QUANTUM MECHANICS

ITZHAK BARS

September 2005

Contents

Chapter 1

Early and Modern QM

Quantum Mechanics is one of the most important fundamental concepts discovered in the 20^{th} century. In this chapter we first review some of the early physical motivations that led to the initial bold steps of quantum mechanics, and then provide a perspective for the status of Quantum Mechanics as one of the pillars of fundamental theory of Nature as understood by the beginning of the 21^{st} century.

1.1 Origins of Quantum Mechanics

1.1.1 Black body radiation

In 1902 Planck suggested that energy is quantized. He was trying to understand the radiation of black bodies by fitting an empirical curve corresponding to the energy density as a function of temperature and frequency. He considered a hot body that emits radiation through a cavity. In the frequency interval ν to $\nu + d\nu$, the radiation energy density per unit volume U is given by

$$dU(\nu, T) = \frac{4\pi\nu^2}{c^3} d\nu \times 2 \times \bar{E} \tag{1.1}$$

where c is the velocity of light. The first factor is the number of degrees of freedom per unit volume in the electromagnetic radiation in the frequency interval ν to $\nu + d\nu$, the second factor of 2 counts the number of polarizations of the emitted photons, and the last factor is the average energy per degree of freedom. The average energy

$$\bar{E} = \frac{\sum_E E \; e^{-E/kT}}{\sum_E e^{-E/kT}} \tag{1.2}$$

is obtained via statistical mechanical considerations by using the Boltzmann distribution. In 1900 Planck had already fitted the experimental curve with an

empirical formula

$$\frac{dU}{d\nu}(\nu,T) = \frac{8\pi h\nu^3}{c^3(e^{h\nu/kT}-1)}. \tag{1.3}$$

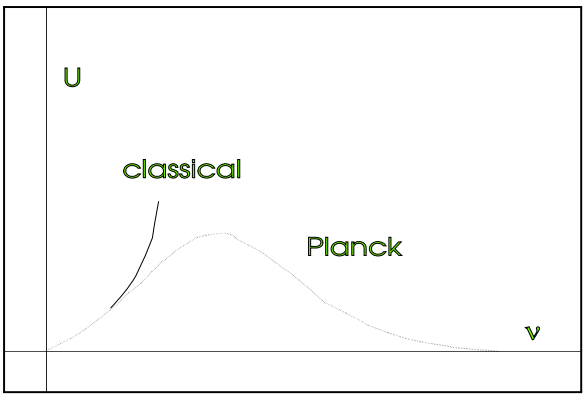


Fig.1.1 : Planck's versus the classical curves.

In Fig.1.1 this curve (in bold) is contrasted with the curve that results from classical considerations, which corresponds to the $h \to 0$ limit of eq.(1.3). Note that there is a logical problem with the classical formula: if one integrates over all values of the frequency to find the total energy carried away by all radiation, one finds an infinite result, which is obviously non-sense. This problem is avoided with Planck's curve. Planck found that the constant $h = 6.63 \times 10^{-15}\ erg/cm^3\,\mathrm{deg}^4$, which is now called the Planck constant, gave the correct experimental fit.

He searched for a theoretical explanation of his empirical formula for the average energy $\bar{E}(\nu,T)$ that would be consistent with eqs.(1.1,1.2,1.3). Previously Wien and Raleigh had constructed models for $\bar{E}$ based on classical physics, but these had failed to give the right answer for all values of the frequency ν. Namely, in classical physics the energy is continuous, and therefore one would perform an integral in eq.(1.2), but this gives $\bar{E} = kT$, which is independent of ν, and leads to the wrong curve. Instead, Planck had the revolutionary idea of postulating that energy comes in quantized units $E = nh\nu$, where n is an integer, and that eq.(1.2) would be computed by performing a sum rather than an integral. He arrived at this view by making an oscillator model for the walls of the cavity that emit the radiation. He then obtained

$$\bar{E} = \frac{\sum_{n=0}^{\infty} e^{-nh\nu/kT} nh\nu}{\sum_{n=0}^{\infty} e^{-nh\nu/kT}} = \frac{h\nu}{e^{h\nu/kT}-1}, \tag{1.4}$$

which is precisely the correct experimental result. Note that the classical result $\overline{E} = kT$ corresponds to the $h \to 0$ limit of eq.(1.4). He therefore came to the conclusion that the walls of the black body cavity must be emitting "quanta" that carried energy in integral multiples of $h\nu$.

1.1.2 Photoelectric effect

Despite Planck's success, the Physics community did not find it easy to accept the idea that energy is quantized. Einstein was the next one to argue that electromagnetic radiation is itself quantized (not just because of the cavity walls), and that it comes in bunches that behave like particles. In his stellar year of 1905 he wrote his article on the photoelectric effect that shows that light behaves like particles. He was interested in explaining the behavior of electrons that are emitted from metals when they are struck by radiation: Electrons were emitted provided the frequency of the light was above some threshold $\nu > \nu_0$, and the number of emitted electrons were determined by the intensity (i.e. the number of incoming photons). Furthermore, the kinetic energy of the emitted electrons, plotted against the frequency, displayed a linear relationship $E_{kin} = h(\nu - \nu_0)$, with a proportionality constant none other than Planck's constant h (see Fig. 1.2).

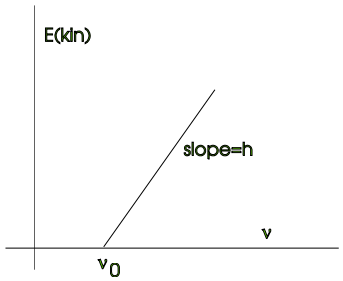


Fig.1.2: Photoelectric effect

To explain these observations Einstein proposed that radiation is made of quanta (photons), with each photon carrying energy

$$E_{photon} = h\nu. \tag{1.5}$$

Furthermore, he postulated that the photons collide with electrons in metals just like billiard balls that conserve energy and momentum, and that it requires a minimum amount of work W to knock out an electron from the metal. Once the electron is struck by a sufficiently energetic quantum of light, it is emitted and carries the excess energy in the form of kinetic energy $E_{kin} = \frac{1}{2}m_e v^2$. According to this billiard ball analogy, the energy conservation equation for each collision reads $E_{photon} = W + E_{kin}$. Using $W = h\nu_0$ this may be rewritten in the form

$$\frac{1}{2}m_e v^2 = h(\nu - \nu_0) \tag{1.6}$$

which gives correctly the experimental curve for the velocity of the emitted electrons. This success, had an important impact on the Physics community for getting closer to the realm of Quantum Mechanics. Millikan performed experimental tests that confirmed Einstein's formula, and Compton successfully extended Einstein's billiard ball approach to the scattering of photons from

electrons. Ironically, Einstein never believed in the probabilistic formulation of Quantum Mechanics that was later developed, and argued against it unsuccessfully throughout his life[?][?].

1.1.3 Compton effect

The idea that light behaves like particles was strengthened by the analysis of the Compton effect. In this case the experiment involves radiation hitting a metal foil and getting scattered to various angles θ. It turns out that the intensity $I(\lambda)$ of the scattered light (number of scattered photons) plotted against the wavelength λ has two peaks: the first is an angle independent peak at a wavelength close to the incident wavelength λ_0, and the second is another peak at a wavelength which is angle dependent at $\lambda = \lambda_0 + \frac{h}{m_e c}(1 - \cos\theta)$ (see Fig. 1.3).

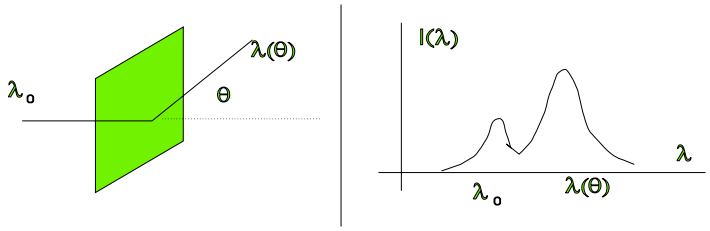


Fig. 1.3 - Compton effect.

Classical physics applied to this problem fails to explain both the angle and wavelength dependence of the observed phenomena. However, by treating light as a particle, and applying momentum and energy conservation to a billiard ball type scattering, Compton explained the observed phenomena correctly.

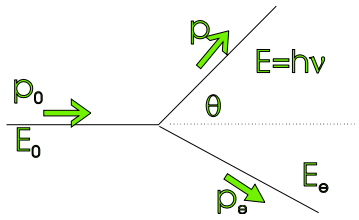


Fig.1.4 - Kinematics

Before the collision the photon and electron have relativistic energy-momentum $(E_0 = h\nu_0,\ \mathbf{p}_0)$ and $(E_{0e} = m_e c^2,\ \mathbf{p}_{0e} = 0)$ respectively, while after the collision they have $(E = h\nu,\ \mathbf{p})$ and $(E_e = \sqrt{m_e^2 c^4 + p_e^2 c^2},\ \mathbf{p}_e)$ (see Fig. 1.4). Momentum and energy conservation require

$$\begin{aligned} \mathbf{p}_0 &= \mathbf{p} + \mathbf{p}_e \\ h\nu_0 + m_e c^2 &= h\nu + \sqrt{m_e^2 c^4 + p_e^2 c^2}. \end{aligned} \tag{1.7}$$

From the first formula one derives ${p_e}^2 = p_0^2 + p^2 - 2p_0 p\cos\theta$, where one may substitute the relativistic photon momenta $p_0 = E_0/c = h\nu_0/c$ and $p = E/c = h\nu/c$. This expression may be replaced in the second formula thus obtaining a relation between the frequencies (ν, ν_0). This relation can be rewritten in terms of the wavelengths $\lambda_0 = \frac{c}{\nu_0}$ and $\lambda = \frac{c}{\nu}$. Finally by isolating the square-root and squaring both sides of the equation one derives

$$\lambda = \lambda_0 + \frac{h}{m_e c}(1 - \cos\theta). \tag{1.8}$$

The combination $\lambda_e = \frac{h}{m_e c} = 3.862 \times 10^{-13} m$ is called the Compton wavelength of the electron. This result explains the second, angle dependent, peak. The first peak occurs due to the scattering of the photon from an entire atom. In this case the mass of the atom must be replaced in place of the mass of the electron so that the Compton wavelength of the atom $\lambda_A = \frac{h}{m_A c} \leq 10^{-16} m$ appears in the formula. Since this is much smaller, the peak appears to be almost angle independent at $\lambda \cong \lambda_0$.

1.1.4 Particle-wave duality

The conclusion from the Planck, Einstein and Compton analyses was that light, which was thought to be a wave classically, could also behave like a particle at the quantum level. It was then natural to ask the question of whether this "particle-wave duality" may apply to other objects that carry energy and momentum? For example, could a classical particle such as an electron behave like a wave at the quantum level? Indeed in 1923 DeBroglie postulated such a particle-wave duality and assigned the wavelength

$$\lambda = \frac{h}{p} \tag{1.9}$$

to any particle that has momentum p. This is consistent with the photon's electromagnetic wave momentum $p = E/c = h\nu/c = h/\lambda$. But he proposed this momentum-wavelength relation to be true for the electron or other classical particles as well, even though the energy-momentum relation for these particles is quite different than the photon's (i.e. $E = p^2/2m$ or $E = (p^2c^2 + m^2c^4)^{1/2}$).

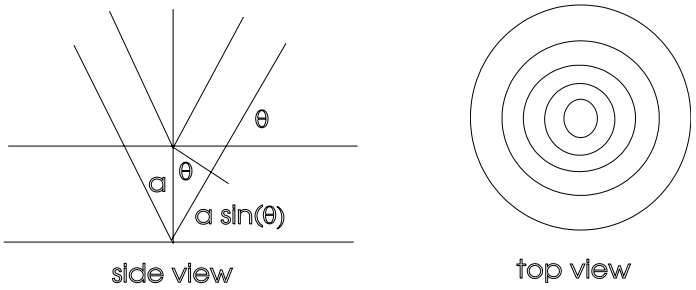


Fig. 1.5 - Interference with electrons.

To check this idea one needed to look for interference phenomena that can happen only with waves, similar to the scattering of light from crystals. Indeed, Davisson and Germer performed experiments in which they scattered electrons from crystals, and found interference patterns of bright and dark rings that have striking similarities to interference patterns of X-rays (see Fig. 1.5). The interference can be understood by considering the difference between the path lengths of electrons that strike two adjacent layers in the crystal. Given that the layers are separated by a distance a and that the rays come and leave at an angle θ, the difference in the path lengths is $2a\sin\theta$. If the electrons behave like a wave, then there will be constructive interference (bright rings) when the path length is a multiple of the wavelength, i.e. $2a\sin\theta = n\lambda$. Thus, for electron beams with momentum p, or kinetic energy $E = p^2/2m_e$, one expects bright rings at angles θ_n that satisfy

$$\sin\theta_n = \frac{n\lambda}{2a} = \frac{nh}{2ap} = \frac{nh}{2a\sqrt{2m_eE}}, \tag{1.10}$$

and dark rings in between these angles. Indeed this is the observation. To distinguish the interference rings from each other the quantity $\lambda/2a$ should not be too small. For electrons in the Davisson-Germer experiment the kinetic energy was about 160 eV, which gave $\lambda/2a \cong 1/4$. One can perform similar interference experiments with molecular beams or slow neutrons with kinetic energies of the order of 0.1 eV. On the other hand for macroscopic objects (e.g. mass of 0.001 mg) moving at ordinary speeds (e.g. 10 $cm/\sec$) the ratio $\lambda/2a$ is too small ($\cong 10^{-14}$) to detect any quantum interference in their behavior.

1.1.5 Bohr atom

While the concept of wave-particle duality was developing, other puzzles about the structure of atoms were under discussion. In 1911 Rutherford had experimentally established that the atom had a positively charged heavy core forming the nucleus of charge Ze and that Z electrons travelled around it like planets around a sun, bound by an attractive Coulomb force. This raised a puzzle: according to the laws of electromagnetism, charged particles that accelerate must radiate energy; since the electrons must accelerate to stay in orbit (radial acceleration) they must gradually loose their energy and fall into the nucleus in about 10^{-10} sec. This reasoning must be false since an atom can live essentially forever, but why? Another puzzle was that when atoms radiated, the emitted photons came in definite quantized frequencies ν, parametrized experimentally by two integers $\nu = const.[(1/n)^2 - (1/n')^2]$, rather than arbitrary continuous frequencies as would be the case in classical physics.

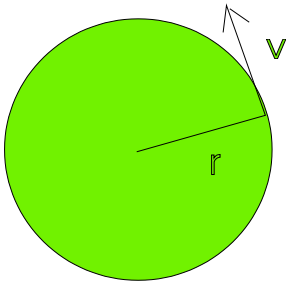


Fig1.6: Circular orbits

The only possibility was to give up classical physics, as was done by Bohr who gave two rules in 1913 to resolve both puzzles. Without an underlying theory Bohr declared the following two principles that explain the observations (see Fig. 1.6)

1) The electron chooses orbits $r = r_n$ in which its angular momentum $L = mvr$ is quantized in units of the Planck constant, $L = mvr = n\hbar$.

As is customary, we have used $\hbar = h/\,2\pi$. To find the orbits and the energies of the electrons one applies the rules of classical physics (force=mass $\times$ acceleration) and the quantization condition

$$\frac{Ze^2}{r^2} = \frac{m_e v^2}{r} \quad , \qquad m_e v r = n\hbar \ , \tag{1.11}$$

and solve for both the radius and the velocity

$$r = \frac{a_0 n^2}{Z} \quad , \qquad v = \frac{Z\alpha c}{n} \ , \tag{1.12}$$

where $\alpha = e^2/\hbar c = 1/137$ is the fine structure constant, and $a_0 = \hbar^2/m_e e^2 \cong 0.53 \ \times 10^{-10} m$ is the Bohr radius. The energy of the electron in such orbits is

$$E_n = \frac{1}{2} m_e v^2 - \frac{Ze^2}{r} = -\frac{m_e c^2 (Z\alpha)^2}{2n^2} \cong -\frac{Z^2}{n^2}(13.6 \ eV) \ . \tag{1.13}$$

Note that $1 \ eV \cong 1.6 \times 10^{-12} \ erg = 1.6 \times 10^{-5} \ J$.

2) Electrons radiate only when they jump from a higher orbit at $r_{n'}$ to a lower one at r_n, and due to energy conservation, the radiation frequency ν is determined by the energy difference in these two orbits, $h\nu = E_{n'} - E_n$.

This gives

$$h\nu = Z^2(13.6 \ eV)[(1/n)^2 - (1/n')^2] \ , \tag{1.14}$$

in accordance with observation[1].

[1] When the quantum numbers get large classical physics results should emerge from quantum mechanics, since large quantum numbers may be regarded as the limit $\hbar \to 0$ while the

The wavelength of the radiation emitted for Hydrogen is in the ultraviolet region, as can be seen by taking the example with $n = 1,\ n' = 2$ which gives $h\nu = (1 - \frac{1}{4})(13.6\ eV)$ and $\lambda = c/\nu \cong 1200$ Angstroms. The success of the Bohr model is additional evidence for the wave-particle duality, since the Bohr quantization rule $mvr = \hbar n$ is rewritten as $2\pi r = nh/mv = nh/p = n\lambda$, where λ is the DeBroglie wavelength (which was proposed later). Therefore it says that an integer number of wavelengths fit exactly into the perimeter of the electron orbit, thus requiring the electron to perform periodic wave motion around the orbit (see Fig. 1.7).

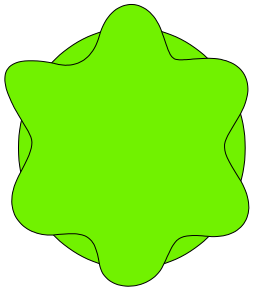


Fig. 1.7 - Waves around the perimeter.

1.1.6 Fundamental principles of QM

With all these hints, it was time to ask the question: what is the wave equation satisfied by particles? One already knew that for radiation the wave equation was given by Maxwell's equations, i.e. $(\nabla^2 - \frac{1}{c^2}\partial_t^2)A_\mu = 0$. Schrödinger first proposed to replace the velocity of light in this equation by the velocity of the particle, but soon afterwards realized that the DeBroglie waves would be correctly described by the Schrödinger equation

$$i\hbar\partial_t\psi = [-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V]\psi\ . \tag{1.15}$$

Then Born proposed to interpret $|\psi(\mathbf{r}, t)|^2$ as the probability of finding the particle at position $\mathbf{r}$ at time t. In the meantime Heisenberg developed a matrix mechanics approach to Quantum Mechanics. Observed quantities such as position or momentum depended on two states and had to be specified in the form x_{ij} and p_{ij}. He extracted multiplication rules for these quantities by analyzing spectral lines for emission and absorption. Finally it was Born and Jordan who realized that these rules could be rewritten as matrix multiplication. It turned

classical quantity remains finite, e.g. $L = \hbar n$. Indeed this correspondance principle may be seen at work in the radiation frequency. The radiation frequency between two neighboring states at large quantum numbers n is $\nu = \frac{m_e c^2}{2h}(Z\alpha)^2[1/n^2 - 1/(n+1)^2] \cong \frac{m_e c^2}{hn^3}(Z\alpha)^2$. On the other hand the classical radiation frequency is $\nu = v/(2\pi r)$ which is seen to be the same once the velocity and radius computed above are substituted in this expression. Thus, for large quantum numbers quantum mechanics reduces to classical mechanics.

out that the matrices that represented position and momentum satisfied the matrix commutation rule

$$(x)(p) - (p)(x) \equiv [(x),(p)] = i\hbar (I) \ , \tag{1.16}$$

where (I) is the identity matrix. It was also understood that in wave mechanics as well there were operators $\mathbf{x}$ and $\mathbf{p} = -i\hbar\boldsymbol{\nabla}$ that satisfy

$$[x_i, p_j] = i\hbar\delta_{ij}, \tag{1.17}$$

when applied on the wavefunction ψ. It can be shown that the non-commutativity of position and momentum leads directly to Heisenberg's uncertainty principle which states that the uncertainties in the measurement of position and momentum must satisfy the inequality $\Delta x \Delta p \geq \hbar/2$ (see next two chapters). Therefore it is not possible to measure both position and momentum simultaneously with infinite accuracy. Quantum Mechanics does not prevent the measurement of the position (or momentum) of a particle with infinite accuracy, but if one chooses to do so then the momentum (or position) of the particle is completely unknown. Finally by 1925, with the work of Born, Heisenberg, Pauli, Jordan, Dirac and Schrödinger, it was understood that all the empirical quantum rules could be derived from the statement that canonical conjugates such as position and momentum do not commute with each other. In fact the commutator is always $i\hbar$ for any set of canonically conjugate observables. Therefore, the rules for Quantum Mechanics boil down to the commutation rules of canonical conjugate pairs. In the 1950's Feynman developed the path integral formalism as an alternative formulation of Quantum Mechanics and showed that it is completely equivalent to the canonical commutation rules. The path integral approach resembles statistical mechanics, and the probabilistic nature of Quantum Mechanics is built in from the beginning. Modern developments in the past couple of decades in several areas of fundamental physics (such as quantum field theory, string theory) have relied heavily on the path integral formulation which turned out to be more convenient for certain computations. These ideas will be explained in more detail in the coming chapters after developing the mathematical formalism of Quantum Mechanics and its physical interpretation in a logical rather than historical sequence.

1.2 QM from one angstrom to the Planck scale

In the 20th century tremendous progress was made in Physics. In Fig.1.8 the evolution of the fundamental theories which describe natural phenomena is given along with the interconnections which exist between these theories. Historically new insight emerged when apparent contradictions arose between theoretical formulations of the physical world. In each case the reconciliation required a better theory, often involving radical new concepts and striking experimental predictions. The major advances were the discoveries of special relativity, quantum mechanics, general relativity, and quantum field theory. All of these are the

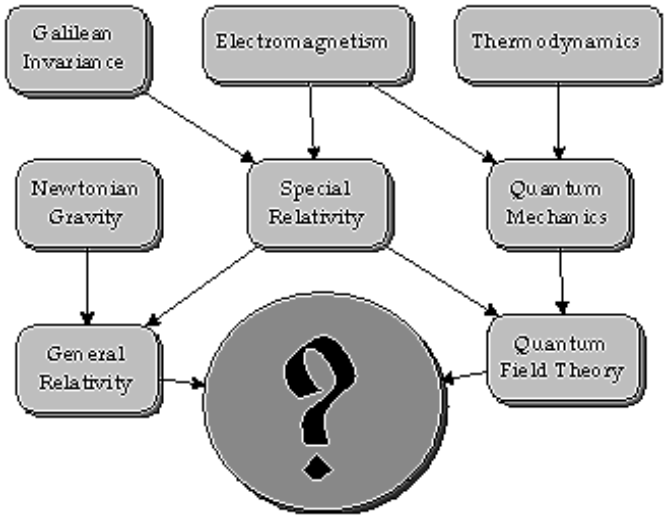


Figure 1.1: Fig.1.8 : Evolution of fundamental theory.

ingredients of the Standard Model, which is a special quantum field theory, that explains natural phenomena accurately down to 10^{-17} cm. The question mark refers to the current status of String Theory which attempts to unify all interactions. These advances were accompanied by an understanding that Nature is described by mathematical equations that have very deep and very beautiful symmetries. In fact, the fundamental physical principles are embodied by the symmetries.

As discussed in this chapter, Quantum Mechanics was born when Planck discovered that he needed to introduce the fundamental constant $\hbar$ in order to understand the thermodynamics and statistical mechanics of black body radiation. To do so he had to abandon certain concepts in classical mechanics and introduce the concept of quantized energy.

Special Relativity developed when Einstein understood the relationship between the symmetries of Maxwell's equations, which describe the properties of light, and those of classical mechanics. He had to introduce the then radical concept that the velocity of light is a constant as observed from any moving frame. In Special Relativity, one considers two observers that are in relative motion to each other with velocity $\mathbf{v}_{rel}$ as in Fig.1.9.

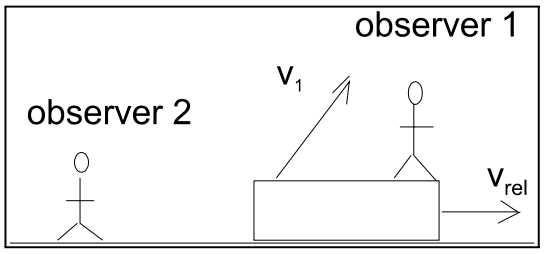


Fig.1.9: Observers in relative motion.

If the first observer measures the velocity of a moving object $\mathbf{v}_1$ in his own frame of reference, then the second observer measures $\mathbf{v}_2$ such that

$$\mathbf{v}_2 = \frac{\mathbf{v}_{rel} + \mathbf{v}_{1\parallel} + \mathbf{v}_{1\perp}\sqrt{1 - \mathbf{v}_{rel}^2/c^2}}{1 + \mathbf{v}_1 \cdot \mathbf{v}_{rel}/c^2}, \tag{1.18}$$

where $\mathbf{v}_{1\parallel}, \mathbf{v}_{1\perp}$ are the components of $\mathbf{v}_1$ that are parallel or perpendicular to $\mathbf{v}_{rel}$ respectively, and c is the velocity of light. One can verify from this formula that, if the particle has the speed of light according to the first observer $|\mathbf{v}_1| = \sqrt{\mathbf{v}_{1\parallel}^2 + \mathbf{v}_{1\perp}^2} = c$, then it also has the speed of light according to the second observer $|\mathbf{v}_2| = c$, for any relative velocity of the two observer $\mathbf{v}_{rel}$. So, the velocity of light is always c as seen by any moving or static observer. This is an example of relativistic invariance (i.e. observations independent of the moving frame characterized by $\mathbf{v}_{rel}$). Furthermore, for slow moving objects and slow moving observers which satisfy $\mathbf{v}_1 \cdot \mathbf{v}_{rel} \ll c^2$ and $\mathbf{v}_{rel}^2 \ll c^2$, the rule for the addition of velocities in eq.(1.18) is approximated by the familiar rule in Newtonian mechanics $\mathbf{v}_2 \to \mathbf{v}_{rel} + \mathbf{v}_1$. Thus, Einstein replaced the Galilean symmetry of Newtonian mechanics (rotations, translations, and Galilean boosts to moving frames) by the Lorentz symmetry of Maxwell's equations and of Special Relativity (rotations, translations, and relativistic boosts to moving frames). Galilean symmetry is just an approximation to Lorentz symmetry when the velocity of the moving frame is much smaller as compared to the velocity of light.

General Relativity emerged from a contradiction between Special Relativity and Newton's theory of gravitation. Newton's gravity successfully explained the motion of the planets and all other everyday life gravitational phenomena. However, it implies instantaneous transmission of the gravitation force between two objects across great distances. On the other hand, according to special relativity no signal can be transmitted faster than the speed of light. The resolution is found in Einstein's general theory of relativity which is based on very beautiful symmetry concepts (general coordinate invariance) and very general physical principles (the equivalence principle). It agrees with the Newtonian theory for low speeds and weak gravitational fields, but differs from it at high speeds and strong fields. In General Relativity gravity is a manifestation of the curvature of space-time. Also, the geometry of space-time is determined by the distribution of energy and momentum. Thus, the fabric of space-time is curved in the vicinity of any object, such that the curvature is greater when its energy (or mass) is bigger. For example space-time is curved more in the vicinity of

the sun as compared to the vicinity of Earth, while it is flat in vacuum. The trajectory of an object in curved space-time corresponds to the shortest curve between two points, called the geodesic. The curving of the trajectory is reinterpreted as due to the action of the gravitational force (equivalence principle). So, gravity acts on any object that has energy or momentum, even if it has no mass, since its trajectory is affected by the curvature of space-time. Thus, the trajectory of light from a star would be curved by another star such as the sun. This tiny effect was predicted with precision by Einstein and observed by Eddington soon afterwards in 1916.

Following the discovery of Quantum Mechanics most of the successful work was done for about thirty years in the context of non-relativistic quantum mechanics. However, another contradiction, known as the Klein paradox, arose in the context of relativistic quantum mechanics. Namely, in the presence of strong fields the formalism of special relativity requires the creation and annihilation of quanta, while the formalism of non-relativistic quantum mechanics cannot describe the physical phenomena. The framework in which quantum mechanics and special relativity are successfully reconciled is quantum field theory. Particle quantum mechanics is itself a limiting case of relativistic quantum field theory. Quantum Electrodynamics (QED), which is a special case of relativistic quantum field theory turned out to explain the interaction of matter and radiation extremely successfully. It agrees with experiment up to 12 decimal places for certain quantities, such as the Lamb shift and the anomalous magnetic moment of the electron and the muon. Such successes provided great confidence that physicists were pretty much on the right track.

There is a special subset of quantum field theories that are especially interesting and physically important. They are called "Yang-Mills" gauge theories, and have a symmetry called gauge invariance based on Lie groups. QED is the first example of such a theory. Gauge invariance turns out to be the underlying principle for the existence of all forces (gravity, electromagnetism, weak and strong forces). The Electroweak gauge theory of Weinberg-Salam and Glashow based on the Lie group SU(2)$\otimes$U(1) is a first attempt to unifying two fundamental interactions (electromagnetic and weak). The Electroweak theory together with QCD , which describes the strong interactions, form the Standard Model of Particle Physics based on the gauge group SU(3)$\otimes$SU(2)$\otimes$U(1). The Standard Model has been shown to be experimentally correct and to describe the fundamental interactions at distances as small as $10^{-19}m$. All phenomena in Nature that occur at larger distances are controlled by the fundamental processes given precisely by the Standard Model. In this theory there exist gauge particles that mediate the interactions. The photon is the mediator of electromagnetic interactions, the $W^{\pm}$ and Z^0 bosons are the mediators of weak interactions and the gluons are the mediators of strong interactions. All known matter is constructed from 6 quarks (each in three "colors") and 6 leptons which experience the forces through their quantum mechanical interactions with the gauge particles. The gauge particles are in one-to-one correspondance with the parameters of the Lie group SU(3)$\otimes$SU(2)$\otimes$U(1), while the quarks and leptons form an array of symmetry patterns that come in three repetitive families. The Stan-

dard Model leaves open the question of why there are three families, why there are certain values of coupling constants and masses for the fundamental fields in this theory, and also it cannot account for quantum gravity. However, it is amazingly successful in predicting and explaining a wide range of phenomena with great experimental accuracy. The Standard Model is the culmination of the discoveries in Physics during the 20th century.

There is one aspect of relativistic quantum field theory, called "renormalization", which bothered theoretical physicists for a while. It involves infinities in the computations due to the singular behavior of products of fields that behave like singular distributions. The process of renormalization removes these infinities by providing the proper physical definition of these singular products. Thus, there is a well defined renormalization procedure for extracting the finite physical results from quantum field theory. This permits the computation of very small quantum corrections that have been measured and thus provides great confidence in the procedure of renormalization. However, renormalization, although a successful procedure, also leaves the feeling that something is missing.

Furthermore, there remains one final contradiction: General relativity and quantum field theory are incompatible because Einstein's General Relativity is not a renormalizable quantum field theory. This means that the infinities cannot be removed and computations that are meant to be small quantum corrections to classical gravity yield infinite results. The impass between two enormously successful physical theories leads to a conceptual crisis because one has to give up either General Relativity or Quantum Mechanics. Superstring theory overcomes the problem of non-renormalizability by replacing point-like particles with one-dimensional extended strings, as the fundamental objects of roughly the size of 10^{-35} meters. In superstring theory there no infinities at all. Superstring theory does not modify quantum mechanics; rather, it modifies general relativity. It has been shown that Superstring theory is compatible with the Standard Model and its gauge symmetries, and furthermore it requires the existence of gravity. At distances of 10^{-32} meters superstring theory is effectively approximated by Supergravity which includes General Relativity. At this stage it is early to know whether superstring theory correctly describes Nature in detail. Superstring theory is still in the process of development.

One can point to three qualitative "predictions" of superstring theory. The first is the existence of gravitation, approximated at low energies by general relativity. Despite many attempts no other mathematically consistent theory with this property has been found. The second is the fact that superstring solutions generally include Yang–Mills gauge theories like those that make up the Standard Model of elementary particles. The third general prediction is the existence of supersymmetry at low energies. It is hoped that the Large Hadron Collider (LHC) that is currently under construction at CERN, Geneva, Switzerland will be able to shed some light on experimental aspects of supersymmetry by the year 2005.

1.3 Problems

1. Consider the Hamiltonian

$$H = \mathbf{p}^2/2m + \gamma |\mathbf{r}| . \tag{1.19}$$

Using Bohr's method compute the quantized energy levels for circular orbits.

2. Consider the Hamiltonian

$$H = c |\mathbf{p}| + \gamma |\mathbf{r}| , \tag{1.20}$$

and its massive version

$$H = \left(\mathbf{p}^2 c^2 + m^2 c^4\right)^{1/2} + \gamma |\mathbf{r}| . \tag{1.21}$$

What are the quantized energy levels for circular orbits according to Bohr's method?

3. According to the non-relativistic quark model, the strong interactions between *heavy* quarks and anti-quarks (charm, bottom and top) can be approximately described by a non-relativistic Hamiltonian of the form

$$H = m_1 c^2 + m_2 c^2 + \mathbf{p}^2/2m + \gamma |\mathbf{r}| - \alpha/|\mathbf{r}| , \tag{1.22}$$

where $m = m_1 m_2/(m_1 + m_2)$ is the reduced mass, and $\mathbf{p}, \mathbf{r}$ are the relative momentum and position in the center of mass (H may be interpreted as the mass of the state, i.e. $H = Mc^2$ since it is given in the center of mass). The combination of linear and Coulomb potentials is an approximation to the much more complex chromodynamics (QCD) interaction which confines the quarks inside baryons and mesons. Note that γ, α are positive, have the units of (energy/distance)=force, and (energy$\times$ distance) respectively, and therefore they may be given in units of $\gamma \sim (GeV/fermi)$ and $\alpha \sim (GeV \times fermi)$ that are typical of strong interactions. Apply Bohr's quantization rules to calculate the energy levels for circular orbits of the quarks. Note that we may expect that this method would work for large quantum numbers. What is the behavior of the energy as a function of n for large n? What part of the potential dominates in this limit?

4. Light quarks (up, down, strange) move much faster inside hadrons. The potential approach is no longer a good description. However, as a simple model one may try to use the relativistic energy in the Hamiltonian (in the rest mass of the system $H = Mc^2$)

$$H = \left(\mathbf{p}^2 c^2 + m_1^2 c^4\right)^{1/2} + \left(\mathbf{p}^2 c^2 + m_2^2 c^4\right)^{1/2} + \gamma |\mathbf{r}| - \alpha/|\mathbf{r}| . \tag{1.23}$$

What are the energy levels (or masses of the mesons) for circular orbits according to Bohr's method? Note that the massless limit $m_1 = m_2 = 0$ is simpler to solve. What are the energy levels in the limits $m_1 = 0,\ m_2 \neq 0$ and $m_1 = m_2 = 0$?

Chapter 2

FROM CM TO QM

In this chapter we develop the passage from classical to quantum mechanics in a semi-informal approach. Some quantum mechanical concepts are introduced without proof for the sake of establishing an intuitive connection between classical and quantum mechanics . The formal development of the quantum theory and its relation to the measurement process will be discussed in the next chapter.

In Classical Mechanics there are no restrictions, in principle, on the precision of simultaneous measurements that may be performed on the physical quantities of a system. However, in Quantum Mechanics one must distinguish between compatible and non-compatible observables. Compatible observables are the physical quantities that may be *simultaneously* observed with any precision. Non-compatible observables are the physical quantities that may not be *simultaneously* observed with infinite accuracy, even in principle. In the measurement of two non-compatible observables, such as momentum p and position x, there will always be some uncertainties Δp and Δx that cannot be made *both* zero simultaneously. As Heisenberg discovered they must satisfy

$$\Delta x \Delta p \geq \hbar/2. \tag{2.1}$$

So, if the position of a particle is measured with infinite accuracy, $\Delta x = 0$, then its momentum would be completely unknown since $\Delta p = \infty$, and vice-versa. In a typical measurement neither quantity would be 100% accurate, and therefore one must deal with *probabilities* for measuring certain values. It must be emphasized that this is not due to the lack of adequate equipment, but it is a property of Nature.

This behavior was eventually formulated mathematically in terms of non-commuting operators that correspond to position $\hat{x}$ and momentum $\hat{p}$, such that

$$[\hat{x}, \hat{p}] = i\hbar. \tag{2.2}$$

All properties of quantum mechanics that are different from classical mechanics are encoded in the non-zero commutation rules of non-compatible observables. The fact that $\hbar$ is not zero in Nature is what creates the uncertainty.

In this chapter we will illustrate the methods for constructing the quantum theory by first starting from the more familiar classical theory of free particles. This will be like a cook book recipe. We will then derive quantum properties such as wave packets, wave-particle duality and the uncertainty principle from the mathematical formalism. It will be seen that the only basic ingredient that introduces the quantum property is just the non-zero commutation rules among non-compatible observables, as above. Classical mechanics is recovered in the limit of $\hbar \to 0$, that is when all observables become compatible.

There is no fundamental explanation for the "classical to quantum recipe" that we will discuss; it just turns out to work in all parts of Nature from one Angstrom to at least the Electroweak scale of $10^{-18}m$, and most likely all the way to the Planck scale of $10^{-35}m$. The distance of 1 Angstrom (10^{-8}cm) naturally emerges from the combination of natural constants $\hbar, e, m_e$ in the form of the Bohr radius $a_0 = \hbar^2/e^2 m_e \cong 0.529 \times 10^{-10}m$. At distances considerably larger than one Angstrom the Classical Mechanics limit of Quantum Mechanics becomes a good approximation to describe Nature. Therefore, Quantum Mechanics is the fundamental theory, while Classical Mechanics is just a limit. The fact that we start the formulation with Classical Mechanics and then apply a recipe to construct the Quantum Theory should be regarded just as an approach for communicating our thoughts. There is another approach to Quantum Mechanics which is called the Feynman path integral formalism and which resembles statistical mechanics. The Feynman approach also starts with the classical formulation of the system. The two approaches are equivalent and can be derived from each other. In certain cases one is more convenient than the other.

It is amusing to contemplate what Nature would look like if the Planck constant were much smaller or much larger. This is left to the imagination of the reader.

2.1 Classical dynamics

The equations of motion of any classical system are derived from an action principle. The action is constructed from a Lagrangian $S = \int dt\ L(t)$. Consider the easiest example of dynamics: a free particle moving in one dimension[1]. Classically, using the Lagrangian formalism, one writes

$$L = \frac{1}{2} m \dot{x}^2. \tag{2.3}$$

Defining the momentum as $p{=}\partial L/\partial \dot{x}$, one has

$$p = m\dot{x} \quad \Longrightarrow \quad \dot{x} = \frac{p}{m}. \tag{2.4}$$

[1] Everything we say in this chapter may be generalized to three dimensions or d dimensions ($d = 2, 3, 4, \cdots$) by simply putting a vector index on the positions and momenta of the particles. We specialize to one dimension to keep the notation and concepts as simple as possible.

The Hamiltonian is given by

$$H = p\dot{x} - L = \quad p\frac{p}{m} - \frac{m}{2}\left(\frac{p}{m}\right)^2 = \frac{p^2}{2m}. \tag{2.5}$$

The Hamiltonian must be expressed in terms of momenta, not velocities. The classical equations of motion follow from either the Lagrangian via Euler's equations

$$\partial_t \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 \quad \Longrightarrow \quad m\ddot{x} = 0 \ , \tag{2.6}$$

or the Hamiltonian via Hamilton's equations

$$\dot{x} = \frac{\partial H}{\partial p} \ , \ \dot{p} = -\frac{\partial H}{\partial x} \quad \Longrightarrow \quad \dot{x} = \frac{p}{m}, \quad \dot{p} = 0. \tag{2.7}$$

In this case the particle moves freely, since its acceleration is zero or its momentum does not change with time.

The relation between momentum and velocity is the familiar one in this case, i.e. $p = m\dot{x}$, but this is not always true. As an example consider the free relativistic particle moving in one dimension whose Lagrangian is

$$L = -mc^2\sqrt{1 - \dot{x}^2/c^2}. \tag{2.8}$$

Applying the same procedure as above, one gets the momentum $p = \partial L/\partial \dot{x}$

$$p = \frac{m\dot{x}}{\sqrt{1 - \dot{x}^2/c^2}} \quad \Longrightarrow \quad \dot{x} = \frac{pc^2}{\sqrt{p^2c^2 + m^2c^4}} \tag{2.9}$$

and the Hamiltonian

$$H = p\dot{x} - L = \sqrt{p^2c^2 + m^2c^4}. \tag{2.10}$$

This energy-momentum relation is appropriate for the relativistic particle. Furthermore, note that even though the momentum takes values in $-\infty < p < \infty$, the velocity cannot exceed the speed of light $|\dot{x}| < c$. Hamilton's equations, $\dot{x} = \partial H/\partial p$, and $\dot{p} = -\partial H/\partial x = 0$, indicate that the particle is moving freely, and that its velocity-momentum relation is the one given above.

Similarly, the classical equations of motion for any system containing any number of free or interacting moving points x_i may be derived from its Lagrangian $L(x_i, \dot{x}_i)$. In non-relativistic mechanics the Lagrangian for N interacting particles is given by the kinetic energy minus the potential energy

$$L = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{x}_i^2 - V(x_1, \cdots, x_N) \tag{2.11}$$

One defines the momentum of each point as $p_i = \partial L/\partial \dot{x}_i$, and from these equations determine the relation between velocities and momenta. Then one may derive the Hamiltonian which is the total energy

$$\begin{aligned} H &= \sum_{i=1}^{N} p_i \dot{x}_i - L(x_i, \dot{x}_i) \\ &= \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(x_1, \cdots, x_N), \end{aligned} \tag{2.12}$$

that must always be written only in terms of positions and momenta $H(x_i, p_i)$.

Although (2.12) is the usual form of the Hamiltonian in non-relativistic dynamics, it is not always the case for more general situations and sometimes it may look more complicated. Nevertheless, through Hamilton's equations we can find out the time evolution of the system once the initial conditions have been specified. One may consider the Hamiltonian as the *generator of infinitesimal time translations* on the entire system, since Hamilton's equations $\dot{x}_i = \partial H/\partial p_i$, and $\dot{p}_i = -\partial H/\partial x_i$ provide the infinitesimal time development of each canonical variable, that is

$$x_i(t+\epsilon) = x_i(t) + \epsilon \dot{x}_i(t) + \cdots \quad \text{and} \quad p_i(t+\epsilon) = p_i(t) + \epsilon \dot{p}_i(t) + \cdots . \tag{2.13}$$

This point of view will generalize to Quantum Mechanics where we will see that the Hamiltonian will play the same role.

2.2 Quantum Dynamics

The classical formalism described above provides the means of defining canonical pairs of positions and momenta (x_i, p_i) for each point i (that is, x_i is associated with the momentum $p_i = \partial L/\partial \dot{x}_i$). It is these pairs that are not compatible observables in Quantum Mechanics. One may observe simultaneously either the position or the momentum of any point with arbitrary accuracy, but if one wants to observe both attributes for the *same point* simultaneously, then there will be some uncertainty for each point i, and the uncertainties will be governed by the equations $\Delta x_i \Delta p_i \geq \hbar/2$. To express the physical laws with such properties one needs to develop the appropriate mathematical language as follows.

A measurement of the system at any instant of time may yield the positions of all the points. Since positions are compatible observables, one may record the measurement with infinite accuracy (in principle) in the form of a list of numbers $|x_1, \cdots, x_N>$. Also, one may choose to measure the momenta of all the points and record the measurement as $|p_1, \cdots, p_N>$. One may also do simultaneous measurements of some compatible positions and momenta and record it as $|x_1, p_2, p_3, \cdots, x_N>$. All of these are 100% precise measurements with no errors. These lists of numbers will be called "eigenstates". They describe the system precisely at any instant of time. We will make up a notation for saying that a particular observable has a 100% accurate measurement that corresponds to an eigenstate. For this purpose we will distinguish the observable from the recorded numbers by putting a hat on the corresponding symbol. So, we have observables $(\hat{x}_i, \hat{p}_i)$ and the statements that they have precise values in some eigenstates are written as

$$\begin{array}{rcl} \hat{x}_i|x_1,\cdots,x_N> &=& x_i|x_1,\cdots,x_N> \\ \hat{p}_i|p_1,\cdots,p_N> &=& p_i|p_1,\cdots,p_N> \\ \hat{x}_1|x_1,p_2,\cdots> &=& x_1|x_1,p_2,\cdots> \\ \hat{p}_2|x_1,p_2,\cdots> &=& p_2|x_1,p_2,\cdots> \end{array} \tag{2.14}$$

etc.. It is said that these operators are "diagonal" on these eigenstates, and the real numbers (x_i, p_i) that appear on the right hand side, or which label the eigenstates, are called "eigenvalues" of the corresponding operators. From now on we will adopt this language of "operator", "eigenvalue" and "eigenstate".

The language that emerged above corresponds to the mathematical language of vector spaces and linear algebra. Let us first give an informal description of the mathematical structure and the corresponding physical concepts. The states are analogous to column matrices or row matrices, while the operators correspond to square matrices. Any square matrix can be diagonalized and it will give an eigenvalue when applied to one of its eigenstates. It is well known that Hermitian matrices must have real eigenvalues. Since only real eigenvalues can correspond to observed quantities, Hermitian matrices will be the candidates for observables. Matrices that commute with each other can be simultaneously diagonalized and will have common eigenstates. These are analogous to compatible observables such as $\hat{x}_1, \hat{x}_2$. Matrices that do not commute with each other cannot be simultaneously diagonalized. They correspond to non-compatible observables such as $\hat{x}_1, \hat{p}_1$. If an observable that is not compatible with the list of measured eigenvalues is applied on the state, as in $\hat{p}_i|x_1,\cdots,x_N>$, the result is not proportional to the same eigenstate, but is some "state" that we will compute later. The vector spaces in Quantum Mechanics are generally infinite dimensional. They are labelled by continuous eigenvalues such as x or p and they are endowed with a dot product and a norm (see below). Such an infinite dimensional vector space is called a Hilbert space.

At this point we can introduce the postulates of Quantum Mechanics. THE FIRST POSTULATE states that: At any instant a physical system corresponds to a "state vector" in the quantum mechanical Hilbert space. THE SECOND POSTULATE states that: To every physical observable there corresponds a linear operator in the Hilbert space. The result of any measurement are real numbers corresponding to eigenvalues of Hermitian operators. These eigenvalues will occur with definite probabilities that depend on the state that is being measured. THE THIRD POSTULATE gives a mathematical expression for the probability as given below. These will be clarified in the discussion that follows.

2.2.1 One particle Hilbert Space

To keep our formalism simple, let us return to the one particle moving in one dimension. In this case there are just two canonical operators $\hat{x}$ and $\hat{p}$. Their eigenstates are $|x>$ and $|p>$ respectively. These will be called "kets" and we will assign to them the mathematical properties of a complex vector space. That is, kets may be multiplied by complex numbers and they may be added to each other. The resulting ket is still a member of the complex vector space.

Kets are analogous to column matrices. For column matrices in n dimensions one may define a complete basis consisting of n linearly independent vectors. Keeping this analogy in mind we want to think of the set of all the kets $\{|x>\}$ as the *complete basis* of an infinite dimensional vector space whose elements are labelled by the continuous number x and whose range is $-\infty < x < \infty$. The idea of completeness here is tied to all the possible measurements of position that one may perform with 100% precision. Similarly, the set of the momentum kets $\{|p>\}$ must form a complete basis for the same particle. Therefore we must think of the position and momentum eigenstates as two complete bases for the same vector space. In the case of an $n-$dimensional vector space one may chose different complete bases, but they must be related to each other by similarity transformations. Similarly, the position and momentum bases must be related to each other by similarity transformations, that is

$$|x>=\int_{-\infty}^{\infty} dp\ |p> F_{p,x} \qquad \text{or} \qquad |p>=\int_{-\infty}^{\infty} dx\ |x> G_{x,p}\ , \tag{2.15}$$

where the functions F, G are to be found.

In general the particle may not be measured in a state of precise position or precise momentum. This will be denoted by a more general vector $|\psi>$ which is a linear superposition of either the position or momentum basis vectors.

$$|\psi>=\int_{-\infty}^{\infty} dx\ |x> \psi(x) = \int_{-\infty}^{\infty} dp\ |p> \widetilde{\psi}(p)\ . \tag{2.16}$$

According to the first postulate of Quantum Mechanics any such state is a physical state, provided it has finite norm (see definition of norm later).

In a vector space it is natural to define an inner product and an outer product. For this purpose one defines a complete set of "bras" $\{<x|\}$ that are in one to one correspondance to the position kets, and similarly another complete set of bras $\{<p|\}$ that are in one to one correspondance to the momentum kets. Bras are analogous to the row vectors of an $n-$dimensional vector space. Formally we relate bras and kets by Hermitian conjugation

$$<x| = (|x>)^\dagger, \qquad <p| = (|p>)^\dagger. \tag{2.17}$$

Therefore, the general vector in bra-space $<\psi| = (|\psi>)^\dagger$has expansion coefficients that are the complex conjugates of the ones that appear for the corresponding kets

$$<\psi| = \int_{-\infty}^{\infty} dx\ \psi^*(x) <x| = \int_{-\infty}^{\infty} dp\ \widetilde{\psi}^*(p) <p|. \tag{2.18}$$

The inner product between a bra $<\phi|$ and a ket $|\psi>$ is analogous to the inner product of a row vector and a column vector. The result is a complex number. The norms of the vectors can be chosen so that the position or momentum bases are orthogonal and normalized to the Dirac delta function

$$< x'|x >= \delta(x' - x) \ , \qquad < p'|p >= \delta(p' - p). \tag{2.19}$$

The outer product between a ket and a bra is analogous to the outer product between a column vector and a row vector, with the result being a matrix. So, the outer product between two states is an operator that will be written as $|\psi_1 >< \psi_2|$. When the same state is involved we will define the Hermitian operator P_ψ

$$P_\psi \equiv |\psi >< \psi|.$$

In particular $|x >< x| \equiv P_x$ is a projection operator to the vector subspace of the eigenvalue x. So, it serves as a "filter" that selects the part of the general state that lies along the eigenstate $|x>$. This can be seen by applying it to the general state

$$\begin{array}{rcl} P_x|\psi > & = & \int_{-\infty}^{\infty} dx' |x >< x|x' > \psi(x') \\ & = & |x > \int_{-\infty}^{\infty} dx' \delta(x - x')\psi(x') \\ & = & |x > \psi(x) \end{array} \tag{2.20}$$

or similarly $< \psi|P_x = \psi^*(x) < x|$. By taking products

$$P_x P_{x'} = |x >< x|x' >< x'| = \delta(x - x')P_x$$

it is seen that the P_x satisfy the properties of projection operators. The complete set of projection operators must sum up to the identity operator

$$\int_{-\infty}^{\infty} dx' \ |x' >< x'| = \mathbf{1} = \int_{-\infty}^{\infty} dp' \ |p' >< p'| \ . \tag{2.21}$$

The consistency of these definitions may be checked by verifying that the symbol $\mathbf{1}$ does indeed act like the number one on any vector. One sees that $\mathbf{1}|x >= |x >$ as follows

$$\mathbf{1}|x >= \int_{-\infty}^{\infty} dx' \ |x' >< x'|x >= \int_{-\infty}^{\infty} dx' \ |x' > \delta(x' - x) = |x >, \tag{2.22}$$

and similarly for any state $\mathbf{1}|\psi >= |\psi >$. Using these properties one may express the functions $F_{p,x}, G_{x,p}, \psi(x), \widetilde{\psi}(p)$ as inner products as follows

$$\begin{array}{lll} |x >= \mathbf{1}|x>= \int_{-\infty}^{\infty} dp \ |p >< p|x > & \Rightarrow & F_{p,x} =< p|x > \\ |p >= \mathbf{1}|p>= \int_{-\infty}^{\infty} dx \ |x >< x|p > & \Rightarrow & G_{x,p} =< x|p > \\ |\psi >= \mathbf{1}|\psi>= \int_{-\infty}^{\infty} dp \ |p >< p|\psi > & \Rightarrow & \widetilde{\psi}(p) =< p|\psi > \\ |\psi >= \mathbf{1}|\psi>= \int_{-\infty}^{\infty} dx \ |x >< x|\psi > & \Rightarrow & \psi(x) =< x|\psi > . \end{array} \tag{2.23}$$

From the above definitions it is evident that

$$\begin{array}{rcl} < & x|p >= (< p|x >)^*, \\ < & \psi|x >= (< x|\psi >)^* = \psi^*(x), \end{array} \tag{2.24}$$

etc.. Also, the inner product between two arbitrary vectors is found by inserting the identity between them

$$\begin{array}{rcl} < \psi_1|\psi_2 > & = & < \psi_1|\mathbf{1}|\psi_2 > \\ & = & \int_{-\infty}^{\infty} dx \ < \psi_1|x >< x|\psi_2 > \quad = \quad \int_{-\infty}^{\infty} dx \ \psi_1^*(x)\psi_2(x) \\ & = & \int_{-\infty}^{\infty} dp \ < \psi_1|p >< p|\psi_2 > \quad = \quad \int_{-\infty}^{\infty} dp \ \widetilde{\psi}_1^*(p)\widetilde{\psi}_2(p). \end{array} \tag{2.25}$$

These properties allows us to interpret the projection operator P_x as the experimental apparatus that measures the system at position x. So, a measurement of the system in the state $|\psi >$ at position $x = 17\,cm$ will be mathematically symbolized by applying the measurement operator $P_{x=17cm}$ on the state $|\psi >$ as above, $P_{17cm}|\psi >= |17\,cm >< 17\,cm|\psi >$. Similarly, a measurement of the system at some momentum with the value p will be symbolized by the projection operator $P_p = |p >< p|$. It is seen that when a measurement is performed on the system $|\psi >$, then its state collapses to an eigenstate of the corresponding observable, and this fact justifies calling this mathematical operation a "measurement". So, the measurement yields an eigenstate of the observable times a coefficient, i.e. $P_x|\psi >= |x > \psi(x)$ or $P_p|\psi >= |p > \tilde{\psi}(p)$. More generally, if the experimental apparatus is not set up to detect an observable with 100% accuracy, but rather it is set up to detect if the system is in some state $|\phi >$, then the measurement operator is $P_\phi = |\phi >< \phi|$, and the measurement yields the collapsed state $P_\phi|\psi >= |\phi >< \phi|\psi >$.

At this point we introduce the THIRD POSTULATE OF QUANTUM MECHANICS: *If the system is in a state $|\psi >$, then the probability that a measurement will find it in a state $|\phi >$ is given by $| < \phi|\psi > |^2$.* The roles of ϕ and ψ may be interchanged in this statement. This probability may be rewritten as expectation values of measurement operators

$$| < \phi|\psi > |^2 =< \psi|P_\phi|\psi >=< \phi|P_\psi|\phi > . \tag{2.26}$$

This leads to the interpretation of the expansion coefficients $\psi(x) =< x|\psi >$ and $\tilde{\psi}(p) =< p|\psi >$ as probability amplitudes for finding the system in state $|\psi >$ at position x or with momentum p respectively. The sums of probabilities such as $|\psi(x_1)|^2 + |\psi(x_2)|^2$ is interpreted as the probability for the system in state ψ to be found at either position x_1 or position x_2. The probability interpretation implies that the system ψ has 100% probability of being found in the same state ψ, i.e. if $\phi = \psi$ then $< \psi|\psi >= 1$, or

$$< \psi|\psi >= \int_{-\infty}^{\infty} dx \ \psi^*(x)\psi(x) = \int_{-\infty}^{\infty} dp \ \widetilde{\psi}^*(p)\widetilde{\psi}(p) = 1, \tag{2.27}$$

for any *physical state* $|\psi >$. The interpretation of this equation makes sense: the total probability that the particle can be found somewhere in the entire universe, or that it has some momentum is 100%.

The interpretation of $|\psi(x)|^2$ as a probability density gives rise to the definition of average position of the system in the state ψ, i.e. $x_\psi = \int_{-\infty}^{\infty} dx \ x \ |\psi(x)|^2$. Similarly the average momentum is $p_\psi = \int_{-\infty}^{\infty} dp \ p \ |\tilde{\psi}(p)|^2$. These may be

rewritten as the expectation values of the position or momentum operators respectively $x_\psi =< \psi|\hat{x}|\psi >$, $p_\psi =< \psi|\hat{p}|\psi >$. This last point may be verified by inserting the identity operator and using the eigenvalue condition as follows

$$< \psi|\hat{x}|\psi >=< \psi|\hat{x}\mathbf{1}|\psi >= \int dx < \psi|\hat{x}|x >< x|\psi >= \int dx\, x\, |\psi(x)|^2 \quad (2.28)$$

and similarly for $< \psi|\hat{p}|\psi >$. Generalizing this observation, the expectation value of any operator $\hat{A}$ in the state ψ will be computed as

$$A_\psi =< \psi|\hat{A}|\psi > . \quad (2.29)$$

The uncertainty of any measurement in the state ψ may be defined for any observable as the standard deviation $(\Delta A)_\psi$ from its average value. Thus

$$(\Delta A)^2_\psi =< \psi|(\hat{A} - A_\psi)^2|\psi >=< \psi|\hat{A}^2|\psi > -(< \psi|\hat{A}|\psi >)^2. \quad (2.30)$$

So, we may write

$$(\Delta x)^2_\psi = \int dx\, (x - x_\psi)^2\, |\psi(x)|^2, \qquad (\Delta p)^2_\psi = \int dp\, (p - p_\psi)^2\, |\tilde{\psi}(p)|^2. \quad (2.31)$$

2.2.2 Quantum rules

Everything that was said so far may apply to Classical Mechanics just as well as to Quantum Mechanics since Planck's constant was not mentioned. Planck's constant is introduced into the formalism by the fundamental commutation rules of Quantum Mechanics. Two canonical conjugate observables such as position and momentum taken *at equal times* are required to satisfy

$$[\hat{x}(t), \hat{p}(t)] = i\hbar \quad (2.32)$$

when applied on any state in the vector space. An important theorem of linear algebra states that *it is not possible to simultaneously diagonalize two non-commuting operators.* As we will see, this means that position and momentum are non-compatible observables and cannot be measured simultaneously with 100% accuracy for both quantities. Their non-compatibility is measured by the magnitude of $\hbar$. Note that it follows that position operators at different times do not commute $[x(t), x(t')] \neq 0$ since the Taylor expansion gives $x(t') = x(t) + (t - t')\dot{x}(t) + \cdots$ and the velocity is related to momentum.

One may now ask: what is the state that results from the action of $\hat{p}$ on a position eigenstate $< x|\hat{p} =?$. It must be consistent with the commutation rule above so that $< x|i\hbar =< x|[\hat{x}, \hat{p}] = x(< x|\hat{p}) - (< x|\hat{p})\hat{x}$. The solution to this equation is

$$< x|\hat{p} = -i\hbar\frac{\partial}{\partial x} < x| \; . \quad (2.33)$$

This can be verified by the manipulation $(< x|\hat{p})\hat{x} = (-i\hbar\frac{\partial}{\partial x} < x|)\hat{x} = -i\hbar\frac{\partial}{\partial x}(x < x|) = -i\hbar < x| + x(< x|\hat{p})$, which leads to the correct result

$< x|i\hbar =< x|[\hat{x},\hat{p}]$. [2] With similar steps one finds also

$$\hat{p}|x >= i\hbar\frac{\partial}{\partial x}|x > \tag{2.34}$$

and

$$< p|\hat{x} = i\hbar\frac{\partial}{\partial p} < p| \ , \qquad \hat{x}|p >= -i\hbar\frac{\partial}{\partial p}|p > \ . \tag{2.35}$$

2.2.3 Computation of $< p|x >$

We are now in a position to calculate the inner product $< p|x >$. Quantities of this type, involving the dot product of different basis vectors, are often needed in quantum mechanics. The following computation may be considered a model for the standard method. One sandwiches an operator between the two states such that its action on either state is known. In the present case we know how to apply either the position or momentum operator, so we may use either one. For example, consider $< p|\hat{x}|x >$ and apply the operator $\hat{x}$ to either the ket or the bra.

$$< p|\hat{x}|x >= \left\{ \begin{array}{l} x < p|x > \\ i\hbar\frac{\partial}{\partial p} < p|x > \end{array} \right. \tag{2.36}$$

The two expressions are equal, so that the complex function $< p|x >$ must satisfy the first order differential equation

$$i\hbar\frac{\partial}{\partial p} < p|x > -x < p|x >= 0. \tag{2.37}$$

This has the solution

$$< p|x >= ce^{-\frac{i}{\hbar}px} \tag{2.38}$$

where c is a constant. Similar steps give the complex conjugate $< x|p >= c^* e^{\frac{i}{\hbar}px}$. To find c consider the consistency with the normalization of the basis and insert identity in the form of the completeness relation

$$\begin{array}{rl} \delta(p-p') =< p|p' >= & \int dx < p|x >< x|p' > \\ & = \int dx|c|^2 e^{-\frac{i}{\hbar}(p-p')x} \\ & = |c|^2 2\pi\hbar\delta(p-p') \end{array} \tag{2.39}$$

Therefore, $|c| = \frac{1}{\sqrt{2\pi\hbar}}$. The phase of c may be reabsorbed into a redefinition of the phases of the states $|x >,\ |p >$, so that finally

$$|x >= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp\ |p > e^{-\frac{i}{\hbar}px} \ , \qquad |p >= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\ |x > e^{\frac{i}{\hbar}px}. \tag{2.40}$$

Therefore position and momentum bases are related to each other through a *Fourier* transformation. Furthermore, one has $\psi(x) =< x|\mathbf{1}|\psi >= \int_{-\infty}^{\infty} dp\ <$

[2] Since these steps are sometimes confusing to the beginner, it may be useful to redo them by defining the derivative of the state by a limiting procedure $(\frac{\partial}{\partial x} < x|)\widehat{x} = (\lim_{a\to 0}\frac{<x+a|-<x|}{a})\widehat{x} = \lim_{a\to 0}\frac{<x+a|(x+a)-<x|x}{a} =< x| + \frac{\partial}{\partial x} < x|$.

$x|p><p|\psi>=\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}dp\ e^{\frac{i}{\hbar}px}\ \widetilde{\psi}(p)$. Similarly, one may insert identity in x-space in $\widetilde{\psi}(p)=<p|\mathbf{1}|\psi>$ and derive that it is the inverse Fourier transform of $\psi(x)$

$$\psi(x)=\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}dp\ e^{\frac{i}{\hbar}px}\ \widetilde{\psi}(p)\ ,\qquad \widetilde{\psi}(p)=\frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}dx\ e^{-\frac{i}{\hbar}px}\ \psi(x) \tag{2.41}$$

Note that Fourier transforms emerged directly from the commutation rules $[\hat{x},\hat{p}]=i\hbar$. Without using Fourier's theorem we have shown that the consistency of the quantum formalism proves that if the first relation in 2.41 is true, then the second one is also true. In other words we proved Fourier's theorem by only manipulating the quantum mechanical formalism. This is just an example of many such mathematical relationships that emerge just from the consistency of the formalism.

2.2.4 Translations in space and time

Momentum as generator of space translations

In a very general way one can see that the momentum operator $\hat{p}$ is the infinitesimal generator of translations in coordinate space. Consider two observers, one using the basis $|x>$ and the other using the basis $|x+a>$ because he measures distances from a different origin that is translated by the amount a relative to the first observer. Both bases are complete, and either one may be used to write an expression for a general state vector. How are the two bases related to each other? Consider the Taylor expansion of $|x+a>$ in powers of a, and rewrite it as follows by taking advantage of $\partial_x|x>=-\frac{i}{\hbar}\hat{p}|x>$

$$\begin{aligned}|x+a> &= \textstyle\sum_{n=0}^{\infty}\frac{a^n}{n!}\frac{\partial^n}{\partial x^n}|x> \\ &= \textstyle\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{-ia\hat{p}}{\hbar}\right)^n|x> \\ &= e^{-ia\hat{p}/\hbar}|x>\end{aligned} \tag{2.42}$$

So, a finite translation by a distance a is performed by the translation operator $\widehat{T}_a=\exp(-ia\hat{p}/\hbar)$, and the infinitesimal generator of translations is the momentum operator. That is, an infinitesimal translation $|x+a>=|x>+\delta_a|x>$, is expressed in terms of the momentum operator as

$$\delta_a|x>=a\frac{\partial}{\partial x}|x>=-\frac{ia}{\hbar}\hat{p}|x>\ . \tag{2.43}$$

Similarly, the position operator performs infinitesimal translations in momentum space. Thus, a shift of momentum by an amount b is given by $\exp(ib\hat{x}/\hbar)|p>=|p+b>$.

Since every state in the basis $\{|x>\}$ is translated by T_a, a general state is also translated by the same operator. Thus, the second observer can relate his state $|\psi'>$ to the state used by the first observer $|\psi>$ by using the translation operator $|\psi'>=T_a|\psi>$, where

$$T_a|\psi>=\int_{-\infty}^{\infty} dx\; T_a|x>\psi(x)=\int_{-\infty}^{\infty} dx\; |x+a>\psi(x)=\int_{-\infty}^{\infty} dx\; |x>\psi(x-a) \tag{2.44}$$

So the probability amplitudes seen by the second observer are related to those of the first observer by

$$\psi(x-a)=< x|T_a|\psi> . \tag{2.45}$$

Hamiltonian as generator of time translations

By analogy to space translations, we may consider an operator that performs time translations

$$U(t,t_0)|\psi,t_0>=|\psi,t> . \tag{2.46}$$

Recall that the Hamiltonian is the generator of infinitesimal time translations in classical mechanics, as explained in eq.(2.13). Therefore, the expansion of $U(t,t_0)$ should involve the Hamiltonian operator. By analogy to the space translation operator above we may write $U(t,t_0)=1-i(t-t_0)\hat{H}/\hbar+\cdots$. Therefore the infinitesimal time translations of an arbitrary state is $\delta t\frac{\partial}{\partial t}|\psi,t>=(-i\delta t\hat{H}/\hbar)|\psi,t>$. This statement is equivalent to

$$i\hbar\frac{\partial}{\partial t}|\psi,t>=\hat{H}|\psi,t>, \tag{2.47}$$

which is the famous Schrödinger equation. Thus, the Schrödinger equation is simply the statement that the Hamiltonian performs infinitesimal time translations. The formal solution of the Schrödinger equation is (2.46), which implies that the operator $U(t,t_0)$ must satisfy the differential equation

$$i\hbar\frac{\partial}{\partial t}U(t,t_0)=\hat{H}U(t,t_0) . \tag{2.48}$$

Therefore, when $\hat{H}$ is time independent one may write the solution

$$U(t,t_0)=e^{-i(t-t_0)\hat{H}/\hbar}. \tag{2.49}$$

If $\hat{H}(t)$ depends on time the solution for the time translation operator is more complicated

$$U(t,t_0)=T\exp[-\frac{i}{\hbar}\int_{t_0}^{t} dt'\hat{H}(t')], \tag{2.50}$$

where T is a time ordering operation that will be studied in a later chapter. For now we consider the case of a time independent Hamiltonian.

The general time dependent state may be expanded in the position basis

$$|\psi,t>=\int_{-\infty}^{\infty} dx\; |x>\psi(x,t)=e^{-it\hat{H}/\hbar}|\psi>, \tag{2.51}$$

where we have chosen the initial state $|\psi, t=0>=|\psi>$. By taking the inner product with $<x|$ we may now derive

$$\psi(x,t) =< x|\psi,t >=< x|e^{-it\hat{H}/\hbar}|\psi> . \tag{2.52}$$

By applying the time derivative on both sides one sees that the time dependent wavefunction satisfies the differential equation

$$i\hbar\,\frac{\partial}{\partial t}\psi(x,t) \;\; = \;\; < x|\hat{H}|\psi,t> \; . \tag{2.53}$$

Up to this point the formalism is quite general and does not refer to a particular Hamiltonian. If we specialize to the non-relativistic interacting particle, its time evolution will be given by

$$\begin{array}{rcl} i\hbar\,\frac{\partial}{\partial t}\psi(x,t) & = & < x|(\frac{\hat{p}^2}{2m}+V(\hat{x}))|\psi,t> \\ & = & [-\frac{\hbar^2}{2m}\partial_x^2+V(x)]\;\psi(x,t). \end{array} \tag{2.54}$$

where the second line follows from the action of $\hat{p}$ on $<x|$. We have thus arrived at the non-relativistic Schrödinger equation!

As another example consider the free relativistic particle for which the Hamiltonian takes the form

$$\begin{array}{rcl} i\hbar\,\frac{\partial}{\partial t}\psi(x,t) & = & < x|\sqrt{m^2c^4+c^2\hat{p}^2}|\psi,t> \\ & = & \sqrt{m^2c^4-\hbar^2c^2\partial_x^2}\;\psi(x,t). \end{array} \tag{2.55}$$

Since it is cumbersome to work with the square roots, one may apply the time derivative one more time and derive

$$[\hbar^2(\partial_t^2-\partial_x^2)+m^2c^2]\;\psi(x,t)=0. \tag{2.56}$$

This is the Klein-Gordon equation in one space dimension. The solutions of these equation will describe the time evolution of the corresponding system. These two examples illustrate that the time development for all systems follows from the fundamental operator equation (2.46,2.53) regardless of the details of the Hamiltonian. Furthermore, the solution of the equation is given by (2.52) for any time independent Hamiltonian, but there remains to compute the matrix element.

2.2.5 Computation of time evolution

For the free particle the Hamiltonian and momentum operators are compatible $[\hat{H},\hat{p}]=0$. Therefore, the Hamiltonian is simultaneously diagonal on the momentum eigenstates

$$\hat{H}|p>=E_p|p> \rightarrow \;\; E_p=\begin{cases} \frac{p^2}{2m} & \text{non-relativistic} \\ (m^2c^4+c^2p^2)^{1/2} & \text{relativistic} \end{cases} \tag{2.57}$$

Then the time evolution of the momentum wavefunctions is easily computed as a phase

$$\widetilde{\psi}(p,t) =< p|e^{-it\hat{H}/\hbar}|\psi >= e^{-itE_p/\hbar}\widetilde{\psi}(p). \tag{2.58}$$

The position wavefunction is obtained by the Fourier transform

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp\ e^{\frac{i}{\hbar}px}\ e^{-itE_p/\hbar}\widetilde{\psi}(p). \tag{2.59}$$

It is easily verified that for any $\widetilde{\psi}(p)$ this is the general solution to the Schrödinger equation above for either the *free* non-relativistic or relativistic particles[3].

In a more general problem, to solve the time evolution of the system one must find the eigenstates of the Hamiltonian

$$\hat{H}|E_i >= E_i|E_i > \tag{2.60}$$

and use these as another orthonormal and complete basis for the same particle. That is

$$< E_i|E_j >= \delta_{ij} \qquad \sum_i |E_i >< E_i| = \mathbf{1}. \tag{2.61}$$

where $\mathbf{1}$ is the same identity operator that has a similar expression in terms of position or momentum bases. Then any state may be expanded in this energy basis instead of the position or momentum basis

$$|\psi, t >= \sum_i |E_i > \psi_{E_i}(t), \tag{2.62}$$

and the time evolution is computed easily as a phase

$$\psi_{E_i}(t) =< E_i|e^{-it\hat{H}/\hbar}|\psi >= c_{E_i}\ e^{-itE_i/\hbar}. \tag{2.63}$$

In particular the position wavefunction may be computed as follows

$$\psi(x,t) =< x|\psi, t >= \sum_i < x|E_i > e^{-itE_i/\hbar} < E_i|\psi > \tag{2.64}$$

provided we specify the basis functions $\psi_{E_i}(x) =< x|E_i >$ and the set of constants $< E_i|\psi >= c_{E_i}$ at the initial time. The energy eigenfunctions may be computed by the same method used for $< x|p >$, i.e. by sandwiching the operator $\hat{H}$ between the states and evaluating the matrix element in two ways

$$< x|\hat{H}|E_i >= \begin{cases} E_i < x|E_i > \\ [-\frac{\hbar^2}{2m}\partial_x^2 + V(x)] \ < x|E_i >, \end{cases} \tag{2.65}$$

[3]However, the *squared* Klein-Gordon equation has a second set of solutions with negative relativistic energies, $E = -E(p)$, which represent anti-particles (although the square-root form has only positive energy solutions). The general solution of the Klein-Gordon equation is the sum of the two types of solutions, with two arbitrary coefficients $\tilde{\psi}_{\pm}(p)$. The negative energy solution has the interpretation of anti-particle.

we see that $\psi_{E_i}(x) =< x|E_i >$ satisfies the time independent Schrödinger equation

$$[-\frac{\hbar^2}{2m}\partial_x^2 + V(x)]\ \psi_{E_i}(x) = E_i\ \psi_{E_i}(x). \tag{2.66}$$

From this we see that from the complete knowledge of the eigenfunctions $\psi_{E_i}(x)$ and eigenvalues E_i we can predict the time development of the system described by that Hamiltonian. The free particle cases worked out above are just special examples of this general procedure.

2.2.6 Wave Packets

According to the third postulate of Quantum Mechanics the quantity $|\psi(x,t)|^2$ is interpreted as the probability for finding the particle in state ψ at position x at time t. In this section we will clarify this point and show its consistency with our intuition from Classical Mechanics.

