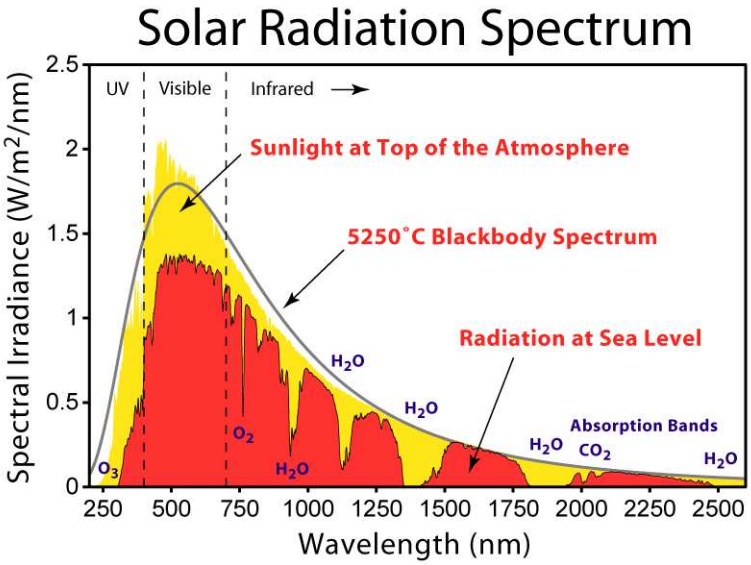


FIG. 1: Left: The radiation spectrum of the sun at the top of the atmosphere (yellow) and at sea level (red) compared with a blackbody spectrum for 5250°C. Right: Max Planck in 1901, around the time he developed the law now named after him. Source: wikipedia

Having solved the atom for one electron, we would like to fill it with more electrons. First, we need to discuss a peculiar aspect of electrons (or fermions in general), namely the exclusion principle. We start with the simplest many-electron atom, Helium. We soon discover that electron-electron interaction have a strong influence and see that things are not as simple as suggested by the Hydrogen atom solution. We continue with a more general discussion of Coulomb interaction between electrons and how to couple the electron wavefunction in a smart way (LS-coupling).

- **Molecules and solids**, Hydrogen molecule, solids, (nearly) free-electron model.
  We then continue by taking two atoms amd bringing them close together, thereby forming a molecule. If we continue adding atoms, we can form a one-dimensional solid, this is known as a tight-binding approximation. We can also approach it from the other side, by viewing a solid as a box of free electrons. The periodic potential has a only a small, but important, effect.

- **Second quantization**, the Coulomb gas.
  We have a look at the ground-state energy of the electron gas and calculate it in the Hartree-Fock limit. We also see what the effect are of screening.

II. OLD QUANTUM MECHANICS

A. Planck's quantum hypothesis

The quantum theory started with the blackbody radiation. All matter emits electromagnetic radiation when its temperature is above abolute zero. The radiation results from the conversion of thermal energy into electromagnetic energy. A black body is a systems that absorbs all incident radiation. On the other hand, this also makes it the best possible emitter. It is hard to find a perfect black body, but, for example, graphite, is a good approximation. However, even when a system is not a perfect black body, the basic phenomena remains the same. The sun can, for example, be described to a good approximation as a black body, see Fig. 1. The spectral distribution has a maximum and decays for small and large wavelengths. For the sun which has an effective temperature of around 5500 K, the maximum lies in the visible region of the electromagnetic spectrum and is inherently related to the fact that the sun is white (and not yellow or even yellowish). Before Max Planck took up the subject there was already a surprising

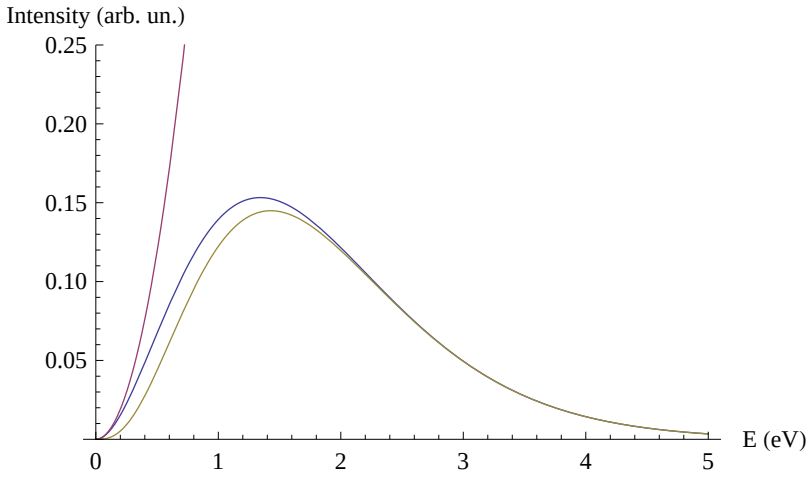


FIG. 2: The blue line shows the Planck distribution as a function of energy for $T = 5523$ K. The brown line shows Wien's distribution curve. The magenta line is the Rayleigh-Jeans law.

amount of understanding of the spectral distribution predominantly due to Wilhelm Wien (1864-1928). For his work, he received the Nobel prize for Physics in 1911. Using general thermodynamic arguments, Wien showed in 1893 that the energy dependence is given by

$$u(\omega) = \omega^3 f(\frac{\omega}{T}), \tag{1}$$

where $\omega = 2\pi\nu$ and f is a function that only depends on the ratio ω/T, where T is the temperature. The function $f(\frac{\omega}{T})$ cannot be derived from these considerations, but it can only be a function of ω/T or λT when expressed in the wavelength λ. The consequence of that is that the maximum occurs at a particular value of T/λ. Therefore, if the temperature changes the maximum occurs at a different wavelength. The human body can also be viewed as a far-from-perfect black body. The maximum occurs for a wavelength

$$\lambda_{\text{max,human}} = \frac{\lambda_{\text{max,Sun}} T_{\text{Sun}}}{T_{\text{human}}} = \frac{525 \times 10^{-9} 5500}{310} \cong 9 \ \ \mu\text{m}, \tag{2}$$

which is in the far infrared. Obviously, we would like to understand that the function $f(\frac{\omega}{T})$ is. Wien made an educated guess that $f(\frac{\omega}{T}) = \exp(-a\omega/T)$, where a is a constant. This dependence was obviously inspired by the Boltzmann factor and is a surprisingly good guess, see Fig. 2. However, there is something strange about using the Boltzmann factor. The factor $\exp(-E/k_B T)$, where k_B is Boltzmann's constant, implies that there is some finite energy associated with exciting electromagnetic radiation. This is in complete contradiction with the classical way of thinking. The energy density of electro-magnetic radiation is given by $\frac{1}{2}\varepsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2$, and therefore determined by the magnitudes of the electric and magnetic fields. Since we can take any possible value of these fields, any possible value of of the energy density is possible.

An attempt to derive a theory for heat radiation was made by Rayleigh and Jeans. We can assume that a black body is an electromagnetic cavity. Inside the cavity the radiation oscillates between the walls. In the x direction, the electromagnetic modes are determined by $\sin k_x x$ which has to be zero at $x = 0$ and $x = L$, i.e., the sides of the cavity. This means that $k_x L = n_x \pi$, with $n_x = 1, 2, 3, \ldots$. This implies that

$$(k_x^2 + k_y^2 + k_z^2)L^2 = k^2 L^2 = \left(\frac{\omega L}{c}\right)^2 = (n_x^2 + n_y^2 + n_z^2)\pi^2. \tag{3}$$

Note that the modes in the cavity are quantized. However, this is something that is understood from classical wavemechanics and already occurs when considering standing waves in a string with fixed boundaries. This does not cause a quantization of the energy. We can compare this with a simple one-dimensional harmonic oscillator. The total energy of an oscillator is given by $\frac{1}{2}m\omega^2 x_0^2$. Even though there is only one mode with frequency ω, this mode can have any arbitrary energy, because the amplitude of the oscillation x_0 can assume any energy. From the equipartition theorem in thermodynamics, this energy can be directly related to the temperature, since, in thermal equilibrium, each independent degree of freedom can be associated with an energy $\frac{1}{2}k_B T$. An oscillator has two degrees of freedom

(position and momentum) and therefore has an average energy $\overline{E} = k_B T$. Note, that the amplitude is not the same. High-frequency oscillators have a smaller amplitude than low-frequency ones. The argument for electro-magnetic cavity modes, which also have two degrees of freedom (the electric and magnetic parts), is very similar. So far, there does not appear to be too much of a problem, until you realize that the number of modes is infinite. If each of the modes has an energy $k_B T$, the total energy would be infinite as well. Let us have a look at the number of modes in more detail. The density of modes can be found from Eqn. (3),

$$\frac{\omega^2 L^2}{c^2} = \pi^2 n^2 \quad \Rightarrow \quad n = \frac{\omega L}{c\pi} \quad \Rightarrow \quad 2\frac{1}{8} 4\pi n^2 dn = \pi \left(\frac{L}{c\pi}\right)^3 \omega^2 d\omega = V g(\omega) d\omega \quad \Rightarrow \quad g(\omega) = \frac{\omega^2}{c^3 \pi^2} \tag{4}$$

with $n^2 = n_x^2 + n_y^2 + n_z^2$. The factor 2 comes from the two different polarization vectors of the light. The factor $\frac{1}{8}$ comes from taking only the positive k values. The Rayleigh-Jeans law can now be derived as

$$u(\omega) = g(\omega)\overline{E} = \frac{\omega^2}{c^3 \pi^2} k_B T. \tag{5}$$

Note that the Rayleigh-Jeans equation has a temperature dependence but just increases as a function of ω, see Fig. 2. Integrating over ω would give an infinite internal energy in our cavity. This was later called the ultraviolet catastrophe. Furthermore, it also deviates from Wien's suggestion at low energies.

This was more or less the state of affairs in 1900, when Max Planck (1858-1947) entered the scene. Planck worked most of his life and his derivation is based on entropy, effectively obtaining an interpolation between the Rayleigh-Jeans law and Wien's expression, which provide the proper limits in the low and high frequency regions, respectively. Later on, Planck realized that the result can be interpreted as a quantization of the energy in quanta of $\hbar\omega$, where, for clarity we use the conventional notation of Planck's constant. This implies that the energy of an harmonic oscillator can be given by $E = n\hbar\omega$, where $n = 0, 1, 2, \cdots$. This looks rather different from the classical result $E = \frac{1}{2} m\omega^2 x_0^2$.

There are therefore two aspects in the expression $E = n\hbar\omega$. First, the energy is directly related to the frequency. Second, the energy is quantized. Let us consider the former statement first. Let us assume that the energy can be varied continuously, i.e. $E = a\omega$, with $a \geq 0$. Since E is proportional to ω, there is an equal number of states in each energy interval of width dE. This allows us to easily calculate the average energy. The probability of finding an oscillator with a certain energy is determined by the Boltzmann factor $\exp(-\beta E)$ where we take $\beta = 1/k_B T$. The average energy can be calculated by integrating the energies multiplied by the probability of finding that energy, and in the end dividing by the total probability. Let us first determine the total probability:

$$Z = \int_0^\infty dE e^{-\beta E} = -\frac{1}{\beta}\left[e^{-\beta E}\right]_0^\infty = \frac{1}{\beta} = k_B T \tag{6}$$

The average energy is now

$$\overline{E} = \frac{1}{Z}\int dE E e^{-\beta E} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\beta \frac{\partial}{\partial \beta}\left(\frac{1}{\beta}\right) = \frac{1}{\beta} = k_B T. \tag{7}$$

This is the classical equipartition effect and would directly reproduce the Rayleigh-Jeans result. Therefore, although $E \sim \omega$ is an interesting finding, it does not solve the ultraviolet catastrophe.

Now let us see what happens if we quantize the energy. In this case, we can no longer integrate over the energies, but we have to perform a summation. Let us again start with the total probability,

$$Z = \sum_{n=0}^{\infty} \exp(-\beta n\hbar\omega) = \frac{1}{1 - \exp(-\beta\hbar\omega)}. \tag{8}$$

The average energy is then

$$\overline{E} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} = \frac{\partial}{\partial \beta}\ln(1 - \exp(-\beta\hbar\omega)) = \frac{\hbar\omega}{1 - \exp(-\beta\hbar\omega)}\exp(-\beta\hbar\omega) = \frac{\hbar\omega}{\exp(\beta\hbar\omega) - 1}. \tag{9}$$

Let us consider the limits of small and large frequencies. For small frequencies, $\hbar\omega \ll k_B T$, and we find

$$\overline{E} = \frac{\hbar\omega}{\exp(\beta\hbar\omega) - 1} \cong \frac{\hbar\omega}{1 + \beta\hbar\omega - 1} = \frac{1}{\beta} = k_B T. \tag{10}$$

Therefore, for small frequencies, the fact that the energy is quantized as $n\hbar\omega$ is irrelevant and we obtain the classical equipartition result. For $\hbar\omega \gg k_B T$, Planck's formula reduces to

$$\overline{E} = \frac{\hbar\omega}{\exp(\beta\hbar\omega) - 1} \cong \hbar\omega \exp(-\beta\hbar\omega), \tag{11}$$

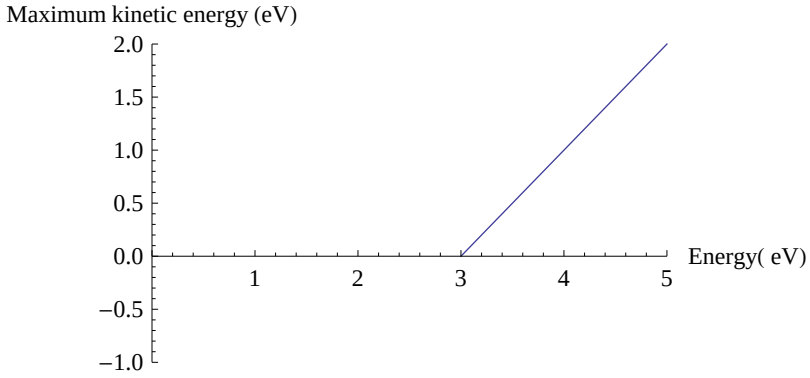


FIG. 3: The maximum kinetic energy of electrons emitted from a solid by electro-magnetic radiation. The work function is Φ =3 eV. Note that no electrons are observed for photon energies less than 3 eV.

which is equivalent to Wien's suggestion. Note that the average energy $\overline{E}$ decreases exponentially for larger frequencies. This effectively solves the ultraviolet catastrophe. The energy distribution is now given by

$$u(\omega) = g(\omega)\overline{E} = \frac{\omega^2}{c^3\pi^2}\frac{\hbar\omega}{\exp(\beta\hbar\omega) - 1}. \tag{12}$$

Although the result is revolutionary, it took some time to gain momentum. First, Planck did not set out to solve the ultraviolet catastrophe, but was mainly looking for an effective formula to describe the black-body radiation. Secondly, there are some subtle arguments about the applicability of the equipartition theorem for a cavity since thermal equilibrium is not easily established due to the lack of coupling between the different frequency modes. Thirdly, Planck did not quantize the electromagnetic field, but was more thinking about microscopic dipole oscillators.

B. Photoelectric effect

The identification of quantization with the electro-magnetic field was done in 1905 by Albert Einstein (1879-1955). Einstein proposed that the energy can only be transferred in discrete steps of $\hbar\omega$. Applying this to a solid gives the very simple expression

$$E_{\rm kin} = \hbar\omega - \Phi. \tag{13}$$

The explanation of the photoelectric effect was the primary reason for Einstein's Nobel prize in 1921. Here Φ is the work function of a solid, i.e., the minimum energy required to remove an electron. The maximum kinetic energy is plotted in Fig. 13. This is independent of the intensity of the light. This is in complete contradiction with classical theories where the amplitude of the incoming light would play an important role in determining the kinetic energy

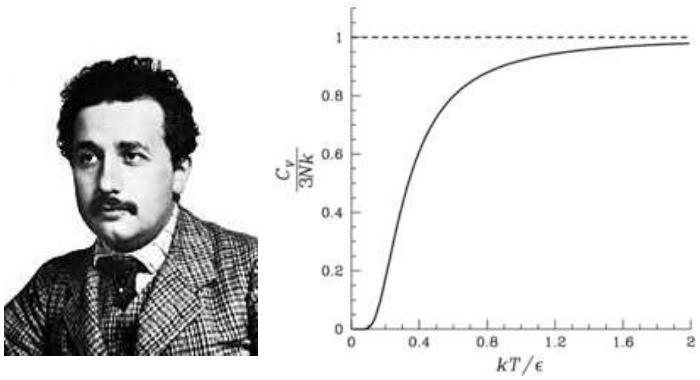


FIG. 4: The specific heat for an Einstein oscillator with an energy $\varepsilon = \hbar\omega$. (source wikipedia)

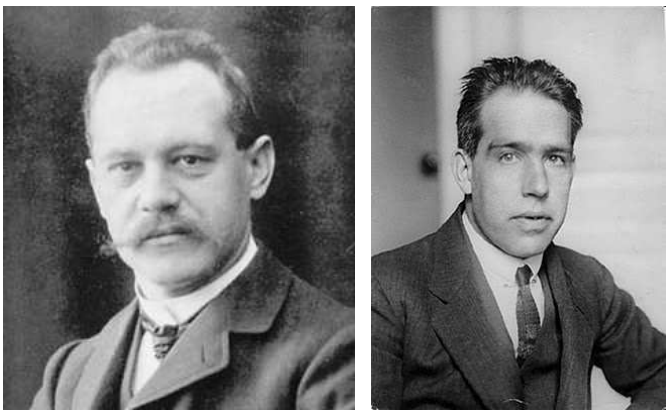


FIG. 5: Arnold Sommerfeld and Niels Bohr, key figures in the establishment of the "old" quantum mechanics.

of the outgoing electron (in a fashion similar that a greater amplitude of your golf swing gives a larger kinetic energy to the golf ball). Obviously, the statement of Einstein would have been pretty obvious if experimental data of the type in Fig. 13 existed. However, Einstein's claims were not experimentally verified until 1914 by Robert Andrews Millikan (1868-1953), which was partially the reason for his Nobel prize in 1923.

C. Specific heat

Up to 1910, quantum physics was mainly limited to the blackbody effect and the photoelectric effect. Its impact on the rest of physics was still rather limited. An important contribution was made by Einstein on the specific heat of solids. Classically this is expected to follow the Dulong-Petit law, i.e., each atom contributes $3k_BT$ (or $3RT$ per mole) to the total energy E, i.e., k_BT per degree of freedom (assuming a three-dimensional system). The specific heat is the quantity that tells how much the total energy of the system changes when the temperature is raised. This gives a specific heat of

$$C_v = \frac{\partial \overline{E}}{\partial T} = 3R. \tag{14}$$

Einstein recognized that each degree of freedom is governed by the same statistics as the oscillators in the blackbody radiation problem, giving an average energy of

$$\overline{E} = \frac{3N\hbar\omega}{\exp\frac{\hbar\omega}{k_BT} - 1} \qquad \Rightarrow \qquad C_v = \frac{3R\left(\frac{\hbar\omega}{k_BT}\right)^2 \exp\frac{\hbar\omega}{k_BT}}{(\exp\frac{\hbar\omega}{k_BT} - 1)^2}, \tag{15}$$

which is plotted in Fig. (4). Note that this reduces to the Dulong-Petit result for high temperatures. This was one of the first applications of quantum mechanics in solid-state physics and chemistry. This convinced the physical chemist Walter Nernst of the importance of quantum theory. He was responsible for convincing the Belgian industrialist Ernest Solvay to organize the famous conferences that bear his name.

D. Bohr-Sommerfeld model

Additional concepts on quantization were introduced by Niels Bohr (1885-1962) and Arnold Sommerfeld (1868-1951) in the period 1913-1920. The idea is essentially that classical physics is still obeyed, but that certain motion are not allowed. Although the theory is now superseded by wave mechanics, it is interesting to have a look into how one tried to grapple with the idea of quantization for about two decades. The quantization conditions were somewhat heuristic and based on the notion that the classic action integral can only assume certain value. For example, in one dimension

$$\int p_x dx = nh \tag{16}$$

Let us consider the case of an harmonic oscillator. The expression for the energy can be rewritten as an ellips in the x,p_x plane:

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 = E \qquad \Rightarrow \qquad \frac{p_x^2}{a^2} + \frac{x^2}{b^2} = 1, \tag{17}$$

with $a = \sqrt{2mE}$ and $b = \sqrt{2E/m\omega^2}$. The action integral is the area of the ellips and should be proportional to nh,

$$\int pdx = \pi ab = \frac{2\pi}{\omega}E = nh \Rightarrow E = n\hbar\omega, \tag{18}$$

giving the right relation between E and ω.

The most impressive result were however obtained by Bohr's treatment of the hydrogen atom for which he received the Nobel prize in 1922 (note that this is before the development of the wavefunction formalism of quantum mechanics). Strong evidence of nonclassical behavior was observed in atomic spectra. Although the spectroscopic data was already available in the late 19th century, it had a strong impact on quantum mechanics from 1910 onwards. Phenomenologically, it was already established that the wavelengths of the spectral lines could be related to each other by rational numbers. It was Balmer, a Swiss school teacher with an interest in numerology who noticed that the first four wavelengths could are proportional to each other by $\frac{9}{5}, \frac{4}{3}, \frac{25}{21}, \frac{9}{8}$. After some playing around, he noticed that this could also be written as $\frac{9}{5}, \frac{16}{12}, \frac{25}{21}, \frac{36}{32}$ or $m^2/(m^2-2^2)$, which reproduces the twelve known Hydrogen lines. This was 1885. Much more spectral information was obtained by, for example, Rydberg. After some attempts by Thompson and others, it was Rutherford who clearly established the model of the atom as a relatively small nucleus with a thin cloud of electrons turning around it. The stability of the electrons in orbits around the nucleus is in contradiction with classical electrodynamics, which states that the electron would emit radiation in its orbit and would therefore and up on the nucleus. This was solved in a some important papers by Bohr in 1913, who wrote down the conditions for stationary states by relating it to Planck's constant (one of the first uses of quantum mechanics not related to oscillators).

The quantization condition states

$$\int pdq = nh, \tag{19}$$

where q is now a generalized coordinate. For an electron in an orbit of radius r, we can write $q = r\varphi$, where φ is the angular coordinate.

$$\int_0^{2\pi} prd\varphi = \int_0^{2\pi} Ld\varphi = 2\pi L = nh \quad \Rightarrow \quad L = n\hbar, \tag{20}$$

where the magnitude of the angular momentum is $L = |\mathbf{r} \times \mathbf{p}|$. The equation of motion is then

$$F = \frac{mv^2}{r} = \frac{Ze_0^2}{r^2} \qquad \text{with} \qquad e_0 = \frac{e}{\sqrt{4\pi\epsilon_0}}, \tag{21}$$

with Z the number of protons in the nucleus. Solving this for v and r gives

$$v_n = \frac{Ze_0^2}{n\hbar} \qquad \text{and} \qquad r_n = \frac{n^2}{Z}\frac{\hbar^2}{mne_0^2}. \tag{22}$$

The energy of the state is then

$$E_n = \frac{1}{2}mv_n^2 - \frac{Ze_0^2}{r_n} = -\frac{Z^2}{n^2}\frac{me_0^4}{2\hbar^2} = -13.6\frac{Z^2}{n^2}\ \text{eV}. \tag{23}$$

The emitted or absorbed photon energy is related to the energy of the stationary states

$$\hbar\omega_{mn} = E_m - E_n = R(\frac{1}{n^2} - \frac{1}{m^2}), \qquad \text{with} \qquad m > n \geq 1, \tag{24}$$

The value of 13.6 eV reproduced the experimentally observed Rydberg constant. An impressive result indeed.

Despite some successes of the Bohr-Sommerfeld model, ultimately it turned out to be deficient. First, there was no standard way of obtaining results, but every problem had to be considered separately. Second, it fails to reproduce many features, in particular for the model of the hydrogen atom. For example, the lowest has in fact $L = 0$ and not $L = \hbar$, it cannot describe fine structure in the atomic levels, or the Zeeman effect.

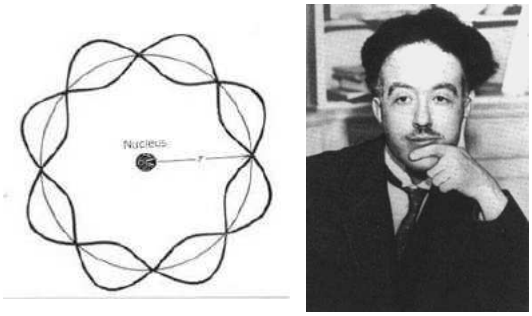


FIG. 6: By assuming that the wavelength λ of a particle is proportional to its inverse momentum h/p, de Broglie was able to show that the Bohr quantization condition corresponded to having a standing wave in the orbit of an electron.

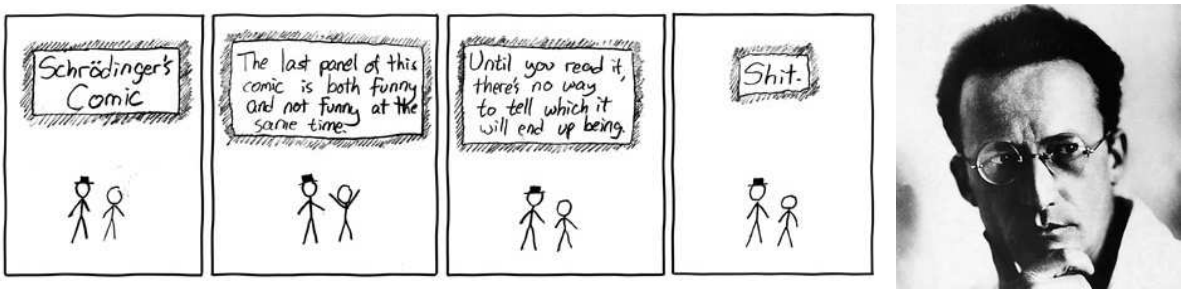


FIG. 7: Erwin Schrödinger developed one of the fundamental equations of quantum mechanics describing matter in terms of waves.

III. WAVE MECHANICS

Up to the early 1920's, quantum mechanics had no solid foundations. Applications were based on applying good physical intuition to the quantization of various problems. This changed when more emphasis was placed on the dual character of particles and radiation, both particle and wave like. The particle character of radiation was clearly demonstrated in the photoelectric effect. It was Louis de Broglie (1892-1987) who introduced in his 1924 Ph.D. thesis the concepts the wave character of particles, identifying the wavelength of a particle with its momentum

$$\lambda = \frac{h}{p} \qquad \text{or} \qquad p = \hbar k, \tag{25}$$

where $k = 2\pi/\lambda$ is the wavenumber. Bohr's quantization of angular momentum then corresponds to $L = rp = n\hbar$, giving $2\pi r = n\lambda$. Implying that an integral number of wavelengths should fit on an orbit, i.e., a standing wave on a circle. This provides additional understanding of the concept of stationary orbits. If the radius of an electron orbit would change, this would result in the destructive interference of the wave. Hence, the radius can only assume very particular values.

The wave character of particles was confirmed in 1927 by Clinton Davisson (1881-1958) and Lester Germer (1896-1971) at Bell Labs and, independently, by George Paget Thomson (1892-1975, the son of Nobel laureate J. J. Thomson) at the University of Aberdeen. Essentially, they performed diffraction experiments, not with X-rays, but with an electron beam. The diffraction pattern could be explained by using the Bragg condition for X-rays but with the de Broglie wavelength. The 1937 Nobel prize in physics was awarded to Davisson and Thomson for their work on electron diffraction.

A. Schrödinger equation

The wave character of the particles as investigated by a number of physicists, most notably Erwin Schrödinger (1887-1961). An obvious starting point is the wave differential equation for particles

$$\nabla^2\psi(\mathbf{r},t) - \frac{1}{v_p^2}\frac{\partial^2\psi(\mathbf{r},t)}{\partial t^2} = 0, \tag{26}$$

where v_p is the phase velocity. The nabla operator is Cartesian coordinates is given by

$$\nabla = \frac{\partial}{\partial x}\hat{\mathbf{x}} + \frac{\partial}{\partial y}\hat{\mathbf{y}} + \frac{\partial}{\partial z}\hat{\mathbf{z}} \quad \text{and} \quad \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{27}$$

For electro-magnetic waves, this velocity is equal to the speed of light, but obvously this is not correct for particles. Solutions to this equation are of the form

$$\psi(\mathbf{r},t) = ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}. \tag{28}$$

Using the quantum mechanical identities $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$ from Planck and de Broglie, respectively, we can also write the solution as

$$\psi(\mathbf{r},t) = ae^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}. \tag{29}$$

Taking the time derivative in the wave equation gives

$$\nabla^2\psi - \frac{E^2}{\hbar^2 v_p^2}\psi = 0, \tag{30}$$

The unknown quantity in the above equation is the phase velocity $v_p = \omega/k$. We are very used to electromagnetic waves, where

$$E = \hbar\omega \qquad \Rightarrow \qquad pc = \hbar kc = \hbar\omega \qquad \Rightarrow c = \omega/k. \tag{31}$$

Therefore the phase velocity is equal to the real velocity (the speed of light). This is not true for particles. Here

$$\frac{1}{2}mv^2 = E - U(\mathbf{r}) \qquad \Rightarrow \qquad v = \sqrt{\frac{2(E-U(\mathbf{r}))}{m}}, \tag{32}$$

where $U(\mathbf{r})$ is the potential energy. For the wavevector we have

$$k = \frac{mv}{\hbar} = \frac{1}{\hbar}\sqrt{2m(\hbar\omega - U)}. \tag{33}$$

For the phase velocity we then have

$$v_p = \frac{\omega}{k} = \frac{\hbar\omega}{\sqrt{2m(\hbar\omega - U)}} = \frac{E}{\sqrt{2m(E-U)}}. \tag{34}$$

This is completely different from the velocity, with $vv_p = E/m$. For electromagnetic waves, we can have comparable changes in phase velocity when the waves travel through a different medium. In vacuum $c^{-2} = \varepsilon_0\mu_0$. In a medium, this changes to $\varepsilon_r\mu_r\varepsilon_0\mu_0 = n^2/c^2$. Moving though a potential landscape is therefore comparable to moving through a medium with a different refractive index. Inserting (34) in Eq. (30) gives the time-independent Schrödinger equation.

$$\nabla^2\psi + \frac{2m(E-U)}{\hbar^2}\psi = 0 \qquad \text{or} \qquad -\frac{\hbar^2}{2m}\nabla^2\psi + U\psi = E\psi. \tag{35}$$

The strange thing is that, whereas in Eq. (30) the energy appears in quadratic form, it now appears linearly in the equation. Since $i\hbar\partial\psi/\partial t = E\psi$ can put the time-dependent derivative back into the Schrödinger equation giving

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t) + U(\mathbf{r})\psi(\mathbf{r},t) = i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} \tag{36}$$

An alternative way to look at the Schrödinger equation is that the usual scalar quantities are now replaced by operators:

$$\mathbf{p} \to -i\hbar\nabla \qquad \text{and} \qquad E \to i\hbar\frac{\partial}{\partial t}. \tag{37}$$

For particles, we start from the usual expression for the energy

$$\frac{\mathbf{p}^2}{2m} + U(\mathbf{r}) = E. \tag{38}$$

Inserting the operators and making them work on a wavefunction leads to the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U(\mathbf{r})\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{39}$$

This can be written in a more condensed form by defining the Hamiltonian

$$H = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r}). \tag{40}$$

The time-independent Schrödinger equation is then simply

$$H\psi = E\psi. \tag{41}$$

For the development of the equation named after him, Schrödinger received the Nobel prize in 1933.

Example: bound states in a delta function

Let us now look at a particular example of a potential. You are probably familiar with a particle in a box, so we will consider the one-dimensional example of bound states in the presence of a delta function potential $V(x) = -\alpha\delta(x)$ with $\alpha >$, where the delta function is defined as

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0 \\ \infty, \text{if } x = 0 \end{cases}, \tag{42}$$

where in addition we have

$$\int_{-\infty}^{\infty} \delta(x)dx = 1. \tag{43}$$

The Schrödinger equation can now be written down as

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} - \alpha\delta(x)\psi(x) = E\psi. \tag{44}$$

In the absence of a potential, we only expect solutions for positive energies, since the kinetic energy is always positive. The solutions are in then plane waves. For positive energies, we still find plane waves, but the presence of the local scattering center causes a phase shift in the plane waves at the origin. We will not study this part in this example. Since the potential is attractive, we also have solutions for negative energies. These are bound states. Let us first look at the Schrödinger equation for $x \neq 0$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = -|E|\psi. \tag{45}$$

Since the energy is negative, the solution are not running waves but exponentials. Using the definition

$$\kappa = \frac{\sqrt{2m|E|}}{\hbar}, \tag{46}$$

the solutions can be written as

$$\psi(x) = Ae^{\kappa x} + Be^{-\kappa x}. \tag{47}$$

Neither of them are solutions in the absence of a potential since the potentials blow up at either plus or minus infinity. However, since there is a potential at $x = 0$, the solution can be written as

$$\psi(x) = \begin{cases} Ae^{\kappa x}, & \text{if } x < 0 \\ Be^{-\kappa x}, & \text{if } x > 0 \end{cases} \tag{48}$$

So what do we do at the $x = 0$. To solve this, we have to look at the two boundary conditions

- ψ is always continuous.
- $d\psi/dx$ is continuous except at the points where the potential is infinite.

The first point is very easy and simply implies that $A = B$. This follows also directly from the symmetry of the problem, since there is no reason why the left side should not be the mirror image of the right side. The second point is more tricky, especially since we are dealing with a situation where the potential is infinite. Let us first see why this is the case. The idea is to integrate the Schrödinger equation around a particular point. Let us focus on $x = 0$. Intgrating gives

$$-\frac{\hbar^2}{2m}\int_{-\epsilon}^{\epsilon} dx\ \frac{d^2\psi(x)}{dx^2} + \int_{-\epsilon}^{\epsilon} dx\ V(x)\psi(x) = E\int_{-\epsilon}^{\epsilon} dx\ \psi(x), \tag{49}$$

where we take the limit $\epsilon \to 0$. The first integral is of course $d\psi/dx$. The integral on the right-hand side goes to zero if $\epsilon \to 0$. So we find for the change in derivative

$$\Delta\left(\frac{d\psi(x)}{dx}\right) = \left.\frac{d\psi(x)}{dx}\right|_{\epsilon} - \left.\frac{d\psi(x)}{dx}\right|_{-\epsilon} = \frac{2m}{\hbar^2}\int_{-\epsilon}^{\epsilon} dx\ V(x)\psi(x). \tag{50}$$

The right-hand side goes to zero unless $V(x)$ goes to infinity. For our δ-function potential we find that the change in derivative is given by

$$\Delta\left(\frac{d\psi(x)}{dx}\right) = -\frac{2m\alpha}{\hbar^2}\psi(0). \tag{51}$$

evaluting gives

$$-A\kappa e^{-\kappa x}\big|_{0^-} - A\kappa e^{\kappa x}|_{0^-} = -2\kappa A = -\frac{2m\alpha}{\hbar^2}A \tag{52}$$

which gives

$$\kappa = \frac{m\alpha}{\hbar^2}. \tag{53}$$

The energies are then given by

$$E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{m\alpha^2}{2\hbar^2}. \tag{54}$$

To find the final result for the wavefunction, we just need to normalize it

$$\int_{-\infty}^{\infty} dx|\psi(x)|^2 = 2|A|^2\int_0^{\infty} dx e^{-2\kappa x} = \frac{|A|^2}{\kappa} = 1 \tag{55}$$

or

$$A = \sqrt{\kappa} = \frac{\sqrt{m\alpha}}{\hbar}. \tag{56}$$

The δ-function well has, regardless of the strength α, one bound state given by

$$\psi(x) = \frac{\sqrt{m\alpha}}{\hbar}\exp\left(\frac{-m\alpha|x|}{\hbar^2}\right); \qquad E = -\frac{m\alpha^2}{2\hbar^2} \tag{57}$$

We return to the problem of a bound state when discussing matrix algebra.

B. Probability density

The discovery of the dual character of particles led to several philosophical problems on how to interpret the waves that follow from the Schrödinger equation. Max Born (1882-1970) emphasized that we should look at $\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$

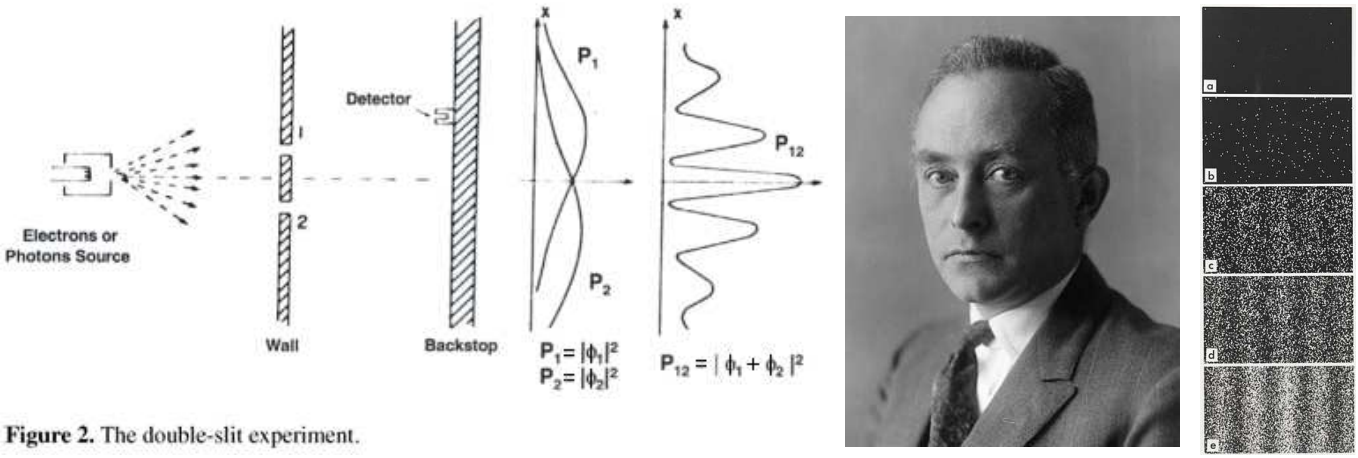


Figure 2. The double-slit experiment.

FIG. 8: Max Born developed the concepts of the interpretation of the wavefunction in terms of probability. An example is the two-split experiment where quantum mechanics correctly predicts the probability distribution of electrons hitting a wall after passing through two slits. The right shows the buildup of the image by electrons hitting the detector over time.

as the probability of finding a particle. The Schrödinger equation can be seen as a diffusion equation for probability. To see this one has to start from the Schrödinger equation in Eqn. (36) and its complex conjugate

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + U(\mathbf{r})\psi^*. \tag{58}$$

Multiplying the equations by ψ^* and ψ, respectively, and substracting gives

$$\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t} = -\frac{\hbar}{2mi}(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*). \tag{59}$$

This can be written as

$$\frac{\partial(\psi^*\psi)}{\partial t} = -\frac{\hbar}{2mi}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*). \tag{60}$$

This can be written as

$$\frac{\partial\rho}{\partial t} = -\nabla\cdot\mathbf{J} \qquad \text{with} \qquad \mathbf{J} = \frac{\hbar}{2mi}(\psi^*\nabla\psi - \psi\nabla\psi^*). \tag{61}$$

Integrating over the volume gives

$$\int d\mathbf{r}\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial t}\int d\mathbf{r}\psi^*\psi = -\int d\mathbf{r}\nabla\cdot\mathbf{J} = -\int_S \mathbf{J}d\mathbf{a} = \frac{1}{2m}\int_S[\psi^*(-i\hbar\nabla)\psi + \psi(-i\hbar\nabla\psi^*)]d\mathbf{a}, \tag{62}$$

where the last term gives the density going in and out of the surface S. Note that if we let the surface go to infinity and ensuring that $\psi \to 0$ when $r \to \infty$, then the right-hand side will be zero. This means that $\frac{\partial}{\partial t}\int d\mathbf{r}\psi^*\psi = 0$. Therefore if we had initially normalized the probability, it will remain normalized in time. Three decades after developing the concepts of probability density, Max Born received the Nobel prize in 1954.

There is a wide variety of interpretations regarding the wavefunction in quantum mechanics. Here, we follow the practical one, namely that in quantum mechanics there is only the wavefunction and the wavefunction correctly predicts many experiments. Generally, there is no need to imagine a particle "inside" the wavefunction. For example, quantum mechanics correctly predicts the probability distribution of a two-slit experiment see Fig. 8. Since particles have a wave character, the wavefunction should show destructive and constructive interference, just as an optical wave passing through two slits as explained by Thomas Young in the 1790s. This was more a thought experiment until it was actually measured by Clauss Jönsson of the University of Tübingen in 1961. Even more surprisingly, this even works with objects such as a C_{60} (bucky balls) molecule with a diameter of 0.7 nm as was demonstrated in 1999. This completely shatters our ideas of particles as little balls. Now there is an additional complicating factor of the measurement. One does not observe the whole electron interference pattern at once. In fact, it requires many electrons

for the image to build up, see Fig. 8. However, practically the "experiment" is one electron passing through two slits giving rise to an interference pattern. The "measurement" is the interaction of the electron with the detector which is a very complicated quantum system consisting of many particles. This causes the "collapse" of the wavefunction in a particular position. The problem of measurement and the transition from a wave to an electron hitting a detector is very difficult. However, often one is only interested in the "experiment": one wants to calculate the interference pattern of one electron passing through a slit. For example, when doing spectroscopy, one calculates a spectrum using quantum mechanics and generally completely ignores the effect of the "measurement".

An additional step to the two-slit experiment is the determination through which slit the electron passes. Here one sees that the interference pattern disappears and that one obtains two classical probability distributions of electrons passing through slits like particles, see Fig. 8. Again, the problem arises because of the "measurement". In quantum mechanics, the "measurement" is never simple, since the "measurement" is of the same order of magnitude as the system. One cannot simply separate the system from the measurement. This is different from classical mechanics, where a measurement is generally an extremely minor perturbance to the system. For example, the measurement of the speed of a car with the use of radio waves. The situation would change drastically is speed was measured with cannon balls. For example, the "detection" of whether an electron passes through a slit can be done by placing an atom inside each slit. However, is this still a two-slit experiment? From a quantum mechanics point of view, this is a system of an electron passing through two slits and while passing through the slits it can interact with atoms positioned in the slits. To really understand what is going on, this is the system you need to calculate. But then it changes from a one-particle problem (electron through two slits) to a many-particle problem, which is very difficult to calculate. Since the problem of "measurement" is highly complicated, we will not discuss it in too much detail since it just confuses things. The "experiment" (an electron passing through two slits) is beautiful enough in itself without further complicating it with the concepts of "measurement" (and in particular with philosophical questions about human involvement).

C. Expectation values

Since particles are no longer little balls, but wavefunctions, it is impossible to say where they are in real space. However, we can calculate the expectation value of the position. In one dimension, we have

$$\langle x\rangle = \int x|\psi(x)|^2 dx, \tag{63}$$

since $|\psi(x)|^2$ gives the probability of finding a particle at position x. However, how do we do this for more general operators. Let us have a look at the momentum, which should be given by

$$\langle p_x\rangle = m\frac{\langle x\rangle}{dt} = m\int x\frac{\partial|\psi(x)|^2}{\partial t}dx. \tag{64}$$

Using Eq. (61) in one dimension, we can write the right-hand side as

$$\langle p_x\rangle = \frac{i\hbar}{2}\int x\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x} - \frac{\partial\psi^*}{\partial x}\psi\right)dx. \tag{65}$$

Using an integration by parts gives

$$\langle p_x\rangle = -\frac{i\hbar}{2}\int\left(\psi^*\frac{\partial\psi}{\partial x} - \frac{\partial\psi^*}{\partial x}\psi\right)dx, \tag{66}$$

using the fact that $\partial x/\partial x = 1$ and omit the integrand since the wavefunction $\psi(\pm\infty) = 0$. Now the second term on the right-hand side is actually equivalent to the first term, which follows from another integration by parts. This gives

$$\langle p_x\rangle = -i\hbar\int\psi^*\frac{\partial\psi}{\partial x}dx = \int\psi^*\left(-i\hbar\frac{\partial}{\partial x}\right)\psi dx = \int\psi^* p_x\psi dx, \tag{67}$$

where we identify p_x as the x-component of $\mathbf{p} = -i\hbar\nabla$. Note that this follows directly from the Schrödinger equation which was derived from the wave equation. This justifies the heuristic assumption based on de Broglie's principle in Eq. (37). It is important to realize that the position of p_x in the integral is important since p_x is an operator and therefore affects the wavefunction. For a general operator $\hat{A}$, the expectation value is given by

$$\langle\hat{A}\rangle = \int d\mathbf{r}\psi^*(\mathbf{r})\hat{A}\psi(\mathbf{r}). \tag{68}$$

Another aspect that seems relatively trivial is that the expectation value is a real number. This will become of importance later on when we start discussing Heisenberg's matrix algebra. Also, the notation for the expectation value is a shorthand. More, generally, we should write

$$\langle \hat{A} \rangle \equiv \langle \psi | \hat{A} | \psi \rangle. \tag{69}$$

There are many problems that involve more than one wavefunction, for example the free particle problem in the next section. You may have notice that on the right-hand side, the ψ appears without the $\mathbf{r}$ dependence. This is because the expectation value does not depend on $\mathbf{r}$, because of the integration over the spatial coordinates. Operators can also have finite values between different wavefunctions

$$\langle \psi_{n'} | \hat{A} | \psi_n \rangle = \int d\mathbf{r} \psi^*_{n'}(\mathbf{r}) \hat{A} \psi_n(\mathbf{r}). \tag{70}$$

where the index n labels the different wavefunctions. In this case, $\langle \psi_{n'} | \hat{A} | \psi_n \rangle$ is called a matrix element. The use of matrix element or expectation value for the case $n = n'$ depends on the context in which it is used.

Some of you might be bothered that the momentum $\mathbf{p}$ is represented by a complex operator. Let us rewrite the expectation value,

$$\langle p_x \rangle \;=\; \int \psi^* \left(-i\hbar \frac{\partial \psi}{\partial x} \right) dx, \tag{71}$$

for the momentum in the x direction by performing an integration by parts

$$\langle p_x \rangle = -i\hbar \left[\psi^* \psi \right]_{-\infty}^{\infty} - \int \left(-i\hbar \frac{\partial \psi^*}{\partial x} \right) \psi dx = \int \left(-i\hbar \frac{\partial \psi}{\partial x} \right)^* \psi dx \tag{72}$$

The integrand is zero if the wavefunction properly goes to zero in $\pm\infty$. Now if ψ is an eigenfunction of p_x with eigenvalue q, or $p_x \psi = q\psi$ then from Eqs. (71) and (X D) it follows that

$$\langle p_x \rangle \;=\; \int \psi^* \left(-i\hbar \frac{\partial \psi}{\partial x} \right) dx = q = \int \left(-i\hbar \frac{\partial \psi}{\partial x} \right)^* \psi dx = q^*, \tag{73}$$

and therefore $q = q^*$ which is only possible if q is real. So that is reassuring that the expectation value of the momentum operator is a real number.

This can be generalized as follows. The operator $A^\dagger$ is the Hermitian conjugate of A if

$$\int (A^\dagger \psi)^* \psi d\mathbf{r} = \int \psi^* A \psi d\mathbf{r}. \tag{74}$$

The operator is Hermitian is $A^\dagger = A$. From Eq. (73), we see that the momentum operator satisfies this condition. Just as for the momentum operator, if ψ is an eigenfunction of A with eigenvalues q, we find that

$$\int \psi^* A \psi d\mathbf{r} = q = \int (A\psi)^* \psi d\mathbf{r} = q^* \tag{75}$$

and therefore that $q = q^*$ and the eigenvalues q must be real for any Hermitian operator.

It also has important consequences for the eigenstates. Suppose we have two different eigenstates with different eigenvalues

$$A\psi_q = q\psi_q \quad \text{and} \quad A\psi_{q'} = q'\psi_{q'}. \tag{76}$$

If A is Hermitian, we have

$$\int \psi^*_{q'} A \psi_q d\mathbf{r} = q \int \psi^*_{q'} \psi_q d\mathbf{r} = \int (A\psi_{q'})^* \psi_q d\mathbf{r} = q' \int \psi^*_{q'} \psi_q d\mathbf{r}. \tag{77}$$

The only way to satisfy this equation if $q \neq q'$ is to have

$$\int \psi^*_{q'} \psi_q d\mathbf{r} = 0. \tag{78}$$

This means that the wavefunctions are orthogonal to each other. Unfortunately, this theorem does not help us when eigenfunctions are degenerate. However, even it that case it is possible to make different eigenfunctions orthogonal to each other.