For the free particle we have seen that the general solution to the Schrödinger equation is given by a general superposition of waves

$$\begin{aligned} \psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dp\,\tilde{\psi}(p)\,e^{i[px-E(p)t]/\hbar}, \\ E(p) &= \begin{cases} p^2/2m & \text{non-relativistic} \\ \sqrt{p^2c^2+m^2c^4} & \text{relativistic} \end{cases} \end{aligned} \tag{2.67}$$

where $\tilde{\psi}(p)$ is arbitrary. For a given momentum p the wave $\exp(i[px-E(p)t]/\hbar)$ is periodic under the translation $x \to x + 2\pi n(\hbar/p)$. Therefore the wavelength is

$$\lambda = \frac{\hbar}{p}, \tag{2.68}$$

in agreement with DeBroglie's idea. Recall that the wave form emerged from the commutation rules $[\hat{x},\hat{p}] = i\hbar$. If $\hbar$ were zero there would be no quantum mechanics, position and momentum would be compatible observables, and the DeBroglie wavelength would vanish.

Let us start at $t = 0$ and choose $\tilde{\psi}(p)$ in such a way that the DeBroglie waves interfere constructively in a certain region and destructively outside of that region. For example, consider the extreme choice $\tilde{\psi}(p) = (1/\sqrt{2\pi\hbar})\exp(-ipx_0/\hbar)$ which yields complete destructive interference everywhere except at one point $\psi(x) = \delta(x - x_0)$. So the particle is located at x_0 without any errors so that $\Delta x = 0$. On the other hand, since the probability distribution in momentum space $|\tilde{\psi}(p)|^2 = 1/2\pi\hbar$ is independent of momentum, the momentum can be anything and therefore $\Delta p = \infty$. However, this wavefunction is not part of the physical Hilbert space since its norm is infinite.

Next consider the Gaussian distribution in momentum space

$$\tilde{\psi}(p) = \left(\frac{2\alpha}{\pi\hbar^2}\right)^{1/4} e^{-ipx_0/\hbar}e^{-\alpha(p-p_0)^2/\hbar^2}, \tag{2.69}$$

which is normalized $\int dp\ |\tilde{\psi}(p)|^2 = 1$. It gives

$$\begin{aligned}
\psi(x) &= \left(\tfrac{2\alpha}{\pi\hbar^2}\right)^{1/4} \int_{-\infty}^{\infty} \tfrac{dp}{\sqrt{2\pi\hbar}}\ e^{ip(x-x_0)/\hbar}\ e^{-\alpha(p-p_0)^2/\hbar^2} \\
&= \left(\tfrac{2\alpha}{\pi\hbar^2}\right)^{1/4} e^{ip_0(x-x_0)/\hbar} \int_{-\infty}^{\infty} \tfrac{dp'}{\sqrt{2\pi\hbar}}\ e^{ip'(x-x_0)/\hbar - \alpha(p'/\hbar)^2} \\
&= \left(\tfrac{2\alpha}{\pi\hbar^2}\right)^{1/4} e^{ip_0(x-x_0)/\hbar} e^{-(x-x_0)^2/4\alpha} \int_{-\infty}^{\infty} \tfrac{dp''}{\sqrt{2\pi\hbar}}\ e^{-\alpha(p''/\hbar)^2} \\
&= \left(\tfrac{1}{2\pi\alpha}\right)^{1/4}\ e^{ip_0(x-x_0)/\hbar}\ e^{-(x-x_0)^2/4\alpha} \quad ,
\end{aligned} \tag{2.70}$$

where $p' = p - p_0$, $\quad p'' = p' - i\hbar x/2\alpha$, and the contour has been deformed in the complex p'' plane back to the real axis since there are no singularities. Note that all the dependence on $x, x_0, p_0, \hbar$ follows from changes of integration variables, and only to determine the overall constant we need the intagral $\int_{-\infty}^{\infty} du\, e^{-u^2} = \sqrt{\pi}$. The probability distributions and the widths in position and momentum space are

$$\begin{aligned}
|\psi(x)|^2 &= \tfrac{1}{\sqrt{2\pi\alpha}}\ e^{-(x-x_0)^2/2\alpha} \qquad & \Delta x &= \sqrt{\alpha} \\
|\tilde{\psi}(p)|^2 &= \sqrt{\tfrac{2\alpha}{\pi\hbar^2}} e^{-2\alpha(p-p_0)^2/\hbar^2} \qquad & \Delta p &= \tfrac{\hbar}{2\sqrt{\alpha}}.
\end{aligned} \tag{2.71}$$

For small α the position distribution is sharply peaked and approaches a delta function, but the momentum is widely distributed (see Fig.2.1).

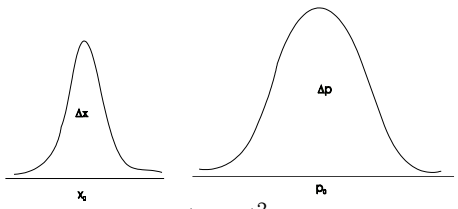


Fig.2.1 - $|\psi(x)|^2$ and $\left|\tilde{\psi}(p)\right|^2$ probability densities.

For large α the opposite is true. Furthermore the uncertainties satisfy the product

$$\Delta x\, \Delta p = \hbar/2. \tag{2.72}$$

Therefore the Gaussian wave packet gives the minimum possible uncertainties according to Heisenberg's relation $\Delta x \Delta p \geq \hbar/2$.

Let us now examine the propagation in time of these Gaussian wave packets according to (2.67). Since the integrand is sharply peaked near p_0 we may approximate

$$E(p) = E(p_0) + (p - p_0)\, E'(p_0) + \frac{1}{2}(p - p_0)^2\, E''(p_0) + \cdots , \tag{2.73}$$

and then perform the integral with the same steps as in (2.70). The result is the probability distribution (see problem)

$$|\psi\,(x,t)\,|^2 = \frac{\exp\left(-(x-x_0(t))^2/2\alpha(t)\right)}{\left(2\pi\sqrt{\alpha\cdot\alpha(t)}\right)^{1/2}} \quad , \qquad \Delta x(t) = \sqrt{\alpha(t)} \tag{2.74}$$

where

$$x_0(t) = x_0 + E'(p_0)\,t, \qquad \alpha(t) = \alpha + \frac{1}{4\alpha}[\hbar t\, E''(p_0)\,]^2. \tag{2.75}$$

Therefore, the position distribution is now peaked at $x = x_0(t)$ and it spreads as time goes on. Notice that $E' = \partial E/\partial p$ is just the velocity (recall $\dot{x} = \partial H/\partial p$), so that the peak moves just like the classical particle for any definition of the energy-momentum relation. This is called the *group velocity* of the wave packet, and this phenomenon now accounts for the **wave-particle duality**. Namely, the wave packet simulates the motion of the particle on the average, while it participates also in wave phenomena such as interference.

Notice that the spreading is less significant for an extremely relativistic particle, since for $cp_0 \gg mc^2$,

$$E''(p_0) = m^2c^6(p_0^2c^2 + m^2c^4)^{-3/2} \approx m^2c^3/p_0^3 \approx 0. \tag{2.76}$$

and there is no spreading at all for particles such as photons or neutrinos that have zero mass $m = 0$ and move at the speed of light. We may compute the amount of time required for doubling the initial uncertainty $(\Delta x)^2 = \alpha$ to 2α. This is given by $\hbar t\, E''(p_0) \geq 2\alpha$, or

$$t \approx 2(\Delta x)^2/(\hbar E''(p_0)\,) \Rightarrow \begin{cases} 2m(\Delta x)^2/\hbar & \text{non-relativistic} \\ 2(\Delta x)^2 p_0^3/\left(\hbar m^2c^3\right) & \text{relativistic} \\ \infty & \text{ultra relativistic} \end{cases} \tag{2.77}$$

So, a non-relativistic electron wave packet localized in the atom within $\Delta x \approx 10^{-10}m$ would spread in about $t \approx 10^{-17}s$. This simply means that in practice the electron's location cannot be pin-pointed within the atom. Furthermore, in this situation one should remember that if one tries to localize a system to distances smaller than its natural size, then the forces acting on the system will dictate a different energy-momentum relation than the one assumed for the free particle. Generally one cannot localize to distances smaller than the natural size, for then the quantum mechanical system gets destroyed, as we shall demonstrate below.

2.2.7 Understanding $\Delta x\ \Delta p \geq \hbar/2$ through gedanken experiments

As we have seen from the wave packet analysis there is an inescapable uncertainty in position and/or momentum as dictated by the uncertainty principle. It is useful to sharpen our intuition about this phenomenon by considering a few gedanken experiments and explaining the role of the uncertainty principle in those situations.

Heisenberg's microscope

Suppose one tries to measure the position of a moving electron by observing it through a microscope. This means that the electron collides with a photon, and the photon is deflected inside the microscope where it is focused to form an image (see Fig.2.2)

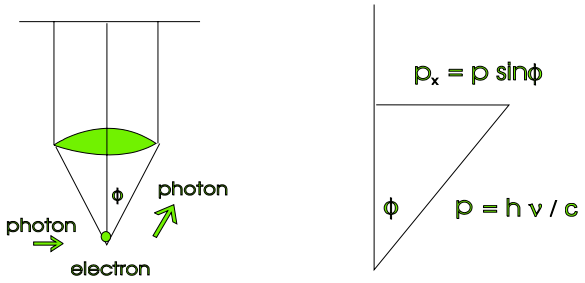


Fig.2.2 - Heisenberg's microscope

According to the laws of optics the resolving power of the microscope is given by the size $\Delta x = \lambda / \sin\phi$. Therefore this is the minimum uncertainty in the position of the observed electron. It would seem that by using very short wavelengths one could locate the electron's position to within any desired accuracy. What happens to the momentum? How accurately can we measure it simultaneously? By momentum conservation, the uncertainty in the x-component of the electron momentum is equal to the uncertainty in x-component of the photon momentum. The photon will be deflected to an angle smaller or equal to ϕ as shown in the figure. The x-component of its momentum will be smaller or equal to $p \sin\phi = (h\nu/c)\sin\phi$. Therefore the uncertainty in the x-component of the photon momentum (and hence in the electron's momentum) will be $\Delta p_x = 2(h\nu/c)\sin\phi$ (the factor of 2 takes into account the full angle). The product of these uncertainties *for the electron* is

$$\Delta x \ \Delta p_x = \frac{\lambda}{\sin\phi} \times 2\frac{h\nu}{c}\sin\phi = 4\pi\hbar \tag{2.78}$$

where we have used $\nu\lambda = c$. So, one could not beat the uncertainty principle. Can one measure the recoil of the screen and determine (Δp_x) photon? No, because one will have to confront (Δx) for the screen, etc.

Two slit experiment:

Consider the two slit experiment. The distance between the two slits is a, and the distance to the screen is d . The path difference between two rays passing through two different slits is $a\sin\theta$, therefore at angles $a\sin\theta_n = \lambda n$ there is constructive interference and bright fringes form (see Fig. 2.3)

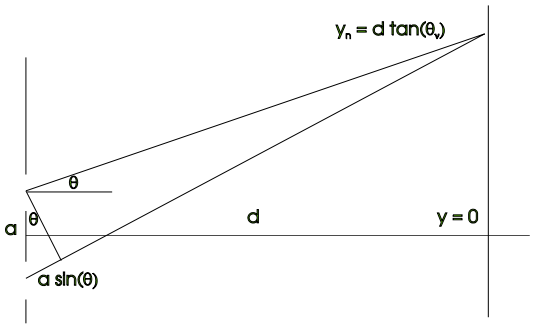


Fig.2.3 - Double slit experiment.

The distance between adjacent fringes is given by

$$(\delta y)_{fringes} = y_{n+1} - y_n = d\tan\theta_{n+1} - d\tan\theta_n \simeq \frac{d\lambda}{a}, \tag{2.79}$$

where we have approximated $\tan\theta \approx \sin\theta$ for small θ. Suppose an observer who is placed at the slits will attempt to tell which slit the electrons pass from. The question is whether this measurement is possible and still have the interference phenomena? If the observer can determine the position of the electron within $\Delta y < a/2$, then he can tell which slit the electron went through. To make an observation he must impart the electron with a momentum which is imprecise by an amount $\Delta p_y > \hbar/a$. But then there will be an uncertainty in the angle of the electron given by

$$\Delta(\sin\theta) = \frac{\Delta p_y}{p} > \frac{\hbar}{ap} = \frac{\lambda}{a} \tag{2.80}$$

This will propagate to an uncertainty on the location the electron hits the screen

$$\Delta y \approx d\ \Delta(\sin\theta) > \frac{d\lambda}{a} = (\delta y)_{fringes}\,. \tag{2.81}$$

The uncertainty Δy is larger than the distance between the fringes $(\delta y)_{fringes}$. Therefore, the interference pattern is lost if the observer tries to localize the electrons at the slits. So, it cannot be done without destroying the system.

Locating the electron position in H-atom

According to Bohr's calculation (see chapter 1) the electron in the H-atom travels in quantized orbits with radii given by $r_n = \hbar n^2/mc\alpha$. The distance between two adjacent radii is $\delta r_n = r_{n+1} - r_n = \hbar(2n+1)/mc\alpha$. Can we make an experiment that measures the position of the electron with an error that is smaller than half this distance, $\Delta r < \delta r_n/2$, so that we can clearly tell where the electron is within the atom?

According to the uncertainty principle, in such an experiment we can determine the momentum of the electron with an uncertainty $\Delta p_r > \frac{\hbar}{2(\delta r_n/2)} =$

$mc\alpha/(2n+1)$. Then the energy will be uncertain by an amount

$$\Delta E \cong p\Delta p/m > \frac{mc\alpha}{n} \cdot \frac{\alpha c}{2n+1} \cong \frac{1}{2}\, mc^2\alpha^2/n^2. \tag{2.82}$$

Since the uncertainty in energy is larger than the binding energy, such a measurement will disturb the atom to the extent of altering its energy level structure. Therefore, such an experiment cannot be performed without destroying the atom.

Quick estimates in quantum theory

One can sometimes use the uncertainty relation to make quick estimates about a quantum system. We give two examples. The first concerns the ground state energy of the H-atom, and the second the mass of the pion.

(i) The H atom is described by the Hamiltonian $H = p^2/2m - Ze^2/r$. The electron is localized within a radius r, so its momentum must be of order $p \approx \hbar/r$. Therefore, the energy is estimated to be $E \sim \hbar^2/2mr^2 - Ze^2/r$. To find the ground state we can minimize the energy with respect to r, $\partial E/\partial r = 0$, and find $r = \hbar^2/me^2Z = a_0/Z$, which is the Bohr radius. Substituting this into the energy equation we get the correct ground state energy $E = -mc^2Z^2\alpha^2/2$

(ii) Yukawa postulated that the nuclear force is due to emission and absorption of the $\pi-$meson. From the knowledge that the range of this force is about one Fermi, we can estimate the mass of the pion by using the energy-time uncertainty relations. Let us first establish this relation. The uncertainty in momentum implies an uncertainty in the energy $\Delta E = p\Delta p/m$, and the uncertainty in position implies an uncertainty in time $\Delta t = \Delta x\,/\dot{x} = \Delta x m/p$. Therefore their product is

$$\Delta E\, \Delta t = \Delta p\, \Delta x \geq \hbar/2. \tag{2.83}$$

This relation was derived for a non-relativistic particle, however it also applies unchanged to a relativistic particle described by the equations (2.9) and (2.10). Now, because of the emission/absorption of the pion the energy is uncertain by an amount $\Delta E = m_\pi c^2$. Therefore the time that it takes for the pion to reach another part of the nucleus may be estimated to be $t \approx \hbar/\ m_\pi c^2$. The distance travelled by the pion during this time will be of the order of the range of the nuclear force, i.e. $10^{-15}m$. So, $r_0 \approx ct \approx c\hbar/\ m_\pi c^2 \approx 10^{-15}m$. This allows the computation

$$m_\pi c^2 = \frac{\hbar c}{r_0} \approx 130 \text{ MeV} \tag{2.84}$$

which is a fairly close estimate of the correct value of $139\, MeV$.

2.3 Problems

1. Consider a particle moving on a circle of radius R instead of the infinite line. The position and momentum eigenstates are $|x>$ and $|p>$, respec-

tively. What are the allowed eigenvalues? Why? Consider probabilities in providing an answer. What is $< x|p >=$? Write down the completeness and orthonormality relations, and show how they work on the set of functions $< x|p >$. What happens in the infinite radius limit? Next consider a particle moving on a torus of radii (R_1, R_2). Generalize your reasoning and provide the allowed eigenvalues for positions and momenta.

2. Consider two operators $\hat{\alpha}$ and $\hat{\beta}$ whose commutator is a c-number $[\hat{\alpha}, \hat{\beta}] = c$. Prove that
$$e^{\hat{\alpha}} e^{\hat{\beta}} = e^{\hat{\alpha}+\hat{\beta}} \times const. \tag{2.85}$$
and determine the constant. Now consider space translations by a distance a that are performed by the operator $A = \exp(-ia\hat{p}/\hbar)$ and momentum shifts by an amount b that are performed by the operator $B = \exp(ib\hat{x}/\hbar)$. Show that the products AB and BA differ from each other by a phase, i.e. $AB = BAe^{i\phi}$, and find the conditions on a, b such that $[A, B] = 0$ become compatible observables.

3. Consider the Hermitian dilatation operator $D = \frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})$. It commutes with the parity operator P. By definition, their eigenvalues and eigenvectors satisfy $D|\lambda, \pm >= \lambda|\lambda, \pm >$, and $P|\lambda, \pm >= \pm|\lambda, \pm >$. These states must form a complete orthonormal basis just like $|x >$ or $|p >$. Therefore one may expand one basis in terms of the others. Find the correctly *normalized* position space wavefunctions (or expansion coefficients) $\psi^{\pm}_{\lambda}(x) =< x|\lambda, \pm >=?$, and prove that they are complete and orthonormal
$$\begin{aligned} \sum_{\pm} \int_{-\infty}^{\infty} d\lambda \; \psi^{\pm *}_{\lambda}(x) \; \psi^{\pm}_{\lambda}(x') &= \delta(x - x'), \\ \int_{-\infty}^{\infty} dx \; \psi^{\pm *}_{\lambda}(x) \; \psi^{\pm}_{\lambda'}(x) &= \delta(\lambda - \lambda'). \end{aligned} \tag{2.86}$$

4. In class we discussed the Gaussian wave packets at arbitrary times $\psi(x, t)$. Using these wavefunctions compute the average position, the average momentum, and the average uncertainty
$$\begin{aligned} x_{\psi}(t) &\equiv < \psi, t|\hat{x}|\psi, t >, \\ p_{\psi}(t) &\equiv < \psi, t|\hat{p}|\psi, t >, \\ \Delta x_{\psi}(t) &= [< \psi, t|(\hat{x} - x_{\psi}(t))^2|\psi, t >]^{1/2} \end{aligned} \tag{2.87}$$
at arbitrary times, for (i) a free non-relativistic particle, and for (ii) a massless ultra-relativistic particle. What relation is there between $\dot{x}_{\psi}(t)$ and $p_{\psi}(t)$?

5. Suppose that at $t = 0$ the wavefunction of a free non-relativistic particle is
$$\psi(x, 0) = Ce^{-a|x|+ikx}. \tag{2.88}$$

What is the correct normalization $C =?$ What is the momentum wavefunction $\tilde{\psi}(p,0) =?$ What is $\psi(x,t) =?$

6. Consider the position and momentum translation operators of problem 2. Assume that they are applied on the wavefunction of a particle that moves on a circle. Such a wavefunction must be periodic $\psi(x + 2\pi R) = \psi(x)$. Suppose that the translation operator A applied N times on this wavefunction is equivalent to a $2\pi R$ translation, i.e. $A^N\psi = \psi$. Assume that the particle is allowed to live only on the discrete points $x_n = (n/N)\, 2\pi R$ connected to each other by the translation A (like particles in a periodic crystal). Labelling the wavefunction at the positions x_n as ψ_n we may write $A\psi_n = \psi_{n+1}$ and $\psi_{n+N} = \psi_n$. There are N independent states, so we may take them as isomorphic to the N dimensional vector space of column or row matrices

$$\psi_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \psi_N = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{2.89}$$

Furthermore, this position wavefunction is obviously an eigenstate the operator B :

$$B\psi_n = \exp\left(ibx_n/\hbar\right)\ \psi_n = \exp\left(ibn2\pi R/N\hbar\right)\ \psi_n. \tag{2.90}$$

Note that $A^N = 1$ is simultaneously diagonal on every state. Let's impose also the condition that $B^N = 1$ on the wavefunctions. Then the allowed eigenvalues for B must satisfy $\exp(iNbx_n/\hbar) = 1$, for every n, which fixes $bR/\hbar = k =$ integer. Using the smallest possible b, we may take $k = 1$. Therefore the eigenvalues of B are $\exp(in2\pi/N)$. So, we can write B as a diagonal $N \times N$ matrix

$$B = \begin{pmatrix} \exp(i2\pi/N) & & & & \\ & \exp(i4\pi/N) & & & \\ & & \exp(i6\pi/N) & & \\ & & & \ddots & \\ & & & & \exp(i2\pi) \end{pmatrix} \tag{2.91}$$

This allows us to interpret ψ_n as position eigenstates. What is the matrix form of the translation operator $A =?$ (note this is called a "circular matrix"). Verify explicitly the relation $AB = BAe^{i\phi}$ by matrix multiplication; what is the phase? Now consider the products of translations in position and momentum space $A^{m_1}B^{m_2}$. How many such independent operators are there? What do get if you try to commute them $[A^{m_1}B^{m_2}, A^{n_1}B^{n_2}] =?$

7. The uncertainty relation $\Delta E \Delta t \geq \hbar/2$ was derived for a non-relativistic particle. Show that it also applies to a relativistic particle described by the eqs.(2.9) and (2.10).

Chapter 3

STRUCTURE OF QM

3.1 Postulates

Quantum Mechanics is based on three postulates that establish the mathematical language corresponding to the physical concepts. In this chapter the mathematical structure of QM will be discussed in general terms. The relation of mathematical concepts to physical concepts, such as Hilbert spaces to physical states, operators and matrices to observables, eigenstates and eigenvalues to physical measurements, etc., will be explained. The meaning of solving a quantum mechanical problem completely, will be clarified. Finally these concepts will be illustrated in a quantum system consisting of only two states.

The three postulates of QM were already introduced in the second chapter in the discussion of the single free particle. We state them once again in general terms:

1. A physical state is represented by a vector in the Hilbert space $|\psi>$.

2. To every physical observable there corresponds a *linear, Hermitian* operator in the Hilbert space. Operators have eigenstates and corresponding eigenvalues. The results of measurements are the eigenvalues of compatible observables. As we will see, the simultaneous eigenstates of a complete set of compatible observables form a complete basis, such that an arbitrary state $|\psi>$ in the Hilbert space can be written as a linear superposition of the basis.

3. When a system is prepared in a state $|\psi>$, and it is probed to find out if it is in a state $|\phi>$, after the measurement the state $|\psi>$ collapses to the state $|\phi><\phi|\psi>$, where $<\phi|\psi>$ is a complex number called the probability amplitude. The probability that a measurement will find the system in a state $|\phi>$ is given by $|<\phi|\psi>|^2$.

In principle, in an ideal measurement the state $|\phi>$ is an eigenstate of the observables being measured. Therefore the observed eigenvalues occur with

definite probabilities depending on the state $|\psi>$. More generally, the state $|\phi>$ may be a superposition of eigenstates of the observables, corresponding to a distribution of eigenvalues. A consequence of the probability interpretation is that a physical state must satisfy $|<\psi|\psi>|=1$, or

$$<\psi|\psi>=\pm 1, \tag{3.1}$$

since a system prepared in a state $|\psi>$ has 100% probability of being found in the same state $|\psi>$. The (-1) case is discarded as part of the postulates of QM. This corresponds to requiring a unitary positive norm Hilbert space. We now define the necessary mathematical concepts, relate them to physical quantities according to these postulates and clarify their meaning.

3.2 States, kets&bras, norm

A physical state has a mathematical description in terms of a vector in a unitary Hilbert space. A unitary Hilbert space is an infinite dimensional vector space with a positive norm that is not infinite. However, there are quantum mechanical systems that are effectively described in terms of finite dimensional vector spaces as well. We will consider a Hilbert space as the $n\to\infty$ limit of a complex vector space in n dimensions. A complex vector space is a set of elements called vectors (which we denote here by the ket symbol $|\bullet>$) that may be multiplied by complex numbers and added to each other such that the set remains closed under these operations. So, $\alpha|\psi>+\beta|\phi>$ is in the set if (α,β) are complex numbers and $(|\psi>,|\phi>)$ are vectors in the set. A complex vector space in n-dimensions boils down to a set of basis vectors $\{|i>,\quad i=1,2,\cdots n\}$. A general vector $|\psi>=\sum_i|i>\psi_i$ is a linear superposition with coefficients ψ_i that are complex numbers. It is evident that general vectors satisfy the definition.

There is a dual vector space whose elements are put in one to one correspondence to the original vector space by Hermitian conjugation. Formally one writes the bra $<\psi|\equiv(|\psi>)^\dagger$, and the dual basis vectors are denoted by bras $\{<i|,\quad i=1,2,\cdots n\}$. Therefore the general dual vector is $<\psi|=\sum_i\psi_i^*<i|$, where the complex conjugate coefficients appear.

It is natural to define inner and outer products $<i|j>$ and $|j><i|$ respectively. By assumption, the basis is complete, and it is always possible to bring it to orthonormal form. These statements are expressed as

$$\sum_{i=1}^{n}|i><i|=\mathbf{1},\qquad <i|j>=\delta_{ij}=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases}, \tag{3.2}$$

where δ_{ij} is the Kronecker delta function. The dot product of two general vectors follows as a complex number

$$<\phi|\psi>=\sum_{ij}\phi_i^*\psi_j<i|j>=\sum_i\phi_i^*\psi_i, \tag{3.3}$$

and the norm of the general vector is real and positive

$$< \psi|\psi >= \sum_{i=1}^{n} \psi_i^* \psi_i, \tag{3.4}$$

and may be normalized to $< \psi|\psi >= 1$.

A finite dimensional vector space can be put in one to one correspondence with column and row matrices. Thus, $|i >$ is represented by a column vector that has zero entries everywhere except at the $i'th$ row, where the entry is 1. Similarly, $< i|$ is represented by the Hermitian conjugate row matrix. The outer product $|j >< i|$ is a square matrix with zeros everywhere except at location (j, i)

$$|1 >= \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \cdots \qquad |n >= \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

$$< 1| = (1\ 0\ \cdots\ 0) \qquad \cdots \qquad < n| = (0\ 0\ \cdots\ 1)$$

$$|1 >< 1| = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & & \cdots & 0 \end{pmatrix} \quad \cdots \quad |1 >< n| = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & & \cdots & 0 \end{pmatrix} \quad \cdots \tag{3.5}$$

Therefore, the general vector $|\psi >$ and $< \psi|$ are general complex columns or rows respectively, while the linear combination of outer products

$$M = \sum_{kl} |k > M_{kl} < l| = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix} \tag{3.6}$$

is a general complex matrix with entries M_{kl}. The matrix elements may be computed by using the dot products

$$< i|M|j >= \sum_{kl} < i|k > M_{kl} < l|j >= \sum_{kl} \delta_{ik} M_{kl} \delta_{lj} = M_{ij}. \tag{3.7}$$

An infinite dimensional vector space may be countable or not. If it is countable it may be labelled by integers $i = 1, 2, \cdots \infty$, so that it is defined by simply taking the limit $n \to \infty$ in the above expressions. It may also be convenient to label it with all integers $i = -\infty, \cdots, -1, 0, 1, \cdots, \infty$

instead of only the positive integers; these are equivalent. If it is not countable it may be labelled by a continuous variable, as e.g. the position basis $\{|x>,\ -\infty < x < \infty\}$. In this case it is useful to think of the continuous variable as the limit of a lattice with the lattice constant going to zero, i.e. $\{x = ia;\quad i = -n, \cdots, -1, 0, 1, \cdots, n;\ a \to 0,\ n \to \infty\}$. This allows us to define vectors with a new normalization

$$|x = ia >= \frac{|i>}{\sqrt{a}} \tag{3.8}$$

such that, in the $a \to 0$ limit, the sums are replaced by integrals (using $\int dx \sim \sum a$) and the Kronecker delta is replaced by a Dirac delta function

$$\begin{aligned} \int_{-\infty}^{\infty} dx\ |x><x| \quad &= \lim_{\substack{a\to 0\\ n\to\infty}} \sum_{i=-n}^{n} a \frac{|i>}{\sqrt{a}} \frac{<i|}{\sqrt{a}} = \mathbf{1} \\ < x|y > \quad &= \lim_{\substack{a\to 0\\ n\to\infty}} \frac{<i|}{\sqrt{a}} \frac{|j>}{\sqrt{a}} = \lim_{\substack{a\to 0\\ n\to\infty}} \frac{\delta_{ij}}{a} = \delta(x-y). \end{aligned} \tag{3.9}$$

Therefore, the norm of the vector is infinite $< x|x >= \delta(0) = \infty$. However, this is what is required for consistency with integration of the delta function

$$\int_{-\infty}^{\infty} dx\, \delta(x-y) = \lim_{\substack{a\to 0\\ n\to\infty}} \sum_{i=-n}^{n} a \frac{\delta_{ij}}{a} = 1 \tag{3.10}$$

and is self consistent as seen in chapter 2. In some applications it is also possible that the interval in the continuous x-variable is finite $x \in [-\alpha, \alpha]$. In this case the limits for n and a are not independent, and must be taken such that $na = \alpha$ always remains finite.

A Hilbert space is an infinite dimensional vector space which admits only vectors which have a finite norm. Therefore, not all vectors that can be constructed from the continuous basis $|x>$ is a member of the Hilbert space. In particular the vector $|x>$ itself is not in the Hilbert space due to its infinite norm. Any general vector expressed in terms of a continuous basis, such as $|\psi >= \int dx\, |x > \psi(x)$, or countably infinite basis such as $|\psi >= \sum_{i=1}^{\infty} |i > \psi_i$, is in the Hilbert space provided its norm is finite

$$< \psi|\psi >= \int dx\ |\psi(x)|^2 < \infty, \qquad \text{or} \qquad < \psi|\psi >= \sum_{i=1}^{\infty} |\psi_i|^2 < \infty. \tag{3.11}$$

Physical states must have finite norm since the norm is interpreted as probability that cannot exceed 1.

3.3 Observables, eigenstates, eigenvalues

An observable in Quantum Mechanics has a corresponding observable in Classical Mechanics. Examples of observables are position, momentum, angular momentum, spin, charge, isospin, color, etc. for every particle in a system.

According to the second postulate an observable A is a linear operator in the Hilbert space. This means that the action on a vector $|\psi>$ in the Hilbert space involves the first power of A, and the result of the action is another vector $|\chi>$ in the Hilbert space. A similar statement holds for the bra space, thus

$$A|\psi>=|\chi>, \quad <\phi|A=<\eta|,$$

such that both $|\psi>$ and $|\chi>$ (or $<\phi|$ and $<\eta|$) have finite norm. The action of the Hermitian conjugate operator is defined through Hermitian conjugation of the states $<\psi| \equiv (|\psi>)^\dagger$

$$<\psi|A^\dagger =<\chi|\,, \quad A^\dagger|\phi>=|\eta>\,. \tag{3.12}$$

Sandwiching the operator between a ket and a bra defines a complex number through the inner product defined in the previous section

$$\begin{array}{c} <\phi|A|\psi>=<\phi|\chi>=<\eta|\psi> \\ <\psi|A^\dagger|\phi>=<\chi|\phi>=<\psi|\eta>\,. \end{array} \tag{3.13}$$

Therefore the result is the same if one acts first on either the left or the right

$$<\phi|A|\psi>=<\phi|(A|\psi>)=(<\phi|A)|\psi>=(<\psi|A^\dagger|\phi>)^*, \tag{3.14}$$

and the last equality follows from the complex conjugation property of the inner product $<\phi|\chi>=(<\chi|\phi>)^*=\sum \phi_i^*\psi_i$.

It is possible to find a set of vectors $\{|\alpha>\}$ which are eigenvectors of the operator A. That is, the action of A reproduces the vector up to an overall constant which is called the eigenvalue. The eigenvector is labelled conveniently by the eigenvalue

$$A|\alpha_i>=\alpha_i|\alpha_i>, \quad <\alpha_i|A^\dagger=\alpha_i^*<\alpha_i|\,. \tag{3.15}$$

The second equation follows by Hermitian conjugation. The results of physical measurements are the eigenvalues. Since these have the same meaning as the classically observed quantities (i.e. position, etc.), the eigenvalues of observables must be real numbers. To guarantee this property the postulate requires that a physical observable is a Hermitian operator $A^\dagger = A$. Indeed, for a Hermitian operator one can show that the eigenvalues must be real, as follows. Consider

$$<\alpha_i|A|\alpha_j>=\begin{cases}\alpha_i^*<\alpha_i|\alpha_j> \\ \alpha_j<\alpha_i|\alpha_j>\end{cases}, \tag{3.16}$$

which is obtained by acting with the Hermitian operator to the left or the right. This equality leads to

$$(\alpha_i^*-\alpha_j)<\alpha_i|\alpha_j>=0. \tag{3.17}$$

Choosing the same states gives $(\alpha_i^*-\alpha_i)<\alpha_i|\alpha_i>=0$, which requires real eigenvalues $\alpha_i^*=\alpha_i$ since the norm cannot vanish. Furthermore, choosing different eigenvalues $\alpha_i \neq \alpha_j$ requires $<\alpha_i|\alpha_j>=0$, which shows that eigenstates

belonging to distinct eigenvalues must be *orthogonal*. There could be more than one state with the same eigenvalue. Such states are called *degenerate* eigenstates. One can always choose a basis in which the degenerate eigenstates are also orthogonal among themselves, and they are distinguished from each other with an additional label on the state. To keep our notation as simple as possible we assume at first that there are no degenerate states, and discuss them later. Since all possible measurements of an observable correspond to these eigenstates, they must be complete. Thus, we have the completeness and orthonormality conditions satisfied for the basis defined by the eigenstates of a Hermitian operator.

$$\sum_i |\alpha_i >< \alpha_i| = \mathbf{1}, \qquad < \alpha_i|\alpha_j >= \delta_{ij}, \tag{3.18}$$

where the sum runs over all the states (including the degenerate ones, if any).

In this basis one can define matrix elements of any operator B by sandwiching it between bras and kets

$$B_{ij} \equiv < \alpha_i|B|\alpha_j >, \tag{3.19}$$

and the operator itself may then be written as a sum over outer products, since $B = \mathbf{1}B\mathbf{1}$ gives

$$B = \sum_{\alpha_i,\alpha_j} |\alpha_i > B_{ij} < \alpha_j|. \tag{3.20}$$

In particular the operator A corresponds to a diagonal matrix in this basis since

$$A_{ij} =< \alpha_i|A|\alpha_j >= \alpha_i < \alpha_i|\alpha_j >= \alpha_i\delta_{ij} \tag{3.21}$$

It is evident that there is an isomorphism between the eigenstates of a Hermitian observable and the basis of vector spaces discussed in the previous section. If the basis of $|\alpha_i >$ eigenstates is countable, we may choose the isomorphism such that $|\alpha_i >\equiv |i >$ corresponds to a column vector in $n-$ dimensions as in (3.5). Then the operator A takes the diagonal matrix form

$$A = \sum_{\alpha_i} |\alpha_i > \alpha_i < \alpha_i| = \begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & \ddots \end{pmatrix} \tag{3.22}$$

and the normalized eigenstates are precisely the ones listed in (3.5).

If there are degenerate eigenstates, then there will be several α_1's or α_2's etc. on the diagonal. Then one may think of the above matrix as blocks containing the degenerate eigenvalues, such that the unit matrix within that block is multiplied by the corresponding eigenvalue. One reason for the occurrence of degeneracies is the presence of more than one compatible observables (e.g. $[A, B] = 0$), since the vector space is labelled by their simultaneous eigenvalues $|\alpha_i, \beta_j >$. Then necessarily there are degenerate eigenstates since for each

eigenvalue α_i there are many states labelled by the eigenvalues β_j. The vectors within each block of degenerate α's would then be distinguished by the eigenvalues of B.

3.4 Compatible and incompatible observables

3.4.1 Compatible observables and complete vector space

Given a physical system, there is a *maximal* number of compatible observables. These form a set of operators $\{A_1, A_2, \cdots A_N\}$ that commute with each other

$$[A_i, A_j] = 0. \tag{3.23}$$

In principle the number N corresponds to the number of degrees of freedom of a system in its formulation in Classical Mechanics. This corresponds to the number of canonical position-momentum pairs, plus spin, charge, color, flavor, etc. degrees of freedom, if any. Thus, for a system of k spinless neutral particles in d dimensions there are $N = k \times d$ degrees of freedom that correspond to compatible positions or momenta needed to describe the system. The operators $\{A_1, A_2, \cdots A_N\}$ could be chosen as the N positions or the N momenta or some other N compatible operators constructed from them, such as energy, angular momentum, etc.. As explained below, for each choice there is a corresponding *complete* set of eigenstates that define a basis *for the same physical system.* A physical state may be written as a linear combination of basis vectors for any one choice of basis corresponding to the eigenstates of compatible observables. Furthermore any basis vector may be expanded in terms of the vectors of some other basis since they all describe the same Hilbert space.

Let us consider the matrix elements of two compatible observables A_1 and A_2 in the basis that diagonalizes $(A_1)_{ij} = \alpha_{1i}\delta_{ij}$. The matrix elements of the zero commutator give

$$0 =< i|[A_1, A_2]|j >= (\alpha_{1i} - \alpha_{1j})\ (A_2)_{ij}. \tag{3.24}$$

Choosing states with different eigenvalues of the first operator $\alpha_{1i} \neq \alpha_{1j}$ we conclude that $(A_2)_{ij} = 0$. Therefore, the second operator must also be diagonal on the same basis $(A_2)_{ij} = \alpha_{2i}\delta_{ij}$. This argument is slightly different if there are degenerate eigenstates with same values of α_{1i}. Then the commutator argument only says that $(A_2)_{ij}$ is block diagonal, where the blocks are defined by the degenerate eigenvalues of A_1, as discussed at the end of the previous section. However, by an appropriate similarity transformation each block in A_2 can be diagonalized since similarity transformations do not change the blocks of degenerate eigenvalues of A_1. Therefore the result is that the basis can always be chosen such that the two compatible operators are simultaneously diagonal. The argument may now be repeated with the pairs (A_1, A_3) and (A_2, A_3), arriving at the result that the three operators are simultaneously diagonal, etc..

So, all observables in the complete maximal compatible set $\{A_1, A_2, \cdots A_N\}$ have common eigenstates on which they are simultaneously diagonal. It is wise

to label the eigenstates with their eigenvalues so that

$$|i> \Longrightarrow |\alpha_{1i_1}\ , \alpha_{2i_2}\ , \cdots > \tag{3.25}$$

corresponds to a possible measurement of the system. As a shorthand notation we will often use the symbol $|i>$ although we mean (3.25). Since these states represent *all possible measurements* that may be performed on the system they form a *complete basis*. Any state of the system may be expressed as a linear superposition of this basis. They can also be chosen orthonormal, as discussed above. Thus, we may write

$$\begin{array}{l} \sum_{\alpha_{1i_1}\ ,\alpha_{2i_2}\ ,\cdots} |\alpha_{1i_1}\ , \alpha_{2i_2}\ , \cdots >< \alpha_{1i_1}\ , \alpha_{2i_2}\ , \cdots | = \mathbf{1} \quad , \\ < \alpha_{1i_1}\ , \alpha_{2i_2}\ , \cdots |\alpha_{1j_1}\ , \alpha_{2j_2}\ , \cdots >= \delta_{\alpha_{1i_1}\ ,\alpha_{1j_1}}\ \delta_{\alpha_{2i_2}\ ,\alpha_{2j_2}}\ \cdots . \end{array} \tag{3.26}$$

All computations may be performed in terms of this basis. Thus, the matrix label "i" is replaced by a composite label "$i_1 i_2 \cdots$", and matrix elements may be denoted by

$$(B)_{ij} \equiv (B)_{i_1\ i_2 \cdots,\ j_1\ j_2 \cdots} \quad , \tag{3.27}$$

etc.. When the eigenvalues of observables are continuous the corresponding labels are also continuous, and matrix elements are then functions of those variables. For example consider an operator acting on position space, with $< x|A|x' >= A(x, x')$ which is a function of two variables. However, it is still useful to consider these functions as infinite dimensional "matrices".

3.4.2 Incompatible observables.

Let us now consider an observable B that is not compatible with the set $\{A_1, A_2, \cdots A_N\}$, that is $[B, A_i] \neq 0$. Then we may derive

$$(\alpha_{k\ i_k} - \alpha_{k\ j_k})\ (B)_{ij} =< i|[A_k, B]|j >= 0\ (?) \tag{3.28}$$

The right hand side may or may not be zero; this depends on the states i, j and operators A, B. Consider the subset of states for which the result is non-zero. In general when $i \neq j$ the eigenvalues of the A_k's are different, and the non-zero result on the right implies that B cannot be diagonal. Therefore, there cannot exist a basis in which incompatible observables would be simultaneously diagonal. If there were such a basis, they would commute by virtue of being diagonal matrices, and this contradicts the assumption.

Let us consider two incompatible observables and their respective sets of eigenstates

$$A|\alpha_i >= \alpha_i|\alpha_i >, \qquad B|\beta_i >= \beta_i|\beta_i > . \tag{3.29}$$

Since each set of eigenstates is complete and orthonormal, and they span the same vector space, they must satisfy

$$\begin{array}{ccc} \sum_k |\alpha_k >< \alpha_k| & = \mathbf{1} = & \sum_k |\beta_k >< \beta_k| \\ < \alpha_i|\alpha_j > & = \delta_{ij} = & < \beta_i|\beta_j > \end{array} . \tag{3.30}$$

This allows the expansion of one set in terms of the other by simply multiplying the state by the identity operator

$$|\alpha_i >= \sum_k |\beta_k >< \beta_k|\alpha_i >, \qquad |\beta_i >= \sum_k |\alpha_k >< \alpha_k|\beta_i > . \tag{3.31}$$

The expansion coefficients satisfy $< \beta_k|\alpha_i >= (< \alpha_i|\beta_k >)^*$ therefore they form a unitary matrix

$$U_{ki} =< \beta_k|\alpha_i > \tag{3.32}$$

since we can write

$$\begin{array}{ll} (UU^\dagger)_{kl} & = \sum_i U_{ki}U^*_{li} = \sum_i < \beta_k|\alpha_i > (< \beta_l|\alpha_i >)^* \\ & = \sum_i < \beta_k|\alpha_i >< \alpha_i|\beta_l >=< \beta_k|\beta_l > \\ & = \delta_{kl} \quad . \end{array} \tag{3.33}$$

We conclude that the bases of non-compatible observables are related to each other by a unitary transformation. The unitary operator that performs the transformation may be written as the outer product of the two bases. By construction we see that the operator U has the following properties:

$$U = \sum_l |\alpha_l >< \beta_l| \ , \quad |\alpha_i >= U|\beta_i >, \quad U_{ki} =< \alpha_k|U|\alpha_i >=< \beta_k|\alpha_i > . \tag{3.34}$$

In the case of multiple observables that form complete sets $\{A_1, A_2, \cdots, A_N\}$ and $\{B_1, B_2, \cdots, B_N\}$ that are not compatible, the unitary transformation from one basis to the other takes the form

$$U = \sum_{l_1\,, l_2\,, \cdots} |\alpha_{1\,l_1}\,, \alpha_{2\,l_2}\,, \cdots >< \beta_{1\,l_1}\,, \beta_{2\,l_2}\,, \cdots| \quad . \tag{3.35}$$

3.5 Measurement

3.5.1 Projection operators

The third postulate of Quantum Mechanics relates to the measurement process. To measure a system one must put it in an eigenstate of the observables that are being measured. If $|\phi >$ represents the quantities that we wish to measure, and the system is in a state $|\psi >$, then the measurement process collapses the state to $|\phi >< \phi|\psi >$. The probability that the system is found with the desired values of the observables is $| < \phi|\psi > |^2$. Therefore we may represent the measurement apparatus by the outer product

$$P_\phi = |\phi >< \phi|. \tag{3.36}$$

The act of measuring is then equivalent to applying this operator on the state

$$P_\phi \ |\psi >= |\phi >< \phi|\psi > . \tag{3.37}$$

The probability $| < \phi|\psi > |^2$ may be expressed as the norm of the new state obtained after the measurement, provided we normalize $|\phi >$ to one $< \phi|\phi >= 1$. Then *the norm of the collapsed state gives the probability* which may also be written as the expectation value of the measurement operator

$$|P_\phi \ |\psi > |^2 =< \psi|P_\phi^2|\psi >=< \psi|P_\phi|\psi >= | < \phi|\psi > |^2. \tag{3.38}$$

Since we have a complete set of states defined by the simultaneous eigenvalues of compatible observables as in (3.25) we may define the set of all possible measurements of the observables $\{A_1, A_2, \cdots, A_N\}$ by the measurement operators

$$P_i = |i >< i|. \tag{3.39}$$

We may think of P_i as a filter that projects the general vector $|\psi >$ along the basis vector $|i >$ and computes the component of the vector. Indeed, the P_i satisfy the properties of projection operators

$$P_i P_j = \delta_{ij} P_i, \quad \sum_i P_i = \mathbf{1}. \tag{3.40}$$

where the last sum follows from (3.26). If one wants to measure only observable A_1, and not the others, one must sum over everything compatible that is not measured. So, one sums over all the projection operators with a fixed eigenvalue of A_1

$$P_{i_1} = \sum_{\alpha_{2i_2}, \cdots} |\alpha_{1i_1} \ , \alpha_{2i_2} \ , \cdots >< \alpha_{1i_1} \ , \alpha_{2i_2} \ , \cdots| \quad . \tag{3.41}$$

This projection operator serves as a filter for the eigenvalue α_{1i_1} . The sum of such projectors is the full identity (3.26)

$$\sum_{i_1} P_{i_1} \ = \sum_{i_1, i_2, \cdots} P_{i_1, i_2, \cdots} = 1 \tag{3.42}$$

According to these rules, measuring position means multiplying the state by the projector $|x >< x|$, and obtaining the state $|x >< x|\psi >$. The coefficient $< x|\psi >= \psi(x)$ is the probability amplitude for finding the system at position x [1].Similarly, measuring momentum means multiplying the

[1] Since the norm of $|x >$ is infinite $\delta(0) = \infty$, rather than 1, we must divide out its norm if we wish to express the probability as the norm of the state $|x >< x|\psi >$ that results after the measurement. To be strictly correct, the measurement apparatus cannot be represented by $|x >< x|$ since the the state $|x >$ is not in the Hilbert space due to its infinite norm. One should really use a normalized wavepacket $|\psi_{x_0} >$ that cooresponds to a sharp probability distribution concentrated around the measured position x_0. Then the measurement apparatus is really $|\psi_{x_0} >< \psi_{x_0}|$ (which is more realistic anyway), and there are no problems with infinite norms. However, in computations it is much more convenient to use the position space vector $|x >$ as an idealization for the sharp wavepacket, and then one must understand that its infinite norm should be divided out. The same remarks apply to momentum measurements $|p >< p|$, or any other observable that has continuous eigenvalues.

state with $|p><p|$,and interpreting $<p|\psi>=\tilde{\psi}(p)$as the probability amplitude for finding the system at momentum p.The same idea applies to energy, angular momentum, spin, charge, etc. So, for example, the projector $|E,l,q><E,l,q|$represents the simultaneous measurement of the compatible observables energy, angular momentum and charge. If only one of these quantities is measured the corresponding projection operator is obtained by summing over the unmeasured eigenvalues.

If one performs a series of measurements one after the other, these filtering operators keep track of the system and compute the state after each measurement. Imagine the two possible series of measurements described by the setups below. where each box represents a measurement.

$$|\psi>\rightarrow \quad \underset{|1><1|\psi>}{[measure\ \#1]} \rightarrow \quad \underset{|2><2|1><1|\psi>}{[measure\ \#2]} \rightarrow \quad \underset{|3><3|2><2|1><1|\psi>}{[measure\ \#3]}$$

$$|\psi>\rightarrow \quad \underset{|1><1|\psi>}{[measure\ \#1]} \rightarrow \qquad \rightarrow \qquad \underset{|3><3|1><1|\psi>}{[measure\ \#3]}$$

Thus, after three measurements of observables described by the measuring systems associated with $|1>,\ |2>,\ |3>$, the system is found in the state

$$P_3P_2P_1|\psi>=|3><3|2><2|1><1|\psi>, \tag{3.43}$$

and the probability of finding those attributes in the state $|\psi>$ is given by the norm of the resulting state $P_3P_2P_1|\psi>$. Using $P_3^2=P_3$, it may be written as

$$<\psi|P_1P_2P_3P_2P_1|\psi>=|<3|2><2|1><1|\psi>|^2. \tag{3.44}$$

On the other hand if one only measures attributes described by $|1>$ and $|3>$, i.e. not measure in between the attributes described by $|2>$, then the resulting state and probability are

$$P_3P_1|\psi>=|3><3|1><1|\psi>, \qquad <\psi|P_1P_3P_1|\psi>=|<3|1><1|\psi>|^2. \tag{3.45}$$

The second result is unrelated to the first one, and in general it cannot be recovered from a sum of all possible measurements of attributes $|2>$. This is because of the decoherence or collapse of the state created by the act of measurement. This would not have happened in classical mechanics.

However, there are cases in which decoherence will not happen. For example if the series of measurements involve only compatible observables decoherence will not happen. Thus, if every attribute in $|2>$ corresponds to compatible observables with those in $|3>$ then $[P_2,P_3]=0$ (prove this statement for two compatible measurements represented by operators of the type (3.41) $P_{\alpha_1 i_1}$ and $P_{\alpha_2 i_2}$. see problem 1:). Then one can write

$$<\psi|P_1P_2P_3P_2P_1|\psi>=<\psi|P_1(P_2)^2P_3P_1|\psi>=<\psi|P_1P_2P_3P_1|\psi> \tag{3.46}$$

This means that after summing over all the probabilities for measuring all values of attributes $|2>$ we can recover the probability for not measuring attributes

$|2>$ at all (i.e. no filtering of any special attribute $|2>$) :

$$\sum_2 <\psi|P_1P_2P_3P_1|\psi> = <\psi|P_1P_3P_1|\psi> . \tag{3.47}$$

This explains in general why there is no decoherence in Classical Physics. It is because in the limit $\hbar \to 0$ all observables are compatible.

3.6 Uncertainty relations

The average value of an observable A in a physical state $|\psi>$ is defined by the "expectation value"

$$A_\psi \equiv <\psi|A|\psi> . \tag{3.48}$$

As an example consider the average position

$$x_\psi = <\psi|\hat{x}|\psi> = \int dx <\psi|\hat{x}|x><x|\psi| > = \int dx\, x\, |\psi(x)|^2, \tag{3.49}$$

which makes sense intuitively given the interpretation of $|\psi(x)|^2$ as a probability distribution. The mean square deviation from the average may be taken as a definition of the uncertainty in a measurement in the state $|\psi>$

$$(\Delta A)^2_\psi \equiv <\psi|(A - A_\psi)^2|\psi> . \tag{3.50}$$

This quantity is positive since it can be rewritten as the norm of the vector $(A - A_\psi)|\psi> \equiv |\psi_A>$. Consider two incompatible observables $[A, B] \neq 0$, and examine the product of the uncertainties in their measurements

$$\begin{aligned}
(\Delta A)^2_\psi\,(\Delta B)^2_\psi &= <\psi|(A - A_\psi)^2|\psi><\psi|(B - B_\psi)^2|\psi> \\
&= <\psi_A|\psi_A><\psi_B|\psi_B> \\
&\geq |<\psi_A|\psi_B>|^2 = |<\psi|(A - A_\psi)(B - B_\psi)|\psi>|^2 \\
&\geq |<\psi|\tfrac{1}{2i}[(A - A_\psi), (B - B_\psi)]|\psi>|^2 \\
&\geq |<\psi|\tfrac{1}{2i}[A, B]|\psi>|^2,
\end{aligned} \tag{3.51}$$

where we have used the following steps: (i) From line 2 to line 3 we used the Schwartz inequality for the vectors $|\psi_A>$ and $|\psi_B>$, (ii) from line 3 to line 4 we used the Schwartz inequality for complex numbers $\alpha = <\psi|(A - A_\psi)(B - B_\psi)|\psi>$ or $\alpha^* = <\psi|(B-B_\psi)(A-A_\psi)|\psi>$, i.e. $|\alpha|^2 \geq (\mathrm{Im}\,\alpha)^2 = [(\alpha-\alpha^*)/2i]^2$, and (iii) in the last line the c-number terms dropped from the commutators. Finally taking the square root we obtain the basic uncertainty relation that applies to any state $|\psi>$

$$(\Delta A)_\psi\,(\Delta B)_\psi \geq |<\psi|\frac{1}{2i}[A, B]|\psi>|. \tag{3.52}$$

If two observables are compatible it is possible that $(\Delta A)_\psi(\Delta B)_\psi = 0$ in some states $|\psi>$, but this is not necessarily true in every state $|\psi>$. However, if

they are incompatible there can be no state in which the product is zero. As an example, we may now apply this relation to the observables $\hat{x}$ and $\hat{p}$ and derive from first principles the well known Heisenberg uncertainty relation

$$\Delta x \Delta p \geq | < \psi | \frac{1}{2i} [\hat{x}, \hat{p}] | \psi > | = \frac{\hbar}{2} < \psi | \psi >= \frac{\hbar}{2}. \tag{3.53}$$

In this final expression the dependence on the state $|\psi >$ dropped out since the norm of any physical state is 1. Obviously this will always happen if the commutator of the two observables is a c-number, which is of course the case for any canonically conjugate pair of generalized "position" and "momentum". Therefore, for any such pair the uncertainty relation is the same as the basic Heisenberg uncertainty relation.

3.7 General solution to a QM problem

3.7.1 Time translations

Just as in Classical Mechanics, in Quantum Mechanics one wants to study the time development of the system once it is prepared in some initial state $|\psi, t_0 >$. Solving this problem is a central goal of physics, since this is how we understand a physical system and predict its behavior. As motivated in chapter 2, the time development of a state is done through the time translation operator

$$|\psi, t >= U(t, t_0) |\psi, t_0 > . \tag{3.54}$$

The state $|\psi, t >$ is the solution to the Schrödinger equation

$$i\hbar \partial_t |\psi, t >= \hat{H} |\psi, t > . \tag{3.55}$$

These two statements combined imply that the time translation operator satisfies the first order equation

$$i\hbar \partial_t U(t, t_0) = \hat{H} U(t, t_0) \tag{3.56}$$

which must be solved with the boundary condition $U(t_0, t_0) = 1$. It is easy to see that

$$U(t, t_0) = \exp[-i\hat{H}(t - t_0)/\hbar] \tag{3.57}$$

when the Hamiltonian is time independent. If the Hamiltonian depends on time, then the differential equation, combined with the boundary condition, can be put into the form of an integral equation

$$U(t, t_0) = 1 - i/\hbar \int_{t_0}^{t} dt_1 \, \hat{H}(t_1) \, U(t_1, t_0). \tag{3.58}$$

An equation of this type is solved by iteration, i.e. by replacing repeatedly the same form inside the integral

$$\begin{aligned} U(t,t_0) &= 1 - i/\hbar \int_{t_0}^{t} dt_1\, \hat{H}(t_1)\, \left[1 - i/\hbar \int_{t_0}^{t_1} dt_2 \hat{H}(t_2)\, U(t_2,t_0)\right] = \cdots \\ &= 1 - i/\hbar \int_{t_0}^{t} dt_1\, \hat{H}(t_1) + (-i/\hbar)^2 \int_{t_0}^{t} \int_{t_0}^{t_1} dt_1 dt_2\, \hat{H}(t_1)\, \hat{H}(t_2) + \cdots \\ &= \sum_{n=0}^{\infty} (-i/\hbar)^n \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n \quad \hat{H}(t_1) \cdots\, \hat{H}(t_n) \\ &= \sum_{n=0}^{\infty} \frac{(-i/\hbar)^n}{n!} \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n \quad T\left\{\hat{H}(t_1) \cdots\, \hat{H}(t_n)\right\} \\ &= T\left\{\exp[-i/\hbar \int_{t_0}^{t} dt_1 \hat{H}(t_1)]\right\} \quad . \end{aligned} \tag{3.59}$$

To be able to write a compact form we have introduced the time ordering operation, which instructs to order the operators according to their chronological order

$$T\{A(t_1)B(t_2)\} = \theta(t_1 - t_2)\; A(t_1)B(t_2) + \theta(t_2 - t_1)\; B(t_2)A(t_1)\;. \tag{3.60}$$

More generally

$$\begin{aligned} T\{A(t_1)B(t_2)C(3)\cdots\} &= A(t_1)B(t_2)C(3)\cdots \qquad t_1 \geq t_2 \geq t_3 \geq \cdots \\ &= B(t_2)A(t_1)C(3)\cdots \qquad t_2 \geq t_1 \geq t_3 \geq \cdots \\ &\quad \vdots \qquad\qquad\qquad\qquad \vdots \end{aligned}$$

Inside the time ordering bracket the order of operators can be changed as if they commute with each other at different times

$$T\{A(t_1)B(t_2)\} = T\{B(t_2)A(t_1)\}, \tag{3.61}$$

since this gives the same result according to its definition. This property allows us to write the steps in lines 4 and 5 in the above equation (see problem 3) by using the fact that $T\left\{\hat{H}(t_1)\cdots\, \hat{H}(t_n)\right\}$ is completely symmetric as a function of the time variables $t_1 \cdots t_n$. However, the meaning of these expressions really boils down to the series on the 3rd line.

Schrödinger and Heisenberg Pictures

We also need to discuss the time development of operators. In Quantum Mechanics there are several setups of time development. The one we have been discussing so far prescribes the time development of the states $|\psi, t> = U(t,t_0)|\psi,t_0>$. Operators are at time t_0 and their matrix elements may be computed in the Hilbert space at time t, for example, $<\chi,t|B(t_0)|\psi,t>$. This approach is called the Schrödinger picture. Another alternative is to peel off the

time dependence from the state and associate it with the operator by rewriting the same equation as follows

$$\begin{aligned} < \chi, t|B(t_0)|\psi, t > &=< \chi, t_0|U^\dagger(t,t_0)B(t_0)U(t,t_0)|\psi, t_0 > \\ &=< \chi, t_0|B(t)|\psi, t_0 > \end{aligned} \tag{3.62}$$

where we have defined a time dependent operator $B(t) = U^\dagger(t,t_0)B(t_0)U(t,t_0)$. The second approach produces the same matrix elements or probability amplitudes for measurements. This is called the Heisenberg picture. In the second picture we may keep the Hilbert space unchanged, but compute the time evolution of all observables by this rule. The Heisenberg picture parallels more the situation in classical mechanics in which there are no states, but there are time dependent observables.

To emphasize that the rules to be followed for time development are different in the two pictures, it is customary to attach additional labels on the states and observables : $|\psi >_S, B_S$ or $|\psi >_H, B_H$ where S, H stand for Schrödinger or Heisenberg pictures respectively. Choosing $t_0 = 0$ for convenience, the two pictures are related to each other by the unitary transformation $U(t) \equiv U(t,0)$

$$B_H(t) = U^\dagger(t)B_S U(t) \qquad |\psi, t >_S = U(t)|\psi >_H \ . \tag{3.63}$$

The time derivative of observables in the Heisenberg picture have a relation to dynamical equations of motion in classical mechanics. To explore this point we compute the time derivatives by using the equation of motion satisfied by $U(t,0)$

$$\begin{aligned} \partial_t B_H(t) &= [\partial_t U^\dagger(t)]B_S U(t) + U^\dagger(t)B_S[\partial_t U(t)] \\ &= \tfrac{i}{\hbar}\left(U^\dagger(t)H_S\ B_S U(t) - U^\dagger(t)B_S\ H_S U(t)\right) \\ &= \tfrac{i}{\hbar}\left(U^\dagger H_S U\ U^\dagger B_S U - U^\dagger B_S U\ U^\dagger H_S U\right) \\ &= \tfrac{i}{\hbar}\left(H_H B_H - B_H H_H\right) \\ &= \tfrac{i}{\hbar}[H_H(t), B_H(t)] \end{aligned} \tag{3.64}$$

where we have defined the Hamiltonian in the Heisenberg picture $H_H(t) = U^\dagger(t)H_S U(t)$. The Heisenberg Hamiltonian has the same form as the Schrödinger Hamiltonian $H_S = \hat{H}(\hat{x}_S, \hat{p}_S)$, but it is built from the Heisenberg position and momenta $H_H = \hat{H}(\hat{x}_H, \hat{p}_H)$ since

$$H_H = U^\dagger \hat{H}(\hat{x}_S, \hat{p}_S)U = \hat{H}(U^\dagger \hat{x}_S U, U^\dagger \hat{p}_S U) = \hat{H}(\hat{x}_H, \hat{p}_H).$$

Furthermore, the equal time commutation rules in the Heisenberg picture are the same as those of Schrödinger picture $[\hat{x}_H(t), \hat{p}_H(t)] = i\hbar$. This is proved by applying the similarity transformation to products of operators and inserting the identity operator $U^\dagger(t)U(t) = 1$ in between factors, e.g.

$$[\hat{x}_H(t), \hat{p}_H(t)] = [U^\dagger(t)\hat{x}_S U(t), U^\dagger(t)\hat{p}_S U(t)] = U^\dagger(t)[\hat{x}_S, \hat{p}_S]U(t) = i\hbar.$$

The last equation together with (3.64) are in one to one correspondance to the Poisson bracket formulation of classical mechanics and therefore one expects

to obtain equations of motion for Heisenberg observables that have the same form as those in classical mechanics (except for taking care that the orders of non-commuting operators in the resulting equations must not be changed). Indeed, as an example, if we apply it to the case of free particles we see that it produces the same equations of motion for the dynamical observables as in classical physics

$$\begin{array}{l} \partial_t \hat{x}_H = \frac{i}{\hbar}[\hat{H}_H, \hat{x}_H] = \frac{i}{\hbar}[\frac{\hat{p}_H^2}{2m}, \hat{x}_H] = \frac{\hat{p}_H}{m} \\ \partial_t \hat{p}_H = \frac{i}{\hbar}[\hat{H}_H, \hat{p}_H] = 0. \end{array} \tag{3.65}$$

Computation of time development

In most cases we only need to discuss the time independent Hamiltonian, so we will stick to the solution in (3.57). To actually perform the time translation of the states in the Schrödinger representation, one must first find the eigenstates of the Hamiltonian operator (we will avoid writing explicitly the extra index S for the Schrödinger representation)

$$\hat{H}|E_i >= E_i|E_i >, \tag{3.66}$$

and expand the general initial state in this energy basis

$$|\psi, t_0 >= \sum_{E_i} |E_i > c_{E_i} \ . \tag{3.67}$$

Then the time development follows from

$$|\psi, t >= U(t, t_0)|\psi, t_0 >= \sum_{E_i} |E_i > e^{-iE_i(t-t_0)/\hbar} c_{E_i} \ . \tag{3.68}$$

It is easy to verify that this solves the time dependent Schrödinger equation with the correct boundary condition.

3.7.2 Complete solution

We now clarify what it takes to have a complete solution to a general problem in Quantum Mechanics. We identify schematically three steps

1. Find a complete Hilbert space. This is defined by means of the set of eigenvectors $\{|\alpha_1, \alpha_2, \cdots >\}$ of a complete *maximal set* of compatible observables $\{A_1, A_2, \cdots\}$. Generally there are several convenient sets of compatible observables . Each such set defines a complete basis which may be convenient for different calculations (e.g., position space, versus momentum space, versus angular momentum space, versus energy space, etc.).

2. Learn how to expand one convenient basis in terms of others, and how to apply all the non-diagonal operators on the chosen basis. We say that

a complete solution for the system is obtained when all expansion coefficients $< \beta_1, \beta_2, \cdots | \alpha_1, \alpha_2, \cdots >$ are known and all operations of the type $B|\alpha_1, \alpha_2, \cdots >$, for any operator B and any $|\alpha_1, \alpha_2, \cdots >$ are computed, at least in principle. In this case one is able to find any desired matrix element of the type $< \alpha_1, \alpha_2, \cdots |B| \alpha_1', \alpha_2', \cdots >$ and then relate them to measurable properties.

3. To predict how a state $|\psi, t >$ will evolve with time one needs to consider the basis that includes the Hamiltonian $\widehat{H} = A_1$, as one of the operators in the set of compatible observables $\{\widehat{H}, A_2, \cdots\}$, so that one of the eigenvalues is the energy. Then the time evolution of the system is solved as discussed above. This is the fundamental reason for computing the eigenvalues and eigenstates of the Hamiltonian. To find the space-time interpretation of the system one needs to also compute the position space wavefunction, $\psi_{E_i}(x_1, x_2, \cdots) = < x_1, x_2, \cdots | E_i, \cdots >$, which is the probability amplitude. To be able to find the energy eigenvalues and eigenstates one may need to solve the differential form of the Schrödinger equation satisfied by this amplitude. In some cases algebraic methods may suffice to find the energy eigenvalues and/or the position space wavefunction as we will see in some examples.

To clarify this program let us see how it applies to solvable problems consisting of just free particles:

One free particle in one dimension

For the free particle this program has already been accomplished in chapter 2, but let us review the essential elements. There are just two basic operators $\hat{x}$ and $\hat{p}$. Therefore at step (1) there are two bases $\{|x >\}$ and $\{|p >\}$. At step (2) we compute $< x|\hat{p} = -i\hbar\partial_x < x|$ and derive the expansion coefficients $< x|p >= (2\pi\hbar)^{-1/2} \exp(ixp/\hbar)$. Furthermore, using these results we know how to compute the matrix elements of any function of the operators $F(\hat{x}, \hat{p})$ in the momentum or position basis. The answers to step (3) are easy because the Hamiltonian is a function of momentum

$$\hat{H} = \begin{cases} \hat{p}^2/2m & \text{non-relativistic} \\ \sqrt{\hat{p}^2c^2 + m^2c^4} & \text{relativistic,} \end{cases} \tag{3.69}$$

and therefore it is already diagonal in momentum space. That is, energy space is the same as momentum space in this case. Therefore the expansion coefficients $< x|p >=< x|E >$ are already computed. The general state may be expanded in the position basis or momentum basis (i.e. energy basis): $|\psi >= \int dx\ |x > \psi(x) = \int dp\ |p > \tilde{\psi}(p)$, and its time evolution is computed easily in momentum space. For definiteness let us concentrate on the non-relativistic particle (the reader may easily generalize the discussion to the relativistic case). Then

$$|\psi, t > \ = \int dp\ |p > \exp(-itp^2/2m\hbar)\ \tilde{\psi}(p)\ . \tag{3.70}$$

It is harder to compute it in position space

$$\begin{aligned} |\psi, t> &= \int dx\ e^{-it\hat{H}/\hbar}|x>\psi(x) \\ &= \int dx \int dx'\ |x'><x'|e^{-it\hat{H}/\hbar}|x>\psi(x) \\ &= \int dx'\ |x'>\psi(x',t) \end{aligned} \tag{3.71}$$

The final $\psi(x,t)$ is obtained by equating the last line to (3.70) and taking the inner product with $<x|$. This produces the Fourier transform

$$\psi(x,t) = \int \frac{dp}{\sqrt{2\pi\hbar}} \exp(ixp/\hbar - itp^2/2m\hbar)\ \tilde{\psi}(p)\ . \tag{3.72}$$

The space-time properties of this probability amplitude was discussed in the previous chapter, where it was interpreted as a wave packet.

Another interesting quantity is the *propagator* defined by

$$\theta(t)\ <x'|e^{-it\hat{H}/\hbar}|x> = \theta(t)\ <x'|x,t> \equiv G(x',x;t). \tag{3.73}$$

It has the interpretation of the probability amplitude, for a particle that was initially at position x, to be found at position x' after some time t. We have included the step function $\theta(t)$ in the definition since this interpretation is valid only for positive times $t \geq 0$. We know by construction that this quantity is a delta function at zero time

$$G(x',x;t=0) = <x'|x> = \delta(x'-x). \tag{3.74}$$

Namely, initially the particle is sharply located at $x' = x$. For later times, the probability amplitude is computed by inserting identity in terms of the momentum basis

$$\begin{aligned} G(x',x;t) &= \theta(t)\ \int dp\ <x'|e^{-it\hat{H}/\hbar}|p><p|x> \\ &= \theta(t)\ \int \tfrac{dp}{2\pi\hbar}\ e^{-itp^2/2m\hbar}e^{-ip(x-x')/\hbar} \\ &= \theta(t)\ \sqrt{\tfrac{m}{2\pi i\hbar\, t}} \exp\left(\tfrac{im(x-x')^2}{2\hbar t}\right)\ . \end{aligned} \tag{3.75}$$

The last integral is performed by the completion of squares method that was illustrated for wave packets in the previous chapter. We expect that, by construction, in the limit $t \to 0$ this expression should reduce to a delta function, and indeed it does (see problem 4). For larger times, the probability that the particle is at some other point increases, since the distribution gets broader. This is related to the spreading of wave packets (see problem 2 for the relativistic case).

The propagator $G(x',x;t)$ is actually the Green function for the Schrödinger equation for the free particle since it satisfies the equation

$$\left(i\hbar\partial_{t'} - \frac{\hbar^2}{2m}\partial^2_{x'}\right) G(x',x;\ t'-t) = i\hbar\delta(x'-x)\ \delta(t'-t). \tag{3.76}$$

This result may be seen in two ways. One may either check the differential equation by substituting the explicit result for $G(x', x; t)$ given above, or note that it must be satisfied by construction (use $< x'|\hat{p}^2 = -\hbar^2 \partial^2_{x'} < x'|$):

$$\begin{aligned}
&\left(i\hbar\partial_t - \tfrac{\hbar^2}{2m}\partial^2_{x'}\right) G(x', x;\ t) = \\
&\quad = \left(i\hbar\partial_t - \tfrac{\hbar^2}{2m}\partial^2_{x'}\right)\left(\theta(t)\ \ < x'|e^{-it\hat{H}/\hbar}|x>\right) \\
&\quad = i\hbar\delta(t) < x'|e^{-it\hat{H}/\hbar}|x> + \theta(t)\ \ < x'|(\hat{H} - \tfrac{\hat{p}^2}{2m})e^{-it\hat{H}/\hbar}|x> \\
&\quad = i\hbar\delta(t) < x'|x> \\
&\quad = i\hbar\delta(x' - x)\ \delta(t).
\end{aligned} \tag{3.77}$$

where we have used $\hat{H} = \hat{p}^2/2m$ and set $t = 0$ for the terms multiplying $\delta(t)$.

From (3.71) we see that the propagator gives the time evolution of the wavefunction

$$\psi(x, t) = \int dx'\, G(x', x;\ t - t_0)\, \psi(x', t_0) \tag{3.78}$$

This is integral is directly in position space and reduces to the same result as the momentum space integral in integral (3.72).