For matrix elements between two wavefunctions ψ_i and ψ_j (that are not necessarily eigenfunctions) for a Hermitian operator $A = A^\dagger$, we find

$$A_{ij}^* = (\langle\psi_i|A|\psi_j\rangle)^* = \left(\int d\mathbf{r}\psi_i^* A\psi_j\right)^* \tag{79}$$

$$= \int d\mathbf{r}\psi_j^* A^\dagger\psi_i = \int d\mathbf{r}\psi_j^* A\psi_i \tag{80}$$

$$= \langle\psi_j|A|\psi_i\rangle = A_{ji}. \tag{81}$$

Note that this implies that $A_{ii}^* = A_{ii}$ and therefore real. However, matrix elements between different states can be complex and of the form $A_{ij} = a + ib$ and $A_{ji} = a - ib$ where a and b are real numbers.

D. Free particles

The simplest problem that can be addressed in quantum mechanics are free particles in the absence of extermal potentials. The equation for a free particle moving in the x-direction is

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(\mathbf{r},t)}{\partial x^2} = i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} \tag{82}$$

The solutions are

$$\psi_k(x,t) = \frac{1}{\sqrt{L}}e^{ikx - i\varepsilon_k t/\hbar} \qquad \text{with an energy} \qquad \varepsilon_k = \frac{\hbar^2 k^2}{2m}, \tag{83}$$

where L is a normalization length, such that the integral $\int_L dx\psi^*\psi = 1$. This can be easily extended to three dimensions giving eigenstates

$$\psi_{\mathbf{k}}(\mathbf{r},t) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r} - i\varepsilon_{\mathbf{k}} t/\hbar} \qquad \text{with an energy} \qquad \varepsilon_{\mathbf{k}} = \frac{\hbar^2\mathbf{k}^2}{2m}, \tag{84}$$

where V is now a normalization volume, such that the integral $\int_V d\mathbf{r}\psi^*\psi = 1$. The running waves obtained from the differential equation are known as the eigenfunctions. We can also separate the position and time dependence in the eigenfunctions

$$\psi_k(\mathbf{r},t) = \psi_k(\mathbf{r})e^{-i\varepsilon_k t/\hbar} \qquad \text{with} \qquad \psi_k(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{85}$$

where the functions $\psi_k(\mathbf{r})$ are eigenfunctions of the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(\mathbf{r})}{\partial x^2} = E\psi(\mathbf{r}) \tag{86}$$

We see that for free particles, the eigenstates are determined by the wavevector $\mathbf{k}$. If we want to maintain running waves as eigenfunctions (as opposed to particles in a box where we take that the wavefunction is zero at the edges of the box), then we can apply periodic boundary conditions which state that

$$e^{ik(x+L)} = e^{ikx} \quad \Rightarrow \quad e^{ikL} = 1 \quad \Rightarrow \quad k = \frac{2\pi n}{L} \quad \text{where } n \text{ is an integer.} \tag{87}$$

Or in three dimensions

$$\mathbf{k} = \frac{2\pi}{L}(n_x\hat{\mathbf{x}} + n_y\hat{\mathbf{y}} + n_z\hat{\mathbf{z}}), \tag{88}$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are the Cartesian unit vectors. The choice of a rectangular box might appear somewhat arbitrary. However, we assume here that the normalization volume is large enough, such that the spacing between the levels

is unimportant for the problem that we are considering. Effectively, the particle should be free, that is we are not looking at a problem where a particle is localized inside a box. The energy is now

$$\varepsilon_{\mathbf{k}} = \frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2). \tag{89}$$

We see that the energy and wavefunction are directly related to the quantum numbers n_x, n_y, and n_z. These quantum numbers are intimately linked to the momentum $\mathbf{p} = \hbar\mathbf{k}$. Therefore, the operator $\mathbf{p} = -i\hbar\nabla$ is the relevant operator for the problem of a free particle. The importance of the momentum operator is clear when we operate it on the eigenfunctions in Eq. (84)

$$\mathbf{p}\psi_{\mathbf{k}}(\mathbf{r}) = -i\hbar\nabla\left(\frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}\right) = \hbar\mathbf{k}\psi_{\mathbf{k}}(\mathbf{r}). \tag{90}$$

The matrix elements of the momentum operator are given by

$$\langle\psi_{\mathbf{k}'}|\mathbf{p}|\psi_{\mathbf{k}}\rangle = \int d\mathbf{r}\ \psi^*_{\mathbf{k}'}(\mathbf{r})\mathbf{p}\psi_{\mathbf{k}}(\mathbf{r}) \tag{91}$$

$$= \hbar\mathbf{k}\frac{1}{V}\int d\mathbf{r}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = \hbar\mathbf{k}\delta_{\mathbf{k},\mathbf{k}'}, \tag{92}$$

where $\delta_{\mathbf{k},\mathbf{k}'}$ is one for $\mathbf{k} = \mathbf{k}'$ and zero otherwise. The wavevectors $\mathbf{k}$ label all the eigenstates. The eigenfunctions are obviously also eigenstates of the Hamiltonian

$$H\psi_{\mathbf{k}}(\mathbf{r}) = \frac{\hbar^2\nabla^2}{2m}\left(\frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}\right) = \varepsilon_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r}). \tag{93}$$

Later, we will find that for other problems, different operators are relevant. For example, for the hydrogen atom, the relevant operator is the angular momentum. The relevant operators are directly related to the symmetry of the problem. We are dealing here with free particles. We now want to shift our coordinate system. However, since space is translationally invariant, nothing should really change by that operation (the situation drastically changes in the presence of a potential). In classical mechanics, a translation of the coordinates can be easily made by adding a translation vector $\boldsymbol{\tau}$ to the position, giving $\mathbf{r}' = \mathbf{r} + \boldsymbol{\tau}$. In quantum mechanices, one is translating wavefunctions and one needs to find the proper function $U_{\boldsymbol{\tau}}$ that performs this operation. Operating this on the wavefunction should give

$$U_{\boldsymbol{\tau}}\psi(\mathbf{r}) = \psi(\mathbf{r} + \boldsymbol{\tau}). \tag{94}$$

The right-hand side can be expanded as

$$\psi(\mathbf{r} + \boldsymbol{\tau}) = \psi(\mathbf{r}) + \boldsymbol{\tau}\nabla\psi + \frac{1}{2!}(\boldsymbol{\tau})^2\nabla^2\psi + \cdots = \exp(\boldsymbol{\tau}\nabla)\,\psi(x). \tag{95}$$

We would like this transformation to be unitary. Unitary transformations essentially transform the system but leaves the norms unchanged. Essentially we are going to a new coordinate system. A good example is a rotation of the Cartesian coordinates system in three-dimensional space. Interesting things can be concluded if the unitary transformation leaves the system unchanged. This says something about the symmetry of the problem. For example, if the Hamiltonian does not change under a unitary transformation ($H' = U^\dagger HU = H$) then the eigenvalues also remain unchanged. For example, a translation does not do anything with the eigenenergies of a free particle

$$\int d\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r})H'\psi_{\mathbf{k}}(\mathbf{r}) = \int d\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r})U^\dagger HU\psi_{\mathbf{k}}(\mathbf{r}) \tag{96}$$

$$= \int d\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r}+\boldsymbol{\tau})H\psi_{\mathbf{k}}(\mathbf{r}+\boldsymbol{\tau}) = \varepsilon_{\mathbf{k}} \tag{97}$$

Now suppose we can write $U = e^{iL}$, then it is easy to see that U is unitary

$$U^\dagger U = e^{-iL^\dagger}e^{iL} = e^{i(L-L^\dagger)} = 1, \tag{98}$$

if L is Hermitian ($L^\dagger = L$) (For a translation this means that U translates the system in one direction and $U^\dagger$ does the same translation, but now in the opposite direction). If we take the momentum $\mathbf{p} = -i\hbar\nabla$, then $U_{\boldsymbol{\tau}} = \exp(\frac{i}{\hbar}\boldsymbol{\tau}\mathbf{p})$ is

unitary, since $\mathbf{p}$ is Hermitian. The invariance of the Hamiltonian under a certain unitary transformation has important consequences, such as the commutation of H and U,

$$UH = UU^\dagger HU = UH. \tag{99}$$

This can also be rewritten as

$$HU - UH = 0 \quad \Rightarrow \quad [H, U] = 0, \tag{100}$$

using the definition of the commutator of two operators

$$[A, B] = AB - BA \tag{101}$$

From this, one can show that an eigenstate ψ_n of H with eigenenergy E_n

$$H(U\psi_n) = UH\psi_n = UE_n\psi_n = E_n(U\psi_n) \tag{102}$$

must also be an eigenstate of U since $U\psi_n$ must be proportional to ψ_n since it is an eigenstate of H with eigenvalue E_n. This directly implies that $\mathbf{p}$ commutes with H and that momentum is a good quantum number when a system is translationally invariant. It also implies that the eigenfunctions of $\mathbf{p}$ are also eigenfunctions of H. Defining the eigenvalue of $\mathbf{p}$ as $\hbar\mathbf{k}$, we obtain the usual plane waves

$$\mathbf{p}\psi_\mathbf{k}(\mathbf{r}) = \hbar\mathbf{k}\psi_\mathbf{k}(\mathbf{r}) \quad \Rightarrow \quad -i\hbar\nabla\psi_\mathbf{k}(\mathbf{r}) = \hbar\mathbf{k}\psi_\mathbf{k}(\mathbf{r}) \quad \Rightarrow \quad \psi_\mathbf{k}(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{103}$$

where V is a normalization volume. So we see that the eigenfunctions of the momentum $\mathbf{p}$ are the same as the eigenfunctions of the Schrödinger equation in the absence of a potential.

E. Eigenfunctions and basis sets

The momentum $\mathbf{p}$ plays a special role when a system has translational symmetry. However, this is no longer the case when the translational symmetry is broken. For example, when a local scattering potential $U(\mathbf{r})$ is placed in the system, momentum is no longer a good quantum number. An electron in a particular plane wave can scatter to a diffent wavefunction. This implies that the matrix element

$$\langle\psi_{\mathbf{k}'}|U|\psi_\mathbf{k}\rangle = \int d\mathbf{r}\ \psi^*_{\mathbf{k}'}(\mathbf{r})U(\mathbf{r})\psi_\mathbf{k}(\mathbf{r}) \neq 0. \tag{104}$$

Although, the plane waves are not eigenfunctions of the system with the local potential, it can still be useful to use the plane waves as a basis set and express the eigenfunctions in terms of plane waves. Since the plane waves form a complete basis (there are an infinite number of them), one can express any function in terms of plane waves. This is the underlying principle of Fourier series. One has to make a clear distinction between a basis set and eigenfunctions. A basis set is a human choice, a decision to use certain functions to express the eigenfunctions. We shall see that the eigenfunctions of the hydrogen atom are expressed in terms of Laguerre polynomials and spherical harmonics. Obviously, the eigenfunctions are a basis set by themselves (again there are an infinite number of them). It is a very convenient basis set since they are eigenstates of the potential of the nucleus. However, when things are added to the Hamiltonian (say, an electric field or the spin-orbit coupling), the spherical harmonics are no longer eigenfunctions. However, they are still a very convenient basis set. The choice of basis set can be very important. For example, trying to describe the hydrogen atom in terms of plane waves is possible, but very inconvenient. Vice versa expressing plane waves in terms of spherical harmonics is also possible. You might laugh and think that noone would want to do that, but this is actually done when one wants to express the interaction of a electromagnetic plane wave with an atom. To better understand the transitions for absorption and emission, it actually makes sense to express the plane wave in terms of spherical harmonics and Bessel functions. We will return to the concepts of basis sets and eigenfunctions when discussing the matrix mechanics developed by Heisenberg (1901-1976), Born, and Jordan (1902-1980).

F. Wave packet

We found that the eigenfunctions for the Schrödinger equation in the absence of a potential are plane waves. That means that although they have a very specific momentum, they are completely delocalized in space. This does not

necessarily correspond to our idea of a particle. So let us try to localize our particle by putting a Gaussian envelope of width a around the probability distribution function. So, for $t = 0$,

$$\Psi(\mathbf{r}, 0) = \frac{1}{\sqrt{a\sqrt{2\pi}}} \exp(-\frac{x^2}{4a^2}) e^{ik_0 x} \qquad \Rightarrow \qquad P(x,0) = \Psi^*(x,0)\Psi(x,0) = \frac{1}{a\sqrt{2\pi}} \exp(-\frac{x^2}{2a^2}). \tag{105}$$

So our particle is nicely localized in space and we have given it the momentum k_0, we think. However, let us look at the Fourier transform

$$b(k) = \frac{1}{\sqrt{2\pi}} \int dx \Psi(x,0) e^{-ikx} = \frac{1}{\sqrt{a}(2\pi)^{3/4}} \int dx \exp(-\frac{x^2}{4a^2}) e^{i(k_0 - k)x} = \sqrt{\frac{2a}{\sqrt{2\pi}}} \exp(-a^2(k-k_0)^2). \tag{106}$$

We thus find that the distribution is not a delta function in k space at k_0, but a Gaussian with width $1/2a$ around k_0. So we have

$$\Delta x \Delta p = \Delta x \hbar \Delta k = a\hbar \frac{1}{2a} = \frac{\hbar}{2}. \tag{107}$$

Apparently it is difficult to localize a wavefunction both in momentum and real space. This will be shown more generally below. This also implies that we cannot say that the particle has a momentum $\hbar k_0$. There is a probability distribution of momenta. The momentum that we measure differs from one experiment to the other. We can, however, speak of an average momentum. The expectation value of an operator $\hat{A}$ is defined as

$$\langle \hat{A} \rangle = \int dx \Psi^* \hat{A} \Psi. \tag{108}$$

So if we could do N experiments, for large N, $\langle \hat{A} \rangle = \frac{1}{N} \sum_i A_i$, where A_i is the value of A in the i'th experiment. For the Gaussian wave packet, we have

$$\langle \hat{p}_x \rangle = \hbar \int dx \Psi^* (k_0 + i\frac{x}{2a^2}) \Psi = \hbar k_0. \tag{109}$$

Let us look briefly at the time development of a Gaussian wave packet (this is treated in more detail in Liboff 6.1). Since each k value has a different energy $E = \hbar\omega = \frac{\hbar^2 k^2}{2m}$, it is more convenient to make the time development in k space and then make a Fourier transform back to real space

$$\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk b(k) \exp(i(kx - \omega t)) \tag{110}$$

Using Eqn. (106), we find

$$\Psi(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dx' \Psi(x',0) e^{-ikx'} \exp(i(kx - \omega t)) = \int_{-\infty}^{\infty} dx' \Psi(x',0) K(x,x',t), \tag{111}$$

with

$$K(x,x',t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \exp(i(k(x-x') - \omega t)) \tag{112}$$

describing the time development of the wavefunction. After integration (see Liboff 6.1), we find for the distribution function at time t

$$P(x,t) = \frac{1}{a'\sqrt{2\pi}} \exp\left(-\frac{(x - v_g t)^2}{2a'^2}\right) \qquad \text{with} \qquad a' = a\sqrt{1 + \frac{t^2\omega^2}{k_0^4 a^4}} \tag{113}$$

and the group velocity is $v_g = \hbar k_0/m$. Therefore, the distribution remains Gaussian, but its width, given by a', increases in time.

FIG. 9: Heisenberg's uncertainty principle has entered the popular imagination, but in the end it is a property of waves and operators.

G. Heisenberg's uncertainty principle

We saw above with the Gaussian wavepacket that it is difficult to localize a particle both in real and momentum space. This is a general aspect of waves and is well known to people familiar with Fourier analysis. A plane wave (i.e., localized in momentum space) is delocalized over the entire space, whereas a δ-function can be represented by an integration over all k values $\delta(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} dx$. The bottom line is that it is impossible to localize a particle without changes the derivatives with respect to x (and hence the momentum) and it is also impossible to reduce the spread in momentum without delocalizing the wavefunction. This can be proven more generally. We know that

$$I(\lambda) = \int_{-\infty}^{\infty} dx \left| x\psi + \lambda\hbar \frac{\partial \psi}{\partial x} \right|^2 \geq 0 \tag{114}$$

for any λ. Thus

$$I(\lambda) = \int_{-\infty}^{\infty} dx \left[|x\psi|^2 + \lambda\hbar (\frac{\partial \psi^*}{\partial x} x\psi + x\psi^* \frac{\partial \psi}{\partial x}) + \lambda^2 \hbar^2 \left| \frac{\partial \psi}{\partial x} \right|^2 \right]. \tag{115}$$

The term with λ can be evaluated by partial integration

$$\int dx (\frac{\partial \psi^*}{\partial x} x\psi + x\psi^* \frac{\partial \psi}{\partial x}) = \int dx x \frac{\partial}{\partial x} (\psi^* \psi) = [x\psi^* \psi] - \int dx \psi^* \psi, \tag{116}$$

where we choose the integration boundaries in such a way that the first term on the right-hand side is zero. Similarly, we can write for the term with λ^2

$$\int dx \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x} = \left[\psi^* \frac{\partial \psi}{\partial x} \right] - \int dx \psi^* \frac{\partial^2 \psi}{\partial x^2}. \tag{117}$$

This gives a total result

$$I(\lambda) = \int_{-\infty}^{\infty} dx \psi^* x^2 \psi - \lambda\hbar \int_{-\infty}^{\infty} dx \psi^* \psi + \lambda^2 (-i\hbar)^2 \int_{-\infty}^{\infty} dx \psi^* \frac{\partial^2 \psi}{\partial x^2}, \tag{118}$$

or

$$I(\lambda) = \langle x^2 \rangle - \lambda\hbar + \lambda^2 \langle p_x^2 \rangle \geq 0. \tag{119}$$

Since this quadratic equation in λ has no solutions, this means that the determinant should be less than zero,

$$\hbar^2 - 4\langle x^2 \rangle \langle p_x^2 \rangle \leq 0 \qquad \Rightarrow \qquad \langle x^2 \rangle \langle p_x^2 \rangle \geq \frac{\hbar^2}{4}. \tag{120}$$

If we look at deviations from an average value, we have

$$\Delta x = \sqrt{\langle x^2\rangle - \langle x\rangle^2} \qquad \text{and} \qquad \Delta p = \sqrt{\langle p_x^2\rangle - \langle p_x\rangle^2}. \tag{121}$$

We can always choose the wavefunction in such a way that $\langle x\rangle = \langle p\rangle = 0$. This gives us the result

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \tag{122}$$

which is known as the Heisenberg uncertainty principle.

Let us have a look again at the term with λ which apparently prevents us from making $\Delta x \Delta p_x = 0$. We can rewrite this slightly using

$$\int dx \frac{\partial \psi^*}{\partial x} x\psi = [\psi^* x\psi] - \int dx \psi^* \frac{\partial}{\partial x}(x\psi), \tag{123}$$

giving, using Eqn. (116),

$$\hbar \int dx \psi^* \left[x\frac{\partial}{\partial x} - \frac{\partial}{\partial x} x \right] \psi = -\hbar \int dx \psi^* \psi = -\hbar. \tag{124}$$

This can also be written as

$$x\left(-i\hbar\frac{\partial \psi}{\partial x}\right) - \left(-i\hbar\frac{\partial}{\partial x}\right)(x\psi) = i\hbar\psi. \tag{125}$$

or

$$xp_x - p_x x = i\hbar \qquad \text{or} \qquad [x, p_x] = i\hbar, \tag{126}$$

with the commutator defined as $[A, B] = AB - BA$. Note, the intricate relationship between the commutation relationship and the Heisenberg uncertainty principle. If two operators do not commute we cannot have ΔA and ΔB zero at the same time.

IV. MATRIX MECHANICS

Matrix mechanics was the first correct representation of quantum mechanics. It was developed by Werner Heisenberg together with Max Born and Pacual Jordan. At first, it was not immediately accepted since matrix mechanics were new concepts in physics at the time. The later introduction of wave mechanics by Schrödinger was generally preferred, since the concept of waves was much more familiar. However, as we shall see, the two formulations are in fact equivalent.

A. Matrix mechanics: Two-level system

Let us consider a simple system consisting of only two levels. Examples of such systems are two spin levels (spin up and spin down), the hydrogen molecule when only the $1s$ levels are considered, etc. Let us consider an electron hopping between two levels. Now let us assume we know the wavefunctions $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ of the two levels. Knowing the wavefunctions and the Hamiltonian allows us to calculate the energy when an electron is in a particular wavefunction

$$\varepsilon_i = \langle \varphi_i | H | \varphi_i \rangle = \int d\mathbf{r}\ \varphi_i^*(\mathbf{r}) H \varphi_i(\mathbf{r}) \quad \text{for} \quad i = 1, 2 \tag{127}$$

Now if there was no coupling between states 1 and 2, then the wavefunctions φ_1 and φ_2 are the eigenfunctions of the system, with eigenenergies ε_1 and ε_2, respectively. However, if there is a coupling

$$t = \langle \varphi_2 | H | \varphi_1 \rangle = \int d\mathbf{r}\ \varphi_2^*(\mathbf{r}) H \varphi_1(\mathbf{r}) \tag{128}$$

$$= \langle \varphi_1 | H | \varphi_2 \rangle^* = \left(\int d\mathbf{r}\ \varphi_1^*(\mathbf{r}) H \varphi_2(\mathbf{r}) \right)^*, \tag{129}$$

then this is no longer correct. However, we can use the wavefunctions as a basis to express the eigenstates of the system. The eigenfunctions can be expressed as a linear combination:

$$\psi(\mathbf{r}) = a\varphi_1(\mathbf{r}) + b\varphi_2(\mathbf{r}). \tag{130}$$

The Schrödinger equation reads now

$$H\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad \Rightarrow \quad H[a\varphi_1(\mathbf{r}) + b\varphi_2(\mathbf{r})] = E[a\varphi_1(\mathbf{r}) + b\varphi_2(\mathbf{r})]. \tag{131}$$

We can multiply this equation from the left by $\varphi_1^*(\mathbf{r})$ and integrate over $\mathbf{r}$. This gives

$$a\int d\mathbf{r}\ \varphi_1^*(\mathbf{r})H\varphi_1(\mathbf{r}) + b\int d\mathbf{r}\ \varphi_1^*(\mathbf{r})H\varphi_2(\mathbf{r}) = Ea\int d\mathbf{r}\ \varphi_1^*(\mathbf{r})\varphi_1(\mathbf{r}) + Eb\int d\mathbf{r}\ \varphi_1^*(\mathbf{r})\varphi_2(\mathbf{r}). \tag{132}$$

Using the orthonormality of the wavefunctions, we find

$$a\langle\varphi_1|H|\varphi_1\rangle + b\langle\varphi_1|H|\varphi_2\rangle = Ea. \tag{133}$$

Similarly, for multiplication with $\varphi_2^*(\mathbf{r})$,

$$a\langle\varphi_2|H|\varphi_1\rangle + b\langle\varphi_2|H|\varphi_2\rangle = Eb. \tag{134}$$

We can combine these two equation by making use of matrices

$$\begin{pmatrix} \langle\varphi_1|H|\varphi_1\rangle & \langle\varphi_1|H|\varphi_2\rangle \\ \langle\varphi_2|H|\varphi_1\rangle & \langle\varphi_2|H|\varphi_2\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = E \begin{pmatrix} a \\ b \end{pmatrix} \tag{135}$$

Replacing the matrix elements by the symbols from Eqs. (127) and (129) gives a Hamiltonian

$$\mathrm{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} \varepsilon_1 & t \\ t^* & \varepsilon_2 \end{pmatrix}, \tag{136}$$

where we use non-italic letters to indicate matrices and $H_{ij} = \langle\varphi_i|H|\varphi_j\rangle$. An important thing to not is that the Hamiltonian is completely independent of the spatial coordinates, since all the $\mathbf{r}$ dependence has been integrated out in deriving the parameters ε_i and t. Therefore, we managed to recast the whole problem into something that only depends on numbers.

In Eq. (135), the wavefunction is written as a 2×1 column

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{137}$$

Again, there is no $\mathbf{r}$ dependence. However, this can be easily obtained if we want to since the whole problem is expressed in the two basis functions $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$. Once we know how much is in each of the basis functions (given by the coefficients a and b, we can always obtain the spatial dependence of the wavefunction. However, in many problems we are only interested in the coefficients (or the occupations $|a|^2$ and $|b|^2$) and not even in the exact spatial dependence of the wavefunction. Especially, for complex problems, the spatial dependence of the wavefunction can sometimes be rather uninformative. Essentially, the wavefunction is vector in a two-dimensional space, expressed in two basis vectors. To bring out the character of the different components of the basis, we can make the following identifications

$$|\varphi_1\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\varphi_2\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{138}$$

The wavefunction is now expressed as

$$|\psi\rangle = a|\varphi_1\rangle + b|\varphi_2\rangle. \tag{139}$$

We can use this notation also to express other properties of the wavefunction. For example, the normalization of a wavefunction expressed in two basis functions is given by

$$\langle\psi|\psi\rangle = \int d\mathbf{r}\psi^*(\mathbf{r})\psi(\mathbf{r}), \tag{140}$$

following the notation in Eq. (70). Using Eq. (130), we find

$$\langle\psi|\psi\rangle = |a|^2 \int d\mathbf{r}\varphi_1^*(\mathbf{r})\varphi_1(\mathbf{r}) + |b|^2 \int d\mathbf{r}\varphi_2^*(\mathbf{r})\varphi_2(\mathbf{r}) + a^*b \int d\mathbf{r}\varphi_1^*(\mathbf{r})\varphi_2(\mathbf{r}) + b^*a \int d\mathbf{r}\varphi_2^*(\mathbf{r})\varphi_1(\mathbf{r}), \tag{141}$$

where the latter two are zero due to the orthogonality of the basis functions. The first to integral are unity, and we therefore have

$$\langle\psi|\psi\rangle = |a|^2 + |b|^2. \tag{142}$$

So far, we have not really specificied what $\langle\psi|$ is. However, let us have a look at what is known in mathematics as the conjugate transposed or Hermitian conjugate. Given a matrix A, with matrix elements A_{ij}, the matrix elements of the Hermitian conjugate matrix $\mathrm{A}^\dagger$ are given by

$$A_{ij}^\dagger = A_{ji}^*. \tag{143}$$

For the vector $\langle\psi|$, we then find

$$\langle\psi| = (|\psi\rangle)^\dagger = \begin{pmatrix} a \\ b \end{pmatrix}^\dagger = (a^* \; b^*) \equiv (a^*, b^*), \tag{144}$$

where the comma in the conjugate of a vector is conventional. Note that the conjugate of the vector $|\psi\rangle$ (a 2×1 matrix) is a 1×2 matrix. Let us now look at matrix multiplication. The matrix elements of the product of a matrix A of dimension $m \times l$ and a matrix B of dimension $l \times n$ are

$$(\mathrm{AB})_{ij} = \sum_{k=1}^{l} A_{ik}B_{kj}. \tag{145}$$

Effectively, we are taking the inner product of the rows of A with the columns of B. The resulting matrix is of dimensions AB is a $m \times n$. Interpreting, $\langle\psi|\psi\rangle$ as a matrix multiplication of $\langle\psi|$ and $|\psi\rangle$, we find

$$\langle\psi|\psi\rangle = (a^*, b^*) \begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 + |b|^2, \tag{146}$$

which is the same result as Eq. (142). The multiplication of a 1×2 matrix and a 2×1 matrix gives a 1×1 matrix, which is essentially a number. The orthogonality of the basis function is directly satisfied

$$\langle\varphi_1|\varphi_1\rangle = (1, 0) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \quad \text{and} \quad \langle\varphi_2|\varphi_1\rangle = (0, 1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \tag{147}$$

B. Eigenvalues and eigenvectors

Solving the Schrödinger equation now becomes equivalent to finding the solution to the equation

$$H|\psi\rangle = E|\psi\rangle \quad \Rightarrow \quad \begin{pmatrix} \varepsilon_1 & t \\ t & \varepsilon_2 \end{pmatrix} |\psi\rangle = E|\psi\rangle, \tag{148}$$

where for brevity we assume that t is a real number. That's nice, but what do we do now? How do we find the eigenvalues E? Let us define the unit matrix

$$\mathrm{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{149}$$

where the subscript 2 indicates the dimension. Using this we can rewrite Eq. (148) as

$$(\mathrm{H} - E\ \mathrm{I}_2)|\psi\rangle = 0, \tag{150}$$

where $\mathrm{H} - E\ \mathrm{I}_2$ is just a new matrix. Now let us assume we can find and inverse of this matrix such that

$$(\mathrm{H} - E\ \mathrm{I}_2)^{-1}(\mathrm{H} - E\ \mathrm{I}_2) = \mathrm{I}_2. \tag{151}$$

Multiplying Eq. (150) on both sides by the inverse gives

$$(\mathrm{H} - E\ \mathrm{I}_2)^{-1}(\mathrm{H} - E\ \mathrm{I}_2)|\psi\rangle = \mathrm{I}_2|\psi\rangle = |\psi\rangle = 0. \tag{152}$$

Therefore if there is an inverse for $\mathrm{H} - E\ \mathrm{I}_2$ then Eq. (150) can only be satisfied by having $|\psi\rangle = 0$ which is a trivial solution but not something that is normalized to unity. Now the inverse of any 2×2 matrix can be found

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad \mathrm{A}^{-1} = \frac{1}{ad - cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{153}$$

unless $ad - cb = 0$. This quantity is also known as the determinant

$$\mathrm{detA} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{154}$$

This gives a condition for our eigenvalue problem and the eigenenergies can be found from

$$\begin{vmatrix} \epsilon_1 - E & t \\ t & \epsilon_2 - E \end{vmatrix} = (\epsilon_1 - E)(\epsilon_2 - E) - t^2 \tag{155}$$

$$= E^2 - (\epsilon_1 + \epsilon_2)E + \epsilon_1\epsilon_2 - t^2 = 0. \tag{156}$$

This quadratic equation is easily solved and gives two eigenvalues

$$E_\pm = \frac{\epsilon_1 + \epsilon_2}{2} \pm \sqrt{\left(\frac{\epsilon_2 - \epsilon_2}{2}\right)^2 + t^2}. \tag{157}$$

We can write the solution of the 2×2 eigenvalue problem in a variety of ways. This is not crucial physics, but it is convenient for a variety of problems. Let us introduce some shorthand notations

$$\overline{\epsilon} = \frac{\epsilon_1 + \epsilon_2}{2} \tag{158}$$

$$\Delta = \epsilon_2 - \epsilon_1 \ \text{ and } \ \delta = \frac{\Delta}{2} \tag{159}$$

$$\mathcal{E} = \sqrt{\delta^2 + t^2} \tag{160}$$

which allows us to write

$$E_\pm = \overline{\epsilon} \pm \mathcal{E}. \tag{161}$$

The eigenvalues for a 2×2 matrix can always be written as

$$|\psi_-\rangle = \cos\theta|1\rangle - \sin\theta|2\rangle \tag{162}$$

$$|\psi_+\rangle = \sin\theta|1\rangle + \cos\theta|2\rangle. \tag{163}$$

Let us determine the value of θ from the equation

$$t\cos\theta + (\epsilon_2 - E_-)(-\sin\theta) = t\cos\theta - [\epsilon_2 - (\overline{\epsilon} - \mathcal{E})]\sin\theta = t\cos\theta - (\mathcal{E} + \delta)\sin\theta = 0,$$

giving

$$\tan\theta = \frac{t}{\mathcal{E} + \delta}. \tag{164}$$

We can also express this as

$$\tan 2\theta = \frac{2\tan\theta}{1 - 2\tan^2\theta} = \frac{2\frac{t}{\mathcal{E}+\delta}}{1 - \left(\frac{t}{\mathcal{E}+\delta}\right)^2} = \frac{2t(\mathcal{E} + \delta)}{(\mathcal{E} + \delta)^2 - t^2} = \frac{2t(\mathcal{E} + \delta)}{\mathcal{E}^2 + 2\mathcal{E}\delta + \delta^2 - t^2}. \tag{165}$$

Using that $\mathcal{E}^2 = \delta^2 + t^2$, we obtain

$$\tan 2\theta = \frac{2t(\mathcal{E} + \delta)}{2\mathcal{E}\delta + 2\delta^2} = \frac{t}{\delta} = \frac{2t}{\Delta} = \frac{2t}{\varepsilon_2 - \varepsilon_1} \tag{166}$$

We can use this to obtain

$$\cos 2\theta = \frac{\delta}{t^2+\delta^2} = \frac{\delta}{\mathcal{E}}. \tag{167}$$

Which we can use to express

$$\begin{aligned} \cos^2\theta &= \frac{1}{2} + \frac{1}{2}\cos 2\theta = \frac{1}{2}\left(1+\frac{\delta}{\mathcal{E}}\right) = \frac{1}{2}\left(1+\frac{\Delta}{\sqrt{\Delta^2+4t^2}}\right) \\ \sin^2\theta &= \frac{1}{2} - \frac{1}{2}\cos 2\theta = \frac{1}{2}\left(1-\frac{\delta}{\mathcal{E}}\right) = \frac{1}{2}\left(1-\frac{\Delta}{\sqrt{\Delta^2+4t^2}}\right) \end{aligned} \tag{168}$$

Let us consider some limiting cases. When $t = 0$, we find that $E_- = \varepsilon_1$ and $E_+ = \varepsilon_2$ (where the minus sign indicates the lowest energy and assuming $\varepsilon_1 < \varepsilon_2$. Since $\theta = 0$, we find that the basis functions are directly the eigenfunctions. In the limit, $\varepsilon_1 = \varepsilon_2 = \varepsilon$, the eigenvalues are $\varepsilon \pm t$ and the eigenvectors are $\frac{1}{\sqrt{2}}(|\varphi_1\rangle \pm |\varphi_2\rangle)$.

After we find the eigenvalues and eigenvectors, one might wonder if we could have chosen a different and more convenient basis set. Obviously it would be nice to work in the basis given by (163), which is the basis of the eigenvectors. It is often useful to make transformations from one basis set to another. The transformations are given by the expressions for the eigenfunctions in Eq. (163). This can also be expressed in terms of a matrix

$$\mathrm{U} = \mathrm{U}_{be} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{169}$$

The tricky part is making sure that the unitary transformation goes in the right direction. Usually in textbooks, no subscripts are given, but they aid in seeing between what bases the unitary transformation is. U_{be} is the unitary transformation that expresses the eigenvectors (e) in terms of the original basis vectors (b) (you have to read the indices from right to left, because that is were the matrix multiplications are going). Therefore, applying $\mathrm{U} = \mathrm{U}_{be}$ on the lowest eigenvector $|\psi_-\rangle$ gives

$$\mathrm{U}_{be}|\psi_-\rangle = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix}. \tag{170}$$

The inverse U_{eb} expresses the basis functions in terms of the eigenfunctions and is given by the Hermitian conjugate

$$\mathrm{U}^\dagger = \mathrm{U}_{eb} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{171}$$

Those somewhat familiar with transformation will have noticed that $\mathrm{U}^\dagger = \mathrm{U}_{eb}$ is a rotation of the coordinate system by an angle θ. Its Hermitian conjugate $\mathrm{U} = \mathrm{U}_{be}$ is a rotation over $-\theta$. It appears obvious that the two operations should cancel each other

$$\mathrm{U}^\dagger\mathrm{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos^2\theta+\sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta+\cos^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2$$

A matrix element between two wavefunctions in the original basis set can be rewritten as

$$\langle\psi'_b|\mathrm{H}|\psi_b\rangle = \langle\psi_b|\mathrm{U}_{be}\mathrm{U}_{eb}\mathrm{H}\mathrm{U}_{be}\mathrm{U}_{eb}|\psi_b\rangle = \langle\psi'_e|\mathrm{U}_{eb}\mathrm{H}\mathrm{U}_{be}|\psi_e\rangle, \tag{172}$$

using that $|\psi_e\rangle = U_{eb}|\psi_b\rangle$, expressing the wavefunction in the basis of eigenfunctions. The Hamiltonian H' in this basis is given by

$$\mathrm{H}' \equiv \mathrm{H}_{ee} = \mathrm{U}_{eb}\mathrm{H}\mathrm{U}_{be} = \mathrm{U}^\dagger\mathrm{H}\mathrm{U}. \tag{173}$$

Note that the original Hamiltonian is expressed in the original basis set $\mathrm{H} \equiv \mathrm{H}_{bb}$. The evaluation of this matrix is a bit of a pain

$$\mathrm{U}^\dagger\mathrm{H}\mathrm{U} = \begin{pmatrix} \cos^2\theta\varepsilon_1 + \sin^2\theta\varepsilon_2 - 2\sin\theta\cos\theta & \sin\theta\cos\theta(\varepsilon_1-\varepsilon_2) + t(\cos^2\theta-\sin^2\theta) \\ \sin\theta\cos\theta(\varepsilon_1-\varepsilon_2) + t(\cos^2\theta-\sin^2\theta) & \sin^2\theta\varepsilon_1 + \cos^2\theta\varepsilon_2 - 2\sin\theta\cos\theta \end{pmatrix}. \tag{174}$$

This looks like a horrible expression, however

$$\sin\theta\cos\theta(\varepsilon_1-\varepsilon_2) + t(\cos^2\theta-\sin^2\theta) = -\frac{\Delta}{2}\sin 2\theta + t\cos 2\theta = 0, \tag{175}$$

because of Eq. (166). Furthermore,

$$\cos^2\theta\varepsilon_1 + \sin^2\theta\varepsilon_2 - 2\sin\theta\cos\theta = \frac{\varepsilon_1+\varepsilon_2}{2} + \frac{\varepsilon_1-\varepsilon_2}{2}\frac{\Delta}{\sqrt{\Delta^2+4t^2}} - 2t\sqrt{\frac{1}{4}\left(1-\frac{\Delta^2}{\Delta^2+4t^2}\right)} \tag{176}$$

because of Eq. (168). Now since $\Delta = \varepsilon_2 - \varepsilon_1$), this can also be written as

$$\frac{\varepsilon_1+\varepsilon_2}{2} - \frac{1}{2}\frac{\Delta^2}{\sqrt{\Delta^2+4t^2}} - 2t\frac{1}{2}\frac{2t}{\sqrt{\Delta^2+4t^2}} = \frac{\varepsilon_1+\varepsilon_2}{2} - \frac{1}{2}\sqrt{\Delta^2+4t^2} = E_-. \tag{177}$$

The other diagonal matrix element is given by E_+. The total Hamiltonian in the eigenvector basis is therefore given by

$$\mathrm{H}' = \mathrm{H}_{ee} = \begin{pmatrix} E_- & 0 \\ 0 & E_+ \end{pmatrix}, \tag{178}$$

as it should be, since $\mathrm{H}'|\psi_\pm\rangle = E_\pm|\psi_\pm\rangle$.

In the end, you might wonder how we are ever going to get somewhere like this, since even solving a 2×2 matrix is quite a bit of work. However, this is because we are doing it analytically, where we can do a 2×2 matrix and maybe a 3×3. When the parameters take on actual values there are a wide variety of codes available to solve eigenvalue problems so this is not really a problem. What is a problem is the size. Finding the complete eigenenergies and eigenvectors can be done (say on your labtop) for matrices with a dimension of the order of a few thousand. This might seem like a lot, but it is actually for many problems rather small. Using approximate methods, one can deal with systems with a dimension of several million, but even that is often not sufficient for real problems.

When looking at the basis vectors, we found that their orthogonality is given by $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$. However, what happens when we multiply the vectors the other way around? This gives, for example,

$$|\varphi_1\rangle\langle\varphi_1| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad |\varphi_1\rangle\langle\varphi_2| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{179}$$

where the multiplication of a 2×1 and a 1×2 matrix gives a 2×2 matrix. This gives some nice features. For example,

$$\sum_{i=1}^{2} |\varphi_i\rangle\langle\varphi_i| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2, \tag{180}$$

this is a completeness relation, since this only works if we sum over all the wavefunctions $|\varphi_i\rangle$ in the basis. Another interesting feature is that one can build any possible matrix using all the combinations of $|\varphi_i\rangle\langle\varphi_j|$, or

$$\mathrm{H} = \sum_{i,j=1}^{2} H_{ij}|\varphi_i\rangle\langle\varphi_j| \tag{181}$$

$$= \sum_{i,j=1}^{2} |\varphi_i\rangle\langle\varphi_i|H|\varphi_j\rangle\langle\varphi_j|. \tag{182}$$

Operating this on $|\varphi_k\rangle$ gives

$$\mathrm{H}|\varphi_k\rangle = \sum_{i,j=1}^{2} H_{ij}|\varphi_i\rangle\langle\varphi_j|\varphi_k\rangle = \sum_{i}^{2} |\varphi_i\rangle\langle\varphi_i|H|\varphi_k\rangle. \tag{183}$$

using the orthogonality relation. Therefore the effect on the Hamiltonian is as follows. First, it removes the state $|\varphi_k\rangle$. This state is then scattering into all states $|\varphi_i\rangle$ with a scattering amplitude given by the matrix element $\langle\varphi_i|H|\varphi_k\rangle$.

C. Basis functions and matrix mechanics

In this section, we generalize the matrix algebra introduced in the previous section. Fortunately, most of it consists of changing the basis from two to any arbitrary number. Matrix algebra depends on choosing a particular basis

consisting of wavefunctions $\varphi_i(\mathbf{r})$. This is often called a Hilbert space. Again, the difference between the basis functions and eigenfunctions should be stressed. Basis functions are a human choice: a set of wavefunctions that we feel are good to describe a certain problem. Eigenfunctions are determined by the physics of the problem. We can approach the same problem with different bases, but in the end the eigenfunctions should be the same. In addition, wavefunctions that are eigenfunctions in one problem can be basis functions in another. For example, plane waves are eigenfunctions of free particles. However, as soon as we switch on some sort of potential that depends on the position, the plane waves are no longer eigenfunctions. We can still use them as a basis. The index/label i can be very general and can be split into several indices. For example, for a basis set of plane waves, the functions are labelled by the momentum $\mathbf{k}$. If spin is included in the problem, the label will become $i \rightarrow \mathbf{k}\sigma$ where $\sigma =\uparrow,\downarrow$ for up and down spins, respectively. For atomic orbitals, the index splits into three indices n, l, and m that label the atomic orbitals. There are very good reasons for choosing indices like that, but in principle, we can label everything with an index i.

A basis has to satisfy, in principle, several conditions. The basis functions have to be orthonormal

$$\langle\varphi_i|\varphi_j\rangle = \delta_{ij}. \tag{184}$$

Some of this is just for convenience and/or mathematical elegance. For example, there is no fundamental problem having the wavefunctions normalized as $\langle\varphi_i|\varphi_i\rangle = a_i$, where a_i is a real number. It is just a pain to keep track of the norms. Since we have the choice, we might as well take $a_i \equiv 1$. Also, people have worked with bases where $\langle\varphi_i|\varphi_j\rangle \neq 0$, for example, in applications involving multiple atoms where the atomic orbitals of one atom are not orthogonal to those of the other atom. Again, this is possible, although it introduces a lot of practical issues.

The second mathematical condition is that the bases are complete, in that we should be able to describe any possible wavefunction. For example, the basis of plane waves in Eq. (84) is complete if all possible wavevectors $\mathbf{k}$ values are included. Again, this is a mathematical condition that physicists like to break usually for obvious practical reasons. It is nice to say that your basis has to be complete and all basis functions have to be taken into account, but there are an infinite number of plane waves and that is just a few too many. Many problems can be solved by restricting oneself to a limited number of basis functions. Unfortunately, many times this restriction has to be done for purely mathematical reasons.

In the end, we should be able to express the eigenfunctions in terms of the basis functions. In the case of planes waves, this gives

$$\psi_n(\mathbf{r}) = \sum_{i=1}^{N} a_i\varphi_i(\mathbf{r}), \tag{185}$$

where $\varphi_{\mathbf{k}}(\mathbf{r})$ is our chosen basis set denoted by quantum numbers i. Our eigenvectors satisfy the Schrödinger equation

$$H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}). \tag{186}$$

with eigenvalues E_n are indexed by n. Restricting our basis to N functions (for example by taking only N $\mathbf{k}$-values into account) directly restricts the number of possible eigenfunctions to N. This is directly related to the fact that going from the our choice of the basis to the eigenfunction basis is a unitary transformation (just as in the previous example for the two-dimensional problem), which does not change the size of the basis. This might bother you at first. However, this is not disastrous if we think about it a little longer. For example, for atoms we talk about different shells. You probably heard of s, p, d, and f shells. However, if we really consider all basis functions, the number of atomic orbitals would be infinite. However, in practice we are mainly interested in the orbitals that we want to put electrons in. Since the atomic number of atoms goes to about 100, a same number of orbitals is needed to describe atoms. Often we can do with even less. For hydrogen ($1s^1$) and helium ($1s^2$), the presence of $4f$ orbitals is generally not relevant for most properties.

The good choice of a basis depends very much on the system that you are studying and its symmetry. For spherical systems, spherical harmonics are a good starting point. For systems with mainly translational symmetry, plane wave are an obvious choice. However, sometimes the choice is not that clear. A solid for example has translational symmetry, but, due to the presence of atoms, many points with spherical symmetry. This has led to the development of bases where spherical harmonics are centered on the atomic positions, but the regions between the atoms are made up of plane waves. Although such bases can be convenient, they also lead to some arbitrariness, such as, where does one switch from atomic regions to plane wave regions? However, such issues are rather advanced and should not bother us for the moment.

The matrix element between to wavefunctions $\Psi(\mathbf{r}) = \sum_{i=1}^{N} a_i\varphi_i(\mathbf{r})$ and $\Phi(\mathbf{r}) = \sum_{i=1}^{N} b_i\varphi_i(\mathbf{r})$ is given by

$$\langle\Psi|A|\Phi\rangle = \sum_{j=1}^{N}\sum_{i=1}^{N}\int d\mathbf{r} b_i^*\varphi_i^*(\mathbf{r})Aa_j\varphi_j(\mathbf{r}) = \sum_{j=1}^{N}\sum_{i=1}^{N} b_j^*A_{ij}a_i \tag{187}$$

where the matrix element is given by

$$A_{ij} = \langle\varphi_i|A|\varphi_j\rangle = \int d\mathbf{r}\varphi_i^*(\mathbf{r})A\varphi_i(\mathbf{r}). \tag{188}$$

The result in Eq. (187) can be written as a matrix. The matrix elements between different basis functions can be put in matrix form like

$$\mathrm{A} = \begin{pmatrix} A_{11} & . & . & A_{1N} \\ . & . & . & . \\ A_{N1} & . & . & A_{NN} \end{pmatrix}. \tag{189}$$

In expression ((187), we see that the detailed nature of the basis function is no longer relevant, since the entire $\mathbf{r}$ dependence is absorped in the matrix element. The basis function can then be expressed in terms of a vector

$$|\varphi_i\rangle = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 0 \\ 1 \\ 0 \\ . \\ . \end{pmatrix} \leftarrow i, \tag{190}$$

where there is only 1 in the i'th row. Its Hermitian conjugate is given by

$$\langle\varphi_i| = (|\varphi_i\rangle)^\dagger = (0, 0, \ldots, 0, 1, 0, \ldots). \tag{191}$$

This notation is known as the bra-ket notation. The inner product of two basis functions is then

$$\langle\varphi_i|\varphi_j\rangle = \delta_{ij}, \tag{192}$$

where we have assumed that our basis set is limited to n functions. A wavefunction can now be written as

$$|\Psi\rangle = \sum_i a_i|\varphi_i\rangle = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{pmatrix}, \tag{193}$$

where the only thing that is relevant now is the relative contributions of the different basis functions given by the coefficients a_i. The wavefunctions followed the same operations as vectors in an N-dimensional space. For example, addition is given by

$$|\Psi\rangle + |\Phi\rangle = \sum_i a_i|\varphi_i\rangle + \sum_i b_i|\varphi_i\rangle = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ . \\ . \\ . \\ a_n + b_n \end{pmatrix}. \tag{194}$$

The inner product of two functions becomes

$$\langle\Phi|\Psi\rangle = \sum_{i=1}^{N} b_i^* a_i, \tag{195}$$

Multiplying the wavefunction $|\Psi\rangle$ by the matrix A gives

$$\mathrm{A}|\Psi\rangle = \begin{pmatrix} \sum_i A_{1i}a_i \\ \sum_i A_{2i}a_i \\ . \\ . \\ \sum_i A_{ni}a_i \end{pmatrix}. \tag{196}$$

The matrix element between to vectors is

$$\langle\Phi|A|\Psi\rangle = (b_1^*, b_2^*, \ldots, b_n^*) \begin{pmatrix} \sum_i A_{1i}a_i \\ \sum_i A_{2i}a_i \\ . \\ . \\ \sum_i A_{ni}a_i \end{pmatrix} = \sum_{ij} b_j^* A_{ji} a_i, \tag{197}$$

which is the same result as in Eq. (187).