N free particles in 1-dimension

To see how the general program applies to a system of particles, let us consider N free non-relativistic particles. Let $x_i(t)$ represent the x-coordinate of the i-th particle at time t. The classical Lagrangian for the system is the total kinetic energy

$$L(x_i, \dot{x}_i) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{x}_i^2. \tag{3.79}$$

Using the standard procedure we define momenta $p_i = \partial L/\partial \dot{x}_i$, and derive the Hamiltonian which has the expected form

$$H(x_i, p_i) = \frac{1}{2}\sum_{i=1}^{N} \frac{p_i^2}{m_i}. \tag{3.80}$$

For relativistic particles the corresponding Hamiltonian is

$$H(x_i, p_i) = \sum_{i=1}^{N} \sqrt{p_i^2 c^2 + m_i^2 c^4}. \tag{3.81}$$

Quantum mechanics is defined for either case by the commutation rules of compatible and non-compatible observables

$$\begin{aligned}
[\hat{x}_i, \hat{p}_j] &= i\hbar\ \delta_{ij} \\
[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] &= 0.
\end{aligned} \tag{3.82}$$

We completely define the Hilbert space of a system of N free particles by giving one of the following Hilbert bases

$$\begin{array}{c} \{|x_1,\ldots,x_N>\} \\ \{|p_1,\ldots,p_N>\} \\ \{|E_1,\ldots,E_N>\} \end{array} \tag{3.83}$$

or any combination of coordinate, momentum and energy eigenvalues that are simultaneously observable, like

$$|x_1,p_2,x_3,E_4\cdots> . \tag{3.84}$$

The properties of orthogonality and completeness are given accordingly

$$\begin{array}{c} < x_1\ldots x_N \ |x_1^{'}\ldots x_N^{'} >= \delta(x_1-x_1^{'})\ldots\delta(x_N-x_N^{'}) \\ \int dx_1\cdots\int dx_N \quad |x_1\ldots x_N >< x_1\ldots x_N| = \mathbf{1} \ . \end{array} \tag{3.85}$$

We notice that we really are working in a direct product vector space $V_{TOT} = V_1\otimes V_2\otimes\cdots\otimes V_N$, where V_i is the vector space of particle i. So, the meaning of $|x_1,\cdots,x_N>$ is the direct product $|x_1>\otimes|x_2>\otimes\cdots\otimes|x_N>$. All the other definitions given before hold, provided we make the necessary modifications; in fact one has

$$< x_1\ldots x_N|p_1\ldots p_N >= \Pi_{i=1}^N < x_i|p_i >= \Pi_{i=1}^N \frac{e^{ip_ix_i/\hbar}}{\sqrt{2\pi\hbar}} \tag{3.86}$$

and the expansion of the position basis in terms of the momentum basis parallels the single particle case

$$\begin{aligned} |x_1\ldots x_N> \quad &= \int dp_1\cdots\int dp_N \quad |p_1\ldots p_N >< p_1\ldots p_N \ |x_1\ldots x_N> \\ &= (2\pi\hbar)^{-N/2}\int dp_1\cdots\int dp_N \ e^{-i\sum p_ix_{i/\hbar}} \ |p_1\ldots p_N> \end{aligned} \tag{3.87}$$

To find the time development we consider the momentum space basis $|p_1\cdots p_N>$ since the Hamiltonian is diagonal on it

$$\hat{H} \ |p_1\cdots p_N >= \left(\sum_{i=1}^N E_i(p_i)\right)|p_1\cdots p_N> . \tag{3.88}$$

where

$$E_i(p_i) = \begin{cases} p_i^2/2m_i & \text{non-relativistic} \\ \sqrt{p_i^2c^2+m_i^2c^4} & \text{relativistic} \end{cases} \tag{3.89}$$

Given any state, we compute its expansion in the energy basis $|\psi,t=0>= \int dp_1\cdots\int dp_N \ |p_1\cdots p_N> \ \tilde{\psi}(p_1\cdots p_N)$ and obtain its time development in the usual manner

$$\begin{array}{c} |\psi,t>= e^{-\frac{i}{\hbar}\ \hat{H}t}|\psi>= \\ = \int dp_1 \ e^{-itE_1(p_1)/\hbar}\int dp_N \ \ e^{-itE_N(p_N)/\hbar} \ |p_1\cdots p_N> \tilde{\psi}(p_1\cdots p_N). \end{array} \tag{3.90}$$

As before, this leads to a collective wave packet for all the particles

$$\psi(x_1, \cdots, x_N, t) = \int dp_1 \cdots \int dp_N \; e^{-it\sum_i E_i(p_i)/\hbar} \; \tilde{\psi}(p_1 \cdots p_N). \tag{3.91}$$

If there are no correlations in momentum space in the initial state, then we may write a specialized wavefunction $\tilde{\psi}(p_1 \cdots p_N) = \tilde{\psi}_1(p_1) \cdots \tilde{\psi}_N(p_N)$ which leads to a product of position space wave packets

$$\psi(x_1, \cdots, x_N, t) = \psi_1(x_1, t) \cdots \psi_N(x_N, t). \tag{3.92}$$

The propagator takes the form of products

$$\begin{aligned} G(x'_1 \cdots x'_N \; , \; x_1 \cdots x_N \, ; \, t) \quad &= < x'_1 \cdots x'_N \, | \, e^{-iHt/\hbar} \, | \, x_1 \cdots x_N \; > \\ &= \prod_{i=1}^{N} < x'_i \, | \, e^{-itH_i/\hbar} \, | \, x_i > \\ &= \prod_{i=1}^{N} G_i(x'_i \, , \, x_i \, ; \, t), \end{aligned} \tag{3.93}$$

where $G_i(x'_i \, , \, x_i \, ; \, t)$ is the free propagator for particle of mass m_i. It was given explicitly above for the non-relativistic particle in eq.(3.75).

In this discussion we assumed that all the particles are distinguishable (e.g. they have different masses). To discuss indistinguishable particles correctly we must take into account their properties under permutations and how the wavefunction behaves. Notice that when the masses of particles are equal the Lagrangian and/or the Hamiltonian remains invariant under the interchange of the particles. The wavefunction must be labelled by the quantum numbers of this symmetry. We will postpone this discussion until we learn about symmetries and their representations. However, we might as well mention that Nature chooses only one representation space for the permutation symmetry of identical particles. If the particles are identical bosons (integer spin) their wavefunction must be completely symmetric. If they are fermions (integer plus 1/2 spin) their wavefunction must be completely antisymmetric. Mixed symmetry states are not allowed for identical particles. So, for identical particles it is not possible to have uncorrolated wave packets of the type (3.92) since this is not either symmetric or antisymmetric. For example for two identical free bosons or fermions one must have

$$\begin{aligned} \psi_{bosons}(x_1, x_2, t) &= \tfrac{1}{\sqrt{2}} \left(\psi_1(x_1, t) \; \psi_2(x_2, t) + \psi_1(x_2, t) \; \psi_2(x_1, t) \right) \\ \psi_{fermions}(x_1, x_2, t) &= \tfrac{1}{\sqrt{2}} \left(\psi_1(x_1, t) \; \psi_2(x_2, t) - \psi_1(x_2, t) \; \psi_2(x_1, t) \right) \end{aligned} \tag{3.94}$$

So, there are automatically correlations even for free identical bosons or fermions. For example two fermions have zero probability for being in the same state since for $\psi_1 = \psi_2 = \psi$ the fermion wavefunction vanishes. This peculiar property of Nature is explained in terms of the fundamental principles of causality, Poincaré symmetry and quantization rules in Relativistic Quantum Field Theory (spin & statistics theorem), but in our formalism it has to be assumed.

1 free particle in 3-dimensions.

The Lagrangian of a free non-relativistic particle in three dimensions is given by its kinetic energy. Using bold characters to denote 3-dimensional vectors we have

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 = \frac{1}{2}m\sum_{i=1}^{3}\dot{x}_i^2 \tag{3.95}$$

This problem is mathematically the same as the one discussed above for 3 equal mass particles moving in 1-dimension. However the physical interpretation is quite different. As the vector notation suggests, there is a rotational symmetry since the Lagrangian depends only on the length of the velocity vector and not its direction. This will be discussed in more detail in later chapters. There is, of course, also a permutation symmetry among the coordinates, but we do not impose symmetry or antisymmetry properties on the wavefunction for the interchanges of the different coordinates, since the interpretation is not the same as identical bosons or fermions.

The vector spaces are labelled as $|\mathbf{x}>$ or $|\mathbf{p}>$. This has the same meaning as direct products

$$|\mathbf{x}>=|x_1,x_2,x_3>=|x_1>\otimes|x_2>\otimes|x_3>\,. \tag{3.96}$$

Therefore, following the same formalism as the multi-particle case we obtain the results for the expansion of the position basis in terms of the momentum basis

$$<\mathbf{x}|\mathbf{p}>=(2\pi\hbar)^{-3/2}\ \exp(i\mathbf{x}\cdot\mathbf{p}/\hbar) \tag{3.97}$$

and all other relevant quantities. It all boils down to the expressions of the single free particle in one dimension, except for substituting vectors everywhere one encounters a position or momentum symbol, and taking dot products among them. For example, for the non-relativistic particle, the free propagator in 3 dimensions is

$$G(\mathbf{x}',\mathbf{x};t) \quad =\theta(t)\ \left(\tfrac{m}{2\pi i\hbar\, t}\right)^{3/2}\exp\left(\tfrac{im(\mathbf{x}-\mathbf{x}')^2}{2\hbar t}\right) \tag{3.98}$$

For the relativistic case the results do not follow from the relativistic multi-particle formalism, since the Hamiltonians are different: For the three dimensional relativistic Hamiltonian we must use

$$H=\sqrt{\mathbf{p}^2c^2+m^2c^4} \tag{3.99}$$

and not eq.(3.81) with $i=1,2,3$. Therefore, the results are similar to those of the single relativistic particle in one dimension except for substituting vectors and taking dot products.

N free particles in 3-dimensions

For the non-relativistic case the Lagrangian is

$$L = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{x}}_i^2 = \frac{1}{2}\sum_{i=1}^{N} m_i [\dot{x}_{i1}^2 + \dot{x}_{i2}^2 + \dot{x}_{i3}^2] \tag{3.100}$$

This problem is mathematically equivalent to a set of $3N$ particles moving freely in 1-dimension. There is a rotational symmetry. To obtain all the relevant results it is sufficient to consider the one dimensional multiparticle case, substitute vectors everywhere and take dot products. For identical particles there is a permutation symmetry, therefore the wavefunctions would have to be completely symmetric for bosons and completely antisymmetric for fermions.

3.8 Matrix QM with 2 states

It is instructive to examine the Quantum Mechanics problem for a two state system, since its complete solution provides a model for clarifying many concepts. Furthermore, it has some rather important applications in various parts of physics.

First we need a basis. We assume that we have diagonalized an observable A . Following the general discussion let us assume that its normalized eigenstates are in one to one correspondence to the vectors $|i>, \quad i = 1, 2$

$$|1> = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |2> = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.101}$$

that are orthonormal and complete

$$\sum_{i=1}^{2} |i><i| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \quad 0) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \quad 1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.102}$$

Then the operator A is diagonal in this basis

$$A = \sum_{i=1}^{2} |i> \alpha_i <i| = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}. \tag{3.103}$$

The matrix elements of any other operator B are given by $<i|B|j> = B_{ij}$, and it can be represented in the form

$$B = \mathbf{1}B\mathbf{1} = \sum_{i,j} |i> B_{ij} <j| = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}. \tag{3.104}$$

The action of B on the vector space is given by

$$B|k> = \sum_{i,j} |i> B_{ij} <j|k> = \sum_{i,j} |i> B_{ik} \tag{3.105}$$

This action on the vector space is identical to matrix multiplication, i.e. $B|1>= |1>B_{11}+|2>B_{21}$ is the same as

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}=\begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix}. \tag{3.106}$$

Thus, any operator may be written as a 2x2 matrix, and its action on the vector space is well defined. An arbitrary state is

$$|\psi>=|1>\psi_1+|2>\psi_2=\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{3.107}$$

where ψ_1, ψ_2 are complex numbers.

Now consider the Hamiltonian operator. Let us assume it does not commute with the observable A, therefore it should take the form of a general Hermitian matrix

$$H=\begin{pmatrix} h_1 & h_3 \\ h_3^* & h_2 \end{pmatrix} \tag{3.108}$$

with h_1, h_2 real and h_3 a complex number. The time translation operator is the matrix

$$U(t,t_0)=\exp\left[\frac{i(t-t_0)}{\hbar}\begin{pmatrix} h_1 & h_3 \\ h_3^* & h_2 \end{pmatrix}\right]. \tag{3.109}$$

This matrix may be constructed by expanding the exponential series, computing the powers H^n as 2x2 matrices and re-summing the matrix elements (see problems 6,7). Then one would know how to act with U on any state $|\psi>$. However, the easier and standard approach is to go into a basis that diagonalizes H,

$$H|E_1>=E_1|E_1>, \quad H|E_2>=E_2|E_2>, \tag{3.110}$$

expand every state in terms of that basis, and then easily compute the time translation. This is an exercise in 2x2 matrix diagonalization. The result is most conveniently given in terms of the following parametrization $h_1=x+y$, $h_2=x-y$, $h_3=re^{i\phi}$, $\tan\theta=2|h_3|/(h_1-h_2)=r/y$, that is,

$$H=\begin{pmatrix} x+y & y\tan\theta\, e^{i\phi} \\ y\tan\theta\, e^{-i\phi} & x-y \end{pmatrix} \tag{3.111}$$

Then it can be verified that the eigenstates and eigenvalues of the Hamiltonian are

$$\begin{aligned} &E_1=x+\tfrac{y}{\cos\theta}, && E_2=x-\tfrac{y}{\cos\theta} \\ &|E_1>=\begin{pmatrix} \cos\theta/2\; e^{i\phi/2} \\ \sin\theta/2\;\; e^{-i\phi/2} \end{pmatrix}, && |E_2>=\begin{pmatrix} -\sin\theta/2\; e^{i\phi/2} \\ \cos\theta/2\;\; e^{-i\phi/2} \end{pmatrix}. \end{aligned} \tag{3.112}$$

This basis is complete and orthonormal

$$|E_1><E_1| + |E_2><E_2| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad <E_i|E_j> = \delta_{ij}. \tag{3.113}$$

An arbitrary state may be expanded in this basis $|\psi> = \sum |E_i> \psi_{(E_i)}$ where we may compute the expansion coefficients by taking the inner product of the new basis with the state in eq.(3.107)

$$\begin{aligned} \psi_{(E_1)} &= <E_1|\psi> = \psi_1 \cos\theta/2 \; e^{-i\phi/2} + \psi_2 \sin\theta/2 \; e^{i\phi/2} \\ \psi_{(E_2)} &= <E_2|\psi> = -\psi_1 \sin\theta/2 \; e^{-i\phi/2} + \psi_2 \cos\theta/2 \; e^{i\phi/2} \end{aligned} \tag{3.114}$$

The time translation of the state is then

$$|\psi, t> = |E_1> \psi_{(E_1)} \; e^{-i(t-t_0)E_1/\hbar} + |E_2> \psi_{(E_2)} \; e^{-i(t-t_0)E_2/\hbar} \tag{3.115}$$

which may be written as a two dimensional column after substituting the above expressions. In particular one may find the time development of the original basis $|i, t>$ by specializing the calculation above to $(\psi_1 = 1, \; \psi_2 = 0)$ or $(\psi_1 = 0, \; \psi_2 = 1)$ respectively

$$\begin{aligned} |1, t> = & \begin{pmatrix} \cos\left(\frac{y(t-t_0)}{\hbar \cos\theta}\right) - i\cos\theta \; \sin\left(\frac{y(t-t_0)}{\hbar\cos\theta}\right) \\ -ie^{-i\phi}\sin\theta \; \sin\left(\frac{y(t-t_0)}{\hbar\cos\theta}\right) \end{pmatrix} e^{-ix(t-t_0)/\hbar} \\ |2, t> = & \begin{pmatrix} -ie^{i\phi}\sin\theta \; \sin\left(\frac{y(t-t_0)}{\hbar\cos\theta}\right) \\ \cos\left(\frac{y(t-t_0)}{\hbar\cos\theta}\right) + i\cos\theta \; \sin\left(\frac{y(t-t_0)}{\hbar\cos\theta}\right) \end{pmatrix} e^{-ix(t-t_0)/\hbar} \end{aligned} \tag{3.116}$$

The matrix elements of the time translation operator are

$$U_{ij}(t, t_0) = <i|U(t, t_0)|j> = <i|j, t> . \tag{3.117}$$

From the above result we see that the operator

$$U(t, t_0) = \sum_{ij} |i> U_{ij} <j| = \sum_j |j, t><j| \tag{3.118}$$

may be written in the form of a matrix $U(t, t_0)$ whose first and second columns are precisely the column vectors of (3.116). One may then compute various quantities at later times, such as transition probabilities, expectation values, etc..In particular, the Green function is

$$G_{ij}(t, t_0) = \theta(t - t_0) < i|j, t> = \theta(t - t_0) U_{ij}(t, t_0). \tag{3.119}$$

Following the same steps as eq. (3.77), it is expected that, by construction, the matrix $G(t, t_0)$ must satisfy the matrix differential equation,

$$[(i\hbar\partial_t - H)G(t, t_0)]_{ij} = i\hbar\delta(t - t_0)\delta_{ij}. \tag{3.120}$$

The reader should verify this equation explicitly by inserting the matrices for H and $G(t, t_0)$.

3.9 Quantum Mechanical puzzles and answers

3.9.1 Schrödinger's cat

3.9.2 Einstein-Rosen-Podolsky paradox

3.9.3 Measuring the phase of the wavefunction

Aharonov-Bohm effect

Recent developments

3.9.4 Quantum Mechanics for the entire universe

3.10 PROBLEMS

1. Consider two compatible observables $[A, B] = 0$ that have simultaneous eigenstates $|\alpha_i, \beta_j >$. When one of the observables is not measured, the measurement apparatus is represented by the projection operators P_{α_i} or P_{β_j} defined by

$$P_{\alpha_i} = \sum_{\beta_j} |\alpha_i\,, \beta_j >< \alpha_i\,, \beta_j|, \quad P_{\beta_j} = \sum_{\alpha_i} |\alpha_i\,, \beta_j >< \alpha_i\,, \beta_j|. \quad (3.121)$$

   (a)Prove that these satisfy the properties of projection operators.

   (b) When two measurements are performed for a system in state $|\psi >$, first for the eigenvalue of A, and then for the eigenvalue of B, or in reverse order, are the results the same or different? What is your answer to part (b) if the observables are incompatible? (Hint: consider $[P_{\alpha_i}\,, P_{\beta_j}]$).

2. A and B are observables and they have one or more simultaneous eigenstates labelled by their eigenvalues as $|a, b >$.

   - If the two observables anti-commute $\{A, B\} = 0$, what can we conclude about their simultaneous eigenvalues a, b?
   - If the states $|a, b >$ are orthonormal and complete, can we always conclude that $[A, B] = 0$? If your answer is yes, prove it. If your answer is no, give a counter example.

3. What is the propagator $G(x', x; t)$ in the case of a massless relativistic particle? What is the behavior of this probability distribution as a function of time?

4. Using the properties of the time ordering prescription, prove the steps in eq.(3.59). Start by proving it for the $n = 2$ term that involves only two integrals, and give convincing arguments for the higher order terms.

5. The propagator $G(x', x; t)$ given in eq.(3.75) is expected to reduce to a delta function in the limit $t \to 0$. By noticing the rapid oscillatory behavior in the limit, prove that it indeed acts like a delta function by showing that its integral with any smooth function satisfies

$$f(x') = \lim_{t\to 0} \int dx\, G(x', x; t)\ f(x). \tag{3.122}$$

6. For the 2-state problem show that the propagator is the Green function by verifying that it satisfies the matrix differential equation of (3.120).

7. For any 2x2 real or complex matrix M show that the powers M^n may be rewritten as a linear combination of the matrix M and the matrix $\mathbf{1}$. Then verify that $M^n = \alpha_n \mathbf{1} + \beta_n M$, where the coefficients are the following functions of the trace $t \equiv tr(M)$ and determinant $d \equiv \det M$

$$\begin{aligned} \alpha_n &= \tfrac{-d}{2^{n-1}\sqrt{t^2-4d}} \left[\left(t + \sqrt{t^2-4d}\right)^{n-1} - \left(t - \sqrt{t^2-4d}\right)^{n-1} \right] \\ \beta_n &= \tfrac{1}{2^n\sqrt{t^2-4d}} \left[\left(t + \sqrt{t^2-4d}\right)^{n} - \left(t - \sqrt{t^2-4d}\right)^{n} \right]. \end{aligned} \tag{3.123}$$

8. Using the results of the previous problem compute the time translation operator $U(t, t_0)$ for the 2-state problem by re-summing the exponential series into the form of a 2x2 matrix. The parametrization of H given in eq.(3.111) simplifies the expressions. Show that you recover the same results as eqs.(3.116-3.118).

Chapter 4

INTERACTIONS

4.1 The framework

In non-relativistic Classical Mechanics the motion of interacting particles is governed by a Lagrangian that is equal to the total kinetic energy minus the potential energy, $L = \sum_i \frac{1}{2} m_i \dot{\mathbf{x}}_i^2 - V(\mathbf{x}_1, \cdots, \mathbf{x}_N)$. The standard canonical formalism gives the usual momenta, $\mathbf{p}_i = \partial L / \partial \dot{\mathbf{x}}_i = m \dot{\mathbf{x}}_i$, and the Hamiltonian is the total energy

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{x}_1, \cdots, \mathbf{x}_N). \tag{4.1}$$

The equations of motion are $\partial_t \mathbf{x}_i = \partial H / \partial \mathbf{p}_i = \mathbf{p}_i / m_i$ and $\partial_t \mathbf{p}_i = -\partial H / \partial \mathbf{x}_i = -\partial V / \partial \mathbf{x}_i$ (force). In Relativistic Mechanics interactions cannot be covariantly discussed in terms of potentials. One needs to consider field theories that are relativistically covariant. Therefore we will limit our discussion here to non-relativistic particle mechanics, or more generally to any Hamiltonian that may be written as a function of positions and momenta $H(\mathbf{x}, \mathbf{p})$.

In quantum Mechanics we promote the canonical variables to operators $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{p}}_i$ and impose the canonical commutation rules $[\hat{x}_{iI}, \hat{x}_{jJ}] = 0 = [\hat{p}_{iI}, \hat{p}_{jJ}]$, and $[\hat{x}_{iI}, \hat{p}_{jJ}] = i\hbar \delta_{ij} \delta_{IJ}$. Here we have used low case indices $i = 1, \cdots, N$ to identify the particle, and capital indices $I = 1, \cdots, d$ to denote the components of a vector in $d-$dimensions. This part is independent of the details of the Hamiltonian and is the same as the free particles case studied in previous chapters. Therefore there can be no difference in the properties of the position $|\mathbf{x}_1, \cdots, \mathbf{x}_N>$ or momentum $|\mathbf{p}_1, \cdots p_N>$ states as compared to free particles. We must have the same setup as before for the position and momentum bases. In particular, we recall that for a single particle in $d-$dimensions we must have

$$\begin{array}{cc} <\mathbf{x}|\hat{x}_I = x_I <\mathbf{x}|, & <\mathbf{x}|\hat{p}_I = -i\hbar\partial_{x_I} <\mathbf{x}| \\ \hat{p}_I|\mathbf{p}> = p_I|\mathbf{p}>, & \hat{x}_I|\mathbf{p}> = -i\hbar\partial_{p_I}|\mathbf{p}> \\ <\mathbf{x}'|\mathbf{x}> = \delta^{(d)}(\mathbf{x}'-\mathbf{x}) & <\mathbf{p}'|\mathbf{p}> = \delta^{(d)}(\mathbf{p}'-\mathbf{p}) \end{array} \qquad (4.2)$$
$$\int d\mathbf{x}\ |\mathbf{x}><\mathbf{x}| = \mathbf{1} = \int d\mathbf{p}\ |\mathbf{p}><\mathbf{p}|$$
$$<\mathbf{x}|\mathbf{p}> = (2\pi\hbar)^{-d/2}\, e^{i\mathbf{x}\cdot\mathbf{p}/\hbar}.$$

The form of the equations of motion of observables in the Heisenberg picture, $\partial_t A_H = i[\hat{H}_H, A_H]/\hbar$, must be the same as those of Classical Mechanics, as discussed in the previous chapter. Indeed, it is easy to check that positions and momenta in the Heisenberg picture satisfy equations of motion that have the same form as the classical equations written above:

$$\begin{array}{c} \partial_t(\hat{x}_{iI})_H = i[\hat{H}_H, (\hat{x}_{iI})_H]/\hbar = \frac{1}{m_i}(p_{iI})_H \\ \partial_t(\hat{p}_{iI})_H = i[\hat{H}_H, (\hat{p}_{iI})_H]/\hbar = -\frac{\partial}{\partial(\hat{x}_{iI})_H}V. \end{array} \qquad (4.3)$$

Therefore, we should expect that the expectation values of observables in the quantum mechanical system will behave somewhat similarly to the classical observables at least in a probabilistic sense. This helps us develop physical intuition about the quantum system from the behavior of the classical one. For an arbitrary potential V it is not easy to solve the equations of motion for the observables in the Heisenberg picture, and generally we must work our way through the diagonalization of the Hamiltonian operator to find the time development of the system in the Schrödinger picture as described generally in the previous chapter.

For a single particle the general problem we would like to solve is reduced to the eigenvalue equation in the Schrödinger representation

$$\hat{H}|E,\cdots> = E|E,\cdots>\,. \qquad (4.4)$$

The dots $\cdots$ correspond to eigenvalues of simultaneous observables that commute with the Hamiltonian. The additional labels represented by the dots are needed to specify a complete Hilbert space. We will see examples below. Furthermore, to give a space-time interpretation, we need to compute the probability amplitude $<\mathbf{x}|E,\cdots> \equiv \psi_{E\cdots}(\mathbf{x})$. There is a distinct wavefunction for each choice of compatible quantum numbers represented by the dots "$\cdots$", but for notational convenience we will suppress the extra dots until needed. To solve for the probability amplitude we sandwich the Hamiltonian between the states $<\mathbf{x}|$ and $|E,\cdots>$, and evaluate it by applying it to the ket or the bra

$$<\mathbf{x}|\hat{H}|E,\cdots> = \begin{cases} <\mathbf{x}|E,\cdots> E \\ \left[-\frac{\hbar^2}{2m}\boldsymbol{\nabla}_x^2 + V(\mathbf{x})\right] <\mathbf{x}|E,\cdots>\,. \end{cases} \qquad (4.5)$$

Therefore the probability amplitude must satisfy the time independent Schrödinger equation.

$$\left[-\frac{\hbar^2}{2m}\,\boldsymbol{\nabla}_x^2 + V\left(\mathbf{x}\right)\right]\psi_E\left(\mathbf{x}\right) = E\;\psi_E\left(\mathbf{x}\right) \quad . \tag{4.6}$$

An arbitrary state at the initial time $|\psi>$ may be expanded in the complete basis $|E,\cdots>$. Then, as discussed in the previous chapter, the probability amplitude $<x|\psi>=\psi(x,t)$ at later times will be given by

$$\psi(\mathbf{x},t) = \sum_{E,\cdots} c_{E\cdots}\;\psi_{E\cdots}(\mathbf{x})\;e^{-iEt/\hbar} \tag{4.7}$$

where $\psi_{E\cdots}(\mathbf{x})$ form the basis of energy eigenstates and $c_{E\cdots} =< E,\cdots|\psi>$ is a set of constants that characterize the initial state $|\psi>$. This structure is designed to be the solution to the time dependent Schrödinger equation

$$i\hbar\partial_t\psi(\mathbf{x},t) = \left[-\frac{\hbar^2}{2m}\boldsymbol{\nabla}_x^2 + V\left(\mathbf{x}\right)\right]\psi(\mathbf{x},t). \tag{4.8}$$

4.1.1 Conservation of probability

As always, the time development of an energy eigenstate

$$\psi_{E,\cdots}(\mathbf{x},t) \;= \psi_{E,\cdots}(\mathbf{x})\;e^{-iEt/\hbar} \tag{4.9}$$

gives a probability density which is stationary in time, i.e. $|\psi_E(\mathbf{x},t)|^2$ is time independent. However, the general superposition of energy eigenstates given above is not stationary, since

$$|\psi(\mathbf{x},t)|^2 = \left|c_{E_1}\;\psi_{E_1}\left(\mathbf{x}\right)\exp(-iE_1t/\hbar) + c_{E_2}\;\psi_{E_2}\left(\mathbf{x}\right)\exp(-iE_2t/\hbar) + \cdots\right|^2 \tag{4.10}$$

depends on time. On the other hand the total probability is conserved, i.e. $\int|\psi\left(\mathbf{x},t\right)|^2\,d\mathbf{x}$ is time independent for any state and any Hamiltonian. To see this consider the probability density $\rho(\mathbf{x},t)$ and the probability current $\mathbf{J}(\mathbf{x},t)$ defined by

$$\rho(\mathbf{x},t) = \psi^*(\mathbf{x},t)\psi(\mathbf{x},t), \qquad \mathbf{J}(\mathbf{x}) = \frac{-i\hbar}{2m}\left[\psi^*(\mathbf{x},t)\boldsymbol{\nabla}\psi(\mathbf{x},t) - \psi(\mathbf{x},t)\boldsymbol{\nabla}\psi^*(\mathbf{x},t)\right]. \tag{4.11}$$

The probability current may be thought of as the velocity ($\hat{\mathbf{p}}/m \to -i\hbar\boldsymbol{\nabla}/m$) "times" the probability density. By using the Schrödinger equation (4.8) for any potential it is easily seen that these satisfy the equation

$$\partial_t\rho(\mathbf{x},t) = \boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{x},t). \tag{4.12}$$

This is called the probability conservation equation because it can be used to derive that the total probability for finding the particle anywhere in space is time independent: i.e.

$$\partial_t\int d\mathbf{x}\;\rho(\mathbf{x},t) = \int d\mathbf{x}\;\boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{x},t) = \int d\Omega\;\mathbf{n}\cdot\mathbf{J}(\mathbf{x},t)|_\infty = 0, \tag{4.13}$$

where we used Stokes's theorem to write the volume integral as a surface integral at infinity, and we assumed that for any localized wave packet the wavefunction vanishes at infinity. Therefore the normalization of the state does not change in time, and can be consistently chosen to be 1

$$\int d\mathbf{x}\ \rho(\mathbf{x},t) =< \psi,t|\psi,t >=< \psi|\psi >= 1. \tag{4.14}$$

Of course the last equation is equivalent to the statement that the time translation operator is unitary

$$U^\dagger(t,t_0)U(t,t_0) = 1, \tag{4.15}$$

since $|\psi,t >= U(t,t_0)|\psi,t_0 >$. We see that conservation of probability and unitarity of the time evolution operator (or hermiticity. of the Hamiltonian) are intimately connected[1].

For a system of N particles the same arguments hold. The probability density $\rho(\mathbf{x}_1,\cdots,\mathbf{x}_N;t)$ and the currents for the individual particles $\mathbf{J}^{(i)}(\mathbf{x}_1,\cdots,\mathbf{x}_N;t)$, $i = 1,\cdots N$

$$\begin{aligned}\rho(\mathbf{x}_1,\cdots,\mathbf{x}_N;t) &= \psi^*(\mathbf{x}_1,\cdots,\mathbf{x}_N;t)\psi(\mathbf{x}_1,\cdots,\mathbf{x}_N;t) \\ \mathbf{J}^{(i)}(\mathbf{x}_1,\cdots,\mathbf{x}_N;t) &= \frac{-i\hbar}{2m}\left[\psi^*\boldsymbol{\nabla}^{(i)}\psi - \psi\boldsymbol{\nabla}^{(i)}\psi^*\right]\end{aligned} \tag{4.16}$$

satisfy the conservation law

$$\partial_t\rho(\mathbf{x}_1,\cdots,\mathbf{x}_N;t) = \sum_{i=1}^{N}\boldsymbol{\nabla}^{(i)}\cdot\mathbf{J}^{(i)}(\mathbf{x}_1,\cdots,\mathbf{x}_N;t). \tag{4.17}$$

(see problem). Then it follows that the total probability is conserved

$$\partial_t\int d\mathbf{x}_1\cdots d\mathbf{x}_N\ \ \rho(\mathbf{x}_1,\cdots,\mathbf{x}_N;t) = 0. \tag{4.18}$$

To discuss the main issue of finding the eigenstates and eigenvalues of the Hamiltonian we will need to develop methods starting with completely solvable problems of one particle in a potential in one dimension and then working our way toward more complicated problems with more particles and more dimensions.

4.2 Particle in a potential in 1 dimension

Let us first develop some physical intuition for the motion of particles in potentials. It would be very useful to consider as a model the motion of a particle in

[1] A dissipative system in which probability is not conserved may be described by a non-unitary time evolution operator, which in turn may correspond to a non-hermitian Hamiltonian.

a gravitational potential since it is easier to understand intuitively the effects of the gravitational forces. The mathematics is the same for many physical situations, therefore the results expected intuitively for the gravitational case can be carried over to other cases. In this way one roughly knows what to expect before plunging into the detailed mathematical formalism of Quantum Mechanics.

So, for our physical model, consider a particle (such as a car or a ball) moving along a road that goes through hills and valleys. The horizontal position of the particle is denoted by x, and its vertical position by y. The topographical shape of the road is described by some function $y = f(x)$. The gravitational potential energy is a linear function of the height, i.e. $V \sim y = f(x)$: When the car climbs a hill along the road in gains gravitational energy and when it comes down into a valley it loses gravitational energy. Therefore, its gravitational potential energy is a function of its location along the road, $V(x)$. This function must follow the actual shape of the hills and valleys. Therefore, the plot in Fig. (4.1) represents the shape of the hills and valleys as well as the potential energy of the particle.

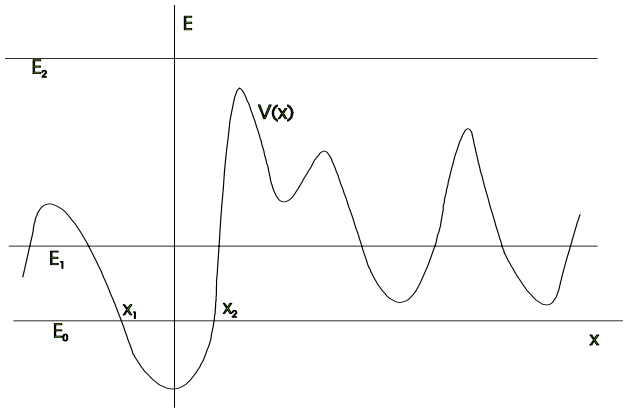


Fig.4.1 - Potential energy with hills and valleys.

The Hamiltonian is the sum of its kinetic and potential energy

$$H = p^2/2m + V(x). \tag{4.19}$$

Let us consider the motion of the particle in Classical Mechanics. Imagine a car, moving ideally without friction, that has total energy $H = E$ at the bottom of a valley located at $x = x_0$, the potential energy is $V(x_0)$, and the kinetic energy is $p^2/2m = E - V(x_0)$. Let the driver turn off his motor at the bottom of the valley. As any car driver knows intuitively, if the total energy E is large enough, it will manage to climb over the hill and reach the next valley, and then the next one, and so on. However, if the total energy is not sufficient, say $E = E_0$ as in the figure, it will climb up to a certain maximum height (or equivalently it will manage to go up to a maximum distance x_2 horizontally). It will momentarily stop at x_2 and then roll back down to the valley, pass the bottom at x_0, and climb the road on the other side. If the energy is not sufficient to go over the hill again, it will reach some x_1 and then return back and repeat

the process. In other words, if the energy is too low it will remain trapped in the valley and perform some sort of oscillatory motion between some extreme positions $x_1 \leq x \leq x_2$. At the extreme positions the kinetic energy vanishes and the total energy is purely potential energy. Thus, at energy level $E = E_0$ there is a relation $E_0 = V(x_2) = V(x_1)$ as pictured in the figure. On the other hand if the total energy is sufficiently large, the car can escape the pull of the valley and can go on travelling (as for $E = E_2$ in the figure). So, the general aspects of the motion are determined by the total energy. The mathematical equations of motion are the same for other physical situations described by the same sort of potential (e.g. electrons moving in atoms, molecules or solids). Therefore, the classical motions of particles will be intuitively the same for some given gravitational potential energy function $V(x)$.

As emphasized above, in Quantum Mechanics observables satisfy the same equations of motion as Classical Mechanics. Therefore, at least in a fuzzy probabilistic sense, one should anticipate that expectation values of observables in various states will follow the same intuitive behavior described in the previous paragraph. In the discussion above, the focus was on the observation of the position of the particle as a function of the energy. The connection should therefore be made with the probability amplitude for observing a particle at position x when it has energy E. That is, we should consider the probability amplitude $< x|E > \equiv \psi_E(x)$. The expected motion should be reflected in the behavior of the probability distribution $|\psi_E(x)|^2$. It should be large where the particle is allowed to be classically, and small where it is not allowed to be, as determined by its energy level. If the particle energy is very low, classically we expect it to spend most of its time close to the bottom of the valley, i.e. in the vicinity of x_0 in a narrow range $x_1 < x < x_2$. Therefore, we should expect that at low energies $|\psi_E(x)|^2$ will be largest near x_0. By the same token, as the energy increases we should expect $|\psi_E(x)|^2$ to be less peaked around x_0, and more spread out. Finally, if the energies are much higher than the peaks of the potential energy, then $|\psi_E(x)|^2$ should approach the behavior of free particles (i.e. wave packets) as discussed in the previous chapter, which permit the particle to travel along x. This intuitive behavior is born out by detailed computations as will be demonstrated in specific examples.

However, quantum mechanical details of the probability distributions that are non-intuitive classically, and which relate to interference phenomena of waves, cannot be obtained without solving for them for a given potential $V(x)$. For example, the quantum mechanical behavior of the particle at energy $E = E_1$ in the figure is very different than the classical behavior. The particle can tunnel from one valley to the next quantum mechanically although it cannot do it classically.

The time independent Schrödinger equation satisfied by the probability amplitude is

$$\left[-\frac{\hbar^2}{2m}\partial_x^2 + V(x)\right]\psi_E(x) = E\psi_E(x) \quad . \tag{4.20}$$

There are two linearly independent solutions to a second order differential equation. The solution is not unique unless one specifies boundary conditions. The boundary conditions must be consistent with the expected intuitive physical behavior discussed above. Thus, if the particle is trapped in the valley, its probability amplitude must vanish as x gets far away from the valley, i.e.

$$|\psi_E(x)| \to 0, \quad as \quad x \to \pm\infty. \tag{4.21}$$

On the other hand if the particle has more energy than any potential energy peak, then $\psi_E(x)$ should behave like a wave packet, i.e. a superposition of plane waves. In other words, the boundary conditions to be imposed depends on the energy.

In order to clarify these issues we will first discuss models that involve some $V(x)$ for which the differential equations are easy to solve. The detailed results serve as models to develop intuition for more complicated cases. Some of the solvable models may also be serious candidates for realistic physical applications.

4.2.1 Piecewise continuous potentials

The first set of models idealizes the valleys or hills with sharply shaped wells or barriers. The potential energy is therefore discontinuous (see Fig.(4.2)).

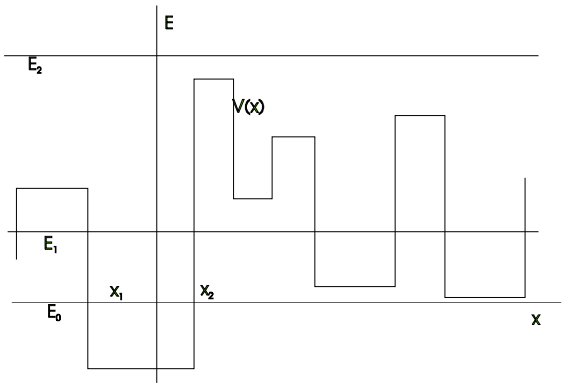


Fig.4.2 - Approximation to the hills and valleys of fig.4.1 with sharp wells and barriers.

By continuous deformation of the shape of the potential of Fig. (4.1) we can extend our intuitive discussion to apply to the case of Fig.(4.2). Thus we expect similar behavior classically and quantum mechanically.

Let us now begin the discussion of the solutions of the differential equation. The wavefunction $\psi_E(x)$ and its derivative $\partial_x\psi_E(x)$ must be continuous at the jumping points of the potential energy (provided it is not an infinitely large potential energy jump!). This continuity is justified by integrating the Schrödinger equation in the vicinity of a jump at $x = a$ as follows:

$$-\frac{\hbar^2}{2m}\int_{a-\varepsilon}^{a+\varepsilon} dx\ \partial_x^2\psi_{_E}(x) = \int_{a-\varepsilon}^{a+\varepsilon} (E-V)\ \psi_{_E}\ dx \underset{\varepsilon\to 0}{\longrightarrow} 0 \tag{4.22}$$

Since the right hand side must vanish as we shrink the region of integration ($\psi_{_E}(x)$ must not diverge since it is a probability amplitude), the integral on the left must tend to zero, thus

$$\partial_x\psi_{_E}(x)|_{a-\varepsilon}^{a+\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} 0. \tag{4.23}$$

Therefore, we conclude that the derivative is continuous. Next consider the first integral of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\ \partial_x\psi_{_E}(x) = -\frac{\hbar^2}{2m}\ \partial_x\psi_{_E}(x_0) + \int_{x_o}^{x} (E-V)\,\psi_{_E}\,(x')\,dx', \tag{4.24}$$

and integrate it once again in the vicinity of the jump . Again using a similar argument for the vanishing of the right side as the region of integration shrinks, one concludes that the wavefunction must also be continuous

$$\psi_{_E}(x)|_{a-\varepsilon}^{a+\varepsilon} \underset{\varepsilon\to 0}{\longrightarrow} 0\ . \tag{4.25}$$

If the jump in the potential energy is infinitely large, then the integrals in the vicinity of the jump will not vanish as $\varepsilon \to 0$. Accordingly, there may be discontinuities in the derivative of the wavefunction, or even in the wavefunction itself, depending on the nature of the jump.

Another way to arrive at the same conclusion is by writing the wavefunction in the form

$$\psi_{_E}(x) = \psi_{_E}^{(1)}(x)\ \theta(a-x) + \psi_{_E}^{(2)}(x)\ \theta(x-a) \tag{4.26}$$

where $\theta(x)$ is the step function. $\psi_{_E}^{(1,2)}(x)$ are the solutions of the Schrödinger equation to the left and to the right of the point $x = a$ where the potential function makes a jump. When this form is inserted into the differential equation, the derivatives are applied to the step functions as well. These produce extra terms containing a delta function and its derivative at $x = a$, i.e. $\partial_x\theta(x-a) = \delta(x-a)$, etc. These terms must vanish for the Schrödinger equation to be satisfied everywhere, including at the jumping point. So, the wavefunction and its derivative must be continuous at the discontinuities of the potential energy provided the potential energy is not itself singular. If the potential energy is singular (such as infinite steps, or delta functions, etc.) then they must cancel against the above terms.

It is convenient to consider the continuity for the ratio of the derivative $\partial_x\psi_{_E}(x)$ to the wavefunction $\psi_{_E}(x)$ since the overall normalization drops out for the combination $(\partial_x\psi_{_E}(x))/\psi_{_E}(x) = \partial_x \ln(\psi_{_E}(x))$. So, often one finds that

it is an easier procedure to impose the continuity conditions on the wavefunction and its logarithmic derivative in the form

$$\left.\frac{\psi_E^{(1)}(x)}{\psi_E^{(2)}(x)}\right|_{x=a} - 1 = 0, \qquad \partial_x \ln\left(\frac{\psi_E^{(1)}(x)}{\psi_E^{(2)}(x)}\right)_{x=a} = 0 \tag{4.27}$$

at each jumping point. Again, if the potential is singular the right hand side of these equations is modified accordingly (see problems 6-9). These conditions are peculiar to piecewise continuous potentials and they must be imposed to obtain legitimate solutions.

Infinite square well The potential energy is given by $V(x) = V_0\theta(|x| - a)$ with $V_0 \to \infty$, or

$$V(x) = \begin{cases} 0 & -a < x < a \\ +\infty & |x| > a \end{cases} \tag{4.28}$$

This corresponds to two infinitely high walls at $x = \pm a$, as in Fig.(4.3).

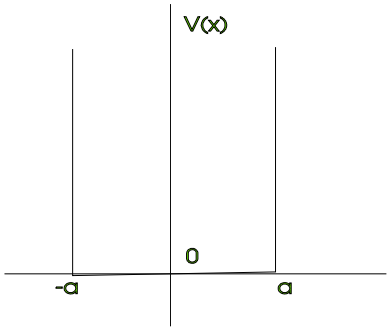


Fig.4.3 - Infinite square well

The classical motion consists of bouncing back and forth against the walls. The particle is trapped in a hole since the kinetic energy is never sufficient to jump over the infinitely high wall. Let us define the wavefunction in the various regions as follows

$$\psi_E(x) = \psi_E^L(x)\,\theta(-x-a) + \psi_E^0(x)\,\theta(a-|x|) + \psi_E^R(x)\,\theta(x-a)\ . \tag{4.29}$$

The Schrödinger equation outside the hole has the form

$$\left[-\frac{\hbar^2}{2m}\partial_x^2 + \infty\right]\psi_E^{L,R}(x) = E\,\psi_E^{L,R}(x)\,. \tag{4.30}$$

Because of the infinite potential the only way to satisfy this equation is

$$\psi_E^L(x) = 0 = \psi_E^R(x) \tag{4.31}$$

Inside the hole the equation is easily solved since $V = 0$

$$-\frac{\hbar^2}{2m}\,\partial_x^2\psi_E^0(x) = E\;\psi_E^0(x) \quad \rightarrow \qquad \psi_E^0(x) = A_E \sin\left(\frac{\sqrt{2mE}}{\hbar}\,(x - x_0)\right) \tag{4.32}$$

The required continuity conditions of the wavefunction at the boundaries are $\psi_E^0(-a) = \psi_E^L(-a) = 0$ and $\psi_E^0(a) = \psi_E^R(a) = 0$ (we do not expect continuity of the derivative since the potential makes an infinite jump). The first of these is satisfied by choosing $x_0 = -a$, then the second one imposes $\sin(\sqrt{2mE}2a/\hbar) = 0$, which is possible only with quantization conditions on the energy, $\sqrt{2mE}2a/\hbar = \pi n$. Therefore, the full solution is

$$\psi_{_E}(x) = \psi_{_n}(x)\;\theta(a - |x|), \qquad E_n = \frac{\hbar^2\pi^2 n^2}{8ma^2} \quad n = 1, 2, 3, \cdots \tag{4.33}$$

where

$$\psi_{_n}(x) = \sqrt{1/a}\;\sin\pi n(\frac{1}{2} - \frac{x}{2a}) = \left\{ \begin{array}{ll} i^{n-1}\sqrt{1/a}\;\cos\left(\frac{\pi n x}{2a}\right) & n = 1, 3, \cdots \\ i^{n-2}\sqrt{1/a}\;\sin\left(\frac{\pi n x}{2a}\right) & n = 2, 4, \cdots . \end{array} \right. \tag{4.34}$$

The probability distribution takes the form

$$|\psi_{_E}(x)|^2 = \theta(a - |x|)\;\frac{1}{2a}\;\left[1 - (-1)^n\cos(\pi n x/a)\right], \tag{4.35}$$

and the normalization $A_E = \sqrt{1/a}$ has been chosen so that $\int dx\,|\psi_{_E}(x)|^2 = 1$. A plot of the probability distribution confirms our fuzzy expectations based on Classical Mechanics (see Fig. (4.4)):

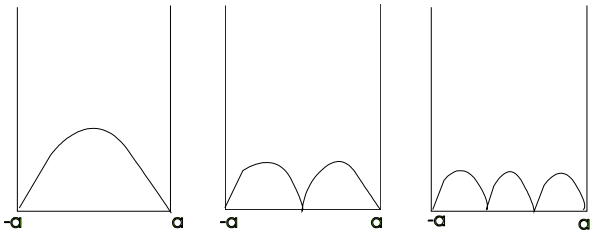


Fig.4.4 - Probability densities for n=1,2,3 eigenfunctions.

For small n the probability of finding the particle near the center $x = 0$ is greatest. As n increases, the probability is more distributed over the region. This may be understood in terms of the classical motions of the particle. During the same amount of time, a particle that is slow moving (low n, or low energy) is mostly found in the middle of the region, while a particle that is fast moving is likely to be found anywhere in the region. However, the detailed results provided by the probability distributions go well beyond any information that

may be extracted from Classical Mechanics. For example, (i) The energies are quantized and (ii) since the probability oscillates, there are positions of maximum and minimum probability. The locations of such extrema $x = x_i$ depend on the energy.

Energy-Parity representation

The solutions are naturally classified as even and odd under the parity transformation $x \to -x$. This is to be expected for the following reasons. In Quantum Mechanics we define the unitary parity operator S_P and its inverse S_P^{-1} such that

$$S_P\,\hat{x}\,S_P^{-1} = -\hat{x}, \quad S_P\,\hat{p}\,S_P^{-1} = -\hat{p}. \tag{4.36}$$

Then, any Hamiltonian with a potential energy that is an even function , $V(-x) = V(x)$, is automatically invariant under parity transformations,

$$S_P \hat{H} S_P^{-1} = +\hat{H}. \tag{4.37}$$

This means that the Hamiltonian commutes with the parity operator, $[\hat{H}, S_P] = 0$, and therefore they are simultaneous observables. The eigenvalues of the parity operator must label the complete set of states along with the eigenvalues of the Hamiltonian. What are the eigenvalues of the parity operator $S_P|\psi >= \lambda_P|\psi >$? First notice that its action on position space is $S_P|x >= |-x >$, as required by its action on the position operator. So, its square acts as the identity operator $S_P^2|x >= |x >$. Since position space is complete, S_P^2 is also identity on any state $S_P^2|\psi >= |\psi >$. Therefore the inverse of S_P is itself $S_P^{-1} = S_P$, and this requires that its eigenvalues satisfy $\lambda_P^2 = 1$. The only possibility is $\lambda_P = \pm 1$. Therefore in the present problem energy and parity eigenvalues combined provide a complete set of labels $|E, \pm >$ for a complete Hilbert space.

As we saw through the explicit solutions, we may summarize all the labels by the integer n, such that the energy is a function of n, while the parity eigenvalues are associated with even or odd integer. Therefore, we may write the operator equations in the energy-parity basis $|n >$ as follows

$$\hat{H}|n >= \frac{\hbar^2\pi^2 n^2}{8ma^2}|n >, \qquad S_P|n >= (-1)^{n-1}|n > . \tag{4.38}$$

This basis is orthonormal and complete by definition, just as the position basis is orthonormal and complete

$$\begin{array}{ll} < n|n' >= \delta_{nn'}, & < x|x' >= \delta(x - x') \\ \sum_{n=1}^{\infty} |n >< n| = \mathbf{1}, & \int_{-\infty}^{\infty} dx\ |x >< x| = \mathbf{1}. \end{array} \tag{4.39}$$

One complete basis may be expanded in terms of the other

$$|n >= \int_{-\infty}^{\infty} dx\ |x >< x|n >, \quad |x >= \sum_{n=1}^{\infty} |n >< n|x > \tag{4.40}$$

The consistency of these conditions can be checked explicitly (see problem 1) by inserting identity in between states and using the expansion coefficients $< x|n >=< x|E, \pm >= \psi_{E,\pm}(x)$ that were already computed above

$$< x|n >= \theta(a - |x|) \ \sqrt{\frac{1}{a}} \sin \pi n(\frac{1}{2} - \frac{x}{2a}). \tag{4.41}$$

In the energy-parity basis the time development of the states is given by

$$| \ n, t >= e^{-iHt/\hbar}|n >= \ |n > e^{-iE_n t/\hbar}, \qquad E_n = \frac{\pi^2\hbar^2 n^2}{8ma^2}. \tag{4.42}$$

The action of various operators on this basis may be computed by taking advantage of the completeness relations

$$\hat{x}|n >= \sum_{m=1}^{\infty} |m >< m|\hat{x}|n >= \sum_{m=1}^{\infty} |m > X_{mn} \tag{4.43}$$

where the matrix elements of the position operator X_{mn} in the energy-parity basis are given by

$$\begin{array}{ll} X_{mn} & =< m|\hat{x}|n > \\ & = \int_{-\infty}^{\infty} dx \ x \ < m|x >< x|n > \\ & = \int_{-a}^{a} dx \ x \ \left(\sqrt{1/a}\right)^2 \sin\left[\pi m(x/2a + 1/2)\right] \ \sin\left[\pi n(x/2a + 1/2)\right] \\ & = \frac{-16amn}{\pi^2(m^2-n^2)^2}\left[\frac{1-(-1)^{m-n}}{2}\right] \end{array} \tag{4.44}$$

Similarly the matrix elements of the momentum operator are

$$\begin{array}{ll} P_{mn} & =< m|\hat{p}|n > \\ & = \int_{-\infty}^{\infty} dx < m|x > (-i\hbar\partial_x) < x|n > \\ & = \frac{-i\hbar}{a} \int_{-a}^{a} dx \ \ \sin\left[\pi m(x/2a + 1/2)\right] \frac{\partial}{\partial x} \sin\left[\pi n(x/2a + 1/2)\right] \\ & = \frac{-2i\hbar mn}{a(m^2-n^2)}\left[\frac{1-(-1)^{m-n}}{2}\right] \end{array} \tag{4.45}$$

They may be written out in matrix notation

$$X_{mn} = -\frac{16a}{\pi^2}\begin{pmatrix} 0 & \frac{2}{9} & 0 & \frac{4}{225} & \cdots \\ \frac{2}{9} & 0 & \frac{6}{25} & 0 & \cdots \\ 0 & \frac{6}{25} & 0 & \frac{12}{49} & \cdots \\ \frac{4}{225} & 0 & \frac{12}{49} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.46}$$

and

$$P_{mn} = \frac{2i\hbar}{a}\begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{4}{15} & \cdots \\ -\frac{2}{3} & 0 & \frac{6}{5} & 0 & \cdots \\ 0 & -\frac{6}{5} & 0 & \frac{12}{7} & \cdots \\ -\frac{4}{15} & 0 & -\frac{12}{7} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.47}$$

Because the position or momentum operators are odd under parity their matrix elements vanish when the parity of both ket and bra are the same. This is the reason for the many zeroes in these matrices. So there is a "selection rule" due to the parity symmetry in this problem. There are several quantities that may be computed exactly, such as the mean square deviations in position and momentum in each state $(\Delta x)_n$ and $(\Delta p)_n$. These are left for the reader as exercises (see problems 2,3).

Finite square well

For the finite square well we will choose the origin of the energy axis such that the potential energy is either zero or negative. Then we have the potential energy

$$V(x) = -V_0 \; \theta(a - |x|) \tag{4.48}$$

corresponding to Fig.(4.5).

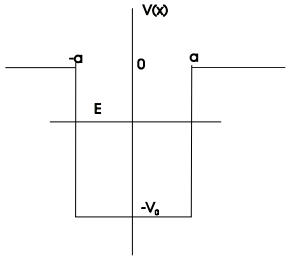


Fig.4.5 - Finite well.

The Schrödinger wavefunction may be written in the form

$$\psi_E(x) = \psi_E^L(x)\,\theta(-x-a) + \psi_E^0(x)\,\theta(a-|x|) + \psi_E^R(x)\,\theta(x-a) \tag{4.49}$$

where the various functions satisfy the Schrödinger equation in the left (L) right (R) and middle (o) regions

$$\begin{array}{ll} (-\frac{\hbar^2}{2m}\,\partial_x^2)\;\psi_E^{L,R} = E\;\psi_E^{L,R} & |x| > a \\ (-\frac{\hbar^2}{2m}\,\partial_x^2 - V_0)\;\psi_E^0 = E\;\psi_E^0 & |x| < a. \end{array} \tag{4.50}$$

For a bound state solution the energy level must be below the top of the well. With our definitions this means that we expect bound states for negative energies and unbound states for positive energies. First we discuss the bound state negative energies. We define the quantities

$$E = \frac{\hbar^2 k^2}{2m} - V_0 = -\frac{\hbar^2 K^2}{2m}, \quad \alpha = \sqrt{\frac{2mV_0}{\hbar^2}}, \quad K = \sqrt{\alpha^2 - k^2} \tag{4.51}$$

and write the solution in the form

$$\begin{array}{l} \psi_E^{even}(x) = A_+ \left(e^{-K(|x|-a)}\,\theta(|x|-a) + \frac{\cos(kx)}{\cos(ka)}\,\theta(a-|x|)\right) \\ \psi_E^{odd}(x) = A_- \left(e^{-K(|x|-a)}\,\varepsilon(x)\,\theta(|x|-a) + \frac{\sin(kx)}{\sin(ka)}\,\theta(a-|x|)\right) \end{array} \tag{4.52}$$

where $\varepsilon(x) = sign(x)$, and we have already imposed the continuity condition for the wavefunction at $x = \pm a$. Since the Hamiltonian commutes with the parity operator we have chosen our basis in terms of even and odd functions. At the boundaries of the well $x = \pm a$ the continuity of the logarithmic derivative gives (see (4.27))

$$\begin{aligned} \tan(ka) &= \sqrt{\tfrac{\alpha^2}{k^2} - 1} \qquad even\ \ solution \\ \cot(ka) &= -\sqrt{\tfrac{\alpha^2}{k^2} - 1} \qquad odd\ \ solution. \end{aligned} \tag{4.53}$$

These transcendental equations cannot be solved for k (or E) analytically, but an approximate graphical solution provides the essential physics (see Fig.(4.6)).

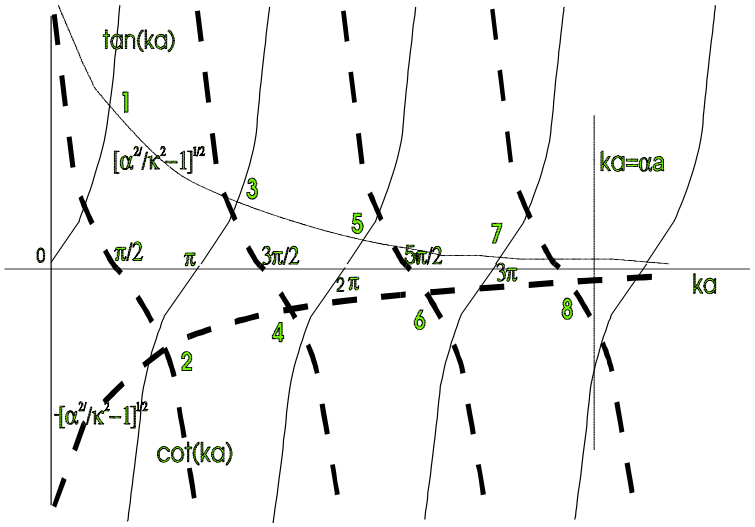


Fig.4.6 - Intersections of curves give the values of ak_n.

More accurate solutions are obtained numerically. The odd numbers in the figure mark the intersections of the two solid curves (even solution) and the even numbers mark the intersection of the two dashed curves (odd solution). At these values of $ka = k_n a$ there are values of the energy $E = E_n$ that solve the equations (4.53). These quantized energies E_n correspond to the bound state energy eigenvalues we are seeking. From the figure we see that the intersections occur when

$$(n-1)\frac{\pi}{2} < k_n a < n\frac{\pi}{2}, \quad \begin{cases} n = 1, 3, \cdots & even\ \ solution \\ n = 2, 4, \cdots & odd\ \ solution. \end{cases} \tag{4.54}$$

The number of bound state solutions N is given by the number of intersections of the two curves and is equal to the number of $\pi/2$ intervals that can fit between $ak_{\min} = 0$ and $ak_{\max} = a\alpha$ (see Fig. ( )). Therefore

$$N = 1 + \text{Int}\left((2a/\pi\hbar)\sqrt{2mV_0}\right) \tag{4.55}$$

where Int(x) is the largest integer contained in the real number x. Finally, the overall constants $A_\pm$ are computed by requiring an overall norm of 1, then

$$A_\pm = \frac{k_n}{\alpha}(a + 1/K_n)^{-1/2} \tag{4.56}$$

where (4.53) and (4.51) have been used to simplify the expression. We have thus found the simultaneous negative energy and parity eigenstates

$$\hat{H}|E_n >= E_n|E_n >, \quad P|E_n >= (-1)^{n-1}|E_n >, \quad n = 1, \cdots, N, \tag{4.57}$$

and computed the probability amplitudes in position space, which may be written neatly as

$$< x|E_n >= \frac{1}{\sqrt{a+1/K_n}} \quad \begin{array}{l} \left[\frac{k_n}{\alpha}\, e^{-K_n(|x|-a)}\, (\varepsilon(x))^{n-1}\ \theta(|x|-a)\ , \right. \\ \left. \quad + \sin\left(\frac{n\pi}{2} - k_n x\right)\ \theta(a-|x|)\right] \end{array} \tag{4.58}$$

where we have used

$$\sin(\frac{n\pi}{2} - k_n a) = k_n/\alpha, \quad \cos(\frac{n\pi}{2} - k_n a) = K_n/\alpha \tag{4.59}$$

as follows from (4.53) and (4.54).

According to (4.55), a deeper or wider well has a larger number of bound states. This makes sense intuitively.. It is interesting to consider some limits of either parameter. One limit is the infinite square well ($V_0 \to \infty$) and another limit is an infinitely narrow and deep one given by a delta function $V(x) = -v_0\delta(x)$ (see problems 5,6).

A new feature for the finite well is that the probability does not vanish in the classically forbidden region beyond the walls. Even though the particle does not have enough energy to jump over the wall, there is a probability to find it beyond the wall. However, this probability is exponentially decreasing with distance. We can estimate the wall penetration distance as the point at which the probability drops by a factor of e^{-1} as compared to its value at the wall boundary, that is

$$\frac{|\psi_n(a+\Delta)|^2}{|\psi_n(a)|^2} = e^{-\sqrt{2m|E_n|}2\Delta/\hbar} \sim e^{-1};\ \Rightarrow \quad \Delta \sim \frac{\hbar}{2\sqrt{2m|E_n|}} \quad . \tag{4.60}$$

Is it possible to perform an experiment to examine the particle while it is beyond the wall? This would be particularly interesting since the kinetic energy of the particle must be negative while it is in that region. Such an experiment would require an accuracy for the measurement of position $\Delta x < \hbar/(2\sqrt{2m|E_n|})$. Then by the uncertainty principle we must probe with momenta $p > \sqrt{2m|E_n|}$ and the uncertainty in the energy is $\Delta E > p^2/2m = |E_n|$. Therefore, it is not possible to perform such a measurement since the energy levels of the system would be destroyed. So, the negative kinetic energy of the particle cannot be observed.

We now consider the positive energy solutions. The complete set of states must include them since we cannot reconstruct the identity operator with only the finite number of bound states. That is, we need the positive energy eigenstates to write

$$\sum_{n=1}^{N} |E_n >< E_n| + \int_0^\infty dE \; |E >< E| \; = 1. \tag{4.61}$$

Classically the particle can move from minus infinity to plus infinity. Therefore in the quantum theory we expect oscillatory solutions corresponding to unbound particles for any positive energy. The boundary conditions to be imposed on the solutions must correspond to the physical process being analyzed. For example, particles may initially come in from the left, get scattered by the potential well and then get reflected to the left or transmitted to the right. Then the corresponding boundary conditions should be right-moving waves $\exp(i\kappa x)$ and left-moving waves $\exp(-i\kappa x)$ in the region $x < -a$ and right moving waves $\exp(i\kappa x)$ in the region $x > a$. With such conditions we can find the appropriate behavior for the probability amplitude $< x|E >$. This is left as an exercise for the reader (see problem 9). Note that the process just described is not left-right symmetric, therefore the energy eigenstate that describes it should be neither even nor odd under parity. There is, of course, another state which is the parity reflection of this state $P|E >$, which describes the mirror image of the physical process. Both of these states must be included in the completeness relation.

Barrier penetration

Let us now discuss the potential barrier given in Fig.(4.7).

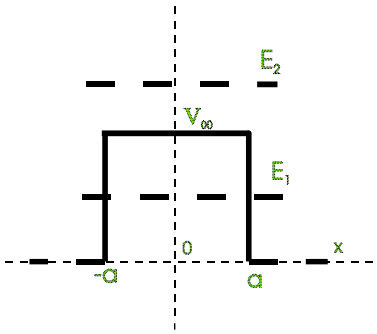


Fig.4.7 - Barrier

It is described by the potential energy function

$$V(x) = V_0 \; \theta(a - |x|). \tag{4.62}$$

In Classical Mechanics, for energies $E < V_0$ we expect that a particle moves freely until it hits the wall, at which point it is reflected and moves again freely in the opposite direction with the same kinetic energy. For energies $E > V_0$ it

goes forward after being slowed down temporarily by the barrier (it is easier to the intuition if one thinks of a smooth hill instead of the sharp barrier). In Quantum Mechanics transmission can happen even if $E < V_0$ through barrier penetration, but the probability of the transmission will be smaller as compared to the probability of the classically expected reflection. Similarly, when the energy is larger than the barrier energy $E > V_0$, there will be reflection, but at a smaller probability than the classically expected transmission. These phenomena occur because of the wave nature of Quantum Mechanics and are similar to the diffraction of ordinary light.

The Schrödinger wavefunction may be written in the form

$$\psi_E(x) = \psi_E^L(x)\,\theta(-x-a) + \psi_E^0(x)\,\theta(a-|x|) + \psi_E^R(x)\,\theta(x-a) \tag{4.63}$$

where the various functions satisfy the Schrödinger equation in the left (L) right (R) and middle (o) regions

$$\begin{array}{ll} (-\frac{\hbar^2}{2m}\,\partial_x^2)\;\psi_E^{L,R} = E\;\psi_E^{L,R} & |x| > a \\ (-\frac{\hbar^2}{2m}\,\partial_x^2 + V_0)\;\psi_E^0 = E\;\psi_E^0 & |x| < a. \end{array} \tag{4.64}$$

We define the quantities

$$E = \frac{\hbar^2 k^2}{2m} = -\frac{\hbar^2\kappa^2}{2m} + V_0\,, \quad \alpha = \sqrt{\frac{2mV_0}{\hbar^2}}, \quad \kappa = \sqrt{\alpha^2 - k^2}. \tag{4.65}$$

The solutions may be written as combinations of oscillatory plane waves

$$A_\pm(k)\exp(\pm ikx) \tag{4.66}$$

in the regions $|x| > a$, and combination of exponentials $B_\pm(\kappa)\exp(\pm\kappa x)$ in the region $|x| < a$. In order to choose the correct combinations we must consider the physical process we are trying to describe and impose boundary conditions consistent with it. For this we pay attention to the meaning of the plane wave solutions. We recall that the time evolution of a free particle is described by a wave packet of the form $\psi(x,t) = \int dp\;\phi(p)\exp[(ip(x-x_0) - iE(p)t)/\hbar]$ where $\phi(p)$ is a momentum distribution concentrated around $p = p_0$. Then the packet is concentrated in x-space around the position $x = x_0 + vt$ and moves with a group velocity $v = \partial E(p_0)/\partial p_0 = p_0/m$. For an idealized sharp momentum distribution $\phi(p) = A(p_0)\,\delta(p-p_0)$ the position wavefunction reduces to a plane wave

$$\psi(x,t) = A(p_0)e^{-i[p_0x_0 + tE(p_0)]/\hbar}\exp(ip_0x/\hbar). \tag{4.67}$$

The overall coefficient of the amplitude is closely related to the momentum-space (or energy-space) wavefunction. The $x-$dependence of the wavefunction, which is a plane wave, contains the information that the particle is moving with momentum p_0. It is this part that is relevant to our present discussion. In particular, the sign of p_0 tells us whether the particle moves to the left or right since it is related to the group velocity of the wavepacket $v = p_0/m$. Following this discussion we must interpret the plane wave solution $A_+(k)\exp(+ikx)$

as related to a right moving packet and $A_-(k)\exp(-ikx)$ as related to a left moving packet. In these solutions $p_0 = \pm\hbar k$ is a momentum determined by the energy eigenvalue E as in (4.65). Wavepackets can indeed be formed by taking superpositions of the energy eigenstates instead of the momentum eigenstates of the free particle.

Thus, if a particle begins its motion in the region $x < -a$ by moving toward the barrier, we expect that on the left of the barrier there will be both right and left moving waves (incoming and reflected waves), but on the right of the barrier there can be only right moving waves (transmitted waves). The last physical input, namely the absence of left moving waves in the region $x > a$ is the boundary condition we must impose. Accordingly we may write a solution of the form

$$\begin{aligned}\psi_E(x) = \quad & \theta(-x-a)\ A(k)\ [e^{ik(x+a)} + r(k)\ e^{-ik(x+a)}] \\ & +\theta(a-|x|)\ A(k)\ \tfrac{1+r(k)}{1+b(k)}\left[e^{\kappa(x+a)} + b(k)\ e^{-\kappa(x+a)}\right] \\ & +\theta(x-a)\ A(k)\ \tfrac{1+r(k)}{1+b(k)}\left[e^{2\kappa a} + b(k)\ e^{-2\kappa a}\right]\ e^{ik(x-a)}\end{aligned} \tag{4.68}$$

where the complicated coefficients are due to the continuity conditions at $x = \pm a$ which we have already imposed on the wavefunction. There remains the continuity of the logarithmic derivatives

$$\partial_x \ln\left(\frac{e^{ik(x+a)} + r(k)\ e^{-ik(x+a)}}{e^{\kappa(x+a)} + b(k)e^{-\kappa(x+a)}}\right)_{x=-a} = 0 = \partial_x \ln\left(\frac{e^{ik(x-a)}}{e^{\kappa(x+a)} + b(k)e^{-\kappa(x+a)}}\right)_{x=a} \tag{4.69}$$

where we have ignored the multiplicative constants since they drop out. These provide two equations from which we can solve for both $r(k)$ and $b(k)$.

$$b(k) = \frac{\kappa - ik}{\kappa + ik}e^{4\kappa a}, \quad r(k) = \frac{\alpha^2(1-e^{-4\kappa a})}{(\kappa-ik)^2 - (\kappa+ik)^2 e^{-4\kappa a}} \tag{4.70}$$

The transmission (or reflection) coefficient is defined as the ratio of the amplitudes multiplying the transmitted wave $\exp(ikx)$ (or reflected wave $\exp(-ikx)$) and the incoming wave $\exp(ikx)$. So, the reflection coefficient is $R(k) = r(k)e^{-2ika}$ and the transmission coefficient $T(k)$ is given by

$$\begin{aligned} T(k) &= \tfrac{1+r(k)}{1+b(k)}\left[e^{2\kappa a} + b(k)e^{-2\kappa a}\right]\ e^{-2ika} \\ &= -\tfrac{4ik\kappa\ e^{-2\kappa a}\ e^{-2ika}}{(\kappa-ik)^2-(\kappa+ik)^2e^{-4\kappa a}}. \end{aligned} \tag{4.71}$$

The probability of reflection (transmission) is the ratio of the probabilities for the reflected (transmitted) wave to the incident wave at the fixed energy eigenstate. Therefore these are given by $|R(k)|^2$ and $|T(k)|^2$. By probability conservation their sum must add to 1, since there is nothing else that can happen to the particle. Indeed we can explicitly check that

$$|R(k)|^2 + |T(k)|^2 = 1.$$

Consider the behavior of the transmission probability for a thick (large a) or tall (large V_0) barrier, defined by $\exp(-4\kappa a) \ll 1$

$$|T|^2 \sim 16\frac{k^2}{\alpha^2}\left(1-\frac{k^2}{\alpha^2}\right)e^{-4\kappa a} \tag{4.72}$$

As expected intuitively, the transmission is exponentially small in these cases. By contrast, a thin or low barrier $\exp(-4\kappa a) \approx 1$ gives

$$|T|^2 \approx e^{-4\kappa a} \tag{4.73}$$

indicating that it is easy to penetrate it. We may also consider the limit of a thin but tall barrier which is analogous to a delta function potential $V(x) = v_0\delta(x)$. This is left as an exercise (see problem 8).

For energies higher than the top of the barrier $E > V_0$ the solution is oscillatory everywhere. We may go through the algebra and discover that all expressions follow by analytic continuation of the energy from the region $E < V_0$ to the new region. This is equivalent to the analytic continuation of $\kappa = -iK$, where $K = (k^2 - \alpha^2)^{1/2}$. For example the transmission coefficient is now

$$T(k) = \frac{4kK\ e^{2iKa}\ e^{-2ika}}{(K+k)^2 - (K-k)^2\, e^{4iKa}} \tag{4.74}$$

At special energies the transmission coefficient becomes a pure phase indicating that transmission is 100%. This occurs when $4Ka = 2\pi n$, or when $E_n = V_0 + \hbar^2\pi^2 n^2/8ma^2$.

α-decay

A general barrier potential may be thought of as made up a large number of thin barriers as depicted in Fig.(4.8).

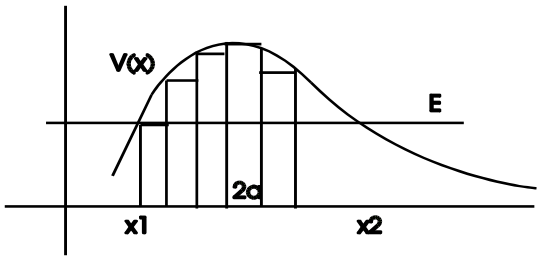


Fig.4.8 - General barrier composed of thin barriers.

The thin barrier centered at x_i has width $2a = dx$ and height $V_0 = V(x_i)$. According to the computation above the transmission probability through the i'th thin barrier is

$$|T_i(k)|^2 \approx e^{-4\kappa_i a} \approx \exp\left[-\frac{2dx}{\hbar}\sqrt{2m(V(x_i)-E)}\right]. \tag{4.75}$$

One may attempt to obtain an approximate expression for the transmission probability through the overall barrier by putting together the transmission probabilities through the individual thin barriers

$$|T(k)|^2 \approx \prod_i |T_i(k)|^2. \tag{4.76}$$

This is not entirely correct since it ignores the reflections forward and backward in between the thin barriers (see for example the double delta-shell potential, problem 9). However, one may hope that this is still a good approximation in certain cases. In fact the approximation turns out to be valid at sufficiently high energies as will be justified when we study the semi-classical WKB approximation. Thus, assuming that we have a valid approximation we may write

$$|T(E)|^2 = \exp\left[-\frac{2\sqrt{2m}}{\hbar}\int_{x_1}^{x_2} dx\sqrt{V(x)-E}\right] \tag{4.77}$$

where $x_{1,2}$ depend on the energy and are given by the relation $V(x_{1,2}) = E$, as seen in the Figure.

We can apply this result to learn something about the alpha decay of a nucleus. Within the range of strong interactions of a radius of about 1 fermi from the center of a nucleus, $r = a \approx 1\ f$, the main force is the attraction due to strong interactions. This may be represented by a well. When an α-particle moves out of this range the main force is the electromagnetic repulsion between the Z protons that remain in the nucleus and the two protons in the α. This is represented by the potential energy $V_{em}(x) = 2Ze^2/r$. The combined potential energy is depicted in Fig.(4.9).

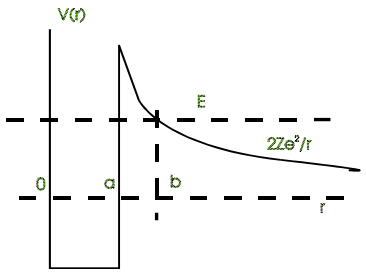


Fig4.9 - Nuclear + Coulomb potential for $\alpha-$ decay.