A special operator is the Hamiltonian of the system. The eigenfunctions of the Hamiltonian can be found by solve the eigenvalue equation

$$\mathrm{H}|\psi_n\rangle = E|\psi_n\rangle, \tag{198}$$

where n labels the eigenfunctions.

1. Function spaces in infinite dimensions

If you have followed everything so far, that's great. However, some of you might be wondering if we can write down a direct connection between $|\psi\rangle$ and $\psi(\mathbf{r})$. For completeness, we add this. However, for some of you, this might be more confusing then helpful. You probably will not use this too often in practice. Let us assign every vector in real space as a basis vector in an infinite-dimensional space. Let us first do this discretely and later replace the summations by integrals. The basis vector for vector $\mathbf{r}_i$ is given by

$$|\mathbf{r}_i\rangle = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 0 \\ 1 \\ 0 \\ . \\ . \end{pmatrix} \leftarrow j \quad . \tag{199}$$

Since the positions in space are independent of each other, we have $\langle\mathbf{r}_{j'}|\mathbf{r}_j\rangle = \delta_{jj'}$. Now we can expand the wavefunction $|\psi\rangle$ in this vector space as

$$|\psi\rangle = \begin{pmatrix} \psi(\mathbf{r}_1) \\ \psi(\mathbf{r}_2) \\ . \\ . \\ \psi(\mathbf{r}_i) \\ . \\ . \end{pmatrix}, \tag{200}$$

which we can also express this as

$$|\psi\rangle = \sum_i \psi(\mathbf{r}_i)|\mathbf{r}_i\rangle = \int d\mathbf{r}\ \psi(\mathbf{r})|\mathbf{r}\rangle, \tag{201}$$

where in the last step, we have replaced the summation by an integral in the limit that the discretization goes to zero. If we want to know the value of the wavefunction at a certain $\mathbf{r}_j$ value, we take the inner product with that basis vector:

$$\langle\mathbf{r}_j|\psi\rangle = \sum_i \psi(\mathbf{r}_i)\langle\mathbf{r}_j|\mathbf{r}_i\rangle = \psi(\mathbf{r}_j)\delta_{ij} = \psi(\mathbf{r}_j) \tag{202}$$

Or if we do it in integral notation

$$\langle\mathbf{r}|\psi\rangle = \int d\mathbf{r}'\ \psi(\mathbf{r}')\langle\mathbf{r}|\mathbf{r}'\rangle = \psi(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') = \psi(\mathbf{r}) \tag{203}$$

Suppose we now want to calculate a matrix element $\langle\psi'|H|\psi\rangle$. The expression $\langle\psi'|H|\psi\rangle$ is nice, but in order to calculate a matrix element, one really needs to now what it looks like in real space. However, in any matrix expression, we can always squeeze in identity matrices, since a multiplication with an identity matrix does not do anything. The identity matrix for $|\mathbf{r}\rangle$ is given by

$$\int d\mathbf{r}\ |\mathbf{r}\rangle\langle\mathbf{r}| = \mathrm{I}_\infty, \tag{204}$$

where infinity-dimensional I_∞ is the identity matrix (all ones on the diagonal, zeros everywhere else). The matrix element can now be written as

$$\langle\psi'|H|\psi\rangle = \int d\mathbf{r}\ \int d\mathbf{r}'\ \langle\psi'|\mathbf{r}'\rangle\langle\mathbf{r}'|H|\mathbf{r}\rangle\langle\mathbf{r}|\psi\rangle = \int d\mathbf{r}\ \int d\mathbf{r}'\ \psi'^*(\mathbf{r}')\langle\mathbf{r}'|H|\mathbf{r}\rangle\psi(\mathbf{r}). \tag{205}$$

Now a potential is only a function of $\mathbf{r}$, so $\langle\mathbf{r}'|H|\mathbf{r}\rangle = H(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$, and we are left with

$$\langle\psi'|H|\psi\rangle = \int d\mathbf{r}\psi'^*(\mathbf{r})H(\mathbf{r})\psi(\mathbf{r}), \tag{206}$$

which is the usual expression for the matrix element.

This might appear somewhat confusing with bases in $\mathbf{r}$ of infinite dimensions. However, it might become clearer if you think of it as a code in your computer. You want to solve a problem using the matrix method. This means you have to diagonalize a matrix to find the eigenvalues and eigenvectors. However, the matrix consists of numbers that you need to evaluate. In order to find the matrix element, you really need to know what the basis functions look like in real space. Therefore, there is a separate routine has the values of the wavefunctions stored. This is essentially the expression in Eq. (201). You only need that when you actually want to calculate the matrix elements. Sometimes people skip the calculation of matrix elements all together. For example, the two-dimensional problem that we started out with can be treated without ever calculating ε_i and t. You can just assume some values. This is then known as a model Hamiltonian, which are used by a large number of physicist/chemists.

2. Impurity scattering

In this problem, we want to look at the effect of a local impurity potential on free particles. This is a rather complex problem that can be treated analytically, although that generally involves complex methods using Green's functions. We will solve the problem numerically. In the absence of the local impurity the eigenfunctions are plane waves

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{207}$$

with an energy $\varepsilon_{\mathbf{k}} = \hbar^2\mathbf{k}^2/2m$. After introducing the impurity potential, the plane waves are no longer eigenfunctions. However, since they form a complete set, we can still use them as basis functions. The eigenfunctions in the presence of an impurity can be expressed in terms of plane waves

$$\psi_n(\mathbf{r}) = \sum_{\mathbf{k}} a_{\mathbf{k}}\varphi_{\mathbf{k}}(\mathbf{r}). \tag{208}$$

In bra-ket notation, we can write the basis $|\varphi_{\mathbf{k}}\rangle$. The Hamiltonian for the kinetic energy is then written as

$$\mathrm{H}_0 = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}|\varphi_{\mathbf{k}}\rangle\langle\varphi_{\mathbf{k}}|, \tag{209}$$

which is diagonal in $\mathbf{k}$.

A local potential in real space can be expressed as

$$H_U = U\delta(\mathbf{r}), \tag{210}$$

which is indeed an impurity potential. The matrix elements for such a potential are given by

$$\langle\varphi'_{\mathbf{k}}|H_U|\varphi_{\mathbf{k}}\rangle = \frac{1}{V}\int d\mathbf{r}' e^{-i\mathbf{k}'\cdot\mathbf{r}}U\delta(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{U}{V}. \tag{211}$$

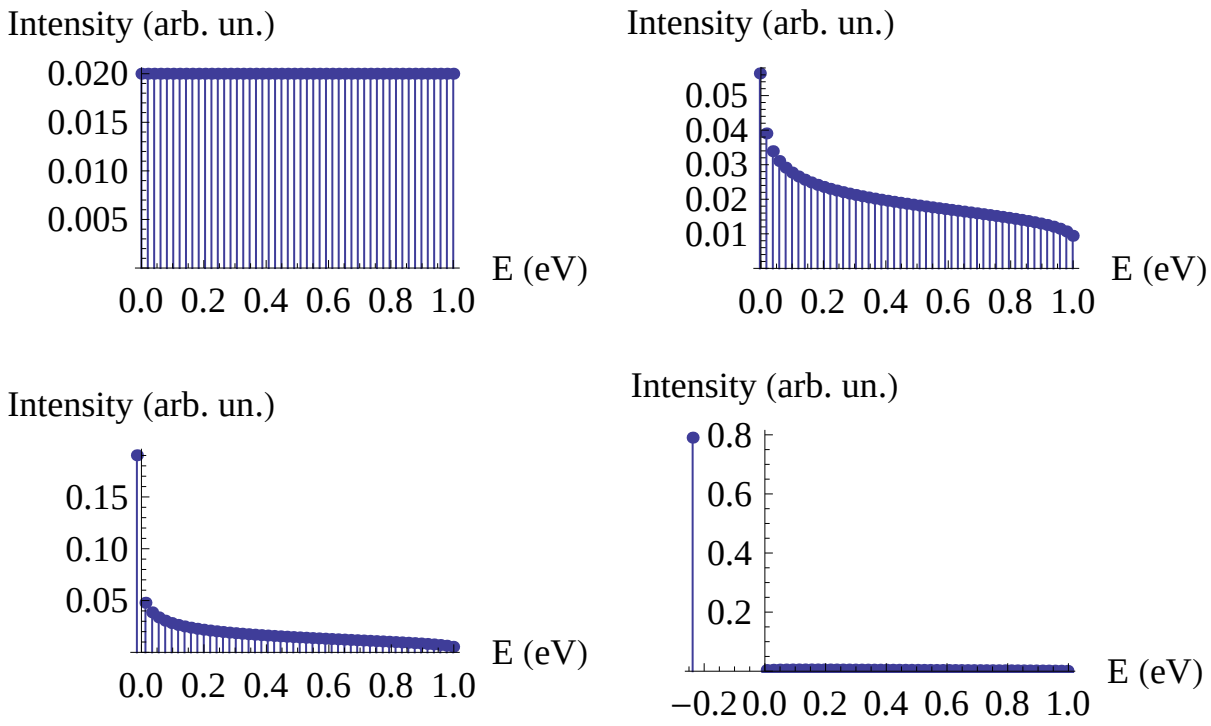


FIG. 10: The value of $|\langle\psi_n|\Psi_{\mathbf{r}=0}\rangle|^2$ where the local state is given $|\Psi_{\mathbf{r}=0}\rangle = \frac{1}{\sqrt{N}}\sum_m |\varphi_m\rangle$ for scattering of a local attractive impurity for $U = 0, -0.1, -0.2$, and -0.6. The number of states that are included in the model is 50.

The Hamiltonian can then be written as

$$\mathrm{H}_U = \frac{U}{V}\int d\mathbf{k}' \int d\mathbf{k} |\varphi'_{\mathbf{k}}\rangle\langle\varphi_{\mathbf{k}}| \rightarrow \mathrm{H}_U = \frac{U}{N}\sum_{\mathbf{k},\mathbf{k}'} |\varphi'_{\mathbf{k}}\rangle\langle\varphi_{\mathbf{k}}|, \tag{212}$$

where in the last step we have replaced the integration by a summation, where we only take N k-values into account. This is more appropiate for numerical work since it is impossible to use an infinite basis set. Note that whereas the kinetic energy is diagonal in the momentum (which is of course a direct result of our choice of plane waves as a basis set), the local potential couples all the different $\mathbf{k}$ values.

To really be able to perform some numerical solution (and in addition simplify our lives somewhat), let us forget about the parabolic dependence of the energy on $\mathbf{k}$, the density of k values and simply assume that there is only an energy dependence. These are pretty big approximations, but not directly essential for the physics that we want to get out. So let us take the kinetic energy term as

$$\mathrm{H}_0 = \sum_{m=1}^{N} \epsilon_m |\varphi_m\rangle\langle\varphi_m| = \begin{pmatrix} \epsilon_1 & 0 & 0 & \cdots & \\ 0 & \epsilon_2 & 0 & \cdots & \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & & & \epsilon_N \end{pmatrix}, \tag{213}$$

with

$$\epsilon_n = \frac{m-1}{N-1}W, \tag{214}$$

where we went from a momentum basis $|\mathbf{k}\rangle$ to an basis in energy $|n\rangle$, where the index n is directly related to the energy ϵ_m. Note that the energy values are equidistant and go from 0 to W, where W can be called a bandwidth.

The scattering term then becomes

$$\mathrm{H}_U = \frac{U}{N}\sum_{m,m'=1}^{N}|\varphi'_m\rangle\langle\varphi_m| = \frac{U}{N}\begin{pmatrix} 1 & 1 & 1 & \cdots & \\ 1 & 1 & 1 & \cdots & \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 1 & 1 & & & 1 \end{pmatrix}. \tag{215}$$

The eigenvalues and eigenvectors satisfy the Schrödinger equation

$$(\mathrm{H}_0 + \mathrm{H}_U)|\psi_n\rangle = E_n|\psi_n\rangle. \tag{216}$$

Obtaining the eigenvalues involves solving the determinant

$$\begin{vmatrix} \epsilon_1 + U/N & U/N & U/N & \cdots & \\ U/N & \epsilon_2 + U/N & U/N & \cdots & \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ U/N & U/N & & & \epsilon_N + U/N \end{vmatrix}. \tag{217}$$

This looks pretty horrific, but with the wide availability of codes or commericial software aimed at solving eigenvalue problems, this problem is acually pretty straightforward (in this case mathematica was used). In the system, there is one special site, namely that with the impurity. We want to know if we put our system in this local state, how will it end up in the eigenstates of the system. When dealing with wavevectors $\mathbf{k}$, the local state can be expressed as a summation over all $\mathbf{k}$ values

$$|\Psi_{\mathbf{r}=0}\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}}|\varphi_{\mathbf{k}}\rangle. \tag{218}$$

This is essentially the definition of a δ-function. In our simplified problem, we take

$$|\Psi_{\mathbf{r}=0}\rangle = \frac{1}{\sqrt{N}}\sum_{m}|\varphi_m\rangle. \tag{219}$$

The weight

$$w_n = |\langle\psi_n|\Psi_{\mathbf{r}=0}\rangle|^2. \tag{220}$$

The results are given in Fig. 10 for $W = 1$, $N = 50$, and $U = 0, -0.1, -0.2$, and -0.6. It is easy to do calculations for larger N, but the physics remains the same but the figure becomes worse. Let us first consider the situation for $U = 0$. This is basically the absence of an impurity and the basis functions are the eigenfunctions of the system. The weight is then

$$w_n = |\langle\varphi_n|\Psi_{\mathbf{r}=0}\rangle|^2 = |\frac{1}{\sqrt{N}}\sum_{m}\langle\varphi_n|\varphi_m\rangle|^2 = \frac{1}{N}, \tag{221}$$

which is $1/50 = 0.02$ if $N = 50$. Therefore, a localized state has an equal probability of ending up in any of the delocalized states. For $U = -0.1$ and -0.2 the weight increases to lower energies. For $U = -0.8$, we see that a state has split off from the other states that has nearly 80% of the weight. For decreasing U even more, the energy becomes close to $-U + W/2$ and the weight approaches 1. What is the underlying physics of this? Actually, this is not too hard to understand. The impurity potential is effectively a local attractive potential. If the potential becomes very attractive, the local state becomes bound to this potential.

V. HARMONIC OSCILLATOR

The harmonic oscillator played a significant role in the development of quantum mechanics. An harmonic oscillator is, for example, a particle which is bound in its position by a force proportional with the displacement x of the particle from its equilibrium position. Classicaly this is easy to solve starting from Newton's equation

$$m\frac{d^2x}{dt^2} = -Kx, \tag{222}$$

where K is the spring constant. The oscillation written as $A\cos\omega t$ assuming that the oscillation is maximum at $t = 0$. Inserting gives

$$-m\omega^2 A\cos\omega t = -KA\cos\omega, \tag{223}$$

or $\omega = \sqrt{K/m}$ (or $K = m\omega^2$). The velocity is given by $v = -\omega A\sin\omega t$. The total energy is therefore

$$E = \frac{1}{2}mv^2 + \frac{1}{2}m\omega^2x^2 = \frac{1}{2}m\omega^2A^2\sin^2\omega t + \frac{1}{2}m\omega^2A^2\cos^2\omega t = \frac{1}{2}m\omega^2A^2 \tag{224}$$

Certainly classically there seems to be no indication of quantization. The energy can be varied continuously by changing the amplitude.

There are two ways to approach this problem in quantum mechanics. This first is writing down the Schrödinger equation and solve it, leading to a solution in terms of Hermite polynomials. This solution directly gives you the probability distribution function as a function of x. The other approach is in terms of effective operators, that tells a lot about the eigenfunctions and eigenenergies of the system, but not much about the probability distribution function. However, quite often we are only interested in eigenenergies anyway, since experiments often yield little information on the probability distribution, but a lot about the spectral, i.e., energy, distribution.

A. Schrödinger equation approach

The time-independent Schrödinger equation in case of a harmonic oscillator is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2x^2\psi = E\psi. \tag{225}$$

This can be brought in a somewhat more compact form by introducing the variables $\lambda = 2E/\hbar\omega$ and $\xi = \sqrt{m\omega/\hbar}x$, giving

$$\frac{d^2\psi}{d\xi^2} + (\lambda - \xi^2)\psi = 0. \tag{226}$$

Let us first look at the $\xi \to \infty$ case for which

$$\frac{d^2\psi_\infty}{d\xi^2} - \xi^2\psi_\infty = 0. \tag{227}$$

Since we are looking at bound particles, the solution has to go to zero when $\xi \to \infty$. An approximate solution is $\psi_\infty(\xi) = \exp(-\xi^2/2)$

$$\frac{d\psi_\infty}{d\xi} = e^{-\xi^2/2}(-\xi) \qquad \Rightarrow \qquad \frac{d^2\psi_\infty}{d\xi^2} = e^{-\xi^2/2}\xi^2 - e^{-\xi^2/2} \cong \xi^2e^{-\xi^2/2}, \tag{228}$$

thereby satisfying Eqn. 227. We now proceed to seek solutions of the form $\psi(\xi) = H(\xi)e^{-\xi^2/2}$ which have the right asymptotic behavior. Inserting this into the differential equation gives

$$\frac{d^2H(\xi)}{d\xi^2} - 2\xi\frac{dH(\xi)}{d\xi} + (\lambda - 1)H(\xi) = 0. \tag{229}$$

This can be solved by a series solution.

What are series solutions?

Let us consider a very simple differential equation

$$\frac{dy}{dx} = cy. \tag{230}$$

Let us assume that the solution can be written as a polynomial

$$y = \sum_{k=0}^{\infty} a_k x^k. \tag{231}$$

We can exclude the negative k values since the function does not diverge at $x =$. The derivative is then given by

$$\frac{dy}{dx} = \sum_{k=1}^{\infty} a_k k x^{k-1}. \tag{232}$$

In order to compare this with the right-hand side of the differential equation, let us shift the indices by introducing $k' = k - 1$ or $k = k' + 1$.

$$\frac{dy}{dx} = \sum_{k'=0}^{\infty} a_{k'+1}(k'+1)x^{k'}. \tag{233}$$

Since k' is just a summation variable, we can also call the index n again. We can now rewrite our differential equation as

$$\sum_{k=0}^{\infty}((k+1)a_{k+1} - ca_k)x^k = 0 \tag{234}$$

The functions x^k are independent. This means that we can never find a combination of x^k that will be zero for all values of x. Therefore, the coefficients in front of the x^n have to be zero. Therefore,

$$a_{k+1} = \frac{c}{k+1}a_k = \frac{c}{k+1}\frac{c}{k}a_{k-1} = \frac{c^{k+1}}{(k+1)!}a_0 \qquad \text{or} \qquad a_k = \frac{c^k}{k!}a_0 \tag{235}$$

Our function is now

$$y = \sum_{k=0}^{\infty} a_0 \frac{1}{k!}(cx)^n = a_0 e^{cx}, \tag{236}$$

where we recognize the series as an expansion of the exponent

$$e^x \equiv \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{237}$$

One might also look at this differently and say that e^x is the function that is defined by this series and we have now proven that the derivative of $e^{cx} = ce^x$. Of course, we could have solved this problem more simply if we had used this result directly.

Let us return to the problem of the harmonic oscillator, where for the differential equation we assume that the solution can be described as a polynomial

$$H(\xi) = \sum_{k=0}^{\infty} a_k \xi^k. \tag{238}$$

The derivatives are given by

$$\frac{dH}{d\xi} = \sum_{k=0}^{\infty} a_k k \xi^{k-1} \qquad \Rightarrow -2\xi\frac{dH}{d\xi} = -\sum_{k=0}^{\infty} a_k 2k\xi^k \tag{239}$$

$$\frac{d^2H}{d\xi^2} = \sum_{k=2}^{\infty} a_k k(k-1)\xi^{k-2} = \sum_{k=0}^{\infty} a_{k+2}(k+2)(k+1)\xi^k, \tag{240}$$

where in the last summation the index has been shifted by 2. Inserting this back into the differential equation gives

$$\sum_{k=0}^{\infty}\left\{(k+1)(k+2)a_{k+2} - (2k+1-\lambda)a_k\right\}\xi^k = 0. \tag{241}$$

Since ξ^k are independent functions, the only way to make this zero is by having all the components in front of the ξ^k equal to zero. This leads to the recursion relation

$$a_{k+2} = \frac{2k+1-\lambda}{(k+1)(k+2)} a_k. \tag{242}$$

For large values of ξ the series is dominated by the terms with a large power of k. In the limit of $k \gg \lambda$, this becomes approximately

$$a_{k+2} \cong \frac{2}{k} a_k = \frac{2}{k}\frac{2}{k-2}\frac{2}{k-4}\cdots\frac{2}{2} a_0 = \frac{1}{(\frac{1}{2}k)!} a_0. \tag{243}$$

For the series with even powers of ξ (there is also one with only odd powers), we have in the large ξ limit

$$H(\xi) = a_0 \sum_{k=0}^{\infty} \frac{1}{k!} \xi^{2k} = a_0 e^{\xi^2}. \tag{244}$$

This gives as total solution $\psi(\xi) = H(\xi) e^{-\xi^2/2} = a_0 e^{\xi^2/2}$. Note that this diverges as $\xi \to \infty$. We cannot have that, because this would mean that the particle would spend most of its time in infinity, whereas we would like to have it close to $\xi = 0$. Therefore, we cannot have the series continue till infinity. However, there is a way to cut off the series. If we take $\lambda = 2n+1$ then the coefficient a_{n+2} will be zero. Note that this implies for the energy

$$\lambda_n = 2n+1 = \frac{2E_n}{\hbar\omega} \qquad \Rightarrow \qquad E_n = (n+\frac{1}{2})\hbar\omega. \tag{245}$$

This implies that the energies of the harmonic oscillator are quantized just as Planck originally proposed. There is a slight difference, since there is what is called the zero-point energy. This is directly related to Heisenberg's uncertainty principle. Since the oscillation is around zero, we have $\langle x \rangle = \langle p \rangle = 0$. This directly implies that $\Delta x^2 = \langle x^2 \rangle$ and $\Delta p^2 = \langle p^2 \rangle$. The total energy is therefore

$$E = \frac{\langle p^2 \rangle}{2m} + \frac{1}{2} m\omega^2 \langle x^2 \rangle = \frac{\Delta p^2}{2m} + \frac{1}{2} m\omega^2 \Delta x^2. \tag{246}$$

Now since $\Delta x \Delta p \geq \frac{\hbar}{2}$, we have

$$E \geq \frac{\hbar^2}{8m\Delta x^2} + \frac{1}{2} m\omega^2 \Delta x^2. \tag{247}$$

Minimizing this energy with respect to Δx^2 gives

$$-\frac{\hbar^2}{8m(\Delta x^2)^2} + \frac{1}{2} m\omega^2 = 0 \qquad \Rightarrow \qquad (\Delta x^2)_{\min} = \frac{\hbar}{2m\omega}. \tag{248}$$

This means that

$$E \geq \frac{\hbar^2 2m\omega}{8m\hbar} + \frac{1}{2} m\omega^2 \frac{\hbar}{2m\omega} = \frac{1}{2}\hbar\omega. \tag{249}$$

This is the mimimum energy an oscillation can have without violating Heisenberg's uncertainty principle. Therefore, unlike the classical oscillator which can just hang still, in the quantum mechanical case the wavefunction always has to spread out to avoid a very large kinetic energy. In general, we find that

$$\Delta x_n^2 = \frac{\hbar}{m\omega}(n+\frac{1}{2}). \tag{250}$$

Classically, the amplitude of the oscillation is given by A. However, Δx^2 is given by

$$\Delta x^2 = \frac{A^2\omega}{\pi} \int_0^{\frac{\pi}{\omega}} dt \cos^2 \omega t = \frac{A^2\omega}{\pi} \int_0^{\frac{\pi}{\omega}} dt \left(\frac{1}{2} + \frac{1}{2}\cos 2\omega t\right) = \frac{A^2}{2} \tag{251}$$

The effective amplitude of the oscillation is therefore quantized

$$A_n^2 = 2\Delta x_n^2 = \frac{2\hbar}{m\omega}(n+\frac{1}{2}). \tag{252}$$

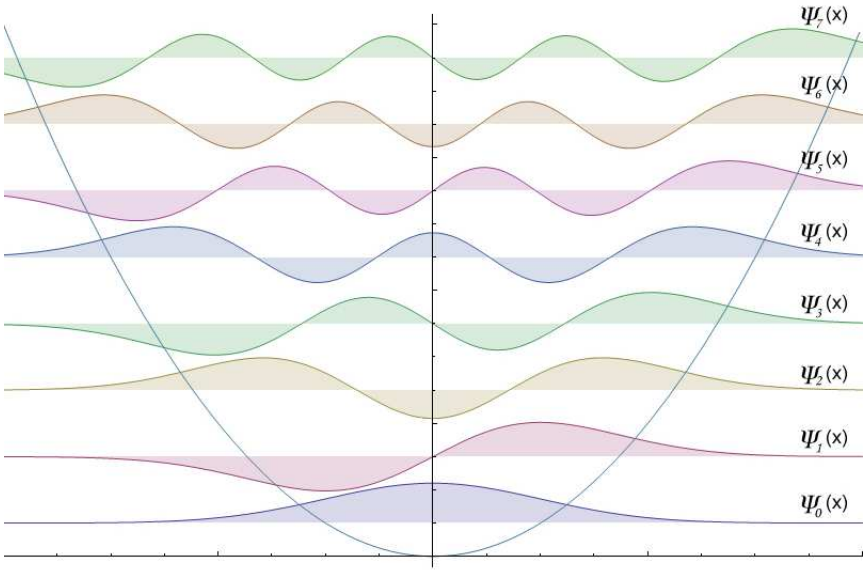


FIG. 11: The solutions of the Schrödinger equation for the harmonic oscillator for different values of n. Source: wikipedia.

where π/ω is half an oscillation period. The energy is then

$$E = \frac{1}{2}m\omega^2 A^2 = \frac{1}{2}m\omega^2 \frac{2\hbar}{m\omega}(n+\frac{1}{2}) = (n+\frac{1}{2})\hbar\omega \tag{253}$$

It can be shown that the function $H_n(\xi)$, which are called Hermite polynomials, can be written as

$$H_n(\xi) = (-1)^n e^{\xi^2}\frac{d^n}{d\xi^n}(e^{-\xi^2}). \tag{254}$$

The first Hermite poynomials are

$$H_0(\xi) = 1 \tag{255}$$
$$H_1(\xi) = 2\xi \tag{256}$$
$$H_2(\xi) = 4\xi^2 - 2 \tag{257}$$
$$H_3(\xi) = 8\xi^3 - 12\xi. \tag{258}$$

Note that the relation between the coefficients can be found from the recursion relation in Eqn. (242). For example, for $n = 2$, we have $\lambda = 5$ and therefore $a_2 = (1-5)/(1\times 2)a_0 = -2a_0$. For $n = 3$, we have $\lambda = 7$ and therefore $a_3 = (3-7)/(2\times 3)a_1 = -\frac{2}{3}a_1$. Then there is a constant factor, which ensures that the Hermite polynomial satisfy the (traditional) normalization relation

$$\int_{-\infty}^{\infty} H_n(\xi)H_m(\xi)e^{-\xi^2}d\xi = \sqrt{\pi}2^n n!\delta_{n,m}. \tag{259}$$

The total normalized wavefunction is then

$$\psi_n(\xi) = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}}H_n(\xi)e^{-\xi^2/2}. \tag{260}$$

The solutions for the lowest values of n are shown in Fig. 11.

B. Solution using operators

An alternative approach to the harmonic oscillator is the use of operators. Classically, the values of x and p_x are intimately related. If $x = x_0\cos\omega t$, then $p_x = mx_0\omega\sin\omega t$, i.e. they differ by a factor and a phase shift. This is also

seen from the Hamiltonian, which can be viewed as the values of x and p_x lying on an ellips, $\frac{p_x^2}{a^2} + \frac{x^2}{b^2} = E$. This was one of the underlying starting points of the 'old' quantum mechanics description of the quantization of the harmonic oscillator. In general, we are less interested in the phase of the oscillation, which just tells you than something is swinging. We are more interested in the absolute value. In classical mechanics, we can easily factorize an equation of the form $u^2 + v^2 = E$, by going to complex coordinates $(u - iv)(u + iv) = ae^{-i\varphi}ae^{i\varphi} = a^*a = |a|^2 = E$, with $a = \sqrt{u^2 + v^2}$ and $\tan\varphi = v/u$. This is roughly the approach that we would like to follow. The classical Hamiltonian can be factorized as

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 = \omega\left(\sqrt{\frac{m\omega}{2}}x - i\frac{p_x}{\sqrt{2m\omega}}\right)\left(\sqrt{\frac{m\omega}{2}}x + i\frac{p_x}{\sqrt{2m\omega}}\right). \tag{261}$$

However, in quantum mechanics, we have to take care of the fact that $\hat{p}_x$ and $\hat{x}$ do not commute

$$\omega\left(\sqrt{\frac{m\omega}{2}}\hat{x} - i\frac{\hat{p}_x}{\sqrt{2m\omega}}\right)\left(\sqrt{\frac{m\omega}{2}}\hat{x} + i\frac{\hat{p}}{\sqrt{2m\omega}}\right) = \frac{\hat{p}_x^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 - \frac{i\omega}{2}(\hat{p}_x\hat{x} - \hat{x}\hat{p}_x) = \hat{H} - \frac{1}{2}\hbar\omega, \tag{262}$$

using Eq. (126). Let us now introduce the operators

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\frac{\hat{p}_x}{\sqrt{2m\omega\hbar}} = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x + \sqrt{\frac{\hbar}{m\omega}}\frac{\partial}{\partial x}\right) = \frac{1}{\sqrt{2}}\left(\xi + \frac{\partial}{\partial\xi}\right), \tag{263}$$

where $\xi = \sqrt{\frac{m\omega}{\hbar}}x$ as defined above. Similarly,

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\frac{\hat{p}_x}{\sqrt{2m\omega\hbar}} = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{\partial}{\partial x}\right) = \frac{1}{\sqrt{2}}\left(\xi - \frac{\partial}{\partial\xi}\right). \tag{264}$$

(Textbooks usually put hats on the operators, however this becomes quite cumbersome and it is often clear that we are dealing with operators. In the majority of scientific papers the hats are simply omitted. We will follow that practice). The operators follow the following commutation relations

$$[a, a^\dagger] = aa^\dagger - a^\dagger a = \frac{1}{2}\left[\left(\xi + \frac{\partial}{\partial\xi}\right)\left(\xi - \frac{\partial}{\partial\xi}\right) - \left(\xi - \frac{\partial}{\partial\xi}\right)\left(\xi + \frac{\partial}{\partial\xi}\right)\right] \tag{265}$$

$$= \frac{1}{2}\left[-\xi\frac{\partial}{\partial\xi} + \frac{\partial}{\partial\xi}\xi - \xi\frac{\partial}{\partial\xi} + \frac{\partial}{\partial\xi}\xi\right] = 1. \tag{266}$$

The Hamiltonian then obtains the simple form

$$H = (a^\dagger a + \frac{1}{2})\hbar\omega. \tag{267}$$

We want to find the eigenfunctions and eigenvalues of the Hamiltonian. Since the eigenfunctions should have the property $H\psi = E\psi$, operating $a^\dagger a$ on the eigenfunction can only lead to a real number n (although, as we now already, this later turns our to be an integer), i.e.,

$$a^\dagger a\psi_n = n\psi_n. \tag{268}$$

What is now the effect of having $a^\dagger a$ work on $a\psi_n$? Using the commutation relation, we find

$$a^\dagger aa\psi_n = (aa^\dagger - 1)a\psi_n = a(a^\dagger a - 1)\psi_n = a(n-1)\psi_n = (n-1)a\psi_n \tag{269}$$

This means that $a\psi_n$ is proportional to ψ_{n-1}, since the eigenvalue for $a^\dagger a$ is $n - 1$. Similarly, we can show that

$$a^\dagger aa^\dagger\psi_n = (n+1)a^\dagger\psi_{n+1}. \tag{270}$$

Thus, we can view $a^\dagger$ and a as step up and step down operators, respectively. They are also called creation and annihilation operators, respectively.

So we have already found that the eigenvalues are separated by $\hbar\omega$, without solving any differential equation. In addition, we know that the eigenvalues of H have to be larger than or equal to zero, since H has only x^2 and a p_x^2 operators in it. Thus

$$H\psi_n = (a^\dagger a + \frac{1}{2})\hbar\omega\psi_n = (n + \frac{1}{2})\hbar\omega\psi_n \qquad \text{with} \qquad (n + \frac{1}{2}) \geq 0 \qquad \Rightarrow \qquad n \geq -\frac{1}{2} \tag{271}$$

So the lowest eigenvector has $-\frac{1}{2} \leq n_{\min} < \frac{1}{2}$. We must have $a\psi_{n_{\text{low}}} = 0$ since we are not allowed to have eigenvalues with $n_{\text{low}} - 1$. From this it follows that

$$a^\dagger a \psi_{n_{\text{low}}} = 0 \qquad \Rightarrow \qquad n_{\text{low}} = 0. \tag{272}$$

Therefore the eigenvalues for the harmonic oscillator are $\frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega, \frac{5}{2}\hbar\omega, \ldots$.

We see that in the operator approach, we can obtain the eigenvalues without solving the differential equation. Therefore, we do not need to know $\psi_n(\xi)$ to obtain understanding of our problem. This is comparable to the classical situation, where it is not necessary to know the complete motion, because everything is determined by the maximum amplitude x_0. We can therefore introduce new eigenstates $|n\rangle$, which have the same properties when applying the operators on them as the wavefunctions $\psi_n(\xi)$, although we do not the what the function looks like in real space. Classically, we can say that we have an oscillation with a maximum oscillation x_0 (or its energy $\frac{1}{2}m\omega^2 x_0^2$), the exact motion is determined by our Hamiltonian (and is a simple cosine). Quantum mechanically, we can say that we have an eigenstate that is given by its energy $n\hbar\omega$. Its motion is determined by the Schrödinger equation and is given by $\psi_n(\xi)$ although often this knowledge is not essential. Let us now denote the lowest eigenstate in the bra-ket notation as $|0\rangle$. Then we can obtain the n'th eigenstate by

$$|\tilde{n}\rangle = (a^\dagger)^n |0\rangle. \tag{273}$$

We put a tilde on this wavefunction to indicate that this state is not normalized. Let us look at some specific examples for the normalization:

$$|\tilde{1}\rangle = a^\dagger|0\rangle \qquad \Rightarrow \qquad \langle\tilde{1}|\tilde{1}\rangle = \langle 0|aa^\dagger|0\rangle = \langle 0|(1 + a^\dagger a)|0\rangle = 1, \tag{274}$$

since $a|0\rangle = 0$ and since $\langle 0|0\rangle = 1$. So $|\tilde{1}\rangle$ is also nicely normalized. However, for $|2\rangle$, we find

$$|\tilde{2}\rangle = a^\dagger a^\dagger|0\rangle \qquad \Rightarrow \qquad \langle\tilde{2}|\tilde{2}\rangle = \langle 0|aaa^\dagger a^\dagger|0\rangle = \langle 0|a(1 + a^\dagger a)a^\dagger|0\rangle = \langle 0|aa^\dagger + aa^\dagger aa^\dagger|0\rangle \tag{275}$$

$$= \langle 0|(1 + a^\dagger a) + (1 + a^\dagger a)(1 + a^\dagger a)|0\rangle = 2, \tag{276}$$

So the normalized wavefunction is $|2\rangle = \frac{1}{\sqrt{2}}a^\dagger a^\dagger|0\rangle$. If we look at it a bit more carefully, we see that the first annihilation operator a from the right has to make a 1 with one of the creation operators $a^\dagger$. For $|n\rangle$, there a n possibilities to do that. The next annihilation operator also has to make a 1, however, now there are only $n - 1$ creation operators left, and so on. Therefore, there are $n!$ different ways of making a non-zero combination. The normalized wavefunction are therefore given by

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \tag{277}$$

We can use this result to determine the factors for the step operators

$$a^\dagger|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^{n+1}|0\rangle = \sqrt{n+1}\frac{1}{\sqrt{(n+1)!}}(a^\dagger)^{n+1}|0\rangle = \sqrt{n+1}|n+1\rangle \tag{278}$$

and

$$a|n\rangle = \frac{1}{\sqrt{n!}}a(a^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}}n(a^\dagger)^{n-1}|0\rangle = \sqrt{n}\frac{1}{\sqrt{(n-1)!}}(a^\dagger)^{n-1}|0\rangle = \sqrt{n}|n-1\rangle, \tag{279}$$

again using the fact that there are n possible ways to apply the commutation relation. From this result, we easily obtain

$$a^\dagger a|n\rangle = a^\dagger\sqrt{n}|n-1\rangle = n|n\rangle. \tag{280}$$

We can also write $\hat{x}$ and $\hat{p}_x$ in terms of the creation and annihilation operators

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a) \qquad \text{and} \qquad \hat{p}_x = i\sqrt{\frac{m\hbar\omega}{2}}(a^\dagger - a). \tag{281}$$

For the expectation values, we find

$$\langle n|\hat{x}|n\rangle = \int_{-\infty}^{\infty} dx\psi_n(x)x\psi_n(x) = \sqrt{\frac{\hbar}{2m\omega}}\langle n|(a^\dagger + a)|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left\{\sqrt{n+1}\langle n|n+1\rangle + \sqrt{n}\langle n|n-1\rangle\right\} = 0, \tag{282}$$

where

$$\langle n'|n\rangle = \int_{-\infty}^{\infty} dx \psi_{n'}(x)\psi_n(x) = \delta_{n,n'} \tag{283}$$

is the inner product between the different wavefunction. This is zero when the quantum numbers are unequal, since the wavefunctions are orthogonal. Similarly, we find $\langle \hat{p}_x \rangle = 0$. This is in agreement with our classical idea that for an oscillation the average x and p_x are zero. However, for $\langle \hat{x}^2 \rangle$, we obtain

$$\langle n|\hat{x}^2|n\rangle = \frac{\hbar}{2m\omega}\langle n|(a^\dagger + a)(a^\dagger + a)|n\rangle = \frac{\hbar}{2m\omega}\langle n|(a^\dagger a^\dagger + a^\dagger a + aa^\dagger + aa)|n\rangle \tag{284}$$

$$= \frac{\hbar}{m\omega}\langle n|(a^\dagger a + \frac{1}{2})|n\rangle = \frac{\hbar}{m\omega}(n + \frac{1}{2}) \tag{285}$$

and for $\langle \hat{p}_x^2 \rangle$, we obtain

$$\langle n|\hat{p}_x^2|n\rangle = -\frac{m\hbar\omega}{2}\langle n|(a^\dagger - a)(a^\dagger - a)|n\rangle = -\frac{m\hbar\omega}{2}\langle n|(-a^\dagger a - aa^\dagger)|n\rangle = m\hbar\omega(n + \frac{1}{2}). \tag{286}$$

Combining these results gives

$$\Delta x \Delta p_x = \sqrt{\langle x^2\rangle\langle p_x^2\rangle} = \hbar(n + \frac{1}{2}) \geq \frac{1}{2}\hbar, \tag{287}$$

which satisfies Heisenberg's uncertainty principle.

C. Operators and wavefunctions

Using the operators it is also possible to obtain the wavefunctions. As noticed above, we should have

$$a\psi_0 = \frac{1}{\sqrt{2}}\left(\xi + \frac{\partial}{\partial\xi}\right)\psi_0 = 0. \tag{288}$$

It is easy to verify that

$$\psi_0 = A_0 e^{-\xi^2/2} \tag{289}$$

satisfies this equation. The requirement that ψ_0 is normalized to unity gives

$$1 = \int_{-\infty}^{\infty} |\psi_0|^2 d\xi = A_0^2 \int_{-\infty}^{\infty} e^{-\xi^2} d\xi = \sqrt{\pi} A_0^2 \qquad \Rightarrow \qquad A_0 = \pi^{1/4}. \tag{290}$$

Having found ψ_0, it is straightforward to find the other wavefunction with the step operators

$$\psi_1 = a^\dagger\psi_0 = \frac{1}{\sqrt{2}\sqrt{\pi}}\left(\xi - \frac{\partial}{\partial\xi}\right)e^{-\xi^2/2} = \frac{1}{\sqrt{2\sqrt{\pi}}} 2\xi e^{-\xi^2/2}. \tag{291}$$

Generally, we have

$$\psi_n = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}}\left(\xi - \frac{\partial}{\partial\xi}\right)^n e^{-\xi^2/2}. \tag{292}$$

This leads to an alternative definition of the Hermite polynomial

$$\left(\xi - \frac{\partial}{\partial\xi}\right)^n e^{-\xi^2/2} = H_n(\xi)e^{-\xi^2/2}. \tag{293}$$

So we find, again, for the total wavefunction

$$\psi_n(\xi) = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}} H_n(\xi) e^{-\xi^2/2}. \tag{294}$$

In the previous section, we obtained the properties of the creation and annihilation operators by studying the characteristics of the operators themselves. However, we can also obtain these by operating them on the wavefunction. Although this is somewhat more complicated, it is instructive to see that we do ge the same results. Let us first look at yet another way of writing defining the Hermite polynomial. The generating function for the Hermite polynomials is

$$g(\xi,t) = e^{-t^2+2t\xi} = \sum_{n=0}^{\infty} H_n(\xi)\frac{t^n}{n!}. \tag{295}$$

This means that is we expand the exponential and collect the terms with the same powers of t then the coefficients in front of them (which are functions of x) are the Hermite polynomials (try it yourself). From the generating function, we can derive a number of relationships between Hermite polynomials of different order. These are called recurrence relations. Differentiating the generating function to ξ gives

$$\frac{dg}{d\xi} = 2te^{-t^2+2t\xi} = \sum_{n=1}^{\infty} H_n'(\xi)\frac{t^n}{n!}. \tag{296}$$

Note that $H_0' = 0$. The left-hand side gives

$$2te^{-t^2+2t\xi} = 2\sum_{n=0}^{\infty} H_n(\xi)\frac{t^{n+1}}{n!} = 2\sum_{n'=1}^{\infty} H_{n'-1}(\xi)\frac{t^{n'}}{(n'-1)!} = \sum_{n=1}^{\infty} 2nH_{n-1}(\xi)\frac{t^n}{n!}. \tag{297}$$

Therefore, we have

$$H_n'(\xi) = 2nH_{n-1}(\xi). \tag{298}$$

Another recursion relation can be obtained by differentiating the generating function with respect to t gives

$$\frac{d}{dt}(e^{-t^2+2t\xi}) \;=\; (-2t+2\xi)e^{-t^2+2t\xi} = \sum_{n=1}^{\infty} H_n(\xi)\frac{nt^{n-1}}{n!}. \tag{299}$$

Then by using the result obtained in Eqn. (297) and by shifting the indices of the right-hand side.

$$-\sum_{n=1}^{\infty}(2n)H_{n-1}\frac{t^n}{n!} + 2\xi\sum_{n=0}^{\infty} H_n\frac{t^n}{n!} = \sum_{n=0}^{\infty} H_{n+1}(\xi)\frac{t^n}{n!} \tag{300}$$

giving

$$H_{n+1} = 2\xi H_n - 2nH_{n-1}. \tag{301}$$

Not let us apply the annihilation operator on the wavefunction in Eqn. (294)

$$a\psi_n(\xi) = \frac{1}{\sqrt{2}}\left(\xi + \frac{d}{d\xi}\right)\psi_n(\xi) \;=\; \frac{1}{\sqrt{2}}\left(\xi\psi_n(\xi) + (-\xi)\psi_n(\xi) + \frac{1}{(2^n n!\pi^{1/2})^{1/2}}e^{-\xi^2/2}H_n'(\xi)\right). \tag{302}$$

Application of the recurrence relation in Eqn. (301) gives

$$\begin{aligned} a\psi_n(\xi) &= \frac{1}{\sqrt{2}}\frac{2n}{(2^n n!\pi^{1/2})^{1/2}}e^{-\xi^2/2}H_{n-1}(\xi) && (303)\\ &= \sqrt{n}\frac{1}{(2^{n-1}(n-1)!\pi^{1/2})^{1/2}}e^{-\xi^2/2}H_{n-1}(\xi) && (304)\\ &= \sqrt{n}\psi_{n-1}(\xi). && (305)\end{aligned}$$

For the creation operator, we find

$$\begin{aligned} a^\dagger\psi_n(\xi) = \frac{1}{\sqrt{2}}\left(\xi - \frac{d}{d\xi}\right)\psi_n(\xi) &= \frac{1}{\sqrt{2}}\frac{1}{(2^n n!\pi^{1/2})^{1/2}}e^{-\xi^2/2}\left(2\xi H_n(\xi) - H_n'(\xi)\right) && (306)\\ &= \frac{1}{\sqrt{2}}\frac{1}{(2^n n!\pi^{1/2})^{1/2}}e^{-\xi^2/2}\left(2\xi H_n(\xi) - 2nH_{n-1}(\xi)\right) && (307)\end{aligned}$$

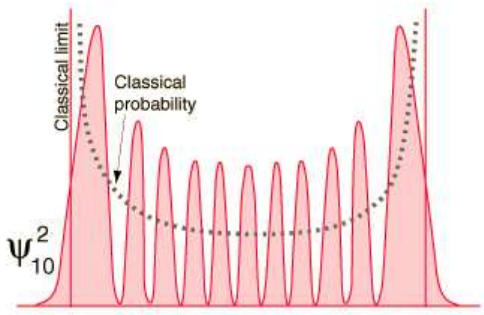


FIG. 12: The classical probability function compared to the probability for a quantum harmonic oscillator for a large value of n. Source: hyperphysics.

Application of the recurrence relation in Eqn. (298) gives

$$a^\dagger \psi_n(\xi) = \frac{1}{(2^{n+1} n! \pi^{1/2})^{1/2}} e^{-\xi^2/2} H_{n+1}(\xi) \tag{308}$$

$$= \sqrt{n+1}\psi_{n+1}(\xi). \tag{309}$$

Note that we obtained the same relationships as in the previous section. However, here we really apply the operator on the wavefunction, which leads to quite a bit of mathematics. In the previous sections, we made use of our knowledge of the behavior of the wavefunctions under the operators without even obtaining the x dependence of the wavefunctions.

D. Correspondence principle

As was shown before, the classical solution of the harmonic oscillator is

$$x = x_0 \sin \omega t \qquad \text{and} \qquad v = \dot{x} = x_0 \omega \cos \omega t, \tag{310}$$

where x_0 is the turning point of the oscillator. We can also write the velocity as

$$v(x) = \omega \sqrt{x_0^2 - x^2}. \tag{311}$$

To find the classical probability $P(x)dx$ of finding a particle at a certain position x, we have to know the time the particle spends in a region dx around x. This is simply given by $dt = dx/v(x)$. Now we have to normalize this to half a period, i.e. π/ω, since we want the particle to pass the interval dx only once. Thus,

$$P(x)dx = \frac{\omega}{\pi} \frac{dx}{v(x)} = \frac{dx}{\pi \sqrt{x_0^2 - x^2}}. \tag{312}$$

This is very close to the quantum-mechanical probability for large n, except for the oscillations. See the figures in Liboff.

Also for large n the energy should become equivalent to the classical energy, i.e $\hbar$ should disappear from the equations. Classicaly, we have $E = \frac{1}{2} m\omega^2 x_0^2$, whereas quantum mechanically, we obtained $E = (n + \frac{1}{2})\hbar\omega$. This looks quite different, but in Eqn. (1212), we found that $\langle x^2 \rangle = \frac{\hbar}{m\omega}(n + \frac{1}{2})$. We can therefore write the energy as

$$E_n = (n + \frac{1}{2})\hbar\omega = m\omega^2 \langle x^2 \rangle_{\rm QM}. \tag{313}$$

The classical expectation value of x^2 is

$$\langle x^2 \rangle_{\rm class.} = \frac{1}{T} \int_0^T dt x_0^2 \sin^2 \omega t = \frac{1}{T} \int_0^T dt x_0^2 (\frac{1}{2} - \frac{1}{2} \cos 2\omega t) = \frac{1}{2} x_0^2. \tag{314}$$

Therefore, the classical energy is $E = m\omega^2 \langle x^2 \rangle_{\rm class.}$. Thus the energies become equal when the classical and quantum mechanical expectation values approach each other. This is indeed true for large n.

VI. THE HYDROGEN ATOM

For the Hydrogen atom we are interested in solving a Schrödinger equation of the form

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = E\psi, \tag{315}$$

with $V(r) = -Ze^2/(4\pi\varepsilon_0 r)$ for a Hydrogen-like atom (a single electron, but the nucleus could still have Z protons). When dealing with atomic physics, the most important operator is angular. Let us first try to understand why this is so. An atom is spherically symmetric. This means that we can rotate the atom along a certain axis and this should not directly affect the eigenstates of the system. That also implies that the Hamiltonian is invariant under rotations. Let us see what happens to a function if we make an infinitesimally small rotation along the z-axis

$$R_{z,d\varphi}\psi(x,y,z) = \psi(x - yd\varphi, y + xd\varphi, z), \tag{316}$$

let us expand this in a Taylor series

$$R_{z,d\varphi}\psi(x,y,z) = \psi(x,y,z) - y\frac{\partial\psi}{\partial x}d\varphi + x\frac{\partial\psi}{\partial y}d\varphi = \left\{1 + \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)d\varphi\right\}\psi(x,y,z) = (1 + \frac{i}{\hbar}L_z d\varphi)\psi(x,y,z) \tag{317}$$

where the angular momentum is given by

$$L_z = (\mathbf{r}\times\mathbf{p})_z = xp_y - yp_x = -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}) \tag{318}$$

Now a rotation over an angle φ can be described by N steps of $d\varphi = \varphi/N$

$$R_{z,\varphi} = \lim_{N\to\infty}\left(1 + \frac{i}{\hbar}L_z\frac{\varphi}{N}\right)^N = \exp(\frac{i}{\hbar}L_z\varphi). \tag{319}$$

Therefore, angular momentum describes a rotation around a certain axis. Obviously, we do not need to make the rotation around the z-axis. A general rotation over $d\boldsymbol{\varphi} = d\varphi_x\hat{\mathbf{x}} + d\varphi_y\hat{\mathbf{y}} + d\varphi_z\hat{\mathbf{z}}$, would lead to a rotation operator,

$$R_{d\boldsymbol{\varphi}} = 1 + \frac{i}{\hbar}d\boldsymbol{\varphi}\cdot\mathbf{L} = 1 + \frac{i}{\hbar}(L_x d\varphi_x + L_y d\varphi_y + L_z d\varphi_z) \tag{320}$$

However, we will see later on that the L_i do not commute with each other and we can only have eigenfunctions for one of the components of L. It is customary to choose L_z.

So now let us suppose we make a unitary transformation of our system (the atom) around the z-axis over an angle $\Delta\varphi$.

$$U\psi(r,\theta,\varphi) = R_{z,\Delta\varphi}\psi(r,\theta,\varphi) = \exp(\frac{i}{\hbar}L_z\Delta\varphi)\psi(r,\theta,\varphi) = \psi(r,\theta,\varphi+\Delta\varphi). \tag{321}$$

Now since a rotation around the z axis is a unitary transformation that does not change the atom, the eigenvalues for the rotated system, should be identical to those of the original system:

$$E_n = \langle\psi_n|H|\psi_n\rangle = \langle\psi_n|U^\dagger HU|\psi_n\rangle. \tag{322}$$

This directly implies that

$$H = U^\dagger HU \qquad \Rightarrow \qquad UH = UU^\dagger HU = HU \qquad \Rightarrow \qquad [H,U] = HU - UH = 0, \tag{323}$$

and therefore that U and H commute, which directly implies that L_z and H commute. However, what does that imply? Suppose we have an operator A that commutes with the Hamiltonian, then $AH = HA$, therefore

$$AH|\psi_n\rangle = E_n A|\psi_n\rangle = H\left(A|\psi_n\rangle\right). \tag{324}$$

This means that $A|\psi_n\rangle$ can only be const $\times\,|\psi_n\rangle$, therefore $|\psi_n\rangle$ must also be an eigenfunction of A.

Since we are dealing with spherical symmetry, i.e.,

$$x = r\sin\theta\cos\varphi \tag{325}$$
$$y = r\sin\theta\sin\varphi \tag{326}$$
$$z = r\cos\theta \tag{327}$$

let us first look at L_z in spherical coordinates. Now since,

$$\frac{\partial}{\partial\varphi} = \frac{\partial x}{\partial\varphi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\varphi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\varphi}\frac{\partial}{\partial z} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}, \tag{328}$$

we can write

$$L_z = -i\hbar\frac{\partial}{\partial\varphi}. \tag{329}$$

We know that the wavefunction ψ must be an eigenfunction of L_z,

$$L_z\psi = m\hbar\psi, \tag{330}$$

where we have taken an eigenvalue of the form $m\hbar$. Since L_z only depends on φ, let us try a separation of variables

$$\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi). \tag{331}$$

This implies that

$$-i\hbar\frac{\partial\Phi(\varphi)}{\partial\varphi} = m\hbar\Phi(\varphi) \qquad \Rightarrow \qquad \frac{\partial\Phi(\varphi)}{\partial\varphi} = im\Phi(\varphi) \qquad \Rightarrow \qquad \Phi(\varphi) = \text{const} \times e^{im\varphi}. \tag{332}$$

Since we want Φ to be single-valued this implies that

$$e^{im(\varphi+2\pi)} = e^{im\varphi} \qquad \Rightarrow \qquad e^{im2\pi} = 1 \qquad \Rightarrow \qquad m = 0, \pm1, \pm2, \cdots, \tag{333}$$

The Φ functions, we want to have orthogonal

$$\int_0^{2\pi} e^{-im'\varphi}e^{im\varphi}d\varphi = 2\pi\delta_{m,m'}, \tag{334}$$

giving $c = 1/\sqrt{2\pi}$.

However, for the $R(r)$ and $\Theta(\theta)$, we need to write down the Hamiltonian in spherical coordinates. From classical arguments we can get some feeling for what the kinetic energy term will look like. Let us have a look at $\mathbf{L}^2$,

$$\mathbf{L}^2 = r^2p^2\sin^2\theta = r^2p^2 - r^2p^2\cos^2\theta = r^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})^2 \qquad \Rightarrow \qquad \frac{\mathbf{p}^2}{2m} = \frac{\mathbf{p}_r^2}{2m} + \frac{\mathbf{L}^2}{2mr^2}, \tag{335}$$

where $p_r = \mathbf{p}\cdot\mathbf{r}$ is the momentum in the r-direction. This means that $\mathbf{L}^2$ is important in the classical case, can we expect something similar in the quantum mechanical situation. We note that the components of L commute with the Hamiltonian. What does this imply for $\mathbf{L}^2$. Since $[L_i, H] = 0$, we can write

$$L_i[L_i,H] + [L_i,H]L_i = L_i^2H - L_iHL_i + L_iHL_i - HL_i^2 = L_i^2H - HL_i^2 = [L_i^2, H]. \tag{336}$$

Thus L_i^2 also commute with the Hamiltonian. A linear combination of L_i^2 should therefore also commute with the Hamiltonian, which implies that $[\mathbf{L}^2, H] = 0$. Therefore, the eigenfunctions of the Hamiltonian for the atom are also eigenfunctions of $\mathbf{L}^2$. Let us see if we can find the quantum mechanical equivalent of Eqn. (335). In quantum mechanics, we are dealing with operators and have to take care that r and p do not commute. First, we have

$$\begin{aligned}
\mathbf{L}^2 &= |\mathbf{r}\times\mathbf{p}|^2 = (xp_y - yp_x)^2 + \text{c.p.} && (337)\\
&= xp_yxp_y + yp_xyp_x - xp_yyp_x - yp_xxp_y + \text{c.p.} && (338)\\
&= x^2p_y^2 + y^2p_x^2 - 2xyp_xp_y + i\hbar(xp_x + yp_y) + \text{c.p.} && (339)
\end{aligned}$$

where use has been made of the commutation relation $[x, p_x] = i\hbar$; c.p. stands for cyclic permutations, i.e., $x \to y \to z$. On the other hand, we have

$$\begin{aligned}
r^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})^2 &= (x^2+y^2+z^2)(p_x^2+p_y^2+p_z^2) - (xp_x + yp_y + zp_z)^2 && (340)\\
&= x^2p_x^2 + x^2p_y^2 + y^2p_x^2 + \text{c.p.} - (xp_xxp_x + xp_xyp_y + yp_yxp_x + \text{c.p.}) && (341)\\
&= x^2p_y^2 + y^2p_x^2 - 2xyp_xp_y + i\hbar xp_x + \text{c.p.}. && (342)
\end{aligned}$$

Collecting the results, we obtain in the quantum mechanical case

$$\mathbf{L}^2 = r^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})^2 + i\hbar(\mathbf{r}\cdot\mathbf{p}), \tag{343}$$

where we obtain an extra term from the fact that r and p do not commute. Let us try to rewrite the result somewhat. For $\mathbf{r}\cdot\mathbf{p}$, we have

$$\mathbf{r}\cdot\mathbf{p} = -i\hbar\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right) = -i\hbar r\frac{\partial}{\partial r}, \tag{344}$$

using the fact that

$$r = \sqrt{x^2+y^2+z^2} \qquad \Rightarrow \qquad x\frac{\partial}{\partial x} = x\frac{\partial r}{\partial x}\frac{\partial}{\partial r} = \frac{x^2}{r}\frac{\partial}{\partial r} \tag{345}$$

We can use this to rewrite

$$\begin{aligned}(\mathbf{r}\cdot\mathbf{p})^2 - i\hbar\mathbf{r}\cdot\mathbf{p} &= (-i\hbar r\frac{\partial}{\partial r})(-i\hbar r\frac{\partial}{\partial r}) - i\hbar(-i\hbar r\frac{\partial}{\partial r}) &(346)\\ &= -\hbar^2 r\frac{\partial}{\partial r}(r\frac{\partial}{\partial r}) - \hbar^2 r\frac{\partial}{\partial r} = -\hbar^2\left\{r^2\frac{\partial^2}{\partial r^2} + 2r\frac{\partial}{\partial r}\right\}. &(347)\end{aligned}$$

Let us now introduce

$$p_r = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r = -i\hbar(\frac{\partial}{\partial r} + \frac{1}{r}). \tag{348}$$

From this, we obtain that

$$r^2p_r^2 = -\hbar^2(r^2\frac{\partial^2}{\partial r^2} + r\frac{\partial}{\partial r} + r^2\frac{\partial}{\partial r}\frac{1}{r} + 1) = -\hbar^2\left\{r^2\frac{\partial^2}{\partial r^2} + 2r\frac{\partial}{\partial r}\right\}. \tag{349}$$

This is equivalent to the result obtained above. Therefore, in the quantum mechanical case, we can also write

$$\mathbf{p}^2 = p_r^2 + \frac{\mathbf{L}^2}{r^2}, \tag{350}$$

if we use the definition for p_r in Eqn. (348).

Could we have seen more easily to use this definition for p_r? We want to know the component of $\mathbf{p}$ in the r direction, i.e., $\hat{\mathbf{r}}\cdot\mathbf{p}$. However, this is not very symmetric since we could also write $\mathbf{p}\cdot\hat{\mathbf{r}}$. So let us try a more symmetric form:

$$p_r = \frac{1}{2}\left\{\frac{\mathbf{r}}{r}\cdot\mathbf{p} + \mathbf{p}\cdot\frac{\mathbf{r}}{r}\right\}. \tag{351}$$

From the definition of the gradient in spherical symmetry,

$$\nabla f = \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial\varphi}\hat{\varphi}, \tag{352}$$

we see that the first term in Eqn. (351) is given by

$$\frac{\mathbf{r}}{r}\cdot\mathbf{p} = -i\hbar\frac{\partial}{\partial r}. \tag{353}$$

The second term would give

$$\mathbf{p}\cdot\mathbf{r}\frac{1}{r}\psi = -i\hbar\nabla(\frac{\mathbf{r}}{r}\psi) = -i\hbar\left\{\nabla\left(\frac{\mathbf{r}}{r}\right)\psi + \frac{\mathbf{r}}{r}\nabla\psi\right\}. \tag{354}$$

Using that

$$\nabla\left(\frac{\mathbf{r}}{r}\right) = \frac{\nabla\cdot\mathbf{r}}{r} + \mathbf{r}\cdot\left(-\frac{\hat{\mathbf{r}}}{r^2}\right) = \frac{1}{r}\left\{\frac{2}{r}r + \frac{\partial}{\partial r}r\right\} - \frac{1}{r} = \frac{3}{r} - \frac{1}{r} = \frac{2}{r}, \tag{355}$$

where use has been made of the definition of the divergence in spherical coordinates. Collecting the terms, we obtain for p_r

$$p_r = -i\hbar \left\{ \frac{\partial}{\partial r} + \frac{1}{r} \right\}, \tag{356}$$

which is identical to Eqn. (348).

Of course, all of this could also have been obtained by starting from the angular momentum in spherical coordinates:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = -i\hbar \begin{vmatrix} \hat{\mathbf{r}} & \hat{\boldsymbol{\theta}} & \hat{\boldsymbol{\varphi}} \\ r & 0 & 0 \\ \frac{\partial}{\partial r} & \frac{1}{r}\frac{\partial}{\partial \theta} & \frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi} \end{vmatrix} = -i\hbar \left(-\hat{\boldsymbol{\theta}} \frac{1}{\sin\theta}\frac{\partial}{\partial \varphi} + \hat{\boldsymbol{\varphi}} \frac{\partial}{\partial \theta} \right). \tag{357}$$

Note that the angular momentum is perpendicular to $\mathbf{r}$ and has therfore no $\hat{\mathbf{r}}$ component. What is now the angular momentum in Cartesian coordinates. Therefore, we use

$$\hat{\boldsymbol{\theta}} = \hat{\mathbf{x}} \cos\theta \cos\varphi + \hat{\mathbf{y}} \cos\theta \sin\varphi - \hat{\mathbf{z}} \sin\theta \tag{358}$$

$$\hat{\boldsymbol{\varphi}} = -\hat{\mathbf{x}} \sin\varphi + \hat{\mathbf{y}} \cos\varphi. \tag{359}$$

This gives

$$\begin{aligned} L_x &= i\hbar \left(\sin\varphi \frac{\partial}{\partial \theta} + \cot\theta \cos\varphi \frac{\partial}{\partial \varphi} \right) \\ L_y &= i\hbar \left(-\cos\varphi \frac{\partial}{\partial \theta} + \cot\theta \sin\varphi \frac{\partial}{\partial \varphi} \right) \\ L_z &= -i\hbar \frac{\partial}{\partial \varphi}. \end{aligned} \tag{360}$$

From what we have seen above, it is obvious to look at $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$,

$$\mathbf{L}^2 \psi = -\hbar^2 \left[\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta \frac{\partial \psi}{\partial \theta}) + \frac{1}{\sin^2\theta}\frac{\partial^2 \psi}{\partial \varphi^2} \right]. \tag{361}$$

Comparing this with

$$\nabla^2 \psi = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial \psi}{\partial r}) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}(\sin\theta \frac{\partial \psi}{\partial \theta}) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 \psi}{\partial \varphi^2}. \tag{362}$$

we note that the angular part can be written as $\mathbf{L}^2/r^2$, and thus

$$-\frac{\hbar^2}{2m}\nabla^2 \psi = -\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\mathbf{L}^2}{2mr^2}\psi = \frac{p_r^2}{2m}\psi + \frac{\mathbf{L}^2}{2mr^2}\psi. \tag{363}$$

Note that we can write the r term in different ways

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial \psi}{\partial r}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) = \frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r}\frac{\partial \psi}{\partial r}. \tag{364}$$

A. Legendre polynomials

Let us try to find the eigenvalues of the eigenvalue equation

$$\mathbf{L}^2 \psi = \alpha \hbar^2 \psi. \tag{365}$$

We assumed already that we could separate the variables in ψ, see Eqn. (331). Since $\mathbf{L}^2$ only depends on θ and φ, and since we know that the φ-dependent part is given by $e^{im\varphi}/\sqrt{2\pi}$, we can write the differential equation for $\Theta(\theta)$ as

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}(\sin\theta \frac{d\Theta}{d\theta}) + \frac{m^2}{\sin^2\theta}\Theta = \alpha\Theta \tag{366}$$

This equation is known as the Legendre equation. This can be put in a different form by substituting $x = \cos\theta$. This implies that

$$\frac{d\Theta}{dx} = \frac{d\Theta}{d\theta}\frac{d\theta}{dx} = \frac{d\Theta}{d\theta}\frac{1}{dx/d\theta} = -\frac{1}{\sin\theta}\frac{d\Theta}{d\theta} \qquad \text{or} \qquad \frac{d\Theta}{d\theta} = -\sin\theta\frac{d\Theta}{dx}. \tag{367}$$

We can write the DE as

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left((1-x^2)\frac{d\Theta}{dx}\right) + (\alpha - \frac{m^2}{1-x^2})\Theta = \frac{d}{dx}\left((1-x^2)\frac{d\Theta}{dx}\right) + (\alpha - \frac{m^2}{1-x^2})\Theta = 0. \tag{368}$$

This can also be written as

$$(1-x^2)\frac{d^2\Theta}{dx^2} - 2x\frac{d\Theta}{dx} + (\alpha - \frac{m^2}{1-x^2})\Theta = 0. \tag{369}$$

Let us first consider $m = 0$. We will find later on that the other solutions can be related to this solution. As with the harmonic oscillator, we try to find the solution in form of a series:

$$\Theta(x) = P(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots, \tag{370}$$

where the Legendre polynomial is conventionally denoted as $P(x)$. This gives

$$\alpha P(x) = \sum_{n=0}^{\infty} \alpha a_n x^n \tag{371}$$

$$-2x\frac{dP}{dx} = -2x\sum_{n=1}^{\infty} n a_n x^{n-1} = -2\sum_{n=1}^{\infty} n a_n x^n \tag{372}$$

$$(1-x^2)\frac{d^2P}{dx^2} = (1-x^2)\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} = \sum_{n=0}^{\infty}(n+1)(n+2)a_{n+2}x^n - \sum_{n=2}^{\infty} n(n-1)a_n x^n. \tag{373}$$

Using this we can write

$$\begin{aligned}(1-x^2)\frac{d^2P}{dx^2} - 2x\frac{dP}{dx} + \alpha P &= \alpha a_0 + \alpha a_1 x - 2a_1 x + 2a_2 + 6a_3 x \\ &\quad + \sum_{n=2}^{\infty}\left\{\alpha a_n - 2n a_n + (n+2)(n+1)a_{n+2} - n(n-1)a_n\right\}x^n = 0\end{aligned} \tag{374}$$

This gives the general equation

$$(\alpha - n(n+1))a_n + (n+2)(n+1)a_{n+2} = 0. \tag{375}$$

The other equations $\alpha a_0 + 2a_2 = 0$ and $\alpha_1 - 2a_1 + 6a_3 = 0$ are the same equation for $n = 0$ and 1, respectively. Thus we have the recursion relation

$$a_{n+2} = \frac{n(n+1) - \alpha}{(n+2)(n+1)}a_n. \tag{376}$$

However, just as we found for the harmonic oscillator, this series diverges. Therefore, we have to cut off the series at a certain value of n. When taking $\alpha = l(l+1)$, the series becomes zero for $n > l$. Note that this implies

$$\mathbf{L}^2\psi = l(l+1)\hbar^2\psi. \tag{377}$$

A Legendre polynomial of order l is denoted as $P_l(x)$. As for the harmonic oscillator, we have again two series. By choosing $\alpha = l(l+1)$, we can cut off the series of the even or the odd powers, but not both. Therefore when chosing to cut off the even powers, we have to choose all the coefficients in front of the odd powers equal to zero, and vice versa. The normalization of the Legendre polynomial is $P(1) = 1$. The first terms can be derived rather easily. For $k = 0$, we only have a_0, which should be unity, giving $P_0 = 1$. For $k = 1$, $\alpha = 2$, and we have $P_1(x) = a_1 x \to x$. For $k = 2$, we have the even series, with $\alpha = k(k+1) = 6$. The recursion relation gives $a_2 = -6/2a_0 = -3a_0$, yielding

$P_2(x) = a_0 - 3a_0x^2$. Using $P_2(1) = -2a_0 = 1$ gives $a_0 = -1/2$, which gives $P_2(x) = 3/2x^2 - 1/2$. For $P_3(x)$, $\alpha = 12$, giving $a_3 = (2-12)/(2.3) = -5/3$. Thus $P_3(x) = a_3x - 5/3a_3x^3$. The normalization gives $P_3(1) = 2/3a_3$. With $a_3 = -3/2$, we find $P_3(x) = 1/2(5x^3 - 3x)$. In order to normalize the wavefunctions, we have to integrate over the Legendre polynomials from -1 to 1 (note, that x comes from $\cos\theta$). It can be shown that this gives

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{m,n}, \tag{378}$$

where Legendre polynomials of different order are orthogonal.

Remember that we had put m to zero. Actually, we want to solve the full equation

$$(1-x^2)\frac{d^2P_n^m(x)}{dx^2} - 2x\frac{dP_n^m(x)}{dx} + \left[n(n+1) - \frac{m^2}{1-x^2}\right]P_n^m(x) = 0. \tag{379}$$

It can be shown that the associated Legendre polynomials are related to the m'th derivative of the Legendre polynomial,

$$P_n^m(x) = (1-x^2)^{m/2}\frac{d^m}{dx^m}P_n(x), \tag{380}$$

for positive m, with $m \leq l$. Note that by solving the differential equation for the associated Legendre polynomials, we find that $|m| \leq l$ or $m = -l, -l+1, \ldots, l-1, l$. The associated Legendre polynomials for negative m are equal to those for positive m apart from a multiplicative constant. In atomic physics, one often makes use of the spherical harmonics which are related to the associated Legendre polynomials as

$$Y_{lm}(\theta,\varphi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\varphi}, \tag{381}$$

where the normalization has been chosen such that

$$\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta Y^*_{l'm'}(\theta,\varphi)Y_{lm}(\theta,\varphi) = \delta_{l,l'}\delta_{m,m'} \tag{382}$$

B. Angular momentum

As was shown before, as a result of the spherical symmetry of the problem L_i and $\mathbf{L}^2$ commute with the Hamiltonian. The commutation relationships of angular momentum are given by

$$\begin{aligned}
[L_x, L_y] &= L_xL_y - L_yL_x = (yp_z - zp_y)(zp_x - xp_z) - (zp_x - xp_z)(yp_z - zp_y) && (383)\\
&= yp_zzp_x - yp_zxp_z - zp_yzp_x + zp_yxp_z - zp_xyp_z + zp_xzp_y + xp_zyp_z - xp_zzp_y && (384)\\
&= xp_y(zp_z - p_zz) - yp_x(zp_z - p_zz) = i\hbar(xp_y - yp_x) = i\hbar L_z. && (385)
\end{aligned}$$

By cyclic permutation, we also obtain

$$[L_y, L_z] = i\hbar L_x \qquad \text{and} \qquad [L_z, L_x] = i\hbar L_y, \tag{386}$$

often $\hbar$ is taken equal to unity or one assumes that L_i is measured in units of $\hbar$. Can we make a wavefunction ψ that has eigenvalues in all operators L_x, L_y, and L_z? Let us apply this to an eigenfunction which has an eigenvalue m_z

$$iL_z\psi = im_z\psi = (L_xL_y - L_yL_x)\psi = (L_xm_y - L_ym_x)\psi = \hbar(m_xm_y - m_ym_x)\psi = 0. \tag{387}$$

When $m_z \neq 0$ this implies ψ=0 or that we cannot have eigenfunctions in L_x, L_y, and L_z at the same time. What about $\mathbf{L}^2$? Making use of the commutation relations, we find

$$\begin{aligned}
\mathbf{L}^2L_z - L_z\mathbf{L}^2 &= L_x^2L_z - L_zL_x^2 + L_y^2L_z - L_zL_y^2 + L_z^2L_z - L_zL_z^2 && (388)\\
&= L_x(L_zL_x - iL_y) - (L_xL_z + iL_y)L_x + L_y(L_zL_y + iL_x) - (L_yL_z - iL_x)L_y = 0
\end{aligned}$$

Thus we have eigenvector with eigen values for L_z and $\mathbf{L}^2$. This means we have

$$L_z\psi_{bm} = m\psi_{bm} \Rightarrow L_z^2\psi_{bm} = L_zm\psi_{bm} = m^2\psi_{bm} \tag{389}$$

Now we also have

$$\mathbf{L}^2\psi_{bm} = b\psi_{bm} \Rightarrow (L_x^2 + L_y^2)\psi_{bm} = (b - m^2)\psi_{bm}. \tag{390}$$

Since the operator $L_x^2 + L_y^2$ is positive, this implies that $|m| \leq \sqrt{b}$. Let us now consider the effect of the operator

$$L_z(L_x \pm iL_y) = L_xL_z + iL_y \pm i(L_yL_z - iL_x) = (L_x \pm iL_y)(L_z \pm 1). \tag{391}$$

Operating this on an eigenfunction gives

$$L_z(L_x \pm iL_y)\psi_{bm} = (L_x \pm iL_y)(L_z \pm 1)\psi_{bm} = (L_x \pm iL_y)(m \pm 1)\psi_{bm} = (m \pm 1)(L_x \pm iL_y)\psi_{bm}.$$

So $(L_x \pm iL_y)\psi_{bm}$ are eigenfunctions of L_z with eigenvalues $m \pm 1$. Moreover, since L_x and L_y commute with $\mathbf{L}^2$, they are also eigenfunctions of $\mathbf{L}^2$. So starting from a particular eigenfunction we can create a set of eigenfunctions that with eigenvalues

$$m' = \cdots, m-2, m-1, m, m+1, m+2, \cdots \tag{392}$$

However, since $|m| \leq \sqrt{b}$, there have to be certain values of m for which

$$\begin{aligned}(L_x - iL_y)\psi_{bl_-} &= 0 &(393)\\ (L_x + iL_y)\psi_{bl_+} &= 0 &(394)\\ &&(395)\end{aligned}$$

From the commutation relations we find

$$\begin{aligned}(L_x + iL_y)(L_x - iL_y) &= L_x^2 + L_y^2 + L_z &(396)\\ (L_x - iL_y)(L_x + iL_y) &= L_x^2 + L_y^2 - L_z &(397)\\ &&(398)\end{aligned}$$

These operators will give us eigenvalues

$$\begin{aligned}(L_x^2 + L_y^2 + L_z)\psi_{bl_-} &= (b - l_-^2 + l_-)\psi_{bl_-} = 0 &(399)\\ (L_x^2 + L_y^2 - L_z)\psi_{bl_+} &= (b - l_+^2 - l_+)\psi_{bl_+} = 0 &(400)\\ &&(401)\end{aligned}$$

giving

$$(b - l_-^2 + l_-) = (b - l_+^2 - l_+) = 0 \tag{402}$$

or

$$(l_+ + l_-)(l_+ - l_- + 1) = 0. \tag{403}$$

Since $l_+ \geq l_-$, we find that $l_- = -l_+$ in order to make this equation zero. This directly implies that $b = l(l+1)$. Therefore we have the

$$\begin{aligned}\mathbf{L}^2\psi_{lm} &= l(l+1)\psi_{lm} &(404)\\ L_z\psi_{lm} &= m\psi_{lm}, &(405)\end{aligned}$$

as we already found from the differential equation. Note that we found the values for m without having to invoke an argument about the single-valuedness of the function.

As we have already seen that operators can connect different eigenstates

$$L_\pm\psi_{lm} = (L_x \pm iL_y)\psi_{lm} = N_\pm\psi_{l,m\pm1}. \tag{406}$$

What is now the value of $N_\pm$

$$\begin{aligned}|N_\pm|^2 &= \langle\psi_{lm}|L_\pm^\dagger L_\pm|\psi_{lm}\rangle &(407)\\ &= \langle\psi_{lm}|L_\mp L_\pm|\psi_{lm}\rangle &(408)\\ &= \langle\psi_{lm}|L_x^2 + L_y^2 \mp L_z|\psi_{lm}\rangle &(409)\\ &= \langle\psi_{lm}|l(l+1) - m^2 \mp m|\psi_{lm}\rangle &(410)\\ &= (l \mp m)(l \pm m + 1) &(411)\end{aligned}$$

giving a value

$$N_{\pm} = \sqrt{(l \mp m)(l \pm m + 1)} e^{i\delta}, \tag{412}$$

by convention we choose $\delta = 0$.

We have learnt already a lot about the eigenvalue problem of the system without actually having to solve a differential equation. Just as for the harmonic oscillator, we can also use the operators to determine the wavefunctions. We know that applying the step-down operator cannot last forever since $|m| \leq l$, thus

$$(L_x - iL_y)Y_{l,-l}(\theta, \varphi) = (L_x - iL_y)\Theta_{l,-l}(\theta, \varphi)\frac{1}{\sqrt{2\pi}}e^{-il\varphi} = 0. \tag{413}$$

Using the expression for L_x and L_y in Eqn.(360), we can show that

$$L_{\pm} = (L_x \pm iL_y) = e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right]. \tag{414}$$

This gives us a differential equation

$$-\frac{d\Theta_{l,-l}}{d\theta} + l\cot\theta\Theta_{l,-l} = 0. \tag{415}$$

Let us try as a solution

$$\Theta_{l,-l} = c\sin^l\theta, \tag{416}$$

see Liboff Table 9.1. This gives

$$-cl\sin^{l-1}\theta\cos\theta + cl\frac{\cos\theta}{\sin\theta}\sin^l\theta = 0, \tag{417}$$

which shows that this is indeed a correct eigenfunction. By applying the step-up operators, all the Θ_{lm} can be obtained. This is a somewhat tedious process and will be omitted here.

C. Angular momentum in matrix form

In order to write down a matrix for angular momentum operators, we first need to choose a basis set. Although we can choose the solution of the Hydrogen atom as a basis set, we have to remember that the angular momentum operators do not work on the radial part, but only on the spherical harmonics $Y_{lm}(\theta, \varphi)$. In addition, angular momentum operators do not couple spherical harmonics with each other that have different l values. Therefore, we can consider the spherical harmonics for one particular l value. Let us begin with the step-operators:

$$L_{\pm}|lm\rangle = \hbar\sqrt{(l \mp m)(l \pm m + 1)}|l, m \pm 1\rangle. \tag{418}$$

In the case that $l = 1$, we can write down the matrices straightforwardly as

$$\mathrm{L}_+ = \hbar\sqrt{2}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathrm{L}_- = \hbar\sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{419}$$

The functions now become the colums

$$|11\rangle \to \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |10\rangle \to \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |1, -1\rangle \to \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{420}$$

where the matrix elements are obtained from

$$(\mathrm{L}_{\pm})_{m'm} = \langle lm'|L_{\pm}|lm\rangle \tag{421}$$

Using these matrices, we obtain, for example,

$$L_+|1,-1\rangle = \hbar\sqrt{2}|10\rangle \qquad \rightarrow \qquad \hbar\sqrt{2}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \hbar\sqrt{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{422}$$

From the step matrices for the step operators, we can obtain the matrices for L_x and L_y,

$$\mathrm{L}_x = \frac{1}{2}(\mathrm{L}_+ + \mathrm{L}_-) = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathrm{L}_y = \frac{1}{2i}(\mathrm{L}_+ - \mathrm{L}_-) = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{423}$$

For L_z, we have

$$\mathrm{L}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{424}$$

Note that the matrix elements can also be evaluated from

$$(\mathrm{L}_z)_{m,m'} = \langle 1m|L_z|1m'\rangle = \int Y^*_{1m}\left(-i\hbar\frac{\partial}{\partial\varphi}\right)Y_{lm'}\sin\theta d\theta d\varphi = m\hbar\int Y^*_{1m}Y_{lm'}\sin\theta d\theta d\varphi = m\hbar\delta_{m,m'}.$$

Note that the commutation relations,

$$[\mathrm{L}_x, \mathrm{L}_y] = i\hbar L_z. \tag{425}$$

can also be derived by simple matrix multiplication.

As in Eqn. (320) we can build a finite rotation over φ of a function can be built up of infinitesimally small rotations. For $l = 1$

$$\mathrm{R}_z(d\varphi) = (\mathbf{1} + id\varphi\mathrm{L}_z) \tag{426}$$

or for $d\varphi = \varphi/N$ with $N \to \infty$,

$$\mathrm{R}_z(\varphi) = (\mathbf{1} + i\frac{\varphi}{N}\mathrm{L}_z)^N = \exp(i\varphi\mathrm{L}_z). \tag{427}$$

The exponential may be expanded in a Taylor series

$$\exp(i\varphi\mathrm{L}_z) = \mathbf{1} + i\varphi\mathrm{L}_z + \frac{1}{2!}(i\varphi\mathrm{L}_z)^2 + \frac{1}{3!}(i\varphi\mathrm{L}_z)^3 + \cdots \tag{428}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\{1 - \frac{1}{2!}\varphi^2 + \frac{1}{4!}\varphi^4 - \cdots\} \tag{429}$$

$$+i\mathrm{L}_z\{\varphi - \frac{1}{3!}\varphi^3 + \frac{1}{5!}\varphi^5 - \cdots\} \tag{430}$$

$$\tag{431}$$

(note that the first matrix is $\mathbf{1} - \mathrm{L}_z^2$). We have used here that

$$\mathrm{L}_z^2 = \hbar^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \mathrm{L}_z^3 = \mathrm{L}_z. \tag{432}$$

The series may be recognized as the Taylor series for a cosine and sine, respectively, and we can also write the rotation as

$$\mathrm{R}_z(\varphi) = \begin{pmatrix} \cos\varphi + i\sin\varphi & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \cos\varphi - i\sin\varphi \end{pmatrix} = \begin{pmatrix} e^{i\varphi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\varphi} \end{pmatrix}. \tag{433}$$

This can be clearly seen as the rotation of a vector over an angle $\Delta\varphi$ around the z-axis. Thus a rotation of the function $\Phi_m(\varphi)$ gives

$$R_z(\Delta\varphi)\Phi_m(\varphi) = R_z(\Delta\varphi)\frac{1}{\sqrt{2\pi}}e^{im\varphi} = e^{im\Delta\varphi}\frac{1}{\sqrt{2\pi}}e^{im\varphi} = \frac{1}{\sqrt{2\pi}}e^{im(\varphi+\Delta\varphi)} = \Phi_m(\varphi + \Delta\varphi) \tag{434}$$

VII. LAGUERRE POLYNOMIALS

For the Hydrogen atom, we are left with the Hamiltonian

$$H = \frac{p_r^2}{2m} + \frac{\mathbf{L}^2}{2mr^2} - \frac{Ze^2}{r}, \tag{435}$$

where, for brevity, $e \to e/\sqrt{4\pi\varepsilon_0}$ for SI units. The $\mathbf{L}^2$ term one sometimes calls the centrifugal term, i.e., it tends to throw the particle outward, as the centrifugal force in quantum mechanics. This gives us a radial equation

$$\left\{-\frac{\hbar^2}{2m}\left(\frac{1}{r}\frac{d^2}{dr^2}r\right) + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{Ze^2}{r} + |E|\right\} R = 0, \tag{436}$$

Introducing the radial function $u = rR$ gives

$$\left\{-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} - \frac{2m}{\hbar^2}\left(\frac{Ze^2}{r} - |E|\right)\right\} u = 0, \tag{437}$$

where we have taken $E = -|E|$. Let us try to simplify the equation by introducing

$$|E| = \frac{\hbar^2\kappa^2}{2m} \qquad \text{and} \qquad \rho = 2\kappa r \tag{438}$$

$$\lambda^2 = \left(\frac{Z}{\kappa a_0}\right)^2 = \frac{Z^2 R_0}{|E|}, \tag{439}$$

with $R_0 = \hbar^2/2ma_0^2$ the Rydberg constant and $a_0 = \hbar^2/me^2$ the Bohr radius. This gives us the following differential equation

$$\frac{d^2u}{d\rho^2} - \frac{l(l+1)}{\rho^2}u + \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)u = 0. \tag{440}$$

Let us take a look at the limiting behavior. When $\rho \to \infty$, we have the approximate differential equation

$$\frac{d^2u}{d\rho^2} - \frac{u}{4} = 0. \tag{441}$$

This can be easily solved:

$$u \cong Ae^{-\rho/2} + Be^{\rho/2}. \tag{442}$$

However, we do not want our particle to sit in infinity so we want $u \to 0$, giving $B = 0$.

For $\rho \to 0$, we obtain

$$\frac{d^2u}{d\rho^2} - \frac{l(l+1)}{\rho^2}u = 0. \tag{443}$$

Trying a solution of $u = \rho^q$ gives

$$(q(q-1) - l(l+1))u = 0 \qquad \Rightarrow \qquad q = -l \quad \text{or} \quad q = l+1. \tag{444}$$

Thus

$$u \sim A\rho^{-l} + B\rho^{l+1}. \tag{445}$$

We also do not want our electron to sit at the nucleus, which gives $A = 0$, or $u \sim \rho^{l+1}$. Let us try to find a solution of the form

$$u(\rho) = e^{-\rho/2}\rho^{l+1}F(\rho). \tag{446}$$

Writing $u = AF$ with $A = e^{-\rho/2}\rho^{l+1}$, the second derivative is given by $AF'' + 2A'F' + A''F$, with

$$A'(\rho) = e^{-\rho/2}\left(-\frac{1}{2}\right)\rho^{l+1} + e^{-\rho/2}\rho^l(l+1) = \left(-\frac{1}{2} + \frac{l+1}{\rho}\right)A(\rho) \tag{447}$$

$$A''(\rho) = e^{-\rho/2}\frac{1}{4}\rho^{l+1} + 2\times\left(-\frac{1}{2}\right)e^{-\rho/2}\rho^l(l+1) + e^{-\rho/2}\rho^{l-1}l(l+1) = \left(\frac{1}{4} - \frac{l+1}{\rho} + \frac{l(l+1)}{\rho^2}\right)A(\rho). \tag{448}$$

Substituting this into the differential equation and dividing by $A(\rho)$ gives

$$\left[F'' + 2\left(-\frac{1}{2} + \frac{l+1}{\rho}\right)F' + \left(\frac{1}{4} - \frac{l+1}{\rho} + \frac{l(l+1)}{\rho^2}\right)F\right] - \frac{l(l+1)}{\rho^2}F + \left(\frac{\lambda}{\rho} - \frac{1}{4}\right)F = 0. \tag{449}$$

Cleaning this up and multiplying with ρ gives

$$\rho F''(\rho) + (2l + 2 - \rho)F'(\rho) - (l + 1 - \lambda)F = 0. \tag{450}$$

Again, we will approach this with a series solution

$$F(\rho) = \sum_{k=0}^{\infty} c_k \rho^k. \tag{451}$$

The first and second derivative are given by

$$F'(\rho) = \sum_{k=1}^{\infty} c_k k \rho^{k-1} \qquad \text{and} \qquad F''(\rho) = \sum_{k=2}^{\infty} c_k k(k-1)\rho^{k-2}. \tag{452}$$

We would like to make all the powers in the summation equal to k. This can be done by shifting the index. So,

$$\rho F''(\rho) = \sum_{k=2}^{\infty} c_k k(k-1)\rho^{k-1} = \sum_{k=1}^{\infty} c_{k+1}(k+1)k\rho^k \tag{453}$$

(note that we can also let the summation start at $k = 0$). and

$$(2l+2)F'(\rho) = (2l+2)\sum_{k=-1}^{\infty} c_{k+1}(k+1)\rho^k. \tag{454}$$

Collecting the term gives

$$\sum_{k=0}^{\infty} \{c_{k+1}(k+1)k + (k+1)(2l+2)c_{k+1} - kc_k - (l+1-\lambda)c_k\}\rho^k \;=\; \tag{455}$$

$$\sum_{k=0}^{\infty} \{(k+1)(2l+2+k)c_{k+1} - (k+l+1-\lambda)c_k\}\rho^k \;=\; 0. \tag{456}$$

Since all the ρ^k are independent function the only way to make this zero is by having all the coefficients in front of all powers of ρ^k zero. This gives the recursion relation

$$c_{k+1} = \frac{(k+l+1)-\lambda}{(k+1)(k+2l+2)}c_k. \tag{457}$$

In the limit that $k \to \infty$, this becomes

$$c_{k+1} \cong \frac{c_k}{k} = \frac{c_k}{k}\frac{c_{k-1}}{k-1}\cdots c_0 = \frac{c_0}{k!}, \tag{458}$$

giving for F

$$F = \sum_{k=0}^{\infty} c_0 \frac{\rho^k}{k!} = c_0 e^{\rho}. \tag{459}$$

The total solution,

$$u = e^{-\rho/2}\rho^{l+1}e^{\rho} = e^{\rho/2}\rho^{l+1} \tag{460}$$

then diverges as $\rho \to \infty$. Therefore, just as in the case of the harmonic oscillator and the Legendre polynomials, we have to end the series at the certain point, so that it remains a simple polynomial. This can be done by taking

$$\lambda = k_{\max} + l + 1, \tag{461}$$

i.e., since both $k_{\max}$ and l are integers, λ is integer. Setting $\lambda_n = n$, we obtain from Eqn. (439)

$$\lambda_n^2 = n^2 = \frac{Z^2 R_0}{|E_n|} \qquad \Rightarrow \qquad E_n = -|E_n| = -\frac{Z^2 R_0}{n^2}, \tag{462}$$

which is in agreement with Bohr's result. The radial wavefunctions are given by

$$u_{nl}(\rho) = e^{-\rho/2}\rho^{l+1}F_{nl}(\rho) = A_{nl}e^{-\rho/2}\rho^{l+1}\sum_{k=0}^{n-l-1} c_k \rho^k. \tag{463}$$

The function F_{nl} is the associated Laguerre polynomial, which is usually denoted as L_{n-l-1}^{2l+1}. Note that the maximum value for l is given by $l_{\max} = n - 1$. We can therefore summarize the quantum number as

$$n = 1, 2, 3, \ldots \tag{464}$$

$$l = 0, 1, 2, \ldots, n-1 \tag{465}$$

$$m = -l, -l+1, \ldots, l \tag{466}$$

and the total wavefunction of an orbital with a set of quantum numbers nlm is

$$\psi(r, \theta, \varphi) = R_{nl}(r)Y_{lm}(\theta, \varphi), \tag{467}$$

with an energy

$$E_n = -\frac{Z^2 R}{n^2} = \frac{mZ^2 e^4}{2\hbar^2 n^2}. \tag{468}$$

VIII. RELATIVISTIC QUANTUM MECHANICS (LIBOFF CHAPTER 15)

Let us have a brief look at relativistic quantum mechanics. We will not deal with relativistic quantum mechanics in detail. Our main interest is to discover the most important correction to the nonrelativistic limit. We will see how the spin is introduced in the relativistic equation, how the spin interacts with the magnetic field (obtaining the gyromagnetic factor $g = 2$), and obtain the spin-orbit interaction. We all know Einstein's famous formula

$$E = mc^2 = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{469}$$

where m_0 is the rest mass of the particle and c the speed of light. We want to transform this into a differential equation similar to the Schrödinger equation. We are therefore not very happy with the presence of the square root, so let us start by squaring the equation

$$E^2 = \frac{m_0^2 c^4}{1 - \frac{v^2}{c^2}} = \frac{m_0^2 c^4 (1 - \frac{v^2}{c^2} + \frac{v^2}{c^2})}{1 - \frac{v^2}{c^2}} = m_0^2 c^4 + \frac{m_0^2 v^2 c^2}{1 - \frac{v^2}{c^2}}. = m_0^2 c^4 + p^2 c^2 \tag{470}$$

A. Klein-Gordon equation

We could start from the previous equation and simply make the substitutions $\mathbf{p} \to -i\hbar\nabla$ and $E = i\hbar\frac{\partial}{\partial t}$, giving

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -c^2\hbar^2\nabla^2\psi + m_0^2 c^4 \psi. \tag{471}$$

or

$$(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2})\psi = \frac{m_0^2 c^2}{\hbar^2}\psi. \tag{472}$$

This is known as the Klein-Gordon equation. We see a second derivative to t. This looks more like a wave equation than a Schrödinger equation, which has a first derivative in t. Unfortunately this equation is not satisfactory as a relativistic version of the Schrödinger equation. Let us first look at the energies

$$E = \pm\sqrt{\mathbf{p}^2 c^2 + m_0^2 c^4}. \tag{473}$$

We find positive and negative energies. This might look problematic at first but these are actually found in nature. The electron has a negative energy equivalent in the positron. However, let us try to write the Klein-Gordon equation as a probability current, as we did for the Schrödinger equation in Eqn. (59). Again, after writing the conjugate wavefunction and substracting after multiplication with ψ^* and ψ, respectively, we find

$$\psi^*\nabla^2\psi - \psi\nabla^2\psi^* = \frac{1}{c^2}\left(\psi^*\frac{\partial^2\psi}{\partial t^2} - \frac{\partial^2\psi^*}{\partial t^2}\psi\right). \tag{474}$$

Trying to write this as a diffusion equation for probability, cf. Eqn. (61), gives

$$-\frac{\hbar}{2m_0 i}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*) = -\frac{\hbar}{2m_0c^2 i}\frac{\partial}{\partial t}\left(\psi^*\frac{\partial\psi}{\partial t} - \frac{\partial\psi^*}{\partial t}\psi\right). \tag{475}$$

Comparison with Eqn. (61) leads to the identification of the probability density as

$$\rho = \frac{i\hbar}{2m_0c^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \frac{\partial\psi^*}{\partial t}\psi\right). \tag{476}$$

For a stationary state, we can take the time dependence $\psi(\mathbf{r},t) = \psi(\mathbf{r})\exp(iEt/\hbar)$, which gives

$$\rho = \frac{E}{m_0c^2}(\psi^*\psi). \tag{477}$$

Note that for particles in the nonrelativistic limit, where $E \cong m_0c^2$ this gives the proper probability density $\rho = \psi^*\psi$. However, for relativistic particles, we saw that E could be less than zero, which would lead to negative probability densities, which are undesirable.

B. Dirac equation

Since the Klein-Gordon equation is not entirely satisfactory (although it has been used successfully in describing spinless bosons,π-mesons), Dirac set out to find a wave equation that is linear in the derivative to t and not quadratic. Bringing the p^2 term in Eqn. (470) to the other side gives

$$\frac{E^2}{c^2} - \mathbf{p}^2 = m_0^2c^2. \tag{478}$$

We are still stuck with an E^2 and we would prefer to end up with something linear in E, just like the Schrödinger equation. We could try to factorize the left-hand side as $a^2 - b^2 = (a-b)(a+b)$. However, we cannot take simply the square root of $\mathbf{p}^2 = p_x^2 + p_y^2 + p_z^2$. We can, however, try to make use of the equation

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b}\mathrm{I} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}), \tag{479}$$

where the matrix I is given by

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{480}$$

and $\boldsymbol{\sigma}$,

$$\boldsymbol{\sigma} = \hat{\mathbf{x}}\sigma_x + \hat{\mathbf{y}}\sigma_y + \hat{\mathbf{z}}\sigma_z = \hat{\mathbf{x}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \hat{\mathbf{y}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \hat{\mathbf{z}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{481}$$

Using this for the momentum $(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p}) = \mathbf{p}\cdot\mathbf{p}\mathrm{I} = \mathbf{p}^2\mathbf{I}$ (note that $\mathbf{p}\times\mathbf{p} = 0$), with

$$\mathbf{p}\cdot\boldsymbol{\sigma} = p_x\sigma_x + p_y\sigma_y + p_z\sigma_z = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}. \tag{482}$$

would enable us to factorize the equation but would in fact give rise to two equations instead of one

$$\frac{E^2}{c^2}\mathbf{I} - \mathbf{p}^2\mathbf{I} = \frac{E^2}{c^2}\mathbf{I} - (\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p}) = \left(\frac{E}{c}\mathbf{I} - (\boldsymbol{\sigma}\cdot\mathbf{p})\right)\left(\frac{E}{c}\mathbf{I} + (\boldsymbol{\sigma}\cdot\mathbf{p})\right) = m_0^2c^2\mathbf{I}. \tag{483}$$

Some of you might recognize the σ_i as Pauli spin matrices. What we have effectively done is introducing an extra degree of freedom, which is the spin of the electron. Although this was first done in the derivation of the relativistic wave equation, we could have done it, in fact, from the non-relativistic Schrödinger equation, by replacing $\mathbf{p}^2$ by $(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p})$.