The probability of α decay of the nucleus is measured by the transmission probability of the α to the outside of the nucleus

$$|T(E)|^2 = \exp\left[-\frac{2\sqrt{2m_\alpha}}{\hbar}\int_a^b dr\sqrt{2Ze^2/r - E}\right] \tag{4.78}$$

where the lower limit on the integral is the range of the nuclear force, and the upper limit $r = b$ is given by $E = 2Ze^2/b$. Therefore

$$|T(E)|^2 = \exp\left[-\frac{2}{\hbar}\sqrt{2m_\alpha E}\int_a^b dr\sqrt{\frac{b}{r}-1}\right] \tag{4.79}$$

The integral may be performed by a change of variables $r = b\ \sin^2\theta$. The result simplifies under the assumption that $b \gg a$

$$|T(E)|^2 \approx \exp\left(-\frac{4\pi}{\hbar}\ \frac{Ze^2}{v_\alpha}\ + \frac{4}{\hbar}\sqrt{4Ze^2m_\alpha a}\right) \tag{4.80}$$

where $v_\alpha = \sqrt{2m_\alpha E}$ is the velocity of the α particle that emerges from the nucleus. The lifetime of the nucleus is inversely proportional to the decay probability $\tau^{-1} \sim |T(E)|^2$. This allows us to make the statement that

$$\ln\tau \sim \frac{c_1}{v_\alpha} + c_2 \tag{4.81}$$

where $c_{1=}\frac{4\pi}{\hbar}Ze^2$ is the Gamow factor. To compare this result to experiment we examine isotopes of heavy nuclei and plot their lifetimes versus the speed of the emitted $\alpha' s$. One finds good general agreement.

4.2.2 Harmonic Oscillator in 1d

The harmonic oscillator problem is one of the most important simple mathematical structures that has multiple applications in many areas of classical and quantum physics. Its applications range from molecular physics to string theory. It is used do describe vibrations of atoms in molecules, or electrons in crystal, and in quantum field theory is it is applied to phonons in solid state physics, photons in electrodynamics, quarks, leptons and other elementary particles. In this section we study the simple harmonic oscillator in one dimension which is at the basis of all the applications.

It is useful to keep in mind the classical problem of a particle attached to the end of a spring, which is the simplest physical system that obeys the laws of harmonic motion. As the spring is stretched (or contracted) away from equilibrium ($x = a$), the restoring force is proportional to the distance $F = -k(x-a)$, where k is the spring constant. The potential energy that obeys $F = -\partial V/\partial x = -k(x-a)$ is

$$V = \frac{1}{2}k(x-a)^2. \tag{4.82}$$

The classical equations of motion $m\ddot{x} = -k(x-a)$ have solutions that display oscillatory motion

$$x(t) = a + \frac{\sqrt{2mE}}{\omega}\sin(\omega t - \omega t_0) \tag{4.83}$$

where $\omega = \sqrt{\frac{k}{m}}$ and E is the total energy of the system. Raising the energy gives a larger amplitude of oscillations. Many other physical systems obey the same equations of motion and have the same vibrating solutions.

The quantum system is described by the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{4.84}$$

One must solve the eigenvalue problem

$$\left(\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\right)|E_n> = E_n|E_n> . \tag{4.85}$$

As usual, the energy basis $|E_n>$ or the position basis $|x>$ are complete, and one may expand one basis in terms of the other by using the expansion coefficients $< x|E_n > \equiv \psi_n(x)$. They represent the probability amplitude for the particle to be located at position x when it is in the energy eigenstate with $E = E_n$. We expect that the gross features of probability distributions described by wave packets

$$\psi(x,t) = \sum_m c_m(n) < x|E_m > \exp(-iE_m t/\hbar) \tag{4.86}$$

with $c_m(n)$ chosen to be sharply peaked around some energy $E = E_n$, is consistent with the classical physics described by the time dependence of the classical solutions given above.

There are several methods of computing the energy eigenstates or eigenvalues. The most elegant approach is an operator formalism that we will discuss in the next chapter. The most direct approach is to solve the Schrödinger equation for the wavefunction in position space

$$\left(-\frac{\hbar^2}{2m}\,\partial_x^2 + \frac{1}{2}m\omega^2x^2\right)\psi_n\left(x\right) = E_n\psi_n\left(x\right) \tag{4.87}$$

and impose the boundary condition $\psi(\pm\infty) = 0$ that is consistent with the physical situation. First we simplify the expressions by defining the dimensionless variable u

$$x = x_0 u, \quad \partial_x = \frac{1}{x_0}\,\partial_u, \tag{4.88}$$

The equation becomes

$$\left(-\frac{\partial^2}{\partial u^2} + u^2\right)\psi_n(u) = \varepsilon_n\psi_n(u) \tag{4.89}$$

provided we choose

$$x_0 = \sqrt{\frac{\hbar}{m\omega}}, \quad E_n = \frac{\hbar\omega}{2}\varepsilon_n \tag{4.90}$$

It is useful to first extract the leading behavior of the solution as $u \to \pm\infty$. Ignoring non-leading terms we see that $\psi(u) \sim \exp(-u^2/2)$ solves the equation and boundary conditions. Therefore, it is useful to extract this behavior by defining $\psi_n(u) = C_n H_n(u)\exp(-u^2/2)$, where C_n is a normalization constant, and the function $H_n(u)$ cannot grow faster than the exponential as $u \to \pm\infty$. Replacing this form in the differential equation one finds that $H_n(u)$ must obey

$$\partial_u^2 H_n - 2u\partial_u H_n + (\varepsilon_n - 1)\, H_n = 0. \tag{4.91}$$

Such equations may be solved by a power series

$$H_n(u) = \sum_0^\infty b_j^{(n)} u^j = b_0^{(n)} + b_1^{(n)} u + b_2^{(n)} u^2 + b_3^{(n)} u^3 + ... \tag{4.92}$$

Replacing it in the differential equation and collecting the coefficients of u^j, one finds that each such coefficient must vanish for every j

$$(j+2)(j+1) b_{j+2}^{(n)} - (2j - \varepsilon_n + 1) b_j^{(n)} = 0 \tag{4.93}$$

This provides a recursion relation that determines the coefficients

$$b_{j+2}^{(n)} = \frac{2j - \varepsilon_n + 1}{(j+2)\,(j+1)} b_j^{(n)}. \tag{4.94}$$

There are two solutions, one starting with b_0, the other starting with b_1

$$H_n^{even} = b_0^{(n)} + b_2^{(n)} u^2 + b_4^{(n)} u^4 + \cdots, \quad H_n^{odd} = b_1^{(n)} u + b_3^{(n)} u^3 + b_5^{(n)} u^5 + \cdots \tag{4.95}$$

To determine the asymptotic behavior we analyze the ratio of consecutive terms, with $j = 2m$ or $j = 2m+1$

$$\frac{b_{j+2}^{(n)} u^{j+2}}{b_j^{(n)} u^j} = \frac{2j - \varepsilon_n + 1}{(j+2)(j+1)} u^2 \underset{j\to\infty}{\longrightarrow} \frac{u^2}{m} \tag{4.96}$$

by comparing to the power series $\sum \frac{\left(u^2\right)^m}{m!} = e^{u^2}$, for which the ratio of consecutive terms is similar, we conclude that the series for $H_n(u)$ will grow faster than permitted for large u, unless the series is cutoff into a polynomial. This is accomplished by imposing

$$\varepsilon_n = 2n + 1, \tag{4.97}$$

so that b_{j+2} vanishes for some value of $j = n$, thus allowing H_n to be a polynomial. There is a solution for every integer n. Therefore the energy eigenstates and eigenvalues are

$$\hat{H}|n> = \hbar\omega\,(n + \frac{1}{2})\,|n>, \quad n = 0, 1, 2, \cdots \tag{4.98}$$

These must be orthonormal and complete

$$< n|n' > = \delta_{nn'}, \qquad \sum_{n=0}^{\infty} |n>< n| = 1. \tag{4.99}$$

Using the recursion relations one finds the coefficients for the even $n = 2N$ or odd $n = 2N+1$ solutions

$$b_{2k}^{(2N)} = \frac{(-4)^k N!\, b_0^{(2N)}}{(2k)!\ (N-k)!}, \quad b_{2k+1}^{(2N+1)} = \frac{(-4)^k (N)!\, b_1^{(2N+1)}}{(2k+1)!\ (N-k)!}, \quad k = 0, 1, 2, \cdots N \tag{4.100}$$

The result may be listed for a few of these solutions

$$H_0 = 1, \quad H_1 = 2u, \quad H_2 = 4u^2 - 2, \quad H_3 = 8u^3 - 12u, \quad \cdots \tag{4.101}$$

where the overall constants $b_0^{(n)}, b_1^{(n)}$ have been chosen so that the overall wavefunction is normalized to one

$$< n|n >= \int dx\, |\psi_n(x)|^2 = 1, \tag{4.102}$$

Therefore the explicit solution becomes

$$\psi_n(x) = \frac{1}{2^{n/2}}\, \frac{1}{\sqrt{n!}}\, \frac{1}{\sqrt{x_0\sqrt{\pi}}} e^{-x^2/2x_0^2} H_n(x/x_0). \tag{4.103}$$

The $H_n(u)$ are recognized as the Hermit polynomials. The quantum mechanical setup guaranties that these wavefunctions are orthonormal and complete. That is, the completeness and orthonormality of the energy basis $|n>$ and position basis $|x>$ require that the expansion coefficients $\psi_n(x) =< x|n >$ be orthonormal and complete

$$\int_{-\infty}^{\infty} dx\, \psi_n^*(x)\, \psi_{n'}(x) = \delta_{nn'}, \quad \sum_{n=0}^{\infty} \psi_n(x)\, \psi_n^*(x') = \delta(x - x'). \tag{4.104}$$

This implies that the Hermit polynomials must have these properties

$$\int_{-\infty}^{\infty} du\, e^{-u^2} H_n(u)\, H_{n'}(u) = 2^n n! \sqrt{\pi}\delta_{nn'}, \quad \sum_{n=0}^{\infty} \frac{H_n(u)\, H_n^*(u')}{2^n n!\sqrt{\pi}} = e^{u^2}\, \delta(u - u'). \tag{4.105}$$

By explicit computation this can be verified (see problem 12).

The probability densities $|\psi_n(x)|^2$ may be plotted. By comparing them to the probability densities for the infinite square well (Fig.(4.4)) one sees the similarities between them. Their interpretation is therefore analogous to that case. All of these results will be re-derived in the next chapter by using purely operator methods, without solving differential equations.

4.3 Problems

1. Show that eq.(4.17) follows from the Schrödinger equation for any potential energy.

2. Consider the infinite well problem. Show that the orthogonality and completeness conditions of eq.(4.39) are consistent by inserting identity between states and using the explicit solutions for the wavefunctions $< x|n >$ as computed in the text.

3. For the infinite well problem compute the mean square root deviations $(\Delta x)_n, (\Delta p)_n$ in every state $|n >$ for both position and momentum. Are they consistent with what you would expect intuitively from classical mechanics considerations? What is the product of the uncertainties, for which state is it a minimum?

4. Compute the propagator in the infinite well problem. Show it can be written in terms of Jacobi Theta functions, and discuss the physical properties by using the properties of the Theta functions.

5. Using the matrix elements X_{mn} and P_{mn} of position and momentum in the infinite square well problem, compute the commutator $[X, P]_{mn} = \sum_k (X_{mk}P_{kn} - P_{mk}X_{kn})$. Do you get the expected result?

6. Consider the finite square well with a potential energy function $V(x) = V_0\,\theta(|x| - a)$ (Fig.4.5). This is the same problem as the one considered in the text except for a shift in the energy axis to a new origin $E_{new} = V_0 + E_{old}$. Therefore all energy levels are positive and the solution for the eigenstates or energies are the same except for the shift. In the limit $V_0 \to \infty$ this becomes the infinite square well problem studied in detail in the text. Show that in this limit the bound state solutions of the finite well indeed tend to those of the infinite well. In particular discuss what happens to the continuity of the derivative?

7. Solve for the bound eigenstates and energies of the infinitely narrow and infinitely deep well represented by the potential energy $V(x) = -v_0\delta(x)$. Compare your direct solution to the one obtained as a limit for the square well $V = -v_0/2a\,\theta(a-|x|)$ which approximates the delta function potential in the limit $a \to 0$ (note that the integral over x is the same for these potentials).

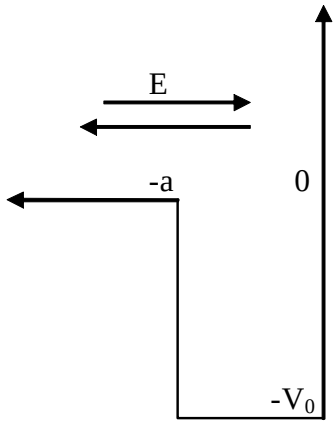
8. Consider the delta shell potential $V(x) = v_0\delta(x)$ with $v_0 > 0$. Compute the reflection and transmission coefficients and show the same result may be recovered as a limit of an infinitely tall and infinitely thin barrier (limit of Fig.4.7).

9. Consider the potential $V(x) = aV_0\delta(|x|-a)$ which describes two infinitely tall peaks at $x = \pm a$. Consider a particle that initially is moving to the *right* in the region $x < -a$. Compute the transmission probability to the region $x > a$. What is the largest value it can attain, at what energies?

10. Consider delta shell potentials that form two peaks $V = v_-\delta(x+a) - v_+\delta(x-a)$, and allow $v_\pm$ to have all possible signs. Compute the transmission coefficient. When is the reflection back and forth between the two peaks negligible? Consider the following limits (i) $a = 0$, (ii) $v_+ = 0$, (iii) $v_- = v_+$ (iv) $v_- = -v_+$, and discuss the agreement with results in the problems above.

11. Compute the transmission coefficient for the square well $V(x) = -V_0\,\theta(a-|x|)$. Show that the same result follows by analytic continuation of the transmission coefficient of the finite barrier $V(x) = V_0\,\theta(a-|x|)$ given in the text, by sending $V_0 \to -V_0$. Then notice that the transmission is 100% at certain energies, what are these energies ? (This is called the Ramsauer effect). What happens as V_0 gets large?

12. Consider a particle coming in from the left side at some positive energy E and then scattered from the one dimensional potential well shown in the figure



$V(x) = 0$ for $x < -a$; $V(x) = -V_0$ for $-a \le x < 0$; Infinite wall at $x = 0$.

a) Is it possible to obtain the reflection probability $|R|^2$ without detailed computation? If no, why not? If yes, what is it? Explain either way.

b) Compute the reflection coefficient R as a function of the energy and the parameters of the potential a, V_0. You may reparametrize E and V_0 in terms of some convenient quantities that simplify your equations while doing your work.

c) If you lower the energy E to negative values, discuss carefully the boundary conditions, and find an equation that determines the permitted values of the quantized energy (do not solve this equation for the energy eigenvalues). You can do this computation easily by modifying your algebra in part (b) appropriately.

d) Do you see any relationship between the equation you derive in part (c) and analytic properties of the reflection coefficient R as a function of the square root of energy $\sqrt{E} \sim k$ in the complex k plane? What is the property, and how is it explained in physical terms.

13. Uranium is bombarded by α particles, what is the probability for the α's to be captured?

14. Show that the Hermit polynomials may be rewritten in the forms

$$H_n = (-1)^n e^{u^2} \partial_u^n \; e^{-u^2} = e^{u^2} \frac{2^{n+1}}{\sqrt{\pi}} \int_0^\infty dt \; e^{-t^2} t^n \cos(2ut - n\pi/2).$$

Using these representations prove the orthogonality and completeness relations given in eq.(4.105)

15. Find the bound state eigenvalues and eigenfunctions for the attractive potential $V(x) = -\gamma^2/x^2$ and compare the behavior of your solutions to the infinite square well.

16. The relativistic generalization of the Schrödinger equation in 1-dimension is the Klein-Gordon equation. If a particle of mass m and charge q interacts with the electromagnetic field, gauge invariance dictates the following equation

$$[(\hbar\partial_t + iqA_0)^2 - (\hbar\partial_x + i\frac{q}{c}A_x)^2 + m^2]\phi(x,t) = 0.$$

Consider a positive energy solution $\phi(x,t) = \psi(x) \exp(-iEt/\hbar)$, and assume that there is no magnetic field, $A_x = 0$. Then the equation reduces to

$$\left[m^2 - (E - V(x))^2 - \hbar^2\partial_x^2\right] \; \psi(x) = 0$$

where $V(x) = qA_0(x)$. This equation is analogous to the 1-dimensional Schrödinger equation that we studied in this chapter. Note that the familiar non-relativistic limit is obtained by replacing $E \approx m + \hbar^2k^2/2m$ and keeping the leading terms for large m. One may analyse the solutions for any of the model potentials. In particular consider a particle trapped in a square well and analyse what happens as the well gets deeper and deeper. You should discover that the behavior is completely counter intuitive and paradoxical. This is called the Klein paradox.

Chapter 5

OPERATOR METHODS

Quantum Mechanics is formulated in terms of Hilbert spaces and operators acting on them. Beyond this, the quantum rules boil down to fundamental commutation rules for the canonical variables. It is often convenient to represent the canonical operators as differential or multiplicative operators, such as $\hat{p} \to -i\hbar\partial_x$, $\hat{x} \to x$, acting in position space, in order to compute the position probability amplitude $\psi(x) = \langle x|\psi\rangle$. However, Quantum Mechanics is actually defined in more abstract terms independent than position or momentum space. Only the basic commutation rules are sufficient, in principle, to solve a Quantum Mechanics problem, although this may not be the simplest method for a given Hamiltonian. However, there are a number of problems for which operator methods are actually much simpler and reveal the structure of the system much better. The harmonic oscillator in any number of dimensions is one of these cases, but there are quite a few more examples, such as angular momentum, the Hydrogen atom, and others. The harmonic oscillator is a fundamental tool that has a large number of applications in physics. Such applications are found in all branches of physics, ranging from condensed matter physics through molecular physics, nuclear physics, quantum field theory and particle physics, to superstring theory. Therefore, we will devote this chapter exclusively to developing the operator methods for the harmonic oscillator and applying them to the solution of a number of problems. In later chapters we will introduce operator methods for other systems.

5.1 Harmonic oscillator in 1 dimension

Recall the Hamiltonian for the simple harmonic oscillator

$$\hat{H} = \frac{1}{2m}(\hat{p}^2 + m^2\omega^2\hat{x}^2). \tag{5.1}$$

Consider the following combination of the canonical operators

$$a = \frac{1}{\sqrt{2}}(\frac{\hat{x}}{x_0} + i\frac{x_0}{\hbar}\hat{p}), \quad a^\dagger = \frac{1}{\sqrt{2}}(\frac{\hat{x}}{x_0} - i\frac{x_0}{\hbar}\hat{p}), \tag{5.2}$$

or equivalently

$$\hat{x} = \frac{x_0}{\sqrt{2}}(a + a^\dagger), \quad \hat{p} = -\frac{i\hbar}{\sqrt{2}x_0}(a - a^\dagger). \tag{5.3}$$

Using the basic commutation rules $[\hat{x}, \hat{p}] = i\hbar$, one finds

$$[a,\ a^\dagger] = 1, \ \ [a,\ a] = 0 = [a^\dagger,\ a^\dagger], \tag{5.4}$$

which hold for *any value* of x_0. If this form of $\hat{x}, \hat{p}$ is substituted into $\hat{H}$, the Hamiltonian becomes a linear combination of $a^2 + a^{\dagger 2}$ and $aa^\dagger + a^\dagger a$. However, if one chooses x_0 as

$$x_0 = \sqrt{\frac{\hbar}{m\omega}} \tag{5.5}$$

to kill the coefficient of $a^2 + a^{\dagger 2}$, then the Hamiltonian takes a simple form

$$H = \frac{\hbar\omega}{2}(aa^\dagger + a^\dagger a) = \hbar\omega(a^\dagger a + \frac{1}{2}) = \hbar\omega(\hat{N} + \frac{1}{2}), \tag{5.6}$$

where we have defined $\hat{N} \equiv a^\dagger a$. So, to find the eigenstates of $\hat{H}$ it is sufficient to find the eigenstates of $\hat{N}$. It is useful to further explore the commutators of $\hat{N}$ with a and $a^\dagger$

$$[\hat{N}, a] = -a, \quad [\hat{N}, a^\dagger] = a^\dagger. \tag{5.7}$$

Therefore, moving the operator $\hat{N}$ from one side of a to the other side gives $\hat{N}a = a(\hat{N} - 1)$, and similarly one finds $\hat{N}a^\dagger = a^\dagger(\hat{N} + 1)$. By repeatedly using this result one can find the rule for moving $\hat{N}$ through any power of a or $a^\dagger$, thus $\hat{N}a^n = a(\hat{N} - 1)a^{n-1} = a^2(\hat{N} - 2)a^{n-2} = \cdots = a^n(\hat{N} - n)$. So, one can write

$$\hat{N}a^n = a^n(\hat{N} - n), \quad \hat{N}a^{\dagger m} = a^{\dagger m}(\hat{N} + m), \quad \hat{N}a^n a^{\dagger m} = a^n a^{\dagger m}(\hat{N} - n + m). \tag{5.8}$$

Let us now assume that we have an eigenstate of $\hat{N}$ with eigenvalue λ

$$\hat{N}|\lambda\rangle = \lambda|\lambda\rangle,$$

and consider the states obtained by applying the operators $a^n|\lambda\rangle, a^{\dagger n}|\lambda\rangle$ on it. By applying $\hat{N}$ on these new states one finds that they are also eigenstates with eigenvalues $(\lambda - n)$ and $(\lambda + n)$ respectively

$$\hat{N}(a^n|\lambda\rangle) = a^n(\hat{N} - n)|\lambda\rangle = (\lambda - n)\ (a^n|\lambda\rangle)$$
$$\hat{N}(a^{\dagger n}|\lambda\rangle) = a^{\dagger n}(\hat{N} + n)|\lambda\rangle = (\lambda + n)\ (a^{\dagger n}|\lambda\rangle).$$

Thus each a annihilates a quantum of $\hat{N}$ and each $a^\dagger$ creates one. For this reason it is appropriate to call a the annihilation operator and $a^\dagger$ the creation operator. They act like ladder down or up operators respectively, moving one state to the neighboring state that differs by one unit of $\hat{N}$. Therefore we must conclude that the eigenvalues of $\hat{N}$ differ from each other by integers. Since $\hat{N}$ is a positive

operator, it must have a lowest eigenvalue $\lambda_0 \geq 0$. The set of eigenvalues must then be λ_0, $\lambda_0 + 1$, $\lambda_0 + 2, \cdots$. Furthermore, the state $a|\lambda_0\rangle$ which would have a lower eigenvalue $(\lambda_0 - 1)$ cannot exist by definition of λ_0, therefore we must impose $a|\lambda_0\rangle = 0$. However, using this condition in the eigenvalue equation for the lowest state $\lambda_0|\lambda_0\rangle = \hat{N}|\lambda_0\rangle = a^\dagger a|\lambda_0\rangle = 0$, shows that $\lambda_0 = 0$. Therefore we have the result that the eigenvalues of $\hat{N}$ are the positive integers $0, 1, 2, \cdots$, and the eigenstates are labelled by them as $|n\rangle$

$$\hat{N}|n\rangle = n|n\rangle, \quad n = 0, 1, 2, \cdots . \tag{5.9}$$

It is appropriate to call $\hat{N}$ the number operator. Each quantum of $\hat{N}$ is a quantum of energy since the energy eigenstates are the eigenstates of $\hat{N}$

$$\hat{H}|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle. \tag{5.10}$$

The lowest state, which is the energy ground state with $E_0 = \hbar\omega/2$, is annihilated by the operator a

$$a|0\rangle = 0. \tag{5.11}$$

The higher states are given by $|n\rangle = C_n a^{\dagger n}|0\rangle$, where C_n is a normalization constant. Therefore the ground state is often called the "vacuum state", meaning it has no excitations $\hat{N} \to 0$, and the excited states are then created out of the vacuum by the creation operators. The bra space is obtained by Hermitian conjugation $\langle n| = C_n^* \langle 0|a^n$. The bra vacuum is annihilated by $a^\dagger$

$$\langle 0|a^\dagger = 0. \tag{5.12}$$

The ground state is assumed to be normalized, $\langle 0|0\rangle = 1$, and the inner product for general states is

$$\langle m|n\rangle = C_m^* C_n \langle 0|a^m a^{\dagger n}|0\rangle. \tag{5.13}$$

To compute matrix elements of this type one uses the commutation rules to push the a's to the right and the $a^\dagger$'s to the left so that they annihilate the ground state. For example, for $n = m = 1$ we get

$$\begin{aligned} \langle 0|aa^\dagger|0\rangle &= \langle 0|a^\dagger a + [a, a^\dagger]|0\rangle \\ &= 0 + \langle 0|0\rangle \\ &= 1. \end{aligned} \tag{5.14}$$

For the more general case one uses the rule

$$[A, BC\cdots DE] = [A, B](C\cdots DE) + \cdots + (BC\cdots D)[A, E] \tag{5.15}$$

to compute

$$\begin{aligned} [a, a^{\dagger n}] &= [a, a^\dagger]a^{\dagger(n-1)} + \cdots + a^{\dagger(n-1)}[a, a^\dagger] \\ &= na^{\dagger(n-1)}. \end{aligned} \tag{5.16}$$

This may be rewritten as $aa^{\dagger n} = a^{\dagger(n-1)}(a^\dagger a + n)$, which gives the rule for moving one power of a from left to right. This may be used repeatedly to move any power of a or $a^\dagger$

$$\begin{array}{lll} a^m a^{\dagger n} = a^{\dagger(n-m)}\ (a^\dagger a + n - m + 1)\cdots(a^\dagger a + n - 1)(a^\dagger a + n) & if & m \leq n \\ a^m a^{\dagger n} = (a^\dagger a + m)(a^\dagger a + m - 1)\cdots(a^\dagger a + m - n + 1)\ a^{m-n} & if & m \geq n \end{array} \tag{5.17}$$

The vacuum expectation value of these expressions are now easy to obtain by using (5.11) and (5.12).

$$\begin{array}{lll} \langle 0|a^m a^{\dagger n}|0\rangle & = 0 \quad if \quad m \neq n \\ & = n! \quad if \quad m = n. \end{array} \tag{5.18}$$

This determines the normalization of the state

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \tag{5.19}$$

and shows that the number states are orthonormal

$$\langle m|n\rangle = \delta_{mn}. \tag{5.20}$$

We can now compute more precisely the action of the ladder operators on the states

$$\begin{array}{lll} a^\dagger|n\rangle & = \frac{(a^\dagger)^{n+1}}{\sqrt{n!}}|0\rangle = \sqrt{n+1}\frac{(a^\dagger)^{n+1}}{\sqrt{(n+1)!}}|0\rangle \\ & = \sqrt{n+1}|n+1\rangle \end{array} \tag{5.21}$$

Similarly

$$\begin{array}{lll} a|n\rangle & = \frac{aa^{\dagger n}}{\sqrt{n!}}|0\rangle = \frac{1}{\sqrt{n!}}a^{\dagger(n-1)}(a^\dagger a + n)|0\rangle = \sqrt{n}\frac{(a^\dagger)^{n-1}}{\sqrt{(n-1)!}}|0\rangle \\ & = \sqrt{n}|n-1\rangle. \end{array} \tag{5.22}$$

The last equation is consistent with the annihilation of the vacuum state for $n = 0$. The infinite dimensional Hilbert space that we have just constructed for the harmonic oscillator is called the Fock space. We may also compute the probability amplitude $\psi_n(x) = \langle x|n\rangle$ for finding the particle in the energy eigenstate $|n\rangle$ at position x. We will need to apply the creation-annihilation operators on the position basis

$$\begin{array}{lll} \langle x|a & = \langle x|\frac{1}{\sqrt{2}}(\frac{\hat{x}}{x_0} + i\frac{x_0}{\hbar}\hat{p}) = \frac{1}{\sqrt{2}}(\frac{x}{x_0} + x_0\partial_x)\langle x| \\ \langle x|a^\dagger & = \langle x|\frac{1}{\sqrt{2}}(\frac{\hat{x}}{x_0} - i\frac{x_0}{\hbar}\hat{p}) = \frac{1}{\sqrt{2}}(\frac{x}{x_0} - x_0\partial_x)\langle x| \end{array} \tag{5.23}$$

We start with the vacuum state, sandwich the operator a between the states and evaluate it by applying it to the left or the right

$$\langle x|a|0\rangle = 0 = \frac{1}{\sqrt{2}}(u + \partial_u)\langle x|0\rangle, \tag{5.24}$$

where we have used $u = x/x_0$. Therefore $\langle x|0\rangle = \psi_0(x)$ satisfies a first order differential equation which is solved by the normalized wavefunction

$$\psi_0(x) = \frac{1}{\sqrt{x_0\sqrt{\pi}}} e^{-u^2/2}. \tag{5.25}$$

The wavefunctions for the excited states are

$$\begin{aligned} \psi_n(x) &= \langle x|n\rangle = \langle x|\tfrac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \\ &= \tfrac{1}{\sqrt{n!}}\left(\tfrac{(u-\partial_u)}{\sqrt{2}}\right)^n \langle x|0\rangle \\ &= \tfrac{1}{\sqrt{x_0\sqrt{\pi}2^n n!}}\,(u-\partial_u)^n\, e^{-u^2/2}. \end{aligned} \tag{5.26}$$

This expression can be simplified by using the following trick. Note that for any function $f(u)$ we may write $(u-\partial_u)f = -e^{u^2/2}\left(\partial_u(e^{-u^2/2}f)\right)$. That is, the differential operator $(u-\partial_u)$ may be rewritten as

$$(u-\partial_u) = e^{u^2/2}(-\partial_u)e^{-u^2/2}, \tag{5.27}$$

where the derivative is applied on everything that follows on the right. Then the powers of this operator are

$$\begin{aligned} (u-\partial_u)^n = &\left(e^{u^2/2}(-\partial_u)e^{-u^2/2}\right)\left(e^{u^2/2}(-\partial_u)e^{-u^2/2}\right)\cdots\left(e^{u^2/2}(-\partial_u)e^{-u^2/2}\right) \\ &= e^{u^2/2}(-\partial_u)^n e^{-u^2/2}. \end{aligned} \tag{5.28}$$

Using this result one finds

$$(u-\partial_u)^n\, e^{-u^2/2} = (-1)^n e^{u^2/2}(\partial_u)^n e^{-u^2}, \tag{5.29}$$

which leads to

$$\psi_n = \frac{1}{\sqrt{x_0\sqrt{\pi}2^n n!}}\; e^{-u^2/2} H_n(u) \tag{5.30}$$

where $H_n(u)$ is the Hermit polynomial written in the form

$$H_n = (-1)^n e^{+u^2}\partial_u^n\; e^{-u^2}. \tag{5.31}$$

The result is identical to the wavefunction computed in chapter 4. The present formalism is well suited for many computations. For example the matrix elements of the position and momentum operators $\hat{x}, \hat{p}$ are

$$\begin{aligned} \langle m|\hat{x}|n\rangle &= \tfrac{x_0}{\sqrt{2}}\langle m|a+a^\dagger|n\rangle \\ &= \tfrac{x_0}{\sqrt{2}}\left(\langle m|n-1\rangle\sqrt{n} + \langle m|n+1\rangle\sqrt{n+1}\right) \\ &= \tfrac{x_0}{\sqrt{2}}\left(\sqrt{n}\;\delta_{m,n-1} + \sqrt{n+1}\delta_{m,n+1}\right) \\ &= \tfrac{x_0}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{aligned} \tag{5.32}$$

while for the momentum operator one gets

$$\langle m|\hat{p}|n\rangle = \frac{-i\hbar}{\sqrt{2}x_0}\left(\sqrt{n}\,\delta_{m,n-1} - \sqrt{n+1}\delta_{m,n+1}\right). \tag{5.33}$$

It is also easy to compute powers of these operators (see problem 1).

5.2 Coherent States

Coherent states are quantum mechanical wavepackets which describe nearly classical behavior of the harmonic oscillator, as will be explained below. They are defined by

$$|z\rangle = e^{za^\dagger}|0\rangle, \quad \langle z| = \langle 0|e^{az^*} \tag{5.34}$$

where z is a complex number. By expanding the exponential, and using the definition of the number states, one can rewrite them as superpositions of Fock space states

$$|z\rangle = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\,|n\rangle, \quad \Rightarrow \quad \langle n|z\rangle = \frac{z^n}{\sqrt{n!}}. \tag{5.35}$$

Using (5.16) one finds $[a, e^{za^\dagger}] = z\,e^{za\dagger}$, which can be used to show that the coherent state ket is an eigenstate of the annihilation operator; similarly the coherent state bra is an eigenstate of the creation operator

$$a|z\rangle = z|z\rangle, \quad \langle z|a^\dagger = z^*\langle z|. \tag{5.36}$$

Since a or $a^\dagger$ are not Hermitian, their eigenvalues are not real. Furthermore, the creation operator acts like a derivative operator on the ket

$$a^\dagger|z\rangle = a^\dagger e^{za^\dagger}|0\rangle = \partial_z|z\rangle, \tag{5.37}$$

and the number operator acts like the dimension operator

$$\hat{N}|z\rangle = a^\dagger a|z\rangle = z\partial_z|z\rangle, \quad \langle z|\hat{N} = z^*\partial_{z^*}\langle z|. \tag{5.38}$$

The Baker-Housedorf formula

$$e^A e^B = e^B e^A e^{[A,B]}, \tag{5.39}$$

which is valid when both A and B commute with $[A, B]$ (see problem 2 in Chapter 2), may be used to compute the inner product

$$\langle z|z'\rangle = \langle 0|e^{az^*}e^{z'a^\dagger}|0\rangle = \langle 0|e^{z'a^\dagger}e^{az^*}e^{z'z^*[a,a^\dagger]}|0\rangle = e^{z^*z'}. \tag{5.40}$$

Therefore, coherent states are not orthonormal, but they are complete

$$1 = \sum_{n=0}^{\infty}|n\rangle\langle n| = \frac{1}{\pi}\int d^2z\,e^{-|z|^2}|z\rangle\langle z|. \tag{5.41}$$

The completeness relation is proven by using (5.35) and performing the integrals in polar coordinates $z = r\exp(i\phi)$, $d^2z = rdrd\phi$. So, any state $|\psi\rangle$ may be expanded as a linear superposition of coherent states by multiplying it with the identity operator

$$|\psi\rangle = \int d^2z\ |z\rangle\psi(z,z^*), \qquad \psi(z,z^*) = \frac{1}{\pi}\, e^{-|z|^2}\langle z|\psi\rangle. \tag{5.42}$$

The expectation value of the position, momentum and energy operators in a coherent state are easily computed by using (5.36)and (5.40).

$$\begin{aligned}
\bar{x} &= \tfrac{\langle z|\hat{x}|z\rangle}{\langle z|z\rangle} = \tfrac{x_0}{\sqrt{2}\langle z|z\rangle}\langle z|a^\dagger + a|z\rangle = \tfrac{x_0}{\sqrt{2}}(z^* + z) \\
\bar{p} &= \tfrac{\langle z|\hat{p}|z\rangle}{\langle z|z\rangle} = \tfrac{i\hbar}{\sqrt{2}x_0\langle z|z\rangle}\langle z|a^\dagger - a|z\rangle = \tfrac{i\hbar}{\sqrt{2}x_0}(z^* - z) \\
\bar{E} &= \tfrac{\langle z|\hat{H}|z\rangle}{\langle z|z\rangle} = \tfrac{1}{\langle z|z\rangle}\langle z|\hbar\omega(a^\dagger a + 1/2)|z\rangle = \hbar\omega(|z|^2 + 1/2).
\end{aligned} \tag{5.43}$$

where we needed to divide by the norm since the coherent state is not normalized (see (5.40)). By solving this equation one may interpret the complex number z in more physical terms by relating it to the average position and momentum.

$$z = \frac{1}{\sqrt{2}}(\frac{\bar{x}}{x_0} + i\frac{x_0}{\hbar}\bar{p}). \tag{5.44}$$

Furthermore, the average energy in the coherent state may be rewritten in the suggestive classical form

$$\bar{E} = \frac{\bar{p}^2}{2m} + \frac{m\omega^2\bar{x}^2}{2} + E_0 \tag{5.45}$$

where $E_0 = \hbar\omega/2$ is the vacuum energy. The time dependence of these and other quantities may be computed in the time translated coherent state (see problem)

$$\begin{aligned}
|z,t\rangle &= \exp(-i\hat{H}t/\hbar)|z\rangle \\
&= e^{-i\omega t/2}e^{-it\omega a^\dagger a}e^{za^\dagger}|0\rangle \\
&= e^{-i\omega t/2}\exp\left(ze^{-i\omega t}a^\dagger\right)|0\rangle \\
&= |ze^{-i\omega t}\rangle e^{-i\omega t/2}.
\end{aligned} \tag{5.46}$$

For example,

$$\begin{aligned}
\bar{x}(t) &= \langle z,t|\hat{x}|z,t\rangle/\langle z,t|z,t\rangle \\
&= \tfrac{x_0}{\sqrt{2}}\langle ze^{-i\omega t}|a^\dagger + a|ze^{-i\omega t}\rangle/\langle z,t|z,t\rangle \\
&= \tfrac{x_0}{\sqrt{2}}(z^*e^{i\omega t} + ze^{-i\omega t}) \\
&= \bar{x}\cos\omega t + \tfrac{\bar{p}}{m\omega}\sin\omega t
\end{aligned} \tag{5.47}$$

This shows that the time dependence of the average position of the particle in a coherent state is just like the oscillatory motion of the particle according to the rules of classical mechanics. Therefore, a coherent state is interpreted as an almost classical particle state (see also problem 2).

5.3 Normal ordering

Consider a product of various powers of creation annihilation operators written in some order, for example $aa^{\dagger 3}a^5a^\dagger a^{17}\cdots a^{\dagger 2}$. Suppose that altogether the product contains n annihilation operators and m creation operators. The normal ordered product of the same set of operators is defined to be $a^{\dagger m}a^n$. That is, one pulls all the creation operators to the left and all annihilation operators to the right as if they commute, and *define* the resulting operator as the normal ordered product. Normal ordering is denoted by placing a column on both sides of the original product, that is,

$$:aa^{\dagger 3}a^5a^\dagger a^{17}\cdots a^{\dagger 2}: \equiv a^{\dagger m}a^n. \tag{5.48}$$

Inside the normal ordering signs the order of the operators does not matter since by definition any order gives the same result

$$:aa^{\dagger 3}a^5a^\dagger a^{17}\cdots a^{\dagger 2}: =: a^{\dagger 3}aa^5a^\dagger a^{\dagger 2}\cdots a^{17}: = \cdots = a^{\dagger m}a^n. \tag{5.49}$$

The vacuum expectation value of a normal ordered product is zero as long as m or n is not zero

$$\langle 0|\left(:aa^{\dagger 3}a^5a^\dagger a^{17}\cdots a^{\dagger 2}:\right)|0\rangle = \langle 0|a^{\dagger m}a^n|0\rangle = 0, \tag{5.50}$$

since the vacuum on the right is annihilated by a and the one on the left is annihilated by $a^\dagger$. Any product may be rewritten as a linear combination of normal ordered products by using the commutation rules to shift creation operators to the left and creation operators to the right. Here are a few examples

$$\begin{aligned} aa^\dagger &= a^\dagger a + 1 =: aa^\dagger : +1, \\ a^2a^\dagger &= a^\dagger a^2 + 2a =: a^2a^\dagger : +2 : a :, \\ a^2a^{\dagger 2} &= a^{\dagger 2}a^2 + 4a^\dagger a + 4 =: a^2a^{\dagger 2} : +4 : a^\dagger a : +4 \end{aligned} \tag{5.51}$$

Thus, in rewriting ordinary products in terms of normal ordered products, in addition to the naive normal ordered term one may find terms with fewer numbers of creation annihilation operators. In particular, the vacuum expectation value of an ordinary product is determined by the last term of its expansion in terms of normal ordered products. Thus, using the examples above,

$$\langle 0|aa^\dagger|0\rangle = 1, \quad \langle 0|a^2a^\dagger|0\rangle = 0, \quad \langle 0|a^2a^{\dagger 2}|0\rangle = 4. \tag{5.52}$$

Normal ordering is a useful tool in quantum field theory, where the field has the structure of a linear combination of creation/annihilation operators, $\phi = \alpha a + \beta^* a^\dagger$, where α, β^* are coefficients. For example the operators $\hat{x}, \hat{p}$ have this structure. In quantum field theory one is often interested in evaluating vacuum expectation values of the form

$$\langle 0|\phi_1\phi_2\cdots\phi_n|0\rangle \tag{5.53}$$

where for each field $\phi_i = \alpha_i a + \beta_i^* a^\dagger$ generally there are different coefficients α_i, β_i, but they all involve the same set of harmonic oscillator operators $a, a^\dagger$. So, the vacuum expectation value is a definite function of the coefficients α_i, β_i. To perform this computation in an organized way one uses the technique of normal ordering. That is, rewriting the operator products in terms of normal ordered products and then taking the vacuum expectation value.

5.3.1 Wick's theorem

Wick derived the following theorem which gives all the terms and their coefficients in the expansion of ordinary products in terms of normal ordered products (see problem at the end of chapter).

$$\begin{aligned} \phi_1\phi_2\cdots\phi_n =: \phi_1\phi_2\cdots\phi_n : + \{\langle\phi_1\phi_2\rangle : \phi_3\cdots\phi_n : + \textit{all permutations}\} \\ + \{\langle\phi_1\phi_2\rangle\langle\phi_3\phi_4\rangle : \phi_6\cdots\phi_n : + \textit{all permutations}\} + \{\cdots\} \\ + \{\langle\phi_1\phi_2\rangle\langle\phi_3\phi_4\rangle\cdots\langle\phi_{n-1}\phi_n\rangle + \textit{all permutations}\} \end{aligned} \tag{5.54}$$

where $\langle\phi_i\phi_j\rangle$ is the vacuum expectation value of a pair of fields. We can easily compute

$$\langle\phi_i\phi_j\rangle = \langle 0|\left(\alpha_i a + \beta_i^* a^\dagger\right)\left(\alpha_j a + \beta_j^* a^\dagger\right)|0\rangle = \alpha_i\beta_j^* \tag{5.55}$$

so that all coefficients are determined. The last line is written as if n is even since all fields have been paired. If n is odd there remains an unpaired field in each term of the last line. Using this theorem we see that when n is odd $\langle 0|\phi_1\phi_2\cdots\phi_n|0\rangle = 0$, and when it is even,

$$\langle 0|\phi_1\phi_2\cdots\phi_n|0\rangle = \alpha_1\beta_2^*\alpha_3\beta_3^*\cdots\alpha_{n-1}\beta_n^* + \textit{all permutations of } (12\cdots n). \tag{5.56}$$

This theorem plays a basic role in doing computations in quantum field theory.

5.4 Harmonic oscillator in 2 and d dimensions

The harmonic oscillator in d dimensions is described by the Hamiltonian

$$H = \frac{\vec{p}^2}{2m} + \frac{m\omega^2\vec{x}^2}{2} \tag{5.57}$$

where $\vec{x} = (x_1, x_2, \cdots, x_d)$ is a $d-$dimensional vector. We may define creation-annihilation operators

$$\vec{a} = \frac{1}{\sqrt{2}}\left(\frac{\vec{x}}{x_0} + i\frac{x_0\vec{p}}{\hbar}\right), \quad \vec{a}^\dagger = \frac{1}{\sqrt{2}}\left(\frac{\vec{x}}{x_0} - i\frac{x_0\vec{p}}{\hbar}\right), \tag{5.58}$$

that are also $d-$dimensional vectors. Their commutation rules follow from those of position and momentum operators

$$[a_I, a_J^\dagger] = \delta_{IJ}, \quad [a_I, a_J] = 0 = [a_I^\dagger, a_J^\dagger]. \tag{5.59}$$

By choosing $x_0 = (\hbar/m\omega)^{1/2}$ as before the Hamiltonian takes the simple form

$$\hat{H} = \hbar\omega(\vec{a}^\dagger \cdot \vec{a} + \frac{d}{2}). \tag{5.60}$$

Since this Hamiltonian is a sum of independent Hamiltonians $\hat{H} = \hat{H}_1 + \hat{H}_2 + \cdots + \hat{H}_d$, that are constructed from operators that commute with each other, the overall eigenstates reduce to direct products of eigenstates of the individual Hamiltonians. From the study of the harmonic oscillator in one dimension one already knows that the eigenstates of the Hamiltonian in each direction $H_I = \hbar\omega(a_I^\dagger a_I + \frac{1}{2})$ are the number states $|n_I\rangle$. Therefore the overall eigenstate is

$$|n_1, n_2, \cdots, n_d\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \cdots \otimes |n_d\rangle, \tag{5.61}$$

and the energy eigenvalue depends only on the sum of the integers n_I

$$E_{n_1 \cdots n_d} = \hbar\omega(n_1 + n_2 + \cdots + n_d + \frac{d}{2}). \tag{5.62}$$

There is a unique ground state with $n_I = 0$, which we denote simply by $|0\rangle$. However, at higher levels there are many ways to construct the same value for the total integer

$$n = \sum_{I=1}^{d} n_I. \tag{5.63}$$

Therefore, there are many states at each level and it can be checked that the number of degenerate states at level n is

$$D_n(d) = \frac{(d+n-1)!}{(d-1)!\; n!}. \tag{5.64}$$

For example for $d = 2$, $n = 0, 1, 2, 3, 4$ there are the following states $|n_1, n_2\rangle$

n	$D_n(2)$	*states* $\quad \lvert n_1, n_2 >$
0	1	$\lvert 0,0\rangle$
1	2	$\lvert 1,0\rangle,\ \lvert 0,1\rangle$
2	3	$\lvert 2,0\rangle,\ \lvert 0,2\rangle,\ \lvert 1,1\rangle$
3	4	$\lvert 3,0\rangle,\ \lvert 2,1\rangle,\ \lvert 1,2\rangle,\ \lvert 0,3\rangle$
4	5	$\lvert 4,0\rangle,\ \lvert 3,1\rangle,\ \lvert 2,2\rangle,\ \lvert 1,3\rangle,\ \lvert 0,4\rangle$

(5.65)

These states are constructed by applying the creation operators on the vacuum state

$$|n_1, n_2\rangle = \frac{a_1^{\dagger n_1}}{\sqrt{n_1!}} \frac{a_2^{\dagger n_2}}{\sqrt{n_2!}} |0\rangle. \tag{5.66}$$

It is useful to re-label the states in terms of the sum and difference of the quantum numbers. Defining

$$j = \frac{1}{2}(n_1 + n_2), \quad m = \frac{1}{2}(n_1 - n_2), \tag{5.67}$$

and renaming $a_1^\dagger = a_+^\dagger$ and $a_2^\dagger = a_-^\dagger$, we have

$$\begin{array}{l} |j,m\rangle = \frac{(a_+^\dagger)^{j+m}}{\sqrt{(j+m)!}} \, \frac{(a_-^\dagger)^{j-m}}{\sqrt{(j-m)!}} \, |0\rangle \\ m = -j,\, -j+1, \cdots,\, j-1,\, j \\ j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \cdots \end{array} \tag{5.68}$$

The total number operator

$$\hat{N} = \vec{a}^\dagger \cdot \vec{a} = a_+^\dagger a_+ + a_-^\dagger a_- \tag{5.69}$$

has eigenvalue $n_1 + n_2 = 2j$

$$\hat{N} \, |j,m\rangle = 2j \, |j,m\rangle, \tag{5.70}$$

The second quantum number m distinguishes the $(2j+1)$ degenerate states from each other. Using eqs.(5.21) and (5.22) we see that $a_\pm$, $a_\pm^\dagger$ act like ladder operators that change both j and m by $1/2$ unit

$$\begin{array}{ll} a_\pm |j,m\rangle & = \sqrt{j \pm m} \, |j - 1/2,\, m \mp 1/2\rangle \\ a_\pm^\dagger |j,m\rangle & = \sqrt{j \pm m + 1} \, |j + 1/2,\, m \pm 1/2\rangle. \end{array} \tag{5.71}$$

The creation operators increase j while the annihilation operators decrease it. The products of a creation with an annihilation operator do not change j, but may change m by at most one unit. This means they must commute with the operator $\hat{N}$. Thus, defining the operators

$$J_+ \equiv a_+^\dagger a_-, \quad J_- \equiv a_-^\dagger a_+, \quad J_0 \equiv \frac{1}{2}(a_+^\dagger a_+ - a_-^\dagger a_-), \tag{5.72}$$

one can verify that they commute with the total number operator

$$[\hat{N}, J_{\pm,0}] = 0 \tag{5.73}$$

and that they change m by at most one unit that corresponds to their labels $(+,0,-)$. Their action on the states follows from (5.71)

$$J_0 \, |j,m\rangle = m \, |j,m\rangle, \quad J_\pm |j,m\rangle = \sqrt{j(j+1) - m(m \pm 1)} \, |j, m \pm 1\rangle. \tag{5.74}$$

From the following rules for commuting pairs of creation-annihilation operators

$$[a_I^\dagger a_J, a_K^\dagger a_L] = a_I^\dagger a_L \, \delta_{JK} - a_K^\dagger a_J \, \delta_{LI} \tag{5.75}$$

one can easily derive that the commutation rules for the $J_{\pm,0}$ are given by

$$[J_0, J_\pm] = \pm J_\pm \ , \quad [J_+, J_-] = 2J_0 \ . \tag{5.76}$$

A set of operators that close under commutation rules to the same set, as above, is called a Lie algebra. The present Lie algebra is the Lie algebra of $SU(2)$. It is sometimes rewritten in terms of

$$J_1 = \frac{1}{2}(J_+ + J_-), \quad J_2 = \frac{1}{2i}(J_+ - J_-), \quad J_3 = J_0 \tag{5.77}$$

in the form

$$[J_1, J_2] = iJ_3 \ , \quad [J_2, J_3] = iJ_1 \ , \quad [J_3, J_1] = iJ_2 \ . \tag{5.78}$$

It is illuminating to note that there is a quadratic function of the $J_{\pm,0}$ that can be rewritten only in terms of the total number operator $\hat{N}$ (problem 6)

$$\begin{aligned} J^2 &= J_0^2 + \tfrac{1}{2}(J_+J_- + J_-J_+) \\ &= J_1^2 + J_2^2 + J_3^2 \\ &= \tfrac{\hat{N}}{2}\left(\tfrac{\hat{N}}{2} + 1\right). \end{aligned} \tag{5.79}$$

Therefore the eigenvalues of J^2 are $j(j+1)$. Even though the total number operator $\hat{N}$ commutes with all three operators $J_{\pm,0}$, it can be simultaneously diagonal with only one of them since they do not commute among themselves. However, they all play a role in clarifying the nature of the degenerate states at a fixed value of j. In fact, as we will see in the coming chapters, the three operators $J_{\pm,0}$ have a close connection to the symmetry group $SU(2)$, which is the same group as the group of rotations $SO(3)$ in three dimensions. In the present case of the two dimensional harmonic oscillator there certainly are no 3-dimensional rotations. But nevertheless there is a hidden symmetry group whose mathematical structure is similar to the group of 3-dimensional rotations. The $d-$dimensional harmonic oscillator also has a symmetry group which is $SU(d)$. There is a Lie algebra associated with $SU(d)$ as is evident from eq.(5.75) taken in $d-$dimensions. Although we will study this topic in more detail in the chapter on symmetries, it is appropriate at this juncture to point out how the symmetry acts. Thus, consider two observers O and O' who use coordinates and momenta $(\vec{x}, \vec{p})$ and $(\vec{x}', \vec{p}')$ respectively. Let us define the relation between these observers through a transformation of the creation-annihilation operators constructed from the respective coordinates and momenta

$$a_I{}' = \sum_{J=1}^{d} (U)_{IJ} \ a_J \ , \quad (a_I^\dagger)' = \sum_{J=1}^{d} a_J^\dagger (U^\dagger)_{JI} \ , \tag{5.80}$$

where U is a $d \times d$ unitary matrix, and $U^\dagger$ is its Hermitian conjugate. It may be clearer to the reader if these relations are explicitly written in matrix notation

$$\begin{pmatrix} a_1' \\ a_2' \\ \vdots \\ a_d' \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1d} \\ U_{21} & U_{22} & \cdots & U_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ U_{d1} & U_{d2} & \cdots & U_{dd} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} \tag{5.81}$$

and

$$\begin{pmatrix} a_1^{\dagger\prime} & a_2^{\dagger\prime} & \cdots a_d^{\dagger\prime} \end{pmatrix} = \begin{pmatrix} a_1^\dagger & a_2^\dagger & \cdots a_d^\dagger \end{pmatrix} \begin{pmatrix} U_{11}^\dagger & U_{12}^\dagger & \cdots & U_{1d}^\dagger \\ U_{21}^\dagger & U_{22}^\dagger & \cdots & U_{1d}^\dagger \\ \vdots & \vdots & \ddots & \vdots \\ U_{d1}^\dagger & U_{d2}^\dagger & \cdots & U_{dd}^\dagger \end{pmatrix} \tag{5.82}$$

Then one finds that the number operators $\hat{N}$ and $\hat{N}'$ used by the two observers are unchanged by the transformation, provided the matrix U is unitary, i.e. $UU^\dagger = U^\dagger U = 1$, since

$$\hat{N}' = \sum_{I=1}^{d} a_I^{\dagger\prime} a_I' = \sum_{I,J,K=1}^{d} a_J^\dagger \, (U^\dagger)_{JI} \, (U)_{IK} \, a_K = \sum_{I,J,K=1}^{d} a_J^\dagger \, \delta_{IK} \, a_K = \hat{N}. \tag{5.83}$$

This result indicates that the Hamiltonians of the two observers may be written in terms of different coordinates and momenta, but the energy does not change. Therefore, there is a symmetry in the system associated with unitary transformations in $d-$dimensions. This symmetry is the underlying reason for the degeneracy of the energy states, as will be more fully explained in the chapter on symmetries. Thus the $SU(2)$ symmetry of the two dimensional oscillator is responsible for the degeneracy discussed explicitly above.

5.5 Fermionic oscillators

There is a fermionic version of creation-annihilation operators $(b_\alpha, b_\alpha^\dagger)$ with $\alpha, \beta = 1, 2, \cdots M$, that are defined with anticommutation relations

$$\{b_\alpha, b_\beta^\dagger\} = \delta_{\alpha\beta}, \quad \{b_\alpha, b_\beta\} = 0 = \{b_\alpha^\dagger b_\beta^\dagger\}. \tag{5.84}$$

The anticommutator between two operators is defined by $\{A, B\} = AB + BA$. The anticommutation rules imply that the square of each operator vanishes, $b_\alpha^2 = 0 = b_\alpha^{\dagger 2}$. The total number operator is defined as before $\hat{N} = \sum b_\alpha^\dagger b_\alpha$, and its commutation rules with the creation-annihilation operators turns out to be the same as the bosonic oscillators

$$[\hat{N}, b_\alpha] = -b_\alpha, \quad [\hat{N}, b_\alpha^\dagger] = b_\alpha^\dagger. \tag{5.85}$$

The vacuum state is annihilated by all the b_α

$$b_\alpha |0\rangle = 0 \tag{5.86}$$

and excited states are constructed by applying the creation operators

$$b_{\alpha_1}^\dagger b_{\alpha_2}^\dagger \, \cdots b_{\alpha_n}^\dagger |0\rangle. \tag{5.87}$$

The total number operator is diagonal on these states, and its eigenvalue is equal to the total number of excitations, just like the ordinary harmonic oscillator. However, there cannot be more than one excitation for each oscillator, since $b_\alpha^{\dagger 2} = 0$ for each α. This explains the Pauli exclusion principle, that says that no two fermions with the same quantum numbers can be in the same state. Therefore the number of states is finite. For M oscillators there are 2^M distinct

states that can be explicitly listed for a given value of M

$$\begin{array}{lcl} M=1 & : & |0\rangle,\ b_1^\dagger|0\rangle. \\ M=2 & : & |0\rangle,\ b_1^\dagger|0\rangle,\ b_2^\dagger|0\rangle,\ b_1^\dagger b_2^\dagger|0\rangle. \\ M=3 & : & \left\{ \begin{array}{c} |0\rangle,\ b_1^\dagger|0\rangle,\ b_2^\dagger|0\rangle,\ b_3^\dagger|0\rangle, \\ b_1^\dagger b_2^\dagger|0\rangle,\ b_2^\dagger b_3^\dagger|0\rangle,\ b_3^\dagger b_1^\dagger|0\rangle,\ b_1^\dagger b_2^\dagger b_3^\dagger|0\rangle. \end{array} \right. \\ \vdots & & \vdots \end{array} \tag{5.88}$$

The set of 2^M states is the complete Hilbert space for a given value of M. A state is either empty or full relative to a given fermion labelled by α. The rules for applying creation-annihilation operators on the number states are easy to figure out: When the state is empty for quantum number α, b_α annihilates it while $b_\alpha^\dagger$ fills it. If the state is full, then b_α empties it while $b_\alpha^\dagger$ annihilates it. So, a completely empty state (the vacuum $|0\rangle$) is annihilated by all b_α, but a completely filled state is annihilated by all $b_\alpha^\dagger$. Thus, the completely filled state has properties very similar to the vacuum state, and the roles of the creation-annihilation operators get interchanged when acting on it. One may consider pairs of creation-annihilation operators $b_\alpha^\dagger b_\beta$ and study their commutators

$$[b_\alpha^\dagger b_\beta\ ,\ b_\lambda^\dagger b_\sigma] = b_\alpha^\dagger b_\sigma\ \delta_{\beta\lambda} - b_\lambda^\dagger b_\beta\ \delta_{\alpha\sigma}, \tag{5.89}$$

as we did in the bosonic oscillator case. Since the set of operators close, they form a Lie algebra $SU(M)$. For $M=2$, renaming $b_1 \equiv b_+$ and $b_2 \equiv b_-$ we may construct again the Lie algebra of $SU(2)$

$$\begin{array}{l} J_+ = b_+^\dagger b_-, \quad J_- = b_-^\dagger b_+, \quad J_0 = \frac{1}{2}(b_+^\dagger b_+ - b_-^\dagger b_-) \\ [J_0, J_\pm] = \pm J_\pm, \qquad [J_+, J_-] = 2J_0. \end{array} \tag{5.90}$$

The eigenvalues of $J^2 = j(j+1)$ can only take a finite number of values since there are only 4 states in this case. In fact the vacuum and the filled state both have $j=0$, while the one particle states $b_\pm^\dagger|0\rangle$ represent the two states for $j=1/2$.

5.6 Quadratic interactions for N particles

Consider the problem of N particles moving in $d-$dimensions and interacting through ideal springs as in Fig.(5.1).

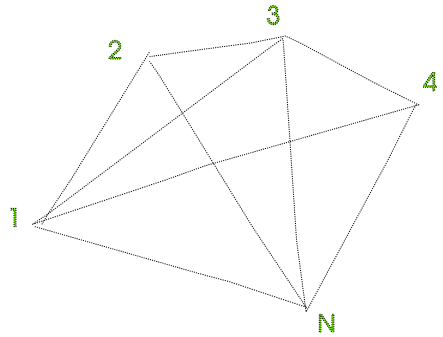


Fig.5.1: Interactions via springs.

The potential energy stored in the springs is proportional to the square of the distance between the particles at its two end points. For example for three springs coupled to three particles in all possible ways the potential energy of the system is

$$V = \frac{k_{12}}{2}(\vec{x}_1 - \vec{x}_2)^2 + \frac{k_{23}}{2}(\vec{x}_2 - \vec{x}_3)^2 + \frac{k_{31}}{2}(\vec{x}_3 - \vec{x}_1)^2 \tag{5.91}$$

where k_{ij} are the spring constants, and the kinetic energy is

$$K = \frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} + \frac{\vec{p}_3^2}{2m_3}. \tag{5.92}$$

More generally the springs may not be isotropic and may pull differently in various directions. To cover all possibilities we will consider a Hamiltonian of the form

$$H = \frac{1}{2}\sum_{i,j=1}^{N}\left(K_{ij}p_ip_j + V_{ij}x_ix_j + W_{ij}^T x_ip_j + W_{ij}\,p_ix_j)\right) + \sum_{i=1}^{N}(\alpha_ip_i + \beta_ix_i) \tag{5.93}$$

where the indices i, j run over the particle types and the various directions, and we will assume a real general matrix W, arbitrary symmetric matrices K, V, and coefficients α,β which may be considered column or row matrices (vectors). The mathematics of this system could model a variety of other physical situations besides the coupled spring problem which we used to motivate this Hamiltonian.. This general problem has an exact solution in both classical and quantum mechanics.

By translating the positions and momenta to $\bar{p}_i = p_i - p_i^0$, $\bar{x}_i = x_i - x_i^0$, one can find $H = \hat{H}(\bar{x}\ , \bar{p}) + E(x^0, p^0)$ such that the linear terms are absent in the new variables. The constant $E(x^0, p^0) = H(x^0, p^0)$ just shifts the energy of every state by the same amount. Since the commutation rules for the new variables are the same as the original ones, one may assume that this step has already been done and start as if $\alpha_i = \beta_i = 0$ without any loss of generality. It is convenient to write this Hamiltonian in matrix notation in the form

$$H = \begin{array}{c}\left(\bar{p}^T\,\bar{x}^T\right)\end{array}\begin{pmatrix} K & W \\ W^T & V \end{pmatrix}\begin{pmatrix} \bar{p} \\ \bar{x} \end{pmatrix} + E(x^0, p^0). \tag{5.94}$$

where we have defined $N-$dimensional column and row matrices

$$\begin{aligned} \bar{p}^T &= (\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_N), \\ \bar{x}^T &= (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N), \end{aligned} \tag{5.95}$$

etc. Next define a $2N \times 2N$ transformation M involving both positions and momenta $\bar{p}_i = \sum_j (A_{ij}\tilde{p}_j + B_{ij}\tilde{x}_j)$, $\bar{x}_i = \sum_j (C_{ij}\tilde{p}_j + D_{ij}\tilde{x}_j)$, or

$$\begin{pmatrix} \bar{p} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \tilde{p} \\ \tilde{x} \end{pmatrix}, \quad M \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{5.96}$$

We require that H has no mixed terms in the new variables $(\tilde{p}, \tilde{x})$, which means it takes the form $H = \tilde{p}^T \tilde{K} \tilde{p} + \tilde{x}^T \tilde{V} \tilde{x} + E(x^0, p^0)$ once we substitute Eq.(5.96) in Eq.(5.94). Therefore we must require the block diagonalization condition

$$\begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} K & W \\ W^T & V \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \tilde{K} & 0 \\ 0 & \tilde{V} \end{pmatrix} \tag{5.97}$$

or

$$\begin{pmatrix} A^T (KA + WC) + C^T \left(W^T A + VC\right) & A^T (KB + WD) + C^T \left(W^T B + VD\right) \\ B^T (KA + WC) + D^T \left(W^T A + VC\right) & B^T (KB + WD) + D^T \left(W^T B + VD\right) \end{pmatrix}$$
$$= \begin{pmatrix} \tilde{K} & 0 \\ 0 & \tilde{V} \end{pmatrix} \tag{5.98}$$

The off diagonal blocks vanish provided A, B, C, D are restricted by the following equation

$$A^T K B + A^T W D + C^T W^T B + C^T V D = 0. \tag{5.99}$$

From this equation we can solve for the parameters of the matrix M that satisfy the condition.

In addition we require that the commutation rules among the new variables are the standard canonical rules, i..e. $[\tilde{x}_i, \tilde{p}_j] = i\hbar\delta_{ij},\ [\tilde{x}_i, \tilde{x}_j] = 0 = [\tilde{p}_i, \tilde{p}_j]$, and that these produce the standard canonical rules for the old variables $\bar{x}_i, \bar{p}_i$. This requirement produces the following conditions on A, B, C, D

$$AB^T - BA^T = 0,\ CD^T - DC^T = 0, \quad AD^T - BC^T = 1. \tag{5.100}$$

These equations may be rewritten in the form $M^T S M = S$, or equivalently

$$M^{-1} = S^{-1} M^T S, \ \text{ with } \ S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ S^{-1} = -S. \tag{5.101}$$

A $2N \times 2N$ linear transformation M that leaves the commutation rules invariant is a canonical transformation that is called a symplectic transformation. Such transformations form the symplectic group $\mathrm{Sp}(2N)$. The general symplectic matrix can be parametrized as follows

$$\begin{aligned} M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} u & bu^{-1T} \\ cu & (cb+1)\, u^{-1T} \end{pmatrix} \\ &= \begin{pmatrix} 1 & b \\ c & (cb+1) \end{pmatrix} \begin{pmatrix} u & 0 \\ 0 & u^{-1T} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} u & 0 \\ 0 & u^{-1T} \end{pmatrix} \begin{pmatrix} 1 & u^{-1} b u^{-1T} \\ 0 & 1 \end{pmatrix} \end{aligned}$$

where b and c are any $N \times N$ real symmetric matrices and u is any $N \times N$ real matrix that has an inverse u^{-1}. Note that the number of parameters in b or c is $\frac{1}{2}N(N+1)$ each, and the number of parameters in u is N^2. Therefore a real $\mathrm{Sp}(2N)$ symplectic matrix has altogether $\frac{1}{2}N(N+1) + \frac{1}{2}N(N+1) + N^2 =$

$N(2N+1)$ real parameters. It can be checked directly that this form of M satisfies automatically the symplectic condition $M^T S M = S$. This implies that the inverse M^{-1} is given by $M^{-1} = S^{-1} M^T S$ for any b, c, u. We mention in passing that u is an element of the subgroup $\mathrm{GL}(N, R)$.

The matrices b, c, u must be further restricted to block diagonalize the Hamiltonian. We insert the A, B, C, D of Eq.(8.63) in the diagonalization condition in Eq.(5.99). This gives

$$Kb + W(cb+1) + cW^T b + cV(cb+1) = 0 \tag{5.102}$$

Therefore we need to solve this equation to find b, c for given K, W, V. This equation imposes N^2 conditions, but we have $N^2 + N$ free parameters in the symmetric b, c to satisfy N^2 equations, which means there is always a solution[1] (see problem 12). Note that u is not restricted by the block diagonalization condition, so it can be taken as the identity matrix 1 for the step above. However, in the following steps we will see that the degrees of freedom in u will be used to further simplify the problem.

We also compute $\tilde{K}$ and $\tilde{V}$ from Eq.(5.98)

$$\begin{aligned} \tilde{K} &= A^T K A + A^T W C + C^T W^T A + C^T V C \\ &= u^T \left(K + Wc + cW^T + cVc \right) u \end{aligned} \tag{5.103}$$

Similarly we have

$$\begin{aligned} \tilde{V} &= B^T K B + B^T W D + D^T W^T B + D^T V D \\ &= u^{-1} \left(W^T b + V(cb+1) \right) u^{-1T} \end{aligned} \tag{5.104}$$

where we have used Eq.(5.102) to simplify the expression for $\tilde{V}$. In the following discussion we will need $\tilde{V}\tilde{K}$, so we compute it as

$$\tilde{V}\tilde{K} = u^{-1} \left[W^T b + V(cb+1) \right] \left[K + Wc + cW^T + cVc \right] u \tag{5.105}$$

Note that we have not used the freedom we have in u, so at this stage u can be set equal to 1.

Assuming that these steps have been performed, the Hamiltonian takes the form

$$\hat{H} = \frac{1}{2} \sum_{i,j=1}^{N} \left(\tilde{K}_{ij} \tilde{p}_i \tilde{p}_j + \tilde{V}_{ij} \tilde{x}_i \tilde{x}_j \right) + E(x_0, p_0). \tag{5.106}$$

It will be convenient to write this expression in matrix notation as

$$\hat{H} = \tilde{p}^T \tilde{K} \tilde{p} + \tilde{x}^T \tilde{V} \tilde{x} + E(x_0, p_0). \tag{5.107}$$

Next introduce an orthogonal transformation R to diagonalize the symmetric matrix $\tilde{K} = R^T k^2 R$, where k^2 is assumed to be a positive diagonal matrix .

[1] In passing, it is also interesting to note that this equation has the following solution for the general matrix W for given symmetric matrices b, c, K, V, namely $W = \sum_{n=0}^{\infty} c^n \left(Kb(cb+1)^{-1} + cV \right) \left(b(cb+1)^{-1} \right)^n (-1)^{n+1}$.

The positivity is required on physical grounds that the kinetic energy in the new variables must be positive, so this will be true in any physically meaningful system. By inserting $R^T R = 1$ the Hamiltonian is rewritten as

$$\begin{aligned} \hat{H} &= \tfrac{1}{2}\left(\tilde{p}^T R^T k^2 R\tilde{p} + \tilde{x}^T R^T R\tilde{V} R^T R\tilde{x}\right) + E(x_0, p_0) \\ &= \tfrac{1}{2}\left(p'^T k^2 p' + x'^T V' x'\right) + E(x_0, p_0) \end{aligned} \tag{5.108}$$

where $p' = R\tilde{p},\ x' = R\tilde{x},\ V' = R\tilde{V}R^T$. Since orthogonal transformations also preserve the canonical commutation rules, we may again assume that this step is accomplished and reduce the problem to a diagonal $K'_{ij} = k_i^2\delta_{ij}$ without loss of generality. Next introduce a rescaling that also preserves the canonical structure

$$\tilde{p}'_i = k_i p'_i, \quad \tilde{x}'_i = k_i^{-1} x'_i. \tag{5.109}$$

This brings the Hamiltonian to the form

$$\hat{H} = \frac{1}{2}\left(\tilde{p}'^T\tilde{p}' + \tilde{x}'^T(kV'k)\tilde{x}'\right) + E(x_0, p_0). \tag{5.110}$$

Finally using an orthogonal transformation that diagonalizes the symmetric matrix, $(kV'k) = U^T\omega^2 U$, the Hamiltonian takes its final simplest form by inserting $U^T U = 1$ and defining $p'' = U\tilde{p}',\ x'' = U\tilde{x}'$

$$\begin{aligned} \hat{H} &= \tfrac{1}{2}\left(\tilde{p}'^T U^T U\tilde{p}' + \tilde{x}'^T U^T\omega^2 U\tilde{x}'\right) + E(x_0, p_0) \\ &= \tfrac{1}{2}\left(p''^T p'' + x''^T\omega^2 x''\right) + E(x_0, p_0) \\ &= \tfrac{1}{2}\sum_{i=1}^{N}\left((p''_i)^2 + \omega_i^2 (x''_i)^2\right) + E(x_0, p_0). \end{aligned} \tag{5.111}$$

This decoupled form defines the normal coordinates that are independent oscillators. The frequencies ω_i^2 correspond to the eigenvalues of the matrix $(kV'k)$. The eigenvalues of $(kV'k)$ are the same as those of the matrix $R^T(kV'k)R$, which can be rewritten as $\left(R^T kR\right)\left(R^T V' R\right)\left(R^T kR\right) = \sqrt{\tilde{K}}\tilde{V}\sqrt{\tilde{K}}$. Therefore, to compute the frequencies we setup the secular equation $\det\left(\sqrt{\tilde{K}}\tilde{V}\sqrt{\tilde{K}} - \lambda\right) = 0$. The solutions for λ correspond to the frequencies ω_i^2. Now, the determinant can be rewritten as follows

$$\begin{aligned} 0 &= \det\left(\sqrt{\tilde{K}}\tilde{V}\sqrt{\tilde{K}} - \lambda\right) = \det\left(\sqrt{\tilde{K}}\left(\tilde{V} - \lambda\tilde{K}^{-1}\right)\sqrt{\tilde{K}}\right) \\ &= \left(\det\sqrt{\tilde{K}}\right)\det\left(\tilde{V} - \lambda\tilde{K}^{-1}\right)\left(\det\sqrt{\tilde{K}}\right) = \left(\det\tilde{K}\right)\det\left(\tilde{V} - \lambda\tilde{K}^{-1}\right) \\ &= \det\left(\tilde{K}\tilde{V} - \lambda\right) \end{aligned} \tag{5.112}$$

Now we insert the product $\tilde{K}\tilde{V}$ from Eq.(5.105)

$$\begin{aligned} 0 &= \det\left(\tilde{K}\tilde{V} - \lambda\right) = \det\left(u^{-1}\left[W^T b + V(cb+1)\right]\left[K + Wc + cW^T + cVc\right]u - \lambda\right) \\ &= \det\left(\left[W^T b + V(cb+1)\right]\left[K + Wc + cW^T + cVc\right] - \lambda\right) \end{aligned} \tag{5.113}$$

So, we have shown that we can compute the frequencies by solving for the roots of this secular equation and identifying the solutions with $\lambda = \omega_i^2$. When $W = 0$ this task is much simpler since then b, c are also zero, and we get simply $\det(VK - \lambda) = 0$, where K, V are the original matrices. Once the frequencies are determined in this way, we have seen that we can write down immediately the simpler form of the Hamiltonian in Eq.(5.111).