So far, we have not done any quantum mechanics, so it is time to introduce the operators

$$E \rightarrow i\hbar\frac{\partial}{\partial t} \qquad \text{and} \qquad \mathbf{p} \rightarrow -i\hbar\nabla. \tag{484}$$

It is common in relativity to introduce a four-vector

$$x_\mu = x_1, x_2, x_3, x_4 = x, y, z, ict, \tag{485}$$

treating space and time on equal footing. The time-dependence then becomes

$$i\hbar\frac{\partial}{\partial t} = -\hbar c\frac{\partial}{\partial x_4}. \tag{486}$$

Equation (483) now becomes

$$\left(-\hbar\frac{\partial}{\partial x_4} + i\hbar\boldsymbol{\sigma}\cdot\nabla\right)\left(-\hbar\frac{\partial}{\partial x_4} - i\hbar\boldsymbol{\sigma}\cdot\nabla\right)\psi = m_0^2c^2\psi, \tag{487}$$

where the unit matrices $\mathbf{I}$ have been omitted. Unfortunately, we still do not have a linear differential equation. For this we have to make another sacrifice. We now have a two-component wavefunction and we have to go to a four-component wavefunction. We can achieve this by introducing the wavefunctions

$$\left(\hbar\frac{\partial}{\partial x_4} + i\hbar\boldsymbol{\sigma}\cdot\nabla\right)\psi^L = -m_0c\psi^R, \tag{488}$$

$$\left(\hbar\frac{\partial}{\partial x_4} - i\hbar\boldsymbol{\sigma}\cdot\nabla\right)\psi^R = -m_0c\psi^L, \tag{489}$$

$$\tag{490}$$

where $\psi^L = \psi$. By taking the sum and difference of the above equations, we obtain

$$\hbar\frac{\partial}{\partial x_4}(\psi^R + \psi^L) - i\hbar\boldsymbol{\sigma}\cdot\nabla(\psi^R - \psi^L) = -m_0c(\psi^R + \psi^L), \tag{491}$$

$$i\hbar\boldsymbol{\sigma}\cdot\nabla(\psi^R + \psi^L) - \hbar\frac{\partial}{\partial x_4}(\psi^R - \psi^L) = -mc(\psi^R - \psi^L). \tag{492}$$

Note that both equations have two spin components. By defining $\psi_\pm = \psi^R \pm \psi^L$, we can write this equation in matrix form

$$\begin{pmatrix} \hbar\frac{\partial}{\partial x_4} & -i\hbar\boldsymbol{\sigma}\cdot\nabla \\ i\hbar\boldsymbol{\sigma}\cdot\nabla & -\hbar\frac{\partial}{\partial x_4} \end{pmatrix}\begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} = -m_0c\begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix}. \tag{493}$$

If we now make the shorthand

$$\psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix}, \tag{494}$$

we can finally write the equation as a linear equation working on the fourvector ψ,

$$\left(\boldsymbol{\gamma}\cdot\nabla + \gamma_4\frac{\partial}{\partial x_4}\right)\psi + \frac{m_0c}{\hbar}\psi = 0 \tag{495}$$

or

$$\left(\gamma_\mu\frac{\partial}{\partial x_\mu} + \frac{m_0c}{\hbar}\right)\psi = 0, \tag{496}$$

where as, using the Einstein notation, the summation over μ is implicitly assumed. The 4×4 matrices are given by

$$\gamma_k = \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix} \qquad \text{and} \qquad \gamma_4 = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & -\mathbf{I} \end{pmatrix}. \tag{497}$$

Often this equation is given in a somewhat different form, which can be obtained by multiplying by $-\hbar c\gamma_4$ from the right side. Using

$$\gamma_4 \frac{\partial}{\partial x_4}(-\hbar c\gamma_4) = i\hbar\frac{\partial}{\partial t} \qquad \text{and} \qquad -\hbar c\gamma_k\gamma_4 = i\hbar c(i\gamma_k\gamma_4) = i\hbar c\alpha_k \tag{498}$$

and

$$\alpha_k = i\gamma_k\gamma_4 = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \tag{499}$$

and $\beta = \gamma_4$, we can explicitly bring out the time dependence

$$H\psi = \left\{c\boldsymbol{\alpha}\cdot(-i\hbar\nabla) + \beta m_0c^2\right\} = i\hbar\frac{\partial\psi}{\partial t} \tag{500}$$

or

$$i\hbar\frac{\partial\psi}{\partial t} = \left\{(\boldsymbol{\alpha}\cdot\mathbf{p})c + \beta m_0c^2\right\}\psi. \tag{501}$$

This is known as the Dirac equation. One might wonder whether this has the right probability density.

$$i\hbar\frac{\partial\psi}{\partial t} = -ic\hbar(\boldsymbol{\alpha}\cdot\nabla\psi)c + \beta m_0c^2\psi. \tag{502}$$

Multiplying with ψ^* and substracting the conjugated equation multiplied by ψ gives

$$\frac{\partial}{\partial t}(\psi^*\psi) + \nabla\cdot(c\psi^\dagger\boldsymbol{\alpha}\psi) = 0, \tag{503}$$

giving a current density of $\mathbf{J} = c\psi^\dagger\boldsymbol{\alpha}\psi$ and the proper probability density of $\rho = \psi^*\psi$.

We can also write this out in four coupled equations. Rewriting Dirac's equation

$$\{E - (\boldsymbol{\alpha}\cdot\mathbf{p})c - \beta m_0c^2\}\psi = 0. \tag{504}$$

Expanding the matrices gives

$$\begin{pmatrix} E-m_0c^2 & 0 & -p_zc & -(p_x-ip_y)c \\ 0 & E-m_0c^2 & -(p_x+ip_y)c & p_zc \\ -p_zc & -(p_x-ip_y)c & E+m_0c^2 & 0 \\ -(p_x+ip_y)c & p_zc & 0 & E+m_0c^2 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = 0 \tag{505}$$

For a plane wave

$$\psi = \mathbf{a}(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{r}/\hbar - i\omega t), \tag{506}$$

the operator p_i can be replaced by c-numbers. We now only need to solve the coeffients of $\mathbf{a}$. This can still depend on $\mathbf{p}$, but not on t and $\mathbf{r}$. These equations have a nonzero solution only if the determinant of the matrix vanishes. This gives

$$(E^2 - m_0^2c^4 - c^2p^2)^2 = 0, \tag{507}$$

which is nothing but the fundamental relativistic energy equation. Solutions for the eigenvectors for $E_\pm = \pm\sqrt{c^2p^2 + m_0^2c^4}$ are

$$\psi^+_\uparrow = \begin{pmatrix} 1 \\ 0 \\ cp_z/\Delta_+ \\ c(p_x+ip_y)/\Delta_+ \end{pmatrix}, \qquad \psi^+_\downarrow = \begin{pmatrix} 0 \\ 1 \\ c(p_x-ip_y)/\Delta_+ \\ -cp_z/\Delta_+ \end{pmatrix} \tag{508}$$

$$\psi^-_\uparrow = \begin{pmatrix} cp_z/\Delta_- \\ c(p_x+ip_y)/\Delta_- \\ 1 \\ 0 \end{pmatrix}, \qquad \psi^-_\downarrow = \begin{pmatrix} c(p_x-ip_y)/\Delta_- \\ -cp_z/\Delta_- \\ 0 \\ 1 \end{pmatrix}, \tag{509}$$

where

$$\Delta_\pm = (E_\pm \pm m_0c^2). \tag{510}$$

Note that the wavefunction are not normalized. The normalization for, e.g., $\psi^+_\uparrow$ is given by

$$N_p = \frac{1}{\sqrt{1+\frac{c^2p^2}{\Delta_+^2}}} \tag{511}$$

Note that in the nonrelativistic limit all the terms with $\Delta_\pm$ in there vanish and we obtain a two-vector plane wave for ψ^+ (with spin up and spin down), and a two-vector plane wave for ψ^-.

Let us now try to make a connection with the non-relativistic theory. We also want to introduce an electromagnetic field and a potential. We can rewrite Eqn. (493) as

$$\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\psi^- = \frac{1}{c}(E-V-m_0c^2)\psi^+ \tag{512}$$

$$\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\psi^+ = \frac{1}{c}(E-V+m_0c^2)\psi^- \tag{513}$$

$$\tag{514}$$

We can solve these equations for ψ^+, giving

$$\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\frac{c^2}{E-V+m_0c^2}\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A})\psi^+ = (E-V-m_0c^2)\psi^+. \tag{515}$$

In the non-relativistic limit $E = E_{\rm nr}+m_0c^2 \cong m_0c^2$ and $v/c \ll 1$. We can therefore expand

$$\frac{c^2}{E-V+m_0c^2} = \frac{1}{2m_0}\frac{2m_0c^2}{2m_0c^2+E_{nr}-V} = \frac{1}{2m_0}\left\{1-\frac{E_{nr}-V}{2m_0c^2}+\cdots\right\} \tag{516}$$

When keeping only the lowest order this becomes $1/2m_0$. This gives for the Hamiltonian

$$\frac{1}{2m_0}(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}))^2\psi^+ = (E_{nr}-V)\psi^+. \tag{517}$$

or, using Eqn. (479),

$$\left\{\frac{(\mathbf{p}-e\mathbf{A})^2}{2m_0}+g\frac{\mu_B}{\hbar}\mathbf{S}\cdot\mathbf{B}+V\right\}\psi^+ = E_{nr}\psi^+, \tag{518}$$

using the fact that

$$\mathbf{p}\times(\mathbf{A}\psi)+\mathbf{A}\times\mathbf{p}\psi = \psi(\mathbf{p}\times\mathbf{A})-\mathbf{A}\times\mathbf{p}\psi+\mathbf{A}\times\mathbf{p}\psi = \psi(\mathbf{p}\times\mathbf{A}) \tag{519}$$

We can now write

$$\frac{ie}{2m_0}\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{A}) = \frac{e\hbar}{2m_0}\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{A}) = \frac{e\hbar}{2m_0}\boldsymbol{\sigma}\cdot\mathbf{B} = \frac{e}{m_0}\mathbf{S}\cdot\mathbf{B} = g\frac{\mu_B}{\hbar}\mathbf{S}\cdot\mathbf{B} \tag{520}$$

with the Bohr magneton $\mu_B = \frac{e\hbar}{2m_0}$ and the gyromagnetic factor $g = 2$. Note that the total interaction with the magnetic field is proportional to $(\mathbf{L}+g\mathbf{S})\cdot\mathbf{B}$. Obtaining the right value of the gyromagnetic factor was one the achievements of the relativistic theory.

Note that the other two-component wavefunction

$$\psi^- = c\frac{(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}))}{E-V+m_0c^2}\psi^+ \cong \frac{(\boldsymbol{\sigma}\cdot(\mathbf{p}-e\mathbf{A}))}{2m_0c}\psi^+ \cong \frac{v}{2c}\psi^+, \tag{521}$$

is an order $p/mc = v/c$ smaller than the ψ^+ component.

We can also keep one term higher in the expansion in Eqn. (516). Transforming it into a decent eigenvalue problem is somewhat elaborate (see Sakurai) and we will therefore only give the final result (for simplicity we set $\mathbf{A}=0$)

$$\left[\frac{\mathbf{p}^2}{2m_0}+V-\frac{\mathbf{p}^4}{8m_0^3c^2}-\frac{e\hbar}{4m_0^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{E}\times\mathbf{p})-\frac{e\hbar^2}{8m_0^2c^2}\nabla\cdot\mathbf{E}\right]\psi = E_{nr}\psi, \tag{522}$$

where the electric field comes from $\mathbf{p}V = -i\hbar e\nabla U = e\hbar\mathbf{E}$, where U is the potential. The first two terms are well known. The first correction terms is the next higher-order correction to the kinetic energy

$$\sqrt{\mathbf{p}^2c^2 + m_0^2c^4} - m_0c^2 = \frac{\mathbf{p}^2}{2m_0} - \frac{\mathbf{p}^4}{8m_0^3c^2} + \cdots . \tag{523}$$

For a Hydrogen like atom, the next term is

$$-\frac{e\hbar}{4m_0^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{E}\times\mathbf{p}) = -\frac{\hbar}{4m_0^2c^2}\left(-\frac{1}{r}\frac{dV}{dr}\right)\boldsymbol{\sigma}\cdot(\mathbf{r}\times\mathbf{p}) = \frac{e^2}{8\pi\varepsilon_0 m_0^2c^2r^3}\mathbf{L}\cdot\mathbf{S} \tag{524}$$

where the electric field from the nucleus is along the $\mathbf{r}$ direction and $S = \frac{1}{2}\hbar\boldsymbol{\sigma}$. This term is known as the spin-orbit coupling and will be discussed in more detail below.

The third term is the Darwin term and is related to the negative solution of the Dirac equation. For the Hydrogen atom, one has

$$-\frac{e\hbar^2}{8m_0^2c^2}\nabla\cdot\mathbf{E} = \frac{\hbar^2}{8m_0^2c^2}\nabla^2V = \frac{e^2\hbar^2}{8\varepsilon m_0^2c^2}\delta(\mathbf{r}). \tag{525}$$

We like to draw the attention again to the existence of antiparticles in the Dirac theory. One can see this as a kind of semiconductor picture, where all the states with negative energy are filled. For electrons the negative energy equivalent is the positron. The energies of the antiparticles start at $-m_0c^2$, and those of the particles at m_0c^2, giving a gap of 1.02 MeV, see Fig. 15.4 in Liboff. A positron can only be created by making an electron-positron pair across the gap.

C. Spin-orbit coupling and relativistic corrections (Liboff 12.2)

The spin-orbit coupling follows directly from the relativistic Dirac equation, but how do we understand this from a semiclassical point of view? Since the electron is moving through the electric field of the nucleus, it will also experience an internal magnetic field given by the Biot-Savart law

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{Ze(\mathbf{r}\times\mathbf{v})}{r^3} = \frac{Ze}{4\pi\varepsilon_0 mc^2}\frac{(\mathbf{r}\times m\mathbf{v})}{r^3} = \frac{\mathbf{L}}{mec^2}\frac{1}{r}\frac{d}{dr}\left(\frac{-Ze^2}{4\pi\varepsilon_0 r}\right) = \frac{1}{mec^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{L}. \tag{526}$$

In a different frame of reference one could see this as the nucleus turning around the electron. The orbiting charged nucleus gives rise to a current, which leads to a magnetic field interacting with the spin-moment of the electron. The change in energy for the electron is then given by the scalar product $\mu_S\cdot\mathbf{B}$. Relativistic kinematic considerations lead to the introduction of a factor $\frac{1}{2}$. This gives an energy of

$$\frac{1}{2}\mu_S\cdot\mathbf{B} = \frac{1}{2}\left(\frac{e}{m}\mathbf{S}\right)\left(\frac{1}{mec^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{L}\right) = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV(r)}{dr}\mathbf{L}\cdot\mathbf{S} = \zeta(r)\mathbf{L}\cdot\mathbf{S}. \tag{527}$$

What are the eigenvalues of the spin-orbit coupling for a single electron? Let us first put the parameter $\zeta = 1$. We would like to couple the orbit $\mathbf{l}$ and spin $\mathbf{s}$ moments together to a total moment $\mathbf{j}$,

$$\mathbf{j} = \mathbf{l} + \mathbf{s}. \tag{528}$$

The norm of the new moment is

$$\mathbf{j}^2 = (\mathbf{l}+\mathbf{s})^2 = \mathbf{l}^2 + \mathbf{s}^2 + 2\mathbf{l}\cdot\mathbf{s}, \tag{529}$$

so the eigenvalues are given by

$$\mathbf{l}\cdot\mathbf{s}\psi_{jm_j} = \frac{1}{2}(\mathbf{j}^2 - \mathbf{l}^2 - \mathbf{s}^2)\psi_{jm_j} = \frac{1}{2}(j(j+1) - l(l+1) - \frac{3}{4})\hbar^2\psi_{jm_j}. \tag{530}$$

Since j can only take the values $l \pm \frac{1}{2}$, we can also write the eigenvalues as

$$\langle jm_j|\mathbf{l}\cdot\mathbf{s}|jm_j\rangle = \begin{cases} \frac{1}{2}l\hbar^2 & j = l + \frac{1}{2} \\ -\frac{1}{2}(l+1)\hbar^2 & j = l - \frac{1}{2} \end{cases} \tag{531}$$

Although this produces the right result, it is a bit based on a trick to obtain the value for $\mathbf{l}\cdot\mathbf{s}$. When evaluating matrix elements with a computer, one would often compute matrix elements such as $\langle m's'_z|\mathbf{L}\cdot\mathbf{S}|ms_z\rangle$ (of course, this depends on the basis set one chooses). To find the matrix elements, it is convenient to rewrite the spin-orbit coupling as

$$\mathbf{L}\cdot\mathbf{S} = L_xS_x + L_yS_y + L_zS_z = \frac{1}{2}(L_+S_- + L_-S_+) + L_zS_z, \tag{532}$$

since we know the matrix elements for the z-component and the step operators, see Eqns. (406) and (412). Let us consider a p electron ($l=1$). For example, the state $|1\uparrow\rangle$ has as diagonal element $ms_z = 1\times\frac{1}{2}\hbar^2 = \frac{\hbar^2}{2}$. It has no off-diagonal element, since lowering the orbital momentum by 1 would require raising the spin by 1. This is however impossible since the state already has an up spin. The state $|1\downarrow\rangle$ has as diagonal element $ms_z = 1\times(-\frac{1}{2})\hbar^2 = -\frac{\hbar^2}{2}$. This state has an off-diagonal element connecting it to $|0\uparrow\rangle$. The matrix element is given by

$$\begin{aligned}\langle 0\uparrow|\frac{1}{2}L_-S_+|1\downarrow\rangle &= \frac{\hbar^2}{2}\sqrt{(l+m)(l-m+1)}\sqrt{(s-s_z)(s+s_z+1)} &(533)\\ &= \frac{\hbar^2}{2}\sqrt{(l+1)(1-1+1)}\sqrt{(\frac{1}{2}-(-\frac{1}{2}))(\frac{1}{2}+(-\frac{1}{2})+1)} = \frac{\hbar^2}{\sqrt{2}}. &(534)\end{aligned}$$

The total eigenvalue problem then becomes

$$\hbar^2\begin{pmatrix} \frac{1}{2} & & & & & \\ & -\frac{1}{2} & \frac{1}{\sqrt{2}} & & & \\ & \frac{1}{\sqrt{2}} & 0 & & & \\ & & & 0 & \frac{1}{\sqrt{2}} & \\ & & & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \\ & & & & & \frac{1}{2}\end{pmatrix}\begin{matrix}|1\uparrow\rangle\\|1\downarrow\rangle\\|0\uparrow\rangle\\|0\downarrow\rangle\\|-1\uparrow\rangle\\|-1\downarrow\rangle\end{matrix}. \tag{535}$$

Although a 6×6 matrix, this matrix is easily reduced into 4 smaller matrices. Two are 1×1 giving eigenvalues of $\hbar^2/2$. The other two are 2×2 matrices,whose eigenvalues are determined by the determinant

$$\begin{vmatrix} -\frac{\hbar^2}{2}-E & \frac{\hbar^2}{\sqrt{2}} \\ \frac{\hbar^2}{\sqrt{2}} & -E\end{vmatrix} = (-\frac{\hbar^2}{2}-E)(-E) - \frac{\hbar^4}{2} = E^2 + \frac{\hbar^2}{2}E - \frac{\hbar^4}{2}. \tag{536}$$

The eigenvalues are easily evaluated to give

$$E_{1,2} = -\frac{\hbar^2}{4} \pm \frac{\hbar^2}{2}\sqrt{\frac{1}{4} - 4\left(-\frac{1}{2}\right)} = \hbar^2(-\frac{1}{4}\pm\frac{3}{4}) = \begin{cases} -\hbar^2 \\ \frac{1}{2}\hbar^2\end{cases}. \tag{537}$$

In total, we find a fourfold degenerate eigenstate with energy $\hbar^2/2$ and a twofold degenerate eigenstate with energy $-\hbar^2$. Since the degeneracy is given by $2j+1$ (since $m_j = -j,\ldots,j$), this gives $j=\frac{3}{2},\frac{1}{2}$, as we would expect. So even if a computer would not know anything about coupling to jm_j and only knows it is working in a basis set $|ms_z\rangle$ and the spin-orbit coupling interaction, it still ends up with eigenstates $|jm_j\rangle$. Note also that the center of gravity is zero since $4\times\hbar^2/2 + 2\times(-\hbar^2) = 0$.

The eigenstates can also be found. Note that the solutions of the 1×1 matrices are $|\frac{3}{2}\frac{3}{2}\rangle = |1\frac{1}{2}\rangle$ and $|\frac{3}{2},-\frac{3}{2}\rangle = |-1,-\frac{1}{2}\rangle$ since m and m_z are coupled parallel. For example, writing

$$|j\frac{1}{2}\rangle = a|1\downarrow\rangle + b|0\uparrow\rangle \tag{538}$$

Substituting E gives

$$\left(-\frac{1}{2}-(-1)\right)a + \frac{1}{\sqrt{2}}b = 0 \qquad \Rightarrow \qquad a = -\sqrt{2}b \qquad \Rightarrow \qquad |\frac{1}{2}\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|1\downarrow\rangle - \sqrt{\frac{1}{3}}|0\uparrow\rangle. \tag{539}$$

The other state has to be orthoganal to this one giving

$$|\frac{3}{2}\frac{1}{2}\rangle = \sqrt{\frac{1}{3}}|1\downarrow\rangle + \sqrt{\frac{2}{3}}|0\uparrow\rangle. \tag{540}$$

Note that the state which is lowest in energy has more $|1\downarrow\rangle$ character since the diagonal element of the state is lower than that for $|0\uparrow\rangle$.

What is now the value of the parameter ζ? We will calculate this using perturbation theory. A perturbation means that there is a small term in the Hamiltonian $H = H_0 + H'$. We know how to solve the problem for H_0. However, we do not know how to solve it for H. Although perturbation theory will be treated later on in more detail, the lowest order correction to the energy is rather intuitive,

$$E_1 = \langle\psi_0|H'|\psi_0\rangle. \tag{541}$$

This means that we calculate the expectation value of the perturbance H' for the wavefunction of the unperturbed state, that follows from $H_0|\psi_0\rangle = E_0|\psi_0\rangle$. The total energy is then given by $E_0 + E_1$. The assumption is that the small perturbance does not really change the wavefunction.

For the spin-orbit coupling, the unperturbed Hamiltonian is that of the Hydrogen atom in Eqn. (315), which we have solved. We want to obtain the expectation value for

$$\zeta(r) = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr} = \frac{Ze^2}{2m^2c^2}\frac{1}{r^3} \tag{542}$$

where we have taken $e/(4\pi\varepsilon_0)^{1/2} \to e$. Since ζ only depends on r, we can calculated ζ from

$$\langle\zeta\rangle_{nl} = \frac{Ze^2}{2m^2c^2}\int_0^\infty \frac{R_{nl}^2(r)}{r^3}r^2dr. \tag{543}$$

The expectation value can be straightforwardly evaluated (see the results of problem 10.49 in Liboff). We find

$$\frac{\hbar^2}{2}\langle\zeta\rangle_{nl} = \left(\frac{me^4Z^2}{2\hbar^2n^2}\right)^2 \frac{n}{mc^2(l+\frac{1}{2})(l+1)l}. \tag{544}$$

Recognizing that $|E_n| = mZ^2e^4/(2\hbar^2n^2)$, we can rewrite the result as

$$\frac{\hbar^2}{2}\frac{\langle\zeta\rangle_{nl}}{|E_n|} = \frac{|E_n|}{mc^2}\frac{n}{(l+\frac{1}{2})(l+1)l} = \frac{(Z\alpha)^2}{n}\frac{1}{(2l+1)(l+1)l}, \tag{545}$$

where

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.037} \tag{546}$$

is known as the fine-structure constant. The total energy can now be written as

$$E_{nlj^\pm} = -|E_n|\left[1 \mp \frac{(Z\alpha)^2}{n}\frac{1}{(2l+1)(j^\pm+\frac{1}{2})}\right] = \begin{cases} -|E_n|\left[1-\frac{(Z\alpha)^2}{n}\frac{1}{(2l+1)(l+1)}\right] & j = l+\frac{1}{2} \\ -|E_n|\left[1+\frac{(Z\alpha)^2}{n}\frac{1}{(2l+1)l}\right] & j = l-\frac{1}{2} \end{cases}. \tag{547}$$

Note that the spin-orbit coupling is not necessarily a small quantity. In atoms with many electrons, the spin-orbit coupling of a deep-lying level can be several to several tens of electron volts. Compared to the 0.1 eV for shallow levels, this cannot be treated as fine structure, i.e., a small perturbation.

Relativistic correction In a similar fashion, we can treat the relativistic correction to the kinetic energy from Eqn. (523),

$$H' = -\frac{\mathbf{p}^4}{8m^3c^2}. \tag{548}$$

The lowest order correction is then given by

$$E_1 = \langle nl|H'|nl\rangle = -\frac{1}{8m^3c^2}\langle\mathbf{p}^4\rangle_{nl}. \tag{549}$$

We do not know how to calculate the expectation value for $\mathbf{p}^4$, but we are more familiar with the $\mathbf{p}^2$ term, since

$$\left(\frac{\mathbf{p}^2}{2m} + V\right)|nl\rangle = E_n|nl\rangle. \tag{550}$$

Rearranging gives

$$\mathbf{p}^2|nl\rangle = 2m(E_n - V)|nl\rangle. \tag{551}$$

Using this, we can calculate the perturbation

$$E_1 = \langle nl|H'|nl\rangle = -\frac{1}{8m^3c^2}\langle nl|\mathbf{p}^2\mathbf{p}^2|nl\rangle \tag{552}$$

$$= -\frac{1}{8m^3c^2}\int_0^\infty R_{nl}[2m(E_n - V)][2m(E_n - V)]R_{nl}r^2dr = -\frac{1}{2mc^2}(E_n^2 - 2E_n\langle V\rangle + \langle V\rangle^2). \tag{553}$$

Using the expectation values for r^{-1} and r^{-2} from problem 10.49 in Liboff gives

$$E_1 = -|E_n|\frac{\alpha^2Z^2}{4n^2}\left(\frac{8n}{2l+1} - 3\right). \tag{554}$$

Adding this to Eqn. (547) gives

$$E_{nlj} = -|E_n|\left[1 + \left(\frac{\alpha Z}{2n}\right)^2\left(\frac{4n}{j+\frac{1}{2}} - 3\right)\right]. \tag{555}$$

This expression gives quite good agreement with the experimentally observed spectra for one-electron atoms. It is now time to see what happens if we start adding more electrons to the atom.

IX. PERTURBATION THEORY AND ABSORPTION AND EMISSION OF PHOTONS

After a brief excursion into relativistic quantum mechanics, we take another small detour. A large amount of data on atoms has been obtained by spectroscopy. We would like to describe the process of absorption and emission of photon. However, we are missing the necessary mathematical equipment. Usually this is done with perturbation theory, which was already introduced (in its simplest form) earlier. We will now discuss it in more detail, since we will need it later on.

A. Time-independent perturbation theory

We want to know the effect of a (small) perturbation H_1 on a Hamiltonian H_0. Often, in quantum mechanics, we are able to solve the problem for a certain Hamiltonian H_0, but are unable to solve it for the complete Hamiltonian $H_0 + H_1$. However, when H_1 is a relatively small effect, there is no need to solve the entire problem, but we can calculate the effect of H_1 to a certain order. To keep track of the order, we introduce a constant λ, which goes to unity when we have sorted out all the different orders. Therefore, we are interested in solving the problem

$$(H_0 + \lambda H_1)\psi_n = E_n\psi_n. \tag{556}$$

We were able to solve the problem $H_0\psi_n^0 = E_n^0\psi_n^0$. We would like to know the corrections to the energy and wavefunction due to H_1:

$$\psi_n = \psi_n^0 + \Delta\psi_n = \psi_n^0 + \lambda\psi_n^1 + \lambda^2\psi_n^2 + \cdots \tag{557}$$

$$E_n = E_n^0 + \Delta E_n = E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \cdots. \tag{558}$$

Equation (556) now becomes

$$(H_0 + \lambda H_1)(\psi_n^0 + \lambda\psi_n^1 + \lambda^2\psi_n^2 + \cdots) = (E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \cdots)(\psi_n^0 + \lambda\psi_n^1 + \lambda^2\psi_n^2 + \cdots) \tag{559}$$

Keeping the terms up to order λ^2 gives

$$H_0\psi_0^n + \lambda(H_1\psi_n^0 + H_0\psi_n^1) + \lambda^2(H^0\psi_n^2 + H_1\psi_n^1) = E_n^0\psi_n^0 + \lambda(E_n^1\psi_n^0 + E_n^0\psi_n^1) + \lambda^2(E_n^2\psi_n^0 + E_n^1\psi_n^1 + E_n^0\psi_n^2). \tag{560}$$

Of course we can go higher in order but that kind of defeats the purpose of doing perturbation theory. (There is perturbation theory which goes to infinite order. This is solving the problem exactly but with perturbation methods.

This is, however, a different story.) We see that we obtain the unperturbed Schrödinger equation $H_0\psi_n^0 = E_n^0$. To first order in λ, we have

$$H_0\psi_n^1 + H_1\psi_n^0 = E_n^0\psi_n^1 + E_n^1\psi_n^0. \tag{561}$$

Multiplying this by $(\psi_n^0)^*$ and integrating over the volume gives

$$\langle\psi_n^0|H_0|\psi_n^1\rangle + \langle\psi_n^0|H_1|\psi_n^0\rangle = E_n^0\langle\psi_n^0|\psi_n^1\rangle + E_n^1\langle\psi_n^0|\psi_n^0\rangle. \tag{562}$$

However $\langle\psi_n^0|H_0|\psi_n^1\rangle = E_n^0\langle\psi_n^0|\psi_n^1\rangle$, and we are therefore left with

$$E_n^1 = \langle\psi_n^0|H_1|\psi_n^0\rangle. \tag{563}$$

This is a result that we used already. It says that the first order correction to the energy is the expectation value of the perturbation in the unperturbed state. This means that one assumes that the perturbation does not really affect the wavefunction.

The first-order correction to the wavefunction is given by

$$(H_0 - E_n^0)\psi_n^1 = -(H_1 - E_n^1)\psi_n^0. \tag{564}$$

Now since the eigenfunctions of H_0 form a complete set, it is always possible to express the perturbed wavefunction as a linear combination of unperturbed wavefunctions

$$\psi_n^1 = \sum_{m\neq n} c_{nm}\psi_m^0. \tag{565}$$

Note that we do not need ψ_n^0 since the total wavefunction (up to first order) is given by $\psi_n^0 + \alpha\psi_n^1$, so ψ_n^0 is already included in the wavefunction. Substituting the linear combination into the Schrödinger equation gives

$$\sum_{m\neq n}(E_m^0 - E_n^0)c_{nm}\psi_m^0 = -(H_1 - E_n^1)\psi_n^0. \tag{566}$$

Taking the inner product with ψ_l^0 gives

$$\sum_{m\neq n}(E_m^0 - E_n^0)c_{nm}\langle\psi_l^0|\psi_m^0\rangle = -\langle\psi_l^0|H_1|\psi_n^0\rangle + E_n^1\langle\psi_l^0|\psi_n^0\rangle. \tag{567}$$

If $l = n$, the left side is zero and we obtain the equation for E_1, when $l \neq n$

$$(E_l^0 - E_n^0)c_{nl} = -\langle\psi_l^0|H_1|\psi_n^0\rangle, \tag{568}$$

we obtain the coefficient

$$c_{nl} = \frac{\langle\psi_l^0|H_1|\psi_n^0\rangle}{E_n^0 - E_l^0}. \tag{569}$$

The perturbed wavefunction is therefore given by

$$\psi_n^1 = \sum_{m\neq n}\frac{\langle\psi_m^0|H_1|\psi_n^0\rangle}{E_n^0 - E_m^0}\psi_m^0. \tag{570}$$

Notice that this works as long as the eigenstates are nondegenerate, i.e. $E_n \neq E_m$. Although the energies are quite good the wavefunctions are usually described less well by perturbation theory.

Example

Calculate the first-order change in energy as a result of the anharmonic potential

$$H' = Kx^4 = \frac{K}{4\beta^4}(a + a^\dagger)^4, \tag{571}$$

with

$$x = \frac{1}{\sqrt{2}\beta}(a + a^\dagger), \qquad \beta^2 = \frac{m\omega}{\hbar} \tag{572}$$

in the state $|n\rangle$.

Answer Since we are calculating an expectation in state $|n\rangle$ the number of step-up operators has to be equal to the number of step-down operators, i.e. term such as $(a^\dagger)^4$ vanish. There are six of those terms, for which we obtain the matrix elements

$$\langle n|(aaa^\dagger a^\dagger + a^\dagger a^\dagger aa + aa^\dagger aa^\dagger + a^\dagger aa^\dagger a + aa^\dagger a^\dagger a + a^\dagger aaa^\dagger)|n\rangle \tag{573}$$
$$= \sqrt{n+1}\sqrt{n+2}\sqrt{n+2}\sqrt{n+1} + \sqrt{n}\sqrt{n-1}\sqrt{n-1}\sqrt{n} + \sqrt{n+1}\sqrt{n+1}\sqrt{n+1}\sqrt{n+1} \tag{574}$$
$$+\sqrt{n}\sqrt{n}\sqrt{n}\sqrt{n} + \sqrt{n+1}\sqrt{n+1}\sqrt{n}\sqrt{n} + \sqrt{n}\sqrt{n}\sqrt{n+1}\sqrt{n+1} \tag{575}$$
$$= (n+1)(n+2) + n(n-1) + (n+1)^2 + n^2 + (n+1)n + n(n+1) = 3[2n(n+1)+1] \tag{576}$$

The total energy is then

$$E_n^1 = \frac{3K}{4\beta}[2n(n+1)+1] \tag{577}$$

Degenerate perturbation theory

Lowest order perturbation theory works well in the situation where the energies are not degenerate. However, perturbation theory also works in the case of degenerate energies. The assumption is that only states of the same energy couple with each other. The splitting between states with different energies is assumed to be significantly larger than the perturbation. We are then left with a number of states at the same energy. Note that we cannot simply use lowest order perturbation theory since the wavefunction in Eqn. (570) would go to infinity. Therefore, we have to solve the problem for our reduced basis set of states at the same energy. This is best illustrated by considering some examples.

The two-dimensional harmonic oscillator

For a two-dimensional harmonic oscillator, we have the Hamiltonian

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{K}{2}(x^2 + y^2). \tag{578}$$

This is easily diagonalized by noting that the x and y coordinates do not mix. Thus,

$$H = (a^\dagger a + b^\dagger b + 1)\hbar\omega, \tag{579}$$

where x is given in the previous example and y

$$y = \frac{1}{\sqrt{2}\beta}(b + b^\dagger). \tag{580}$$

The eigenstates are now products of eigenstates in one direction

$$\varphi_{mn}(x,y) = \varphi_m(x)\varphi_n(y) \qquad \text{or} \qquad |mn\rangle = \frac{1}{\sqrt{m!n!}}(a^\dagger)^m(b^\dagger)^n|0\rangle \tag{581}$$

with energies

$$E_{mn} = (n+m+1)\hbar\omega \tag{582}$$

Therefore, the energy $2\hbar\omega$ is two-fold degenerate: $E_{10} = E_{01} = 2\hbar\omega$.

We would now like to see how the degereneracy of the state with energy $2\hbar\omega$ is lifted by a perturbation

$$H' = K'xy = \frac{K'}{2\beta^2}(a + a^\dagger)(b + b^\dagger) = \frac{K'}{2\beta^2}(ab + a^\dagger b + ab^\dagger + a^\dagger b^\dagger). \tag{583}$$

Note that the first and last term change the total excited states $n+m$ into $n+m-2$ and $n+m+2$. If we now restrict ourselves to a basis set of $|10\rangle$ and $|01\rangle$ these terms vanish and we are only left with the $a^\dagger b$ and $ab^\dagger$ terms. This leads to off-diagonal matrix elements. The matrix element coupling $|10\rangle$ and $|01\rangle$ is

$$T = K'\langle 10|xy|01\rangle = \frac{K'}{2\beta^2}\langle 10|(a^\dagger b + ab^\dagger)|01\rangle = \frac{K'}{2\beta^2}\langle 10|a^\dagger b|01\rangle = \frac{K'}{2\beta^2}. \tag{584}$$

This leaves us with a simple 2×2 problem

$$\begin{vmatrix} 2\hbar\omega - E & T \\ T & 2\hbar\omega - E \end{vmatrix} = (2\hbar\omega - E) - T^2 = 0 \qquad \Rightarrow \qquad E = 2\hbar\omega \pm T = 2\hbar\omega \pm \frac{K'}{2\beta^2} \tag{585}$$

The Stark effect

The Stark effect is the splitting of the levels of the Hydrogen atom under an electric field. This effect is smaller than the Zeeman splitting resulting from a magnetic field and was only observed in 1913, whereas the Zeeman splitting was discovered in 1897. Since this energy is small it will only have a significant effect of states that are degenerate in energy, i.e. with the same quantum number n. The perturbing Hamiltonian is written as

$$H' = -eE_z z, \tag{586}$$

where E_z is the strength of the electric field in the z-direction. Let us now consider the case for $n = 2$. There a four degenerate state, (in $|nlm\rangle$ notation): $|200\rangle$, $|211\rangle$, $|210\rangle$, and $|21,-1\rangle$. What is now the effect of z? In spherical coordinates $z = r\cos\theta$. This can be written as a spherical harmonic $z = rY_0^1$. Note that this is the term with $m = 0$, therefore z cannot change the quantum number m (unlike x and y. We will discuss that when treating the interaction of atoms with radiation). Therefore the matrix elements have the form $\langle nl'm'|z|nlm\rangle = \langle nl'm|z|nlm\rangle\delta_{m,m'}$. The diagonal terms are also zero. Note that this would be the expectation value of z in that orbital. For an atom $\langle z\rangle = 0$. In our basis set, this means we only have matrix elements

$$\langle 210|H'|200\rangle = \langle 200|H'|210\rangle = -\frac{ea_0E_z}{32\pi}\int_0^\infty \rho^4(2-\rho)e^{-\rho}d\rho\int_{-1}^1\cos^2\theta d\cos\theta\int_0^{2\pi}d\varphi \tag{587}$$

$$= \frac{eE_z\hbar^2}{mZe} = -3|e|E_z a_0 = -T \tag{588}$$

with $\rho = r/2a_0$. This gives again a 2×2 problem as in the previous problem. The eigenenrgies of the perturbing Hamiltonian are therefore $E' = 0, 0, -T, T$. The eigenvectors are given by $|21\rangle = |211\rangle$ and $|2,-1\rangle = |21,-1\rangle$, and $|20^\pm\rangle = \frac{1}{\sqrt{2}}(|200\rangle \mp |210\rangle)$ with $H'|20^\pm\rangle = \pm T|20^\pm\rangle$. Note that as a result of the Stark effect l is no longer a good quantum number, but m still is.

Second order energies. Continuing collecting terms, we find for λ^2

$$H_0\psi_n^2 + H_1\psi_n^1 = E_n^0\psi_n^2 + E_1\psi_n^1 + E_n^2\psi_n^0. \tag{589}$$

Again taking the inner product with ψ_n^0 gives

$$\langle\psi_n^0|H_0|\psi_n^2\rangle + \langle\psi_n^0|H_1|\psi_n^1\rangle = E_n^0\langle\psi_n^0|\psi_n^2\rangle + E_n^1\langle\psi_n^0|\psi_n^1\rangle + E_n^2\langle\psi_n^0|\psi_n^0\rangle. \tag{590}$$

Since $\langle\psi_n^0|H_0|\psi_n^2\rangle = E_n^0\langle\psi_n^0|\psi_n^2\rangle$, we find

$$E_n^2 = \langle\psi_n^0|H_1|\psi_n^1\rangle - E_n^1\langle\psi_n^0|\psi_n^1\rangle. \tag{591}$$

But ψ_n^0 is orthogonal to ψ_n^1 since

$$\langle\psi_n^0|\psi_n^1\rangle = \sum_{m\neq n} c_{nm}\langle\psi_n^0|\psi_m^0\rangle = 0. \tag{592}$$

So,

$$E_n^2 = \langle\psi_n^0|H_1|\psi_n^1\rangle = \sum_{m\neq n} c_{nm}\langle\psi_n^0|H_1|\psi_m^0\rangle = \sum_{m\neq n}\frac{\langle\psi_m^0|H_1|\psi_n^0\rangle\langle\psi_n^0|H_1|\psi_m^0\rangle}{E_n^0 - E_m^0} = \sum_{m\neq n}\frac{|\langle\psi_m^0|H_1|\psi_n^0\rangle|^2}{E_n^0 - E_m^0}. \tag{593}$$

Two-level system. One of the simplest, but nevertheless very important, problems is that of a two level system. Let us take the energies of the two states $E_1 = 0$ and $E_2 = \Delta$. This is our unperturbed Hamiltonian H_0. Now let us introduce a coupling between the two states $\langle 2|H_1|1\rangle = T$. According to second-order perturbation theory the energy should be equal to

$$E_1^2 = \frac{|\langle 2|H_1|1\rangle|^2}{E_1 - E_2} = -\frac{T^2}{\Delta}. \tag{594}$$

Similarly, we find that $E_2^2 = T^2/\Delta$. We can also solve this problem actually by solving the determinant

$$\begin{vmatrix} 0 & T \\ T & \Delta \end{vmatrix} = -E(\Delta - E) - T^2 = E^2 - \Delta E - T^2 = 0. \tag{595}$$

This gives

$$E_{1,2} = \frac{\Delta}{2} \pm \frac{1}{2}\sqrt{\Delta^2 + 4T^2} = \frac{\Delta}{2} \pm \frac{\Delta}{2}\sqrt{1 + \frac{4T^2}{\Delta^2}} \cong \frac{\Delta}{2} \pm \frac{\Delta}{2}(1 + \frac{2T^2}{\Delta^2}) = \begin{cases} \Delta + \frac{T^2}{\Delta} \\ -\frac{T^2}{\Delta} \end{cases}, \tag{596}$$

which is in agreement with perturbation theory in the limit $T \gg \Delta$.

Example

Let us consider a three-level system given by the matrix

$$H = \begin{pmatrix} 0 & -t & -T \\ -t & 0 & -T \\ -T & -T & \Delta \end{pmatrix} \quad \begin{matrix} |1\rangle \\ |2\rangle \\ |3\rangle \end{matrix} \tag{597}$$

In the case that $t \ll \Delta$ and $T \ll \Delta$, we can see this as a Hamiltonian plus two different perturbations:

$$H_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta \end{pmatrix}, \qquad H_t = \begin{pmatrix} 0 & -t & 0 \\ -t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad H_T = \begin{pmatrix} 0 & 0 & -T \\ 0 & 0 & -T \\ -T & -T & 0 \end{pmatrix}. \tag{598}$$

So we have three levels of which two a degenerate. So we have two choices: first applying degenerate perturbation theory and then nondegenerate perturbation theory or the other way around. The former is simpler, so let's do that first. Degenerate perturbation theory for H_t involves solving a 2×2 eigenvalue problem:

$$\begin{vmatrix} -E & -t \\ -t & -E \end{vmatrix} \qquad \Rightarrow \qquad E = \pm t \qquad \Rightarrow \qquad |\mp t\rangle = \frac{1}{\sqrt{2}}(|1\rangle \pm |2\rangle). \tag{599}$$

Although we can call this degenerate perturbation theory, we can also say we are solving the problem but restricting ourselves to the basis set containing $|1\rangle$ and $|2\rangle$. We then can write down the matrix for the new basis set $|-t\rangle$, $|t\rangle$, and $|3\rangle$. For this we need to know the matrix elements

$$\langle 3|H_T|\mp t\rangle = \frac{1}{\sqrt{2}}(|1\rangle \pm |2\rangle) = \begin{cases} -\sqrt{2}T \\ 0 \end{cases}. \tag{600}$$

Let us consider a three-level system given by the matrix

$$H = \begin{pmatrix} -t & 0 & -\sqrt{2}T \\ 0 & t & 0 \\ -\sqrt{2}T & 0 & \Delta \end{pmatrix} \quad \begin{matrix} |-t\rangle \\ |t\rangle \\ |3\rangle \end{matrix} \tag{601}$$

This matrix reduces to a 2×2 and a 1×1 matric since $|t\rangle$ does not couple to any other states. Solving the 2×2 matrix to lowest order gives and energy shift of $2T^2/\Delta$. So we find the following eigenvalues (to lowest order): $-t - 2T^2/\Delta, t, \Delta + 2T^2/\Delta$. Note that we take the difference $E_3^0 - E_{1,2}^0 = \Delta$ and not $E_3^0 - E_{\pm t}^1 = \Delta + t$. This would give a more accurate description but is an order higher in perturbation. This higher-order correction would be $2T^2/(\Delta + t) \cong \frac{2T^2}{\Delta}(1 - \frac{t}{\Delta})$

We could also have started with the perturbation H_T. To lowest order, this will gives energies

$$E_3 = E_3^0 + \sum_{i=1,2} \frac{|\langle i|H_T|3\rangle|^2}{E_3^0 - E_i^0} = \Delta + \frac{2T^2}{\Delta}, \quad E_1 = -\frac{T^2}{\Delta}, \quad E_2 = -\frac{T^2}{\Delta}, \quad . \tag{602}$$

We now need to apply the perturbation H_t. This will couple the perturbed states 1 and 2:

$$|1'\rangle = |1\rangle + \frac{T}{\Delta}|3\rangle, \qquad |2'\rangle = |2\rangle + \frac{T}{\Delta}|3\rangle. \tag{603}$$

Note that it appears that these states are not normalized. However, the normalization is $(1 + \left(\frac{T}{\Delta}\right)^2)^{-1/2} \cong 1 - \frac{1}{2}\left(\frac{T}{\Delta}\right)^2$. This means that the wavefunction is normalized to order T/Δ the normalization is a higher-order effect. We have to take some care in calculating the matrix element between $|1'\rangle$ and $|2'\rangle$,

$$\langle 1'|H|2'\rangle = \langle 1|H_t|2\rangle + \frac{T}{\Delta}(\langle 1|H_T|3\rangle + \langle 3|H_T|2\rangle) + \left(\frac{T}{\Delta}\right)^2 \langle 3|H_0|3\rangle = -t - 2\frac{T^2}{\Delta} + \frac{T^2}{\Delta} = -t - \frac{T^2}{\Delta}. \tag{604}$$

This leads to the following eigenvalue problem

$$\begin{vmatrix} -\frac{T^2}{\Delta} - E & -t - \frac{T^2}{\Delta} \\ -t - \frac{T^2}{\Delta} & -\frac{T^2}{\Delta} - E \end{vmatrix} \qquad \Rightarrow \qquad E = -\frac{T^2}{\Delta} \pm (-t - \frac{T^2}{\Delta}) = \begin{cases} -t - \frac{2T^2}{\Delta} \\ t \end{cases}, \tag{605}$$

giving the same result as above.

B. Time-dependent perturbation theory

Another important type of perturbations are time-dependent ones. Suppose at a certain point a perturbation $H_1(t)$ is turned on (say electromagnetic waves) that leads to transition between different states. Thus our total Hamiltonian is given by

$$H(t) = H_0 + H_1(t). \tag{606}$$

The eigenfunctions of the unperturbed Hamiltonian are

$$|\psi_n(t)\rangle = e^{-i\omega_n t}|\psi_n\rangle \qquad \text{with} \qquad H_0|\psi_n\rangle = E_n|\psi_n\rangle = \hbar\omega_n|\psi_n\rangle. \tag{607}$$

The real eigenstates (with perturbation) can be written as a linear combination of the unperturbed states

$$|\psi(t)\rangle = \sum_n c_n(t)|\psi_n(t)\rangle, \tag{608}$$

where the time-dependence of the wavefunction due to the pertubation is put in the coefficients $c_n(t)$. This wavefunction is a solution of the Schrödinger equation

$$\{H_0 + H_1(t)\}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \sum_n \left[E_n|\psi_n(t)\rangle + i\hbar\frac{dc_n(t)}{dt}|\psi_n(t)\rangle\right]. \tag{609}$$

Taking the inner product with $\langle\psi_k(t)|$ gives

$$i\hbar\frac{dc_k(t)}{dt} = \sum_n \langle\psi_k(t)|H_1(t)|\psi_n(t)\rangle c_n(t). \tag{610}$$

Now let us suppose that initially our system is in state m. So the unperturbed coefficients are $c_m^0 = 1$ and $c_k^0 = 0$ for $k \neq m$. To lowest order the coefficients can be written as $c_n(t) = c_n^0 + c_n^1(t)$, which means for $k \neq m$

$$i\hbar\frac{dc_k(t)}{dt} = i\hbar\frac{dc_k^1(t)}{dt} \cong \sum_n \langle\psi_k(t)|H_1(t)|\psi_n(t)\rangle c_n^0 = \langle\psi_k(t)|H_1(t)|\psi_m(t)\rangle. \tag{611}$$

Integrating gives

$$c_k(t) = \frac{1}{i\hbar}\int_{-\infty}^{t} \langle\psi_k(t')|H_1(t')|\psi_m(t')\rangle dt'. \tag{612}$$

Now suppose $H_1(t)$ can be factorized as $H_1 f(t)$, then we can write

$$c_k(t) = \frac{1}{i\hbar}\int_{-\infty}^{t} \langle\psi_k(t')|H_1|\psi_m(t')\rangle f(t')dt' = \frac{1}{i\hbar}\langle\psi_k|H_1|\psi_m\rangle \int_{-\infty}^{t} e^{-i(\omega_m - \omega_k)t'} f(t')dt'. \tag{613}$$

The probability of finding the system in state k is given by $|c_k|^2$.

C. Transition probabilities

Let us now consider the case of an electromagnetic field that is switched on at $t = 0$. The perturbation is then

$$H_1(t) = \begin{cases} 0 & t < 0 \\ 2H_1 \cos\omega t & t \geq 0 \end{cases}, \tag{614}$$

where the factor 2 is introduced to obtain the traditional expression for Fermi's Golden Rule in the end. For the transition from an initial state i to a final state f, we find

$$c_f(t) = \frac{1}{i\hbar}\langle f|H_1|i\rangle \int_0^t e^{i\omega_{fi}t'}(e^{-i\omega t'} + e^{i\omega t'})dt' \tag{615}$$

$$= -\frac{1}{\hbar}\langle f|H_1|i\rangle \left[\frac{e^{i(\omega_{fi}-\omega)t}-1}{\omega_{fi}-\omega} + \frac{e^{i(\omega_{fi}+\omega)t}-1}{\omega_{fi}+\omega}\right] \tag{616}$$

where $\omega_{fi} = \omega_f - \omega_i$. Using

$$e^{i\theta} - 1 = e^{i\theta/2}(e^{i\theta/2} - e^{-i\theta/2}) = 2ie^{i\theta/2}\sin(\theta/2), \tag{617}$$

we can rewrite the equation as

$$c_f(t) = -\frac{2i}{\hbar}\langle f|H_1|i\rangle \left[\frac{e^{i(\omega_{fi}-\omega)t/2}\sin[(\omega_{fi}-\omega)t/2]}{\omega_{fi}-\omega} + \frac{e^{i(\omega_{fi}+\omega)t/2}\sin[(\omega_{fi}+\omega)t/2]}{\omega_{fi}+\omega}\right]. \tag{618}$$

We can distinguish two cases where c_f peaks as a function of ω. In the first case $\hbar\omega = \hbar\omega_{fi} = E_f - E_i$ the final state energy is higher than the initial state energy. Therefore, the atom absorbs energy and goes into a higher state. In the second case $\hbar\omega = -\hbar\omega_{fi} = E_i - E_f$ the initial state has a higher energy. This means that a photon is emitted from the system. Note that this decay process is stimulated by the presence of radiation of the same frequency.

Let us consider the case of aborption, i.e., $\omega_{fi} > 0$. We find that the transition probability is given by

$$P_{i\to f} = |c_f(t)|^2 = \frac{4|\langle f|H_1|i\rangle|^2}{\hbar^2(\omega_{fi}-\omega)^2}\sin^2[\frac{1}{2}(\omega_{fi}-\omega)t], \tag{619}$$

see Fig. 13.9 in Liboff. The states in the region

$$|\hbar\omega_{fi} - \hbar\omega| = |E_f - (E_i + \hbar\omega)|^2 \leq \frac{2\pi\hbar}{t} \cong \Delta E \tag{620}$$

have greatest probability of being excited. When we study the long-time evolution of the, we can use the approximation

$$\delta(\omega) = \frac{2}{\pi}\lim_{t\to\infty}\frac{\sin^2(\omega t/2)}{t\omega^2}, \tag{621}$$

so that

$$P_{i\to f} = \frac{2\pi t}{\hbar^2}|\langle f|H_1|i\rangle|^2\delta(\omega_{fi}-\omega). \tag{622}$$

The transition probability per time unit is now

$$p_{i\to f} = \frac{dP_{i\to f}}{dt} = \frac{2\pi}{\hbar}|\langle f|H_1|i\rangle|^2\delta(E_f - E_i - \hbar\omega), \tag{623}$$

where the δ function is now expresses the conservation of energy. This is known as Fermi's golden rule.

D. Interaction of radiation and atom

The Hamiltonian including the electromagnetic field described by the vector potential $\mathbf{A}$ is given by

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + V(\mathbf{r}). \tag{624}$$

Expanding the kinetic energy gives the interaction terms

$$H_1 = -\frac{e}{2m}(\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}) \qquad \text{and} \qquad H_1' = \frac{e}{2m}\mathbf{A}^2. \tag{625}$$

The latter term is squared in $\mathbf{A}$. For a nonzero matrix element, the number of photons is conserved. It therefore scatters the radiation. This is known as Thompson scattering. We shall neglect it in the further discussion.

The matrix element is now given by

$$\langle f|H_1|i\rangle = -\frac{e}{2m}\int d\mathbf{r}\psi_f^*(\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p})\psi_i. \tag{626}$$

Using that

$$\nabla\cdot(\mathbf{A}\psi) = \nabla\cdot\mathbf{A}\psi + \mathbf{A}\nabla\psi. \tag{627}$$

If we take the so-called Coulomb gauge, we have $\nabla\cdot\mathbf{A}=0$, and

$$\langle f|H_1|i\rangle = -\frac{e}{m}\int d\mathbf{r}\psi_f^*\mathbf{A}\cdot\mathbf{p}\psi_i. \tag{628}$$

We can distinguish two different matrix elements

$$\langle f|H_1|i\rangle \qquad \rightarrow \qquad \langle f|H_1|\hbar\omega;i\rangle \qquad \text{or} \qquad \langle\hbar\omega;f|H_1|i\rangle, \tag{629}$$

corresponding to the absorption and emission of a photon, respectively.

Now suppose we describe the vector potential by plane wave

$$\mathbf{A}(\mathbf{r},t) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\sum_{\alpha=1,2}\left[e_{\mathbf{k}\alpha}\boldsymbol{\varepsilon}^{(\alpha)}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \text{complex conjugate}\right], \tag{630}$$

where $\boldsymbol{\varepsilon}^{(\alpha)}$ denotes the two polarization vectors of the radiation and $e_{\mathbf{k}\alpha}$ complex coefficients. Using this potential gives a complicated problem. Often calculation are restricted to the lowest orders of the expansion of the vector potential in $\mathbf{r}$

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 1 + i\mathbf{k}\cdot\mathbf{r} + \cdots \tag{631}$$

The first term gives the (electric) dipole approximation. This gives a matrix element proportional to $\mathbf{r}$. This can be seen as follows. The matrix element is now proportional to

$$\frac{e}{m}\boldsymbol{\varepsilon}^{(\alpha)}\cdot\langle f|\mathbf{p}|i\rangle \tag{632}$$

Making use of the commutation relation

$$[\mathbf{p}^2,\mathbf{r}] = -2i\hbar\mathbf{p} \qquad \Rightarrow \qquad \mathbf{p} = \frac{im}{\hbar}[H_{\text{atom}},\mathbf{r}], \tag{633}$$

we find

$$\frac{e}{m}\frac{im}{\hbar}\boldsymbol{\varepsilon}^{(\alpha)}\cdot\langle f|(H_{\text{atom}}\mathbf{r}-\mathbf{r}H_{\text{atom}})|i\rangle = \frac{ie}{\hbar}(E_f-E_i)\boldsymbol{\varepsilon}^{(\alpha)}\cdot\langle f|\mathbf{r}|i\rangle = ie\omega\boldsymbol{\varepsilon}^{(\alpha)}\cdot\langle f|\mathbf{r}|i\rangle, \tag{634}$$

which is an expectation value of $\mathbf{r}$, hence the name dipole approximation. Now let us make use of the fact that we can write the components in spherical coordinates as

$$r_q = r\sqrt{\frac{4\pi}{3}}Y_q^1(\theta,\varphi) \rightarrow \begin{cases} r_1 = -r\sin\theta e^{i\varphi} = -\frac{1}{\sqrt{2}}(x+iy) \\ r_0 = r\cos\theta = z \\ r_1 = r\sin\theta e^{-i\varphi} = \frac{1}{\sqrt{2}}(x-iy) \end{cases}. \tag{635}$$

Note that the direction of the polarization vectors are $\boldsymbol{\varepsilon}_+ = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x+i\hat{\mathbf{e}}_y)$, $\boldsymbol{\varepsilon}_z = \hat{\mathbf{e}}_z$, $\boldsymbol{\varepsilon}_- = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x-i\hat{\mathbf{e}}_y)$, where $\hat{\mathbf{e}}_i$ are the axis of the system, i.e., the atom in our case. $\boldsymbol{\varepsilon}_+$ and $\boldsymbol{\varepsilon}_-$ are left and right circularly polarized light, respectively; $\boldsymbol{\varepsilon}_z$ is linearly polarized light along the z axis. Obviously, for a free atom, the choice of our axis does not make any difference for the spectrum that we obtain. The situation changes when we put the atom in a magnetic field. This will make the spectra strongly depend on the magnetic field. We can now evaluate the matrix elements for atomic orbitals

$$\langle n'l'm'|r_q|nlm\rangle \sim \langle l'm'|r|lm\rangle\langle n'l'|Y_q^1|nl\rangle = \int_0^\infty dr\ r^3R_{n'l'}(r)R_{nl}(r)\int(Y_{m'}^{l'})^*Y_q^1Y_m^l d\Omega. \tag{636}$$

The evaluation is rather complicated but leads to some simple selection rules

$$\Delta l = l - l' = \pm 1 \tag{637}$$

$$\Delta m = m - m' = 0, \pm 1. \tag{638}$$

The integration over the angular coordinates is related to the vector coupling or Clebsch-Gordan coefficients, that we will meet again when we start coupling angular momenta. The integration over φ is straightforward and gives $\int d\varphi \exp(i(-m' + q + m)\varphi)$, leading to the m selection rules. The integration over θ gives l selection rules. The second term in Eqn. (631) gives electric quadrupole (proportional to $\mathbf{r}^2$) and magnetic dipole transitions (proportional to $\mathbf{r} \times \mathbf{p} = \mathbf{L}$). These are small, but certainly not always negligible.

The combination of energy conservation and the selection rules provide a powerful tool to study the electronic states of atoms but also molecules and solids.

X. MANY-ELECTRON ATOM

Before we can start talking about filling the orbtitals we have to introduce an important concept: Pauli's exclusion principle. Without it, the most obvious thing to do is putting all the electrons in the lowest level, i.e., $nl = 10$. Experimentally, this is not observed, but one sees that each orbital from the Hydrogen atom contains only one electron. This is quite unlike, for example, photons. We can put as many photons in a particular photon state (say, a mode in an electromagnetic cavity or a plane wave mode) as we want. This corresponds simply to increasing the intensity of the radiation.

A. Pauli's exclusion principe

In quantum mechanics, particles are indistinguishable. This means that we should be able to interchange two particles without seeing any effect on the probability function. We can distinguish two fundamentally different types of particles, namely those with a symmetric and antisymmetric wavefunction. The quantization of the harmonic oscillator, photon, and a number of fundamental particles have symmetric wavefunctions. These particles are called bosons. Here we find that a permutation between particle i and j for a N particle wavefunction

$$P_{ij}\psi(k_1\mathbf{r}_1,\ldots,k_i\mathbf{r}_i,\ldots,k_j\mathbf{r}_J,\ldots,k_N\mathbf{r}_N) = \psi(k_1\mathbf{r}_1,\ldots,k_j\mathbf{r}_j,\ldots,k_i\mathbf{r}_i,\ldots,k_N\mathbf{r}_N), \tag{639}$$

where k_i means that a particle is in a certain quantum state k_i (for example a plane wave). Note that for the harmonic oscillator we considered only one type of quantum state, but in principly we can take a set of oscillators. Let us look at two bosons. The wavefunction is given by

$$\psi_S(k_1r_1,k_2r_2) = \frac{1}{\sqrt{2}}(\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) + \varphi_{k_2}(\mathbf{r}_1)\varphi_{k_1}(\mathbf{r}_2)). \tag{640}$$

Obviously the interchange of the particles does not affect the wavefunction. However, the other type of particles, called Fermions, have antisymmetric wavefunctions. These are, e.g., electrons and protons. This gives

$$P_{ij}\psi(k_1\mathbf{r}_1,\ldots,k_i\mathbf{r}_i,\ldots,k_j\mathbf{r}_j,\ldots,k_N\mathbf{r}_N) = -\psi(k_1\mathbf{r}_1,\ldots,k_j\mathbf{r}_j,\ldots,k_i\mathbf{r}_i,\ldots,k_N\mathbf{r}_N). \tag{641}$$

Although this looks like a simple phasefactor and would make absolutely no difference in the probability density ψ^2, it has very important consequences. Let us take $k_i = k_j$, this would mean that the wavefunction on the right and left hand side are equal. However, this implies that $\psi = -\psi$, which is only possible if $\psi = 0$. Therefore, two Fermions can never be in the same quantum state. A two-fermion state would look like this

$$\psi_A(k_1r_1,k_2r_2) = \frac{1}{\sqrt{2}}(\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) - \varphi_{k_1}(\mathbf{r}_2)\varphi_{k_2}(\mathbf{r}_1)). \tag{642}$$

The symmetric and antisymmetric wavefunctions give two fundamentally different particles. Examples of bosons are photons, π and K mesons. These are particles with integral spin. Also the harmonic oscillator behaves like a "boson"-like particle. Particles with half integral spin behave like fermions. Examples are electrons, protons, and neutrons.

For bosons the general wavefunction is

$$\Psi_{k_1,\ldots,k_N} = \sqrt{\frac{\Pi_i n_{k_i}!}{N!}}\sum_P \psi_{k_1}(\mathbf{r}_1)\cdots\psi_{k_N}(\mathbf{r}_N), \tag{643}$$

where the summations runs over all the possible permutations of *different* indices k_i. This gives the normalization constant. n_{k_i} indicates how many k-values are k_i. Note that the $\sum_i n_{k_i} = N$. For two particles with $S = 0$, we have

$$\Psi_{k_1,k_2} = \frac{1}{\sqrt{2}}\left\{\psi_{k_1}(\mathbf{r}_1)\psi_{k_2}(\mathbf{r}_2) + \psi_{k_2}(\mathbf{r}_1)\psi_{k_1}(\mathbf{r}_2)\right\}. \tag{644}$$

For $k_1 = k_2 = k$, we find

$$\Psi_{k_1,k_2} = \sqrt{\frac{2!}{2!}}\psi_k(\mathbf{r}_1)\psi_k(\mathbf{r}_2) = \psi_k(\mathbf{r}_1)\psi_k(\mathbf{r}_2), \tag{645}$$

using the fact that $n_k = 2$.

Remember that for the harmonic oscillator we could write the wave function like

$$|n\rangle = \frac{1}{\sqrt{N!}}(a^\dagger)^N|0\rangle. \tag{646}$$

Now forget for the moment the relation of $a^\dagger$ with x and p, since this result is far more general. We can in principle make different types of bosons. These operators have some nice characteristics that we would like to see in wavefunctions: they are indistinguishable. Also they are symmetric. This follows from the more general commutation relations

$$[a_k, a_{k'}^\dagger] = \delta_{k,k'} \qquad \text{and} \qquad [a_k^\dagger, a_{k'}^\dagger] = 0 \qquad \text{and} \qquad [a_k, a_k'] = 0 \tag{647}$$

Thus for $k \neq k'$, we can write

$$a_k^\dagger a_{k'}^\dagger = a_{k'}^\dagger a_k^\dagger. \tag{648}$$

Therefore, interchange of two particles does not change the sign of the wavefunction. Let us look at the two-particle wavefunctions. If the two particles are in the same state (i.e. have the same quantum number) then the wavefunction should be similar to that of the harmonic oscillator:

$$|kk\rangle = \frac{1}{\sqrt{2}}a_k^\dagger a_k^\dagger|0\rangle. \tag{649}$$

This wavefunction is now properly normalized since

$$\begin{aligned}
\langle kk|kk\rangle &= \frac{1}{2}\langle 0|a_k a_k a_k^\dagger a_k^\dagger|0\rangle = \frac{1}{2}\langle 0|a_k(1+a_k^\dagger a_k)a_k^\dagger|0\rangle && (650)\\
&= \frac{1}{2}\langle 0|a_k a_k^\dagger|0\rangle + \frac{1}{2}\langle 0|a_k a_k^\dagger a_k a_k^\dagger|0\rangle && (651)\\
&= \frac{1}{2}\langle 0|(1+a_k^\dagger a_k)|0\rangle + \frac{1}{2}\langle 0|(1+a_k^\dagger a_k)(1+a_k^\dagger a_k)|0\rangle = 1 && (652)\\
& && (653)
\end{aligned}$$

since $a_k|0\rangle = 0$. For two different wavevectors, we have

$$|k_1 k_2\rangle = a_{k_1}^\dagger a_{k_2}^\dagger|0\rangle. \tag{654}$$

This gives the right normalization since

$$\langle k_1k_2|k_1k_2\rangle = \langle 0|a_{k_1}a_{k_2}a_{k_2}^\dagger a_{k_1}^\dagger|0\rangle = \langle 0|a_{k_1}(1+a_{k_2}^\dagger a_{k_2})a_{k_1}^\dagger|0\rangle = \langle 0|a_{k_1}a_{k_1}^\dagger|0\rangle = \langle 0|(1+a_{k_1}^\dagger a_{k_1})|0\rangle = 1 \tag{655}$$

For wavefunctions there is a dependence on the particle by the coordinate $\varphi_k(\mathbf{r}_i)$. The operators are identical, which means that we do not have to make all the different permutations. For more than one k value, we can generalize the wavefunction as

$$|\{n_{k_i}\}\rangle = \frac{1}{\sqrt{\Pi_i n_{k_i}}}\Pi_i(a_{k_i}^\dagger)^{n_{k_i}}|0\rangle. \tag{656}$$

The situation for electrons (and other spin-1/2 particles) is fact a bit more complicated since the total wavefunction, i.e., orbital and spin, has to be asymmetric. This means $\psi_S(\mathbf{r}_1,\mathbf{r}_2)\xi_A(1,2)$ or $\psi_A(\mathbf{r}_1,\mathbf{r}_2)\xi_S(1,2)$, where $\xi(1,2)$ is the wavefunction for the spin. There are four ways to combine two spins $\xi_\uparrow(1)\xi_\uparrow(2)$, $\xi_\uparrow(1)\xi_\downarrow(2)$, $\xi_\downarrow(1)\xi_\uparrow(2)$, and $\xi_\downarrow(1)\xi_\downarrow(2)$. $\xi_\uparrow(1)\xi_\uparrow(2)$ and $\xi_\downarrow(1)\xi_\downarrow(2)$ are clearly symmetric spin functions, since interchange of the particles does not change the sign of the wave function. These are the $\xi_{11}(1,2)$ and $\xi_{1,-1}(1,2)$ components of the triplet ($S=1$). Since $S_z = 1, 0, -1$ for $S=1$, there should also be a $S=0$ component. This can be obtained by using the step operator for the spin

$$\begin{aligned}
S_-\xi_{11} &= \sqrt{(1+1)(1-1+1)}\xi_{10} = \sqrt{2}\xi_{10} = \xi_\uparrow(1)\xi_\downarrow(2) + \xi_\downarrow(1)\xi_\uparrow(2) && (657)\\
&\Rightarrow \qquad \xi_{10} = \frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2) + \xi_\downarrow(1)\xi_\uparrow(2)\right] && (658)
\end{aligned}$$

where Eqn. (412) has been used. Note that this is also a symmetric wavefunction. The remaining wave function has to be orthogonal to the previous one

$$\xi_{00}(1,2) = \frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2) - \xi_\downarrow(1)\xi_\uparrow(2)\right] \tag{659}$$

which is antisymmetric. Since there is only one component (called a singlet), this has to be $S = 0$. The four total wavefunctions for two electrons are therefore $\Psi_{1S_z}(\mathbf{r}_1, \mathbf{r}_2) = \psi_A(\mathbf{r}_1, \mathbf{r}_2)\xi_{1S_z}(1,2)$ with $S_z = 1, 0, -1$ and $\Psi_{00}(\mathbf{r}_1, \mathbf{r}_2) = \psi_S(\mathbf{r}_1, \mathbf{r}_2)\xi_{00}(1,2)$. The $S = 1$ and $S_z = 1$ is given by

$$\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \psi_A(\mathbf{r}_1, \mathbf{r}_2)\xi_{11}(1,2) = \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) - \varphi_{k_2}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_1)\right]\xi_\uparrow(1)\xi_\uparrow(2) \tag{660}$$

$$= \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) - \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(2)\right] \tag{661}$$

$$= \frac{1}{\sqrt{2}}\begin{vmatrix} \varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2) \\ \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) \end{vmatrix} \tag{662}$$

The $S = 0$ and $S_z = 0$ wavefunction is given by

$$\Psi_{00}(\mathbf{r}_1, \mathbf{r}_2) = \psi_S(\mathbf{r}_1, \mathbf{r}_2)\xi_{00}(1,2) = \frac{1}{\sqrt{2}}\left[\varphi_{k_1}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_2) + \varphi_{k_2}(\mathbf{r}_1)\varphi_{k_2}(\mathbf{r}_1)\right]\frac{1}{\sqrt{2}}\left[\xi_\uparrow(1)\xi_\downarrow(2) - \xi_\downarrow(1)\xi_\uparrow(2)\right] \tag{663}$$

$$= \frac{1}{2}\left[\varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\downarrow(2) - \varphi_{k_2}(\mathbf{r}_1)\xi_\downarrow(1)\varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2)\right. \tag{664}$$

$$\left. -\varphi_{k_1}(\mathbf{r}_1)\xi_\downarrow(1)\varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) + \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1)\varphi_{k_1}(\mathbf{r}_2)\xi_\downarrow(2)\right] \tag{665}$$

Note that this can also be written as a combination of Slater determinants

$$\Psi_{00}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\begin{vmatrix} \varphi_{k_1}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\uparrow(2) \\ \varphi_{k_2}(\mathbf{r}_1)\xi_\downarrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\downarrow(2) \end{vmatrix} - \frac{1}{\sqrt{2}}\begin{vmatrix} \varphi_{k_1}(\mathbf{r}_1)\xi_\downarrow(1) & \varphi_{k_1}(\mathbf{r}_2)\xi_\downarrow(2) \\ \varphi_{k_2}(\mathbf{r}_1)\xi_\uparrow(1) & \varphi_{k_2}(\mathbf{r}_2)\xi_\uparrow(2) \end{vmatrix}\right) \tag{666}$$

$$= \frac{1}{\sqrt{2}}[\psi_{k_1\uparrow,k_2\downarrow}(\mathbf{r}_1, \mathbf{r}_2) - \psi_{k_1\downarrow,k_2\uparrow}(\mathbf{r}_1, \mathbf{r}_2)] \tag{667}$$

In fact, we are really entering already the realm of many-body physics (although you might not directly realize the significance of it). Suppose, $\Psi_{k_1k_2,1S_z}$ and $\Psi_{k_1k_2,00}$ have different energies, then the energy of the two-particle state depends strongly on the relation of the two particles in the wavefunction. A lot of theories (band structure calculations in particular, a very large number of people are working in that area in solid-state physics) do not take into account that difference. In so-called independent-particle models, Ψ_{10} and Ψ_{00} are essentially equivalent. In that case, one might as well use $\xi_\uparrow(1)\xi_\downarrow(2)$ and $\xi_\downarrow(1)\xi_\uparrow(2)$, which are simpler to handle.

Note that in general we can built up antisymmetric states from Slater determinants

$$\Psi_{k_1\sigma_1,\ldots,k_N\sigma_N}(\mathbf{r}_1 \cdots \mathbf{r}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \varphi_{k_1}(\mathbf{r}_1)\xi_{\sigma_1}(1) & \cdots & \varphi_{k_1}(\mathbf{r}_N)\xi_{\sigma_1}(N) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \varphi_{k_N}(\mathbf{r}_1)\xi_{\sigma_N}(1) & \cdots & \varphi_{k_N}(\mathbf{r}_N)\xi_{\sigma_N}(N) \end{vmatrix} \tag{668}$$

where $\sigma_i = \uparrow, \downarrow$. As a result of the properties of the determinant, having two columns with the same k and σ implies that the determinant is zero, just as for the exclusion principle. Interchange of two columns also leads to a change in sign of the determinant, giving us directly the antisymmetry. This is obvious for the $S_z = 1, -1$ states of the triplet. For the $S_z = 0$ states, one needs a combination of Slater determinants

$$\Psi_{k_1k_2,00} = \frac{1}{\sqrt{2}}\left(\psi_{k_1\uparrow,k_2\downarrow} - \psi_{k_1\downarrow,k_2\uparrow}\right) \qquad \text{and} \qquad \Psi_{k_1k_2,10} = \frac{1}{\sqrt{2}}\left(\psi_{k_1\uparrow,k_2\downarrow} + \psi_{k_1\downarrow,k_2\uparrow}\right). \tag{669}$$

Working with Slater determinants becomes quite elaborate. Often it is much more convenient to work with operators in a similar way as we saw for bosons. However, unlike bosons that satisfy commutation relations, fermions satisfy anticommutation relations

$$\{c_k, c_{k'}^\dagger\} = \delta_{k,k'}, \qquad \{c_k^\dagger, c_{k'}^\dagger\} = 0, \qquad \{c_k, c_{k'}\} = 0, \tag{670}$$

where the term in braces is defined as

$$\{A, B\} = AB + BA. \tag{671}$$

Suppose, we have a "vacuum" state $|0\rangle$. We can then put a fermion in there by

$$c_k^\dagger|0\rangle = |k\rangle, \tag{672}$$

where $|k\rangle$ indicates that there is one electron in state k and zero in all the others (note that according to Pauli's exclusion principle there can only be zero or one electron in a state).

However, we cannot put two particles in the same state, since

$$c_k^\dagger c_k^\dagger |0\rangle = \frac{1}{2}\{c_k^\dagger, c_k^\dagger\}|0\rangle = 0. \tag{673}$$

We can put them in different quantum states, thus for $k \neq k'$

$$c_k^\dagger c_{k'}^\dagger |0\rangle = -c_{k'}^\dagger c_k^\dagger |0\rangle. \tag{674}$$

Note that the operators directly take care of the antisymmetry of the wavefunctions. Including spin, we can write the wavefunctions $|k_1 k_2, SS_z\rangle$ for two particles in states labeled with quantum numbers k_1 and k_2 with different spin S and S_z as

$$|k_1 k_2, 11\rangle = c_{k_2\uparrow}^\dagger c_{k_1\uparrow}^\dagger |0\rangle \quad , \quad |k_1 k_2, 10\rangle = \frac{1}{\sqrt{2}}\left(c_{k_2\downarrow}^\dagger c_{k_1\uparrow}^\dagger + c_{k_2\uparrow}^\dagger c_{k_1\downarrow}^\dagger\right)|0\rangle \quad , \quad |k_1 k_2, 1, -1\rangle = c_{k_2\downarrow}^\dagger c_{k_1\downarrow}^\dagger |0\rangle \tag{675}$$

$$|k_1 k_2, 00\rangle = \frac{1}{\sqrt{2}}\left(c_{k_2\downarrow}^\dagger c_{k_1\uparrow}^\dagger - c_{k_2\uparrow}^\dagger c_{k_1\downarrow}^\dagger\right)|0\rangle. \tag{676}$$

In general, we can write the many-body Fermion wavefunction as

$$|\{n_{k_i}\}\rangle = \Pi_i (c_{k_i}^\dagger)|0\rangle. \tag{677}$$

Note, that the normalization is unity since we can never have more than one fermion in the same quantum state.

B. The Helium atom

Let us start with the Schrödinger equation for the Helium atom

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} + \frac{e^2}{4\pi\varepsilon_0}\left(-\frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}},\right) \tag{678}$$

takin $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. Note that this is already a problem which is impossible to solve exactly. Since we increase the complexity of the system, we simplify the Hamiltonian by leaving out the relativistic terms. Without the last term, this problem could be solved easily, giving Hydrogen-like solutions. The two electrons would simply go into the $1s$ level ($nl = 10$) giving a ground state singlet

$${}^1\Psi_{10}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)\xi_{00}, \tag{679}$$

for $nlm = 100$ and where the spin singlet ξ_{00} is given in Eqn. (659). The superscript in ${}^1\Psi_{10}$ stands for $2S + 1 = 1$, the subscript for $nl = 10$. The energy would simply be

$$E_0 = 2 \times R\frac{Z^2}{n^2} = 2 \times (-13.6)\frac{2^2}{1^2} = -8 \times 13.6 = -108.8 \quad \text{eV}. \tag{680}$$

This is quite far of the experimentally observed value of -78.975 eV. Let us therefore try to get an estimate of the electron-electron interaction using perturbation theory.

Perturbation theory. The lowest order correction is

$$E_1 = \langle {}^1\Psi_{10}|\frac{e^2}{4\pi\varepsilon_0 r_{12}}|{}^1\Psi_{10}\rangle = \frac{e^2}{4\pi\varepsilon_0}\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{|\psi_{100}(\mathbf{r}_1)|^2|\psi_{100}(\mathbf{r}_2)|^2}{r_{12}} \tag{681}$$

$$= \frac{e^2}{4\pi\varepsilon_0}\left(\frac{Z^3}{\pi a_0^3}\right)^2 \int \frac{1}{r_{12}} \exp\left[-\frac{2Z}{a_0}(r_1 + r_2)\right] r_1^2 dr_1 d\Omega_1 r_2^2 dr_2 d\Omega_2. \tag{682}$$

Let us do the integral over $\mathbf{r}_2$ first. Let us fix $\mathbf{r}_1$ along the z axis. Then we can write

$$|\mathbf{r}_1 - \mathbf{r}_2| = \sqrt{r_1^2 - 2r_1 r_2 \cos\theta_2 + r_2^2} \tag{683}$$

We can now write

$$I_2 = \int \frac{e^{-2Zr_2/a_0}}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_2 = 2\pi \int_0^\infty \int_0^\pi \frac{e^{-2Zr_2/a_0}}{\sqrt{r_1^2 - 2r_1 r_2 \cos\theta_2 + r_2^2}} r_2^2 \sin\theta_2 dr_2 d\theta_2 d\varphi_2. \tag{684}$$

The integral over θ_2 is straightforward

$$\int_0^\pi \frac{\sin\theta_2}{\sqrt{r_1^2 - 2r_1 r_2 \cos\theta_2 + r_2^2}} d\theta_2 = \frac{1}{r_1 r_2} \left[\sqrt{r_1^2 - 2r_1 r_2 \cos\theta_2 + r_2^2} \right]_0^\pi \tag{685}$$

$$= \frac{1}{r_1 r_2}(r_1 + r_2 - |r_1 - r_2|) = \begin{cases} \frac{2}{r_1} & r_2 < r_1 \\ \frac{2}{r_2} & r_2 > r_1 \end{cases}. \tag{686}$$

For I_2, we have to separate the integral into two regions

$$I_2 = 4\pi \left[\frac{1}{r_1} \int_0^{r_1} e^{-2Zr_2/a_0} r_2^2 dr_2 + \int_{r_1}^\infty e^{-2Zr_2/a_0} r_2 dr_2 \right] = \frac{\pi a^3}{Z^3 r_1} \left\{ 1 - e^{-2Zr_1/a_0}(1 + \frac{Zr_1}{a_0}) \right\}. \tag{687}$$

Putting this back into the equation for E_1 gives

$$E_1 = \frac{e^2}{4\pi\varepsilon_0} \frac{Z^3}{\pi a^3} \int \left\{ 1 - e^{-2Zr_1/a_0}(1 + \frac{Zr_1}{a_0}) \right\} e^{-2Zr_1/a_0} r_1 \sin\theta_1 dr_1 d\theta_1 d\varphi_1 = \frac{5}{8} \frac{Ze^2}{4\pi\varepsilon_0 a_0}. \tag{688}$$

From this we can calculate the energy

$$E = E_0 + E_1 = -\frac{Ze^2}{4\pi\varepsilon_0 a_0} \left(Z - \frac{5}{8} \right) = -108.8 + 34 = -74.8 \quad \text{eV}. \tag{689}$$

This is a lot closer to the -79 eV, but we have overshot the ground-state energy somewhat. This is understandable. We have just taken the unperturbed ground-state wavefunctions to calculate the electron-electron interaction. Since the electrons repel each other, this is unlikely to be the lowest state. We can gain some energy by increasing the radial extent of the wavefunction somewhat to try to find a new minimum (obviously, from the point of view of the potential energy from the nucleus, expanding the wave function is not a good idea). A larger radial extent of the wavefunction corresponds effectively to a smaller Z of the nucleus. This effect is called screening. An electron feels a smaller attractive force from the nucleus as a result of the repulsive force of the other electron(s). Let us try to find an optimal value for Z by means of a variational approach.

Variational approach. Now let us suppose the electron has a wavefunction in which it sees an effective charge of not $Z = 2$, but $\tilde{Z} < 2$. The presence of the other electron makes the charge slightly lower. Our modified wavefunction then becomes

$${}^1\tilde{\Psi}_{10}(\mathbf{r}_1, \mathbf{r}_2) = \tilde{\psi}_{100}(\mathbf{r}_1)\tilde{\psi}_{100}(\mathbf{r}_2)\xi_{00} = \frac{1}{\pi} \left(\frac{\tilde{Z}}{a_0} \right)^3 \exp(-\tilde{Z}(r_1 + r_2)/a_0)\xi_{00}. \tag{690}$$

Using the above results, we can write down the energy of the Hamiltonian

$$\tilde{H} = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{\tilde{Z}e^2}{4\pi\varepsilon_0 r_1} - \frac{\tilde{Z}e^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 r_{12}}, \tag{691}$$

up to first order in the electron-electron interaction

$$\langle {}^1\tilde{\Psi}_{10} | \tilde{H} | {}^1\tilde{\Psi}_{10} \rangle = -\frac{\tilde{Z}e^2}{4\pi\varepsilon_0 a_0} \left(\tilde{Z} - \frac{5}{8} \right). \tag{692}$$

This leaves us still to evaluate the rest of the Hamiltonian

$$\Delta H = H - \tilde{H} = (\tilde{Z} - Z) \frac{e^2}{4\pi\varepsilon_0} \left(\frac{1}{r_1} + \frac{1}{r_2} \right). \tag{693}$$

Since this is close to the expectation value of the Coulomb potential of the nucleus, we find

$$\Delta E = \langle {}^1\tilde{\Psi}_{10} | \Delta H | {}^1\tilde{\Psi}_{10} \rangle = (\tilde{Z} - Z) \frac{e^2}{4\pi\varepsilon_0} \int d\mathbf{r}_1 d\mathbf{r}_2 \tilde{\psi}_{100}^*(\mathbf{r}_1) \tilde{\psi}_{100}^*(\mathbf{r}_2) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \tilde{\psi}_{100}(\mathbf{r}_1) \tilde{\psi}_{100}(\mathbf{r}_2) \tag{694}$$

$$= (\tilde{Z} - Z) \frac{e^2}{4\pi\varepsilon_0} 2 \times \left(\frac{\tilde{Z}}{a_0} \right) = 2(\tilde{Z} - Z)\tilde{Z} \frac{e^2}{4\pi\varepsilon_0 a_0}, \tag{695}$$

using the fact that $\langle 1/r\rangle = Z/a_0$. The total energy is then given by

$$E = \tilde{E} + \Delta E = -\frac{e^2}{4\pi\varepsilon_0 a_0}\left[\tilde{Z}^2 - \frac{5}{8}\tilde{Z} - 2(\tilde{Z}-Z)\tilde{Z}\right] = -\frac{e^2}{4\pi\varepsilon_0 a_0}\left[-\tilde{Z}^2 - \frac{5}{8}\tilde{Z} + 2Z\tilde{Z}\right]. \tag{696}$$

We want to minimize this energy with respect to $\tilde{Z}$,

$$\frac{dE}{d\tilde{Z}} = -\frac{e^2}{4\pi\varepsilon_0 a_0}\left[-2\tilde{Z} - \frac{5}{8} + 2Z\right] = 0 \qquad \Rightarrow \qquad \tilde{Z} = Z - \frac{5}{16} = \frac{27}{16} = 1.6875\ldots, \tag{697}$$

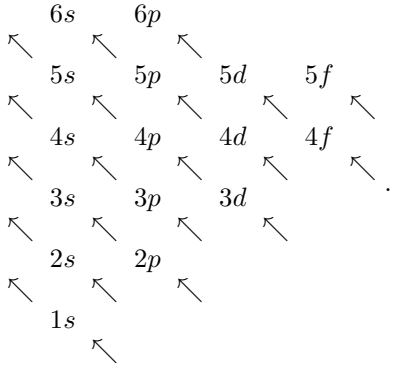
with $Z = 2$ for a Helium atom. Therefore the electron effectively sees a somewhat smaller charge of the nucleus as a result of the presence of the other electron. The energy of the ground state for Helium is then

$$E_{10} = -\left(\tilde{Z} - \frac{5}{16}\right)^2 \frac{e^2}{4\pi\varepsilon_0 a_0^2}. \tag{698}$$

For $Z = 2$, we otbain $E_{10} = -77.38$ eV which is quite close to the experimentally observed value of -79 eV. Further refinements are possible but less didactic.

C. Periodic table

After the discussion of the Helium atom, we are ready to fill the atoms with electrons. Since we are not quite able to calculate everything, we are still left with some phenemenology. From the simple Hydrogen, we would expect that the orbitals with the same quantum number n have the same energy. However, electron-electron interactions severely change this picture. Orbitals with a small radial extent suffer from larger Coulomb interactions and will therefore be filled later. A phenemenological scheme of filling the orbitals is



(699)

This is just some scheme to remember it. We have used here the common notation nl, where for l a symbols is used instead of an integer, i.e., $l = 0, 1, 2, 3 = s, p, d, f$. Note that there are no g orbitals. The rumor goes that when God came to the g orbitals, he found them so complicated that he decided to stop there (forgive the political incorrectness). One starts at the bottom and follows the diagonals, i.e., $1s, 2s, 2p, 3s, 3p, 4s, 3d, \ldots$. We can put $2(2l+1)$ electrons in an orbital. Note that there are exceptions. First, we observe that the higher l numbers are filled later than lower l. A physical argument for that is that the orbitals have to be orthogonal to each other, also in radial extent. That means the R_{nl} has an extra wiggle when l increases by one (see Fig. 10.16 of Liboff). This pushes the electron density to larger r and reduces the electron-electron interactions. Even in a solid, the orbitals where a certain l appears for the first time ($3d$ and $4f$) are more atomic. (the larger Coulomb interaction will also have as a result that the electrons like to have their spins parallel to reduce the Coulomb interactions, as we will see below. These materials often show interesting magnetic properties [first row of transition metals for 3d, and rare earths for $4f$]).

We have already treated Hydrogen and Helium. Helium has a filled $1s$ shell and is the stablest atom, a so-called noble element.

Because of their wider radial extent, the ns levels (with $n > 1$) see less of the nucleus and are therefore note very strongly bound. (See the figure with first ionization energies, Liboff 12.11). For example, for Li, the $1s$ electrons screen the nucleus from the $2s$ electron. Elements with an open s-shell (and no other shells) are unstable. This is also true for solids. Alkali and also Alkaline metals (the first and second row in the periodic table are very unstable). For example, Lithium and Sodium react with water, creating H_2 gas.

Filling the np shell makes the elements stabler. Note the little dip at p^4. Actually this is not a dip but it results from an increased stability of the p^3 configuration making the p^4 configuration look less stable. p^3 is a half-filled shell, and all three the electrons have their spin parallel making an extra stable configuration. This will be discussed in more detail below.

The elements where the p shell is filled are very stable and called noble elements (Ne, Ar, Kr, Xe). They react very little with other elements and usually appear in the gas phase since they interact only via the weak Van der Waals forces, making it very difficult to form a solid. Also a filled d shell causes extra stability. That is why we like silver and gold so much (actually, as you look more closely you see that the d shell "borrows" an electron from the s shell to form a nice closed shell. The same you see for Chromium and Molybdenum, but here to get a nice half-filled shell. Tungsten on the other hand cares less. Here the $6d$ orbital has a larger radial extent and electron-electron interactions in the d shell are not strong enough to be able to "borrow" a s electron.

D. Coulomb interaction

Now that we have filled the shells, we would like to find out what the state is within a certain open shell. This is mainly determined by the Coulomb interaction between the electrons. One important problem in physics are the Coulomb interaction between electrons,

$$V = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{e^2}{\sqrt{r_1^2 - 2r_1 r_2 \cos\omega + r_2^2}}, \tag{700}$$

where ω is the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$. We would like to expand the square-root term in a binomial series:

$$\begin{aligned}(1 - 2xt + t^2)^{-1/2} &= \sum_{n=0}^{\infty} \frac{(2n)!}{2^{2n}(n!)^2}(2xt - t^2)^n = \frac{0!}{2^0(0!)^2}(2xt - t^2)^0 + \frac{2!}{2^2(1!)^2}(2xt - t^2)^1 + \frac{4!}{2^4(2!)^2}(2xt - t^2)^2 + \cdots \\ &= 1 + xt + (\frac{3}{2}x^2 - \frac{1}{2})t^2 + \mathcal{O}(t^3) + \cdots \end{aligned} \tag{701}$$

where we have taken $x = \cos\omega$. The term in front of the t^n are the Legendre polynomials of order n. This function is also known as the generating function for the Legendre polynomials,

$$(1 - 2xt + t^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x) t^n. \tag{702}$$

However, we can only use this theorem for $|t| < 1$. We can achieve this in our potential by defining $r_>$ and $r_<$ the greater, respectively, the smaller of r_1 and r_2:

$$V = \frac{e^2}{r_>}\left(1 + \frac{r_<^2}{r_>^2} - 2\frac{r_<}{r_>}\cos\omega\right)^{-1/2} = \frac{e^2}{r_>}\sum_{k=0}^{\infty}\frac{r_<^k}{r_>^k}P_k(\cos\omega). \tag{703}$$

For Hydrogen-like wavefunctions $\psi_i = R_{n_i l_i}(r_i)Y_{l_i m_i}(\theta_i, \varphi_i)$, the Coulomb matrix element is given by

$$(\psi_1\psi_2|V|\psi_3\psi_4) = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \psi_1^*(\mathbf{r}_1)\psi_2^*(\mathbf{r}_2) V \psi_3(\mathbf{r}_1)\psi_4(\mathbf{r}_2) \tag{704}$$

$$= (\psi_1\psi_2|\frac{e^2}{r_>}\sum_{k=0}^{\infty}\frac{r_<^k}{r_>^k}P_k(\cos\omega)|\psi_3\psi_4) \tag{705}$$

$$= \sum_{k=0}^{\infty}(R_{n_1 l_1}R_{n_2 l_2}|\frac{e^2 r_<^k}{r_>^{k+1}}|R_{n_3 l_3}R_{n_4 l_4})(Y_{l_1 m_1}Y_{l_2 m_2}|P_k(\cos\omega)|Y_{l_3 m_3}Y_{l_4 m_4})\delta_{s_1,s_3}\delta_{s_2,s_4}, \tag{706}$$

where the delta function imply spin conservation. Here we use $(\cdots)$ for the matrix element. The different use from $\langle\cdots\rangle$ will become clear in a minute. We can now use the spherical addition theorem,

$$P_l(\cos\omega) = \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_{lm}^*(\theta_1, \varphi_1)Y_{l,m}(\theta_2, \varphi_2), \tag{707}$$

which connect the Legendre polynomial to the spherical harmonics. We can now write the latter integral as two integrals over the coordinates of $\mathbf{r}_1$ and $\mathbf{r}_2$.

$$\begin{aligned}(Y_{l_1m_1}Y_{l_2m_2}|P_k(\cos\omega)|Y_{l_3m_3}Y_{l_4m_4}) &= \frac{4\pi}{2k+1}\sum_{q=-k}^{k}(Y_{l_1m_1}Y_{l_2m_2}|Y_{kq}(\theta_1\varphi_1)Y^*_{kq}(\theta_2\varphi_2)|Y_{l_3m_3}Y_{l_4m_4}) \\ &= c^k(l_1m_1,l_3m_3)\delta_{m_1-m_3,q}c^k(l_4m_4,l_2m_2)\delta_{m_2-m_4,q}, \end{aligned} \tag{708}$$

where c^k are simple numbers which can be obtained by evaluating the integral

$$c^k(lm,l'm')\delta_{q,m-m'} = \sqrt{\frac{4\pi}{2k+1}}\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta\sin\theta Y^*_{lm}Y_{kq}Y_{l'm'} \tag{709}$$

where the conservation of momentum $q = m - m'$ follows from the integration over φ.

Now suppose we look at one shell and consider diagonal matrix elements only. The radial integral is then always

$$F^k = (R_{nl}R_{nl}|\frac{e^2r_<^k}{r_>^{k+1}}|R_{nl}R_{nl}). \tag{710}$$

Now we are dealing with many-body wavefunction where the particles are indistinguishable. So in reality, we have two different parts in the matrix element

$$\langle nlm\sigma,nlm'\sigma'|V|nlm\sigma,nlm'\sigma'\rangle = (nlm\sigma,nlm'\sigma'|V|nlm\sigma,nlm'\sigma') - (nlm\sigma,nlm'\sigma'|V|nlm'\sigma',nlm\sigma) \tag{711}$$

$$= \int d\mathbf{r}_1\int d\mathbf{r}_2|\psi_{nlm\sigma}(\mathbf{r}_1)|^2V|\psi_{nlm'\sigma'}(\mathbf{r}_1)|^2 \tag{712}$$

$$- \int d\mathbf{r}_1\int d\mathbf{r}_2\psi^*_{nlm\sigma}(\mathbf{r}_2)\psi^*_{nlm'\sigma'}(\mathbf{r}_1)V\psi_{nlm'\sigma'}(\mathbf{r}_2)\psi_{nlm\sigma}(\mathbf{r}_1) \tag{713}$$

$$= \sum_k\left[c^k(lm,lm)c^k(lm',lm') - \delta_{\sigma,\sigma'}(c^k(lm,lm'))^2\right]F^k, \tag{714}$$

where we have introduced the spin $\sigma =\uparrow,\downarrow$ of the electron. The $(\cdots)$ now indicates the evaluation of the matrix element in which the first wavefunction is that of particle 1 and the second that of particle 2. The second term is the exchange term. Note that the electrons can only be exchanged if they have equal spins.