At this stage we can quantize the normal modes and introduce the creation-annihilation operators

$$x_i'' = \sqrt{\frac{\hbar}{2\omega_i}}(a_i + a_i^\dagger), \quad p_i'' = -i\sqrt{\frac{\hbar\omega_i}{2}}(a_i - a_i^\dagger), \tag{5.114}$$

and finally have

$$H = \sum_{i=1}^{N} \hbar\omega_i(a_i^\dagger a_i + 1/2) + E(x^0, p^0). \tag{5.115}$$

The eigenstates are the usual number states $|n_1, \cdots, n_N\rangle$, and the energy eigenvalues are

$$E_{n_1 \cdots n_N} = \sum_{i=1}^{N} \hbar\omega_i(n_i + 1/2) + E(x^0, p^0). \tag{5.116}$$

There would be degeneracies only if the frequencies of the normal modes are accidentally the same. Thus, the solution of the original quantum problem is reduced to the computation of the frequencies. The series of steps above can always be accomplished in order to compute them (see problem). The relation of the creation-annihilation operators to the original position-momentum variables in (5.93) is obtained by putting together the transformations applied at each step

$$\begin{array}{r} a = \sqrt{\frac{\omega}{2\hbar}}\left\{Uk^{-1}R\left(A(x - x^0) + B(p - p^0)\right)\right\} \\ +i\sqrt{\frac{\hbar}{2\omega}}\left\{UkR\left(-B(x - x^0) + A(p - p^0)\right)\right\} \end{array} \tag{5.117}$$

where a, x, p, x^0, p^0 are column matrices, ω, k are diagonal matrices, U, R are orthogonal matrices, and A, B form the $2N \times 2N$ symplectic matrix that satisfy the conditions described above. For any choice of the parameters in these symplectic, orthogonal, diagonal or column matrices, the original position-momentum commutation rules lead to the standard creation-annihilation commutation rules. Furthermore, the Hamiltonian is diagonalized when the parameters in these matrices are chosen according to the steps above, so they end up being functions of the original parameters in K, V, W, α, β (see problem).

There is a quicker way to reach the final result. Starting with $H = \tilde{p}^T\tilde{K}\tilde{p} + \tilde{x}^T\tilde{V}\tilde{x} + E(x_0, p_0)$ we define the canonical transformation $\tilde{p}' = \sqrt{\tilde{K}}\tilde{p}$ and $\tilde{x}' = \left(\sqrt{\tilde{K}}\right)^{-1}\tilde{x}$. The Hamiltonian takes the form

$$\hat{H} = \tilde{p}'^T\tilde{p}' + \tilde{x}'^T\left[\sqrt{\tilde{K}}\tilde{V}\sqrt{\tilde{K}}\right]\tilde{x}' + E(x_0, p_0). \tag{5.118}$$

The eigenvalues of the matrix $\sqrt{\tilde{K}}\tilde{V}\sqrt{\tilde{K}}$ gives the eigenvalues ω_i^2 in agreement with the discussion above.

5.7 An infinite number of particles as a string

Consider a system of coupled particles and springs as in the previous section, but with only nearest neighbor interactions with N springs whose strengths are the same, and $N+1$ particles whose masses are the same. The Lagrangian in $d-$dimensions written in vector notation is

$$L(\mathbf{x}_i, \dot{\mathbf{x}}_i) = \frac{1}{2}m\sum_{i=0}^{N}(\dot{\mathbf{x}}_i \cdot \dot{\mathbf{x}}_i) - \frac{k}{2}\sum_{i=1}^{N}(\mathbf{x}_i - \mathbf{x}_{i-1})^2. \tag{5.119}$$

The index $i = 0, \cdots, N$ refers to the $i-$th particle. We have argued above that one can always solve a problem like this. We will see, in fact, that the solution of such a system for $N \to \infty$ will describe the motion of a string moving in d-dimensions. Let us visualize the system in $d = 3$. We have an array of particles whose motion is described by the solution of the coupled equations for $\mathbf{x}_i(t)$. Suppose that the N particles are initially arranged as in Fig.(5.2) at $t = t_0$

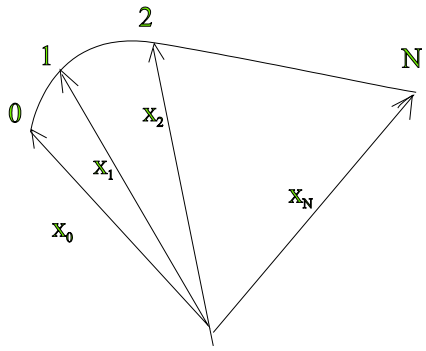


Fig.5.2: N positions at time t=t_0.

As t increases, the configuration of such an array of particles changes. Taking pictures at $t = t_1,\ t = t_2,\ t = t_3$ we can trace the trajectories as in Fig.(5.3).

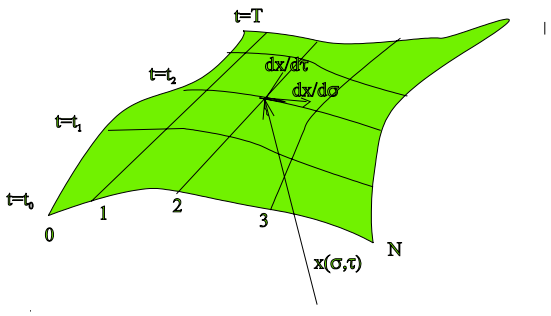


Fig.5.3: Trajectories of N particles.

The overall motion looks like a motion of a discretized string of beads which sweeps a discretized surface embedded in d dimensions. Let us now suppose that we have a continuous string parametrized as $\mathbf{x}(\sigma, t)$. At a fixed time, as the parameter σ changes in some range, say $0 \leq \sigma \leq \pi$, the vector $\mathbf{x}(\sigma, t)$ represents the location of different points on the string in a continuous fashion. Allowing also the time to change, one sees that the vector $\mathbf{x}(\sigma, t)$ points at any location on the two dimensional surface swept by the string. This surface is called the world sheet of the string. At any point on this surface we may identify the two tangents to the surface as $\partial_t \mathbf{x}(\sigma, t) \equiv \dot{\mathbf{x}}(\sigma, t)$ and $\partial_\sigma \mathbf{x}(\sigma, t) \equiv \mathbf{x}'(\sigma, t)$. Suppose we divide the range of σ into N discrete parts of equal length $a = \pi/N$, identify the $N+1$ points $\sigma = ia$ with $i = 0, 2, \cdots, N$ and assign their location $\mathbf{x}(\sigma = ia, t) \to \mathbf{x}_i(t)$. Then the continuous open string is approximated by $N+1$ particles. In the limit $a \to 0$, and $N \to \infty$, with $Na = \pi$ fixed, we recover the string from the collection of an infinite number of particles. A closed string that forms a loop may be described similarly. In this case it is more convenient to double the range of σ to $0 \leq \sigma \leq 2\pi$, discretize it in the same way, but also add the condition that the first and last point parametrized by σ are really the same point $x_0 = x_N$. In this discretized version the two tangents to the string world sheet are replaced or approximated by

$$\dot{\mathbf{x}}(\sigma, t) \to \dot{\mathbf{x}}_i(t), \qquad \mathbf{x}'(\sigma, t) \to \frac{1}{a}\left(\mathbf{x}_i(t) - \mathbf{x}_{i-1}(t)\right). \tag{5.120}$$

We now see the interpretation of the Lagrangian written above in terms of a string. Namely, in the limit $N \to \infty$ it represents the kinetic energy plus the potential energy of a moving open string:

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \frac{\mu}{2}\int_0^\pi d\sigma \ \left(\dot{\mathbf{x}}(\sigma, t)\right)^2 - \frac{T}{2}\int_0^\pi d\sigma \ \left(\mathbf{x}'(\sigma, t)\right)^2 \tag{5.121}$$

where we have defined a mass density along the string $\mu = m/a$ and a tension $T = ka$ that remain constant as $a \to 0$. This Lagrangian, together with the boundary conditions for open or closed strings

$$\begin{array}{ll} \mathbf{x}'(0, t) = \mathbf{x}'(\pi, t) = 0 & (\textit{open string}) \\ \mathbf{x}(0, t) = \mathbf{x}(2\pi, t) & (\textit{closed string}) \end{array} \tag{5.122}$$

recover the $N \to \infty$ limit of its discretized version (the boundary conditions for open strings follow from the minimal action principle by allowing free variation of the end points; they prevent momentum leakage at the end points). We have already seen that we can completely solve the classical or quantum mechanics problem in the discretized version by finding the normal modes. We may also study it directly in the continuous string version and interpret the normal modes as the normal modes of string motions (vibrations, rotations, etc.). The quantum theory is easily solved in terms of the normal modes that are represented as oscillators. Let us first study this Lagrangian classically. The Euler-Lagrange equation takes the form

$$(\mu\partial_t^2 - T\partial_\sigma^2)\mathbf{x}(\sigma, t) = 0. \tag{5.123}$$

This equation is equivalent to the Klein-Gordon equation in one "space" (σ) and one "time" (t) direction. The general solution is

$$\mathbf{x}(\sigma,t) = \mathbf{f}(\omega t+\sigma) + \mathbf{g}(\omega t-\sigma) \tag{5.124}$$

where $\omega = \sqrt{T/\mu}$ and $\mathbf{f},\mathbf{g}$ are arbitrary functions of their arguments. For an open string, imposing the boundary condition $\mathbf{x}'(0,t) = 0$ requires that the two functions be the same up to a constant $\mathbf{f}(\omega t) = \mathbf{g}(\omega t) + \mathbf{c}$. After using this restriction, imposing the boundary condition $\mathbf{x}'(\pi,t) = 0$ demands that the derivative of the function be periodic with period 2π: i.e. $\mathbf{f}'(\omega t+\pi) = \mathbf{f}'(\omega t-\pi)$. The general real periodic function can be expanded in terms of a Fourier series in the form

$$\mathbf{f}^{'}(u) = \frac{\mathbf{a}_0}{2} + \frac{i}{2}\sum_{n=1}^{\infty}\sqrt{n}(\mathbf{a}_n e^{iun} - \mathbf{a}_n^{\dagger}e^{-iun}), \tag{5.125}$$

where the Fourier modes $\mathbf{a}_n$ are arbitrary constants and the extra factor of $\sqrt{n}$ is inserted for later convenience. By integrating this function we have

$$\begin{array}{ll}\mathbf{f}(u) & = \mathbf{c}_1 + \frac{\mathbf{a}_0 u}{2} + \frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\left(\mathbf{a}_n e^{iun} + \mathbf{a}_n^{+}e^{-iun}\right) \\ \mathbf{g}(u) & = \mathbf{c}_2 + \frac{\mathbf{a}_0 u}{2} + \frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\left(\mathbf{a}_n e^{iun} + \mathbf{a}_n^{+}e^{-iun}\right)\end{array} \tag{5.126}$$

leading to the general solution

$$\mathbf{x}(\sigma,t) = (\mathbf{c}_1+\mathbf{c}_2) + \mathbf{a}_0\omega t + \sum_{n=1}^{\infty}\frac{\cos n\sigma}{\sqrt{n}}\left(\mathbf{a}_n e^{i\omega nt} + \mathbf{a}_n^{\dagger}e^{-i\omega nt}\right) \tag{5.127}$$

or

$$\mathbf{x}(\sigma,t) = \mathbf{x}_0(t) + \sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\mathbf{x}_n(t)\ \cos n\sigma \tag{5.128}$$

Here the zero mode which may be rewritten as

$$\mathbf{x}_0(t) = \mathbf{x}_0 + \mathbf{v}_0 t \tag{5.129}$$

is interpreted as the free motion of the center of mass. In fact note that the integral over σ yields just the zero mode, and also the integral is interpreted as the center of mass variable in the discretized version (with $Na = \pi$)

$$\mathbf{x}_0 + \mathbf{v}_0 t = \frac{1}{\pi}\int_0^{\pi}\mathbf{x}(\sigma,t)\ d\sigma \simeq \frac{Na}{\pi}\left[\frac{m(\mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_N)}{Nm}\right] \tag{5.130}$$

So, the center of mass behaves like a free particle. The remaining normal modes which may be rewritten as

$$\mathbf{x}_n(t) = \left(\mathbf{r}_n\cos\omega nt + \frac{\mathbf{v}_n}{\omega n}\sin\omega nt\right) \tag{5.131}$$

perform oscillatory motion with frequencies that are the multiples of the basic frequency $\omega_n = n\omega$. The reader is invited to solve the discretized problem

directly for any N and then take the large N limit to show that the result is the same. It is useful to get a feeling of the motion performed by a pure normal mode. For example consider the solution for which $\mathbf{x}_0 = 0$ and $\mathbf{v}_0 = 0$ so that the center of mass is at rest, and take all the normal modes equal to zero except for the first one

$$\mathbf{x}(\sigma, t) = \left(\mathbf{r}\cos\omega t + \frac{\mathbf{v}}{\omega}\sin\omega t\right)\cos\sigma. \tag{5.132}$$

At $t = 0$, $\mathbf{x}(\sigma, t) = \mathbf{r}\cos(\sigma)$ represents a string stretched from $-\mathbf{r}$ to $+\mathbf{r}$. When t increases, the string vibrates and rotates like a rod about its center of mass at a rate ω. If $\mathbf{v}$ is initially along $\mathbf{r}$ there are only longitudinal vibrations, and if it is initially perpendicular to $\mathbf{r}$ there is rotation while the length of string changes. For $t = \frac{2\pi}{\omega}$, the string goes back to its original position. If the string is purely in the second normal mode

$$\mathbf{x}(\sigma, t) = \left(\mathbf{r}_2\cos 2\omega t + \frac{\mathbf{v}_2}{2\omega}\sin 2\omega t\right)\cos 2\sigma \tag{5.133}$$

the string folds on itself, since $\cos 2\sigma$ covers the same values twice while σ changes from 0 to π. As time goes on it vibrates and rotates while it stays folded. With a more general superposition of many modes the string performs rather complicated vibrational, rotational, folding and unfolding motions in addition to moving as a whole following its center of mass. A closed string is described in the same way. The boundary conditions allow basically two independent functions $\mathbf{f}$ and $\mathbf{g}$, except for the zero mode that must be the same for periodicity in the sigma variable to be satisfied for the total $\mathbf{x} = \mathbf{f} + \mathbf{g}$

$$\begin{aligned}\mathbf{f}(\omega t + \sigma) &= \tfrac{1}{2}\left(\mathbf{x}_0 + \mathbf{v}_0(\omega t + \sigma)/\omega\right) + \tfrac{1}{2}\textstyle\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\left(\mathbf{a}_n e^{in(\omega t+\sigma)} + \mathbf{a}_n^\dagger e^{-in(\omega t+\sigma)}\right)\\ \mathbf{g}(\omega t - \sigma) &= \tfrac{1}{2}\left(\mathbf{x}_0 + \mathbf{v}_0(\omega t - \sigma)/\omega\right) + \tfrac{1}{2}\textstyle\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\left(\tilde{\mathbf{a}}_n e^{in(\omega t-\sigma)} + \tilde{\mathbf{a}}_n^\dagger e^{-in(\omega t-\sigma)}\right)\end{aligned} \tag{5.134}$$

In the sum the sigma dependence drops out for the center of mass motion $\mathbf{x}_0(t) = \mathbf{x}_0 + \mathbf{v}_0 t$, hence periodicity in the sigma variable is achieved. The amplitudes of oscillations $\mathbf{a}_n, \tilde{\mathbf{a}}_n$ are independent from each other. They represent waves traveling clockwise or anticlockwise along the closed string. Let us now analyze the quantum theory. The aim is to express the Hamiltonian in terms of the normal modes. The classical analysis has already identified them. For the open string we simply substitute in the Lagrangian (5.121) the derivatives of the expression in (5.128)

$$\dot{\mathbf{x}}(\sigma, t) = \dot{\mathbf{x}}_0(t) + \sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}\dot{\mathbf{x}}_n(t)\cos(n\sigma), \quad \mathbf{x}'(\sigma, t) = -\sum_{n=1}^{\infty}\sqrt{n}\mathbf{x}_n(t)\sin(n\sigma) \tag{5.135}$$

and do the integrals over σ. The result is

$$L = \frac{\pi\mu}{2}\left[\dot{\mathbf{x}}_0^2 + \sum_{n=1}^{\infty}\frac{1}{n}\left(\dot{\mathbf{x}}_n^2 - \omega^2 n^2\mathbf{x}_n^2\right)\right]. \tag{5.136}$$

We see that the normal modes are indeed separated from each other and that we have the Lagrangian for the harmonic oscillator for each of the normal modes

except for the zero mode. Canonical conjugate variables are defined as usual

$$\mathbf{p}_0 = \pi\mu\dot{\mathbf{x}}_0 \qquad \mathbf{p}_n = \pi\mu\dot{\mathbf{x}}_n/n \tag{5.137}$$

and the Hamiltonian follows

$$\begin{aligned} H_{open} &= \frac{\mathbf{p}_0^2}{2\pi\mu} + \sum_{n=1}^{\infty} n\left(\frac{\mathbf{p}_n^2}{2\pi\mu} + \frac{\pi\mu\omega^2\mathbf{x}_n^2}{2}\right) \\ &= \frac{\mathbf{p}_0^2}{2\pi\mu} + \sum_{n=1}^{\infty}\hbar\omega n\left(\mathbf{a}_n^\dagger\cdot\mathbf{a}_n + \frac{d}{2}\right) \end{aligned} \tag{5.138}$$

where the creation-annihilation operators are constructed as usual,

$$\mathbf{a}_n = \frac{1}{\sqrt{2}}(\frac{\mathbf{x}_n}{x_0} + i\frac{x_0\,\mathbf{p}_n}{\hbar}), \quad x_0 = \sqrt{\frac{\hbar}{\pi\mu\omega}}, \tag{5.139}$$

and they coincide with the $\mathbf{a}_n$ introduced in the Fourier expansion (5.125). For the closed string we follow the same procedure, but we can see that the result is similar except for the doubling in the number of oscillators

$$H_{closed} = \frac{\mathbf{p}_0^2}{2\pi\mu} + \sum_{n=1}^{\infty}\hbar\omega n\left(\mathbf{a}_n^\dagger\cdot\mathbf{a}_n + \tilde{\mathbf{a}}_n^\dagger\cdot\tilde{\mathbf{a}}_n + d\right). \tag{5.140}$$

There is one additional requirement for the closed string. It should not matter where the origin of the sigma coordinate is chosen along the string since there really is no distinguishable starting point on a loop. Therefore, the energy of the string should remain invariant under replacements of σ by $\sigma+\sigma_0$ for any σ_0. In other words, the physically acceptable states of the closed string are those that do not change under $\sigma-$translations. This is possible if the momentum along the string, given by the difference of left moving energy and right moving energy, vanishes. Therefore a physical state of a closed string is one that has an equal amount of left moving as right moving energy

$$\sum_{n=1}^{\infty} n\;\mathbf{a}_n^\dagger\cdot\mathbf{a}_n = \sum_{n=1}^{\infty} n\;\tilde{\mathbf{a}}_n^\dagger\cdot\tilde{\mathbf{a}}_n. \tag{5.141}$$

We see that the vacuum energy for either open or closed strings is infinite due to the vacuum energies of an infinite number of harmonic oscillators. However, only energy differences of excited levels relative to the ground state are physically significant or measurable. So we may redefine the ground state energy as being a constant or zero, by subtracting the infinity (or redefining the Hamiltonian from the beginning). This process is an example of "renormalization" which is commonly used in quantum field theory. It is a process that subtracts infinities that are not measurable. As a matter of fact the string theory discussed in this section is an example of a free field theory. The quantum states of the string are now quite evident. The simultaneously diagonalizable operators are the center of mass momentum and the number operators for each mode in each direction. The ground state, which is the vacuum of the oscillators corresponds to a string contracted to a single point and moving with momentum $\mathbf{p}_0 = \hbar\mathbf{k}$.

$$|\mathbf{k},0\rangle \tag{5.142}$$

Excited states are obtained by applying creation operators for any normal mode. The larger the index n the higher is the energy quantum associated with that normal mode. There are degeneracies at various energy levels. For example, for the open string, at energy level 2 the states

$$a^\dagger_{2I}\,|\mathbf{k},0\rangle, \quad a^\dagger_{1I}\;a^\dagger_{1J}\;|\mathbf{k},0\rangle \tag{5.143}$$

are degenerate and they both have energy

$$E = \frac{\hbar^2\mathbf{k}^2}{2\pi\mu} + 2\hbar\omega. \tag{5.144}$$

There are d states of the first kind and $d(d+1)/2$ of the second kind. The degeneracy due to the vector indices $I = 1, 2, \cdots d$ is related to rotation invariance, but the remaining degeneracies are properties of the quantum string theory. The states of the closed string must obey also the constraint (5.141). A few of the lowest energy closed string states and their energies are

$$\begin{array}{ll} |\mathbf{k},0\rangle & E = \frac{\hbar^2\mathbf{k}^2}{2\pi\mu} \\ a^\dagger_{1I}\;\tilde{a}^\dagger_{1J}\;|\mathbf{k},0\rangle & E = \frac{\hbar^2\mathbf{k}^2}{2\pi\mu} + 2\hbar\omega \\ \left.\begin{array}{l} a^\dagger_{2J}\;\tilde{a}^\dagger_{2K}\;|\mathbf{k},0\rangle \\ \tilde{a}^\dagger_{1I}\;\tilde{a}^\dagger_{1J}\;a^\dagger_{2K}\;|\mathbf{k},0\rangle \\ a^\dagger_{1I}\;a^\dagger_{1J}\;\tilde{a}^\dagger_{2K}\;|\mathbf{k},0\rangle \end{array}\right\} & E = \frac{\hbar^2\mathbf{k}^2}{2\pi\mu} + 4\hbar\omega. \end{array} \tag{5.145}$$

The occurrence of degeneracy in the last three states, and at higher states, is a property of string theory. In this section we have studied some aspects of the non-relativistic quantum string. The relativistic quantum string which is at the basis of the fundamental String Theory that attempts to unify all interactions in Nature is studied with very similar methods. The additional ingredient in that case is the use of relativistic string vectors $x_\mu(\sigma,\tau)$ that include a string-like time coordinate $x_0(\sigma,\tau)$, while τ is the proper time. The formalism is supplemented with additional invariances and constraints that seek to eliminate "ghosts" or negative norm states from the Hilbert space. These additional complications are present because of the time coordinate $x_0(\sigma,\tau)$. However, after proper manipulation of these issues, the relativistic string theory, treated in the "light-cone" gauge, reduces to a formalism which is very close to the one studied in this chapter.

5.8 Problems

1. Compute the matrix elements of the operators $(\hat{x}^2)_{mn}$ and $(\hat{p}^2)_{mn}$ in the harmonic oscillator states $|n\rangle$ and extract the uncertainties $(\Delta x)_n$, $(\Delta p)_n$ and their product in each state . Is the result consistent with what you expect on the basis of classical mechanics considerations?

2. Compute the time dependence of the uncertainties $\Delta x(t)$ and $\Delta p(t)$ in a coherent state $|z,t\rangle$. Interpret the result in light of the almost classical behavior of the coherent state.

3. Show that the state $\exp\left(-a^{\dagger 2}/2 + \lambda a^\dagger\right)|0\rangle$ is an eigenstate of the position operator $\hat{x}$. What is the relation between λ and the eigenvalue of $\hat{x}$? Similarly, construct an eigenstate of the momentum operator $\hat{p}$ with eigenvalue p.

4. Show that Wick theorem eq.(5.54) is correct for $n = 2, 3, 4$ by explicitly rearranging the operator products $\phi_1\phi_2\cdots\phi_n$ in terms of normal ordered products.

5. Consider the operator $U = \exp(\alpha a^\dagger a)$. Show that

$$UaU^{-1} = e^{-\alpha}a, \quad Ua^\dagger U^{-1} = e^{\alpha}a^\dagger. \tag{5.146}$$

(recall $e^A B e^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \cdots$.) Using this result show that, for any function of the creation-annihilation operators, one has $Uf(a,a^\dagger)U^{-1} = f(e^{-\alpha}a, e^{\alpha}a^\dagger)$.

6. Verify that the $SU(2)$ Lie algebra of eq.(5.76) is satisfied by the fermionic construction in (5.90). What is J^2 in terms of fermion number operators? Using this result verify that the only possible values for j are $0, 1/2$.

7. Consider the three operators

$$A_0 = \frac{1}{2}a^\dagger a + \frac{1}{4}, \quad A_+ = \frac{1}{2}a^\dagger a^\dagger, \quad A_- = \frac{1}{2}aa\ . \tag{5.147}$$

Obviously A_0 is diagonal on Fock space $|n\rangle$. Show that the $A_{\pm,0}$ close under commutation rules that differ from those of $J_{\pm,0}$ in eq.(5.76) by a minus sign. These are the commutation rules of the Lie algebra $SL(2,R)$ (equivalent to the Lorentz group $\mathrm{SO}(2,1)$ in 2-space + 1-time dimensions). Show that $A^2 = A_0^2 - (A_+A_- + A_-A_+)/2$ is a constant (not an operator), and by writing it in the form $j(j+1)$ identify two fixed values of j =?. Next find the action of $A_{\pm,0}$ on Fock space and show that it can be written in a form similar to eq.(5.74) with *some sign modifications* and identify again the fixed values of j =?.

8. Apply the $A_{\pm,0}$ operators of the previous problem on coherent states $|z\rangle$ and find their action in terms of differential operators. Verify that the resulting differential operators satisfy the same commutation rules as the previous problem. What is the action of these operators on the functions $\psi_m(z) = \langle m|z\rangle = z^m/\sqrt{m!}$? Rewrite the result in terms of the ψ_m and compare to the previous problem.

9. Consider the coherent states $|z_1, z_2\rangle$ for the two dimensional harmonic oscillator. What is the differential operator form of the $J_{\pm,0}$ and $\hat{N}$ of eqs.(5.72) acting on these states ?

10. Verify that $J^2 = (\hat{N}/2)(\hat{N}/2+1)$ in terms of oscillators and in terms of the differential operators constructed in the previous problem

11. Solve the frequencies for the three body problem given in eqs.(5.91,5.92) in d-dimensions . Hint: the solution is simplified by noticing that one of the frequencies must be zero due to the translation invariance of the problem. The corresponding new position and momentum variables must be the center of mass variables. After taking into account this fact from the beginning, the problem reduces to diagonalizing a 2×2 matrix involving only two independent relative variables.

12. Consider a Hamiltonian of the form

$$H = \frac{1}{2}\sum_{i,j=1}^{N=2}\left(K_{ij}\vec{p}_i \cdot \vec{p}_j + V_{ij}\vec{x}_i \cdot \vec{x}_j\right) + \sum_{i=1}^{N=2}\left(\vec{\alpha}_i \cdot \vec{p}_i + \vec{\beta}_i \cdot \vec{x}_i\right)$$

For the $N = 2$ case, once the the normal modes are obtained, the Hamiltonian takes the form

$$H = E\left(x_0, p_0\right) + \hbar\omega_1\left(\vec{a}_1^{\dagger} \cdot \vec{a}_1 + \frac{d}{2}\right) + \hbar\omega_2\left(\vec{a}_2^{\dagger} \cdot \vec{a}_2 + \frac{d}{2}\right).$$

Find $E\left(x_0, p_0\right), \omega_1, \omega_2$ in terms of the original parameters in $K_{ij}, V_{ij}, \vec{\alpha}_i, \vec{\beta}_i$.

13. The matrices A, B, C, D introduced in Eq.(5.100) in terms of b, c, u, form a $2N \times 2N$ symplectic matrix M which satisfies $M^T S M = S$. This M is also required to block diagonalize the Hamiltonian

$$\begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix}\begin{pmatrix} K & W \\ W^T & V \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \tilde{K} & 0 \\ 0 & \tilde{V} \end{pmatrix}.$$

Consider the $N = 2$ case, use explicit 2×2 matrices K, V, W,

$$K = \begin{pmatrix} k_1 & k_3 \\ k_3 & k_2 \end{pmatrix},\ V = \begin{pmatrix} v_1 & v_3 \\ v_3 & v_2 \end{pmatrix},\ W = \begin{pmatrix} w_1 & w_3 + w_4 \\ w_3 - w_4 & w_2 \end{pmatrix}$$

and find A, B, C, D (equivalently, b, c, u) and $\tilde{K}, \tilde{V}$, in terms of the parameters in K, V, W.

14. Write out explicitly all the equations that relate the original parameters in (5.93) and the transformation matrices at each step and show the overall relations between the frequencies of the normal modes and the original parameters. Take into account the results of the previous problem.

15. Consider the 3-dimensional harmonic oscillator and its 3 creation-annihilation operators. In addition to the total number operator $\hat{N} = a_1^{\dagger}a_1 + a_2^{\dagger}a_2 + a_3^{\dagger}a_3$ define the following 8 operators that commute with it

$$\begin{array}{l} J_0 = \frac{1}{2}(a_1^{\dagger}a_1 - a_2^{\dagger}a_2), \quad J_+ = a_1^{\dagger}a_2, \quad J_- = a_2^{\dagger}a_1 \\ U_+ = a_1^{\dagger}a_3, \quad U_- = a_3^{\dagger}a_1, \quad V_+ = a_2^{\dagger}a_3, \quad V_- = a_3^{\dagger}a_2 \\ Y = \frac{1}{3}(a_1^{\dagger}a_1 + a_2^{\dagger}a_2 - 2\,a_3^{\dagger}a_3). \end{array} \tag{5.148}$$

Using the commutation rules in (5.75) it is not difficult to show that these 8 operators close under commutation. Hence they form the SU(3) Lie algebra. By analogy to the SU(2) states in (5.68) the number states $|n_1, n_2, n_3\rangle$ may be rewritten in terms of the eigenvalues of the commuting operators $\hat{N}, J_0, Y$ with eigenvalues n, m, y respectively

$$|n, m, y\rangle = \frac{(a_1^\dagger)^{n/3+y/2+m}}{\sqrt{(n/3+y/2+m)!}} \frac{(a_2^\dagger)^{n/3+y/2-m}}{\sqrt{(n/3+y/2-m)!}} \frac{(a_3^\dagger)^{n/3-y}}{\sqrt{(n/3-y)!}} |0\rangle \tag{5.149}$$

where we have defined

$$n = n_1 + n_2 + n_3, \quad m = \frac{1}{2}(n_1 - n_2), \quad y = \frac{1}{3}(n_1 + n_2 - 2n_3). \tag{5.150}$$

Furthermore, by concentrating only on the first two oscillators and comparing to the two dimensional case it is useful to define the quantum number j

$$j = \frac{1}{2}(n_1 + n_2) = n/3 + y/2. \tag{5.151}$$

So, the state may be labelled $|n, j, m, y\rangle$ with the additional quantum number j, but keeping in mind that j is a function of n, y but not m. From the positivity of the integers $n_I \geq 0$ one derives the allowed values and ranges of these quantum numbers

$$\begin{array}{rcl} n & = & 0, 1, 2, 3, \cdots . \\ (j, y) & = & (\frac{n}{2}, \frac{n}{3}), \cdots , \ (\frac{n-k}{2}, \frac{n-3k}{3}), \cdots , (0, -\frac{2n}{3}) \\ m & = & -j, -j+1, \cdots , \ j-1, \ j \end{array} \tag{5.152}$$

- Compute the action of the 3 creation and annihilation operators on the states, writing the resulting state in terms of the labels of the type $|n, j, m, y\rangle$.
- Using this result now obtain the action of the 8 operators which do not change the value of n, and again give your result in terms of the states labelled by $|n, j, m, y\rangle$. Since the value of n cannot change you have found irreducible representations of SU(3) for each value of n. You will compare your result later in the course to the general solution of the SU(3) problem and find out that it corresponds to a *subclass* of irreducible representations (see problems in Chapter 8).

16. Consider the 8 operators of SU(3) of the previous problem but now constructed with fermionic oscillators. Verify that the Lie algebra closes with identical coefficients as the bosonic case. Show that the list of all possible values of the quantum numbers (n, j, y) that may be constructed with fermions is (take into account the SU(2) result discussed in the text)

$$\begin{array}{rcl} (n, j, y) & = & (0, 0, 0), \ (1, 0, -2/3), \ (1, 1/2, 1/3) \\ & & (2, 0, 2/3), \ (2, 1/2, -1/3), \ (3, 0, 0) \\ j & = & \frac{1}{2}\left[\left(\frac{2}{3}(n \bmod 3) + y\right) \bmod 2\right] . \end{array} \tag{5.153}$$

Of course, for each value of j the remaining quantum number must take values $-j \leq m \leq j$. If we interpret j as isospin and y as hypercharge, then the three states with $n = 1$ have the quantum numbers of up, down, strange quarks, while the three states with $n = 2$ have the quantum numbers of antiquarks (also of di-quarks). The states with $n = 0, 3$ are SU(3) singlets. There are other applications and/or interpretations of the mathematics. For example the three fermions $b^\dagger_\alpha$ with $\alpha = 1, 2, 3$ may represent the three colors of quarks. Then we see that we can make a color singlet state $(3, 0, 0)$ by putting three quarks of different colors together $b^\dagger_1 b^\dagger_2 b^\dagger_3|0\rangle$ since the 8 SU(3) generators annihilate this state.

17. Compute the vacuum expectation value

$$\langle 0|V(k_1, z_1)V(k_2, z_2)\cdots V(k_n, z_n)|0\rangle$$

where $V(k, z) =: \exp(ik \cdot x(z)) :$ is the normal ordered product of a string plane wave defined by

$$\mathbf{x}(z) = \mathbf{q} - i\mathbf{p}\ln z + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left(\mathbf{a}_n z^n + \mathbf{a}^\dagger_n z^{-n}\right), \quad z = e^{i\omega t}$$

$$V(k, z) = \left[\exp \sum_{n=1}^{\infty} \frac{iz^{-n}}{\sqrt{n}} \mathbf{k} \cdot \mathbf{a}^\dagger_n\right] e^{i\mathbf{k}\cdot\mathbf{q} + \mathbf{k}\cdot\mathbf{p}\ln z} \left[\exp \sum_{n=1}^{\infty} \frac{iz^n}{\sqrt{n}} \mathbf{k} \cdot \mathbf{a}_n\right]$$

Hint: use $e^A e^B = e^B e^A e^{[A,B]}$ to move creation operators to the left and annihilation operators to the right. The final result depends only on the c-number exponential $e^{[A,B]}$. The sums in the exponents of $e^{[A,B]}$ can be easily performed and simplified. This computation is related to n strings scattered from each other.

Chapter 6

CENTRAL FORCE PROBLEM

The central force problem describes the interaction of two bodies through a force that depends only on the distance between them and acts in the direction of the vector that joins them. The gravitational interactions of planets moving around the sun, the electromagnetic interactions of electrons around nuclei, the nuclear interactions of two bodies in collision, etc. are all systems described by the central force problem. Because of the large number of applications, it is worth studying the system in some generality without specifying the precise nature of the force. The Lagrangian and Hamiltonian are given by

$$\begin{aligned} L &= \tfrac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \tfrac{1}{2}m_1\dot{\mathbf{r}}_1^2 - V(|\mathbf{r}_1 - \mathbf{r}_2|) \\ H &= \tfrac{\mathbf{p}_1^2}{2m_1} + \tfrac{\mathbf{p}_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|), \end{aligned} \tag{6.1}$$

where the potential energy $V(r)$ depends only on the distance between the two bodies $r = |\mathbf{r}_1 - \mathbf{r}_2|$. We do not need to specify this function until we discuss a particular system, so we will treat it generally for the time being. The classical Hamilton's equations of motion $\dot{\mathbf{r}}_i = \partial_{\mathbf{p}_i} H$, $\dot{\mathbf{p}}_i = -\partial_{\mathbf{r}_i} H$ give

$$\dot{\mathbf{r}}_1 = \frac{\mathbf{p}_1}{m_1}, \quad \dot{\mathbf{r}}_2 = \frac{\mathbf{p}_2}{m_2}, \quad \dot{\mathbf{p}}_1 = -\dot{\mathbf{p}}_2 = -\frac{\mathbf{r}_1 - \mathbf{r}_2}{r}V'(r). \tag{6.2}$$

where $V'(r) = \frac{\partial}{\partial r}V$. One may then note that, independent of the details of the potential energy, there are several quantities that remain constant throughout the motion. These are the total momentum $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$, the total angular momentum $\mathbf{L}_T = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2$, the total energy H, the relative orbital angular momentum, and the energy of the relative system. To see this and provide simple expressions for the relative quantities, it is useful to introduce center of mass and relative coordinates ($\mathbf{R}$, $\mathbf{r}$) and rewrite the Hamiltonian and equations of motion in terms of them. We define

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{6.3}$$

Inverting these expressions one has

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{M}\mathbf{r}\ , \quad \mathbf{r}_2 = \mathbf{R} - \frac{m_1}{M}\mathbf{r} \tag{6.4}$$

where $M = m_1 + m_2$ is the total mass. The Lagrangian and Hamiltonian take the form

$$\begin{array}{l} L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 - V(|\mathbf{r}|) \\ H = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} + V(|\mathbf{r}|) \end{array} \tag{6.5}$$

where μ is the reduced mass $\mu = m_1 m_2/M$, and

$$\begin{array}{l} \mathbf{P} = \frac{\partial L}{\partial \dot{\mathbf{R}}} = M\dot{\mathbf{R}} = \mathbf{p}_1 + \mathbf{p}_2, \\ \mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = \mu\dot{\mathbf{r}} = (m_2\mathbf{p}_1 - m_1\mathbf{p}_2)/M, \\ \mathbf{p}_1 = \frac{m_1}{M}\mathbf{P} + \mathbf{p}, \quad \mathbf{p}_2 = \frac{m_1}{M}\mathbf{P} - \mathbf{p} \end{array} \tag{6.6}$$

The equations of motion for the center of mass and relative variables follows from the new form of the Hamiltonian, or from the original equations of motion

$$\dot{\mathbf{R}} = \frac{\mathbf{P}}{M}, \quad \dot{\mathbf{P}} = 0, \quad \dot{\mathbf{r}} = \frac{\mathbf{p}}{\mu}, \quad \dot{\mathbf{p}} = -\frac{\mathbf{r}}{r}V'(r). \tag{6.7}$$

Thus, the center of mass moves with a constant velocity, like a free particle, while the relative motion depends on the force $\mathbf{F} = -\frac{\mathbf{r}}{r}V'(r)$.

6.1 Separation of center of mass

It is clear that the center of mass and relative motions are decoupled from each other and the conserved quantities are expressed as constants of either the center of mass or relative motions. They are

$$\begin{array}{ll} \mathbf{P}_{cm} = \mathbf{p}_1 + \mathbf{p}_2, & \mathbf{L}_{cm} = \mathbf{R} \times \mathbf{P} \\ H_{rel} = \frac{\mathbf{p}^2}{2\mu} + V(r), & \mathbf{L}_{rel} = \mathbf{r} \times \mathbf{p}. \end{array} \tag{6.8}$$

For example we can check

$$\partial_t \mathbf{L}_{rel} = \dot{\mathbf{r}} \times \mathbf{p} + \mathbf{r} \times \dot{\mathbf{p}} = \frac{\mathbf{p}}{\mu} \times \mathbf{p} + \mathbf{r} \times \left(-\frac{\mathbf{r}}{r}V'(r)\right) = 0. \tag{6.9}$$

In a later chapter on symmetries we will see that the fundamental reason for the conservation of these quantities are symmetries of the system under space translations ($\dot{\mathbf{P}} = 0$), time translations ($\dot{H} = 0$) and rotations ($\dot{\mathbf{L}}_{cm} = \dot{\mathbf{L}}_{rel} = 0$). These symmetries are evident when we examine the Lagrangian or Hamiltonian. Either one of these expressions remains unchanged under these transformations (see symmetry chapter), therefore the equations of motions derived from them carry the information about the symmetries.

The importance of this fact for Quantum Mechanics is that the operators corresponding to these physical quantities must commute with the Hamiltonian, since the time derivative of an operator is obtained by commuting it with the

Hamiltonian. To see this result directly one may try the total momentum as an example

$$\begin{aligned}[H,\mathbf{P}] &= [H,\mathbf{p}_1+\mathbf{p}_2] \\ &= [V(\mathbf{r}_1-\mathbf{r}_2),\mathbf{p}_1]+[V(\mathbf{r}_1-\mathbf{r}_2),\mathbf{p}_2] \\ &= i\hbar(\boldsymbol{\nabla}_1+\boldsymbol{\nabla}_2)\ V(\mathbf{r}_1-\mathbf{r}_2) \\ &= 0.\end{aligned} \tag{6.10}$$

It works similarly for the other conserved quantities (see problems 1,2). This means that the conserved quantities are simultaneous observables with the Hamiltonian and their eigenvalues can serve to label the states of a complete Hilbert space. It is imperative to take advantage of this fact in order to exhibit the symmetries of the quantum system and obtain a more tractable and understandable solution of the quantum system.

What should be the labels of the complete Hilbert space? To begin, it is evident that we need 6 labels corresponding to 6 canonical degrees of freedom. These can be any of the following $|\mathbf{r}_1,\mathbf{r}_2\rangle$ or $|\mathbf{R},\mathbf{r}\rangle$, or $|\mathbf{P},\mathbf{p}\rangle$, or $|\mathbf{P},\mathbf{r}\rangle$, etc. Among these it is wise to pick the label for total momentum since it is conserved and also diagonalizes the center of mass energy. In fact, since the Hamiltonian is constructed from two decoupled parts $H = H_{cm}(\mathbf{P}) + H_{rel}(\mathbf{p},\mathbf{r})$ we are completely finished with diagonalizing the first part, and can concentrate only on the relative Hamiltonian for any fixed value of the total momentum. The objective is then to diagonalize $H_{rel}(\mathbf{r},\mathbf{p})$, and we must find operators constructed from $(\mathbf{r},\mathbf{p})$ that commute with it, to label a complete Hilbert space. By virtue of the conservation of the relative angular momentum $\mathbf{L}_{rel}= \mathbf{r}\times\mathbf{p}$, and the fact that it is constructed only from relative canonical operators, we have automatically (see problem)

$$[\mathbf{L}_{rel}, H_{rel}] = 0. \tag{6.11}$$

Therefore, any function of $\mathbf{L}_{rel}$ commutes with H_{rel}, and we can find two such functions $(\mathbf{L}^2_{rel},(\mathbf{L}_{rel})_z)$ that commute not only with H_{rel} but also with each other. This will be discussed below in more detail, but for now we treat it as a given in order to complete the description of the general structure. Thus, a complete set of 6 simultaneous observables, that includes the Hamiltonian, is

$$\mathbf{P},\ H_{rel},\ \mathbf{L}^2_{rel},\ (\mathbf{L}_{rel})_z. \tag{6.12}$$

we may label the complete Hilbert space with their eigenvalues

$$|\mathbf{k},E,l,m\rangle. \tag{6.13}$$

Another complete set is, of course, position space $|\mathbf{R},\mathbf{r}\rangle$. One set may be expanded in terms of the other, and the expansion coefficients are the position space probability amplitudes $\langle\mathbf{R},\mathbf{r}|\mathbf{k},E,l,m\rangle$. Because of the separate nature of the center of mass and relative Hamiltonians, their corresponding states form direct product spaces $|\mathbf{R},\mathbf{r}\rangle \equiv |\mathbf{R}\rangle\otimes|\mathbf{r}\rangle$, and $|\mathbf{k},E,l,m\rangle \equiv |\mathbf{k}\rangle\otimes|E,l,m\rangle$. For this reason the wavefunction separates into two factors

$$\langle\mathbf{R},\mathbf{r}|\mathbf{k},E,l,m\rangle = \langle\mathbf{R}|\mathbf{k}\rangle\langle\mathbf{r}|E,l,m\rangle \equiv \frac{e^{i\mathbf{k}\cdot\mathbf{R}}}{(2\pi\hbar)^{3/2}}\ \psi_{Elm}(\mathbf{r}). \tag{6.14}$$

The first factor describes the free motion of the center mass. It is well known from the study of the free particle since the operators $\mathbf{R}, \mathbf{P}$ and the center of mass Hamiltonian constructed from them satisfy the same mathematics as the free particle. The second factor is the non-trivial probability amplitude for the relative motion.

We may examine the Schrödinger equation with the total Hamiltonian $H_{tot}|\psi\rangle = E_{tot}|\psi\rangle$. The equation in position space is obtained by taking the inner product with $\langle \mathbf{R}, \mathbf{r}|$

$$\left(\frac{-\hbar^2}{2M}\boldsymbol{\nabla}_R^2 + \frac{-\hbar^2}{2\mu}\boldsymbol{\nabla}_r^2 + V(r)\right)\psi(\mathbf{R},\mathbf{r}) = E_{tot}\psi(\mathbf{R},\mathbf{r}). \tag{6.15}$$

As argued above this equation is solved by the product

$$\psi(\mathbf{R},\mathbf{r}) = \psi_{cm}(\mathbf{R})\ \psi_{rel}(\mathbf{r}), \tag{6.16}$$

where each factor satisfies its own independent equation

$$\begin{array}{c} \frac{-\hbar^2}{2M}\boldsymbol{\nabla}_R^2\ \psi_{cm}(\mathbf{R}) = E_{cm}\ \psi_{cm}(\mathbf{R}); \quad \psi_{cm}(\mathbf{R}) = \frac{e^{i\mathbf{k}\cdot\mathbf{R}}}{(2\pi\hbar)^{3/2}}\,, \quad E_{cm} = \frac{\hbar^2 k^2}{2M} \\ \left(\frac{-\hbar^2}{2\mu}\boldsymbol{\nabla}_r^2 + V(r)\right)\psi_{rel}(\mathbf{r}) = E\psi_{rel}(\mathbf{r}); \quad \psi_{rel}(\mathbf{r}) = \psi_{Elm}(\mathbf{r}), \quad E_{tot} = E_{cm} + E \end{array} \tag{6.17}$$

The relative energy E, the wavefunction $\psi_{Elm}(\mathbf{r})$, and the quantum numbers (l, m) still remain to be computed.

6.2 Angular momentum commutators

In the previous section the central force problem was reduced to the study of an effective one-particle problem for the relative motion. From here on the subscript "*rel*" will be omitted to save some writing. The equivalent system describes a single particle interacting with a spherically symmetric potential. The Hamiltonian and angular momentum operators commute

$$H = \frac{\mathbf{p}^2}{2\mu} + V(r), \quad \mathbf{L} = \mathbf{r}\times\mathbf{p}, \quad [\mathbf{L}, H] = 0. \tag{6.18}$$

According to the discussion in the previous section the two operators must commute since angular momentum is conserved. That is, $\mathbf{L}$ does not change under infinitesimal time translations, and since H is the generator of time translations, $\mathbf{L}$ must commute with it. However, another interpretation of the zero commutator $[\mathbf{L}, H] = 0$ is that the Hamiltonian does not change under infinitesimal rotations. That point of view will be better appreciated after understanding that $\mathbf{L}$ is the generator of infinitesimal rotations, as follows.

The commutation rules of the angular momentum operators L_I with other operators may be explored. For this purpose the Levi-Civita symbol in three dimensions ϵ_{IJK} plays a useful role since cross products of vectors are written

succinctly in the form[1]

$$L_I = \epsilon_{IJK}\ r_J\, p_K \tag{6.19}$$

where a summation convention is used for repeated indices. From the basic canonical variables $(\mathbf{r}, \mathbf{p})$ it is possible to construct three linearly independent vectors $(\mathbf{r}, \mathbf{p}, \mathbf{L})$, and three independent scalars $\mathbf{r}^2, \mathbf{p}^2, \mathbf{r}\cdot\mathbf{p}$. Any function of the canonical variables may be constructed from these basic entities. Therefore, to find the commutator of L_I with any operator in the theory it is useful to compute the commutator with the basic vectors and scalars. The commutator of angular momentum with either position or momentum look similar

$$\begin{array}{l} [L_I, r_J] = \epsilon_{IMN}[r_M p_N, r_J] = i\hbar\epsilon_{IJK}\ r_K \\ [L_I, p_J] = \epsilon_{IMN}[r_M p_N, p_J] = i\hbar\epsilon_{IJK}\ p_K. \end{array} \tag{6.20}$$

Two angular momentum operators also produce a similar result

$$\begin{array}{ll} [L_I, L_J] & = \epsilon_{IKL}\ \epsilon_{JMN}[r_K p_L, r_M p_N] \\ & = \epsilon_{IKL}\ \epsilon_{JMN}\,(i\hbar\delta_{LM}\ r_K p_N - i\hbar\delta_{KN}\ r_M p_L) \\ & = i\hbar(r_I p_J - r_J p_I) \\ & = i\hbar\epsilon_{IJK}\ L_K. \end{array} \tag{6.21}$$

The commutator with any of the scalar functions is zero. For example, it follows from (6.20) that

$$[L_I, \mathbf{r}.\mathbf{p}] = [L_I, r_J]p_J + r_J[L_I, p_J] = i\hbar\epsilon_{IJK}\,(\,r_K p_J + r_J p_K) = 0.$$

Therefore, any scalar function $S(\mathbf{r}^2, \mathbf{p}^2, \mathbf{r}\cdot\mathbf{p})$ must commute with angular momentum

$$[L_I, S(\mathbf{r}^2, \mathbf{p}^2, \mathbf{r}\cdot\mathbf{p})] = 0. \tag{6.22}$$

[1] The Levi-Civita symbol can be expressed as a determinant

$$\epsilon_{ijk} = \begin{vmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \\ \delta_{1k} & \delta_{2k} & \delta_{3k} \end{vmatrix}.$$

This makes it evident that it is completely antisymmetric. It takes the value 0 if any two indices are the same, the value $\epsilon_{123} = 1$ when i, j, k are in the order $1, 2, 3$, or any cyclic permutation, and the value -1 for any odd permutation of $1, 2, 3$, such as $\epsilon_{213} = -1$. Furthermore the product of two such symbols has simple properties that follow from

$$\epsilon_{ijk}\ \epsilon_{mnk} = \begin{vmatrix} \delta_{im} & \delta_{in} & \delta_{il} \\ \delta_{jm} & \delta_{jn} & \delta_{jl} \\ \delta_{km} & \delta_{kn} & \delta_{kl} \end{vmatrix}.$$

In particular when one of the indices is summed it yields

$$\epsilon_{ijk}\ \epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm},$$

when two indices are summed it gives

$$\epsilon_{ijk}\ \epsilon_{mjk} = 2\delta_{im},$$

and when three indices are summed the result is

$$\epsilon_{ijk}\ \epsilon_{ijk} = 3!\ .$$

Furthermore, the commutator with any vector function $\mathbf{V} = \mathbf{r}S_1 + \mathbf{p}S_2 + \mathbf{L}S_3$, where the $S_{1,2,3}$ are scalar functions, follows from (6.20,6.21) and (6.22)

$$[L_I, V_J] = i\hbar\epsilon_{IJK}\ V_K. \tag{6.23}$$

So, there is a universal rule for the commutation rules of angular momentum: L_I commutes with any scalar and it gives the same commutator with any vector. What is the reason? To answer the question introduce an infinitesimal constant vector $\boldsymbol{\omega}$ and examine the combination $i\boldsymbol{\omega}\cdot\mathbf{L}/\hbar$. Its commutator with any vector $\mathbf{V}$ may be written as an infinitesimal rotation[2] of the vector $\delta_\omega\mathbf{V} = (\boldsymbol{\omega}\times\mathbf{V})$ by an infinitesimal angle $\boldsymbol{\omega}$

$$[\frac{i\boldsymbol{\omega}\cdot\mathbf{L}}{\hbar}, V_J] = \frac{i\omega_I}{\hbar}[L_I, V_J] = -\epsilon_{IJK}\ \omega_I V_K = (\boldsymbol{\omega}\times\mathbf{V})_J = \delta_\omega\mathbf{V}_J\ . \tag{6.24}$$

Therefore L_I is the generator of infinitesimal rotations around the axis I. It is now clear why the commutator of angular momentum must be zero with any scalar: it is because the scalar does not rotate. It is also equally clear why its commutator with any vector must be universal: it is because all vectors must rotate with the same formula $\delta_\omega\mathbf{V} = (\boldsymbol{\omega}\times\mathbf{V})$.

These results now give a better insight on $[\mathbf{L}, H] = 0$. It means that the Hamiltonian is a scalar, and therefore it does not rotate. By understanding that $\mathbf{L}$ is the generator of rotations we know the result of the commutation even before computing it explicitly, since by construction the Hamiltonian is rotationaly invariant. The same general rule applies to the commutation with any vector: the vector property of $\mathbf{V}$ guarantees that its commutator with angular momentum must be (6.23). Also, the commutator of L_I with other components of angular momentum and with the square of the vector follow from the general remarks without any explicit computation. So, just from the *property of rotations* we must have

$$[L_I, L_J] = i\hbar\epsilon_{IJK}\ L_K, \quad [L_I, \mathbf{L}^2] = 0. \tag{6.25}$$

From this analysis we conclude that we can simultaneously diagonalize the operators $H, \mathbf{L}^2, L_3$, and label the complete set of states with their eigenvalues as $|E, l, m\rangle$, as outlined in the previous section. By definition we write the eigenvalue equations

$$\begin{aligned} \mathbf{L}^2\,|E,l,m\rangle &= \hbar^2 l(l+1)\ |E,l,m\rangle \\ L_3\,|E,l,m\rangle &= \hbar m\ |E,l,m\rangle \\ H\ |E,l,m\rangle &= E\ |E,l,m\rangle \end{aligned} \tag{6.26}$$

The eigenvalue of $\mathbf{L}^2$ is parametrized as $\hbar^2 l(l+1)$ for later convenience, but at this stage since the values of l are not specified, this form is to be considered just

[2] An infinitesimal rotation of a vector $\mathbf{V}$ around some axis by an infinitesimal angle ω may be represented by the vector $\boldsymbol{\omega}$ whose magnitude is the angle of rotation and whose direction is the axis of rotation. The infinitesimal change in the vector is then given by the cross product $\delta_\omega\mathbf{V} = \boldsymbol{\omega}\times\mathbf{V}$. The reader who may not be familiar with this formula is urged to try it for $\boldsymbol{\omega}$ pointing in the z-direction and the vector $\mathbf{V}$ lying in the (x, y) plane.

a symbol. The angular momentum eigenvalue problem for $\mathbf{L}^2, L_3$ is independent from the details of the Hamiltonian. Its solution applies quite generally to any problem that has rotational symmetry. It will be studied by itself in section (6.4).

6.3 Radial and angular operators

6.3.1 Radial operators

To go further introduce spherical coordinates (r, θ, ϕ). The direction of the vector $\mathbf{r}$ will be denoted by the solid angle symbol $\Omega = (\theta, \phi)$ and we will also use the bold character $\mathbf{\Omega}$ to denote the unit vector

$$\mathbf{\Omega} = \frac{\mathbf{r}}{r}. \tag{6.27}$$

Introduce the *Hermitian* operator p_r that represents the radial momentum

$$p_r = \frac{1}{2}(\frac{\mathbf{r}}{r} \cdot \mathbf{p} + \mathbf{p} \cdot \frac{\mathbf{r}}{r}). \tag{6.28}$$

Its commutator with the position operator is computed from $[\mathbf{r}_I, \mathbf{p}_J] = i\hbar\delta_{IJ}$, which gives

$$[\mathbf{r}_I, p_r] = i\hbar\mathbf{\Omega}_I. \tag{6.29}$$

Furthermore, recall that $[f(r), \mathbf{p}] = i\hbar\boldsymbol{\nabla} f(r)$. From this one can derive that p_r is the canonical conjugate to $r = \sqrt{\mathbf{r}^2}$ and that it commutes with the unit vector (or the angles θ, ϕ) , i.e.

$$[r, p_r] = i\hbar, \ [\mathbf{\Omega}, p_r] = 0.$$

Furthermore, since (r, p_r) are scalars constructed from dot products, they both commute with angular momentum

$$[L_I, r] = 0 = [L_I, p_r].$$

Thus, the basic commutation rules among the cartesian coordinates $(\mathbf{r}_I, \mathbf{p}_J)$ have been transformed to the following basic commutation rules among the spherical coordinates $(r, p_r, \mathbf{\Omega}, \mathbf{L})$

$$[r, p_r] = i\hbar, \quad [L_I, L_J] = i\hbar\epsilon_{IJK}L_K, \quad [L_I, \mathbf{\Omega}_J] = i\hbar\epsilon_{IJK}\mathbf{\Omega}_K, \tag{6.30}$$

while all other commutators among $(r, p_r, \mathbf{\Omega}, \mathbf{L})$ vanish.

Next we are interested in the inverse relations. Using the definitions of these operators in terms of the basic cartesian operators $(\mathbf{r}, \mathbf{p})$ it is straightforward to obtain the decomposition of position and momentum in terms of *Hermitian* radial and angular operators in spherical coordinates (see problems 4,5)

$$\mathbf{r} = r\mathbf{\Omega}, \quad \mathbf{p} = \mathbf{\Omega}\, p_r - \frac{1}{2r}(\mathbf{\Omega} \times \mathbf{L} - \mathbf{L} \times \mathbf{\Omega}). \tag{6.31}$$

Furthermore, keeping track of order of operators, one finds

$$\begin{aligned}\mathbf{L}^2 &= (\mathbf{r}\times\mathbf{p})\cdot(\mathbf{r}\times\mathbf{p}) \\ &= \mathbf{r}^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})(\mathbf{p}\cdot\mathbf{r}) - 2i\hbar\mathbf{r}.\mathbf{p} \\ &= \mathbf{r}^2(\mathbf{p}^2 - p_r^2)\end{aligned} \tag{6.32}$$

from which it follows that

$$\mathbf{p}^2 = p_r^2 + \frac{1}{r^2}\mathbf{L}^2. \tag{6.33}$$

The Hamiltonian now takes the form[3]

$$H = \frac{1}{2\mu}(p_r^2 + \frac{1}{r^2}\mathbf{L}^2) + V(r). \tag{6.34}$$

When H is applied on the common eigenvectors $|E,l,m\rangle$, $\mathbf{L}^2$ becomes a number and then the Schrödinger equation involves only the radial operators r, p_r

$$H|E,l,m\rangle = \left(\frac{1}{2\mu}(p_r^2 + \frac{\hbar^2 l(l+1)}{r^2}) + V(r)\right)|E,l,m\rangle = E\ |E,l,m\rangle. \tag{6.35}$$

Therefore, the solution of the central force problem is reduced to the solution of this simplified eigenvalue equation and the decoupled equations involving only angular momentum operators as in (6.26).

To investigate the probability amplitude $\psi_{Elm}(\mathbf{r}) = \langle\mathbf{r}|E,l,m\rangle$ in position space we label the bra in spherical coordinates

$$\begin{gathered}\langle\mathbf{r}| \equiv \langle r,\Omega| \equiv \langle r,\theta,\phi| \\ 0\le r<\infty, \quad 0\le\theta\le\pi, \quad 0\le\phi\le 2\pi.\end{gathered} \tag{6.36}$$

Since the angular and radial systems of operators commute with each other, the Hilbert space reduces to direct products $|r\rangle\otimes|\Omega\rangle$ and $|El\rangle\otimes|lm\rangle$, while the wavefunction decomposes to radial and angular factors[4]

$$\psi_{Elm}(\mathbf{r}) = R_{El}(r)\ Y_{lm}(\Omega), \tag{6.37}$$

where

$$R_{El}(r) = \langle r|El\rangle, \quad Y_{lm}(\Omega) = \langle\Omega|lm\rangle. \tag{6.38}$$

[3] For a generalization of this result to d dimensions, including two dimensions or higher dimension, see section (6.7) and problem 9.

[4] In d dimensions, including $d=3$, this is generalized to $\psi(\mathbf{r}) = R_{El}(r)T_{I_1I_2\cdots I_l}(\mathbf{\Omega})$, where the angular momentum wavefunction $Y_{lm}(\mathbf{\Omega})$ for angular momentum l is replaced by the tensor $T_{I_1I_2\cdots I_l}(\mathbf{\Omega})$ which is the completely symmetric tensor with l indices constructed from the unit vector $\mathbf{\Omega}_I$ in d dimensions, $I=1,2,\cdots d$. The tensor must also be "traceless" in the sense that when any pair of indices are set equal to each other and summed the result must be zero, such as, $\sum_{I=1}^{d} T_{III_3\cdots I_l} = 0$. Therefore, $T_{I_1I_2\cdots I_l}(\mathbf{\Omega}) = \mathbf{\Omega}_{I_1}\mathbf{\Omega}_{I_2}\mathbf{\Omega}_{I_3}\cdots\mathbf{\Omega}_{I_l} - \frac{1}{a(l,d)}\left(\delta_{I_1I_2}\mathbf{\Omega}_{I_3}\cdots\mathbf{\Omega}_{I_l} + \textit{permutations}\right)$, where the permutations insure that the subtracted factor is completely symmetric and the coefficient $a(l,d)$ is chosen so that $\sum_{I=1}^{d} T_{III_3\cdots I_l} = 0$. Examples are given in Eqs.(6.90). These tensors are the solutions of the angular part of the Schrödinger equation in d dimensions just as the spherical harmonics $Y_{lm}(\Omega)$ are the corresponding solutions in $d=3$ dimensions. Indeed, when specialized to $d=3$, the components of the traceless tensor with l indices are proportional to the spherical harmonics $Y_{lm}(\Omega)$ (see Eq.(6.91) and problem 7).

The quantum numbers (E, l) did not fully separate because they are coupled through the Schrödinger equation above. We need to compute these functions. It is evident that the angular part $Y_{lm}(\Omega)$ is universal for any central force problem, and is independent from the potential, while the radial part depends on the specific Hamiltonian.

The three spherical orthogonal unit vectors will be denoted by bold characters $\mathbf{\Omega}, \mathbf{\Theta}, \mathbf{\Phi}$.

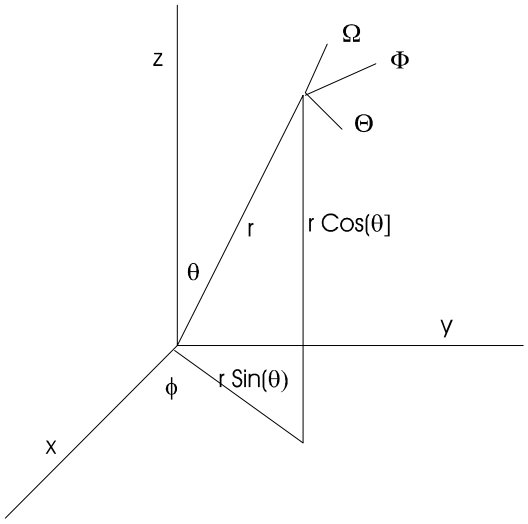


Fig.6.1- Spherical unit vectors $\vec{\Omega}, \vec{\Theta}, \vec{\Phi}$.

From Fig.(6.1) one can construct graphically the Cartesian components of the unit vectors

$$\begin{aligned} \mathbf{\Omega} &= (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta) \\ \mathbf{\Theta} &= (\cos\theta\cos\phi,\ \cos\theta\sin\phi, -\sin\theta) \\ \mathbf{\Phi} &= (-\sin\phi,\ \cos\phi,\ 0) \end{aligned} \tag{6.39}$$

Using

$$\begin{aligned} &\frac{\partial}{\partial\theta}\mathbf{\Omega} = \mathbf{\Theta}, && \frac{\partial}{\partial\phi}\mathbf{\Omega} = \sin\theta\ \mathbf{\Phi} \\ &\frac{\partial}{\partial\theta}\mathbf{\Theta} = -\mathbf{\Omega}, && \frac{\partial}{\partial\phi}\mathbf{\Theta} = \cos\theta\ \mathbf{\Phi} \\ &\frac{\partial}{\partial\theta}\mathbf{\Phi} = 0, . && \frac{\partial}{\partial\phi}\mathbf{\Phi} = -(\cos\theta\ \mathbf{\Theta} + \sin\theta\ \mathbf{\Omega}) \end{aligned} \tag{6.40}$$

one finds the small change in the vector $\mathbf{r}$ under small variations $(\delta r, \delta\theta, \delta\phi)$,

$$\begin{aligned} \delta\mathbf{r} &= \delta\left(r\mathbf{\Omega}\right) = \delta r\mathbf{\Omega} + r\delta\mathbf{\Omega} = \delta r\mathbf{\Omega} + r\delta\theta\frac{\partial}{\partial\theta}\mathbf{\Omega} + + r\delta\phi\frac{\partial}{\partial\phi}\mathbf{\Omega} \\ &= \delta r\,\mathbf{\Omega} + r\delta\theta\ \mathbf{\Theta}\ + r\sin\theta\ \delta\phi\,\mathbf{\Phi}. \end{aligned}$$

This is useful to compute the gradient in spherical coordinates by demanding that any wavefunction $\psi(\mathbf{r}) = \psi(r, \theta, \phi)$ must give the same infinitesimal variation in either the cartesian or spherical coordinates

$$\delta\psi = \delta\mathbf{r}\cdot\boldsymbol{\nabla}\psi = \delta r\frac{\partial\psi}{\partial r} + \delta\theta\frac{\partial\psi}{\partial\theta} + \delta\phi\frac{\partial\psi}{\partial\phi}.$$

Taking into account $\delta\mathbf{r}$ as given above, we see that we must require

$$\boldsymbol{\nabla} = \boldsymbol{\Omega}\ \frac{\partial}{\partial r} + \boldsymbol{\Theta}\ \frac{1}{r}\frac{\partial}{\partial\theta} + \boldsymbol{\Phi}\ \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\ . \tag{6.41}$$

We may now find differential operator expressions in spherical coordinates for the momentum and angular momentum operators when they act in position space $\langle\mathbf{r}|\mathbf{p} = -i\hbar\boldsymbol{\nabla}\langle r,\Omega|$, etc. We find (for a generalization of p_r and p_r^2 to d dimensions, see Eq.(6.97))

$$\begin{array}{l}
\mathbf{p} = -i\hbar\boldsymbol{\nabla} = -i\hbar\left(\boldsymbol{\Omega}\ \frac{\partial}{\partial r} + \boldsymbol{\Theta}\ \frac{1}{r}\frac{\partial}{\partial\theta} + \boldsymbol{\Phi}\ \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\\
\mathbf{L} = -i\hbar\mathbf{r}\times\boldsymbol{\nabla} = -i\hbar\left(-\boldsymbol{\Theta}\ \frac{1}{\sin\theta}\frac{\partial}{\partial\phi} + \boldsymbol{\Phi}\ \frac{\partial}{\partial\theta}\right)\\
p_r = -i\hbar(\frac{\partial}{\partial r} + \frac{1}{r}) = -i\hbar\frac{1}{r}\ \frac{\partial}{\partial r}\ r\\
p_r^2 = -\hbar^2\frac{1}{r}\frac{\partial^2}{\partial r^2}\ r
\end{array} \tag{6.42}$$

We can now derive differential equations satisfied by the radial and spherical wavefunctions by sandwiching some convenient operators between position energy or angular momentum states and evaluating them on both spaces. In particular by using $\langle r|H|El\rangle$ with the Hamiltonian (6.34) we find

$$\left[\frac{\hbar^2}{2\mu}\left(-\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{l(l+1)}{r^2}\right) + V(r)\right]\ R_{El} = E\ R_{El}. \tag{6.43}$$

This simplifies further by multiplying through by r and redefining some quantities as follows

$$\begin{array}{l}
\left(-\frac{\partial^2}{\partial r^2} + \frac{l(l+1)}{r^2} + v(r) - \varepsilon\right) f_{El}(r) = 0\\
f_{El}(r) = rR_{El}(r), \quad v(r) = \frac{2mV(r)}{\hbar^2}, \quad \varepsilon = \frac{2mE}{\hbar^2}
\end{array} \tag{6.44}$$

To compute the eigenvalue E we will return to this "radial equation" in a later section when we discuss specific potentials $v\,(r)$.

6.4 General properties of angular momentum

We have seen in the previous section that the infinitesimal generator of rotations is angular momentum and it is conserved if the Hamiltonian is invariant under rotations. The expression for the conserved total angular momentum is different for different systems. For example it is $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ for the single free or interacting particle, $\mathbf{L} = \sum_{i=1}^{5}\mathbf{r}_i\times\mathbf{p}_i$ for a system of 5 free or interacting particles, and even more complicated for a system of spinning particles. We want to consider the general properties of angular momentum for all cases. For the general case we will denote angular momentum with the symbol $\mathbf{J}$.

There is one common property in all cases, and that is $\mathbf{J}$ is *defined* as the generator of infinitesimal rotations. This implies that the rotation of any operator A by an angle $\boldsymbol{\omega}$ is given, in any physical system, by $A' = exp(i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar)\ A\ exp(-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar)$, or infinitesimally

$$\delta_\omega A = \frac{i}{\hbar}[\boldsymbol{\omega}\cdot\mathbf{J}, A]\ . \tag{6.45}$$

There are immediate general consequences of this statement which apply universally to any system:

(i) When $\mathbf{J}$ is commuted with any scalar, the result must be zero,

$$[\mathbf{J}, scalar] = 0. \tag{6.46}$$

This is because a scalar does not change under rotations, $\delta_\omega(scalar) = 0$. In particular, in a rotationaly symmetric system the Hamiltonian is a scalar, and $[\mathbf{J}, H] = 0$. Also, the dot product of any two vectors is a scalar, and $\mathbf{J}$ must commute with it. In particular the square of the angular momentum vector, $\mathbf{J}\cdot\mathbf{J}$ is a scalar and one must have

$$[\mathbf{J}_I, \mathbf{J}^2] = 0. \tag{6.47}$$

(ii) When $\mathbf{J}$ is commuted with any vector $\mathbf{V}$, it must generate the infinitesimal rotation appropriate for a vector,

$$\delta_\omega V^J = \frac{i}{\hbar}[\boldsymbol{\omega}\cdot\mathbf{J}, V^J] = (\boldsymbol{\omega}\times\mathbf{V})^J.$$

This result can emerge only if, in any theory, the commutation of angular momentum with the vector is

$$[J^I, V^J] = i\hbar\ \epsilon^{IJK}\ V^K\ . \tag{6.48}$$

Candidates of vectors that satisfy this commutation rule are the positions $\mathbf{r}_i$, the momenta $\mathbf{p}_i$ and the spins $\mathbf{S}_i$ of the particles i in any theory, as well as any other vector $\mathbf{V}(\mathbf{r}_i, \mathbf{p}_i, \mathbf{S}_i)$ constructed from them. Since the angular momentum operator itself is a vector, a corollary of (6.48) is

$$[J_I, J_J] = i\hbar\ \epsilon_{IJK}\ J_K\ . \tag{6.49}$$

From the commutation rules in (6.48) one can immediately verify explicitly that $\mathbf{J}$ commutes with the dot product of any two vectors, and thus prove statements made in item (i).

From the general property of rotations, we have derived on very general grounds the commutation rules of angular momentum (6.49). This approach makes it clear that angular momentum is the generator of infinitesimal rotations. The reason for the non-zero commutator is the fact that rotations in different directions do not commute with each other. One can go back to specific examples and use the basic canonical commutators to verify that (6.49) is indeed true, as we did for the central force problem. It is evident that there are a multitude of operator representations of angular momentum; namely, one for every multiparticle system that can be imagined. Similarly there are a multitude of matrix

representations of these commutation rules. These will be studied systematically in the next chapter.

It is useful to note one matrix representation that follows from the universal formula for rotating a vector infinitesimally $\delta_\omega V^J = (\boldsymbol{\omega} \times \mathbf{V})^J$. By writing the vector as a column this formula may be rewritten in matrix notation

$$\begin{pmatrix} \delta_\omega V_1 \\ \delta_\omega V_2 \\ \delta_\omega V_2 \end{pmatrix} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}. \tag{6.50}$$

Thus, under an infinitesimal rotation we have the new vector $V' = V + \delta_\omega V = (1 + A(\omega))V$, where $A(\omega)$ is the antisymmetric matrix containing the ω_I and V is the column vector representing the vector. To apply a finite rotation with non-infinitesimal $\boldsymbol{\omega}$, we can divide it into N equal parts where N is large, apply N infinitesimal rotations, and take the limit for $N \to \infty$. This gives the exponential of the matrix A

$$V' = \lim_{N\to\infty} \left(1 + \frac{A}{N}\right)^N V = e^A V \tag{6.51}$$

Therefore the matrix $R = \exp(A(\omega))$ represents a general rotation of a vector. A close examination of this matrix reveals that, due to the antisymmetry of A, it is the most general orthogonal matrix of determinant $+1$

$$R = e^A \quad \Rightarrow \quad R^T = R^{-1}, \quad \det R = 1. \tag{6.52}$$

Such a matrix is a member of the group $SO(3)$ formed by special orthogonal matrices (special means that the determinant is $+1$). Therefore rotations have the group property of $SO(3)$. We will return to a discussion of group theoretical properties in the chapter on symmetries. Now let us examine the infinitesimal rotation represented by the matrix A

$$A = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \tag{6.53}$$

This may be written as $A = \omega_1 A_1 + \omega_2 A_2 + \omega_3 A_3$, where the matrix elements of A_I are given by $(A_I)_{JK} = \epsilon_{IJK}$

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; A_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{6.54}$$

This set of matrices may be viewed as generators of infinitesimal transformations that implement the infinitesimal rotation of a vector, that is,

$$r'_I = R_{IJ} r_J = \left(e^A\right)_{IJ} \; r_J \;\text{ or }\; \mathbf{r}' = \exp\left(\tilde{\boldsymbol{\omega}}\times\right) \; \mathbf{r}$$

since infinitesimal for infinitesimal $\vec{\omega}$ we have $\delta_\omega r_I = A_{IJ} r_J = (\boldsymbol{\omega} \times \mathbf{r})_I$. Therefore, the matrices A_I are intimately connected to a representation of angular

momentum . Indeed, by comparing $R = \exp(\tilde{\boldsymbol{\omega}} \cdot \vec{A}) = \exp(i\tilde{\boldsymbol{\omega}} \cdot \tilde{\mathbf{J}}/\hbar)$ one can identify a matrix representation of the rotation generators

$$\mathbf{J} = -i\hbar\mathbf{A} \ . \tag{6.55}$$

By explicit matrix commutation, $[A_1, A_2] = -A_3$ and cyclic $(1, 2, 3)$, one can check that (6.49) is satisfied:

$$[(-i\hbar\mathbf{A})_I, (-i\hbar\mathbf{A})_J] = i\hbar \ \epsilon_{IJK} \ (-i\hbar\mathbf{A})_K \ .$$

This emphasizes once again that the commutation rules (6.49) are a property of the rotation group, and are satisfied independently of the details of any quantum theory.