When we have more than one shell, we can have to different radial matrix elements

$$F^k = (R_{nl}R_{n'l'}|\frac{e^2r_<^k}{r_>^{k+1}}|R_{nl}R_{n'l'}) \qquad \text{and} \qquad G^k = (R_{nl}R_{n'l'}|\frac{e^2r_<^k}{r_>^{k+1}}|R_{n'l'}R_{nl}). \tag{715}$$

The matrix element now becomes

$$\begin{aligned}\langle nlm\sigma,n'l'm'\sigma'|V|nlm\sigma,n'l'm'\sigma'\rangle &= (nlm\sigma,n'l'm'\sigma'|V|nlm\sigma,n'l'm'\sigma') - (nlm\sigma,n'l'm'\sigma'|V|n'l'm'\sigma',nlm\sigma) \\ &= \sum_k c^k(lm,lm)c^k(l'm',l'm')F^k - \delta_{\sigma,\sigma'}(c^k(lm,l'm'))^2G^k. \end{aligned} \tag{716}$$

Now let us look at the simplest example: two electrons in an s orbital. We have to know now the coefficients $c^k(00,00) = \delta_{k,0}$ since $lm = 00$. The only non zero k turns out to be $k = 0$. We can see that by noting that the maximum k one can make by adding l and l' is $2l = 0$. If the two electron are in the same s orbital our wavefunction is $|n00\uparrow n00\downarrow\rangle$. The matrix element is

$$\langle n00\uparrow,n00\downarrow|V|n00\uparrow,n00\downarrow\rangle = c^0(00,00)^2F^0_{n0,n0} = F^0_{n0,n0}, \tag{717}$$

where we have used the notation $F^k_{nl,n'l'}$. Where F^0 is the roughly 34 eV that we calculated for the He atom in the previous section. Things become more interesting when we put the electrons in different shells. Again we can take the spins antiparallel, this gives a very similar result

$$\langle n00\uparrow,n'00\downarrow|V|n00\uparrow,n'00\downarrow\rangle = c^0(00,00)^2F^0_{n0,n'0} = F^0_{n0,n'0}. \tag{718}$$

However, when we put the spins parallel, we can also have the exchange term

$$\langle n00\uparrow,n'00\uparrow|V|n00\uparrow,n'00\uparrow\rangle = c^0(00,00)^2F^0_{nl,n'l'} - c^0(00,00)c^0(00,00)G^0_{n0,n'0} = F^0_{n0,n'0} - G^0_{n0,n'0}. \tag{719}$$

Since there is only one way to make a total spin of $M_S = 1$ this must be and eigenenergy of the system with two electrons in two different s orbitals. In this case this is the triplet with $SM_S = 11$. Similarly, we obtain the term with both spins antiparallel. The situation for $M_S = 0$ is more complicated. Above we found one term, however there are more matrix elements:

$$\langle n00\downarrow, n'00\uparrow|V|n00\downarrow, n'00\uparrow\rangle = c^0(00,00)^2 F^0_{n0,n'0} = F^0_{n0,n'0} \tag{720}$$

and also an off-diagonal term

$$\langle n00\uparrow, n'00\downarrow|V|n00\downarrow, n'00\uparrow\rangle = (n00\uparrow, n'00\downarrow|V|n00\downarrow, n'00\uparrow) - (n00\uparrow, n'00\downarrow|V|n'00\uparrow, n00\downarrow) = -G^0_{n0,n'0} \tag{721}$$

Note that the first term on the right-hand side is zero since the integration over the spin-part of the wavefunction is zero. However, for the exchange term, we have a non-zero matrix element. This leads to a 2×2 matrix:

$$\begin{pmatrix} F^0 & -G^0 \\ -G^0 & F^0 \end{pmatrix} \begin{matrix} n00\uparrow, n'00\downarrow \\ n00\downarrow, n'00\uparrow \end{matrix}, \tag{722}$$

leading to eigenvalues $F^0 \pm G^0$. The value $F^0 - G^0$ comes from the $M_S = 0$ term of the triplet. The $F^0 + G^0$ term has only a $M_S = 0$ component (note that this energy does not appear for $M_S = \pm 1$. Note that the sum of the two eigenenergies is $2F^0$, which is equal to the sum of the diagonal elements (the trace of the matrix). This trace is a constant under a transformation to a different basis set. So we would have argues differently. We know that the trace is $2F^0$. We have found an eigenvalue for a triplet $F^0 - G^0$ for $M_S = 1$. However, the triplet should also have a $M_S = 0$ component. Therefore, the other eigenstate for $M_S = 0$ should be $2F^0 - (F^0 - G^0) = F^0 + G^0$.

Let us consider the simple example of two electrons in a p shell. The different electron configurations are shown in Table X D. We use the notation for a configuaration $(m^{\pm}m'^{\pm})$, where the $\pm$ indicates $\sigma = \pm\frac{1}{2}$. For the Coulomb interaction L, M_L, S, and M_S are good quantum numbers. The different terms can be found by looking at the extreme M_L and M_S. The configuration (1^+1^-) belongs to the term with $L = 2$ and $S = 0$. This is denoted by ^1D (say singlet D). The letter indicates the value of L, which corresponds to the notation The different values for l are often denoted by

$$\begin{matrix} L: & 0 & 1 & 2 & 3 & 4 \\ & S & P & D & F & G \end{matrix} \tag{723}$$

The superscript is $2S + 1$. Note that the ^1D term appears in the configurations with $M_L = 2, \cdots, -2$ and $M_S = 0$. This energy level is $(2L+1)(2S+1)$ fold degenerate. The configuration (1^+0^+) belongs to a ^3P term (say triplet P). Then there is only on term unaccounted for with $M_L = M_S = 0$, i.e., a ^1S.

The diagonal matrix elements are given by

$$\sum_k \left[c^k(lm,lm)c^k(l'm',l'm')F^k - \delta_{\sigma\sigma'}(c^k(lm,l'm'))^2 G^k\right], \tag{724}$$

where for p^2, we have $l = l' = 1$. For ^1D, we find

$$E(^1D) = E(1^+1^-) = \sum_k (c^k(11,11))^2 F^k = F^0 + \frac{1}{25}F^2, \tag{725}$$

where the values of c^k are given in Table XIV A. In the case of ^3P, there is also an exchange contribution

$$\begin{aligned} E(^3P) = E(1^+0^+) &= \sum_k \left[c^k(11,11)c^k(10,10)F^k - (c^k(11,10))^2 G^k\right] \\ &= F^0 - \frac{2}{25}F^2 - (0G^0 + \frac{3}{25}G^2) = F^0 - \frac{5}{25}F^2, \end{aligned} \tag{726}$$

TABLE I: *Different configurations for p^2.*

p^2	$M_S = 1$	0	-1
$M_L = 2$		(1^+1^-)	
1	(1^+0^+)	$(1^+0^-)(1^-0^+)$	(1^-0^-)
0	$(1^+ - 1^+)$	$(1^+ - 1^-)(1^- - 1^+)(0^+0^-)$	$(1^- - 1^-)$
-1	$(0^+ - 1^+)$	$(0^+ - 1^-)(0^- - 1^+)$	$(0^- - 1^-)$
-2		$(-1^+ - 1^-)$	

TABLE II: *Values for* $c^k(lm,l'm')$ *for* $l=l'=1$.

m	m'	$c^0(lm,l'm')$	$c^2(lm,l'm')$
±1	±1	1	$-\frac{1}{5}$
±1	0	0	$\frac{\sqrt{3}}{5}$
0	0	1	$\frac{2}{5}$
±1	∓1	0	$-\frac{\sqrt{6}}{5}$

TABLE III: *Different configurations for* p^3.

p^3	$M_S=\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$
$M_L=2$		$(1^+1^-0^+)$	$(1^+1^-0^-)$	
1		$(1^+0^+0^-)(1^+1^- - 1^+)$	$(1^-0^+0^-)(1^+1^- - 1^-)$	
0	$(1^+0^+ - 1^+)$	$(1^-0^+ - 1^+)(1^+0^- - 1^+)(1^+0^+ - 1^-)$	$(1^-0^- - 1^+)(1^-0^+1^-)(1^+0^- - 1^-)$	$(1^-0^- - 1^-)$
-1		$(0^+0^- - 1^+)(1^+ - 1^+ - 1^-)$	$(0^+0^- - 1^-)(1^- - 1^+ - 1^-)$	
-2		$(0^+ - 1^+ - 1^-)$	$(0^- - 1^+ - 1^-)$	

where use has been made of the fact that within one shell $G^k = F^k$. The energy of the ^1S term cannot be calculated this way. However, here we can make use of the fact that the sum of the diagonal energies is unchanged when transforming from our basis set into the LS states,

$$E(1^+-1^-)+E(1^--1^+)+E(0^+0^-) = 2\sum_k c^k(11,11)c^k(1-1,1-1)F^k + \sum_k c^k(10,10)F^k \quad (727)$$

$$= 2(F^0+\frac{1}{25}F^2)+F^0+\frac{4}{25}F^2 \quad (728)$$

$$= E(^1S)+E(^3P)+E(^1D) = E(^1S)+\frac{2}{25}F^0-\frac{4}{25}F^2,$$

which gives $E(^1S)=F^0+\frac{10}{25}F^2$. G enerally, one finds that for electrons in one shell the lowest term is the one with first, the highest spin multiplicity, and, secondly, the highest orbital multiplicity. In the case of p^2, the ^3P term. This can be understood by noting that for the highest spin electrons have to occupy different orbitals, therefore reducing the Coulomb interaction. The same applies, in a more complicated manner for the angular momentum.

Thus we find in general that that

$$^2P \otimes^2 P =^{1,3} D +^{1,3} P +^{1,3} S \quad (729)$$

giving a total degeneracy of $1\times5+3\times5+1\times3+3\times3+1\times1+3\times1=36$, which is equal to 6×6. This would be the case for two electrons in different p shells, i.e. *pp*. However, for p^2, the Pauli exclusion principle states that electrons cannot be in the same orbital. This makes matters a bit more complicated. We can easily see that ^3D is forbidden, since this would require a configuration (1^+1^+). We know that the total number of configurations is 15. So we still have 10 left. The other terms are 3, 9, 3, and 1 fold degenerate. Apparently the only way to make 10 is to take the ^3P and the ^1S configurations, being 9 and 1-fold degenerate.

p^3 *configuration* The procedure described above can be extended for other configurations. Let us consider p^3. The configuration are given in Table XIV A. In total we have

$$\left(\begin{array}{c}6\\3\end{array}\right)=\frac{6\times5\times4}{3\times2}=20 \quad (730)$$

configurations. Let us now see what different LS terms are in these 20 configurations. The configuration $(1^+1^-0^+)$ must belong to an LS term with $L=2$ and $S=\frac{1}{2}$ since there are no configurations with higher L and S that could give rise to this term. This term is denotes as a ^2D. The configuration $(1^+0^+-1^+)$ can only come from an LS term with $S=\frac{3}{2}$ and $L=0$, therefore a ^4S. The first M_L and M_S terms for which we see more than one configuration is $M_L=1$ and $M_S=\frac{1}{2}$. The LS states are linear combinations of the configurations $(1^+0^+0^-)$ and $(1^+-1^+-1^-)$. One of these combinations is part of the ^2D term. The other combination belongs to the LS term with $L=1$ and $S=\frac{1}{2}$, i.e. ^2P. The total degeneracy is $5\times2+1\times4+3\times2=20$ so we have found all the LS terms.

Let us now calculate the eigenenergies. We have calculated the matrix elements between two particles. Now we have three particles. However, the Coulomb interaction is still a two-particle interaction so we need to find all the possible pairs of electrons. Therefore, for the ^2D

$$\begin{aligned}\langle 1^+1^-0^+|V|1^+1^-0^+\rangle &= \langle 1^+1^-|V|1^+1^-\rangle + \langle 1^+0^+|V|1^+0^+\rangle + \langle 1^-0^+|V|1^-0^+\rangle &(731)\\ &= F^0 + \frac{1}{25}F^2 + F^0 - \frac{5}{25}F^2 + F^0 - \frac{2}{25}F^2 = 3F^0 - \frac{6}{25}F^2, &(732)\end{aligned}$$

where results for the p^2 configuration have been used. Note that the 3 in front of the monopole term F^0 just counts the number of different pairs. For the ^4S, we have

$$\begin{aligned}\langle 1^+0^+ - 1^+|V|1^+0^+ - 1^+\rangle &= \langle 1^+0^+|V|1^+0^+\rangle + \langle 1^+ - 1^+|V|1^+ - 1^+\rangle + \langle 0^+ - 1^+|V|0^+ - 1^+\rangle\\ &= F^0 - \frac{5}{25}F^2 + \left(F^0 + \frac{1}{25}F^2 - (0G^0 + \frac{6}{25}G^2)\right) + \left(F^0 - \frac{2}{25}F^2 - (0G^0 + \frac{3}{25}G^2)\right)\\ &= 3F^0 - \frac{15}{25}F^2. &(733)\end{aligned}$$

where we have used $G^k = F^k$ for the same shell since $nl = n'l'$ and therefore the radial wavefunction (which enter in the evaluation of the radial matrix element) are the same.

For the ^2P term, we have to make use again of the diagonal sum rule (or calculate the off-diagonal elements, but this is more complicated). We have two configurations with matrix elements

$$\begin{aligned}\langle 1^+0^+0^-|V|1^+0^+0^-\rangle &= \langle 1^+0^+|V|1^+0^+\rangle + \langle 1^+0^-|V|1^+0^-\rangle + \langle 0^+0^-|V|0^+0^-\rangle &(734)\\ &= F^0 - \frac{5}{25}F^2 + F^0 - \frac{2}{25}F^2 + F^0 + \frac{4}{25}F^2 &(735)\\ &= 3F^0 - \frac{3}{25}F^2 &(736)\end{aligned}$$

and

$$\begin{aligned}\langle 1^+1^- - 1^+|V|1^+1^- - 1^+\rangle &= \langle 1^+1^-|V|1^+1^-\rangle + \langle 1^+ - 1^+|V|1^+ - 1^+\rangle + \langle 1^- - 1^+|V|1^- - 1^+\rangle &(737)\\ &= F^0 + \frac{1}{25}F^2 + F^0 - \frac{5}{25}F^2 + F^0 + \frac{1}{25}F^2 &(738)\\ &= 3F^0 - \frac{3}{25}F^2. &(739)\end{aligned}$$

The fact that the trace is constant in going from our $(m_1^\pm m_2^\pm m_3^\pm)$ basis set to the LS basis, implies that the sum over the diagonal matrix elements is the same, i.e.

$$2(3F^0 - \frac{3}{25}F^2) = 6F^0 - \frac{6}{25}F^2 = E(^2D) + E(^2P) = 3F^0 - \frac{3}{25}F^2 + E(^2P), \quad (740)$$

giving $E(^2P) = 3F^0$.

Therefore, for the p^3 configuration, we find that the ^4S state is the lowest eigenstate, in agreement with Hund's rule that says that S should be maximum. This results directly from the fact that they have to be in different orbitals. Therefore, we like to put the spins parallel in an orbital.

Note also that this whole process of writing done matrices can be done very well by computers. Numerical evaluations of this type are a substantial part of theoretical physics.

E. Addition of angular momenta

We often find it useful to add angular momenta. As we saw with the Coulomb interaction, the total L and S are what they call "good quantum numbers", i.e., all eigenfunctions with the same L and S, but with different M_L and M_S have the same eigenvalues for the Coulomb interaction. Obviously, we would like to be able to create these eigenfunction. We know that the angular part of the eigenfunction can be expressed as spherical harmonics Y_{lm}. So we want to express the wavefunction for L and M (omitting the subscript L)in product of the wavefunctions,

$$|lml'm'\rangle = |lm\rangle|l'm'\rangle, \quad (741)$$

where the $|lm\rangle$ stands for the spherical harmonic Y_{lm}. We now want to couple these to a new total wavefunction, which are eigenstates of the operators

$$\mathbf{L}^2 = L_1^2 + L_2^2 + 2\mathbf{L}_1 \cdot \mathbf{L}_2 \tag{742}$$
$$L_z = L_{1z} + L_{2z}, \tag{743}$$

where L_1 and L_2 are angular momentum operators working on particle 1 and 2, respectively. We can therefore express the wavefunction as

$$|ll'LM\rangle = \sum_{mm'} \langle lml'm'|LM\rangle |lml'm'\rangle. \tag{744}$$

The coefficient $\langle lml'm'|LM\rangle$ is called the Clebsch-Gordan coefficient or Wigner coefficient or vector-coupling coeffcient. And they come in a zoo of different notations

$$\langle lml'm'|LM\rangle, \qquad V(ll'L;mm'M), \qquad C^{LM}_{lm,l'm'}, \tag{745}$$

or as a $3j$ symbol

$$\begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix}. \tag{746}$$

This would be bearable, if they didn't all have different phase factors, and coeffients $\sqrt{2L+1}$. Probably the most elegant are the $3j$ symbols, where the l, l', and L are all equivalent. However, the Clebsch-Gordan coefficients have a longer history and are used very regularly. Analytical expressions for the Clebsch-Gordan coefficients in terms of $ll'Lmm'M$ can be derived, but the derivation is rather complex. However, they are strongly related to the integrals over three spherical harmonics, as we saw in the evaluation of the Coulomb interaction.

The situation simplifies if we look at a particular problem. For two p electrons, we now that the we have an eigenstate $|11\rangle|11\rangle$, since the is no other way to reach $M = 2$. This directly implies that $L = 2$. So the wavefunction is $|1122\rangle = |11\rangle|11\rangle$. For the combined wavefunction we use the notation $|ll'LM\rangle$, for the one-particle wavefunction we have $|lm\rangle$. However, for $M = 1$, we can have $|10\rangle|11\rangle$ and $|11\rangle|10\rangle$, so what are the eigenstates of $\mathbf{L}^2$ and L_z. Here we can make use of the step-down operator

$$L_- = L_{1-} + L_{2-} \tag{747}$$

where

$$L_-|lm\rangle = \sqrt{(l+m)(l-m+1)}|l, m-1\rangle. \tag{748}$$

Therefore,

$$L_-|1122\rangle = \sqrt{(2+2)(2-1+1)}|1121\rangle = 2|1121\rangle = \tag{749}$$
$$= (L_{1-} + L_{2-})|11\rangle|11\rangle = \sqrt{2}(|10\rangle|11\rangle + |11\rangle|10\rangle). \tag{750}$$

This gives

$$|1121\rangle = \frac{1}{\sqrt{2}}(|10\rangle|11\rangle + |11\rangle|10\rangle). \tag{751}$$

This directly gives us the Clebsch-Gordan coefficients $C^{21}_{10,11} = 1/\sqrt{2}$ and $C^{21}_{11,10} = 1/\sqrt{2}$. Since wavefunction are orthonormal, we can directly obtain the other one

$$|1111\rangle = \frac{1}{\sqrt{2}}(|10\rangle|11\rangle - |11\rangle|10\rangle). \tag{752}$$

This must be the wavefunction for $L = 1$ and $M = 1$. Again applying the step-down operator

$$L_-|1121\rangle = \sqrt{(2+1)(2-1+1)}|1120\rangle = \sqrt{6}|1120\rangle \tag{753}$$
$$= (L_{1-} + L_{2-})\frac{1}{\sqrt{2}}(|10\rangle|11\rangle + |11\rangle|10\rangle) = \sqrt{2}\frac{1}{\sqrt{2}}(|1,-1\rangle|11\rangle + 2|10\rangle|10\rangle + |11\rangle|1,-1\rangle). \tag{754}$$

This gives

$$|1120\rangle = \frac{1}{\sqrt{6}}(|1,-1\rangle|11\rangle + 2|10\rangle|10\rangle + |11\rangle|1,-1\rangle). \tag{755}$$

In the case of $L = 1$, the "double product" cancels and we obtain

$$|1110\rangle = \frac{1}{\sqrt{2}}(|1-1\rangle|11\rangle - |11\rangle|1-1\rangle). \tag{756}$$

However, we started out with three basis functions $|1,-1\rangle|11\rangle$, $|10\rangle|10\rangle$, $|11\rangle|1,-1\rangle$. So the must be a third coupled wave function left. Since we did not find this one at higher L, it must be the $L = 0$ term. This one must be orthogonal to the other two, by inspection we see that

$$|1100\rangle = \frac{1}{\sqrt{3}}(|1,-1\rangle|11\rangle - |10\rangle|10\rangle + |11\rangle|1,-1\rangle). \tag{757}$$

So we found the angular momentum part of our p^2 problem. The S, P, and D term. There are also spin terms. The triplet state has as highest component $|\frac{1}{2}\frac{1}{2}11\rangle = |\frac{1}{2}\frac{1}{2}\rangle|\frac{1}{2}\frac{1}{2}\rangle$ Applying the step down operator one can obtain the $M_S = 0$ component

$$|\frac{1}{2}\frac{1}{2}10\rangle = \frac{1}{\sqrt{2}}\left(|\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2}\frac{1}{2}\rangle + |\frac{1}{2}\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\right). \tag{758}$$

From orthogonality we obtain the singlet term

$$|\frac{1}{2}\frac{1}{2}00\rangle = \frac{1}{\sqrt{2}}\left(|\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2}\frac{1}{2}\rangle - |\frac{1}{2}\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle\right). \tag{759}$$

The third term of the triplet is given by $|\frac{1}{2}\frac{1}{2}1,-1\rangle = |\frac{1}{2},-\frac{1}{2}\rangle|\frac{1}{2},-\frac{1}{2}\rangle$.

XI. SPIN-ORBIT COUPLING

So again we find that $L_1 = 1$ and $L_2 = 1$ add up to a total $L = 2, 1, 0$. Note that starting from a certain L value we can now apply a perturbation without really destroying the symmetry. The simplest perturbation is probably a magnetic field, which is proportional to $\mathbf{B}\cdot\mathbf{L}$. If $\mathbf{B}$ is along the z-axis. then this becomes $B_z M$. So for example the $L = 1$ state, the M valueswould be no longer degenerate but would split in $M = -1, 0, 1$. For the Hydrogen atom, we found that further level splitting is caused by the spin-orbit coupling. Another perturbance is the spin-orbit coupling, which behaves as $\mathbf{L}\cdot\mathbf{S}$. This can be written differently as

$$\mathbf{L}\cdot\mathbf{S} = L_x S_x + L_y S_y + L_z S_z = \frac{1}{2}(L_+ S_- + L_- S_+) + L_z S_z. \tag{760}$$

Now we see that L and S are no longer conserved, but the total L and S is. Every time L goes up by one, the S value goes down, and vices versa. In the case where the spin-orbit coupling is much smaller than the Coulomb interaction, we can treat the spin-coupling as a perturbation. This is true for the valence shell. However, for deep lying core level (high binding energy), the spin-orbit coupling can be much larger than the Coulomb interaction. Let suppose we can apply perturbation theory. In this case one uses the LS coupled wavefunction as the unperturbed states. One assumes that the splitting between the energies of the LS states is much larger than the spin-orbit coupling, so that these states do not couple. In lowest-order perturbation, the spin-orbit coupling is given by

$$\langle LS|\mathbf{L}\cdot\mathbf{S}|LS\rangle. \tag{761}$$

Now this is not a very correct expression since the M_L and M_S values are missing. Since we assume that the spin-orbit coupling is small, L and S are still good quantum numbers. However, M_L and M_S are not (note that all the M_L and M_S states of one LS term are degenerate). However, L and S are not good quantum numbers for the spin-orbit coupling. The corresponding $J = |L-S|, \ldots, L+S$ values are. Therefore, we can write the lowest order perturbation of the spin-orbit coupling as

$$\langle LSJM_J|\mathbf{L}\cdot\mathbf{S}|LSJM_J\rangle = \frac{1}{2}[J(J+1) - L(L+1) - S(S+1)]. \tag{762}$$

Let us consider as an example the p^2 configuration. The LS terms are ${}^1D$, ${}^3P$, and ${}^1S$. It is common to put the J values as a subscript, i.e. ${}^1D_2$, ${}^3P_{0,1,2}$, and ${}^1S_0$. Note that the total degeneracy (the sum of all the $2J+1$ values) should remain the same, i.e. $5+1+3+5+1=15$. Hund's rule now says that for less than half-filled shells, the J value with the lowest energy is the one where L and S are coupled to $L-S$. For more than half filled shells, $L+S$ is the lowest. Therefore, the ground state for p^2 is ${}^3P_0$. For p^4 the LS term are exactly the same, but we the lowest state is ${}^3P_2$.

XII. MOLECULES AND SOLIDS

After spending already a considerable time in the world of atomic physics, we would like to know what happens if we start brining the atoms together to create molcules and solids (we will not be discussing liquids here). Let us first have a look at a simple two-atom molecule.

A. Hydrogen molecule

Let us first write the general Hamiltonian for a molecule as a matrix. Suppose we have an eigenstate $|\psi\rangle$ which we want to write in a certain basis set (atomic orbitals is a very obvious choice, let us denote these as $|\varphi_n\rangle$. Note, that for a molecule the atomic orbitals can be centered around different nuclei. Let us assume our basis set consists of N atomic-like orbitals).

$$|\psi\rangle = \sum_n a_n |\varphi_n\rangle. \tag{763}$$

Since $|\psi\rangle$ is an eigenstate, we have $H|\psi\rangle = E|\psi\rangle$. For the n''th component, we then find

$$\langle\varphi_{n'}|H|\psi\rangle = E\langle\varphi_{n'}|\psi\rangle \qquad \Rightarrow \qquad \sum_n \langle\varphi_{n'}|H|\varphi_n\rangle a_n = E\sum_{n'}\langle\varphi_{n'}|\varphi_n\rangle a_n. \tag{764}$$

This can also be written as a matrix

$$\begin{pmatrix} H_{11} & \cdots & H_{1N} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ H_{N1} & \cdots & H_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ a_N \end{pmatrix} = E \begin{pmatrix} S_{11} & \cdots & S_{1N} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_{N1} & \cdots & S_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ a_N \end{pmatrix} \qquad \text{or} \qquad H|\psi\rangle = ES|\psi\rangle, \tag{765}$$

with $H_{n'n} = \langle\varphi_{n'}|H|\varphi_n\rangle$. We have also introduced here the overlap integral. Note that on an atom the wavefunction are orthogonal. However, when bringing two atoms together, their wavefunction do not necessarily have zero overlap. To take that into account we have $S_{ij} = \langle\varphi_i|\varphi_j\rangle$, not that $S_{ii} = 1$. (Later, we will assume, for simplicity S to be the unit matrix). Now let us consider the simplest case with two atoms where only one orbital is involved (say, for Hydrogen the $1s$ orbital). Our Hamiltonian then becomes

$$\begin{pmatrix} \varepsilon & T \\ T & \varepsilon \end{pmatrix} |\psi\rangle = E \begin{pmatrix} 1 & S \\ S & 1 \end{pmatrix} |\psi\rangle. \tag{766}$$

The eigenvalues can be found by solving the determinant

$$\begin{vmatrix} \varepsilon - E & T - ES \\ T - ES & \varepsilon - E \end{vmatrix} = (\varepsilon - E)^2 - (T - ES)^2 = 0 \tag{767}$$

This gives two eigenvalues

$$\varepsilon - E = \mp(T - ES) \qquad \Rightarrow \qquad E_{1,2} = \frac{\varepsilon \pm T}{1 \pm S}. \tag{768}$$

The calculation could be a simplified a lot by taking ε and S zero, giving two eigenvalues $\pm T$. For $T < 0$ this gives eigenstates $|\psi_B\rangle = \frac{1}{\sqrt{2}}(|\varphi^1_{1s}\rangle + |\varphi^2_{1s}\rangle)$ for $E = -|T|$ and $|\psi_A\rangle = \frac{1}{\sqrt{2}}(|\varphi^1_{1s}\rangle - |\varphi^2_{1s}\rangle)$ for $E = |T|$ (the superscript denotes the two atoms). These are known as bonding and antibonding wavefunctions. Note that in the bonding combination, the electron wavefunction spreads over both atoms of the molecule. This relaxation causes a lowering of the kinetic

energy. Actually we could have guesses these eigenstates beforehand since $|\varphi^1_{1s}\rangle$ and $|\varphi^2_{1s}\rangle$ have to play equivalent roles in the wavefunction. There is a symmetry axis at the center of the molecule.

So far we have considered the situation with one electron, say a H_2^+ molecule. However, a H_2 molecule has two electrons. If the electrons did not interact with each other, the energy would be simply $-2|T|$. However, as we saw in the previous sections, electrons do interact with each other. Now let us take $S = 0$ and put two electrons in the orbitals. When taking into account the spin of the electron, we find six different combinations: $|\psi_B \uparrow \psi_B \downarrow\rangle$, $|\psi_A \uparrow \psi_B \downarrow\rangle$, $|\psi_B \uparrow \psi_A \downarrow\rangle$, $|\psi_A \uparrow \psi_B \uparrow\rangle$, $|\psi_B \downarrow \psi_A \downarrow\rangle$, and $|\psi_A \uparrow \psi_A \downarrow\rangle$ at energies $2\varepsilon - 2T, 2\varepsilon, 2\varepsilon, 2\varepsilon, 2\varepsilon, 2\varepsilon + 2T$, respectively. Let us look at one particular combination

$$|\psi_B \uparrow \psi_B \downarrow\rangle = \frac{1}{2}\left(|\varphi^1_{1s\uparrow}\varphi^1_{1s\downarrow}\rangle + |\varphi^1_{1s\uparrow}\varphi^2_{1s\downarrow}\rangle + |\varphi^2_{1s\uparrow}\varphi^1_{1s\downarrow}\rangle + |\varphi^2_{1s\uparrow}\varphi^2_{1s\downarrow}\rangle\right). \tag{769}$$

We see that the electrons spend just as much time on the same atom as they to on different atoms. This is quite a bit different from the situation were the two Hydrogen atoms were far apart. In that case the eigenstates would be all degenerate and given by $|\varphi^1_{1s\sigma}\varphi^2_{1s\sigma'}\rangle$ with $\sigma, \sigma' = \uparrow, \downarrow$ with energy twice the binding energy of the $1s$ orbital, i.e. $2\varepsilon = -2 \times 13.6$ eV. In the remainder, to simplify the expressions, let us measure the energy with respect to 2ε.

The wavefunctions with two spins parallel are different. For example,

$$|\psi_B \uparrow \psi_A \uparrow\rangle = |\varphi^1_{1s\uparrow}\varphi^2_{1s\uparrow}\rangle. \tag{770}$$

Here we see that the Hydrogen atoms cannot bind to each other because of Pauli's principle. This looks then like the situation of two separate Hydrogen atoms.

Now in the previous sections, we saw a lot about the Coulomb interaction between electrons. That would lead you to think that maybe the electrons do not like to be on the same atom (of course, there is also a Coulomb interaction between the electrons when they are not on the same site, but this is much smaller). Let us suppose that this Coulomb interaction can be written as

$$\langle\varphi^i_{1s\sigma}\varphi^j_{1s\sigma'}|\frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}|\varphi^i_{1s\sigma}\varphi^j_{1s\sigma'}\rangle = U\delta_{i,j}, \tag{771}$$

i.e. it is U when the electrons are on the same atom and zero otherwise. (Note that we do not have to deal with all different kinds of LS terms, since for s^2 there is only one: ^1S). This means that the configuration $|\psi_B \uparrow \psi_A \uparrow\rangle$ is at energy 2ε. So the hopping does not enter into the problem. Let us for simplicity take $\varepsilon = 0$ (we can always shift our absolute energy scale). The question now rises, in the situation where the spins are opposite, can we still gain something from the hopping? You might think that there is no reason anymore for the electrons to be at the same atom because of this Coulomb repulsion which can be quite large. However, this is not true.

To solve this problem let us start again with states of the Hydrogen atom (we can in principle also start from the bonding and anitbonding states, but it somewhat more cumbersome). Let us simplify the notation somewhat since we know we are only dealing with Hydrogen atoms. Let us write $|i\sigma j\sigma'\rangle = |\varphi^i_{1s\sigma}\varphi^j_{1s\sigma'}\rangle$. We can now write the Hamiltonian in matrix form as

$$H = \begin{pmatrix} U & T & T & 0 \\ T & 0 & 0 & T \\ T & 0 & 0 & T \\ 0 & T & T & U \end{pmatrix} \begin{matrix} |1\uparrow 1\downarrow\rangle \\ |1\uparrow 2\downarrow\rangle \\ |2\uparrow 1\downarrow\rangle \\ |2\uparrow 2\downarrow\rangle \end{matrix}. \tag{772}$$

We can in principle solve this 4×4 Hamiltonian on the computer, we which give us the eigenvalues.

However, we can get a bit further by making clever combination. Note that atom 1 and 2 should be entirely equivalent. So maybe we get a bit further by taking linear combinations where this condition is satisfied:

$$|0_\pm\rangle = \frac{1}{\sqrt{2}}\left(|1\uparrow 2\downarrow\rangle \pm |2\uparrow 1\downarrow\rangle\right) \tag{773}$$

$$|U_\pm\rangle = \frac{1}{\sqrt{2}}\left(|1\uparrow 1\downarrow\rangle \pm |2\uparrow 2\downarrow\rangle\right), \tag{774}$$

where the diagonal elements are 0 and U, respectively.

The matrix elements for H now greatly simplifies

$$\langle 0_\pm|H|U_\pm\rangle = \frac{1}{2}\left(\langle 1\uparrow 2\downarrow| \pm \langle 2\uparrow 1\downarrow|\right) H \left(|1\uparrow 1\downarrow\rangle \pm |2\uparrow 2\downarrow\rangle\right). \tag{775}$$

This matrix elements is $\langle 0_+|H|U_+\rangle = 2T$ and zero otherwise. In our new basis set, we can write our Hamiltonian as

$$H = \begin{pmatrix} 0 & 2T & 0 & 0 \\ 2T & U & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & U \end{pmatrix} \begin{matrix} |0_+\rangle \\ |U_+\rangle \\ |0_-\rangle \\ |U_-\rangle \end{matrix}. \tag{776}$$

This can easily be solved by noting that it consist of two 1×1 matrices giving eigenvalues 0 and U and a 2×2 matrix with eigenvalues

$$\begin{vmatrix} -E & 2T \\ 2T & U-E \end{vmatrix} = E^2 - UE - 4T^2 \qquad \Rightarrow \qquad E_\pm = \frac{U}{2} \pm \frac{1}{2}\sqrt{U^2 + 16T^2}. \tag{777}$$

Note that E_- is less than zero, implying that in the lowest eigenstate (the ground state) the electrons are something on the same atom. This costs them the Coulomb energy, but they gain kinetic energy.

It is always instructive to look at certain limits. In the case that $U = 0$, we find $E_\pm = \pm 2T$. This is what we already found before by assuming the particles to have no Coulomb interaction. This corresponds two having the electrons in the bonding $(-2T)$ and in the antibonding $(2T)$ orbitals. The limit $T = 0$ gives eigenvalues $0, U$. All the eigenvalues are now $0, 0, 0, 0, U, U$. This corresponds exactly to the possibilities of putting the electrons on diffrent and the same atom. Let us now take $U \ll t$. We can then expand the eigenenergies with a Taylor series

$$E_- = -\frac{4T^2}{U} \qquad \text{and} \qquad E_+ = U + \frac{4T^2}{U}. \tag{778}$$

In the case for parallel spins, we found an eigenenergy 0. For antiparallel spins, we find a lowest energy of $-2T^2/U$. Since U is large the electrons spend most of their time at separate atoms. However, the can still lower their energy by having their spins antiparallel so that they can undertake excursion to the neighboring atom, thereby lowering their energy.

B. Other molecules

In the same way as two s orbitals bind with each other, two p orbitals can also form bonding and antibonding combinations. However, here we have to distinguish two different possibilities. Let us take the z axis to be the axis of the diatomic molecule. The p_z orbitals then have a large overlap with each other, and form so-called σ bonding orbitals. The p_x and p_y orbitals are farther apart and form π molecular orbitals. In both cases, we have bonding and antibonding orbitals. Note that the energy difference between π and π^* is smaller than that between σ and σ^*. This would all be the case for one electron in an molecular orbital. However, when discussing diatomic molecules one has to take into account the Coulomb interaction when two electrons are in the same orbital. A different notation that you often find come from the terms *gerade* (even) and *ungerade* (odd). This is defined with respect to the center of symmetry of the molecule. Therefore, the bonding combination of two s orbitals is *gerade* and the antibonding is *ungerade*. For p_z orbitals the situation is the same, but for the bonding and antibonding of the p_x and p_y, the situation is reverse. "Experimentally", we find the order of molecular orbitals in diatomic molecules

$$1\sigma_g(1s) \to 1s \tag{779}$$
$$1\sigma_u^*(1s) \to 2p_z \tag{780}$$
$$2\sigma_g(2s) \to 2s \tag{781}$$
$$2\sigma_u^*(2s) \to 3p_z \tag{782}$$
$$1\pi_{xu}(2p_x) \qquad 1\pi_{yu}(2p_y) \to 2p_{x,y} \tag{783}$$
$$3\sigma_g(2p_z) \to 3s \tag{784}$$
$$1\pi_{xg}^*(2p_x) \qquad 1\pi_{yg}^*(2p_y) \to 3d \tag{785}$$
$$3\sigma_u^*(2p_z) \to 4p_z \tag{786}$$
$$\tag{787}$$

(note that there is a slight redundancy in the notation here, often the asterisk is left out of the notation). Between parantheses the atomic orbitals from which the molecular orbitals originate are given. Now this looks a bit different from what you would expect from the overlap. Since the hopping between the p_z orbitals is larger than that between the p_x and p_y orbitals, you would expect the $3\sigma_g$ to be lower in energy than the $1\pi_{xu}$ and $1\pi_{yu}$ molecular orbitals.

However, for $3\sigma_g^2$, both electrons are between the two atoms and have a large Coulomb interaction. For the $1\pi_{xg}^*$ and $1\pi_{yg}^*$ molecular orbitals the electrons are farther apart. In addition, if we have only two electrons to put in the 1π orbitals, we can put one in the $1\pi_{xg}^*$ and the other in the $1\pi_{xu}^*$. This will lead to a further reduction in the Coulomb interaction. We already considers the situation of H_2, which has a dissociation energy of 4.75 eV (2 bonding electrons).

Let us consider some diatomic molecules for different elements:

1. He_2, $(1\sigma_g)^2(1\sigma_u)^2$. Both $1s$ orbitals are full, so effectively no bonding occurs.

2. Li_2 $KK(2\sigma_g)^2$. The two $1s$ (or K) shells are complete and do not contribute in the bonding. The bonding is done by the $2s$ electrons forming a stable species. Dissociation energy 1.14 eV (two bonding electrons).

3. Be_2, $KK(2\sigma_g)^2(2\sigma_u)^2$. Both $2s$ orbitals are full, so effectively no bonding occurs. The situation is analogous to He.

4. B_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^2$. We have two degenerate orbitals that contribute to the bonding $1\pi_{xu}$ and $1\pi_{yu}$. In order to minimize the Coulomb interaction, a configuration of $1\pi_{xu\uparrow}1\pi_{yu\uparrow}$ is formed (i.e. a triplet is formed, in total three combinations). Dissociation energy 3.0 eV (two bonding electrons).

5. C_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4$. With a total of six bonding and two antibonding states, bonding occurs. However, Carbon forms stabler bondings with Hydrogen, Oxygen, etc. Dissociation energy 6.24 eV (four bonding electrons).

6. N_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2$. With a total of eight bonding and two antibonding states, a very stable bonding occurs. N_2 has the largest dissociation energy of he molecules in this list. Dissociation energy 9.76 eV (six bonding electrons).

7. O_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2(1\pi_g)^2$. Eight bonding electrons and four antibonding electrons. Again we have two electrons in two degenerate $1\pi_g$ orbitals. These will again form a triplet as in the case of B_2. The oxygen molecule is therefore paramagnetic, which was one of the great triumph of the molecular orbital theory. Dissociation energy 5.12 eV (four bonding electrons).

8. F_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2(1\pi_g)^4$. Dissociation energy 1.60 eV (two bonding electrons).

9. Ne_2, $KK(2\sigma_g)^2(2\sigma_u)^2(1\pi_u)^4(3\sigma_g)^2(1\pi_g)^4(3\sigma_u)^2$. Unstable since all the $2s$ and $2p$ orbitals are full.

C. Hybrids

The configuration of the Carbon atom is given by $(1s)^2(2s)^2(2p_x)^1(2p_y)^1$. Therefore, Carbon would have two unpaired electrons which are at an angle of 90° with each other. However, this is not what is observed experimentally in many Carbon compounds such as methane CH_4. Here Carbon appears to be tetravalent. This can be seen if we represent the configuration of Carbon somewhat differently $(1s)^2(2s)^1(2p_x)^1(2p_y)^1(2p_z)^1$. This treats the s and p orbitals on a similar footing. Now we can consider the mixing of s and p orbitals.

sp hybrids Let us consider the formation of hybrid functions composed of one s and one p_z function. The two hybrid wavefunction can be expressed as

$$\phi_1 = a_1\psi_s + a_2\psi_{p_z} \qquad \text{and} \qquad \phi_2 = b_1\psi_s + b_2\psi_{p_z}. \tag{788}$$

The coefficients are determined such that the wavefunctions are orthogonal and normalized

$$\int \psi_1^*\psi_1 d\mathbf{r} = 1 \qquad \text{and} \qquad \int \psi_2^*\psi_2 d\mathbf{r} = 1 \qquad \text{and} \qquad \int \psi_1^*\psi_2 d\mathbf{r} = 0 \tag{789}$$

We want that the new wavefunctions contain both an equal amount of s character. Combining all the equations leads to the solution

$$\phi_1 = \frac{1}{\sqrt{2}}(\psi_s + \psi_{p_z}) \qquad \text{and} \qquad \phi_2 = \frac{1}{\sqrt{2}}(\psi_s - \psi_{p_z}). \tag{790}$$

An example of an *sp* hybrid is acetylene (C_2H_2). Here the two *sp* hybrids of the Carbon atoms form a σ bonding molecular orbital (if we take the z axis along the molecule, the *sp* hybrid is formed by the s and p_z orbitals. This leaves the p_x and p_y orbitals which are perpendicular to the z axis. These orbitals of the two Carbon atoms form π bonding molecular orbitals. This creates a triple bond.

sp^2 hybrids

Following the same procedure as described above, we can find three new wavefunction consisting of s, p_x, and p_y

orbitals:

$$\psi_1 = \frac{1}{\sqrt{3}}\psi_s + \sqrt{\frac{2}{3}}\psi_{p_x} \tag{791}$$

$$\psi_2 = \frac{1}{\sqrt{3}}\psi_s - \frac{1}{\sqrt{6}}\psi_{p_x} + \frac{1}{\sqrt{2}}\psi_{p_y} \tag{792}$$

$$\psi_2 = \frac{1}{\sqrt{3}}\psi_s - \frac{1}{\sqrt{6}}\psi_{p_x} - \frac{1}{\sqrt{2}}\psi_{p_y}. \tag{793}$$

Note that every orbital has $1/3$ s orbital and $2/3$ p orbital character. The first hybrid has on the positive x side both the positive s orbital and the positive part of the p orbital. There the wave functions add up. On the negative x side, we have a positive s orbital and the negative lobe of the p orbital. There the wavefunction partially cancel each other. For the second hydrid, the situation is very similar. Note that the wave function can be written as

$$\psi_2 = \frac{1}{\sqrt{3}}\psi_s + \sqrt{\frac{2}{3}}\left\{-\frac{1}{2}\psi_{p_x} + \frac{1}{2}\sqrt{3}\psi_{p_y}\right\} \tag{794}$$

$$= \frac{1}{\sqrt{3}}\psi_s + \sqrt{\frac{2}{3}}\left\{\cos 120^\circ \psi_{p_x} + \sin 120^\circ \psi_{p_y}\right\} = \frac{1}{\sqrt{3}}\psi_s + \sqrt{\frac{2}{3}}\psi_{p_{120^\circ}}. \tag{795}$$

The combination of p_x and p_y orbtital is effectively a p orbital at 120° degrees with respect to the x-axis. This sp^2 hybrid is therefore equivalent to the first hybrid except it is rotated by 120° around the z-axis.

A typical example of an sp^2 hybrid is Ethylene (C_2H_4). One $2s$ and two $2p$ orbitals form three sp^2 hybrids at a 120° degree angle with each other. Two of those couple two the Hydrogen atom, the other connects the two Carbon atoms, forming a planar molecule. The remaining p is perpendicular to the plane of the molecule (the p_z orbital). The two p_z orbitals of the Carbon atoms for a π bonding with each other.

sp^3 hybrids In an sp^3 hybrid, all the orbitals assume equivalent roles, this leads to wavefunction of the type

$$\psi_1 = \frac{1}{2}(\psi_s + \psi_{p_x} + \psi_{p_y} + \psi_{p_z}) \tag{796}$$

$$\psi_2 = \frac{1}{2}(\psi_s + \psi_{p_x} - \psi_{p_y} - \psi_{p_z}) \tag{797}$$

$$\psi_3 = \frac{1}{2}(\psi_s - \psi_{p_x} + \psi_{p_y} - \psi_{p_z}) \tag{798}$$

$$\psi_4 = \frac{1}{2}(\psi_s - \psi_{p_x} - \psi_{p_y} + \psi_{p_z}). \tag{799}$$

The orbitals form a tertrahedron and have mutual angles of $109^\circ.28'$. One of the simplest examples of an sp^3 hybrid can be found in Methane (CH_4). The four single bonds form σ bonding with the Hydrogen s orbitals. A somewhat more complicated example is Ethane (C_2H_6). Four sp^3 orbitals are formed at each C atom. One of these connects the two Carbon atoms, the other three connect to the Hydrogen s orbitals.

D. The Hückel molecular orbital method

It has long been clear that numerous chemical and spectroscopic proporties of hydrocarbons involve primarily π electrons. Apparently, the energy spearation between the π and σ orbitals is large, so we can consider them spearately. For ethylene, we find a typical bonding-antibonding solution for the π molecular bond. If we take $\varepsilon_0 = \langle p_{zi}|H|p_{zi}\rangle$ where $i = 1, 2$ and $T = \langle p_{z1}|H|p_{z2}\rangle$, we have the eigenvalue problem

$$\begin{vmatrix} E_0 - E & T \\ T & E_0 - E \end{vmatrix} = 0 \qquad \Rightarrow \qquad E = E_0 \pm T. \tag{800}$$

For Ethylene, there are a total of 12 electrons (8 from the two Carbon atoms, 4 from the Hydrogen atoms). All the σ bonding molecular orbitals are occupied. This accounts for ten of the electrons (8 will go into the bonding orbitals between the Carbon and the Hydrogen, the other two will go into the σ binding orbital between the two Carbons). This leaves two electrons, that will go into the bonding π molecular orbital. The π^* orbital is empty.

Let us consider a somewhat more complicated molecule, such as butadiene (C_4H_6), whose chemical formula is

$$\begin{array}{ccccccccccc} & & H & & H & & H & & H & & \\ & & | & & | & & | & & | & & \\ H & - & C_1 & = & C_2 & - & C_3 & = & C_4 & - & H \end{array} \tag{801}$$

There are a total of 22 electrons. The stongly-bound σ bonds are: 6 between the Carbon and the Hydrogen, 3 between the different Carbons. This accounts for $(6+3)\times 2 = 18$ of the electrons. So the four other electrons go into the more losely bound π molecular orbital states. The π electrons can be assumed to be delocalized. Let us take the $T = \langle p_{zi}|H|p_{zj}\rangle < 0$, where i and j are nearest neighbors. The eigenvalue problem can then be written as

$$\begin{vmatrix} E_0-E & T & 0 & 0 \\ T & E_0-E & T & 0 \\ 0 & T & E_0-E & T \\ 0 & 0 & T & E_0-E \end{vmatrix}. \tag{802}$$

The roots can be found to be

$$E = E_0 \pm 1.6180T, \qquad E_0 \pm 0.6180T. \tag{803}$$

The four π electrons go into the lowest two orbitals, giving a total energy of

$$E = 2(E_0 + 1.6180T) + 2(E_0 + 0.6180T) = 4E_0 + 4.4720T. \tag{804}$$

Now let us assume that the π orbitals are not delocalized but localized between Carbon atoms 1,2 and 3,4. This leads to a determinant

$$\begin{vmatrix} E_0-E & T & 0 & 0 \\ T & E_0-E & 0 & 0 \\ 0 & 0 & E_0-E & T \\ 0 & 0 & T & E_0-E \end{vmatrix}. \tag{805}$$

This is easily solved since it splits into two 2×2 matrices. giving eigenvalues $E = E_0+T, E_0+T, E_0-T, E_0-T$. The total energy is then $E = 4E_0 + 4T$, which is higher than for the delocalized case.

Benzene The Benzene molecule (C_6H_6) consists of 6 Carbon atom that form a circle by the sp^2 bonds. There are a total of $6\times 4+6 = 30$ electrons. There are 6 σ bonds of the sp^2 hybrid between the different Carbon atoms. There are another six σ bonding orbitals between Carbon and Hydrogen. This accounts for $(6+6)\times 2 = 24$ of the electrons. The other 6 electrons delocalize in the p_z orbitals of the 6 Carbon atoms. The secular matrix is given by

$$\begin{vmatrix} E_0-E & T & 0 & 0 & 0 & T \\ T & E_0-E & T & 0 & 0 & 0 \\ 0 & T & E_0-E & T & 0 & 0 \\ 0 & 0 & T & E_0-E & T & 0 \\ 0 & 0 & 0 & T & E_0-E & T \\ T & 0 & 0 & 0 & T & E_0-E \end{vmatrix}. \tag{806}$$

The eigenvalues are found to be

$$E = E_0+2T, E_0+T, E_0+T, E_0-T, E_0-T, E_0-2T. \tag{807}$$

The 6 π electrons go into the lowest molecular orbitals giving a total energy of $E = 6E_0 + 8T$.

The wavefunction for the eigenstate with energy E_0+2T is completely symmetric

$$\psi_1 = \frac{1}{\sqrt{6}}(\psi_{p_{1z}} + \psi_{p_{2z}} + \psi_{p_{3z}} + \psi_{p_{4z}} + \psi_{p_{5z}} + \psi_{p_{6z}}), \tag{808}$$

and contains two electrons. The two wavefunctions at energy E_0+T are

$$\psi_2 = \frac{1}{\sqrt{3}}(\frac{1}{2}\psi_{p_{1z}} - \frac{1}{2}\psi_{p_{2z}} - \psi_{p_{3z}} - \frac{1}{2}\psi_{p_{4z}} + \frac{1}{2}\psi_{p_{5z}} + \psi_{p_{6z}}), \tag{809}$$

and

$$\psi_3 = \frac{1}{2}(\psi_{p_{1z}} + \psi_{p_{2z}} - \psi_{p_{4z}} - \psi_{p_{5z}}). \tag{810}$$

These two wavefunctions contain the other four electrons. The wavefunctions for E_0-T are

$$\psi_4 = \frac{1}{\sqrt{3}}(\frac{1}{2}\psi_{p_{1z}} + \frac{1}{2}\psi_{p_{2z}} - \psi_{p_{3z}} + \frac{1}{2}\psi_{p_{4z}} + \frac{1}{2}\psi_{p_{5z}} - \psi_{p_{6z}}), \tag{811}$$

and

$$\psi_5 = \frac{1}{2}(\psi_{p_{1z}} - \psi_{p_{2z}} + \psi_{p_{4z}} - \psi_{p_{5z}}). \tag{812}$$

Finally the wavefunction at energy $E_0 - 2T$ is given by

$$\psi_6 = \frac{1}{\sqrt{6}}(\psi_{p_{1z}} - \psi_{p_{2z}} + \psi_{p_{3z}} - \psi_{p_{4z}} + \psi_{p_{5z}} - \psi_{p_{6z}}). \tag{813}$$

The last three orbitals are empty. Although at first sight this might seem like some coefficients which just happen to come out of a calculation, on further inspection you can see that there is some system to the coefficients. A rather trivial one is

$$\psi_1 = \frac{1}{\sqrt{6}} \sum_n \psi_{p_{nz}}. \tag{814}$$

But ψ_6 we can write as

$$\psi_6 = \frac{1}{\sqrt{6}} \sum_n \psi_{p_{nz}} \cos n\pi. \tag{815}$$

Looking more carefulle at the wavefunctions, we see that

$$\psi_2 = \frac{1}{\sqrt{3}} \sum_n \psi_{p_{nz}} \cos n\frac{\pi}{3} \tag{816}$$

and

$$\psi_3 = \frac{1}{\sqrt{3}} \sum_n \psi_{p_{nz}} \sin n\frac{\pi}{3}. \tag{817}$$

The same thing we can do for ψ_4 and ψ_5, but now using $2\pi/3$. In addition, since ψ_2 and ψ_3 are degenerate, we can make linear combinations of them that are still eigenfunctions

$$\psi^+_{2,3} = \frac{1}{\sqrt{2}}(\psi_2 + i\psi_3) = \frac{1}{\sqrt{6}} \sum_n \psi_{p_{nz}} e^{in\frac{\pi}{3}}, \tag{818}$$

and

$$\psi^+_{2,3} = \frac{1}{\sqrt{2}}(\psi_2 - i\psi_3) = \frac{1}{\sqrt{6}} \sum_n \psi_{p_{nz}} e^{-in\frac{\pi}{3}}. \tag{819}$$

The same can be done for ψ_4 and ψ_5. So it appears that the eigenstates look like a Fourier transform of the states in real space

$$\psi_k = \frac{1}{\sqrt{6}} \sum_n \psi_{p_{nz}} e^{-ikn}, \qquad k = 0, \pm\frac{\pi}{3}, \pm\frac{2\pi}{3}, \pi. \tag{820}$$

In fact, even the eigenvalues can be written as $E_k = E_0 + 2T\cos k$. Is this surprising? Of course not entirely since we have a repeating function, and we could have expected that making a Fourier transform would yield some results. This is a very important concept when going to solids.

E. Solids

Let us try to extend our molecule into a chain [which one could see as a polymer, or a one-dimensional solid, or even a three-dimensional solid in certain cases (note that, in a cubic solid, the p_z orbital has only overlap with a neigboring p_z orbital along the z-axis)]. Our Hamiltonian will then look like (in its simplest form)

$$H = \begin{pmatrix} 0 & T & & & & \\ T & 0 & T & & & \\ & T & 0 & T & & \\ & & T & 0 & T & \\ & & & \cdot & \cdot & \cdot \end{pmatrix}. \tag{821}$$

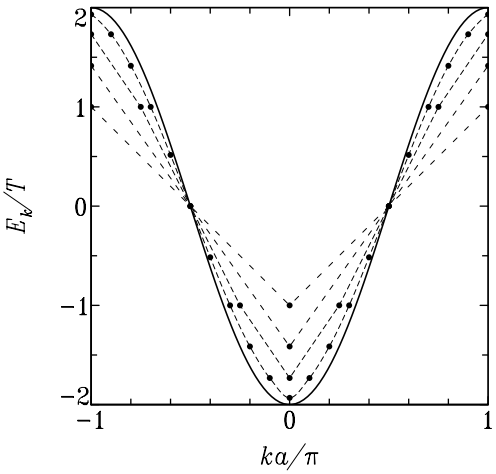


FIG. 13: Eigenvalues for a linear chain of atoms. For periodic boundary conditions, eigenvalues always lie on the $2T\cos ka$ curve. For fixed boundary conditions, eigenvalues are shown for $N = 2, 3, 5, 11$. The dashed lines are drawn as a guide to the eye.

The solutions can be easily found by means of some computer program. Let us just plot the eigenvalues equidistant on a scale from 0 to 1, see Fig. 13 (the same eigenvalues are plotted from -1 to 0). For two atoms, we had $\pm T$. So we plot T at 0 and $-T$ at 1 (note that $T < 0$). For three atoms, we have eigenvalues $\sqrt{2}T, 0, -\sqrt{2}T$, which we plot at $0, \frac{1}{2}, 1$. Also plotted in Fig. 13 are the eigenvalues for 5 and 11 atoms. However, quite quickly the maximum eigenvalues approach $\pm 2T$. Very quickly, the eigenvalues seem to behave a bit like a cosine. Already for 11 atoms, the behavior is quite close to an inifinite number of atoms. Apparently, we are missing something in our approach by connecting all the atoms together. It looks like we should be able to solve the problem more easily.

The input that we are missing here is translational symmetry. Suppose the atom are a distance a apart. In a molecule not all the atoms are equivalent. Although we see from the eigenvalues that if we increase the length of the molecule the eigenvalues follow more or less the same curve (except that we find more and more points of the curve). Things become easier when we connect the first and the last atom with each other. This means that if the chain is N atoms long, then $N+1$ is equivalent to 1. So an electron that tries to hop from the N'th atom to $(N+1)$'th atom ends up at the first atom. This makes all the atoms in the chain equivalent, since on every atom we can go in both directions. This is called a periodic boundary condition (or Born-Von Karman boundary condition). For a linear chain we might see this as a kind of a circle, but we can also apply periodic boundary conditions to two and three dimensional systems, where such pictures are a bit more difficult to imagine. Now are boundary condition becomes different. We can use a complex Fourier transform, and impose the condition that if we go L further we end up at the same atom:

$$e^{ik(x+L)} = e^{ikx} \Rightarrow kL = \pm 2n\pi \Rightarrow k = \pm\frac{2n\pi}{L} \qquad \text{and} \qquad n = 0, \ldots, N/2 \tag{822}$$

or

$$k = 0, \pm\frac{2\pi}{L}, \pm\frac{4\pi}{L}, \pm\frac{6\pi}{L}, \ldots, \pm\frac{N}{2}\frac{2\pi}{L} \tag{823}$$

What we are actually saying is that the ends of the chain or the surface of a solid is not really important for the bulk properties. For our linear chain, we see that it already does not make much difference whether it is 11 or inifinity atoms long. In general, we are having of the order of 10^{23} atoms, so we are probably okay. It has some other mathematical advantages, since we do not have to take bound solutions ($\sin kx$), but we can take running waves e^{ikx}.

A Fourier transform can now be given as

$$|k\rangle = \frac{1}{\sqrt{N}}\sum_j e^{-ikx_j}|j\rangle \qquad \text{and} \qquad |j\rangle = \frac{1}{\sqrt{N}}\sum_k e^{ikx_j}|k\rangle, \tag{824}$$

where N is the number of sites (with $N \to \infty$), $|j\rangle$ is the wavefunction at site j, and $|k\rangle$ the wavevector in reciprocal space. The Hamiltonian can be written as

$$H = \sum_{j,\delta} |j+\delta\rangle\langle j+\delta|H|j\rangle\langle j| = \sum_{j,\delta} |j+\delta\rangle T\langle j|, \tag{825}$$

where $\delta = \pm 1$ indicate that the electron can hop to a neighboring site. Making the Fourier transform gives

$$H = \frac{1}{N}\sum_{j\delta}\sum_{kk'} T|k'\rangle e^{ik'x_{j+\delta}}\langle k|e^{-ikx_j} \tag{826}$$

$$= \frac{1}{N}\sum_{j\delta}\sum_{kk'} T|k'\rangle e^{ik'(j+\delta)a}\langle k|e^{-ikja} \tag{827}$$

$$= \sum_{k,k'}\frac{1}{N}\sum_{j} e^{-i(k-k')ja}\sum_{\delta} e^{ika\delta}T|k'\rangle\langle k|, \tag{828}$$

where $x_j = ja$ with a the distance between the particles. The first term gives a delta function

$$H = \sum_{k,k'}\delta_{k,k'}\varepsilon_k|k'\rangle\langle k| = \sum_k \varepsilon_k|k\rangle\langle k|, \tag{829}$$

where

$$\varepsilon_k = \sum_{\delta=\pm 1} e^{ika\delta}T = 2T\cos(ka). \tag{830}$$

We find that the Hamiltonian is diagonal in k. Therefore, k is obviously a good quantum number for translational symmetry. Note, that in principle the matrix does not have to be entirely diagonal. Here we took only one type of orbital, but in principle there could be different kinds. This approach to a solid is what is usually known as a tight-binding scheme or Linear Combination of Atomic Orbitals (LCAO). Also it does not necessarily have to be a circle. It can also be some linear chain of N atoms, where the N'th atom is connected to the first. Note that if the system is large then what happen at the ends of the chain should be unimportant. Note that this solution works regardless of how many atoms there are. For two atoms, we can also apply periodic boundary conditions (although it is a bit artificial). From atom two, we should be able to hop to atoms $2 \pm 1 = 1, 3$. However, since we have periodic boundary conditions atom 3 is atom 1. This is a bit like have a circle of length $2a$ with two atoms on it. So in both cases we hop to atom 1. Instead of a hopping matrix element T, we now have $2T$, finding eigenvalues $\pm 2T$. We have wavevectors $k = 0, \pm\frac{2\pi}{L} = 0, \pm\frac{\pi}{a}$. This looks a bit like three wavevectors whereas we only had two sites, but the maximum wavevectors only count for a half. This gives eigenenergies $2T\cos 0 \times a = 2T$ and $2T\cos\frac{\pi}{a}a = -2T$. For $N = 3$ with periodic boundary conditions, solving a 3×3 matrix would give eigenvalues $E = 2T, -T, -T$ (do it yourself). The wavevectors are $0, \pm\frac{2\pi}{3a}$. (We should be a bit careful here with the wavevectors in Eqn. (823). $N/2$ is not an integer, so these wavevectors should not be included.) The eigenvalues are now $2T\cos 0 = 2T$ and $2T\cos\pm\frac{2\pi}{3a}a = -T$. This is in agreement with eigenvalues of the 3×3 matrix. Show yourself that you can also find the eigenvalues for benzene this way.

We can also write the Eqns. (825) and (829) in a somewhat different way which is more often observed in the scientific literature. Making use of the operators defined earlier, we can write

$$H = \sum_{j,\delta=\pm 1} T c^\dagger_{j+\delta}c_j. \tag{831}$$

The c_j removes an electron from site j, whereas the $c_{j+\delta}$ creates and electron at site $j + \delta$, thereby moving effectively an electron one site to left or to the right. In k space we can write

$$H = \sum_k \varepsilon_k c^\dagger_k c_k = \sum_k \varepsilon_k n_k. \tag{832}$$

Note that the wavevector k is not changed the operators $c^\dagger_k$ and c_k, they just measure the presence of an electron in the state with energy ε_k.

F. (Nearly) free-electron model

Usually discussion on solids do not start with the tight-binding model (which consider the electrons to be in states that resemble closely the orbitals in atoms). Generally one starts at the opposite end. One often considers the solid to be a box of free electrons. The nuclei are there to provide a positive background and to prevent the electrons from flying out of the box. (Usually it takes some minimum energy to remove an electron from the solid, see Eqn.(photoelectric)

and the discussion on the photoelectric effect). In fact, it is not at all obvious that we can just ignore the potential landscape in the solid, but it works well for so-called "simple" metals, such as Be, Mg, Ca, Al, Ga, or Pb. These are all systems with free s or p electrons. For d and f electrons this approximation works much less (remember our discussion on the radial extent of the wavefunction for atoms).

Application of period boundary conditions means that an increase in x by the length of the chain should result in the same function,

$$e^{ik(x+L)} = e^{ikx} \qquad \Rightarrow \qquad k = 0, \pm\frac{2\pi}{L}, \pm\frac{4\pi}{L}, \pm\frac{6\pi}{L}, \ldots, \pm\frac{N\pi}{L}. \tag{833}$$

Note that the maximum k value is $\pm\frac{N\pi}{L} = \pm\frac{\pi}{a}$. We now consider a three-dimensional solid of free electrons. (we assume it is cubic, but that does not really matter since N is so large that k can be generally assumed continous. A free electron satisfies the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi = E\psi \qquad \Rightarrow \qquad |\mathbf{k}\rangle = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}} \qquad \text{with} \qquad E = \varepsilon_k = \frac{\hbar^2k^2}{2m}. \tag{834}$$

Note that the normalization is chosen such that integration of $\langle\mathbf{k}|\mathbf{k}\rangle$ over the volume $V = L^3$ gives unity. The Hamiltonian can now directly be written as

$$H = \sum_k \varepsilon_k c_k^\dagger c_k = \sum_k \varepsilon_k n_k \qquad \text{and} \qquad \varepsilon_k = \frac{\hbar^2k^2}{2m}. \tag{835}$$

Note the difference between the tight-binding model and the free electron model. In the tight-binding model, we start from atoms and then let the electrons move between the atoms. In the free-electron model, we start letting the electrons move, and put the atoms in as a perturbation, as we shall see below. Note that the resemblance between the tight-binding model and the free-electron model is greatest near $k = 0$. In that case we can expand the tight-binding energy as

$$2T\cos ka = 2T\left\{1 + \frac{k^2a^2}{a} + \cdots\right\}, \tag{836}$$

where the second is comparable to a free electron with an effective mass m^*

$$Tk^2a^2 = \frac{\hbar^2k^2}{2m^*} \qquad \Rightarrow \qquad m^* = \frac{\hbar^2}{2a^2T}. \tag{837}$$

Note that if the hopping matrix element T goes to zero, the effective mass goes to infinity.

Now we have to start filling the levels with electrons. This is similar to what we have done for Planck's distribution law with the major difference that we are now dealing with Fermion instead of bosons in Eqn. (8). For Fermions the partition function (the sum of all the probabilities) is

$$Z = \sum_{n_E=0,1}^{\infty} \exp(-n_E\beta E) = 1 + \exp(-\beta E), \qquad \text{with} \qquad \beta = 1/k_BT \tag{838}$$

The chance of finding a state occupied is then simply the nonnormalized chance of finding that state occupied divided by the total chance (Z)

$$\langle n_E\rangle = \frac{\exp(-\beta E)}{1+\exp(-\beta E)} = \frac{1}{1+\exp(\beta E)}. \tag{839}$$

However, this does not entirely work, since for $T \to 0$ we find $f(0) = \frac{1}{2}$ and $f(E) = 0$ for $E > 0$, which means that we could only put one particle in our solid. Therefore, E should be measured with respect to a certain energy known as the chemical potential μ, i.e., $E = \varepsilon - \mu$. The chemical potential should be chosen in such a way that the number of particles is equal to N. In general μ is temperature dependent.

$$f(\varepsilon) = \langle n_\varepsilon\rangle = \frac{1}{1+e^{(\varepsilon-\mu)/k_BT}}. \tag{840}$$

This is known as the Fermi-Dirac function. In general, the Fermi-Dirac function is one up to close to the chemical potential, where it is broadened over a region of $\sim k_BT$. The chemical potential for $T \to 0$ is also known as the

Fermi energy E_F (although the distinction between Fermi energy and chemical potential is a bit blurred in everyday language).

Also we want to know the number of k states. When k becomes large there are many different ways to form a wavevector with a length between k and $k + dk$:

$$2\left(\frac{L}{2\pi}\right)^3 \frac{4}{3}\pi k^3 = \frac{V}{3\pi^2}k^3 \tag{841}$$

Note that in Eqn. (4) there is a factor $\frac{1}{8}$, but the density of k states is twice as large in each direction, i.e., $2^3 = 8$ in total. We can now write down the Fermi wave vector and energy by noting that the number of k states should be equal to the number of electrons N_e:

$$k_F = \left(\frac{3\pi^2 N_e}{V}\right)^{1/3} \qquad \text{and} \qquad E_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 N_e}{V}\right)^{2/3}. \tag{842}$$

A very important quantity in solid state physics is the density of states. Note that we can write the number of particles as

$$N_e = \frac{V}{3\pi^2}\left(\frac{2mE}{\hbar^2}\right)^{3/2}. \tag{843}$$

The density of states is now defined from the number of states in a region dE,

$$dN_e = \rho(E)dE \qquad \Rightarrow \qquad \rho(E) = \frac{dN_e}{dE} = \frac{V}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E}. \tag{844}$$

Obviously, free electron are an oversimplification of the wavefunctions in a solid. The real eigenfunctions are a linear combination of the plane-wave solutions (let us, for simplicity, restrict ourselves to one dimension).

$$|\psi\rangle = \sum_k c_k e^{ikx}. \tag{845}$$

Therefore, we have almost plane waves in a potential landscape, which can also be given in Fourier components as $U(x) = \sum_K U_K e^{iKx}$ Now the most important thing to note about the potential is that it is periodic, i.e.,

$$U(x+a) = U(x) \qquad \Rightarrow \qquad e^{iKa} = 1 \qquad \Rightarrow \qquad K = 0, \pm\frac{2\pi}{a}, \pm\frac{4\pi}{a}, \pm\frac{6\pi}{a}, \ldots, \tag{846}$$

i.e., K can only assume, what are called, reciprocal lattice vectors. Reciprocal lattice vectors are found when making a Fourier transform of the lattice (Some of you might know them from Bragg reflection). Thus we can write

$$U(x) = \sum_K U_K e^{iKx} \qquad \text{with} \qquad K = 0, \pm\frac{2\pi}{a}, \pm\frac{4\pi}{a}, \pm\frac{6\pi}{a}, \ldots. \tag{847}$$

Now let us try to solve the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U(x)\psi = E\psi. \tag{848}$$

Writing out the components gives

$$\frac{\hbar^2 k^2}{2m}\sum_k c_k e^{ikx} + \sum_{K,k} U_K c_k e^{i(k+K)x} = E\sum_k c_k e^{ikx}. \tag{849}$$

Making the substitution $k + K \to k$ in the second term gives the equation for the coefficients (since all the e^{ikx} are independent the coefficient in front of each e^{ikx} must be zero)

$$(\varepsilon_k - E)c_k + \sum_K U_K c_{k-K} = 0, \tag{850}$$

with $\varepsilon_k = \hbar^2 k^2/2m$. It is important to note that a plane wave only couples with other plane waves whose k differ by a reciprocal lattice vector.

Now let us suppose that we only couple to the plane wave with the smallest reciprocal lattice vector (which is not a bad assumption in a large number of cases), we are then left with only two coupled equations:

$$(\varepsilon_k - E)c_k + U_K c_{k-K} = 0; \tag{851}$$

$$(\varepsilon_{k-K} - E)c_{k-K} + U_K c_k = 0. \tag{852}$$

This can be written as a matrix whose eigenvalues can be found by solving the determinant

$$\begin{vmatrix} \varepsilon_k - E & U_K \\ U_K & \varepsilon_{k-K} - E \end{vmatrix} = 0 \qquad \Rightarrow \qquad E_k^{1,2} = \frac{1}{2}(\varepsilon_{k-K} + \varepsilon_k) \pm \sqrt{(\frac{1}{2}(\varepsilon_{k-K} - \varepsilon_k))^2 + U_K^2}. \tag{853}$$

When U is small compared to $\varepsilon_{k-K} - \varepsilon_k$, the effect of the perturbance will be small. However, when $k = \frac{1}{2}K$, $\varepsilon_{k-K} = \varepsilon_k$ and we find that $E_k^{1,2} = \varepsilon_{\frac{1}{2}K}$ and U opens up a gap of of $2U$.

What is so special about the reciprocal lattice vector $\frac{1}{2}K$? This is exactly the wavevector for which one would find a Bragg reflection. Note that the eigenfunction for $E_k^{1,2}$ are approximately $\frac{1}{\sqrt{2}}(e^{ikx} \pm e^{-ikx}) = \sqrt{2}\cos kx, \sqrt{2}\sin kx$ with $k \cong \frac{\pi}{a}$. Note that the cosine has its extrema (i.e., maxima in the probability distribution function) at the atoms, whereas the sine is zero. The particle in the cosine wavefunction therefore feels more of the attractive potential of the atoms and is at $-U$ and the sine is at $+U$.

Note that this can also be understood in a more orbital-like approach. The cosine has its maximum at the atom and we can therefore look at it as a kind of s orbital. The sine changes sign and is therefore similar to a p orbital. The results are therefore in agreement with our expectation that a s like wavefunction is lower in energy than a p like wavefunction.

Since the potential only couples plane waves that differ by a reciprocal lattice vector, the eigenfunctions can be written as

$$\psi_k(x) = \sum_K c_{k-K} e^{i(k-K)x}. \tag{854}$$

This can be rearranged as

$$\psi_k(x) = \left(\sum_K c_{k-K} e^{iKx} \right) e^{ikx} = e^{ikx} u_k(x). \tag{855}$$

Note that the wavefunction now consists of a product of a function with the periodicity of the lattice ($u_k(x+a) = u_k$, since $e^{iKa} = 1$) and a plane wave. Generalizing this to three dimensions, it says that we can write the wavefunction in a periodic potential as

$$\psi_{\mathbf{k}}(x) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}). \tag{856}$$

This is known as Bloch's theorem.

Also the effect if the periodic potential can be generalized. Note that the potential only couples wavevectors that differ by a reciprocal lattice vector, therefore

$$\mathbf{k}' = \mathbf{k} + \mathbf{G}. \tag{857}$$

Since the energy of the electron is conserved, we must have $k = k'$. We can therefore write

$$2\mathbf{k}\cdot\mathbf{G} + G^2 = 0. \tag{858}$$

This is the condition for Bragg reflection. You might not recognize it directly, so let us try to rewrite this condition for x-rays. We can always replace $\mathbf{G}$ by $-\mathbf{G}$, since both are reciprocal lattice vectors. This gives us

$$2\mathbf{k}\cdot\mathbf{G} = G^2 \qquad \Rightarrow \qquad 2\frac{2\pi}{\lambda}\sin\theta = n\frac{2\pi}{d} \qquad \Rightarrow \qquad 2d\sin\theta = n\lambda. \tag{859}$$

In a usual experiment θ is the angle that the incoming and outgoing wave make with a crystal plane. (Note that in Fig. 8.15 in Liboff, the angle with the normal is taken, this differs from the usual practice in diffraction). By satisfying the Bragg condition, the x-ray or electron feels the repeated action of the periodic potential.

XIII. SECOND QUANTIZATION

As you might remember, for the atom we treated the effects of the Coulomb interaction. We would like to do something similar for the electron gas. However, now we are entering a really complicated world. Electron correlations in solid are one of the most complicated areas in physics. We will try to do a bit. Before doing this, let us introduce, in a somewhat hand-waving way, the ideas of second quantization.

Remember that for the harmonic oscillator, we had this nice Hamiltonian

$$H = (a^\dagger a + \frac{1}{2})\hbar\omega, \tag{860}$$

which contains an enormous amount of information without having to worry about Hermite polynomials (life is already difficult enough). This Hamiltonian applies to a large number of possibible oscillators. It is easy to generalize it to include more than one type of oscillator. For example,

$$H = \sum_i (a_i^\dagger a_i + \frac{1}{2})\hbar\omega_i \qquad \text{or} \qquad H = (a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \frac{1}{2})\hbar\omega_{\mathbf{q}}, \tag{861}$$

where we can have oscillators on different sites (which could in principle have different ω) or oscillators in momentum space (for example, creation of photons in plane waves ($e^{i\mathbf{k}\cdot\mathbf{r}}$) or lattice oscillations [also called phonons (for those of you that are already know some solid-state physics: in principle we could add another index $\mathbf{k}\alpha$, since we have two types of phonon optical and acoustic)]. We can now proceed in the same as Eqns. (278) and (279) except that we have to extend our bookkeeping slightly

$$a_i^\dagger |n_1 n_2 \cdots n_i \cdots\rangle = \sqrt{n_i + 1}|n_1 n_2 \cdots, n_i + 1, \cdots\rangle \tag{862}$$

$$a_i |n_1 n_2 \cdots n_i \cdots\rangle = \sqrt{n_i}|n_1 n_2 \cdots, n_i - 1, \cdots\rangle, \tag{863}$$

where we have raised and lowered the occupation of state i by one (we could do exactly the same thing for any index). These operators follow the commutation relation

$$[a_i, a_j^\dagger] = \delta_{i,j}. \tag{864}$$

We can do exactly the same thing for Fermions, which we in fact already described a bit in Section X A on Pauli's exclusion principle, see Eqns. (670)-(674). We can use the fermion operators c and $c^\dagger$ to rewrite the Hamiltonians in Dirac notation in Eqns. (825) and (829) as

$$H_0 = T\sum_{j\delta} c_{j+\delta}^\dagger c_j \qquad \text{and} \qquad H_0 = \sum_k \varepsilon_k c_k^\dagger c_k, \tag{865}$$

respectively. Which looks similar to the Hamiltonian we had for the harmonic oscillator. Let us first look at the second Hamiltonian. Note that $c_k^\dagger c_k$ just measures the number of electrons in state k. The first term is somewhat more complicated: c_j annihililates an electron at site j. If there is no electron, nothing happens (i.e., c_j gives zero). If there is an electron, then the $c_{j+\delta}^\dagger$ term creates an electron at site $j + \delta$, which is a neighboring site since $\delta = \pm 1$. Which means that effectively an electron has hopped from site j to site $j \pm 1$.

We can do the same thing for a two-particle Hamiltonians $\mathcal{H}_1(\mathbf{r}_1, \mathbf{r_2})$ such as the Coulomb interaction. Let us write it down in k space

$$H_1 = \frac{1}{2}\sum_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} V_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} c_{\mathbf{k}_4}^\dagger c_{\mathbf{k}_3}^\dagger c_{\mathbf{k}_2} c_{\mathbf{k}_1} \tag{866}$$

with

$$V_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \varphi_{\mathbf{k}_4}^*(\mathbf{r}_1)\varphi_{\mathbf{k}_3}^*(\mathbf{r}_2) H_1(\mathbf{r}_1, \mathbf{r_2})\varphi_{\mathbf{k}_2}(\mathbf{r}_2)\varphi_{\mathbf{k}_1}(\mathbf{r}_1). \tag{867}$$

The factor 1/2 is to avoid double counting of the electron interactions. This looks pretty horrible and that is exactly what many-body interactions are. Now let us assume that the φ_i are plane waves $V^{-1/2}e^{i\mathbf{kr}}$. It is important to remember that these are the basis functions and not the eigenfunctions. Remember, that for the case of the periodic

potential, we choose plane waves, but the resulting Bloch waves (the eigenfunctions) were linear combinations of plane waves.

$$V_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} = \frac{1}{V^2}\int d\mathbf{r}_1 \int d\mathbf{r}_2 e^{i(-\mathbf{k}_4\cdot\mathbf{r}_1 - \mathbf{k}_3\cdot\mathbf{r}_2 + \mathbf{k}_2\cdot\mathbf{r}_2 + \mathbf{k}_1\cdot\mathbf{r}_1)}\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}, \tag{868}$$

taking $e^2/4\pi\varepsilon_0 \to e^2$. Now let us put $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$ and $\mathbf{r} = \mathbf{r}_2$, this gives

$$V_{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4} = \frac{1}{V}\left(\frac{1}{V}\int d\mathbf{r}\; e^{i(-\mathbf{k}_4-\mathbf{k}_3+\mathbf{k}_2+\mathbf{k}_1)\cdot\mathbf{r}}\right)\left(\int d\mathbf{R}\;\frac{e^2}{R}e^{i(\mathbf{k}_1-\mathbf{k}_4)\cdot\mathbf{R}}\right). \tag{869}$$

The first term between parentheses gives $\delta(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3-\mathbf{k}_4)$, which means that if we choose $\mathbf{k}_1 = \mathbf{k}$, $\mathbf{k}_2 = \mathbf{k}'$, and , $\mathbf{k}_3 = \mathbf{k}-\mathbf{q}$, that $\mathbf{k}_4 = \mathbf{k}+\mathbf{q}$. The second term in parantheses then only depends on $\mathbf{k}_1 - \mathbf{k}_3 = -\mathbf{q}$.

For V_q we obtain

$$V_q = \int d\mathbf{R}\;\frac{e^2}{R}e^{-i\mathbf{q}\cdot\mathbf{R}} = 2\pi e^2\int_0^\infty dR\int_0^\pi d\theta R^2\sin\theta e^{-iqR\cos\theta}\frac{1}{R} = 2\pi e^2\int_0^\infty dRR\int_{-1}^1 d\cos\theta e^{-iqR\cos\theta} \tag{870}$$

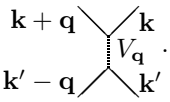
$$= \frac{2\pi e^2}{iq}\int_0^\infty dR(e^{i(q+i\lambda)R} - e^{-i(q-i\lambda)R}) = \frac{4\pi e^2}{q^2+\lambda^2}, \tag{871}$$

where a λ has been introduced to let the oscillations in the sine damp out slowly. In the limit $\lambda \to 0$, we obtain $V_q = 4\pi e^2/q^2$. Later on the λ will actually take on a physical role.

The Hamiltonian now has the form

$$H = H_0 + H_1 = \sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}c^\dagger_{\mathbf{k}}c_{\mathbf{k}} + \frac{1}{2}\sum_{\mathbf{q}\mathbf{k}\mathbf{k}'}\frac{V_q}{V}c^\dagger_{\mathbf{k}+\mathbf{q}}c^\dagger_{\mathbf{k}'-\mathbf{q}}c_{\mathbf{k}'}c_{\mathbf{k}}. \tag{872}$$

Sometimes you might see the Coulomb term as a Feynman diagram language (it is a little bit more complicated than this, but...)



(873)

(874)

Here we see two electrons that have an interaction with each other via the Coulomb potential thereby exchanging momentum q.

A. The electron gas

If we forget about the Coulomb interactions, the ground state at $T = 0$ is the Fermi sphere, i.e. all the independent-particle states (plane wave) up to the Fermi energy are filled:

$$|F\rangle = \prod_{|\mathbf{k}|<k_F} c^\dagger_{\mathbf{k}\downarrow}c^\dagger_{\mathbf{k}\uparrow}|0\rangle, \tag{875}$$

where all the plane wavefunction with a $\mathbf{k}$ whose norm is less than $k_F = \sqrt{2mE_F}/\hbar$ are filled; $|0\rangle$ is the state with no electrons. We fill both spin states. So far we did not include spin in our Hamiltonian. This can be easily introduced

$$H = H_0 + H_1 = \sum_{\mathbf{k}\sigma}\varepsilon_{\mathbf{k}}c^\dagger_{\mathbf{k}\sigma}c_{\mathbf{k}\sigma} + \frac{1}{2}\sum_{\mathbf{q}\mathbf{k}\mathbf{k}'\sigma\sigma'}\frac{V_q}{V}c^\dagger_{\mathbf{k}+\mathbf{q},\sigma}c^\dagger_{\mathbf{k}'-\mathbf{q},\sigma'}c_{\mathbf{k}'\sigma'}c_{\mathbf{k}\sigma}. \tag{876}$$

We would now like to now the ground state energy. Let us first have a look at the kinetic energy H_0. We need to calculate the expectation value of H_0,

$$E_0 = \langle F|H_0|F\rangle = \frac{\hbar^2}{2m}\sum_{|\mathbf{k}|<k_F,\sigma}\mathbf{k}^2. \tag{877}$$

$$E_0 = \langle F|H_0|F\rangle = \frac{\hbar^2}{2m}\sum_{|\mathbf{k}|<k_F,\sigma}\mathbf{k}^2. \tag{878}$$

Let us replace the summation by an integral

$$E_0 = \frac{\hbar^2}{2m}2\frac{V}{8\pi^3}\int_{|\mathbf{k}|<k_F} d\mathbf{k}\; k^2 = \frac{\hbar^2}{2m}\frac{V}{4\pi^3}4\pi\int_{|\mathbf{k}|<k_F} dk\; k^4 = \frac{\hbar^2}{2m}\frac{V}{\pi^2}\frac{1}{5}k_F^5 = \frac{3}{5}N\frac{\hbar^2 k_F}{2m} = \frac{3}{5}NE_F, \tag{879}$$

where we have used $N = Vk_F^3/3\pi^2$ from Eqn. (841).

Let us now try to calculate the first order correction resulting from H_1. Again we assume that is can be calculated as an expectation value in the unperturbed ground state:

$$E_1 = \langle F|H_1|F\rangle = \sum_{\sigma\sigma'}\sum_{\mathbf{k}\mathbf{k}'\mathbf{q}}\frac{V_q}{V}\langle F|c^\dagger_{\mathbf{k}+\mathbf{q},\sigma}c^\dagger_{\mathbf{k}'-\mathbf{q},\sigma'}c_{\mathbf{k}'\sigma'}c_{\mathbf{k}\sigma}|F\rangle. \tag{880}$$

Since the electrons are indistinguishable there are two different ways to make pairs. The most obvious one is to make pairs $c^\dagger_{\mathbf{k}+\mathbf{q},\sigma}c_{\mathbf{k}\sigma}$ and $c^\dagger_{\mathbf{k}'-\mathbf{q},\sigma'}c_{\mathbf{k}'\sigma'}$ which implies that $\mathbf{q} = 0$. This is called the Hartree contribution. (Note what happens if we try to make $\mathbf{q} \neq 0$. Then we annihilate a electron (create a hole) with $\mathbf{k}\sigma$, which necessarily has to be below the Fermi surface and we create an electron with $\mathbf{k}+\mathbf{q},\sigma$ which has to be above the above the Fermi surface. So, effectively, we have created a Fermi surface with an electron-hole pair. However, this is orthogonal to the state $|F\rangle$ and we obtain a zero expectation value). Let us consider the contribution of the $q = 0$ term. Since $V_q = 4\pi e^2/(q^2+\lambda^2)$ this would diverge if we take the limit $q \to 0$. However, for a finite λ, we have

$$\frac{1}{2}\frac{V_0}{V}\sum_{\sigma\sigma'}\sum_{\mathbf{k}\mathbf{k}'}\langle F|c^\dagger_{\mathbf{k},\sigma}c^\dagger_{\mathbf{k}',\sigma'}c_{\mathbf{k}'\sigma'}c_{\mathbf{k}\sigma}|F\rangle = -\frac{1}{2}\frac{V_0}{V}\sum_{\sigma\sigma'}\sum_{\mathbf{k}\mathbf{k}'}\langle F|c^\dagger_{\mathbf{k}',\sigma'}(\delta_{\mathbf{k},\mathbf{k}'}\delta_{\sigma,\sigma'} - c_{\mathbf{k}'\sigma'}c^\dagger_{\mathbf{k},\sigma})c_{\mathbf{k}\sigma}|F\rangle \tag{881}$$

$$= \frac{1}{2}\frac{V_0}{V}\sum_{\sigma\sigma'}\sum_{\mathbf{k}\mathbf{k}'}\langle F|(n_{\mathbf{k}',\sigma'}n_{\mathbf{k},\sigma} - n_{\mathbf{k}',\sigma'})|F\rangle = \frac{1}{2}\frac{V_0}{V}(N^2 - N) \cong \frac{1}{2}\frac{V_0}{V}N^2 = \frac{2\pi e^2 N^2}{V\lambda^2}, \tag{882}$$

where $n_{\mathbf{k},\sigma} = c^\dagger_{\mathbf{k},\sigma}c_{\mathbf{k}\sigma}$ gives the occupation of the state with $\mathbf{k},\sigma$. This does not need to surprise us terribly. It is rather difficult to take of the order of 10^{23} electrons and put them together with all the negative charge present. So far we left out the hole positive background of the nuclei making the total system neutral. Suppose we can describe this background by a uniform density $\rho(\mathbf{r}) = N/V$. We then have two contributions:

$$V_{\text{background}} = \frac{1}{2}e^2\int\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\frac{e^{-\lambda|\mathbf{r}_1-\mathbf{r}_2|}}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 = \frac{2\pi e^2 N^2}{V\lambda^2}, \tag{883}$$

describing the interaction between the nuclei, and

$$V_{\text{e-background}} = -e^2\sum_i\int\rho(\mathbf{r})\frac{e^{-\lambda|\mathbf{r}-\mathbf{r}_i|}}{|\mathbf{r}-\mathbf{r}_i|}d\mathbf{r} = -\frac{4\pi e^2 N^2}{V\lambda^2}, \tag{884}$$

describing the interaction of the electrons with the nuclei. Note the absence of the factor 1/2, since we do not have to correct for double counting the interactions here. The contributions from Eqns. (882), (883), and (884) cancel each other. From now on, we implicitly assume that the summation excludes $\mathbf{q} = 0$.

However, there is another way of combining the creation and annihilation operators $c^\dagger_{\mathbf{k}+\mathbf{q},\sigma}c_{\mathbf{k}'\sigma'}$ and $c^\dagger_{\mathbf{k}'-\mathbf{q},\sigma'}c_{\mathbf{k}\sigma}$ which implies that $\mathbf{k}' = \mathbf{k}+\mathbf{q}$ and $\sigma' = \sigma$ (thus only parallel spins). This is called the Hartree-Fock contribution and is given by

$$\sum_\sigma\sum_{\mathbf{k}\mathbf{q}}\frac{V_q}{V}\langle F|c^\dagger_{\mathbf{k}+\mathbf{q},\sigma}c^\dagger_{\mathbf{k},\sigma}c_{\mathbf{k}+\mathbf{q}\sigma}c_{\mathbf{k}\sigma}|F\rangle = -\sum_\sigma\sum_{\mathbf{k}\mathbf{q}}\frac{V_q}{V}\langle F|n_{\mathbf{k}+\mathbf{q},\sigma}n_{\mathbf{k},\sigma}|F\rangle. \tag{885}$$

It is also known as the exchange term, since the electron are exchanged (therefore they need to have the same spin).

We want to write this in a somewhat different form. We can write the kinetic energy per particle as

$$\langle E_k\rangle = \frac{1}{N}\sum_{\mathbf{k},\sigma}\varepsilon_{\mathbf{k}}\langle F|n_{\mathbf{k},\sigma}|F\rangle. \tag{886}$$

We can rewrite the exchange term as

$$\langle E_{\rm exch}\rangle = \frac{1}{N}\sum_{\mathbf{k},\sigma}\langle F|n_{\mathbf{k},\sigma}|F\rangle\Sigma^{\rm exch}_{\mathbf{k}} \qquad \text{with} \qquad \Sigma^{\rm exch}_{\mathbf{k}} = -\frac{1}{V}\sum_{\mathbf{q}} V_q\langle F|n_{\mathbf{k}+\mathbf{q}}|F\rangle. \tag{887}$$

$\Sigma^{\rm exch}_{\mathbf{k}}$ is known as the self energy of the particle. We see that it is negative. This is a result of Pauli's principle. Since electrons with the same spin cannot be in the same quantum state, they tend to avoid each other. This effectively leads to a somewhat lower Coulomb interaction between electrons with parallel spins.

Let us take $\mathbf{k}' = \mathbf{k} + \mathbf{q}$. We can evaluate the self energy term as

$$\Sigma^{\rm exch}_{\mathbf{k}} = -\int \frac{d\mathbf{k}'}{(2\pi)^3}\frac{4\pi e^2}{|\mathbf{k}-\mathbf{k}'|^2+\lambda^2}n_{\mathbf{k}'} = -\frac{e^2}{\pi}\int_0^{k_F} dk' k'^2 \int_{-1}^{1} da \frac{da}{k^2+k'^2-2kk'a+\lambda^2} \tag{888}$$

$$= -\frac{e^2k_F}{\pi}\left(1+\frac{1-\overline{k}^2}{4\overline{k}}\ln\frac{(1+\overline{k})^2+\overline{\lambda}^2}{(1-\overline{k})^2+\overline{\lambda}^2} - \overline{\lambda}\arctan\frac{1-\overline{k}}{\overline{\lambda}} - \overline{\lambda}\arctan\frac{1+\overline{k}}{\overline{\lambda}}\right) = -\frac{e^2k_F}{\pi}F_{\overline{\lambda}}(\overline{k}), \tag{889}$$

with $\overline{k} = k/k_F$ and $\overline{\lambda} = \lambda/k_F$. Thus the energy is given by

$$E_{\mathbf{k}} = \varepsilon_{\mathbf{k}} + \Sigma^{\rm exch}_{\mathbf{k}} = \frac{\hbar^2k_F}{2m}\left(\overline{k}^2 + \frac{1}{\pi a_0 k_F}F_{\overline{\lambda}}(\overline{k})\right), \tag{890}$$

where $a_0 = \hbar^2/me^2$ is the Bohr radius. Figure 14 shows the energies as a function of k. Unfortunately the results are physically incorrect for $\lambda = 0$. Suppose that the modified dispersion of the electron could be described by an effective mass m^*

$$E_{\mathbf{k}} = E_{\rm shift} + \frac{\hbar^2k^2}{2m^*} \qquad \Rightarrow \qquad dE_{\mathbf{k}} = \frac{\hbar^2k}{m^*}dk. \tag{891}$$

Since this also applies for the nonperturbed case (with $m^* = m$), we can write for the ratio

$$\frac{m}{m^*} = 1 + \frac{\partial\Sigma^{\rm exch}_{\mathbf{k}}}{\partial\varepsilon_{\mathbf{k}}} = \frac{e^2m}{2\pi k_F}\frac{1}{\overline{k}^2}\left(\frac{1+\overline{k}^2}{\overline{k}}\ln\left|\frac{1+\overline{k}}{1-\overline{k}}\right| - 2\right), \tag{892}$$

which has a logarithmic divergence at the Fermi energy $\overline{k} = k/k_F \to 1$. This would have serious consequences at low temperature, where the Fermi gas would become unstable. For example, the specific heat would diverge as $T/\ln T$, which is not observed experimentally. The divergence is a result of the very long range of the Coulomb interaction. If can be removed by a non zero λ. In Fig. 14, one sees that, although the divergence is removed, the energies change relatively little. The effective Coulomb interaction is the $e^2(\mathrm{e}^{-\lambda r}/r)$. The limited range of the Coulomb interaction is known as screening.

B. Thomas-Fermi screening

How can we obtain such a screening of the Coulomb potential? It is possible to obtain it from the Fermi sphere and allow the creation of electron-hole pairs (known as the random phase approximation). However, this requires a knowledge of Green's functions and would go beyond the scope of this course.

The results are very similar to the Thomas-Fermi model. We can understand screening physically by noting that the electron cloud reacts to the electron and will move as a result of the potential created by the electron. However, in a system in equilibrium we assume that the chemical potential $\mu = E_F(\mathbf{r}) - eV(\mathbf{r})$ remains constant. Thus the change in potential has to be compensated by a change in number of electrons

$$\frac{dE_F}{dn}\Delta n - eV = 0. \tag{893}$$

As we saw before, in a free electron model, the Fermi energy is given by

$$k_F = (3\pi^2 n)^{1/3} \Rightarrow E_F = \frac{\hbar^2k_F^2}{2m} = \frac{\hbar^2}{2m}(3\pi n)^{2/3}, \tag{894}$$

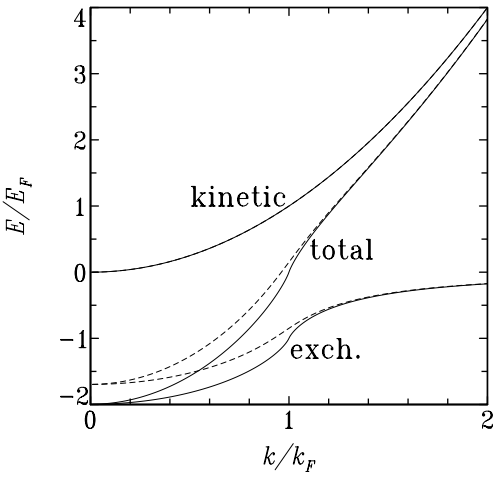


FIG. 14: The energies $\varepsilon_{\mathbf{k}}$ and $\Sigma_{\mathbf{k}}^{\rm exch}$ plus the total energy for $a_0 k_F/\pi = 1$, where $E_F = \hbar^2 k_F^2/2m$. The solid line are the results for $\lambda = 0$, the dashed line for $\lambda = 0.1$.

with the electron density $n = N/V$. This gives $dE_F/dn = \frac{2}{3}E_F/n$. The induced charge is therefore

$$\rho_{\rm ind} = -e\Delta n = -\frac{3}{2}\frac{e^2 nV}{E_F} \tag{895}$$

where use has been made of Laplace's equation $\nabla^2 V = \rho_{\rm tot}/\epsilon_0$. If we write it in Fourier components, we can use $q^2 V = \rho_{\rm tot}/\epsilon_0$, leading to

$$\rho_{\rm ind}(\mathbf{q}) = -\frac{3}{2}\frac{e^2 n}{E_F \epsilon_0 q^2}\rho_{\rm tot}(\mathbf{q}) \tag{896}$$

We now like to relate the effect of our "applied field" (the electron) and the real field

$$\mathbf{E} = \frac{1}{\varepsilon_0}(\mathbf{D} - \mathbf{P}) \qquad \Rightarrow \qquad \nabla\cdot\mathbf{E} = \frac{1}{\varepsilon_0}(\nabla\cdot\mathbf{D} - \nabla\cdot\mathbf{P}) \qquad \Rightarrow \qquad \rho_{\rm tot} = \rho_{\rm ext} + \rho_{\rm ind} \tag{897}$$

Which means that the real electric field is the field from the electron minus a polarization field as a result of the reaction of the electron gas to that field. Now if we assume that the real electric field is directly proportional to the applied field $\varepsilon\varepsilon_0 = \mathbf{D}/\mathbf{E}$, we can write

$$\varepsilon\varepsilon_0\nabla\cdot\mathbf{E} = \nabla\cdot\mathbf{D} \qquad \text{or} \qquad \varepsilon\rho_{\rm tot} = \rho_{\rm ext} = \rho_{\rm tot} - \rho_{\rm ind} \tag{898}$$

For the dielectric function we then find

$$\varepsilon(\mathbf{q}) = 1 - \frac{\rho_{\rm ind}}{\rho_{\rm tot}} = 1 + \frac{3}{2}\frac{e^2 n}{E_F\epsilon_0 q^2} = 1 + \frac{q_{TF}^2}{q^2} \tag{899}$$

with $q_{TF} = \frac{3}{2}\frac{e^2 n}{E_F\epsilon_0}$. The Coulomb potential in momentum space is then

$$V_q = \frac{1}{\epsilon}\frac{e^2}{q^2} = \frac{1}{\epsilon_0}\frac{e^2}{q^2 + q_{TF}^2}, \tag{900}$$

which shows that our $\lambda = q_{TF}$.

XIV. LATTICE VIBRATIONS

Let us consider a (linear) solid consisting of equal atoms. The force pulling the atom back into its equilibrium position is linearly proportional to the distance with the neighboring atom. This leads to the Hamiltonian for the atoms

$$H = \sum_n \frac{p_n^2}{2M} + \frac{K}{2}\sum_n (x_n - x_{n+1})^2. \tag{901}$$