6.5 Hilbert space for angular momentum

What are the properties of rotations that can be observed simultaneously? Since rotations in three dimensions can be specified with two angles (θ, ϕ), one must label the Hilbert space with two rotation quantum numbers. One possible basis is "position space", which is angle space $|\theta, \phi\rangle$. Another possible basis is angular momentum space. Therefore, angular momentum space must have two labels. There is a parallel between angle-angular momentum spaces, and the ordinary position-momentum spaces, respectively. It is useful to keep this analogy in mind. Similar to ordinary Fourier expansions, one can expand the angle basis in terms of the angular momentum basis, and vice versa. Instead of the set of functions $\exp(i\mathbf{p} \cdot \mathbf{x}/\hbar) \sim \langle \mathbf{r}|\mathbf{p}\rangle$ that form a complete basis for a Fourier expansion, one uses the spherical harmonics $\langle \mathbf{\Omega}|lm\rangle \sim Y_{lm}(\mathbf{\Omega})$ as a complete basis for an expansion of angle space into angular momentum space, and vice-versa, as will be seen below.

The two labels for angular momentum correspond to commuting operators constructed from functions of the generators $\mathbf{J}$. We have seen that rotation invariance of scalars guarantee that $[\mathbf{J}, \mathbf{J}^2] = 0$. Therefore the two simplest commuting operators can be chosen as $(\mathbf{J}^2, \ J_3)$, and their eigenvalues may label the states as $|\lambda, m\rangle$

$$J_3|\lambda, m\rangle = \hbar m|\lambda, m\rangle \ , \qquad \mathbf{J}^2|\lambda, m\rangle = \hbar^2\lambda^2|\lambda, m\rangle \tag{6.56}$$

Since $\mathbf{J}^2$ is a positive operator, its eigenvalue must be positive $\lambda^2 \geq 0$. Furthermore, both λ^2 and m must be real since they are eigenvalues of Hermitian operators. The action of all the generators $\mathbf{J}$ on these states will provide all possible representations of the Lie algebra (6.49) .

We now need to find out how J_1 and J_2 operate on the state vectors $|\lambda, m\rangle$. For this purpose define $J_\pm = J_1 \pm iJ_2$, which are called raising/lowering operators. It is also conventional to rename $J_3 = J_0$. Their commutation rules follows from (6.49)

$$[J_0, J_\pm] = \pm\hbar J_\pm \ , \qquad [J_+, J_-] = 2\hbar J_0 \ . \tag{6.57}$$

We are looking for $J_\pm|\lambda,m\rangle = |?\rangle$. Let us act on $|?\rangle$ with J_0 and $\mathbf{J}^2$

$$\begin{aligned}\mathbf{J}^2(J_\pm|\lambda,m\rangle) &= J_\pm\mathbf{J}^2|\lambda,m\rangle = \hbar^2\lambda^2(J_\pm|\lambda,m\rangle)\\ J_0(J_\pm|\lambda,m\rangle) &= (J_\pm J_0 \pm \hbar J_\pm)|\lambda,m\rangle = \hbar(m\pm1)(J_\pm|\lambda,m\rangle)\end{aligned} \tag{6.58}$$

This shows that $|?\rangle$ is an eigenstate of $(\mathbf{J}^2,\ J_0)$ with eigenvalues $\hbar^2\lambda^2$ and $\hbar(m\pm1)$ respectively. This justifies writing

$$J_\pm|\lambda,m\rangle = \beta_\pm|\lambda,m\pm1\rangle\ , \tag{6.59}$$

where the proportionality constants β_+ and β_- are obtained through normalization and hermiticity. For $J_\pm|\lambda,m\rangle$ one has $(J_\pm|\lambda,m\rangle)^\dagger = \langle\lambda,m|J_\mp$, so that

$$|J_\pm|\lambda,m\rangle|^2 = \langle\lambda,m|J_\mp J_\pm|\lambda,m\rangle = |\beta_\pm|^2\langle\lambda,m\pm1|\lambda,m\pm1\rangle = |\beta_\pm|^2\ . \tag{6.60}$$

Noticing that $[J_+,J_-] = 2\hbar J_0$, and $\mathbf{J}^2 = J_\mp J_\pm \pm \hbar J_0 + J_0^2$, we obtain

$$\begin{aligned}&|\beta_\pm|^2 = \langle\lambda,m|(\mathbf{J}^2 \mp \hbar J_0 - J_0^2)|\lambda,m\rangle = \hbar^2[\lambda^2 - m(m\pm1)],\\ &|\beta_+(\lambda,m)| = \hbar\sqrt{\lambda^2 - m(m+1)},\quad |\beta_-(\lambda,m)| = \hbar\sqrt{\lambda^2 - m(m-1)}\end{aligned} \tag{6.61}$$

Since the left hand side is positive and real, we conclude from the right hand side that

$$\lambda^2 \geq m\,(m\pm1)\,. \tag{6.62}$$

To determine the allowed values of (λ,m), consider some starting value $m = m_0$, and apply the rasing/lowering operators n times, $(J_\pm)^n$, to get to a state with $m = m_0 \pm n$. Because of the bound (6.62), there must be a maximum and a minimum value of m for a fixed value of λ,

$$m_{\min}\,(\lambda) \leq m \leq m_{\max}\,(\lambda) \tag{6.63}$$

Applying $J_\pm$ to the maximal/minimal state must produce $\beta_+ = 0$ or $\beta_- = 0$, since otherwise one reaches a state labelled by an eigenvalue outside of the allowed set of m's. Let the maximum value $m_{\max}\,(\lambda)$ be denoted by j. Then from $|\beta_+(\lambda,m_{\max})| = 0$ in (6.59,6.61) we see that

$$\lambda^2 = m_{\max}(m_{\max}+1) = j(j+1). \tag{6.64}$$

Similarly, from $|\beta_-(\lambda,m_{\min})| = 0$ we conclude that

$$m_{\min}(m_{\min}-1) = \lambda^2 = j(j+1).$$

where the result of (6.64) is used. The solution of this equation is $m_{\min} = -j$ or $j+1$. Since $m_{\max} \rangle m_{\min}$, the unique solution is $j \geq 0$ and therefore

$$\begin{aligned}m_{\max} = j,\quad m_{\min} = -j\\ \lambda^2 = j(j+1).\end{aligned}$$

From now on we label the states as $|j,m\rangle$ instead of $|\lambda,m\rangle$. We have obtained $J_+|jm\rangle$ and $J_-|jm\rangle$ up to a phase factor, which can always be chosen, by adjusting the relative phases of states, to be 1. Therefore,

$$J_\pm|jm\rangle = \hbar\sqrt{j(j+1)-m(m\pm1)}\;|j,m\pm1\rangle\;, \tag{6.65}$$

where m must lie in the range $-j \le m \le j$. Furthermore, one can start with the maximal state, and by applying J_- an integer number of times, reach the minimal state. Hence, $(m_{max}-m_{min}) = 2j$ must be an integer. Thus, one must conclude that j, and hence also m, is a half integer or an integer.

In summary, the allowed values are

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \cdots \qquad -j \le m \le j\;. \tag{6.66}$$

The complete orthonormal set of states are specified by these eigenvalues

$$\sum_j \sum_{m=-j}^{j} |jm\rangle\langle jm| = 1, \qquad \langle jm|j'm'\rangle = \delta_{jj'}\delta_{mm'}. \tag{6.67}$$

A significant physical result that emerges from this general analysis is the list of allowed eigenvalues of angular momentum given in (6.66). A physical quantum state cannot exist for other values of angular momentum. In particular, the smallest non-zero quantum of angular momentum is $j = 1/2$. Consequently, for a spinning particle the smallest intrinsic spin is $s = j = 1/2$.

According to the Standard Model of elementary particles the smallest building blocks of all matter fall into two classes: the fundamental fermions, quarks and leptons, that have spin 1/2, and the fundamental gauge bosons, gluons, photon, $W^\pm, Z^0$, all of which have spin 1. There may also be a Higgs particle of spin 0.

In specific quantum mechanical systems only certain values of j listed in (6.66) may be realized, or the values of j need not extend all the way to infinity. This will be determined case by case by the specific system. In particular, we will see in the next section that *orbital* angular momentum can have only *integer* eigenvalues $j = 0, 1, 2, \cdots$. Also, we have already seen in Chapter 5 other constructions of $\mathbf{J}$ in the case of the 2-dimensional harmonic oscillator[5]. For the bosonic construction in eq.(5.63) j takes all the values in (6.66) only once. For the fermionic construction in eq.(5.81) j takes only the values $j = 0, 1/2$, with $j = 0$ repeated twice. This illustrates that the specific system dictates which of the allowed values occur, and how many times.

6.6 Spherical harmonics

We have defined orbital angular momentum as $\mathbf{L} = \mathbf{r}\times\mathbf{p} = \mathbf{J}$. We already know that the possible eigenvalues are in the set listed in (6.66), but we must

[5] Of course, in the 2-dimensional harmonic oscillator the meaning of $\mathbf{J}$ is different than angular momentum in 3-dimensions studied in this section, but it satisfies the same mathematics.

still find out which of the j eigenvalues occur when $\mathbf{J}$ has the specific form of orbital angular momentum. To distinguish this case from the general problem discussed in the previous section we will label the states with the eigenvalues l, m, and search for the allowed values of l. We have seen in section (6.3) that the operator $\mathbf{L}$ acts on angle space as a differential operator

$$\langle\Omega|\mathbf{L} = -i\hbar\left(-\mathbf{\Theta}\,\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} + \mathbf{\Phi}\;\frac{\partial}{\partial\theta}\right)\langle\Omega|. \tag{6.68}$$

It is straightforward to verify that this differential operator representation of angular momentum satisfies the algebra $[L_I, L_J] = i\hbar\varepsilon_{IJK}\;L_K$. From this form we extract the formulas for applying $L_0 \equiv L_3$ and $L_\pm \equiv L_1 \pm iL_2$ on angle space

$$\begin{array}{c}\langle\Omega|L_0 = -i\hbar\frac{\partial}{\partial\phi}\langle\Omega| \\ \langle\Omega|L_\pm = \hbar e^{\pm i\phi}(\pm\frac{\partial}{\partial\theta} + i\cot\theta\;\frac{\partial}{\partial\phi})\langle\Omega|.\end{array} \tag{6.69}$$

Again one can easily verify that these differential operators satisfy $[L_+, L_-] = 2\hbar L_0$ and $[L_0, L_\pm] = \pm\hbar L_\pm$, as expected. By taking this information into account we will find that the allowed spectrum for the angular momentum quantum l is only the integers $l = 0, 1, 2, \cdots$, implying that half integer quantum numbers are excluded for *orbital* angular momentum.

Angle space $|\Omega\rangle$ as well as angular momentum space $|l, m\rangle$ are orthonormal and complete

$$\begin{array}{l}\mathbf{1} = \sum_{l,m}|l,m\rangle\langle l,m|, \quad \langle l,m|l'm'\rangle = \delta_{ll'}\delta_{mm'} \\ \mathbf{1} = \int d\Omega\;|\Omega\rangle\langle\Omega|, \quad \langle\Omega|\Omega^{'}\rangle = \delta(\Omega - \Omega').\end{array} \tag{6.70}$$

The volume element in angle space is the solid angle element

$$d\Omega = d\phi\;d(\cos\theta) = \sin\theta\;d\phi\,d\theta \tag{6.71}$$

which is inherited from the volume element in three dimensions $d^3\mathbf{r} = r^2dr\,d\Omega$. Similarly, the delta function in angle space is inherited from the delta function in 3-dimensions $\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{r^2}\delta(r - r')\;\delta(\Omega - \Omega')$, where

$$\delta(\Omega - \Omega') = \delta(\phi - \phi^{'})\;\delta(\cos\theta - \cos\theta^{'}) = \frac{1}{\sin\theta}\delta(\phi - \phi^{'})\;\delta(\theta - \theta^{'}). \tag{6.72}$$

This delta function satisfies the usual condition $\int d\Omega\;\delta(\Omega - \Omega') = 1$, which is consistent with Eq.(6.70).

Now, we can expand one complete set of states in terms of the other set

$$|l, m\rangle = \int d\Omega\;|\Omega\rangle\langle\Omega|l, m\rangle. \tag{6.73}$$

The expansion coefficients $Y_{lm}(\Omega) = \langle\Omega|l, m\rangle$ are interpreted as the probability amplitudes for finding the particle pointing in the direction $\Omega = (\theta, \phi)$ when it is in the angular momentum eigenstate $|l, m\rangle$. By sandwiching the operator $\mathbf{L}$ between the ket and the bra $\langle\Omega|\mathbf{L}|lm\rangle$, and using (6.65) and (6.69) to evaluate it

in two ways, we derive the first order differential equation constraints satisfied by Y_{lm}

$$\begin{gathered} -i\tfrac{\partial}{\partial\phi}\, Y_{lm}(\theta,\phi) = m\, Y_{lm}(\theta,\phi) \\ e^{\pm i\phi}(\pm\tfrac{\partial}{\partial\theta} + i\cot\theta\ \tfrac{\partial}{\partial\phi})\ Y_{lm}(\theta,\phi) = \sqrt{l(l+1)-m(m\pm1)}\ Y_{l,m\pm1}(\theta,\phi). \end{gathered} \tag{6.74}$$

If we sandwich the operator $\mathbf{L}^2$ we derive a second order differential equation

$$\left(-\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\,\frac{\partial}{\partial\theta} - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) Y_{lm}(\theta,\phi) = l(l+1)\ Y_{l,m}(\theta,\phi). \tag{6.75}$$

This equation may be recognized as the eigenvalue equation for the Laplacian $\boldsymbol{\nabla}^2$ restricted on a sphere of unit radius. The first order differential equations (6.74) are the first integrals of the second order differential equation (6.75), with correct boundary conditions. Therefore, the solution of the first order equations (6.74) will automatically satisfy the second order one (6.75).

It is straightforward to solve the first equation in (6.74)

$$Y_{lm}(\theta,\phi) = e^{im\phi}\ y_{lm}(\theta) \tag{6.76}$$

where $y_{lm}(\theta)$ is unknown. The second equation in (6.74) then gives a recursion relation

$$\sqrt{l(l+1)-m(m\pm1)}\ y_{l,m\pm1}(\theta) = \left(\pm\frac{\partial}{\partial\theta} - m\cot\theta\right) y_{lm}(\theta) \tag{6.77}$$

Specializing to $m=\pm l$, we have

$$\left(\frac{\partial}{\partial\theta} - l\cot\theta\right) y_{l,\pm l}(\theta) = 0. \tag{6.78}$$

The solution is

$$y_{l,\pm l}(\theta) = \frac{1}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}\ (\sin\theta)^l \tag{6.79}$$

Where the correct normalization factor is included up to a sign so that $\int d\Omega\ |Y_{l,\pm l}|^2 = 1$, as follows

$$\begin{aligned} 1 &= \langle l,\pm l|l,\pm l\rangle = \int d\Omega\langle l,\pm l|\Omega\rangle\langle\Omega|l,\pm l\rangle = \int d\Omega|Y_{l,\pm l}|^2 \\ &= \left(\frac{1}{2^l l!}\right)^2\frac{(2l+1)!}{4\pi}\int_0^{2\pi}\int_0^{\pi} d\phi d\theta(\sin\theta)^{2l+1} \\ &= \left(\frac{1}{2^l l!}\right)^2\frac{(2l+1)!}{4\pi}2\pi\int_{-1}^{1}\left(1-x^2\right)^l dx = 1 \end{aligned}$$

where we have used the change of variables $x=\cos\theta$ to do the last integral.

We can now see that only integer values of l are consistent with the recursion relation. To prove that l cannot be half integer we check the consistency of the recursion formula for $l=\frac{1}{2}$ with this result. So, on the one

hand we must require from (6.79) that $y_{1/2,\pm 1/2}(\theta) = C\,(\sin\theta)^{1/2}$ but on the other hand, the recursion relation (6.74) for $l = \frac{1}{2}, \quad m = \frac{1}{2}$ demands $y_{1/2,-1/2}(\theta) = \left(-\frac{\partial}{\partial\theta} - \frac{1}{2}\cot\theta\right) y_{\frac{1}{2},\frac{1}{2}}(\theta)$. One finds that the θ dependence of the two sides of the equation are inconsistent with each other, thus excluding $l = 1/2$ as a solution of these equations. A similar inconsistency arises for all values of l that are of the form integer plus $1/2$, but there is no inconsistency for $l = integer$.

To obtain y_{lm} through the recursion relation it is useful to rewrite (6.77) in the form

$$y_{l,m+1}(\theta) = \gamma_{lm}\,(\sin\theta)^{m+1}\,\left(-\frac{\partial}{\partial\cos\theta}\right)\,(\sin\theta)^{-m} y_{lm} \tag{6.80}$$

where $\gamma_{lm} = [(l-m)(l+m+1)]^{-1/2}$, and apply it repeatedly $l+m$ times by starting with $m = -l$. This gives

$$\begin{array}{rl} y_{lm}(\theta) & = \left[\prod_{k=-l}^{m-1}\gamma_{lk}\left((\sin\theta)^{k+1}\,\left(-\frac{\partial}{\partial\cos\theta}\right)\,(\sin\theta)^{-k}\right)\right]\; y_{l,-l}(\theta) \\ & = \frac{(-1)^m}{2^l\,l!}\left(\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}\right)^{1/2}(\sin\theta)^m\left(\frac{\partial}{\partial(\cos\theta)}\right)^{l+m}(\cos^2\theta-1)^l \end{array} \tag{6.81}$$

This expression is recognized as the associated Legendre polynomial $P_l^m(\cos\theta)$ except for an overall normalization factor. Therefore, we can write the complete solution in terms of known functions

$$Y_{lm}(\theta,\phi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\theta)\,e^{im\phi}. \tag{6.82}$$

This set of functions are the *spherical harmonics*. Using well known properties of associated Legendre polynomials such as

$$P_l^{-m}(\theta) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(\theta), \tag{6.83}$$

or by direct examination of our formulas, we can derive relations among the spherical harmonics

$$Y_{l,-m} = (-1)^m Y_{lm}^*. \tag{6.84}$$

The completeness relations in (6.70) may be sandwiched between angle states $\langle\Omega|\cdots|\Omega'\rangle$ or angular momentum states $\langle lm|\cdots|l'm'\rangle$ to derive completeness and orthonormality relations for the spherical harmonics

$$\begin{array}{c} \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\Omega)Y_{lm}^*(\Omega') = \delta(\Omega-\Omega') \\ \int d\Omega\; Y_{l'm'}^*(\Omega)\; Y_{lm}(\Omega) = \delta_{ll'}\delta_{mm'}\;. \end{array} \tag{6.85}$$

These non-trivial equations that are guaranteed by the consistency of the quantum mechanical setup can be laboriously verified by using the properties of associated Legendre polynomials (see problem).

We list a few Y_{lm} explicitly and analyze the probability distribution in angle space

$$\begin{aligned} Y_{00} &= \tfrac{1}{\sqrt{4\pi}} & Y_{20} &= \sqrt{\tfrac{5}{4\pi}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right)} \\ Y_{10} &= \sqrt{\tfrac{3}{8\pi}}\cos\theta & Y_{2,\pm 1} &= \mp\sqrt{\tfrac{15}{8\pi}}\sin\theta\,\cos\theta\; e^{\pm i\phi} \\ Y_{1,\pm 1} &= \mp\sqrt{\tfrac{3}{8\pi}}\sin\theta\; e^{\pm i\phi} & Y_{2,\pm 2} &= \sqrt{\tfrac{15}{2\pi}}\sin^2\theta\;\; e^{\pm 2i\phi} \end{aligned} \tag{6.86}$$

A polar plot is useful to get a feeling of the probability distribution. Thus, along the direction $\mathbf{\Omega}$ draw a vector whose length is the probability amplitude $|Y_{lm}(\Omega)|$. The longer the vector the more probable it is to find the particle at those angles. A polar plot of $Y_{lm}(\theta, 0)$ at $\phi = 0$ is given in Fig.(6.2).

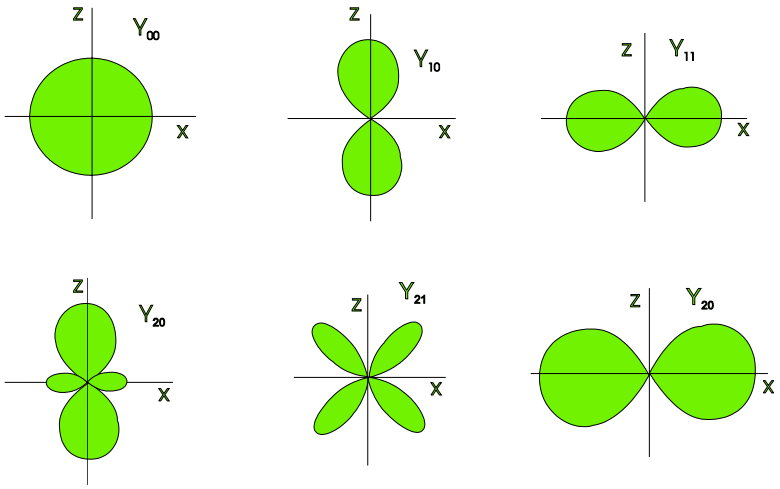


Fig.(6.2) - Polar plot of spherical harmonics.

This plot is the cross section of the probability amplitude in the (x, z) plane, such that the numerical value of $|Y_{lm}(\Omega)|$ at any angle in the (z, x) plane (with θ measured from the z axis) corresponds to the *boundary* of the figures at the given angle. To obtain the full plot one has to rotate the picture around the z-axis since the probability $|Y_{lm}(\Omega)|$ is independent of the angle ϕ. Through these pictures one develops a physical intuition of where the particle is located in angle space when it is in an angular momentum eigenstate $|lm\rangle$. We can verify that, except for the fuzzyness imposed by quantum mechanics, the particle is located roughly where we expect it to be on the basis of its classical motion. For example, for zero angular momentum, it must be equally distributed in any direction, as is the case with Y_{00}. For angular momentum $|l = 1, m = 1\rangle$ we may think of the particle as rotating around a loop close to the equator so that its angular momentum would point toward the north pole. This is indeed what we see in the picture since the probability amplitude Y_{11} is largest near the equator. Similarly, for $|l = 1, m = 0\rangle$, the classical motion of the particle is rotation in the vicinity of a meridian so that the z-component of its angular momentum

is zero. Again, this is what is indicated by the plot of Y_{10}. Of course, the full details provided by the exact quantum mechanical functions cannot be guessed on the basis of classical considerations.

6.6.1 Tensors and spherical harmonics

It is also useful to understand that the Y_{lm}'s represent tensors of rank l, constructed from the unit vector $\mathbf{\Omega}$. For example, neglecting the overall normalization one can verify that for $l = 1$ the Y_{1m} are rewritten as

$$Y_{10} \sim \cos\theta = \frac{z}{r} = \mathbf{\Omega}_3 \equiv \mathbf{\Omega}_0, \tag{6.87}$$

$$Y_{1,\pm 1} \sim \frac{1}{\sqrt{2}} \sin\theta \; e^{\pm i\theta} = \frac{x \pm iy}{r\sqrt{2}} = \frac{\mathbf{\Omega}_1 \pm i\mathbf{\Omega}_2}{\sqrt{2}} \equiv \mathbf{\Omega}_\pm \tag{6.88}$$

Thus, one may think of Y_{1m} as the 3 components of the vector $\mathbf{\Omega}_I$ taken in the basis $I = (+, 0, -)$ instead of the conventional cartesian basis labelled by 1,2,3. The relation between the bases $(1, 2, 3)$ and $(+, 0, -)$ is given by the unitary transformation U that satisfies $UU^\dagger = 1$ as follows

$$\begin{pmatrix} \mathbf{\Omega}_+ \\ \mathbf{\Omega}_0 \\ \mathbf{\Omega}_- \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{\Omega}_1 \\ \mathbf{\Omega}_2 \\ \mathbf{\Omega}_3 \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{\Omega}_1 \\ \mathbf{\Omega}_2 \\ \mathbf{\Omega}_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{\Omega}_+ \\ \mathbf{\Omega}_0 \\ \mathbf{\Omega}_- \end{pmatrix}$$

The dot product between any two vectors $\vec{V} \cdot \vec{W}$ in 3 dimension can be written either in the $(1, 2, 3)$ basis or in the $(+, 0, -)$ basis. In the first case the metric for the dot product is δ_{IJ} while in the latter case it is g_{IJ} given by transforming the cartesian metric δ_{IJ} to the $(+, 0, -)$ basis

$$g_{IJ} = \left(U 1 U^T\right)_{IJ} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Thus we can verify that

$$\mathbf{\Omega} \cdot \mathbf{\Omega} = \mathbf{\Omega}_1\mathbf{\Omega}_1 + \mathbf{\Omega}_2\mathbf{\Omega}_2 + \mathbf{\Omega}_3\mathbf{\Omega}_3 = \mathbf{\Omega}_+\mathbf{\Omega}_- + \mathbf{\Omega}_0\mathbf{\Omega}_0 + \mathbf{\Omega}_-\mathbf{\Omega}_+ = 1$$

Therefore, if we wish to work in the $(1, 2, 3)$ basis we use as the metric δ_{IJ}, and if we wish to work in $(+, 0, -)$ basis we use the metric g^{IJ} as in $\mathbf{\Omega} \cdot \mathbf{\Omega} = \mathbf{\Omega}_I\mathbf{\Omega}_J g^{IJ}$ with

$$g_{IJ} = g^{IJ} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \text{ where } g_{IJ} \text{ is the inverse of } g^{IJ}. \tag{6.89}$$

Following this notation let us consider *traceless and completely symmetric* tensors of second rank, third rank, fourth rank, and so on, constructed from the unit vector $\mathbf{\Omega}$. For the sake of generality we write them in a form valid in d-dimensions, and we can use $g_{IJ}{=}\delta_{IJ}$ if we wish to work directly in the cartesian basis in any dimension

$$\begin{aligned}
T_{IJ} &= \mathbf{\Omega}_I\mathbf{\Omega}_J - \tfrac{\tilde{\mathbf{\Omega}}^2}{d} g_{IJ} \\
T_{IJK} &= \mathbf{\Omega}_I\mathbf{\Omega}_J\mathbf{\Omega}_K - \tfrac{\tilde{\mathbf{\Omega}}^2}{d+2}\left(g_{IJ}\mathbf{\Omega}_K + g_{KI}\mathbf{\Omega}_J + g_{JK}\mathbf{\Omega}_I\right) \\
T_{IJKL} &= \left[\begin{array}{l} \mathbf{\Omega}_I\mathbf{\Omega}_J\mathbf{\Omega}_K\mathbf{\Omega}_L - \tfrac{\tilde{\mathbf{\Omega}}^2}{d+4}\left(\begin{array}{c} g_{IJ}\mathbf{\Omega}_K\mathbf{\Omega}_L + g_{IK}\mathbf{\Omega}_L\mathbf{\Omega}_J + g_{IL}\mathbf{\Omega}_J\mathbf{\Omega}_K \\ +g_{JK}\mathbf{\Omega}_L\mathbf{\Omega}_I + g_{JL}\mathbf{\Omega}_I\mathbf{\Omega}_K + g_{KL}\mathbf{\Omega}_I\mathbf{\Omega}_J \end{array}\right) \\ +\tfrac{d\ \tilde{\mathbf{\Omega}}^2\tilde{\mathbf{\Omega}}^2}{(d+2)(d+4)}\left(g_{IJ}g_{KL} + g_{IK}g_{LJ} + g_{IL}g_{JK}\right) \end{array}\right]
\end{aligned} \tag{6.90}$$

Since $\tilde{\mathbf{\Omega}}^2 = 1$ the second or third terms in these equations may be simplified, but the traceless tensor is defined even if $\tilde{\mathbf{\Omega}}^2$ is arbitrary. In fact, we will later use tensors constructed from other vectors, therefore we may as well consider $\mathbf{\Omega}$ as a general vector in these definitions. Then the tensor of rank l is a homogeneous polynomial of degree l in powers of the vector $\mathbf{\Omega}$.

These tensors are traceless in the sense $g^{IJ}T_{IJ} = g^{IJ}T_{IJK} = g^{IJ}T_{IJKL} = 0$, as can be verified by using $g^{IJ}g_{IJ} = d$. Of course, contraction of the metric with *any pair* of indices in $T_{IJK\ldots}$ also gives zero because of the permutation symmetry of the tensor indices.

By comparing the 5 independent components of the second rank tensor in $d = 3$ dimensions to the 5 spherical harmonics Y_{2m}, we see the correspondence up to normalization factors

$$Y_{2,\pm 2} \sim T_{\pm\pm} = \frac{\sin^2\theta}{2}\, e^{\pm 2i\theta}, \quad Y_{2,\pm 1} \sim T_{\pm 0} = \cos\theta\frac{\sin\theta}{\sqrt{2}}\, e^{\pm i\theta}, \tag{6.91}$$

$$Y_{20} \sim T_{00} = -2T_{+-} = \cos^2\theta - \frac{1}{3} \tag{6.92}$$

where $-$ indices are replaced by $+$ indices under complex conjugation $(T_{--})^* = T_{++}$ etc.. Similarly for $l = 3$ one finds $Y_{33} \sim T_{+++}$, etc. (see problem 7). This observation can be carried on to tensors of higher rank, so indeed spherical harmonics Y_{lm} are just the independent components of symmetric traceless tensors of rank l.

6.7 Radial & angular equations in d-dims.

It is possible to generalize all the results to d-dimensions by following the general operator approach. In $d-$dimensions the generator of rotations in the (I, J) plane is written as

$$L_{IJ} = r_I p_J - r_J p_I, \quad I, J = 1, 2, \cdots d. \tag{6.93}$$

The L_{IJ} commute with all dot products constructed from $(\mathbf{r}, \mathbf{p})$. The commutators of these operators close into the same set (see problem)

$$[L_{IJ}, L_{KL}] = \delta_{JK}L_{IL} - \delta_{IK}L_{JL} - \delta_{IL}L_{JK} + \delta_{JL}L_{IK}. \tag{6.94}$$

This set of commutation rules is the Lie algebra for $SO(d)$. Note that the 3-dimensional case is a special case that permits the rewriting in terms of the familiar $L_{1,2,3}$ as $L_{IJ} = \epsilon_{IJK} L_K$. The quadratic Casimir operator is defined by

$$\begin{aligned} L^2 &= \tfrac{1}{2}\textstyle\sum_{I,J=1}^{d} L_{IJ}^2 \\ &= \mathbf{r}^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})(\mathbf{p}\cdot\mathbf{r}) - 2i\hbar\mathbf{r}\cdot\mathbf{p}. \end{aligned} \tag{6.95}$$

Since it is constructed from dot products it must commute with all L_{IJ}; this result may also be verified abstractly by only using the commutation rules of the $SO(d)$ Lie algebra. The combination $(\mathbf{r}\cdot\mathbf{p})(\mathbf{p}\cdot\mathbf{r}) + 2i\hbar\mathbf{r}\cdot\mathbf{p}$ may be rewritten in terms of the Hermitian operator $p_r = \sum_I \left(\frac{1}{r} r_I p_I + p_I r_I \frac{1}{r}\right)$ and then solve for $\mathbf{p}^2$ in the form (see problem)

$$\mathbf{p}^2 = p_r^2 + \frac{1}{r^2}\left[L^2 + \frac{\hbar^2}{4}(d-1)(d-3)\right]. \tag{6.96}$$

Furthermore, using $\mathbf{r}\cdot\mathbf{p} = -i\hbar r\,\frac{\partial}{\partial r}$ on wavefunctions $\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle$, one finds that p_r acts as

$$p_r = r^{-(d-1)/2}\,\left(-i\hbar\frac{\partial}{\partial r}\right)\,r^{(d-1)/2},\;\; p_r^2 = r^{-(d-1)/2}\,\left(-\hbar^2\frac{\partial^2}{\partial r^2}\right)\,r^{(d-1)/2} \tag{6.97}$$

This form allows one to separate the Schrödinger equation by defining

$$\psi(\mathbf{r}) = r^{-(d-1)/2}\, f_{El}(r) Y_{l\mathbf{m}}(\Omega), \tag{6.98}$$

where $Y_{l\mathbf{m}}(\Omega) = T_{I_1\cdots I_l}$ is the tensor of rank l. By rotation invariance, one may argue that this tensor is an eigenfunction of L^2 for any set of indices. The eigenvalue may be computed by taking a special set of indices; for example by applying the operator in (6.95) to $T_{++\cdots+} = (\Omega_1 + i\Omega_2)^l = (r_1/r + ir_2/r)^l$ one gets the general eigenvalue

$$L^2 Y_{l\mathbf{m}}(\Omega) = \hbar^2 l(l+d-2)\, Y_{l\mathbf{m}}(\Omega). \tag{6.99}$$

The eigenvalue $l = 0, 1, 2, \cdots$ is an integer, and there are several eigenvalues m_i that distinguish the degenerate states. The number of linearly independent angular functions $Y_{l\mathbf{m}}(\Omega)$ is given by the formula

$$D_l = \frac{(d+l-3)!}{(d-2)!\; l!}(d+2l-2) \tag{6.100}$$

which is the number of independent components of the traceless symmetric tensors of rank l in $d-$dimensions. In fact the simplest way to compute D_l is through the dimension of this tensor (see problem). The number of distinct values of $\mathbf{m}_i$, for a fixed value of l, is precisely D_l.

The radial equation now takes the form (see problem for 2 dimensions)

$$\left[-\frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\left(l(l+d-2) + \frac{1}{4}(d-1)(d-3)\right) + v(r) - \varepsilon\right] f_{El}(r) = 0, \tag{6.101}$$

where $v(r) = \frac{2mV(r)}{\hbar^2}$, $\varepsilon = \frac{2mE}{\hbar^2}$, as before. This can be put into the same form as the 3-dimensional radial equation of eq.(6.44) by defining the parameter

$$l_d = l + \frac{d-3}{2} \tag{6.102}$$

where l, d are integers. Then the radial equation is

$$\left[-\frac{\partial^2}{\partial r^2} + \frac{l_d\,(l_d+1)}{r^2} + v(r) - \varepsilon\right] f_{El}(r) = 0, \tag{6.103}$$

and obviously $l_d = l$ when $d = 3$. So, the solutions of the radial equation in various dimensions are related through analytic continuation $l \to l_d$ for the allowed values given above.

In solving the radial equation one must impose physical boundary condition at $u = 0$ and $u = \infty$. The boundary condition at $u = 0$ is

$$f_{El}(0) = 0 \tag{6.104}$$

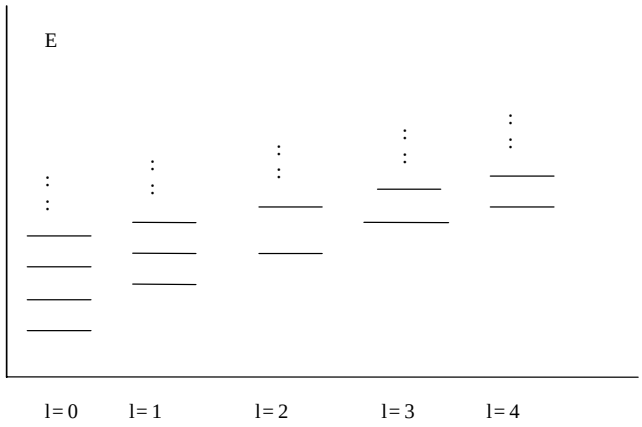
This is because the space is defined only for $r > 0$, which means we must erect an infinite wall at $r = 0$ that does not permit the particle to penetrate to negative values of r. The boundary condition at $r = \infty$ depends on whether one considers a bound state or a scattering state. For a scattering state $f_{El}\,(r)$ is an oscillating function whose normalization must be made consistent with the delta function in $d-$dimensions (see below). For a bound state $f_{El}\,(r)$ must vanish fast enough as $r \to \infty$ and must be integrable so that it can be normalized

$$\int_0^\infty dr\ |f_{El}|^2 = 1. \tag{6.105}$$

This normalization already assumes that the $Y_{l\mathbf{m}}\,(\mathbf{\Omega})$ or $T_{I_1\cdots I_l}\,(\mathbf{\Omega})$ are also normalized correctly in d-dimensions. Note also that this normalization is consistent in d-dimensions with $\int d^d\mathbf{r}\ |\psi(\mathbf{r})|^2 = 1$, where $d^d\mathbf{r} = (r^{d-1}dr)\,(\,d^{d-1}\Omega)$.

The mathematical form of the radial equation (6.103) is similar in all dimensions. In the following sections we study the radial equation for several systems in 3 dimensions by taking $l_d = l$, but it is evident that identical methods would be applied to solve the radial equation in every dimension by using the appropriate value of l_d (see problems for 2 dimensions). Exact solutions of the radial equation is possible for a few systems as discussed below. However, since most cases will not be exactly solvable it is worth noting that some "quick and dirty" estimates of lowest state energies at any fixed angular momentum l can be obtained for bound states by using the uncertainty principle as follows.

First, we note that the general structure of excited levels of bound state systems can be summarized by the following diagram



Energy levels: structure of excited states for each l.

For each fixed value of angular momentum l there is a lowest energy level. The absolute ground state normally occurs for $l = 0$, but here we will aim to get a rough estimate of the energy for the lowest state at each fixed l. According to the uncertainty principle, if the particle is bound within some radius r, then its radial momentum is at least as large as $p_r \geq \hbar/2r$. This leads to an estimate of the derivative of the radial wavefunction $-i\hbar\partial_r f(r) \sim \frac{\lambda\hbar}{r} f(r)$ with some constant λ or order 1. Inserting this into the radial equation (6.103) we get

$$\left[\frac{\lambda^2}{r^2} + \frac{l_d(l_d+1)}{r^2} + v(r) - \varepsilon\right] f_{El}(r) = 0$$

This leads to the heuristic estimate of the energy ε as a function of r at a fixed value of angular momentum

$$\varepsilon(r) = \frac{\lambda^2 + l_d(l_d+1)}{r^2} + v(r)$$

To obtain the lowest energy (i.e. corresponding to the lowest state for fixed l in the figure above) we minimize this expression with respect to r

$$\partial_r \varepsilon(r) = -2\frac{\lambda^2 + l_d(l_d+1)}{r^3} + \partial_r v(r) = 0.$$

Solving this equation we obtain the minimum at some $r = r_l$ for a fixed l, which leads to the estimate for the lowest energy

$$\varepsilon_{lowest}(l) = \frac{\lambda^2 + l_d(l_d+1)}{r_l^2} + v(r_l).$$

From the figure above we expect excited energy levels stacked up as a tower above $\varepsilon_{lowest}(l)$ for each angular momentum l. The excitations are labelled

by an additional quantum number which has not emerged yet in the present discussion. This scheme gives an idea of what to expect roughly for some general potential $v(r)$.

For example, let us consider the quark-antiquark bound states for heavy quarks described in the problems at the end of Chapter 1. As an illustration, we simplify the problem by ignoring the spins of the quarks and by taking only the linear potential $v(r) = \gamma r$ which represents the effective confining color force produced by quantum chromo-dynamics. This example does not have an exact solution (except for $l = 0$ which is solved by $f_0(r)$ related to the Airy function), but we can get a quick estimate as described above, namely

$$\begin{aligned} \partial_r \varepsilon(r) &= -2\frac{\lambda^2 + l_d(l_d+1)}{r^3} + \gamma = 0 \\ r_l &= \gamma^{-1/3}\left(2\lambda^2 + 2l_d^2 + 2l_d\right)^{1/3} \\ \varepsilon_{lowest}(l) &= 3\left(\frac{\gamma}{2}\right)^{\frac{2}{3}}\left(\lambda^2 + l_d^2 + l_d\right)^{1/3}. \end{aligned}$$

Here λ is not known, but we can only guess that it is a pure number of order 1. We see from this result the general trend that the energy of the lowest state increases as a function of $l_d = l + \frac{(d-3)}{2}$ as indicated. When this method is applied to the exactly solvable problems (e.g. harmonic oscillator, hydrogen atom) discussed in the following sections it can be verified that it gives a fairly reasonable estimate of the correct behavior. Of course, this approach is not a substitute for more reliable approximation methods, such as the variational approach that will be discussed later in the book, but it should be taken only as a first quick and easy analysis of a given problem.

6.8 Free particle

For the free particle, the Schrödinger equation in Cartesian coordinates in any dimension has an energy eigenfunction and eigenvalue of the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r}|\mathbf{p}\rangle = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi\hbar)^{d/2}}, \quad E = \frac{\hbar^2 k^2}{2m}, \quad \mathbf{p} = \hbar\mathbf{k}. \tag{6.106}$$

Therefore we already know that the energy eigenvalue is given by $\varepsilon = k^2$. Considering next the Schrödinger equation in spherical coordinates, it is convenient to define the rescaled variable $u = kr$ so that the radial equation becomes

$$\left[-\partial_u^2 + \frac{l(l+1)}{u^2}\right] f_l(u) = f_l(u). \tag{6.107}$$

The dependence on the energy has been absorbed into the rescaled u. This second order radial differential equation may be recognized as being related to the spherical Bessel function and hence is solved by consulting a book on this function. However, here we will use a method of solution that is related to

"Supersymmetric Quantum Mechanics". This is a topic we will develop later, so for now it will seem as if we are just introducing a trick. Thus, define first order differential operators with the following properties

$$\begin{array}{cc} a_l^- = -i\partial_u - \frac{i(l+1)}{u}, & a_l^+ = -i\partial_u + \frac{i(l+1)}{u} \\ a_l^- a_l^+ = -\partial_u^2 + \frac{l(l+1)}{u^2}, & a_l^+ a_l^- = -\partial_u^2 + \frac{(l+1)(l+2)}{u^2}. \end{array} \tag{6.108}$$

These operators are Hermitian conjugates of each other $a_l^- = (a_l^+)^\dagger$. The radial equation takes the form

$$a_l^- a_l^+ \ f_l = f_l. \tag{6.109}$$

If we apply a_l^+ on both sides of the equation, and notice that $a_l^+ a_l^-$ is the same operator as $a_l^- a_l^+$ except for shifting $l \to l+1$, we see that we generate a new solution of the radial equation with l shifted by one unit,

$$a_l^+ a_l^- (a_l^+ f_l) = (a_l^+ f_l) \tag{6.110}$$

Therefore, we must identify

$$f_{l+1} = (a_l^+ f_l) = \left[-i\partial_u + \frac{i(l+1)}{u} \right] f_l(u). \tag{6.111}$$

This equation may be regarded as a recursion relation that allows us to write the full solution in terms of f_0

$$f_l = a_{l-1}^+ a_{l-2}^+ \cdots a_0^+ \ f_0. \tag{6.112}$$

This expression is simplified by noting

$$a_k^+ = u^{k+2} (\frac{-i}{u} \partial_u) \ u^{-k-1}. \tag{6.113}$$

Hence the product of differential operators becomes

$$f_l(u) = u^{l+1} (\frac{-i}{u} \partial_u)^l \ \left(\frac{f_0(u)}{u} \right), \tag{6.114}$$

(what is the analog of this equation in d-dimensions?). We see that we only need to solve the second order differential equation for $f_0(u)$ with the correct boundary conditions (6.104) at $u = 0$, and correct normalization

$$\begin{array}{c} -\partial_u^2 \ f_0(u) = f_0(u) \\ f_0(u) = \sqrt{\frac{2}{\pi\hbar^3}} \sin u. \end{array} \tag{6.115}$$

The normalization is chosen for consistency with the normalization of the plane wave in eq.(6.106) in 3 dimensions (see below). Putting everything together we finally have the radial wavefunction $R_l = f_l/u$, which is recognized as the spherical Bessel function $j_l(u)$ up to a factor

$$R_l(u) = \sqrt{\frac{2}{\pi}} \ u^l (\frac{-i}{u} \partial_u)^l \ \left(\frac{\sin u}{u} \right) = i^l \sqrt{\frac{2}{\pi\hbar^3}} \ j_l(u). \tag{6.116}$$

$$uD_u\left(\frac{1}{u}\left(D_u\left(\frac{\sin u}{u}\right)\right)\right) \tag{6.117}$$

: $\frac{1}{u^3}\left(-3u\left(\cos u\right)+3\sin u-\left(\sin u\right)u^2\right)=\frac{3}{u^3}\left(u-\frac{1}{6}u^3+\frac{1}{5!}u^5\right)-\frac{1}{u}\left(u-\frac{1}{6}u^3\right)-\frac{3}{u^2}\left(1-\frac{1}{2}u^2+\frac{1}{4!}u^4\right)=\frac{1}{15}u^2+O\left(u^3\right):\ \frac{1}{15}u^2$

It is useful to list a few of these functions, and plot them, in order to get a feeling of the probability distribution they represent

$$\begin{array}{c} j_0(u)=\frac{\sin u}{u},\quad j_1(u)=\frac{\sin u}{u^2}-\frac{\cos u}{u} \\ j_2(u)=(\frac{3}{u^3}-\frac{1}{u})\sin u-\frac{3}{u^2}\cos u,\quad \cdots \end{array} \tag{6.118}$$

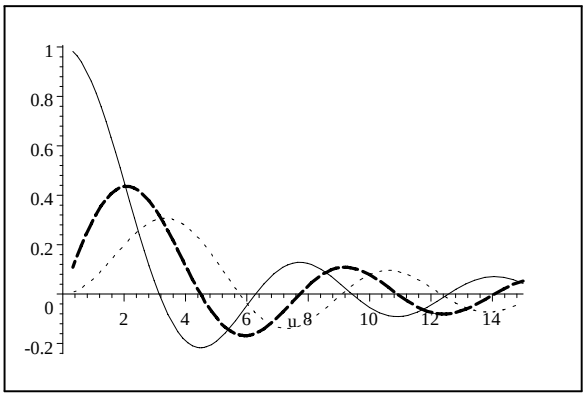


Fig.(6.3): j_0=solid line, j_1=dashed line, j_2=dotted line

These are plotted in the figure above. These vanish near the origin is $j_l(u)\rightarrow u^l$ and oscillate at infinity with a decreasing overall envelop $j_l(u)\rightarrow\frac{1}{u}\sin(u-l\pi/2)$. These are interpreted as spherical waves such that for $l=0$ the probability is largest at the origin $u=0$, and the maximum is farther and farther away for increasing values of l. This is in accordance with our intuition based on the motion of the free particle in classical mechanics. Namely, as seen from the center at a given energy, the freely moving particle has angular momentum $\left|\vec{L}\right|=|\vec{r}\times\vec{p}|=|\vec{p}|\,|\vec{r}|\sin\theta=|\vec{p}|\,b$, where $b=|\vec{r}|\sin\theta$ is the impact parameter. Taking $\left|\vec{L}\right|\sim\hbar l$ and $|\vec{p}|\sim\hbar k$ we estimate $l=kb$ where k is the wave number which determines the energy $E=\hbar^2k^2/2m$. Thus, at a fixed energy, the larger l corresponds to the larger impact parameter b. This indicates that for larger l the particle must be on a straight line trajectory farther away from the origin. The particle can hit the center only if it has zero angular momentum . Thus, the quantum mechanical description keeps the main features similar to the intuitive classical motion, but makes the picture fuzzy by spreading the probability in detailed ways that are not describable in classical mechanics.

To get an overall intuitive idea of the probability distribution in 3-dimensional space, at fixed energy and fixed angular momentum, we consider the full wave-function $|\psi_{Elm}(\mathbf{r})|=|R_{El}(r)|\,|Y_{lm}(\Omega)|$ and do a polar plot. This means that the pictures given in Fig.(6.2) will be modulated by the factor $|R_{El}(r)|$ which corresponds to Fig.(6.3) in the case of a free particle (and other curves in the

case of bound particles). The resulting polar plot looks like a cloud with higher concentration of probability where the functions $|R_{El}(r)|\,|Y_{lm}(\Omega)|$ have peaks and less concentration of probablity where they have minima or vanish. The overall picture obtained in this way is a fuzzy representation of where we expect the particle to be. On the average the result is consistent with the motion expected according to the rules of classical mechanics, but in detail the quantum wavefunction definitely differs from classical mechanics. The difference is greater the lower energies and/or lower angular momenta, as expected from the correspondance principle (quantum mechanics approaches classical mechanics at large quantum numbers). Such detail can be measured in all phenomena from the level of atoms down to the level of quarks and the results confirm the validity of quantum mechanics to great accuracy.

The wavefunction at a fixed momentum and energy in position is the plane wave $\psi_{\mathbf{p}}(\mathbf{r}) = \langle \mathbf{r}|\mathbf{p}\rangle$. It can be expanded in terms of the basis functions in angular momentum by introducing identity in the following form and using the probability distributions $j_l(r), Y_{lm}(\Omega)$ computed above, as applied to the free particle in $d = 3$ dimensions $|\mathbf{p}\rangle = |E(k), \mathbf{\Omega}_k\rangle$

$$\begin{aligned}
\langle \mathbf{r}|\mathbf{p}\rangle &= \tfrac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi\hbar)^{3/2}} \\
&= \textstyle\sum_{l=0}^{\infty}\ \sum_{m=-l}^{l} \langle \mathbf{r}|lm\rangle\langle lm|\mathbf{p}\rangle \\
&= \textstyle\sum_{l=0}^{\infty}\ \sum_{m=-l}^{l} \langle \mathbf{\Omega}_r|lm\rangle\langle lm|\mathbf{\Omega}_k\rangle\ \langle r|El\rangle \\
&= \textstyle\sum_{l=0}^{\infty}\ \sum_{m=-l}^{l} Y_{lm}(\mathbf{\Omega}_r) Y^*_{lm}(\mathbf{\Omega}_k)\ i^l \sqrt{\frac{2}{\pi\hbar^3}}\ j_l(kr) \\
&= \textstyle\sum_{l=0}^{\infty}\ P_l(\mathbf{\Omega}_r\cdot\mathbf{\Omega}_k)\,\frac{2l+1}{4\pi}\ i^l \sqrt{\frac{2}{\pi\hbar^3}}\ j_l(kr)
\end{aligned} \tag{6.119}$$

where $\cos\theta = \mathbf{\Omega}_r\cdot\mathbf{\Omega}_k = \mathbf{k}\cdot\mathbf{r}/kr$. After cancelling some factors we see that the plane wave may be expanded in terms of spherical waves as follows[6]

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{l=0}^{\infty}\ i^l\,(2l+1)\ \ P_l(\mathbf{\Omega}_r\cdot\mathbf{\Omega}_k)\ \ j_l(kr). \tag{6.120}$$

This expression is arrived at as a consistency condition of the quantum mechanical setup. It can be verified by using one of the special integral representations of spherical Bessel functions $j_l(u) = \frac{1}{2\,i^l}\int_{-1}^{1} dx\ P_l(x)\ e^{iux}$. This is one way of justifying the overall normalization chosen for $f_0(0)$ above.

[6] One way of seeing that the sum over m produces the Legendre polynomial is to choose one of the unit vectors along the z-direction. For example if the momentum points along the $z-$direction then the polar angles are both zero $\theta_k = 0$, $\phi_k = 0$, and $Y_{lm}(0,0) = y_{l0}(0)\delta_{m,0} = \delta_{m,0}\sqrt{(2l+1)/4\pi}$. Therefore only $m = 0$ survives in the sum for the special direction. Since $\Omega_r = (\theta,\phi)$ are now the angles between the two unit vectors, we finally see

$$\sum_{m=-l}^{l} Y_{lm}(\mathbf{\Omega}_r) Y^*_{lm}(\mathbf{\Omega}_k) = Y_{l0}(\mathbf{\Omega}_r)\sqrt{(2l+1)/4\pi} = P_l(\cos\theta)\frac{2l+1}{4\pi}.$$

6.9 Harmonic oscillator in 3 dimensions

From our study of the Harmonic oscillator in $d-$dimensions in the previous chapter we already know the energy eigenvalue

$$E = \hbar\omega\,(n + \frac{d}{2}) \tag{6.121}$$

where the integer $n = n_1 + n_2 + \cdots n_d$ is the sum of the excitations in the various dimensions and takes the values $n = 0, 1, 2, \cdots$. Furthermore, the normalized wavefunction is also immediately constructed in Cartesian coordinates by using the direct product form of position space $\langle \mathbf{r}| = \langle r_1|\langle r_2|\cdots\langle r_d|$ and number space $|\mathbf{n}\rangle = |n_1\rangle|n_2\rangle\cdots|n_2\rangle$

$$\psi_{n_1\cdots n_d}(\mathbf{r}) = \langle \mathbf{r}| \prod_{i=1}^{d} \frac{a_i^{\dagger n_i}}{\sqrt{n_i!}}\,|0\rangle = e^{-\mathbf{r}^2/2x_0^2} \prod_{i=1}^{d} \frac{H_{n_i}(r_i/x_0)}{\sqrt{x_0\sqrt{\pi}2^{n_i}\ n_i!}} \tag{6.122}$$

where $x_0 = (\hbar/m\omega)^{1/2}$, and we have used the result for $\langle x|n_\rangle$ obtained for the one dimensional harmonic oscillator as given in eq.(5.30).

To compute the wavefunction in spherical coordinates in 3-dimensions, we need to solve the radial equation for the potential $V(r) = \frac{1}{2}m\omega^2 r^2$. After rescaling the coordinates with x_0, the equation simplifies somewhat by writing it in terms of the variable $u = r/x_0$ and the energy parameter $\lambda = 2mEx_0^2/\hbar^2 = 2E/\hbar\omega$

$$\left(-\partial_u^2 + \frac{l(l+1)}{u^2} + u^2 - \lambda\right) f_{El}(u) = 0, \quad E = \frac{\hbar\omega\lambda}{2}. \tag{6.123}$$

On the basis of the known result (6.121) in Cartesian coordinates, we are expecting to find the quantized values $\lambda = 2n + 3$. It is useful to first find the leading behavior of the solution as $u \to 0$ and $u \to \infty$ that is consistent with physical boundary conditions. As u approaches zero the leading term in the differential equation is the angular momentum term $l(l+1)/u^2$. Neglecting the other terms one solves the equation by $f \sim u^{l+1}$. Similarly, for $u \to \infty$ the leading term is u^2, and the leading behavior consistent with the boundary conditions is $f \sim e^{-u^2/2} \times (polynomial)$. Therefore one expects that

$$f_{El} = u^{l+1} e^{-u^2/2} h_{El}(u), \tag{6.124}$$

with $h_{El}(u)$ a polynomial, solves the equation. By replacing this form in the radial equation one derives the differential equation satisfied by h_{El}

$$\partial_u^2 h_{El} - 2(u - \frac{l+1}{u})\partial_u h_{El} + (\lambda - 2l - 3)h_{El} = 0. \tag{6.125}$$

To solve it with the "series method" we substitute the series form

$$h_{El}(u) = \sum_{k=0}^{\infty} a_k^{(l)}(\lambda)\ u^k, \tag{6.126}$$

collect the coefficients of u^k and set it to zero for each k. This generates a recursion relation for the coefficients

$$a^{(l)}_{k+2}(\lambda) = \frac{2k+2l+3-\lambda}{(k+2)(k+2l+3)}\; a^{(l)}_k(\lambda). \tag{6.127}$$

Since the leading behavior is already known, one must have $a^{(l)}_0 \neq 0$. Therefore, the solution is given by finding the coefficients $(a^{(l)}_0, a^{(l)}_2, a^{(l)}_4, \cdots)$ that involve only the *even* powers of k. For large values of $k = 2n$ the ratio of two consecutive terms in the series is u^2/k, which is similar to the series $\sum u^{2n}/n! \sim e^{u^2}$. Unless the series is cutoff into a polynomial this will destroy the good physical asymptotic behavior. Therefore, one must have a quantized value of λ

$$\lambda = 2n + 3 \tag{6.128}$$

so that the series is cutoff at some even value of $k = 2N$, with a_{2N} the last non-zero coefficient. The relation between N and n is then

$$n = l + 2N. \tag{6.129}$$

Therefore for fixed n the allowed values of l and energy are

$$\begin{aligned} E_n &= \hbar\omega\; (n + \frac{3}{2}) \\ n &= 0, 1, 2, \cdots \\ l &= n,\; n-2,\; n-4,\; \cdots, \begin{cases} 0 & n = even \\ 1 & n = odd\; . \end{cases} \end{aligned} \tag{6.130}$$

The first few eigenstates are listed

$$\begin{aligned} f_{00} &= a_{00}\, u\, e^{-u^2/2}, \\ f_{11} &= a_{11}\, u^2\, e^{-u^2/2}, \\ f_{20} &= a_{20} \left(u - \frac{2}{3} u^3 \right)\; e^{-u^2/2}, \\ f_{22} &= a_{22}\, u^3\, e^{-u^2/2}. \end{aligned} \tag{6.131}$$

6.9.1 Degeneracy and SU(3) symmetry

The energy is independent of l. With these restrictions we plot the energy eigenstates on a n versus l plot in Fig.(6.4)

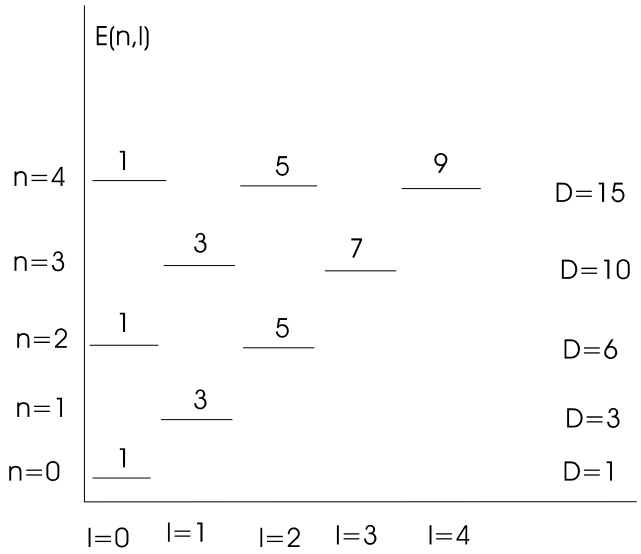


Fig (6.4) - Energy levels and degeneracies for the d=3 harmonic oscillator.

The degeneracy of each state with fixed l is due to the rotation invariance of the problem and is explained by the angular momentum quantum number $-1 \leq m \leq l$. This accounts for $2l+1$ states at the same energy level. However, as we see from the plot there are additional degeneracies for different values of l. Systematic degeneracies of this type can be explained only with a higher symmetry. In this case the symmetry is $SU(3)$. It was pointed out in the last chapter and it will be partially discussed now, but more deeply later. Let us first explain why $SU(3)$ is responsible for the degeneracy. The approach with the creation-annihilation operators explains the reason for the degeneracy because of the $SU(3)$ symmetry: Under an $SU(3)$ transformation $a'_I = U_{IJ}\, a_J$ the Hamiltonian is invariant, therefore the energy does not change. If the unitary matrix is expanded in terms of infinitesimal parameters h_{IJ} in the form $U = 1 + ih + \cdots$, with h Hermitian, then the infinitesimal transformation may be written in the form

$$\delta_h a_I = i h_{IJ} a_J. \tag{6.132}$$

The quantum generators of this transformations are the 9 operators $a^\dagger_I a_J$. This can be seen as follows. If we combine the parameters with the generators in the form $a^\dagger_I h_{IJ} a_j = a^\dagger h a$ (analogous to the case of translations $\mathbf{a}\cdot\mathbf{p}$, or rotations $\boldsymbol{\omega}\cdot\mathbf{J}$) then the infinitesimal transformation may be written as a commutator

$$\delta_h a_I = -i\left[a^\dagger h a, a_I\right] = i h_{IJ} a_J. \tag{6.133}$$

The commutator of two infinitesimal transformations follows from the formula

$$\left[a^\dagger h a, a^\dagger h' a\right] = a^\dagger \left[h, h'\right] a \tag{6.134}$$

This shows that these operators form a Lie algebra (see chapter on symmetry). One may also ask for the infinitesimal transformation on the quantum states

(similar to applying infinitesimal translations or rotations)

$$\delta_h|\psi\rangle = ia^\dagger h a|\psi\rangle. \tag{6.135}$$

In particular consider the transformation of the energy eigenstates in the basis given by the number states $|n_1, n_2, n_3\rangle$. Evidently, under the transformation one oscillator is annihilated but another one is created. Therefore, these states get mixed up with each other under this transformation, but without changing the *total* number n of creation operators applied on the vacuum. Therefore all the states with the same number n belong together in an $SU(3)$ multiplet and are indistinguishable from each other as far as the symmetry is concerned. Thus, different values of l that correspond to the same n must belong to the $SU(3)$ multiplet and hence must be degenerate.

The number of degenerate states D_n at the same level n was derived before for any dimension in Chap.5. In three dimensions it is given by

$$\begin{aligned} D_n(d) &= \frac{(d+n-1)!}{n!\ (d-1)!}, \\ D_n(3) &= \frac{(n+2)!}{2\ n!}. \end{aligned} \tag{6.136}$$

The correspondence between the number states $|n_1, n_2, n_3\rangle$ and the states $|n, l, m\rangle$ in the spherical basis, consistent with the number of degenerate states, are as follows

$$\begin{array}{lrl} n=0 & |0\rangle \ \Rightarrow & |0,0,\ m=0\rangle \\ n=1 & a_I^\dagger|0\rangle \ \Rightarrow & |1,1,\ m=0,\pm1\rangle \\ n=2 & a_I^\dagger a_J^\dagger|0\rangle \ \Rightarrow & |2,2,\ m=0,\pm1,\pm2\rangle \oplus |2,0,m=0\rangle \\ n=3 & a_I^\dagger a_J^\dagger a_K^\dagger|0\rangle \ \Rightarrow & |3,3,\ m=0,\pm1,\pm2,\pm3\rangle \oplus |3,1,m=0,\pm1\rangle \\ n=4 & a_I^\dagger a_J^\dagger a_K^\dagger a_L^\dagger|0\rangle \ \Rightarrow & \begin{array}{l} |4,4,\ m=0,\pm1,\pm2,\pm3,\pm4\rangle \\ \oplus|4,2,m=0,\pm1,\pm2\rangle \oplus |4,0,m=0\rangle \end{array} \\ \vdots & \vdots & \vdots \end{array} \tag{6.137}$$

We see that the creation-annihilation states are completely symmetric tensors of rank n in 3 dimensions, constructed from the vector $a_I^\dagger$. On the other hand we learned that the spherical harmonics are completely symmetric tensors of rank l that are also *traceless*. Therefore we expect that a decomposition of symmetric tensors into symmetric traceless tensors should correspond to the $|n, l, m\rangle$ labelling of the states. Thus

$$\begin{array}{ll} a_I^\dagger a_J^\dagger & = T^{(2)}_{IJ} + \frac{1}{d} g_{IJ} T_0^{(2)} \\ a_I^\dagger a_J^\dagger a_K^\dagger & = T^{(3)}_{IJK} + \frac{1}{d+2}(g_{IJ}T^{(3)}_K + g_{JK}T^{(3)}_I + g_{KI}T^{(3)}_J) \\ a_I^\dagger a_J^\dagger a_K^\dagger a_L^\dagger & = T^{(4)}_{IJKL} + \frac{1}{d+4} T_0^{(2)} \left(\begin{array}{c} g_{IJ}T^{(2)}_{KL} + g_{IK}T^{(2)}_{LJ} + g_{IL}T^{(2)}_{JK} \\ + g_{JK}T^{(2)}_{LI} + g_{JL}T^{(2)}_{IK} + g_{KL}T^{(2)}_{IJ} \end{array} \right) \\ & \quad + \frac{1}{d(d+2)} \left(g_{IJ}g_{KL} + g_{IK}g_{LJ} + g_{IL}g_{JK} \right) T_0^{(2)} T_0^{(2)} \\ \vdots & \vdots \end{array} \tag{6.138}$$

where $T^{(n)}_{I_1\cdots I_l}$ is a symmetric traceless tensor of rank l in d dimensions, and is constructed from n oscillators at energy level n. To compare it to the $|n,l,m\rangle$ notation above we set $d=3$. Here g_{IJ} may be taken as δ_{IJ} if one wishes to work in the $(1,2,3)$ basis, or it can be taken as in (6.89) if one prefers to work in the $(+,0,-)$ basis. Explicitly we have

$$\begin{aligned}
&T^{(0)}_0 = 1\\
&T^{(1)}_I = a^\dagger_I\\
&T^{(2)}_0 = g^{KL}a^\dagger_K a^\dagger_L\\
&T^{(2)}_{IJ} = a^\dagger_I a^\dagger_J - \tfrac{1}{d}g_{IJ}\ T^{(2)}_0\\
&T^{(3)}_I = a^\dagger_I\ T^{(2)}_0\\
&T^{(3)}_{IJK} = a^\dagger_I a^\dagger_J a^\dagger_K - \tfrac{1}{d+2}(g_{IJ}T^{(3)}_K + g_{JK}T^{(3)}_I + g_{KI}T^{(3)}_J)\\
&\vdots
\end{aligned} \tag{6.139}$$

The systematics of this decomposition shows that the degenerate states of level n are organized into angular momentum multiplets that correspond to tensors of rank $l = n,\ n-2,\ n-4,\ \cdots 0\ or\ 1$. This is precisely the information obtained from the radial equation above. Therefore, we have explained how the degenerate states with different l belong to the same $SU(3)$ multiplet. Furthermore, except for an overall normalization, we have also explicitly constructed the states with definite $|n,l,m\rangle$ in the form

$$T^{(n)}_{I_1\cdots I_l}\ |0\rangle, \quad l = n,\ n-2,\ n-4,\ \cdots 0\ or\ 1. \tag{6.140}$$

6.10 Hydrogen atom

We will consider a hydrogen-like atom consisting of a nucleus with charge Ze and an electron of charge $-e$. The Coulomb potential is attractive and given by

$$V = -\frac{Ze^2}{r}. \tag{6.141}$$

For a bound state the energy is negative $E = -|E|$. It will be convenient to rescale the radial variable $r = r_0 u$ and choose r_0 such as to make the energy term equal to $-1/4$, that is, $2\mu E r_0^2/\hbar^2 = -1/4$. Then the radial equation takes the form

$$\left(-\partial_u^2 + \frac{l(l+1)}{u^2} - \frac{\lambda}{u} + \frac{1}{4}\right) f_{El}(u) = 0 \tag{6.142}$$

where $\lambda = Ze^2 2\mu r_0/\hbar^2$. By eliminating r_0 we relate the energy and λ

$$E = -\frac{1}{2\lambda^2}\mu c^2 Z^2\alpha^2, \tag{6.143}$$

where the fine structure constant is used $\alpha = e^2/\hbar c$. Note also that r_0 is related to the Bohr radius a_0

$$r_0 = \lambda a_0/2Z, \quad a_0 = \frac{\hbar^2}{\mu e^2}. \tag{6.144}$$

The value of λ is to be determined as a quantum number that solves the radial equation.

The small $u \to 0$ leading behavior of the solution is determined by the $1/u^2$ in the equation. With the correct boundary conditions one finds $f \sim u^{l+1}$, as in the free particle case. The dominant term for $u \to \infty$ is determined by the constant $1/4$ term in the equation, which gives $f \sim e^{-u/2}$ for correct physical boundary conditions. Combining these two behaviors we expect a solution of the form

$$f_{El}(u) = u^{l+1} e^{-u/2} h_{El}(u). \tag{6.145}$$

Substituting this form into the radial equation gives a differential equation for h_{El}

$$\left[\partial_u^2 + \left(\frac{2l+2}{u} - 1\right)\partial_u + \frac{\lambda - l - 1}{u}\right] h_{El}(u) = 0. \tag{6.146}$$

Trying the series approach with

$$h_{El}(u) = \sum_{k=0}^{\infty} a_k(\lambda)\ u^k, \quad a_0 \neq 0, \tag{6.147}$$

and setting the coefficient of u^k equal to zero produces the recursion relation

$$a_{k+1}(\lambda) = \frac{k + l + 1 - \lambda}{(k+1)(k+2l+2)} a_k(\lambda). \tag{6.148}$$

If the series is not truncated to a polynomial the solution sums up to an unphysical asymptotic behavior. Therefore $\lambda = n$ must be an integer so that the series is truncated to a polynomial, with the highest non-zero coefficient being a_N for some $k = N$. The relation between the two integers is

$$n = N + l + 1 \tag{6.149}$$

where both l and N can start at zero, so that the lowest value of n is 1. N is called the radial quantum number and n is called the total quantum number. For a fixed value of n the allowed values of angular momentum and energy have emerged as

$$\begin{aligned} &E_n = -\tfrac{1}{n^2}\tfrac{\mu c^2}{2} Z^2\alpha^2 = -\tfrac{1}{n^2}\tfrac{Z^2 e^2}{2a_0} \\ &n = 1, 2, 3, \cdots \\ &l = 0, 1, 2, \cdots (n-1) \end{aligned} \tag{6.150}$$

We see that at the same value of energy there are many states with different values of l. For each l there are $(2l+1)$ rotational states distinguished from each other with $-l \leq m \leq l$. The total number of states is therefore

$$D_n = \sum_{l=0}^{n-1} (2l+1) = n^2. \tag{6.151}$$

These states are plotted in a n versus l plot in Fig.(6.5).

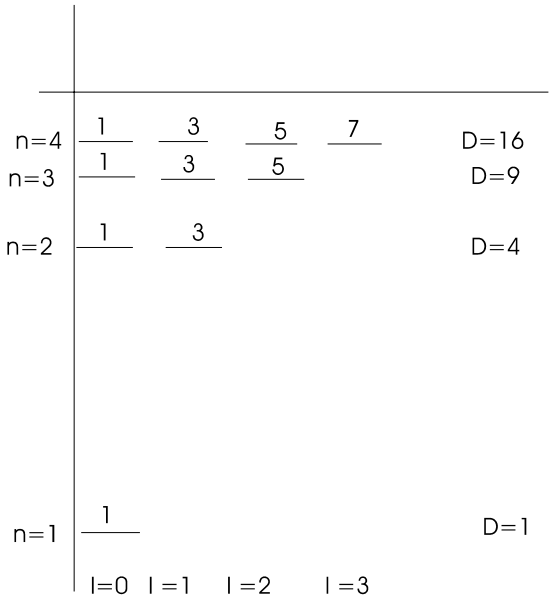


Fig.(6.5) - Degeneracies for H-atom.

As in the case of the Harmonic oscillator of the previous section we must look for a symmetry to explain systematically the reason for the degeneracy. This time it is not very easy to see the symmetry, but there is one, and it is $SU(2) \otimes SU(2) = SO(4)$. We need to learn about addition of angular momentum before we can explain how it works, therefore that discussion will be postponed till the symmetry chapter.

The recursion relations produce a polynomial which is recognized as the associated Laguerre polynomial with indices

$$h_{nl}(u) = L^{2l+1}_{n-l-1}(u) \tag{6.152}$$

up to an overall normalization. Putting everything together we list a few of the solutions for the normalized radial wavefunctions $R_{nl} = f_{nl}/u$

$$\begin{array}{ll} n=1 & R_{10}(r) = 2\left(Z/a_0\right)^{3/2} e^{-Zr/a_0} \\ n=2 & \begin{cases} R_{20}(r) = 2\left(Z/2a_0\right)^{3/2} e^{-Zr/2a_0}\ (1 - Zr/2a_0) \\ R_{21}(r) = \frac{2}{\sqrt{3}}\left(Z/2a_0\right)^{3/2} e^{-Zr/2a_0}\ (Zr/2a_0) \end{cases} \\ n=3 & \begin{cases} R_{30}\left(r\right) = 2\left(Z/3a_0\right)^{3/2} e^{-Zr/3a_0}\ \left[1 - 2Zr/3a_0 + \frac{2}{3}\left(Zr/3a_0\right)^2\right] \\ R_{31}\left(r\right) = \frac{4\sqrt{2}}{3}\left(Z/3a_0\right)^{3/2} e^{-Zr/3a_0}\ \left[1 - \frac{1}{2}Zr/3a_0\right] \\ R_{32}\left(r\right) = \frac{2\sqrt{2}}{3\sqrt{5}}\left(Z/3a_0\right)^{3/2} e^{-Zr/3a_0}\ \left(Zr/3a_0\right)^2 \end{cases} \end{array} \tag{6.153}$$

The full wavefunction

$$\psi_{nlm}(\mathbf{r}) = \langle \mathbf{r}|nlm\rangle = R_{nl}(r)\ Y_{lm}(\Omega) \tag{6.154}$$

is used to compute matrix elements of operators such as

$$\begin{aligned}\langle nlm|r^k|n'l'm'\rangle &= \int r^{k+2} dr d\Omega\; Y^*_{lm}(\Omega)\, Y_{l'm'}(\Omega)\; R^*_{nl}(r)\, R_{n'l'}(r) \\ &= \delta_{ll'}\delta_{mm'} \int_0^\infty dr\; r^{k+2} R^*_{nl}(r) R_{n'l}(r).\end{aligned} \tag{6.155}$$

Plots of probability distributions $|rR_{nl}(u)|^2$ provide an intuitive feeling for where the particle is located. In Fig.6.6 the $n = 3$ and $l = 0, 1, 2$ cases are plotted.

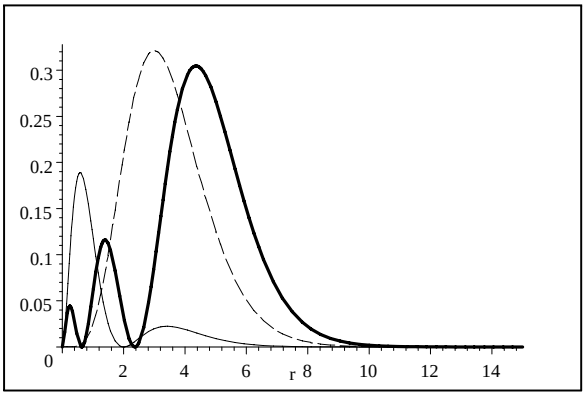


Fig.6.6: $l = 0$ thick,$l = 1$ thin,$l = 2$ dots

By examining such plots one easily sees that the number of peaks in the probability density is $n - l$. Only for the highest value of $l = n - 1$ there is a single peak and then the atom behaves roughly like the Bohr atom, but for lower values of l the behavior of the probability distribution is rather different. For fixed l, the highest peak occurs at larger values of r as n increases. This is in agreement with the intuition that the more excited atom is bigger. Also, for fixed n, the peak closest to the origin occurs at a greater value of r as the value of l increases. Thus the probability of finding the particle near the origin diminishes as l increases, as expected intuitively.

The expectation value of the radius is of interest since it gives information on the average position of the electron. For $n = n'$ the last integral in (6.155) is defined as the average value $\langle r^k\rangle_{nl}$. We give the result for a few values of k

$$\begin{aligned}\langle r\rangle_{nl} &= \tfrac{a_0}{2Z}\left(3n^2 - l(l+1)\right) \\ \langle r^2\rangle_{nl} &= \tfrac{a_0^2 n^2}{2Z^2}\left(5n^2 + 1 - 3l(l+1)\right) \\ \langle 1/r\rangle_{nl} &= \tfrac{Z}{na_0} \\ \langle 1/r^2\rangle_{nl} &= \tfrac{Z^2}{a_0^2 n^3 (l+1/2)}.\end{aligned} \tag{6.156}$$

From these expressions one sees that the average radius is larger as the energy (or n) increases, as expected intuitively.. The dependence on the angular momentum at a fixed value of l may be understood from the behavior of the wavefunction $|f_{El}|$ as in the figures above.

6.11 Problems

1. Using the equations of motion (6.2) or (6.7) prove that all the constants of motion listed in (6.8) are indeed time independent.

2. Using the basic commutation rules in the laboratory frame or in the center of mass frame, show that all the constants of motion listed in (6.8) commute with the Hamiltonian.

3. Using the commutation rules $[r_I, p_J] = i\hbar\delta_{IJ}$ show that the commutation relation among the spherical variables given in Eq.(6.30) follow, while all other commutators among $(r, p_r, \mathbf{\Omega}, \mathbf{L})$ vanish.

4. Use the definition of angular momentum $\mathbf{L}_I$ and the unit vector $\mathbf{\Omega}_I$ in terms of the original Cartesian operators $\mathbf{r}_I, \mathbf{p}_I$ and, while keeping track of orders of operators, prove the decomposition of the momentum operator into radial and angular parts as given in (6.31).

5. By using the relations in (6.31) prove that the commutation rules for the radial and angular operators given in (6.30) lead to the Cartesian commutation rules $[\mathbf{r}_I, \mathbf{p}_J] = i\hbar\delta_{IJ}$.

6. Prove the completeness and orthogonality relations for spherical harmonics of eq.(6.85) by using the properties of associated Legendre polynomials.

7. Write out the 7 independent components of the symmetric traceless tensor T_{IJK} explicitly in terms of (θ, ϕ) and compare them to the 7 spherical harmonics Y_{3m} given by the general formula. Verify that they agree with each other up to a normalization. (You may cut down your work to 4 functions by taking into account complex conjugation).

8. In two dimensions there is only one component of angular momentum $L_0 = r_1 p_2 - r_2 p_1$ that corresponds to rotations in the (1,2) plane. What is the differential operator form of L_0 in cylindrical coordinates, what are its eigenfunctions and eigenvalues, how many states correspond to the same eigenvalue? Analyze the Laplacian in 2 dimensions in cylindrical coordinates (i.e. $\mathbf{p}^2 = -\hbar^2\boldsymbol{\nabla}^2$), and find the radial equation. How do your results compare to the general expressions for $d-$dimensions given in the text?

9. In $d-$dimensions the generator of rotations in the (I, J) plane is written as
$$L_{IJ} = r_I p_J - r_J p_I, \quad I, J = 1, 2, \cdots d.$$
   - Show that the L_{IJ} commute with all dot products constructed from $(\mathbf{r}, \mathbf{p})$.
   - Show that the commutators of these operators satisfy the Lie algebra for $SO(d)$ given in Eq.(9.14).

- Prove the expression for L^2 in Eq.(6.95).
- Show that that Eq.(6.33) obtained in three dimensions is generalized to Eq.(6.96) in d dimensions.

10. Consider the traceless symmetric tensors T, T_I, T_{IJ}, T_{IJK} of rank $l = 0, 1, 2, 3$ in d-dimensions. List explicitly the independent components and verify that their number D_l agrees with eq.(6.100). For higher values of l how does the general formula for D_l compare with what you know in 2 and 3 dimensions from other approaches?

11. Solve the radial equation for the harmonic oscillator in $2-$dimensions. Discuss the boundary conditions and derive a recursion relation for determining the radial wavefunction. Putting all factors together, make sure to give the *full* wavefunction $\psi(\mathbf{r})$ with its radial part $R(r)$ and its angular dependence. Plot your states in the space of energy versus angular momentum, and label the degeneracy of each state. Let your solution guide you in keeping track of the number of states at each level.

12. Consider the solution for the two dimensional harmonic oscillator you obtained in the previous problem. Show how your result agrees with the creation-annihilation approach from the point of view of the energy eigenvalue and the number of states at each level. Then using the following $(+,-)$ basis for the harmonic oscillators $a_{\pm}^{\dagger} = (a_1^{\dagger} \pm i a_2^{\dagger})\sqrt{2}$ (which differs from the one in the text of chapter 5), find the correspondence between the harmonic oscillator states created by $a_{\pm}^{\dagger}$ and the angular momentum basis discussed in problem 8 above. Specifically, for *each state* at level n=4, show how you relate states of definite angular momentum with states constructed in terms of creation operators. This is similar to the correspondence between symmetric traceless tensors and spherical harmonics discussed in the text, but now it is in 2-dimensions instead of 3.

13. Solve the radial equation for the attractive spherical square well potential $V = -V_0\theta(a-r)$ in $d-$dimensions, assuming V_0 is positive. Give numerical values for the bound state energies. How many bound states are there?

14. Solve for the bound states in the delta shell potential $V = -V_0\delta(r/a - 1)$ in $d-$dimensions, assuming V_0 is positive.

15. Consider a spherically symmetric potential $V(r)$ in 3 dimensions, which consists of the infinite square well of width $2a$ and the delta-shell potential of strength V_0 located at $r = a$,

$$V(r) = \begin{cases} V_0\,\delta\left(\frac{r}{a} - 1\right) & for \quad 0 \le r \le 2a \\ +\infty & for \quad 2a \le r \end{cases} .$$

$V(r)$ is shown in the figure. Consider bound states at some energy E as indicated in the figure.

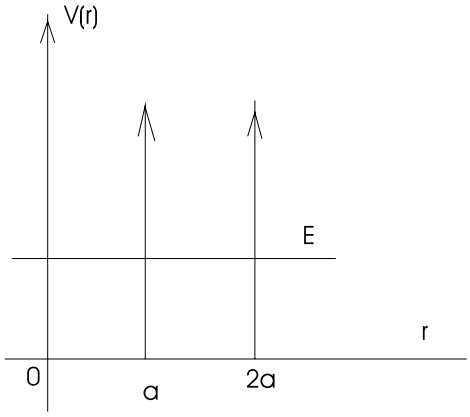


Figure 6.1: Fig.(6.7) - Delta shell at $r = a$, infinite walls at $r = 0$ and $r = 2a$.

(a) Give the solution of the radial Schrödinger equation in each region, consistent with boundary conditions, and continuity.

(b) What is the transcendental equation that determines the quantization of energy for any l?

(c) For $l = 0$ these equations take a simple form. Give a plot that determines roughly the energy of the first few levels for zero angular momentum.

16. Prove that in d dimensions $\mathbf{p}^2 = p_r^2 + \frac{1}{r^2}\left[L^2 + \frac{\hbar^2}{4}(d-1)(d-3)\right]$ where L^2 is given by $L^2 \equiv \frac{1}{2}\sum_{I,J=1}^{d} L_{IJ}^2$ and the radial momentum is the hermitian operator given by $p_r = \frac{1}{2}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \mathbf{p}\cdot\hat{\mathbf{r}}\right) = \frac{1}{2}\left(\frac{1}{r}\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r}\frac{1}{r}\right)$. Hint: first show that $L^2 = \mathbf{r}^2\mathbf{p}^2 - (\mathbf{r}\cdot\mathbf{p})(\mathbf{p}\cdot\mathbf{r}) - 2i\hbar\mathbf{r}\cdot\mathbf{p}$ and next show that $(\mathbf{r}\cdot\mathbf{p})(\mathbf{p}\cdot\mathbf{r}) + 2i\hbar\mathbf{r}\cdot\mathbf{p} = r^2p_r^2 + \frac{\hbar^2}{4}(d-1)(d-3)$ paying attention to the orders of operators in all steps.

17. Prove that in d dimensions the radial momentum acting on the wavefunction in position space is given by

$$p_r\psi(\mathbf{r}) = -i\hbar r^{-(d-1)/2}\,\frac{\partial}{\partial r}\left(r^{(d-1)/2}\psi(\mathbf{r})\right) = -i\hbar\left[\frac{\partial}{\partial r}\psi(\mathbf{r}) + \frac{1}{2}(d-1)\psi(\mathbf{r})\right],$$

where by definition $p_r\psi(\mathbf{r})$ means $p_r\psi(\mathbf{r}) \equiv \langle\mathbf{r}|p_r|\psi\rangle$.

18. Using the wavefunctions for the hydrogen atom, verify the average values $\langle r^k\rangle_{nl}$ for $k = \pm1, \pm2$, as given in Eq.(6.156), and interpret the results.

Chapter 7

PROPERTIES OF ROTATIONS

In the chapter on the central force problem we have studied some of the basic properties of angular momentum. We have seen that the angular momentum operators are the generators of infinitesimal rotations in 3-dimensions. In this chapter we will study rotations in more detail. We will discuss the classification of states according to their properties under rotations, the matrix representations of rotations on wavefunctions of all possible spins or angular momentum, the addition of angular momentum for products of states or operators, and related applications in physical systems.

Rotations form the Lie group SU(2), which is the simplest non-trivial example of a non-Abelian Lie group. Therefore the study of angular momentum or rotations is intimately connected to the study of representation theory of Lie groups. From this point of view the present discussion may be considered as an example of representation theory of Lie groups from a physical or quantum mechanical point of view.

7.1 The group of rotations

A group is characterized by a set G, and a "product" between group elements, with the properties of (i) closure, (ii) unity, (iii) inverse, (iv) associativity. These properties will be discussed in more detail in the chapter on symmetries. For now, let us consider the set G that consists of all possible rotations R in 3-dimensions, and define the product as one rotation followed by another rotation as applied on any object, for example on a chair, or a vector, or a spinning electron, or a complex molecule, or a quantum wavefunction, etc.. Then, at the intuitive level, one can visualize geometrically the four group properties as follows.

Closure means that some rotation followed by another rotation combine together to some overall rotation that is included in the set G. If we write the

product symbolically as

$$R_1 \cdot R_2 = R_3 \ , \tag{7.1}$$

then closure requires that R_3 must be in the set G. Note that the product of two rotations is not commutative: two rotations applied in different orders on the same object do not produce the same configuration of the object, unless the two rotations are applied along the same axis (try applying rotations on a book).

Unity means that applying no rotation at all will be considered to be a "rotation" that is an element of the set G. The unit element is represented by 1,and its properties are

$$1 \cdot R = R \cdot 1 = R. \tag{7.2}$$

That is, a rotation followed by no rotation is equivalent the original rotation, etc..

Inverse means that for every rotation that is included in the set G, the inverse of that rotation, that takes the object back to its original configuration, is also included in the set G. The inverse of the rotation R is denoted by R^{-1}. The property of the inverse is

$$R \cdot R^{-1} = R^{-1} \cdot R = 1. \tag{7.3}$$

Therefore, a rotation followed by its inverse gives *unity*, i.e. it is equivalent to doing nothing on the object.

Associativity is the property of the product that allows the grouping of the rotations as follows

$$R_1 \cdot R_2 \cdot R_3 = (R_1 \cdot R_2) \cdot R_3 = R_1 \cdot (R_2 \cdot R_3) \, . \tag{7.4}$$

It implies that when three rotations are applied on an object in the sequence R_1 followed by R_2 followed by R_3, that the overall rotation may be viewed as the product of two rotations, where the two rotations are either R_1 followed by $(R_2 \cdot R_3)$, or $(R_1 \cdot R_2)$ followed by R_3. It is easy to verify this property geometrically by applying various rotations on a book.

Rotations in 3-dimensions are implemented mathematically on quantum wavefunctions by the rotation operator

$$R(\boldsymbol{\omega}) = \exp(\frac{i}{\hbar}\boldsymbol{\omega} \cdot \mathbf{J}) \tag{7.5}$$

where the vector $\boldsymbol{\omega}$ represents an anti-clockwise angle of rotation of magnitude $\omega = |\boldsymbol{\omega}|$, with an axis of rotation in the direction $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega}/\omega$ (use the right-hand rule to visualize the rotation: thumb along axis, and fingers in the direction of rotation. See problem 1.)

$$\langle \psi |' = \langle \psi | \, e^{\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}} \ . \tag{7.6}$$

The same rotation on a ket is implemented by the Hermitian conjugate operator $\exp(-\frac{i}{\hbar}\boldsymbol{\omega} \cdot \mathbf{J})$. As seen in the previous chapter, the generator of infinitesimal rotations on any system (single particle or a complicated molecule) is the total

angular momentum operator $\mathbf{J}$. When $\boldsymbol{\omega}$ is infinitesimal, the Taylor expansion gives the first order term

$$\delta_\omega \langle\psi| = \frac{i}{\hbar}\langle\psi|\,\boldsymbol{\omega}\cdot\mathbf{J}\ . \tag{7.7}$$

When applied on quantum states, the *"product"* of rotations is inherited from the usual product of quantum operators. That is, a rotation of a quantum state with the angle $\boldsymbol{\omega}_1$, followed by a second rotation $\boldsymbol{\omega}_2$ is given by applying the appropriate rotation operators in the given sequence. Then $R_1 \cdot R_2 = R_3$ is represented on a bra in the form

$$\langle\psi|\, e^{\frac{i}{\hbar}\boldsymbol{\omega}_1\cdot\mathbf{J}}\ e^{\frac{i}{\hbar}\boldsymbol{\omega}_2\cdot\mathbf{J}} = \langle\psi|\, e^{\frac{i}{\hbar}\boldsymbol{\omega}_3\cdot\mathbf{J}} \tag{7.8}$$

With this definition of the "product", *closure* means that the right hand side can indeed be written as an exponential with an exponent that has a *single power* of $\mathbf{J}$ and an $\boldsymbol{\omega}_3$ that is a function of $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$. This is something that we need to prove, and furthermore we need to obtain the formula for $\boldsymbol{\omega}_3$. If we accept this property, then the remaining properties that characterize a group follows. The *unit element* for the group is the usual unit operator in the quantum mechanical Hilbert space, and it corresponds to zero rotation $1 = e^{\frac{i}{\hbar}\mathbf{0}\cdot\mathbf{J}}$. The *inverse* is obtained just by replacing $\boldsymbol{\omega}$ with $-\boldsymbol{\omega}$, since this represents the inverse rotation i.e.

$$R^{-1}(\boldsymbol{\omega}) = R(-\boldsymbol{\omega}) = \exp(-\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}). \tag{7.9}$$

Associativity of the group multiplication follows from the associativity of ordinary products of operators in quantum mechanics.

Therefore, provided we can prove closure, all the properties of a group are satisfied in the quantum mechanical setup. To prove closure consider the product of two exponentials $\exp A \times \exp B$, expand each one in powers and re-arrange terms to show that it is possible to re-sum the series into a single exponential $\exp C$ of the following form

$$\begin{aligned} e^A e^B &= (1 + A + \frac{A^2}{2!} + \cdots)(1 + B + \frac{B^2}{2!} + \cdots) \\ &= 1 + A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \frac{A^2B + AB^2}{2!} + \cdots \\ &= \exp\left(A + B + \frac{1}{2}\,[A,B] + \frac{1}{12}\,[A,[A,B]] + \frac{1}{12}\,[B,[B,A]] + \cdots\right) \end{aligned} \tag{7.10}$$

The crucial observation is that all the terms in the exponent $\exp C$ are constructed *only in terms of commutators*. In the present case

$$A = \frac{i}{\hbar}\boldsymbol{\omega}_1\cdot\mathbf{J}, \quad B = \frac{i}{\hbar}\boldsymbol{\omega}_2\cdot\mathbf{J}, \tag{7.11}$$

and the commutators constructed from these two operators are always proportional to $\mathbf{J}$ since we can use the commutator $[J_I, J_J] = i\hbar\varepsilon_{IJK}\,J_K$ to reduce the

multiple commutators to a single power of $\mathbf{J}$ as follows

$$\begin{aligned}\frac{1}{2}[A,B] &= \frac{1}{2}\left[\frac{i}{\hbar}\boldsymbol{\omega}_1\cdot\mathbf{J}\ ,\frac{i}{\hbar}\boldsymbol{\omega}_2\cdot\mathbf{J}\right] \\ &= \frac{-i}{2\hbar}\left(\boldsymbol{\omega}_1\times\boldsymbol{\omega}_2\right)\cdot\mathbf{J}\end{aligned} \tag{7.12}$$

and similarly

$$\begin{aligned}\frac{1}{12}[A[A,B]] &= \frac{i}{12\hbar}\left(\boldsymbol{\omega}_1\times\left(\boldsymbol{\omega}_1\times\boldsymbol{\omega}_2\right)\right)\cdot\mathbf{J} \\ \frac{1}{12}[B[B,A]] &= \frac{i}{12\hbar}\left(\boldsymbol{\omega}_2\times\left(\boldsymbol{\omega}_2\times\boldsymbol{\omega}_1\right)\right)\cdot\mathbf{J}\end{aligned} \tag{7.13}$$

etc.. Therefore, the exponent has the expected form $\exp C = \exp\left(\frac{i}{\hbar}\boldsymbol{\omega}_3\cdot\mathbf{J}\right)$ with

$$\begin{aligned}\boldsymbol{\omega}_3 = \boldsymbol{\omega}_1+\boldsymbol{\omega}_2 &- \frac{1}{2}\left(\boldsymbol{\omega}_1\times\boldsymbol{\omega}_2\right) \\ &+ \frac{1}{12}\left(\boldsymbol{\omega}_1\times\left(\boldsymbol{\omega}_1\times\boldsymbol{\omega}_2\right)\right) \\ &+ \frac{1}{12}\left(\boldsymbol{\omega}_2\times\left(\boldsymbol{\omega}_2\times\boldsymbol{\omega}_1\right)\right)+\cdots\end{aligned} \tag{7.14}$$

The full $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ will be given below in (7.14) to all orders by using a convenient representation of the commutation rules. It is important to realize that $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ is independent of the state $\langle\psi|$ and only depends on the structure constants ε_{IJK} that characterize the infinitesimal rotation, or the commutation rules of angular momentum. The formula for $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ is an intrinsic property of the rotation group; indeed it can be considered as the definition of the rotation group. The fact that $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ can be built up only from commutators (or the cross product characterized by ε_{IJK}) indicates that the infinitesimal rotations contain all the necessary information about all rotations. Therefore, understanding the properties of the infinitesimal generators $\mathbf{J}$, or the Lie algebra, amounts to understanding the properties of all rotations.

This discussion shows that the group closure property is satisfied, and furthermore, that it is equivalent to the closure of the Lie algebra. Moreover, using this result one can also show immediately that $R(-\boldsymbol{\omega}) = \exp\left(-\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}\right)$ is the inverse of $R(\boldsymbol{\omega})$. Indeed for $A = -B = \frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}$, the multiple commutators vanish $[A,B] = [A,-A] = 0$ and $A+B=0$, proving $R(\boldsymbol{\omega})R(-\boldsymbol{\omega}) = 1$.

Therefore, the rotation operators $R(\boldsymbol{\omega}) = \exp\left(\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}\right)$ acting on the quantum mechanical Hilbert space do form a Lie group.

7.2 Representations of angular momentum

It is useful to classify the quantum states according to their properties under rotations. This amounts to classifying the states according to their spin. This is something we have already done in the previous chapter by diagonalizing $\mathbf{J}^2$

and J_0 simultaneously and labelling the states by their eigenvalues $|j,m\rangle$. Recall that

$$\begin{aligned}\mathbf{J}^2|j,m\rangle &= \hbar^2 j(j+1)\,|j,m\rangle, \quad J_0|j,m\rangle = \hbar m|j,m\rangle \\ J_\pm|j,m\rangle &= \hbar\sqrt{j(j+1)-m(m\pm1)}\;|j,m\pm1\rangle.\end{aligned} \tag{7.15}$$

This means that the infinitesimal rotation of the state $\langle j,m|$ is computable, from

$$\begin{aligned}\delta_\omega\langle jm| &= \frac{i}{\hbar}\langle j,m|\boldsymbol{\omega}\cdot\mathbf{J} \\ &= \frac{i}{\hbar}\sum_{j'm'}\langle jm|\boldsymbol{\omega}\cdot\mathbf{J}|j'm'\rangle\,\langle j'm'|\end{aligned} \tag{7.16}$$

Similarly a finite rotation is computed as follows

$$\begin{aligned}\langle jm|' &= \langle jm|\,e^{\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}} \\ &= \sum_{j'm'}\langle jm|e^{\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}}|j'm'\rangle\langle j'm'|.\end{aligned} \tag{7.17}$$

To implement the rotation we need to compute the matrix elements of the generators

$$\langle jm|J_I|j^{'}m^{'}\rangle = (J_I)_{jm,j'm'}. \tag{7.18}$$

They are arranged in a block diagonal matrix, since J_I does not mix states with different values of spin j

$$\langle jm|J_I|j'm'\rangle = \delta_{jj'}\langle jm|J_I|jm'\rangle = \delta_{jj'}\;\hbar\,(\Im_I)^j_{mm'} \tag{7.19}$$

where $(\Im_I)^j_{mm'}$ are $(2j+1)\times(2j+1)$ matrices. We can then construct a block diagonal matrix for each J_I as follows:

$$\begin{array}{c} \\ \left\langle \begin{smallmatrix} j=0 \\ m=0 \end{smallmatrix}\right| \\ \left\langle \begin{smallmatrix} j=1/2 \\ m=\pm 1/2 \end{smallmatrix}\right| \\ \left\langle \begin{smallmatrix} j=1 \\ m=1,0,-1 \end{smallmatrix}\right| \\ \left\langle \begin{smallmatrix} j=\frac{3}{2} \\ m=\pm\frac{3}{2},\pm\frac{1}{2} \end{smallmatrix}\right| \\ \vdots \end{array} \begin{array}{c} \left|\begin{smallmatrix} j'=0 \\ m'=0 \end{smallmatrix}\right\rangle \quad \left|\begin{smallmatrix} j'=\frac{1}{2} \\ m'=\pm\frac{1}{2} \end{smallmatrix}\right\rangle \quad \left|\begin{smallmatrix} j'=1 \\ m'=1,0,-1 \end{smallmatrix}\right\rangle \quad \left|\begin{smallmatrix} j'=\frac{3}{2} \\ m'=\pm\frac{3}{2},\pm\frac{1}{2} \end{smallmatrix}\right\rangle \quad \cdots \\ \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & \boxed{2\times 2} & 0 & 0 & \cdots \\ 0 & 0 & \boxed{\begin{matrix}3\times & 3\end{matrix}} & 0 & \cdots \\ 0 & 0 & 0 & \boxed{\begin{matrix}4\times & 4\end{matrix}} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array} \tag{7.20}$$

The above block structure of the matrix representation of angular momentum tells us that only states that have the same total spin (same value of j) can

rotate into each other. For example, consider a state $\langle \frac{1}{2} m|$ which represent a particle with spin $j = 1/2$ and $(m = \pm\frac{1}{2})$. When the matrix $(\omega_I \Im_I)^{j=\frac{1}{2}}_{mm'}$ is applied to such a state, then the resulting state will be of the form

$$\delta_\omega \langle \frac{1}{2}, m| = \alpha(m)\ \langle \frac{1}{2}, \frac{1}{2}| + \beta(m)\ \langle \frac{1}{2}, -\frac{1}{2}|. \tag{7.21}$$

Thus, the spin $j = 1/2$ states do not mix with the states of another spin under rotations. The same is obviously true for any spin. Therefore, for a fixed value of j the states $|jm\rangle$, $j \geq m \geq -j$ are said to form an *irreducible multiplet* under rotations.

Let us see in more detail what the matrix elements look like for the operators J_0, J_+, J_-. Let us fix j, and let us examine first $\langle jm|J_0|jm'\rangle = \hbar m' \langle jm|jm'\rangle = \hbar m \delta_{mm'}$, so that

$$(\Im_0)^{\,j}_{mm'} = \begin{pmatrix} j & 0 & \dots & \dots & & \dots & 0 \\ 0 & j-1 & 0 & \dots & & \dots & 0 \\ \vdots & 0 & \ddots & 0 & & \dots & 0 \\ \vdots & \vdots & 0 & m & & \ddots & \vdots \\ & & & & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & & -j+1 & 0 \\ 0 & \dots & \dots & \dots & & 0 & -j \end{pmatrix} \tag{7.22}$$

For $\langle jm|J_+|jm'\rangle$ we have

$$\begin{aligned} \langle jm|J_+|jm'\rangle &= \hbar\sqrt{j(j+1) - m'(m'+1)}\langle jm|j, m+1\rangle \\ &= \hbar\sqrt{j(j+1) - m'(m'+1)}\delta_{m,m'+1} \end{aligned} \tag{7.23}$$

and so

$$(\Im_+)^{j}_{mm'} = \begin{pmatrix} 0 & \sqrt{2j} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sqrt{4j-2} & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \sqrt{4j-2} & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 & \sqrt{2j} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \tag{7.24}$$

Similarly for J_-

$$\begin{aligned} \langle jm|J_-|jm'\rangle &= \hbar\sqrt{j(j+1) - m'(m'-1)}\langle jm, jm'-1\rangle \\ &= \hbar\sqrt{j(j+1) - m'(m'-1)}\delta_{m,m'-1} \end{aligned} \tag{7.25}$$

so that

$$(\Im_-)^j_{mm'} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \sqrt{2j} & 0 & 0 & 0 & \ddots & \vdots & \vdots \\ 0 & \sqrt{4j-2} & 0 & \ddots & 0 & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & 0 & \sqrt{4j-2} & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \sqrt{2j} & 0 \end{pmatrix} \tag{7.26}$$

For every j the matrices $(\Im_{0,\pm})^j_{mm'}$ satisfy the same commutation rules as the operators $\frac{1}{\hbar}J_{0,\pm}$. That is, under matrix multiplication we have

$$\begin{aligned} \left[(\Im_I)^j, (\Im_J)^j\right]_{mm'} &= \textstyle\sum_{m''} \left\{(\Im_I)^j_{mm''}(\Im_J)^j_{m''m'} - (\Im_J)^j_{mm''}(\Im_I)^j_{m''m'}\right\} \\ &= \textstyle\sum_{j''m''} \begin{pmatrix} \langle jm|\frac{1}{\hbar}J_I|j''m''\rangle\langle j''m''|\frac{1}{\hbar}J_J|jm\rangle \\ -\langle jm|\frac{1}{\hbar}J_J|jm''\rangle\langle jm''|\frac{1}{\hbar}J_I|jm\rangle \end{pmatrix} \\ &= \tfrac{1}{\hbar^2}\langle jm|J_IJ_J - J_JJ_I|jm'\rangle \\ &= i\varepsilon_{IJK}\langle jm|\tfrac{1}{\hbar}J_K|jm'\rangle \\ &= i\varepsilon_{IJK}\ (\Im_K)^j_{mm'} \end{aligned} \tag{7.27}$$

In the second line only $j'' = j$ contributes, but the full sum over all j'', m'' allows us to use the completeness relation $\sum_{j''m''}|jm''\rangle\langle jm''| = \mathbf{1}$, to derive the third line. The final result shows that we have obtained an infinite number of *matrix representations* of the commutation rules of angular momentum, one for every value of $j = 0, 1/2, 1, 3/2, \cdots$.

We can see in particular what these matrices look like for $j = \frac{1}{2}$

$$\begin{aligned} (\Im_0)^{j=1/2}_{mm'} &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_{mm'} \\ (\Im_+)^{j=1/2}_{mm'} &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_{mm'} \\ (\Im_-)^{j=1/2}_{mm'} &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_{mm'} \end{aligned} \tag{7.28}$$

From these we construct the 2×2 matrix representation of $\mathbf{J}$

$$(\vec{\Im})^{j=1/2}_{mm'} = \frac{1}{\hbar}\langle 1/2, m|\mathbf{J}|1/2, m'\rangle = \left(\frac{\boldsymbol{\sigma}}{2}\right)_{mm'} \tag{7.29}$$

where $\boldsymbol{\sigma}$ are the three Pauli matrices. Therefore, for spin $j = 1/2$, rotations are represented by $\mathbf{J} \to \hbar\left(\frac{\boldsymbol{\sigma}}{2}\right)$.

An additional property of the Pauli matrices that we will use below is

$$\sigma_I\sigma_J = \delta_{IJ}1 + i\varepsilon_{IJK}\,\sigma_K. \tag{7.30}$$

That is, while the commutator of two Pauli matrices reproduces the commutation rules of angular momentum, the anti-commutator is proportional to the identity matrix $\{\sigma_I, \sigma_J\} = 2\delta_{IJ}1$. Unlike the commutation property which is the same for all j, the anti-commutation property is special for the $j = 1/2$ representation and it does not hold for higher values of j.

Similarly, for $j = 1$ the matrix representation is

$$
\begin{aligned}
(\Im_0)^{j=1}_{mm'} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
(\Im_+)^{j=1}_{mm'} &= \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \\
(\Im_-)^{j=1}_{mm'} &= \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}
\end{aligned}
\tag{7.31}
$$

These give the infinitesimal rotations of a vector $\mathbf{V}$ when written in the $(+, 0, -)$ basis instead of the $(1, 2, 3)$ basis. That is $\delta_\omega \mathbf{V} = \boldsymbol{\omega} \times \mathbf{V}$ can also be written in one of the following forms

$$
\begin{pmatrix} \delta_\omega V_x \\ \delta_\omega V_y \\ \delta_\omega V_z \end{pmatrix} = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \tag{7.32}
$$

$$
\begin{pmatrix} \delta_\omega V_+ \\ \delta_\omega V_0 \\ \delta_\omega V_- \end{pmatrix} = -i \begin{pmatrix} \omega_0 & \frac{\omega_+}{\sqrt{2}} & 0 \\ \frac{\omega_-}{\sqrt{2}} & 0 & \frac{\omega_+}{\sqrt{2}} \\ 0 & \frac{\omega_-}{\sqrt{2}} & -\omega_0 \end{pmatrix} \begin{pmatrix} V_+ \\ V_0 \\ V_- \end{pmatrix} \tag{7.33}
$$

where we have used the definitions $V_m \equiv \langle j = 1,\, m|V\rangle$, with

$$
\begin{aligned}
V_0 &= V_z, \quad & V_\pm &= \mp \left(V_x \mp i V_y\right)/\sqrt{2}, \\
\omega_0 &= \omega_z, \quad & \omega_\pm &= \left(\omega_x \mp i\omega_y\right).
\end{aligned}
\tag{7.34}
$$

So the 3×3 matrices that appear in the (x, y, z) or $(+, 0, -)$ bases in (7.32,7.33) are just $(-i\vec{\omega}\cdot\vec{\Im})^{j=1}_{mm'} = -i(\omega_0\Im_0 + \frac{1}{2}\omega_+\Im_+ + \frac{1}{2}\omega_-\Im_-)^{j=1}_{mm'}$ respectively.

7.3 Finite rotations and the D^j matrices

Let us now compute the finite rotation defined by the matrix in (7.17) which is nonzero only when $j' = j$

$$
\langle jm|' = \sum_{m'} D^j_{mm'}(\boldsymbol{\omega})\langle jm'|, \quad D^j_{mm'}(\boldsymbol{\omega}) = \langle jm|e^{\frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}}|jm'\rangle \tag{7.35}
$$

The exponential may be expanded $D^j_{mm'}(\boldsymbol{\omega}) = \sum_n \frac{i^n}{n!}\langle jm|\left(\frac{\boldsymbol{\omega}\cdot\mathbf{J}}{\hbar}\right)^n|jm'\rangle$ and then the identity is inserted in the form $\mathbf{1} = \sum_{j''m''}|j''m''\rangle\langle j''m''|$. However,

since only $j'' = j$ can contribute, only the sums over the $m_1, m_2, \cdots$ survive as follows

$$\begin{aligned}
&\langle jm| \left(\tfrac{\boldsymbol{\omega}\cdot\mathbf{J}}{\hbar}\right)^n |jm'\rangle \\
&= \textstyle\sum_{m_1,\cdots m_{n-1}} \left(\begin{array}{c} \langle jm|\frac{\boldsymbol{\omega}\cdot\mathbf{J}}{\hbar}|jm_1\rangle\langle jm_1|\frac{\boldsymbol{\omega}\cdot\mathbf{J}}{\hbar}|jm_2\rangle\langle jm_2\cdots \\ \cdots |jm_{n-1}\rangle\langle jm_{n-1}|\frac{\boldsymbol{\omega}\cdot\mathbf{J}}{\hbar}|jm'\rangle \end{array} \right) \\
&= \textstyle\sum_{m_1,\cdots m_{n-1}} (\vec{\omega}\cdot\vec{\Im})^j_{mm_1} (\vec{\omega}\cdot\vec{\Im})^j_{m_1\,m_2} \cdots (\vec{\omega}\cdot\vec{\Im})^j_{m_{n-1}\,m'} \\
&= \left(\left[(\vec{\omega}\cdot\vec{\Im})^j \right]^n \right)_{mm'}
\end{aligned} \tag{7.36}$$

This shows that one way of computing the matrix $D^j_{mm'}(\boldsymbol{\omega})$ is by summing the exponential series of matrices

$$\begin{aligned}
D^j_{mm'}(\boldsymbol{\omega}) &= \textstyle\sum_n \frac{1}{n!} \left(\left[i(\vec{\omega}\cdot\vec{\Im})^j \right]^n \right)_{mm'} \\
&= \left(\exp\left[i(\vec{\omega}\cdot\vec{\Im})^j \right] \right)_{mm'}
\end{aligned} \tag{7.37}$$

These matrices satisfy the group property for each value of j

$$\sum_{m''} D^j_{mm''}(\boldsymbol{\omega}_1) D^j_{m''m'}(\boldsymbol{\omega}_2) = D^j_{mm'}(\boldsymbol{\omega}_3) \tag{7.38}$$

This can be shown directly by using the same steps as eqs.(7.10-7.14), but now using matrices $A = i(\tilde{\boldsymbol{\omega}}_1\cdot\vec{\Im})^j$ and $B = i(\tilde{\boldsymbol{\omega}}_2\cdot\vec{\Im})^j$. Another proof is obtained by inserting the identity operator and using directly the closure property of the operators

$$\begin{aligned}
&\sum_{m''} D^j_{mm''}(\boldsymbol{\omega}_1) D^j_{m''m'}(\boldsymbol{\omega}_2) \\
&= \sum_{j''m''} \langle jm| e^{\frac{i}{\hbar}\boldsymbol{\omega}_1\cdot\mathbf{J}} |j''m''\rangle\langle j''m''| e^{\frac{i}{\hbar}\boldsymbol{\omega}_2\cdot\mathbf{J}} |jm'\rangle \\
&= \langle jm| e^{\frac{i}{\hbar}\boldsymbol{\omega}_1\cdot\mathbf{J}} e^{\frac{i}{\hbar}\boldsymbol{\omega}_2\cdot\mathbf{J}} |jm'\rangle \\
&= \langle jm| e^{\frac{i}{\hbar}\boldsymbol{\omega}_3\cdot\mathbf{J}} |jm'\rangle \\
&= D^j_{mm'}(\boldsymbol{\omega}_3).
\end{aligned} \tag{7.39}$$

The same $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$ is obtained for every value of the spin j since this follows from only the commutation rules among the operators $\mathbf{J}$ that are independent of the value of j (see also (7.27)). Therefore, the $D^j(\boldsymbol{\omega})$ matrices are said to form a matrix *representation* of the rotation group. They simply represent the action of rotations on the multiplet of spin j.

7.3.1 Relation to spherical harmonics

For the special value $m = 0$ the entries in the column $D^l_{0m'}$ in the rotation matrix are related to the spherical harmonics $D^l_{0m'} \sim Y_{lm'}$. To see this, recall the definition of the spherical harmonics in the form $Y_{lm'}(\theta,\phi) = \langle\theta,\phi|lm'\rangle$ where

the position space state $\langle\theta,\phi|$ is associated with the unit vector $\hat{\mathbf{r}} = \cos\phi\,\sin\theta\,\hat{\mathbf{x}} + \sin\phi\,\sin\theta\,\hat{\mathbf{y}} + \cos\theta\,\hat{\mathbf{z}}$. This unit vector is obtained by applying a rotation to the unit vector in the $\hat{\mathbf{z}}$ direction. So, the state $\langle\theta,\phi|$ can be rewritten by rotating the state $\langle\hat{\mathbf{z}}| = \langle\theta=0,\phi=0|$

$$\langle\theta,\phi| = \langle\hat{\mathbf{z}}|\, e^{i\theta L_2/\hbar}\, e^{i\phi L_3/\hbar}. \tag{7.40}$$

Therefore, we have

$$\begin{aligned} Y_{lm'}(\theta,\phi) &= \langle\hat{\mathbf{z}}|\, e^{i\theta L_2/\hbar}\, e^{i\phi L_3/\hbar}|lm'\rangle \\ &= \sum_m \langle\hat{\mathbf{z}}|lm\rangle\langle lm|\, e^{i\theta L_2/\hbar}\, e^{i\phi L_3/\hbar}|lm'\rangle \\ &= \sum_m Y_{lm}(0,0)\, D^l_{mm'}(0,\theta,\phi). \end{aligned} \tag{7.41}$$

Furthermore, we have $Y_{lm}(0,0) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m,0}$, which leads to

$$Y_{lm'}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}} D^l_{0m'}(0,\theta,\phi). \tag{7.42}$$

So, through the knowledge of the spherical harmonics we have obtained some of the matrix elements of rotations

$$\begin{aligned} D^l_{0m'}(0,\theta,\phi) &= \sqrt{\frac{4\pi}{2l+1}} Y_{lm'}(\theta,\phi) \\ &= (-1)^m \sqrt{\frac{(l-m')!}{(l+m')!}} P_l^{m'}(\cos\theta)\, e^{im'\phi}, \end{aligned} \tag{7.43}$$

where $P_l^{m'}(\cos\theta)$ is the associated Legendre polynomial.

7.4 Computation of the D^j matrices

There are several methods for computing $D^j(\boldsymbol{\omega})$ for general j. The first one is by summing the matrix series $D^j(\boldsymbol{\omega}) = e^{i\mathfrak{J}\cdot\boldsymbol{\omega}}$ which works well for small values of j, as illustrated below. However, this becomes cumbersome for larger values of j. The second method is to introduce Euler angles and derive a differential equation which is solved by hypergeometric functions. The third method is to use the two dimensional harmonic oscillator as a model for $SU(2)$,which readily yields a nice result in the form of a polynomial. The latter two methods are explained below. The relation among these methods provides a first look at representation theory of Lie groups, the rotation group being the simplest case of a non-Abelian Lie group.

7.4.1 Spin j=1/2 case

To compute explicitly the spin-1/2 representation $D^{\frac{1}{2}}(\boldsymbol{\omega})$ by direct exponentiation one uses the properties of Pauli matrices that follow from (7.30), namely

$(\hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma})^2 = 1$, or $(\hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma})^{even} = 1$, $(\hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma})^{odd} = \hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma}$. Then

$$\begin{aligned} D^{\frac{1}{2}}(\boldsymbol{\omega}) = e^{\frac{i}{2}\boldsymbol{\omega}\cdot\boldsymbol{\sigma}} &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{i\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}{2}\right)^n \\ &= \sum_{n=even}\frac{1}{n!}\left(\frac{i\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}{2}\right)^n + \sum_{n=odd}\frac{1}{n!}\left(\frac{i\boldsymbol{\omega}\cdot\boldsymbol{\sigma}}{2}\right)^n \\ &= \sum_{n=even}\frac{1}{n!}\left(\frac{i\omega}{2}\right)^n + \hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma}\sum_{n=odd}\frac{1}{n!}\left(\frac{i\omega}{2}\right)^n \\ &= \cos\frac{\omega}{2} + i\hat{\boldsymbol{\omega}}\cdot\boldsymbol{\sigma}\ \sin\frac{\omega}{2} \\ &= \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \end{aligned} \tag{7.44}$$

where

$$\begin{aligned} \alpha &= \cos\frac{\omega}{2} + i\hat{\boldsymbol{\omega}}_z\ \sin\frac{\omega}{2} \\ \beta &= (\hat{\boldsymbol{\omega}}_y + i\hat{\boldsymbol{\omega}}_x)\ \sin\frac{\omega}{2} \\ 1 &= |\alpha|^2 + |\beta|^2 . \end{aligned} \tag{7.45}$$

An arbitrary matrix of the form

$$\begin{aligned} U &= \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} = a + i\mathbf{b}\cdot\boldsymbol{\sigma}, \\ \det U &= |\alpha|^2 + |\beta|^2 = a^2 + \mathbf{b}^2 = 1 \end{aligned} \tag{7.46}$$

is a special unitary matrix that is a member of the set $SU(2)$, where "special" means $detU = 1$, but otherwise arbitrary. Such matrices form a group, since they close into the same set under matrix multiplication (try it!). But we have seen by construction that $D^{j=1/2}(\boldsymbol{\omega})$ is the most general such matrix, and that it forms the rotation group $D^{j=1/2}(\boldsymbol{\omega}_1)\,D^{j=1/2}(\boldsymbol{\omega}_2) = D^{j=1/2}(\boldsymbol{\omega}_3)$. Therefore, modulo a subtlety that is pointed out below, the group $SU(2)$ may be interpreted physically as the rotation group in 3 dimensions.

7.4.2 Spin j=1 case

In the case of $j = 1$, a unitary transformation from the $(+, 0, -)$ basis to the (x, y, z) basis, as in (7.34), shows that the rotation in the (x, y, z) basis is given by the matrix $V'_I = \left(D^{j=1}(\boldsymbol{\omega})\right)_{IJ} V_J$ in the form

$$\begin{pmatrix} V'_x \\ V'_y \\ V'_z \end{pmatrix} = \exp\begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix}\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \tag{7.47}$$

This matrix notation is equivalent to the rotation of a vector in the usual vector notation in the form

$$\begin{aligned}\mathbf{V}' &= \exp(\boldsymbol{\omega}\times)\,\mathbf{V} = \mathbf{V} + \boldsymbol{\omega}\times\mathbf{V} + \frac{1}{2!}\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{V}) + \cdots \\ &= \mathbf{V} + \hat{\boldsymbol{\omega}}\times\mathbf{V}\ \sin\omega + \hat{\boldsymbol{\omega}}\times(\hat{\boldsymbol{\omega}}\times\mathbf{V})\ (1-\cos\omega)\,.\end{aligned} \tag{7.48}$$

where the second line is obtained by summing the series. If this rotation is written in 3×3 matrix form, $V'_I = R_{IJ}V_J$, it yields $D^{j=1}(\boldsymbol{\omega}) = R$ in the (x,y,z) basis (see problem 2 for direct exponentiation in $(+,0,-)$ basis). It can be verified that the resulting matrix $R_{IJ}(\boldsymbol{\omega})$ is a parametrization of the most general 3×3 orthogonal matrix with determinant $+1$. Orthogonal 3×3 matrices form the group $SO(3)$ (prove it!), which is the rotation group in 3-dimensions as we have seen in this discussion, by direct construction.

There is a subtle difference between $SU(2)$ and $SO(3)$. In the $D^{j=1/2}(\boldsymbol{\omega})$ representation it takes an angle of rotation of 4π to come back to the same point. For example, for a rotation of $\omega = 2\pi$ around any axis, we see from (7.45) that $D^{j=1/2}(2\pi\hat{\boldsymbol{\omega}}) = -1$. More generally, for the same 2π rotation we get $D^{j}(2\pi\hat{\boldsymbol{\omega}}) = (-1)^{2j}$ (see e.g. the rotation of a vector (7.48). Thus, states with integer spin come back to the same state after a 2π rotation, but it takes a 4π rotation to get back to the same state if $j = \frac{1}{2}$+integer. The difference between $SU(2)$ and $SO(3)$ is whether or not the representation space contains spin-1/2+integer or not, and consequently whether we should consider the maximum range of rotations to be 4π or 2π respectively. This difference becomes apparent globally. In physical applications the value of the angular momentum determines the range of the angles.

7.4.3 Product of two rotations

The explicit matrix form of $D^{j=1/2}(\boldsymbol{\omega})$ as in (7.44,7.45) is very useful for many purposes. In particular it provides the simplest way of computing $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$. Recall that according to (7.38) we obtain the same $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$ for any value of j. By using the map between rotations and $SU(2)$ matrices provided by (7.45) we can easily compute the product of two rotations and extract $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1,\boldsymbol{\omega}_2)$. The result is (use (7.46) and (7.30) to prove it)

$$\begin{aligned}\cos\frac{\omega_3}{2} &= \cos\frac{\omega_1}{2}\cos\frac{\omega_2}{2} - (\hat{\boldsymbol{\omega}}_1\cdot\hat{\boldsymbol{\omega}}_2)\ \sin\frac{\omega_1}{2}\sin\frac{\omega_2}{2} \\ \hat{\boldsymbol{\omega}}_3\sin\frac{\omega_3}{2} &= \begin{pmatrix} \hat{\boldsymbol{\omega}}_1\sin\frac{\omega_1}{2}\cos\frac{\omega_2}{2} + \hat{\boldsymbol{\omega}}_2\sin\frac{\omega_2}{2}\cos\frac{\omega_1}{2} \\ -(\hat{\boldsymbol{\omega}}_1\times\hat{\boldsymbol{\omega}}_2)\ \sin\frac{\omega_1}{2}\sin\frac{\omega_2}{2} \end{pmatrix}\end{aligned} \tag{7.49}$$

The series expansion of these expressions for small $\boldsymbol{\omega}_1,\boldsymbol{\omega}_2$ reproduces the first few terms given earlier in (7.14) (see problem).

7.4.4 Euler angles

An arbitrary rotation may be parametrized in terms of Euler angles which are defined as follows. Consider first an anti-clockwise rotation of angle γ along the

positive z-axis, followed by an anti-clockwise rotation of angle θ along the new y'-axis, and then followed by anti-clockwise rotation of angle ϕ along the new z''-axis. These are pictured in Fig.(7.1).

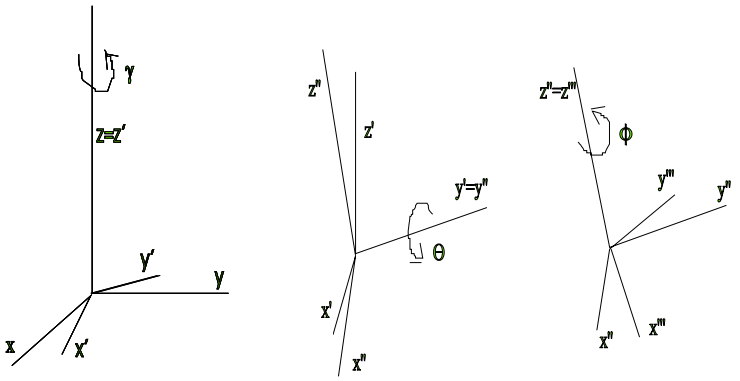


Fig.(7.1) - Euler angles.

On a bra the Euler rotations are implemented by the following operations

$$\langle\psi|' = \langle\psi| e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} = \langle\psi| e^{i\gamma J_3/\hbar} e^{i\theta J_2'/\hbar} e^{i\phi J_3''/\hbar} \tag{7.50}$$

The operators $J_2' \equiv \mathbf{J}\cdot\hat{y}'$ and $J_3'' \equiv \mathbf{J}\cdot\hat{z}''$ are given by

$$\begin{aligned} J_2' &= J_2\cos\gamma - J_1\sin\gamma = e^{-i\gamma J_3/\hbar} J_2\, e^{i\gamma J_3/\hbar} \\ J_3'' &= J_3'\cos\theta + J_1'\sin\theta = e^{-i\theta J_2'/\hbar} J_3' e^{i\theta J_2'/\hbar} \end{aligned} \tag{7.51}$$

where $J_1' \equiv \mathbf{J}\cdot\hat{x}'$, $J_3' = J_3$. These are obtained by inspecting Fig.(7.1) and noting the decomposition of the unit vectors $\hat{y}'$ and $\hat{z}''$ into components in the appropriate orthogonal basis. The inverse of these relations are

$$J_2 = e^{i\gamma J_3/\hbar} J_2' e^{-i\gamma J_3/\hbar} \tag{7.52}$$

$$J_3' = e^{i\theta J_2'/\hbar} J_3'' e^{-i\theta J_2'/\hbar} = J_3. \tag{7.53}$$

The Euler rotations may be rewritten in terms of the original basis J_1, J_2, J_3 as follows:

$$\begin{aligned} e^{i\gamma J_3/\hbar} e^{i\theta J_2'/\hbar} e^{i\phi J_3''/\hbar} &= e^{i\gamma J_3\hbar} \left\{ e^{i\theta J_2'/\hbar} e^{i\phi J_3''/\hbar} e^{-i\theta J_2'/\hbar} \right\} e^{i\theta J_2'/\hbar} \\ &= e^{i\gamma J_3/\hbar} e^{i\phi J_3/\hbar} e^{i\theta J_2'/\hbar} \\ &= e^{i\phi J_3/\hbar} \left\{ e^{i\gamma J_3/\hbar} e^{i\theta J_2'/\hbar}\, e^{-i\gamma J_3/\hbar} \right\} e^{i\gamma J_3/\hbar} \\ &= e^{i\phi J_3/\hbar} e^{i\theta J_2/\hbar} e^{i\gamma J_3/\hbar} \end{aligned} \tag{7.54}$$

In line 1, first insert $1 = e^{-i\theta J_2'/\hbar} e^{i\theta J_2'/\hbar}$ and then use eq.(7.53) to go to line 2. To go from line 2 to line 3, first interchange the order of the left-most two

commuting factors and insert $1 = e^{-i\gamma J_3/\hbar}\, e^{i\gamma J_3/\hbar}$ on the right. Finally to go from line 3 to line 4 use eq.(7.52). Thus, the general rotation takes the form

$$e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} = e^{i\phi J_3/\hbar} e^{i\theta J_2/\hbar} e^{i\gamma J_3/\hbar}. \tag{7.55}$$

According to the group property of rotations (7.38) this operator relation must also be true for every matrix representation, hence

$$D^j\left(\boldsymbol{\omega}\right)_{mm'} = \left[D^j\left(\phi\hat{\mathbf{z}}\right) D^j(\theta\hat{\mathbf{y}}) D^j(\gamma\hat{\mathbf{z}})\right]_{mm'} \ , \tag{7.56}$$

which corresponds merely to taking the matrix elements and inserting identity in between factors. The matrix elements of this relation give

$$\begin{aligned} D^j\left(\boldsymbol{\omega}\right)_{mm'} &= \langle jm| e^{i\phi J_3/\hbar} e^{i\theta J_2/\hbar} e^{i\gamma J_3/\hbar} |jm'\rangle \\ &= e^{i\phi m} \langle jm| e^{i\theta J_2/\hbar} |jm'\rangle e^{i\gamma m'} \\ &= e^{i\left(\phi m + \gamma m'\right)} d^j_{mm'}(\theta) \end{aligned} \tag{7.57}$$

where we have defined the small $d-function$

$$d^j_{mm'}(\theta) \equiv \langle jm| e^{i\theta J_2/\hbar} |jm'\rangle = \left(e^{i\theta \Im^j_2}\right)_{mm'} = D^j_{mm'}(\theta\hat{\mathbf{y}}). \tag{7.58}$$

The $d-$functions are easily computed and tabulated (see below). Therefore, the Euler parametrization provides a second approach for computing $D^j\left(\boldsymbol{\omega}\right)_{mm'}$ for general j.

However, one needs to know how to find ϕ, θ, γ for a given general rotation $\boldsymbol{\omega}$. Since (7.56) is true for any j, the spin $j = 1/2$ representation provides the simplest computation. Using (7.44-7.46) we have

$$\begin{aligned} D^{j=1/2}\left(\boldsymbol{\omega}\right) &= D^{j=1/2}\left(\phi\hat{\mathbf{z}}\right) D^{j=1/2}(\theta\hat{\mathbf{y}}) D^{j=1/2}(\gamma\hat{\mathbf{z}}) \\ &= e^{i\phi\sigma_3/2} e^{i\theta\sigma_2/2} e^{i\gamma\sigma_3/2} \\ &= \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{pmatrix} \\ &= \begin{pmatrix} e^{i(\phi+\gamma)/2}\cos\frac{\theta}{2} & e^{i(\phi-\gamma)/2}\sin\frac{1}{2}\theta \\ -e^{-i(\phi-\gamma)/2}\sin\frac{1}{2}\theta & e^{-i(\phi+\gamma)/2}\cos\frac{\theta}{2} \end{pmatrix} = \\ e^{i\boldsymbol{\omega}\cdot\frac{\boldsymbol{\sigma}}{2}} &= \begin{pmatrix} \cos\frac{\omega}{2} + i\hat{\boldsymbol{\omega}}_z \sin\frac{\omega}{2} & \left(\hat{\boldsymbol{\omega}}_y + i\hat{\boldsymbol{\omega}}_x\right)\sin\frac{\omega}{2} \\ -\left(\hat{\boldsymbol{\omega}}_y - i\hat{\boldsymbol{\omega}}_x\right)\sin\frac{\omega}{2} & \cos\frac{\omega}{2} - i\hat{\boldsymbol{\omega}}_z \sin\frac{\omega}{2} \end{pmatrix} \end{aligned} \tag{7.59}$$

where the last line is $D^{j=1/2}\left(\boldsymbol{\omega}\right) = e^{i\boldsymbol{\omega}\cdot\frac{\boldsymbol{\sigma}}{2}}$ as given in (7.45). By comparing the two matrix expressions for $D^{j=1/2}\left(\boldsymbol{\omega}\right)$ we can relate ϕ, θ, γ to $\boldsymbol{\omega}$ and vice versa.

7.4.5 Differential equation for $d^j_{mm'}$

To derive a differential equation for $d^j_{mm'}(\theta)$ we use the standard quantum mechanics trick of evaluating a matrix element of an operator in two different

ways. In the present case we consider the Casimir operator $\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2$ as follows

$$\begin{array}{l} j(j+1)\, d^j_{mm'}(\theta) = \langle jm|e^{i\theta J_2/\hbar}\frac{1}{\hbar^2}\left(J_3^2+J_2^2+J_1^2\right)|jm'\rangle \\ = (m')^2\ d^j_{mm'}(\theta) - \frac{d^2}{d\theta^2}\, d^j_{mm'}(\theta) + \langle jm|e^{i\theta J_2/\hbar}\frac{1}{\hbar^2}J_1^2|jm'\rangle \end{array} \tag{7.60}$$

where we have used $-\hbar^2\partial_\theta^2\langle jm|e^{i\theta J_2/\hbar}|jm'\rangle = \langle jm|e^{i\theta J_2/\hbar}J_2^2|jm'\rangle$. To evaluate the last term we rewrite $e^{i\theta J_2/\hbar}J_1^2$ by using the following rotation $e^{-\frac{i}{\hbar}\theta J_2}J_3 e^{\frac{i}{\hbar}\theta J_2} = J_3\cos\theta + J_1\sin\theta$. Multiply this relation from the left with $e^{\frac{i}{\hbar}\theta J_2}$ and solve for the combination $e^{\frac{i}{\hbar}\theta J_2}J_1$

$$e^{\frac{i}{\hbar}\theta J_2}J_1 = \frac{1}{\sin\theta}J_3 e^{\frac{i}{\hbar}\theta J_2} - \frac{\cos\theta}{\sin\theta}e^{\frac{i}{\hbar}\theta J_2}J_3. \tag{7.61}$$

Multiply this equation from the right with J_1

$$e^{\frac{i}{\hbar}\theta J_2}J_1^2 = \frac{1}{\sin\theta}J_3 e^{\frac{i}{\hbar}\theta J_2}J_1 - \frac{\cos\theta}{\sin\theta}e^{\frac{i}{\hbar}\theta J_2}J_3 J_1, \tag{7.62}$$

use $J_3J_1 = J_1J_3 + i\hbar J_2$ to rearrange the second term, and then replace $e^{\frac{i}{\hbar}\theta J_2}J_1$ with the right hand side of (7.61):

$$\begin{array}{l} e^{\frac{i}{\hbar}\theta J_2}J_1^2 = \frac{1}{\sin\theta}J_3\left[e^{\frac{i}{\hbar}\theta J_2}J_1\right] - \frac{\cos\theta}{\sin\theta}\left[e^{\frac{i}{\hbar}\theta J_2}J_1\right]J_3 \\ \qquad -\frac{\cos\theta}{\sin\theta}e^{\frac{i}{\hbar}\theta J_2}\ i\hbar J_2 \\ = \frac{1}{\sin^2\theta}J_3^2 e^{\frac{i}{\hbar}\theta J_2} + \frac{\cos^2\theta}{\sin^2\theta}e^{\frac{i}{\hbar}\theta J_2}J_3^2 - 2\frac{\cos\theta}{\sin^2\theta}J_3 e^{\frac{i}{\hbar}\theta J_2}J_3 \\ \quad -\hbar^2\frac{\cos\theta}{\sin\theta}\frac{d}{d\theta}\ e^{\frac{i}{\hbar}\theta J_2}. \end{array} \tag{7.63}$$

The matrix elements of this equation gives the last term in (7.60)

$$\begin{array}{l} \langle jm|e^{i\theta J_2/\hbar}\frac{1}{\hbar^2}J_1^2|jm'\rangle \\ = \left(\frac{1}{\sin^2\theta}m^2 + \frac{\cos^2\theta}{\sin^2\theta}\,(m')^2 - 2mm'\frac{\cos\theta}{\sin^2\theta} - \frac{\cos\theta}{\sin\theta}\partial_\theta\right) d^j_{mm'}(\theta). \end{array} \tag{7.64}$$

Inserting this expression in (7.60) we obtain the desired differential equation

$$\left\{\frac{d^2}{d\theta^2} + \cot\theta\frac{d}{d\theta} + j(j+1) - \frac{m^2 - 2mm'\cos\theta + m'^2}{\sin^2\theta}\right\} d^j_{mm'}(\theta) = 0. \tag{7.65}$$

This is converted to a hypergeometric differential equation in the variable $-(\cot\frac{\theta}{2})^2$. It can be solved by using series methods or by computing the indices for the hypergeometric function ${}_2F_1(a,b;c;-(\cot\frac{\theta}{2})^2)$ as follows

$$\begin{array}{c} d^j_{mm'}(\theta) = N_{jmm'}\ \left(\sin\frac{\theta}{2}\right)^{2j}\left(\cot\frac{\theta}{2}\right)^{m+m'} \\ \times\, {}_2F_1(m-j, m'-j, m+m'+1; -\cot^2\frac{\theta}{2}) \\ N_{jmm'} = \frac{(-1)^{j+m'}}{(m+m')!}\left[\frac{(j+m)!\ (j+m')!}{(j-m)!\ (j-m')!}\right]^{1/2}. \end{array} \tag{7.66}$$

It can also be given in the form of a polynomial valid for all allowed values of j, m, m'

$$d^j_{mm'}(\theta) = (-1)^{j+m'}\,\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!} \times(\sin\tfrac{\theta}{2})^{2j}\sum_k \frac{(-1)^k(-\cot\frac{\theta}{2})^{2k+m+m'}}{k!(k+m+m')!(j-k-m)!(j-k-m')!}, \tag{7.67}$$

The constant factor $N_{jmm'}$ insures consistent normalization with the states $|jm\rangle$. It is determined by demanding $d^j_{mm'}(0) = \delta_{mm'}$ as well as completeness in the form

$$\begin{aligned}&\textstyle\sum_{m'} d^j_{mm'}(\theta)\left(d^j_{m''m'}(\theta)\right)^* \\ &= \textstyle\sum_{m'}\langle jm|e^{iJ_2\theta/\hbar}|jm'\rangle\langle jm'|e^{-iJ_2\theta/\hbar}|jm''\rangle \\ &= \textstyle\sum_{m'}\langle jm|jm''\rangle = \delta_{mm''}.\end{aligned} \tag{7.68}$$

It is useful to list a few $d^j_{mm'}(\theta)$ explicitly

$$\begin{array}{lll}
d^{1/2}_{1/2,1/2}=\cos\tfrac{\theta}{2} & d^{3/2}_{3/2,3/2}=\tfrac{1}{2}(1+\cos\theta)\cos\tfrac{\theta}{2} & d^2_{2,2}=\tfrac{1}{4}(1+\cos\theta)^2 \\
d^{1/2}_{1/2,1/2}=\sin\tfrac{\theta}{2} & d^{3/2}_{3/2,1/2}=\tfrac{\sqrt{3}}{2}(1+\cos\theta)\sin\tfrac{\theta}{2} & d^2_{2,1}=\tfrac{1}{2}(1+\cos\theta)\sin\theta \\
 & d^{3/2}_{3/2,-1/2}=\tfrac{\sqrt{3}}{2}(1-\cos\theta)\cos\tfrac{\theta}{2} & d^2_{2,0}=\tfrac{\sqrt{6}}{4}\sin^2\theta \\
d^1_{1,1}=\tfrac{1}{2}(1+\cos\theta) & d^{3/2}_{3/2,-3/2}=\tfrac{1}{2}(1-\cos\theta)\sin\tfrac{\theta}{2} & d^2_{2,-1}=\tfrac{1}{2}(1-\cos\theta)\sin\theta \\
d^1_{1,0}=\tfrac{1}{\sqrt{2}}\sin\theta & d^{3/2}_{1/2,1/2}=\tfrac{1}{2}(3\cos\theta-1)\cos\tfrac{\theta}{2} & d^2_{2,-2}=\tfrac{1}{4}(1-\cos\theta)^2 \\
d^1_{1,-1}=\tfrac{1}{2}(1-\cos\theta) & d^{3/2}_{1/2,-1/2}=\tfrac{1}{2}(3\cos\theta+1)\sin\tfrac{\theta}{2} & d^2_{1,1}=\tfrac{1}{2}(1+\cos\theta)(2\cos\theta-1) \\
d^1_{0,0}=\cos\theta & & d^2_{1,0}=\sqrt{\tfrac{3}{2}}\sin\theta\cos\theta \\
 & & d^2_{1,-1}=\tfrac{1}{2}(1-\cos\theta)(2\cos\theta+1) \\
 & & d^2_{0,0}=\tfrac{3}{2}\cos^2\theta-\tfrac{1}{2}
\end{array} \tag{7.69}$$

To obtain the solution for other values of m, m' one can use the following identities that follow from the general polynomial expression given above

$$\begin{aligned}d^j_{mm'}(\theta) &= (-1)^{m-m'}\,d^j_{m'm}(\theta) = d^j_{-m,-m'}(\theta) \\ d^j_{mm'}(\theta) &= d^j_{m'm}(-\theta) = (-1)^{m-m'}\,d^j_{mm'}(-\theta).\end{aligned} \tag{7.70}$$

7.4.6 2D harmonic oscillator and SU(2) rotations D^j

It is beneficial to take advantage of the $SU(2)$ symmetry of the two dimensional harmonic oscillator in order to obtain information on the rotation group. Of course, the $SU(2)$ symmetry has nothing to do with three dimensional rotations, but the mathematics is the same, so the 2D harmonic oscillator may be used as a model for computations[1]. Recall that the Hamiltonian $H =$

[1] The same approach applied to the $d-$dimensional harmonic oscillator is beneficial for $SU(d)$. However, only a subset of the representations for $SU(d)$ can be described by the $d-$dimensional harmonic oscillator. This is because only the completely *symmetric tensors* of the fundamental oscillators $a^\dagger_I$ can occur. For $d = 2$ the fundamental oscillators have spin 1/2, and all higher spins can be obtained only from the complete symmetrization of spin 1/2 wavefunctions. So, for $d = 2$ there are no other $SU(2)$ representations.

$\hbar\omega\left(a_1^\dagger a_1 + a_2^\dagger a_2 + 1\right)$ is invariant under the $SU(2)$ transformations $a_i' = U_{ij}a_j$, and $a_i'^\dagger = a_k^\dagger \left(U^\dagger\right)_{ki}$ and that the states are classified in degenerate $SU(2)$ multiplets

$$|jm\rangle = \frac{(a_1^\dagger)^{j+m}(a_2^\dagger)^{j-m}}{\sqrt{(j+m)!(j-m)!}}|0\rangle. \tag{7.71}$$

The generators of the symmetry that commute with the Hamiltonian, and which satisfy the commutation rules of angular momentum, are $J_+ = a_1^\dagger a_2$, $J_- = a_2^\dagger a_1$, $J_0 = (a_1^\dagger a_1 - a_2^\dagger a_2)/2$. These may also be rewritten in terms of Pauli matrices in the convenient form

$$\begin{aligned} J_I &= \hbar a_k^\dagger \left(\frac{\sigma_I}{2}\right)_{kl} a_l, \\ [J_I, J_J] &= \hbar^2 a_k^\dagger \left[\frac{\sigma_I}{2}, \frac{\sigma_I}{2}\right]_{kl} a_l \\ &= i\hbar\,\varepsilon_{IJK}\,\hbar a_k^\dagger \left(\frac{\sigma_K}{2}\right)_{kl} a_l \\ &= i\hbar\,\varepsilon_{IJK}\,J_K. \end{aligned} \tag{7.72}$$

Under the transformation $\exp\left(i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar\right)$ the states rotate as expected

$$\langle jm|' = \langle jm|e^{i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar} = \sum_{m'} D^j_{mm'}(\boldsymbol{\omega})\,\langle jm'|\,. \tag{7.73}$$

The aim is to use the properties of the creation-annihilation operators to extract the properties of these rotations.

First note that the SU(2) transformation of the oscillators is written in operator form as follows

$$\begin{aligned} a_i' &= U_{ij}a_j \\ &= e^{-i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar} a_i e^{i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar} \\ &= \left(e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\omega}/2}\right)_{ij} a_j. \\ \therefore\ U &= e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\omega}/2} = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}. \end{aligned} \tag{7.74}$$

This may be verified easily for infinitesimal values of $\boldsymbol{\omega}$, or for any $\boldsymbol{\omega}$ by using the identity $e^A B e^{-A} = \sum \frac{1}{n!}\left[A, [A, \cdots, [A, B]\cdots]\right]$. Hence, the oscillators transform as the spin $j = 1/2$ doublet of $SU(2)$

$$a_1' = \alpha a_1 + \beta a_2, \quad a_2' = -\beta^* a_1 + \alpha^* a_2, \tag{7.75}$$

where the definitions of the parameters α, β in terms of rotation angles coincide with those given in (7.44,7.45,7.46). Then the harmonic oscillator states $\langle jm|$ with spin j may be viewed as if they have been constructed by putting together $2j$ spin 1/2 quanta. The $SU(2)$ transformation of the state may now be rewritten

in the form

$$\begin{aligned}
&\textstyle\sum_{m'} D^j_{mm'}(\boldsymbol{\omega})\,\langle jm'| = \langle jm|e^{i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar} \\
&= \langle 0|e^{i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar}e^{-i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar}\frac{(a_1)^{j+m}(a_2)^{j-m}}{\sqrt{(j+m)!(j-m)!}}e^{i\mathbf{J}\cdot\boldsymbol{\omega}/\hbar} \\
&= \langle 0|\frac{(a_1')^{j+m}(a_2')^{j-m}}{\sqrt{(j+m)!(j-m)!}} \\
&= \langle 0|\frac{(\alpha a_1+\beta a_2)^{j+m}(-\beta^* a_1+\alpha^* a_2)^{j-m}}{\sqrt{(j+m)!(j-m)!}}.
\end{aligned} \tag{7.76}$$

It is now a simple matter to expand the binomials and identify the normalized states $\langle jm'|$ on the right hand side of the expansion by collecting powers of the oscillators. Then comparing the coefficients on both sides of the equation we can write

$$\begin{aligned}
D^j_{mm'}(\boldsymbol{\omega}) \;&= \sqrt{\tfrac{(j+m')!\,(j-m')!}{(j+m)!\,(j-m)!}} \times \textstyle\sum_k \binom{j+m}{k}\binom{j-m}{j+m'-k} \\
&\quad \times\alpha^k(\alpha^*)^{k-m-m'}\beta^{j+m-k}(-\beta^*)^{j+m'-k}.
\end{aligned} \tag{7.77}$$

Setting now $\alpha = \cos\frac{\theta}{2}e^{i\frac{\phi+\gamma}{2}}$ and $\beta = \sin\frac{\theta}{2}e^{i\frac{\phi-\gamma}{2}}$ we find

$$\begin{aligned}
D^j_{mm'}(\phi\theta\gamma) \;&= e^{i(\phi m+\gamma m')}\sqrt{\tfrac{(j+m')!\,(j-m')!}{(j+m)!\,(j-m)!}} \times \textstyle\sum_k \binom{j+m}{k}\binom{j-m}{j+m'-k} \\
&\quad \times (\cos\tfrac{\theta}{2})^{2k-m-m'}\,(\sin\tfrac{\theta}{2})^{2j+m+m'-2k}\;(-1)^{j+m'-k}
\end{aligned} \tag{7.78}$$

The sum extends over the values of k for which the binomial coefficients make sense. For example, for $j = 1/2,\ m = m' = 1/2$ we get

$$\begin{aligned}
D^{1/2}_{1/2,1/2}(\phi\theta\gamma) &= \sqrt{\frac{1!0!}{1!0!}}e^{\frac{i}{2}(\phi+\gamma)}\sum_k \binom{1}{k}\binom{0}{1-k} \\
&\quad \times (\sin\frac{\theta}{2})^{2-2k}(\cos\frac{\theta}{2})^{2k-1}\,(-1)^{1-k} \\
&= e^{\frac{i}{2}(\phi+\gamma)}\cos\frac{\theta}{2} = \alpha,
\end{aligned} \tag{7.79}$$

which is the expected result according to (7.44). By shifting $k \to k + m + m'$ one can show that the sum in (7.78) reproduces the polynomial in (7.67) identically. The student should verify that this general result agrees with the other approaches given above.

7.5 Addition of angular momentum

In many physical applications one needs to consider combining the angular momenta of two or more parts of the system. In classical physics this is done by the adding the vectors that represent the angular momenta of the various parts. For example the total angular momentum of the Earth orbiting around the sun is obtained by adding the spin of the Earth around itself to the orbital angular momentum for rotating around the sun. How is this done in quantum mechanics? We must answer this question in order to understand a host of problems in

quantum systems such as atoms, molecules, solids, nuclei, particles composed of quarks, scattering, etc., where the spins of particles must be combined with their orbital angular momenta, and the total angular momentum of the system is obtained by adding all the spins and all the orbital angular momenta.

To understand the process of addition of angular momentum it may be helpful to first consider the addition of ordinary linear momentum. Thus consider two particles (or two parts of a system) that have momenta $\mathbf{p}_1$ and $\mathbf{p}_2$. In classical physics the total momentum of the system is $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$. In quantum mechanics each particle is described by a state $|\mathbf{k}_1\rangle, |\mathbf{k}_2\rangle$ labelled with the eigenvalues of the commuting operators $\mathbf{p}_1 \to \hbar\mathbf{k}_1,\ \ \mathbf{p}_2 \to \hbar\mathbf{k}_2$. The total system is described by the direct product state

$$|\mathbf{k}_1\rangle \otimes |\mathbf{k}_2\rangle \equiv\ |\mathbf{k}_1, \mathbf{k}_2\rangle. \tag{7.80}$$

When a function of the operators $f(\mathbf{p}_1, \mathbf{p}_2)$ acts on the direct product state, each operator acts on its corresponding label, leaving the other label untouched. In particular, the total momentum operator $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$ may act on this state, and pick up an overall eigenvalue $\mathbf{p} \to \hbar\mathbf{k}_1 + \hbar\mathbf{k}_2 = \hbar\mathbf{k}$. Thus, the state $|\mathbf{k}_1, \mathbf{k}_2\rangle$ is already an eigenstate of the total momentum, and therefore, it is possible to relabel it in terms of total angular momentum, plus other labels corresponding to the eigenvalues of operators that commute with the total momentum operator $\mathbf{p}$ (e.g. relative angular momentum, if this is convenient for the application)

$$|\mathbf{k}_1\rangle \otimes |\mathbf{k}_2\rangle \equiv\ |\mathbf{k}_1, \mathbf{k}_2\rangle = |\mathbf{k}, \cdots\rangle. \tag{7.81}$$

Addition of angular momentum is conceptually similar to the above process, except for the last step, because unlike linear momentum, angular momentum operators in different directions do not commute, and the labelling of the states involves the squares of operators. Consequently, the direct product states are not eigenstates of the total angular momentum, but instead they can be expanded as a linear combination of total angular momentum states. The expansion coefficients are called the Clebsch-Gordan coefficients, as will be explained below.

7.5.1 Total angular momentum

Consider two rotating systems with angular momentum operators $\mathbf{J}^{(1)}$ and $\mathbf{J}^{(2)}$ respectively. Some examples are: the orbital angular momentum of an electron $\mathbf{J}^{(1)} = \mathbf{L}$ and its spin $\mathbf{J}^{(2)} = \mathbf{S}$, the spins of two electrons in a multi-electron atom $\mathbf{J}^{(1)} = \mathbf{S}^{(1)},\ \mathbf{J}^{(2)} = \mathbf{S}^{(2)}$, etc. The commutation rules are

$$\begin{aligned}
[J_i^{(1)}, J_j^{(1)}] &= i\hbar\varepsilon_{ijk} J_k^{(1)} \\
[J_i^{(2)}, J_j^{(2)}] &= i\hbar\varepsilon_{ijk} J_k^{(2)} \\
[J_i^{(1)}, J_j^{(2)}] &= 0.
\end{aligned} \tag{7.82}$$

Each system is described by a state $|j_1 m_1,\rangle$, $|j_2 m_2\rangle$, while the combined system has the direct product state

$$|j_1 m_1\rangle \otimes |j_2 m_2\rangle \equiv |j_1 m_1 j_2 m_2\rangle. \tag{7.83}$$

We will study how to express this state in terms of total angular momentum states, where the definition of total angular momentum is consistent with classical mechanics $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$.

How does the state $|j_1 m_1 j_2 m_2\rangle$ rotate? This must be obtained from the rotation of each part, thus

$$\begin{aligned} |j_1 m_1\rangle' \otimes |j_2 m_2\rangle' &= e^{-\frac{i}{\hbar}\mathbf{J}^{(1)}\cdot\boldsymbol{\omega}} |j_1 m_1\rangle \otimes e^{-\frac{i}{\hbar}\mathbf{J}^{(2)}\cdot\boldsymbol{\omega}} |j_2 m_2\rangle \\ &\equiv (e^{-\frac{i}{\hbar}\mathbf{J}^{(1)}\cdot\boldsymbol{\omega}}\ e^{-\frac{i}{\hbar}\mathbf{J}^{(2)}\cdot\boldsymbol{\omega}})\, |j_1 m_1, j_2 m_2\rangle \end{aligned} \tag{7.84}$$

where in the second line each operator $\mathbf{J}^{(1)}, \mathbf{J}^{(2)}$ is understood to act on the corresponding labels (1 or 2), leaving the others untouched. Note that the same rotation angles $\boldsymbol{\omega}$ must appear in each exponent since the same rotation is applied on the entire system. The exponents may be combined because $[\mathbf{J}^{(1)}, \mathbf{J}^{(2)}] = 0$,

$$(e^{-\frac{i}{\hbar}\mathbf{J}^{(1)}\cdot\boldsymbol{\omega}}\ e^{-\frac{i}{\hbar}\mathbf{J}^{(2)}\cdot\boldsymbol{\omega}}) = e^{-\frac{i}{\hbar}(\mathbf{J}^{(1)}+\mathbf{J}^{(2)})\cdot\boldsymbol{\omega}}. \tag{7.85}$$

So, we see that a rotation on the total system is performed by the total angular momentum $(\mathbf{J}^{(1)} + \mathbf{J}^{(2)}) = \mathbf{J}$. This is consistent with the concept that the angular momentum operators for the total system should be the generators of rotations on the total system. This is verified by showing that the total $\mathbf{J}$ has the commutation rules of angular momentum, as follows

$$\begin{aligned} [J_i, J_j] &= [J_i^{(1)} + J_j^{(2)}, J_i^{(1)} + J_j^{(2)}] \\ &= [J_i^{(1)}, J_j^{(1)}] + [J_i^{(2)}, J_j^{(2)}] + 0 \\ &= i\hbar\varepsilon_{ijk}(J_k^{(1)} + J_k^{(2)}) \\ &= i\hbar\varepsilon_{ijk}\, J_k. \end{aligned} \tag{7.86}$$

Furthermore, the total angular momentum rotates $\mathbf{J}^{(1)}, \mathbf{J}^{(2)}$ as vectors,

$$[J_i, J_j^{(1)}] = i\hbar\varepsilon_{ijk} J_k^{(1)}, \quad [J_i, J_j^{(2)}] = i\hbar\varepsilon_{ijk} J_k^{(2)}. \tag{7.87}$$

Since $(\mathbf{J}^{(1)})^2, (\mathbf{J}^{(2)})^2$ are scalars under rotations they must commute with the total $\mathbf{J}$. Hence, the operators $\mathbf{J}^2, J_3, (\mathbf{J}^{(1)})^2, (\mathbf{J}^{(2)})^2$ are mutually commuting operators, and one can choose a basis $|jmj_1j_2\rangle$ labelled by their eigenvalues

$$\begin{aligned} \mathbf{J}^2\ |jmj_1j_2\rangle &= \hbar^2 j(j+1)\ |jmj_1j_2\rangle \\ J_3\ |jmj_1j_2\rangle &= \hbar m\ |jmj_1j_2\rangle \\ (\mathbf{J}^{(1)})^2\ |jmj_1j_2\rangle &= \hbar^2 j_1(j_1+1)\ |jmj_1j_2\rangle \\ (\mathbf{J}^{(2)})^2\ |jmj_1j_2\rangle &= \hbar^2 j_2(j_2+1)\ |jmj_1j_2\rangle. \end{aligned} \tag{7.88}$$

The complete set of states $|jmj_1j_2\rangle$ span the same vector space as $|j_1m_1j_2m_2\rangle$, and therefore it should be possible to express one basis in terms of the other. The expansion is obtained by using the completeness relation

$$\sum_{jm} |jmj_1j_2\rangle\langle jmj_1j_2| = \mathbf{1}_{j_1j_2} = \sum_{m_1m_2} |j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2| \tag{7.89}$$

The symbol $\mathbf{1}_{j_1j_2}$ is the identity operator within the subspace of fixed values of j_1, j_2 (it really is a projection operator onto the subspace of fixed values of j_1, j_2; the sum over j_1,j_2 gives the full identity).

$$\begin{aligned} |j_1m_1j_2m_2\rangle &= \sum_{jm} |jmj_1j_2\rangle\langle jmj_1j_2|j_1m_1j_2m_2\rangle \\ |jmj_1j_2\rangle &= \sum_{m_1m_2} |j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2|jmj_1j_2\rangle. \end{aligned} \tag{7.90}$$

The allowed values of j for given j_1, j_2 must be specified in the sum $\sum_{jm}$. One constraint on j is that the total number of states $|jmj_1j_2\rangle$ must match the total number of states $|j_1m_1j_2m_2\rangle$. The numbers of allowed values of m, m_1, m_2 are $(2j+1)$, $(2j_1+1)$, $(2j_2+1)$ respectively. The total number of states $N_{j_1j_2}$ in the subspace $\mathbf{1}_{j_1j_2}$

$$N_{j_1j_2} = \sum_{j=j_{\min}}^{j_{\max}} (2j+1) = (2j_1+1)(2j_2+1) \tag{7.91}$$

restricts j to the range $j_{\min} \le j \le j_{\max}$. Performing the sum gives the relation

$$(j_{\max} + j_{\min} + 1)(j_{\max} - j_{\min} + 1) = (2j_1+1)(2j_2+1) \tag{7.92}$$

Another simple constraint is obtained by considering the maximum eigenvalue of total angular momentum in the z-direction $J_3 = J_3^{(1)} + J_3^{(2)}$. In the $|j_1m_1j_2m_2\rangle$ labelling there is a unique maximal eigenstate $|j_1j_1j_2j_2\rangle$ with $J_3 = \hbar(j_1+j_2)$. In the $|jmj_1j_2\rangle$ labelling there is also a unique maximal eigenstate $|j_{\max}j_{\max}j_1j_2\rangle$ with $J_3 = \hbar j_{\max}$. Since it is the unique state among the $N_{j_1j_2}$ states in either labelling, it must be the same state. Therefore

$$\begin{aligned} |j_{\max}j_{\max}j_1j_2\rangle &= |j_1j_1j_2j_2\rangle \\ j_{\max} &= m_{\max} = j_1 + j_2. \end{aligned} \tag{7.93}$$

Combining this result with (7.92) gives $j_{\min} = |j_1 - j_2|$. Therefore the range of j is

$$|j_1 - j_2| \le j \le j_1 + j_2. \tag{7.94}$$

7.5.2 Reduction to irreducible representations

The expansion coefficients $\langle jmj_1j_2|j_1m_1j_2m_2\rangle$ are called Clebsh-Gordan coefficients. Since there are an equal number of states with either labelling, these

coefficients may be viewed as the matrix elements of a $N_{j_1 j_2} \times N_{j_1 j_2}$ square unitary matrix $C^{j_1 j_2}$

$$\langle j m j_1 j_2 | j_1 m_1 j_2 m_2 \rangle \equiv \left(C^{j_1 j_2} \right)_{j m, m_1 m_2} \tag{7.95}$$

that relates one basis to the other. To simplify the notation the j_1, j_2 labels will be omitted

$$\langle j m j_1 j_2 | j_1 m_1 j_2 m_2 \rangle \equiv \langle j m | m_1 m_2 \rangle \equiv C_{j m, m_1 m_2}. \tag{7.96}$$

The completeness relations of the states correspond to a proof of the unitarity of the matrix C

$$\begin{array}{l} \sum_{m_1 m_2} \langle j m | m_1 m_2 \rangle \langle m_1 m_2 | j' m' \rangle = \delta_{j j'} \delta_{m m'}, \\ \sum_{j m} \langle m_1 m_2 | j m \rangle \langle j m | m_1' m_2' \rangle = \delta_{m_1 m_2} \delta_{m_1' m_2'}. \end{array} \tag{7.97}$$

Now consider a rotation of the state $\langle j_1 m_1 j_2 m_2 |$

$$\begin{array}{rl} \langle j_1 m_1 j_2 m_2 |' & = \langle j_1 m_1 j_2 m_2 |\, e^{i \boldsymbol{\omega} \cdot \mathbf{J} / \hbar} \\ & = \sum_{m_1' m_2'} \langle j_1 m_1 j_2 m_2 |\, e^{i \boldsymbol{\omega} \cdot \mathbf{J} / \hbar} | j_1 m_1' j_2 m_2' \rangle \langle j_1 m_1' j_2 m_2' | \\ & \equiv \sum_{m_1' m_2'} D^{j_1 j_2}_{m_1 m_2; m_1' m_2'}(\boldsymbol{\omega}) \langle j_1 m_1' j_2 m_2' | \\ D^{j_1 j_2}_{m_1 m_2; m_1' m_2'}(\boldsymbol{\omega}) & = \langle j_1 m_1 |\, e^{i \boldsymbol{\omega} \cdot \mathbf{J}^{(1)} / \hbar} | j_1 m_1' \rangle \langle j_2 m_2 |\, e^{i \boldsymbol{\omega} \cdot \mathbf{J}^{(2)} / \hbar} | j_2 m_2' \rangle \\ & = D^{j_1}_{m_1 m_1'}(\boldsymbol{\omega})\ D^{j_2}_{m_2 m_2'}(\boldsymbol{\omega}) \end{array} \tag{7.98}$$

If the state is expanded in the total angular momentum basis, the same rotation is written as

$$\begin{array}{l} \langle j_1 m_1 j_2 m_2 |' = \langle j_1 m_1 j_2 m_2 |\, e^{i \boldsymbol{\omega} \cdot \mathbf{J} / \hbar} \\ \quad = \sum_{j m} \langle m_1 m_2 | j m \rangle \langle j_1 j_2 j m | e^{i \boldsymbol{\omega} \cdot \mathbf{J} / \hbar} \\ \quad = \sum_{j m m'} \langle m_1 m_2 | j m \rangle D^{j}_{m m'}(\boldsymbol{\omega}) \langle j_1 j_2 j m' | \\ \quad = \sum_{m_1' m_2'} \left(\sum_{j m m'} \langle m_1 m_2 | j m \rangle D^{j}_{m m'}(\boldsymbol{\omega}) \langle j m' | m_1' m_2' \rangle \right) \\ \qquad\qquad \times \langle j_1 m_1' j_2 m_2' | \\ \quad = \sum_{m_1' m_2'} D^{j_1 j_2}_{m_1 m_2; m_1' m_2'}(\boldsymbol{\omega}) \langle j_1 m_1' j_2 m_2' | \end{array} \tag{7.99}$$

Comparing the two forms of the rotation matrix $D^{j_1 j_2}_{m_1 m_2; m_1' m_2'}(\boldsymbol{\omega})$ gives

$$D^{j_1}_{m_1 m_1'}(\boldsymbol{\omega})\ D^{j_2}_{m_2 m_2'}(\boldsymbol{\omega}) = \sum_{j m m'} (C^{\dagger})_{m_1 m_2, j m} D^{j}_{m m'}(\boldsymbol{\omega}) C_{j m', m_1' m_2'} \tag{7.100}$$

This shows that the Clebsch-Gordan matrix C performs a unitary transformation on the $N_{j_1 j_2} \times N_{j_1 j_2}$ rotation matrix $D^{j_1 j_2}_{m_1 m_2; m_1' m_2'}(\boldsymbol{\omega})$ in order to reduce it to a block diagonal form, where the various blocks are the $(2j+1) \times (2j+1)$ matrices $D^{j}_{m m'}(\boldsymbol{\omega})$ corresponding to the irreducible representations of rotations. This reduction is written symbolically in the forms

$$\begin{array}{c} D^{j_1} \otimes D^{j_2} = \displaystyle\sum_{j=|j_1 - j_2|}^{j_1 + j_2} \oplus D^{j} \\ j_1 \otimes j_2 = |j_1 - j_2| \oplus \cdots \oplus (j_1 + j_2) \end{array} \tag{7.101}$$

7.5.3 Computation of the Clebsch-Gordan coefficients

It is useful to consider a graphical picture of the two types of labelling of the states. In Fig.(7.2) each cross in the picture on the left corresponds to the states $|m_1 m_2\rangle$ for $j_1 = 3/2$ and $j_2 = 5/2$. Similarly, each cross in the picture on the right corresponds to the states $|jm\rangle$ with $j_{\min} = 5/2 - 3/2 = 1$ and $j_{\max} = 5/2 + 3/2 = 4$. Altogether there are $(2 \times \frac{3}{2} + 1) \times (2 \times \frac{5}{2} + 1) = 24$ states in this example. Similar pictures may be considered for other values of j_1, j_2. Our discussion will be for general j_1, j_2 but we will refer to the figure as an example.

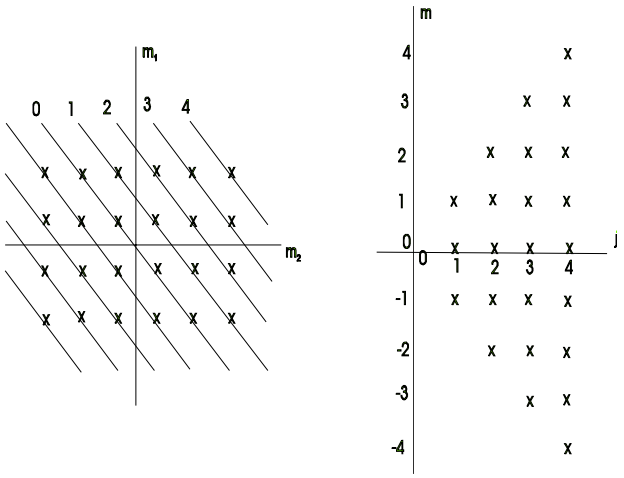


Fig.(7.2) - Lattices of angular momentum states.

A simple constraint follows from the fact that $J_3 = J_3^{(1)} + J_3^{(2)}$ is diagonal on both sets of states. So, when this operator is sandwiched between the two kinds of states it gives a relation

$$\langle jmj_1j_2|J_3|j_1m_1j_2m_2\rangle \begin{cases} = \hbar m\langle jm|m_1m_2\rangle \\ = \hbar(m_1+m_2)\langle jm|m_1m_2\rangle \end{cases}$$

which implies

$$\langle jm|m_1m_2\rangle = \delta_{m,m_1+m_2}\langle j, m_1+m_2|m_1m_2\rangle. \tag{7.102}$$

The total value of $m_1 + m_2 = m$ is shown in the picture on the left on the diagonal lines $m = 4, 3, 2, 1, 0, \cdots, -4$. Similarly, the total value of m is on the vertical axis in the picture on the right. By comparing the two sets at fixed values of m we see that we have the same number of states level by level, as listed below for the example in the figure. The plan is to compute the relation

between these states level by level.

$$\begin{array}{lll} \underline{m} & \underline{|m_1 m_2>} & \underline{|jm>} \\ 4 & |\tfrac{3}{2}\tfrac{5}{2}\rangle & |4,4\rangle \\ 3 & |\tfrac{3}{2}\tfrac{3}{2}\rangle, |\tfrac{1}{2}\tfrac{5}{2}\rangle & |4,3\rangle, |3,3\rangle \\ 2 & |\tfrac{3}{2}\tfrac{1}{2}\rangle, |\tfrac{1}{2}\tfrac{3}{2}\rangle, |\tfrac{-1}{2}\tfrac{5}{2}\rangle & |4,2\rangle, |3,2\rangle, |2,2\rangle \\ 1 & |\tfrac{3}{2}\tfrac{-1}{2}\rangle, |\tfrac{1}{2}\tfrac{1}{2}\rangle, |\tfrac{-1}{2}\tfrac{3}{2}\rangle, |\tfrac{-3}{2}\tfrac{5}{2}\rangle & |4,1\rangle, |3,1\rangle, |2,1\rangle, |1,1\rangle \\ \vdots & \vdots & \vdots \\ -4 & |\tfrac{-3}{2}\tfrac{-5}{2}\rangle & |4,-4\rangle \end{array} \tag{7.103}$$

For a given value of total j the top state with maximum $m = j$ will be called the *highest state*. Thus, for $j_{\min} \leq j \leq j_{\max}$, the states $|jj\rangle$ are "highest". In the example above the states $|4,4\rangle, |3,3\rangle, |2,2\rangle, |1,1\rangle$ are "highest" as seen pictorially in the figure on the right. The remaining states are called *descendants*, and are obtained by applying the lowering operator J_- on the highest states. So, by using $J_-|jm\rangle = \sqrt{(j+m)(j-m+1)}|j,m-1\rangle$ repeatedly one obtains the descendants from the highest states

$$|jm\rangle = (J_-)^{j-m}|jj\rangle \sqrt{\frac{(j+m)!}{(2j)!(j-m)!}}. \tag{7.104}$$

If the highest state $|jj\rangle$ is expressed in terms of the $|m_1 m_2\rangle$ states, then the descendant states are computed through this equation by replacing $(J_-)^{j-m} = (J_-^{(1)} + J_-^{(2)})^{j-m}$ and applying it on the $|m_1 m_2\rangle$ states. This gives all the Clebsch-Gordan coefficients. Let's describe how the process goes.

At the highest level $m_{\max} = m_{1_{\max}} + m_{2\,\max} = j_1 + j_2 = j_{\max}$ there is a unique highest state, as seen by comparing the two figures

$$\begin{aligned} |j_1,\ j_2\rangle &= |j_{\max},\ j_{\max}\rangle \\ |3/2, 5/2\rangle &= |4,4\rangle \end{aligned} \tag{7.105}$$

So, the Clebsch-Gordan coefficient is

$$\langle j_1 j_2 | j_{\max} j_{\max}\rangle = 1. \tag{7.106}$$

In the probability interpretation, this implies that for the total system to spin at the maximum level in the z-direction, the two parts of the system must also be spinning at their maximum levels in the z-direction with 100% probability.

At the second to highest level, $m = j_1 + j_2 - 1$, there is one descendant and one highest state. In the example, $|4,3\rangle$ is the descendant, and $|3,3\rangle$ is the highest state as seen in the figure. The descendant is obtained by applying the lowering operator on $|j_{\max},\ j_{\max}\rangle$

$$\begin{aligned} |j_{\max},\ j_{\max}-1\rangle &= \sqrt{\tfrac{1}{2j_{\max}}} J_- |j_{\max},\ j_{\max}\rangle \\ &= \sqrt{\tfrac{1}{2j_{\max}}} \left(J_-^{(1)} + J_-^{(2)}\right) |j_1,\ j_2\rangle \\ &= \sqrt{\tfrac{j_1}{j_{\max}}} |j_1 - 1, j_2\rangle + \sqrt{\tfrac{j_2}{j_{\max}}} |j_1, j_2 - 1\rangle, \\ |4,3\rangle &= \sqrt{\tfrac{3}{8}} |1/2, 5/2\rangle + \sqrt{\tfrac{5}{8}} |3/2, 3/2\rangle. \end{aligned} \tag{7.107}$$

The highest state must be chosen orthogonal to this one since it has a different value of $j = j_{\max} - 1$

$$\begin{aligned} |j_{\max}-1,\ j_{\max}-1\rangle &= -\sqrt{\tfrac{j_2}{j_{\max}}}|j_1-1,j_2\rangle + \sqrt{\tfrac{j_1}{j_{\max}}}|j_1,j_2-1\rangle \\ |3,3\rangle &= -\sqrt{\tfrac{5}{8}}|1/2,5/2\rangle + \sqrt{\tfrac{3}{8}}|3/2,3/2\rangle. \end{aligned} \tag{7.108}$$

Thus, we have computed the Clebsch-Gordan coefficients

$$\begin{aligned} \langle j_1{-}1,j_2|j_{\max},j_{\max}{-}1\rangle &= \sqrt{\tfrac{j_1}{j_1+j_2}} = \langle j_1,j_2{-}1|j_{\max}{-}1,j_{\max}{-}1\rangle \\ \langle j_1,j_2{-}1|j_{\max},j_{\max}{-}1\rangle &= \sqrt{\tfrac{j_2}{j_1+j_2}} = -\langle j_1{-}1,j_2|j_{\max}{-}1,j_{\max}{-}1\rangle \end{aligned} \tag{7.109}$$

which have a probability interpretation.

At the next lower level $m = j_1 + j_2 - 2$ there are three states, but two of those are descendants, and only one of them is a highest state. In the example above the states $|4,2\rangle, |3,2\rangle$ are the descendants and $|2,2\rangle$ is the highest state, as seen in the figure. The two descendants are obtained by applying $J_- = J_-^{(1)} + J_-^{(2)}$ on the two states of level $m = j_1 + j_2 - 1$ that were constructed above, and the highest state is constructed by making it orthogonal to the descendants. The process continues until level $m = j_{\min} = |j_1 - j_2|$ is reached, and the highest state $|j_{\min}, j_{\min}\rangle$ and its descendants are constructed.

Having explained the process, we can turn to the full computation of the Clebsch-Gordan coefficients. They obey a recursion relation which is derived by sandwiching $J_\pm = J_-^{(1)} \pm J_-^{(2)}$ in $\langle m_1, m_2 \pm 1| \cdot |jm\rangle$ and evaluating it to the right or to the left

$$\langle m_1,m_2{\pm}1|J_\pm|jm\rangle \begin{cases} = \sqrt{(j\mp m)(j\pm m+1)}\langle m_1,m_2{\pm}1|j,m{\pm}1\rangle \\ = \begin{cases} \sqrt{(j_1\pm m_1)(j_1\mp m_1+1)}\langle m_1\mp 1,m_2{\pm}1|j,m\rangle \\ \quad +\sqrt{(j_2\pm m_2)(j_2\mp m_2+1)}\langle m_1,m_2|j,m\rangle \end{cases} \end{cases} \tag{7.110}$$

Comparing the two answers on the right one obtains a recursion relation. Using the J_+ equation for the highest state $m = j$, the first line vanishes and gives a simpler relation among the Clebsch-Gordan coefficients for the highest state

$$\langle m_1{-}1,m_2{+}1|j,j\rangle = -\sqrt{\tfrac{(j_2-m_2)(j_2+m_2+1)}{(j_1+m_1)(j_1-m_1+1)}}\langle m_1,m_2|j,j\rangle \tag{7.111}$$

This equation involves all the states with $m_1 + m_2 = j$ on a diagonal line in the first figure of Fig.(7.2). Starting with $m_1 = j_1$, $m_2 = j - j_1$ and applying this recursion n times one obtains all the coefficients $\langle m_1, m_2|j, j\rangle$ in terms of one unknown. This result may be written in the form

$$\langle m_1,\ m_2|j,j\rangle = \delta_{j,m_1+m_2}\ (-1)^{j_1-m_1}\ C(j,j_1,j_2)\sqrt{\tfrac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!}}. \tag{7.112}$$

The unknown coefficient $C(j, j_1, j_2)$ is proportional to $\langle j_1,\, j - j_1 | j, j\rangle$. To determine it the normalization condition is imposed

$$\langle jj|jj\rangle = \sum_{m_1,m_2}^{j_1,j_2} |\langle m_1, m_2|j, j\rangle|^2 = 1. \tag{7.113}$$

The result is

$$C(j, j_1, j_2) = \sqrt{\frac{(2j+1)!\ (j_1+j_2-j)!}{(j+j_1+j_2+1)!(j+j_1-j_2)!\ (j+j_2-j_1)!}}. \tag{7.114}$$

Note that this is consistent with the results obtained above in (7.106,7.109). Finally the descent equation (7.104) is used to obtain the general coefficient as follows

$$\begin{aligned}
\langle m_1 m_2|jm\rangle &= \sqrt{\tfrac{(j+m)!}{(2j)!(j-m)!}}\langle m_1 m_2|(J_-)^{j-m}|jj\rangle \\
&= \sqrt{\tfrac{(j+m)!}{(2j)!(j-m)!}}\langle m_1 m_2|(J_-^{(1)} + J_-^{(2)})^{j-m}|jj\rangle \\
&= \sqrt{\tfrac{(j+m)!}{(2j)!(j-m)!}}\sum_{k_1=0}^{j-m}\begin{pmatrix} j-m \\ k_1\end{pmatrix}\langle m_1 m_2|\left(J_-^{(1)}\right)^{k_1}\left(J_-^{(2)}\right)^{j-m-k_1}|jj\rangle \\
&= \sqrt{\tfrac{(j+m)!}{(2j)!(j-m)!}}\sum_{k_1,k_2=0}^{j-m}\begin{pmatrix} j-m \\ k_1\end{pmatrix}\langle m_1+k_1,\, m_2+k_2|jj\rangle \\
&\qquad \times\delta_{k_1+k_2,j-m}\sqrt{\tfrac{(j_1-m_1)!\,(j_1+m_1+k_1)!}{(j_1+m_1)!\,(j_1-m_1-k_1)!}}\sqrt{\tfrac{(j_2-m_2)!\,(j_2+m_2+k_2)!}{(j_2+m_2)!\,(j_2-m_2-k_2)!}}
\end{aligned} \tag{7.115}$$

where $\left(J_-^{(1)}\right)^{k_1}\left(J_-^{(2)}\right)^{k_2}$ have been applied to the left in the last step. Replacing the highest state coefficients $\langle m_1 + k_1,\, m_2 + k_2|jj\rangle$ from (7.112) one gets the final result explicitly

$$\begin{aligned}
\langle m_1 m_2|jm\rangle &= \delta_{m,m_1+m_2}(-1)^{j_1-m_1} \\
&\quad\times\sqrt{\tfrac{(2j+1)\ (j_1+j_2-j)!\ (j+m)!\,(j-m)!}{(j+j_1+j_2+1)!(j+j_1-j_2)!\ (j+j_2-j_1)!}} \\
&\quad\times\sqrt{\tfrac{(j_1-m_1)!\,(j_2-m_2)!}{(j_1+m_1)!\,(j_2+m_2)!}} \\
&\quad\times\sum_{k_1,k_2=0}^{j-m}\tfrac{(-1)^{k_1}\delta_{k_1+k_2,j-m}}{k_1!\ k_2!}\tfrac{(j_1+m_1+k_1)!\,(j_2+m_2+k_2)!}{(j_1-m_1-k_1)!\,(j_2-m_2-k_2)!}
\end{aligned} \tag{7.116}$$

It is useful to list the cases for general j_1 but specialized values for $j_2 = 1/2$ and 1, as in tables (7.1, 7.2). In particular table (7.1) is useful in the addition of orbital angular momentum $\mathbf{L}$ and intrinsic spin $\mathbf{S}$, when the spin is 1/2.

7.6 Wigner's 3j symbols

The addition of angular momentum $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$ may be rewritten in the form

$$\mathbf{J}^{(1)} + \mathbf{J}^{(2)} + \mathbf{J}^{(3)} = 0, \tag{7.117}$$

by renaming $\mathbf{J}^3 = -\mathbf{J}$. Therefore the expansion of $|j_1 m_1\rangle \otimes |j_2 m_2\rangle$ in terms of $|jmj_1j_2\rangle$ must be related to the direct product of three states $|j_1 m_1\rangle \otimes |j_2 m_2\rangle \otimes$

j $\downarrow$	$m_2 = \frac{1}{2}$	$m_2 = -\frac{1}{2}$
$j_1 + \frac{1}{2}$	$\sqrt{\frac{j_1+m+\frac{1}{2}}{2j_1+1}}$	$\sqrt{\frac{j_1-m+\frac{1}{2}}{2j_1+1}}$
$j_1 - \frac{1}{2}$	$-\sqrt{\frac{j_1-m+\frac{1}{2}}{2j_1+1}}$	$\sqrt{\frac{j_1+m+\frac{1}{2}}{2j_1+1}}$

Table 7.1: Clebsch-Gordan coefficients $< j_1 m_1 \frac{1}{2} m_2 | jm >$

j $\downarrow$	$m_2 = 1$ $m_1 = m-1$	$m_2 = 0$ $m = m_1$	$m_2 = -1$ $m_1 = m+1$
$j_1 + 1$	$\sqrt{\frac{(j_1+m)(j_1+m+1)}{(2j_1+1)(2j_1+2)}}$	$\sqrt{\frac{(j_1-m+1)(j_1+m+1)}{(2j_1+1)(j_1+1)}}$	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{(2j_1+1)(2j_1+2)}}$
j_1	$-\sqrt{\frac{(j_1+m)(j_1-m+1)}{2j_1(j_1+1)}}$	$\frac{m}{\sqrt{j_1(j_1+1)}}$	$\sqrt{\frac{(j_1-m)(j_1+m+1)}{2j_1(j_1+1)}}$
$j_1 - 1$	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{2j_1(2j_1+1)}}$	$-\sqrt{\frac{(j_1-m)(j_1+m)}{j_1(2j_1+1)}}$	$\sqrt{\frac{(j_1+m)(j_1+m+1)}{2j_1(2j_1+1)}}$

Table 7.2: Clebsch-Gordan coefficients $< j_1 m_1 1 m_2 | jm >$

$|j_3 m_3\rangle$, where the corresponding total angular momentum is zero. The triple product may be expanded in terms of total angular momentum states $\mathbf{J}_{Total}$ by combining the states in pairs in two steps as follows

$$\begin{aligned} |j_1 m_1\rangle \otimes |j_2 m_2\rangle \otimes |j_3 m_3\rangle &= \sum_{jm} \left(|jm j_1 j_2\rangle \otimes |j_3 m_3\rangle\right) \\ &\quad \times \langle jm | m_1 m_2\rangle \\ &= \sum_{j_T m_T} \sum_{jm} |j_T m_T; j j_3; j_1 j_2\rangle \\ &\quad \times \langle j_T m_T | m m_3\rangle \langle jm | m_1 m_2\rangle \end{aligned} \tag{7.118}$$

Now, since $\mathbf{J}_{Total} = 0$, the only possible values for j_T and m_T are $j_T = 0 = m_T$. Furthermore, the coefficient $\langle 00 | m m_3\rangle$ can be non-zero only if $j = j_3$ and $m = -m_3$, and can be obtained from (7.116)

$$\langle 00 | m m_3\rangle = \delta_{j,j_3} \delta_{m,-m_3} \ (-1)^{-j_3 - m_3} \sqrt{\frac{1}{2j_3+1}}. \tag{7.119}$$

Therefore the triple product reduces to

$$|j_1 m_1\rangle \otimes j_2 m_2\rangle \otimes |j_3 m_3\rangle = |00, j_1 j_2 j_3\rangle \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{7.120}$$

where the $3j$ symbol is defined in terms of Clebsch-Gordan coefficients as

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \equiv (-1)^{j_1 - j_2 - m_3} \sqrt{\frac{1}{2j_3+1}} \langle j_3, -m_3 | j_1 m_1 j_2 m_2\rangle. \tag{7.121}$$

An extra phase $(-1)^{j_3 + j_1 - j_2}$ has been absorbed into this definition in order to define a state $|00, j_1 j_2 j_3\rangle$ that is symmetric under permutations of the indices 1,2,3.

The $3-j$ symbols have some symmetry properties:

- The overall $3-j$ symbol should be the same (up to a minus sign) if the indices $1,2,3$ are interchanged. By using the properties of the Clebsch-Gordan coefficients one finds the following symmetry properties under cyclic and anti-cyclic permutations

$$\begin{aligned}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix}, \\ \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix},\end{aligned} \tag{7.122}$$

- Under a reflection $\mathbf{J}^{(1,2,3)} \to -\mathbf{J}^{(1,2,3)}$ the magnetic quantum numbers change sign, but the total angular momentum remains zero, therefore the $3j$ symbol can differ only up to a sign, indeed one finds

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{7.123}$$

These symmetry properties together (7.121) are useful to relate various Clebsch-Gordan coefficients. For example from the knowledge of the coefficients for $(\frac{3}{2}\otimes\frac{1}{2}\to 2)$ one can obtain the coefficients for $\left(2\otimes\frac{1}{2}\to\frac{3}{2}\right)$ and $\left(2\otimes\frac{3}{2}\to\frac{1}{2}\right)$, etc.(see problem). Many useful properties of the 3j symbols may be found in Wigner's book [**?**].

Generalizations of the $3j$ symbols are the more complex $6j$ symbols, used in atomic or nuclear physics in the study of systems composed of a larger number of particles. Their properties may be found in [**?**].

7.7 Integrals involving the $D^j_{mm'}$-functions

The D-functions satisfy some general relations under integration that follow from simple group theoretical properties. Many familiar identities of spherical harmonics, or Legendre functions follow from these more fundamental identities, since the spherical harmonics are a special case of the D-functions. Consider integrals of the form

$$\int d\mu(\phi\theta\gamma)\; D^{j_1}_{m_1 m_1'}(\phi\theta\gamma) D^{j_2}_{m_2 m_2'}(\phi\theta\gamma)\cdots D^{j_n}_{m_n m_n'}(\phi\theta\gamma). \tag{7.124}$$

These can be evaluated in terms of Clebsch-Gordan coefficients by using the properties of the rotation group, without really doing any integration, as will be shown below. Here $d\mu(\phi,\theta,\gamma)$ is the so called Haar measure for the $SU(2)$ group

$$\begin{gathered} d\mu(\phi\theta\gamma) = \sin\theta \tfrac{d\theta}{2}\tfrac{d\phi}{4\pi}\tfrac{d\gamma}{4\pi}, \\ 0\le\theta\le 2\pi, \quad 0\le\phi\le 4\pi, \quad 0\le\gamma\le 4\pi, \\ \textstyle\int d\mu(\phi\theta\gamma) = 1. \end{gathered} \tag{7.125}$$

The integration is extended over the whole range of ϕ,θ and γ, so that it gives a double covering of the sphere (recall that $SU(2)$ is a double cover of $SO(3)$, so that spin $1/2$ representations are included in the analysis).

A very important property of the Haar measure is that it is invariant under two coordinate transformations (see problem)

$$d\mu(\phi\theta\gamma) = d\mu(\phi'\theta'\gamma') = d\mu(\phi''\theta''\gamma''), \tag{7.126}$$

where $(\phi', \theta', \gamma')$, $(\phi'', \theta'', \gamma'')$ are defined by group multiplication from the left or the right

$$\begin{aligned}(D^j(\boldsymbol{\omega})D^j(\phi\theta\gamma))_{mm'} &= D^j_{mm'}(\phi'\theta'\gamma'), \\ (D^j(\phi\theta\gamma)D^j(\boldsymbol{\omega}))_{mm'} &= D^j_{mm'}(\phi''\theta''\gamma'').\end{aligned} \tag{7.127}$$

where $\boldsymbol{\omega}$ correspond to parameters of rotations. Due to the group property, the same $(\phi', \theta', \gamma')$, $(\phi'', \theta'', \gamma'')$ occur for any representation j. The expressions for $(\phi', \theta', \gamma')$, $(\phi'', \theta'', \gamma'')$ are complicated, and they can be obtained most directly by using any of the 2×2 matrix forms for the $j = 1/2$ representation given in (7.44-7.45). However, these details are not needed here, and it is sufficient to start from this invariance property to derive the following results.

Start with the integral over a single D-function, which is a $(2j+1)\times(2j+1)$ matrix $M^j_{mm'}$

$$M^j_{mm'} \equiv \int d\mu(\phi\theta\gamma)\, D^j_{mm'}(\phi\theta\gamma). \tag{7.128}$$

Multiply the matrix M^j from the left or right with $D^j(\boldsymbol{\omega})$, and then use the invariance property of the Haar measure to write

$$\begin{aligned}\left(D^j(\boldsymbol{\omega})M^j\right)_{mm'} &= \textstyle\int d\mu(\phi\theta\gamma)[D^j(\boldsymbol{\omega})D^j(\phi\theta\gamma)]_{mm'} \\ &= \textstyle\int d\mu(\phi\theta\gamma)D^j_{mm'}(\phi'\theta'\gamma') \\ &= \textstyle\int d\mu(\phi'\theta'\gamma')D^j_{mm'}(\phi'\theta'\gamma') = M^j_{mm'}\end{aligned} \tag{7.129}$$

So the constant matrix M^j must satisfy the identities

$$M^j = D^j(\boldsymbol{\omega})M^j = M^j D^j(\boldsymbol{\omega}), \tag{7.130}$$

for any rotation $\boldsymbol{\omega}$. Obviously, this is possible only if $M^j = 0$, unless $j = 0$. This gives the first result

$$\int d\mu(\phi\theta\gamma)\, D^j_{mm'}(\phi\theta\gamma) = \delta_{j,0}. \tag{7.131}$$

It is also possible to verify this result laboriously by using the explicit form of $D^j_{mm'}(\phi\theta\gamma)$ given in the previous sections.

Consider now the integral

$$\int d\mu(\phi\theta\gamma)\; D^{j_1}_{m_1m_1'}(\phi\theta\gamma)\, D^{j_2}_{m_2m_2'}(\phi\theta\gamma). \tag{7.132}$$

We have already shown in (7.100) that the direct product of D-functions can be reduced to a single one by means of the Clebsch-Gordan coefficients

$$D^{j_1}_{m_1m_1'}(\boldsymbol{\omega})\; D^{j_2}_{m_2m_2'}(\boldsymbol{\omega}) = \sum_{jmm'} \langle m_1m_2|jm\rangle D^j_{mm'}(\boldsymbol{\omega})\langle jm'|m_1'm_2'\rangle \tag{7.133}$$

combining this with (7.131) gives the result of the integral

$$\begin{aligned}
&\int d\mu(\phi\theta\gamma)\; D^{j_1}_{m_1 m_1'}(\phi\theta\gamma)\, D^{j_2}_{m_2 m_2'}(\phi\theta\gamma) \\
&= \langle m_1 m_2|00\rangle\langle 00|m_1' m_2'\rangle \\
&= \delta_{m_1+m_2,0}\,\delta_{m_1'+m_2',0}\frac{(-1)^{2j_1-m_1-m_1'}\;\delta_{j_1 j_2}}{2j_1+1}.
\end{aligned} \tag{7.134}$$

From this, and the property of the D-functions in (7.70),

$$[D^{j_1}_{m_1 m_1'}(\phi\theta\gamma)]^* = (-1)^{-m_1+m_1'} D^{j_1}_{-m_1,-m_1'}(\phi\theta\gamma), \tag{7.135}$$

it also follows that

$$\int [D^{j_1}_{m_1 m_1'}(\phi\theta\gamma)]^* D^{j_2}_{m_2 m_2'}(\phi\theta\gamma) d\mu(\phi\theta\gamma) = \frac{\delta_{j_1 j_2}\delta_{m_1 m_2}\delta_{m_1' m_2'}}{(2j_1+1)}. \tag{7.136}$$

where we have used $(-1)^{-m_1+m_1'}(-1)^{2j_1-m_1-m_1'} = (-1)^{2(j_1-m_1)} = 1$. This is a fundamental result of orthogonality.

We can continue on this path, and write

$$\begin{aligned}
&\textstyle\int d\mu(\phi\theta\gamma)\; [D^{j_3}_{m_3 m_3'}(\phi\theta\gamma)]^*\; D^{j_1}_{m_1 m_1'}(\phi\theta\gamma)\; D^{j_2}_{m_1 m_1'}(\phi\theta\gamma) \\
&= \textstyle\sum_{jmm'}\langle m_1 m_2|jm\rangle\langle jm'|m_1' m_2'\rangle \\
&\qquad \times \textstyle\int d\mu(\phi\theta\gamma)\; [D^{j_3}_{m_3 m_3'}(\phi\theta\gamma)]^* D^{j}_{mm'}(\phi\theta\gamma) \\
&= \textstyle\sum_{jmm'}\langle m_1 m_2|jm\rangle\langle jm'|m_1' m_2'\rangle \frac{\delta_{jj_3}\delta_{mm_3}\delta_{m'm_3'}}{(2j_3+1)} \\
&= \frac{\langle m_1 m_2|j_3 m_3\rangle\langle j_3 m_3'|m_1' m_2'\rangle}{(2j_3+1)}
\end{aligned} \tag{7.137}$$

This is equivalent to

$$\begin{aligned}
&\int d\mu(\phi\theta\gamma)\; D^{j_1}_{m_1 m_1'}(\phi\theta\gamma)\; D^{j_2}_{m_1 m_1'}(\phi\theta\gamma)\; D^{j_3}_{m_3 m_3'}(\phi\theta\gamma) \\
&\qquad = \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3' \end{pmatrix}
\end{aligned} \tag{7.138}$$

and so on with more factors of D-functions. Evidently, all such integrals are evaluated in terms of Clebsch-Gordan coefficients.

7.8 Tensor operators

Tensor operators are operators with definite transformation properties under rotations. For example, as we have seen before, the operators $\mathbf{r}, \mathbf{p}, \mathbf{J}$ transform like vectors under rotations

$$e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}\,\mathbf{r}_I\, e^{-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} = \left(e^{\boldsymbol{\omega}\times}\,\mathbf{r}\right)_I = \mathbf{r}_I + (\boldsymbol{\omega}\times\mathbf{r})_I + \;\cdots, \quad etc. \tag{7.139}$$

In the $(+,0,-)$ basis this equation takes the form

$$e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}\,\mathbf{r}_m\,e^{-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}=\sum_{m'}\mathbf{r}_{m'}D^{j=1}_{m'm}(\boldsymbol{\omega}) \tag{7.140}$$

The first order term in the infinitesimal expansion of this equation is

$$\left[\frac{1}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J},\mathbf{r}_m\right]=\sum_{m'}\mathbf{r}_{m'}\langle 1m'|\frac{1}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}|1m\rangle \tag{7.141}$$

which gives

$$\begin{aligned}[J_0,\mathbf{r}_m]&=\hbar m\mathbf{r}_m\\ [J_\pm,\mathbf{r}_m]&=\hbar\sqrt{2-m(m\pm1)}\mathbf{r}_{m\pm1}.\end{aligned} \tag{7.142}$$

Motivated by this form, we define *irreducible tensor operators* more generally as operators that transform irreducibly under rotations, and classify them in one to one correspondance with the states $|jm\rangle\longleftrightarrow T_{jm}$. Thus, the definition of an irreducible tensor operator is really the set of operators $T_{jm},\ -j\le m\le j$, which satisfy any one of the following equivalent statements

$$\begin{array}{ll}(1) & e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}\,T_{jm}\,e^{-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}=\sum_{m'}T_{jm'}D^{j}_{m'm}(\boldsymbol{\omega})\\ (2) & \left[\frac{1}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J},T_{jm}\right]=\sum_{m'}T_{jm'}\langle jm'|\frac{1}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}|jm\rangle\\ (3) & \begin{cases}[J_0,T_{jm}]=\hbar mT_{jm},\\ [J_\pm,T_{jm}]=\hbar\sqrt{j(j+1)-m(m\pm1)}T_{j,m\pm1}.\end{cases}\end{array} \tag{7.143}$$

In Chapter-6 we have seen (eqs.(6.83, 6.126), problem 10) that tensors with l vector indices $T_{I_1I_2\cdots I_l}$, in d-dimensions $I_k=1,2,\cdots d$, are irreducible tensors of rank l under $SO(d)$ rotations, provided they are completely symmetric and traceless with respect to any pair of indices, $g^{I_1I_2}T_{I_1I_2\cdots I_l}=0$. Similarly, any tensor that is completely antisymmetric is also irreducible. There are also tensors with mixed symmetry, that are irreducible under certain rules associated with Young tableaux, a topic that will not be discussed here.

We give some examples in d-dimensions of tensors constructed from the products of two vector operators $\mathbf{p}$ and $\mathbf{x}$

$$\begin{array}{ll}\text{scalar} & T^{l=0}=\frac{1}{2}(\mathbf{x}\cdot\mathbf{p}+\mathbf{p}\cdot\mathbf{x})\\ \text{anti-symmetric} & (T^{l=1})_{[IJ]}=x_Ip_J-x_Ip_J\qquad(=\varepsilon_{IJK}L_K\ \text{ in d=3})\\ \text{symmetric traceless} & (T^{l=2})_{(IJ)}=x_Ip_J+x_Jp_I-\frac{1}{d}(\mathbf{x}\cdot\mathbf{p}+\mathbf{p}\cdot\mathbf{x})\delta_{IJ}\\ \text{symmetric traceless} & (T^{l=3})_{(IJK)}=x_Ix_Jx_K-\frac{\mathbf{x}^2}{d+2}(\delta_{IJ}x_K+\delta_{KI}x_J+\delta_{JK}x_I)\end{array} \tag{7.144}$$

In three dimensions the independent components of all such tensors can be rewritten in the form T_{jm}. In particular, if the tensor is constructed from the same vector $\mathbf{v}$, it is necessarily a *symmetric homogeneous* polynomial of degree l. Furthermore, if the length of the vector is restricted to $\mathbf{v}^2=1$ (unit vector $\boldsymbol{\Omega}$) then the components of the symmetric traceless tensors are in one to one correspondance to the spherical harmonics $Y_{lm}(\boldsymbol{\Omega})$ in 3-dimensions.

Without the restriction to a unit vector one still has an irreducible tensor since the length of the vector is rotationaly invariant and becomes an overall factor $|\mathbf{v}|^l$ in front of the *homogeneous* polynomial. So, in 3-dimensions the independent components of a completely symmetric traceless tensor constructed from the vector $\mathbf{v}$ are

$$T_{lm}(\mathbf{v}) = |\mathbf{v}|^l \, Y_{lm}(\hat{\mathbf{v}}). \tag{7.145}$$

The $l = 1$ example of this equation is the initial vector itself $T_{1m}(\mathbf{v}) = |\mathbf{v}| \, \hat{\mathbf{v}}_m = \mathbf{v}_m$.

As seen above in (7.140), if the vector $\mathbf{v}$ is an operator, we expect the tensor $T_{lm}(\mathbf{v})$ to satisfy precisely the conditions for being a tensor *operator* with $j = l$. Indeed this is true, since the spherical harmonics for a c-number $\hat{\mathbf{v}}$ satisfy the equations

$$\begin{aligned} Y_{lm}(e^{\boldsymbol{\omega}\times}\hat{\mathbf{v}}) &= \langle \mathbf{e}^{\boldsymbol{\omega}\times}\hat{\mathbf{v}}|lm\rangle \\ &= \langle \hat{\mathbf{v}}|e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}|lm\rangle \\ &= \sum_{m'} \langle \hat{\mathbf{v}}|lm'\rangle D^l_{m'm}(\boldsymbol{\omega}) \\ &= \sum_{m'} Y_{lm'}(\hat{\mathbf{v}}) D^l_{m'm}(\boldsymbol{\omega}) \end{aligned} \tag{7.146}$$

For an operator $\mathbf{v}$, we may write

$$\begin{aligned} e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} \left(|\mathbf{v}|^l \, Y_{lm}(\hat{\mathbf{v}}) \right) e^{-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} &= |\mathbf{v}|^l \, Y_{lm}(e^{i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar} \, \hat{\mathbf{v}} \, e^{-i\boldsymbol{\omega}\cdot\mathbf{J}/\hbar}) \\ &= |\mathbf{v}|^l \, Y_{lm}(e^{\boldsymbol{\omega}\times}\hat{\mathbf{v}}) \\ &= \sum_{m'} \left(|\mathbf{v}|^l \, Y_{lm'}(\hat{\mathbf{v}}) \right) D^l_{m'm}(\boldsymbol{\omega}) \end{aligned} \tag{7.147}$$

where in the last step we have used the previous equation that applies even if $\mathbf{v}$ is an operator, as long as v_I commute with each other. Thus, the operators $T_{lm}(\mathbf{v}) = |\mathbf{v}|^l \, Y_{lm}(\hat{\mathbf{v}})$ are examples of tensor operators.

7.8.1 Combining tensor operators and states

In many computations products of operators occur. If each operator is an irreducible tensor operator, then the product is generally not an irreducible tensor operator. One may reduce the product into a sum of irreducible tensor operators in the same way a product of states is reduced to total angular momentum states

$$|j_1 m_1\rangle |j_2 m_2\rangle = \sum_{jm} |jm j_1 j_2\rangle \langle jm|m_1 m_2\rangle \tag{7.148}$$

$$T_{j_1 m_1} T_{j_2 m_2} = \sum_{jm} T_{jm} \langle jm|m_1 m_2\rangle \tag{7.149}$$

The equations may be inverted by using the completeness relation of the Clebsch-Gordan coefficients

$$|jm\rangle = \sum_{m_1 m_2} |j_1 m_1\rangle |j_2 m_2\rangle \langle m_1 m_2 | jm\rangle \tag{7.150}$$

$$T_{jm} = \sum_{m_1 m_2} T_{j_1 m_1} T_{j_2 m_2} \langle m_1 m_2 | jm\rangle \tag{7.151}$$

The rotation properties of T_{jm} obtained with this prescription is precisely the ones obeyed by a tensor operator. This is guaranteed by its parallel structure to the states of total angular momentum. In particular, the example $T_{lm}(\mathbf{v}) = |\mathbf{v}|^l Y_{lm}(\hat{\mathbf{v}})$ may be regarded as a combination of products of l tensor operators $T_{1m}(\mathbf{v}) = \mathbf{v}_m$ reduced to an irreducible one by the multiple use of Clebsch-Gordan coefficients.

Similarly, we may consider products of operators and states $T_{j_1 m_1}|j_2 m_2\rangle$. Without knowing the details of either the operator or the state, we may still combine them into states that transform irreducibly under rotations

$$T_{j_1 m_1}|j_2 m_2\rangle = \sum_{jm} |jm\rangle \langle jm|m_1 m_2\rangle \ \langle j||T_{j_1}||j_2\rangle \tag{7.152}$$

The coefficients $\langle j||T_{j_1}||j_2\rangle$ are called *reduced matrix elements*. They are inserted because the states on the left and right are assumed to be normalized; otherwise the reduced matrix elements $\langle j||T_{j_1}||j_2\rangle$ could have been absorbed into a redefinition of the states. They depend only on j, j_1, j_2, but are independent of the magnetic quantum numbers m_1, m_2, m. Therefore they are scalars under rotations, and their presence does not alter the rotation properties of the states. They may be computed by evaluating the matrix elements

$$\langle j_3 m_3|T_{j_1 m_1}|j_2 m_2\rangle = \langle j_3 m_3|m_1 m_2\rangle \ \langle j_3||T_{j_1}||j_2\rangle, \tag{7.153}$$

and they obviously depend on the detailed construction of the states and operators. To obtain the reduced matrix element the most convenient values of m_1, m_2, m_3 may be chosen on the left hand side, as long as the corresponding Clebsch-Gordan coefficient on the right hand side is non-zero.

The content of eq.(7.153) is called the *Wigner-Eckhart theorem*. It is a very powerful result of rotation covariance. It says that all the matrix elements $\langle j_3 m_3|T_{j_1 m_1}|j_2 m_2\rangle$ are proportional to a single overall constant $\langle j_3||T_{j_1}||j_2\rangle$ that depends on the details of the system, while the ratios of the matrix elements depend only on the Clebsch-Gordan coefficients that are completely known and independent of the details of the system

$$\frac{\langle j_3 m_3|T_{j_1 m_1}|j_2 m_2\rangle}{\langle j_3 m_3'|T_{j_1 m_1'}|j_2 m_2'\rangle} = \frac{\langle j_3 m_3|m_1 m_2\rangle}{\langle j_3 m_3'|m_1' m_2'\rangle}. \tag{7.154}$$

This result has many interesting applications in all branches of physics. In particular we mention *selection rules* in decays or scattering of quantum systems,

such as atoms, nuclei or particles. The transition amplitude has the same structure as $\langle j_3 m_3 | T_{j_1 m_1} | j_2 m_2 \rangle$ where the states are the initial and final states, while $T_{j_1 m_1}$ is the transition operator. Obviously, certain transitions are not allowed if the corresponding Clebsch-Gordan coefficient vanishes! Also, the allowed transitions occur with definite ratios that are predetermined by the values of the Clebsch-Gordan coefficients. The application of this basic idea is very broad.

It must be emphasized that in Eqs.(7.149,7.151) above the operators T_{jm} are not necessarily normalized according to any particular rule, so that any possible rotationally invariant extra factors are absorbed into a redefinition of the operators. Thus, if the operators $T_{jm}, T_{j_1 m_1}, T_{j_2 m_2}$ are each multiplied with arbitrary constant or operator factors which are rotationally invariant (independent of m, m_1, m_2, but possibly dependent on j, j_1, j_2), we can still write the same equations as above for the new operators provided we absorb all the m, m_1, m_2 independent extra factors as part of an appropriate redefinition of the new operators $\tilde{T}_{jm}, \tilde{T}_{j_1 m_1}, \tilde{T}_{j_2 m_2}$. However, in some circumstances, if one wishes to define operators $T_{jm}, T_{j_1 m_1}, T_{j_2 m_2}$ normalized according to some rule (see e.g. problem 10), then more generally one may include factors $C_{j_1 j_2 j}$ that may depend on j, j_1, j_2 in combinaning operators, as follows

$$T_{j_1 m_1} T_{j_2 m_2} = \sum_{jm} C_{j_1 j_2 j} T_{jm} \langle jm | m_1 m_2 \rangle, \tag{7.155}$$

$$T_{jm} = (C_{j_1 j_2 j})^{-1} \sum_{m_1 m_2} T_{j_1 m_1} T_{j_2 m_2} \langle m_1 m_2 | jm \rangle. \tag{7.156}$$

Here $C_{j_1 j_2 j}$ may be constants or operators. The important point is that the extra factors $C_{j_1 j_2 j}$ are invariant under rotations, and if it were not for some desired normalization they could have been absorbed into a redefinition of the $T_{jm}, T_{j_1 m_1}, T_{j_2 m_2}$. By specializing to some values of m, m_1, m_2 and using the corresponding Clebsch-Gordan coefficient one could compute the extra factors once and for all, so that the equation becomes fully determined for all other values of m, m_1, m_2.

7.9 Problems

1. Consider the rotation of the bra in position space $\langle \mathbf{r} |' = \langle \mathbf{r} | e^{i\boldsymbol{\omega} \cdot \mathbf{L}/\hbar}$. By using $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ show that the result is $\langle \mathbf{r} |' = \langle \mathbf{r}' |$, where

$$\mathbf{r}' = e^{\boldsymbol{\omega} \times} \mathbf{r} = \mathbf{r} + \boldsymbol{\omega} \times \mathbf{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + \cdots \tag{7.157}$$

2. Compute the rotation matrices $D^{j=1}(\boldsymbol{\omega})$ by summing up the series. Note that $\left(\vec{\omega} \cdot \vec{\Im}^{j=1}\right)^n$ is proportional to $\left(\vec{\omega} \cdot \vec{\Im}^{j=1}\right)$ or $\left(\vec{\omega} \cdot \vec{\Im}^{j=1}\right)^2$ depending on $n =$even, odd, respectively, and with coefficients that are powers of only $|\boldsymbol{\omega}|$. Compare your result to eq.(7.48).

3. The computation of $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$ may be carried out by considering two rotations applied on a vector as follows

$$e^{\boldsymbol{\omega}_2 \times} \left(e^{\boldsymbol{\omega}_1 \times} \mathbf{r} \right) = e^{\boldsymbol{\omega}_3 \times} \mathbf{r}.$$

Compute it up to 4th order (i.e. no more than 2 powers in either $\boldsymbol{\omega}_1$ or $\boldsymbol{\omega}_2$) and show that your results are in agreement with the expansion in eq.(7.10).

4. The product of two rotations that yields $\boldsymbol{\omega}_3(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$ may be computed exactly by using the $j = 1/2$ representation as outlined just before eq.(7.14). Verify this result and then compute the expansion for small $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2$ and show that it agrees with eq.(7.10).

5. Consider rotations applied on a quantum state $\langle\psi|$. A first rotation of $\pi/3$ around the $\hat{\mathbf{z}}$ axis is followed by another rotation of $\pi/3$ around the *original* $\hat{\mathbf{y}}$ axis,

   (a) If the original state $\langle\psi|$ has angular momentum $j = 2$ and $m = 1$, what are the probabilities that the rotated state has $(j = 1, m = 0)$, $(j = 2,\ m = 0)$, $(j = 3,\ m = 0)$?

   (b) If the original state is the linear superposition

$$\langle\psi| = \frac{1}{\sqrt{2}}\langle j = 1, m = 1| + \frac{1}{\sqrt{2}}\langle j = 2, m = 1|, \tag{7.158}$$

   what is the probality that the rotated state has $(j = 1, m = 0)$, $(j = 2,\ m = 0)$, $(j = 3,\ m = 0)$?

   (c) What is the overall rotation angle $\boldsymbol{\omega}$ (magnitude ω and direction $\hat{\boldsymbol{\omega}}$= x, y, z componets).
   You should provide analytic answers in terms of D-functions for angular momentum, and then plug in the numbers and give complete numerical answers.

6. Compute the rotation vector $\vec{\omega}$ that is equivalent the rotation produced by the Euler angles α, β, γ

7. Use the symmetry properties of the $3j$ symbols to derive the Clebsch-Gordan coefficients for

$$\begin{gathered} j_1 \times (j_1 + \tfrac{1}{2}) \to \tfrac{1}{2} \\ j_1 \times (j_1 + 1) \to 1 \\ j_1 \times j_1 \to 0 \end{gathered}$$

   by using the coefficients for $j_1 \times \frac{1}{2} \to j_1 + \frac{1}{2}$ and $j_1 \times 1 \to (j_1 + 1)$ listed in Tables (7.1,7.2) respectively. Compare your results to the ones obtained from the general formula (7.116).

8. By using any of the spin-1/2 representations given in (7.44-7.45) obtain the expressions for $\left(\phi', \theta', \gamma'\right)$, $\left(\phi'', \theta'', \gamma''\right)$ and verify the left/right invariance property of the Haar measure, $d\mu(\phi, \theta, \gamma) = d\mu(\phi', \theta', \gamma') = d\mu(\phi'', \theta'', \gamma'')$ given in (7.125).

9. As an example of a tensor operator consider the creation operators $a_1^\dagger, a_2^\dagger$ of the harmonic oscillator in two dimensions. Under commutation with the SU(2) operators $J_+ = a_1^\dagger a_2,\ J_- = a_2^\dagger a_1,\ J_0 = \frac{1}{2}\left(a_1^\dagger a_1 - a_2^\dagger a_2\right)$ show that they satisfy the conditions for a tensor operator with $j = 1/2$. Thus, we may identify $T_{\frac{1}{2}m} = a_m^\dagger$ where we may use $m = \pm\frac{1}{2}$, i.e. $a_m^\dagger = a_{\pm\frac{1}{2}}^\dagger$ instead of using the labels $a_1^\dagger, a_2^\dagger$. Then show that the annihilation operators also form a tensor operator a_m where we need to identify the doublet as $a_{+\frac{1}{2}} = -a_2$ and $a_{-\frac{1}{2}} = a_1$.

10. Generalizing the ideas of the previous problem, show that the following operator is a tensor operator

$$T_{jm} = \frac{\left(a_1^\dagger\right)^{j+m}\left(a_2^\dagger\right)^{j-m}}{\sqrt{(j+m)!}\,\sqrt{(j-m)!}}$$

11. Consider the product of two tensor operators of the type given in the previous problem $T_{j_1 m_1} T_{j_2 m_2}$. Show that their product is another tensor operator with $j = j_1 + j_2$. How does this fit with the general formula $T_{j_1 m_1} T_{j_2 m_2} = \sum_{jm} \langle j_1 m_1 j_2 m_2 | jm \rangle T_{jm}$? From the resulting expression derive the Clebsch-Gordan coefficient for the angular momentum addition $j_1 \times j_2 \to (j_1 + j_2)$, and compare to the results for the special cases $j_1 = j$ and $j_2 = 1/2$ or $j_2 = 1$ given in tables (7.1) and (7.2) respectively. Do you agree with those tables?

12. Consider the states of the two dimensional harmonic oscillator and the matrix element $\langle j_1 m_1 | T_{jm} | j_2 m_2 \rangle$. Show how the Wigner-Echart theorem applies in this case, and compute the reduced matrix element.

13. Application of Wigner-Eckhart theorem to selection rules.

14. The fundamental theory of Quantum Chromodynamics produces an effective interaction between two heavy spin 1/2 quarks. The dominant part of this interaction is the linear potential $V(\mathbf{r}_1, \mathbf{r}_2) = \gamma\,|\mathbf{r}_1 - \mathbf{r}_2|$ proportional to the distance between them, as shown in Fig.(6.8a). There is also a $1/r$ term, as in the problems at the end of Chapter 1, as well as spin dependent terms, but for the sake of this excercise we will suppress these terms, except that the spin effects will be taken into account roughly as

described below

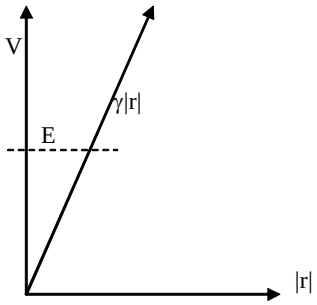


Fig.(7.3a) . $V(\mathbf{r}) = \gamma |\mathbf{r}|$

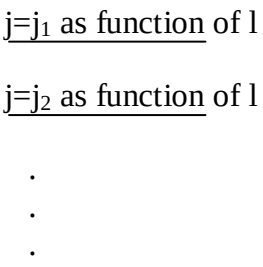


Fig.(7.3b) - Energy as function of $j\,(l)$

a) Write down the **radial** Schrödinger equation in relative radial coordinates explaining carefully the relation of each term in your equation to the two-body Hamiltonian H.

b) Each state is labelled with a quantum number for energy, orbital angular momentum, and spin. In the **total** angular momentum scheme, if the orbital quantum number of some state is l give all the values of total angular momentum j for a fixed l. Then take the spin effects into consideration as follows: If $j_1 > j_2$ then the level labelled by the greater j_1 has a higher energy as shown in Fig.(6.8b). Taking this into account, draw an energy level diagram similar to Fig.(6.8b) for **all** levels with $l = 0, 1, 2, 3$. Give separate diagrams for $l = 0$ and $l \neq 0$, and indicate clearly how many states with same j there are at each level.

c) Using the uncertainty principle, or any other rough approximation scheme, **estimate** the energy E of the lowest state for each fixed value of orbital angular momentum l, and give your answer as a function of the parameters $m, \gamma, \hbar$ and l.

d) Neglecting the spin effects, solve the radial differential equation for the ground state with $l = 0$ and give the exact condition for the quantization of energy for the levels with $l = 0$. Hint: the equation can be solved in Fourier space. Then find the relation of the radial wavefunction in r space to the Airy function and use the properties of this function to find the energy eigenvalues, either numerically or through a plot.

15. Consider the excited levels of an open string in d=3 dimensions whose Hamiltonian is written in terms of its normal modes, as discussed in chapter 5

$$H = \frac{\tilde{\mathbf{p}}^2}{2\pi\mu} + \hbar\omega \sum_{n=1}^{\infty} n\, N_n, \quad N_n = \tilde{\mathbf{a}}_n^\dagger \cdot \tilde{\mathbf{a}}_n = \text{number operator for level } n$$

The ground state is an eigenstate of the center of mass momentum operator $\tilde{\mathbf{p}}|\mathbf{k}, 0\rangle = |\mathbf{k}, \mathbf{0}\rangle \hbar\tilde{\mathbf{k}}$. Excited states are obtained by applying creation

operators for any normal mode. These are arranged by the energy levels, where level is defined by the sum $\sum nN_n$. Thus, the ground state is at level zero, the states $a^\dagger_{1I}\,|\mathbf{k},0\rangle$ are at level one, the states $a^\dagger_{2I}\,|\mathbf{k},0\rangle$, and $a^\dagger_{1I}\;a^\dagger_{1J}\,|\mathbf{k},0\rangle$ are at level two, and so on. In this problem we will consider the excited states that can be constructed up to level 4.
The total angular momentum of the string is given by $\vec{J}=\vec{L}+\vec{S}$ where $\vec{L}=\vec{r}\times\vec{p}$ is the orbital angular momentum of the center of mass, and $\vec{S}=\sum_{n=1}^{\infty}\left(\tilde{\mathbf{a}}^\dagger_n\times\tilde{\mathbf{a}}_n\right)$ is the spin contributed by the excitations. Under rotations the creation operators $\tilde{\mathbf{a}}^\dagger_n$ transform as vectors, and the various excited levels can be classified as *traceless symmetric or antisymmetric tensors* constructed from these creation operators. For SO(3) an antisymmetrizing two vectors gives again a vector due to the existence of the Levi-Civita tensor (e.g. $V_IW_J-V_JW_I=\varepsilon_{IJK}(\vec{V}\times\vec{W})_K$). After taking this effect into account there remains symmetrizing vectors (or composite vectors such as $\vec{V}\times\vec{W}$) to construct symmetric *traceless* tensors that correspond to higher spins. The rank of a symmetric traceless tensor is the spin of the state. For example at level 2 there are states with spin 0,1,2 that correspond to the traceless tensors of rank zero $a^\dagger_1\cdot\,a^\dagger_1\,|\mathbf{k},0\rangle$, rank one $a^\dagger_{2I}\,|\mathbf{k},0\rangle$, and rank two $\left(a^\dagger_{1I}\;a^\dagger_{1J}\;-\frac{1}{d}\delta_{IJ}a^\dagger_1\cdot\,a^\dagger_1\right)|\mathbf{k},0\rangle$, respectively.

The object of the current problem is to classify the spins at every level up to level 4 by constructing explicitly the traceless tensors. Thus, first argue on the basis of addition angular momentum, $j_1\times j_2=|j_1-j_2|+\cdots+(j_1+j_2)$, which spins you expect at every level, and how many states with the same spin would occur. Then construct explicitly the traceless tensors, and specify the spin of each state up to level 4. Then create a plot of the spin versus the level and insert in the plot the degeneracy of states that have the same spin and level, at every point of the plot.

16. Consider a spin 2 nucleus placed in a field such that its energy is given by the Hamiltonian

$$H=a\left(2\tilde{\mathbf{S}}^2-4\left(S_0\right)^2\right)+b\left(\left(S_+\right)^3+\left(S_-\right)^3\right),$$

where $\tilde{\mathbf{S}}$ is the spin operator and a,b are constants. Evidently this Hamiltonian is not diagonal in the $|sm_s\rangle$ basis with $s=2$. In this problem we are interested in the eigenstates and eigenvalues of this Hamiltonian, where the eigenstates are to be expressed as a linear combination of the $|sm_s\rangle$. Note that parts (d,e,f) can be done without needing the results of parts (a,b,c).

(a) Give an argument that a 180^o rotation around the x-axis is a symmetry of this Hamiltonian (formulas not needed, a picture may be helpful)).

(b) Construct the corresponding symmetry operator R and show that it commutes with H. From staring at R give a quick argument for what the eigenvalues of R can be?

(c) Show how the states $|sm_s\rangle$ transform under R, and find the transformed states $|sm_s\rangle'$ in terms of the $|sm_s\rangle$ *explicitly (i.e. not a formal answer).* [Hint: by considering what happens to $\tilde{\mathbf{S}}$ under the 180^o rotation R you can essentially guess the answer; but for a proof use Euler angles, or use a rotation of the $\hat{x}$ axis into the $\hat{y}$ axis, to take advantage of known $d^j_{mm'}$ functions]. From this result construct the eigenstates of R.

(d) Construct the matrix elements of the Hamiltonian in the $|sm_s\rangle$ basis.

(e) Find the eigenvalues E_n of the Hamiltonian (for this step, and the next, it is useful to first reorganize the matrix into a block diagonal form).

(f) Find the eigenstates $|E_n\rangle$ as a linear combination of the states $|sm_s\rangle$.

(g) Find the linear combination of the states $|sm_s\rangle$ that are simultaneous eigenstates of H and R.

17. Consider a spin 1/2 electron constrained to move on a spherical shell of a fixed radius $r = r_0$ (this constraint requires also a vanishing radial momentum $p_r = 0$). Its interactions include spin-orbit coupling $\mathbf{L} \cdot \mathbf{S}$ due to the magnetic moment of the electron, and an interaction proportional to $\tilde{\mathbf{r}} \cdot \tilde{\mathbf{E}} = RE\cos\theta$ due to an external constant electric field $\tilde{\mathbf{E}}$. Let us assume the Hamiltonian is organized into three terms with constant coefficients a, b, c

$$H = a\left(\mathbf{L} + \mathbf{S}\right)^2 + b\mathbf{L} \cdot \mathbf{S} + c\cos\theta$$

where $\vec{L}$, $\vec{S}$ are the orbital and spin angular momenta respectively. In this problem we are concerned with computing the energy levels and states of the system in certain limits of the parameters a, b, c. You only need to use algebraic methods involving angular momentum to answer the following questions.

(a) In the limit $b = c = 0$ give the exact energy eigenstates and eigenvalues by providing a complete set of labels including the range of all eigenvalues of simultaneous observables. Provide an energy level diagram including the lowest 3 energy levels in which you clearly indicate all the states with all their quantum numbers. Give the degeneracy of each of these energy levels, specifically note that the ground state is degenerate.

(b) Next, consider $a, b \neq 0$ while $c = 0$, and again find the exact energy eigenvalues and eigenstates. On the energy level diagram indicate how the degeneracies of the previous case are broken and again clearly show the energies, other quantum numbers, and degeneracies of the first three **new** ground and excited levels.

(c) As the value of b changes from very small to very large consider what happens to the ordering of the energy levels. At which critical value of b/a does the order of energy levels begins to change?

(d) We would like to compute the matrix elements of the Hamiltonian in the limit $b = 0$, but $a, c \neq 0$. Note that $\cos\theta \sim Y_{10}(\theta)$ is a vector operator. Consider only the degenerate states at the ground level you found in part (1), and take advantage of the Wigner-Eckhart theorem to compute the matrix elements among these states. For your answer give the full matrix explicitly, including all coefficient computed in *all detail*.

18. Consider the operator $R = \exp(i\pi J_1)$ for a rotation around x-axis by the angle π. Compute the action of R on the states $|jm\rangle$ for any j, m. You can compute this with brute force by using the D^j functions, but in this exercise you are required to do the computation using a more clever method that takes advantage of the fact that $RJ_1R^{-1} = +J_1$, $RJ_2R^{-1} = -J_2$, $RJ_3R^{-1} = -J_3$, and $R^2|jm\rangle = (-1)^{2j}|jm\rangle$ since R^2 is a 2π rotation. Write out $R_{m'm} = \langle jm'|R|jm\rangle$ explicitly in the form of a matrix. Verify that the D^j functions give the same result. How would you compute the square root of this matrix? give an expression for it.

34. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND *d* FUNCTIONS

Note: A square-root sign is to be understood over *every* coefficient, *e.g.*, for $-8/15$ read $-\sqrt{8/15}$.

Notation:

J	J	...
M	M	...
m_1	m_2	Coefficients
m_1	m_2	
.	.	
.	.	
.	.	

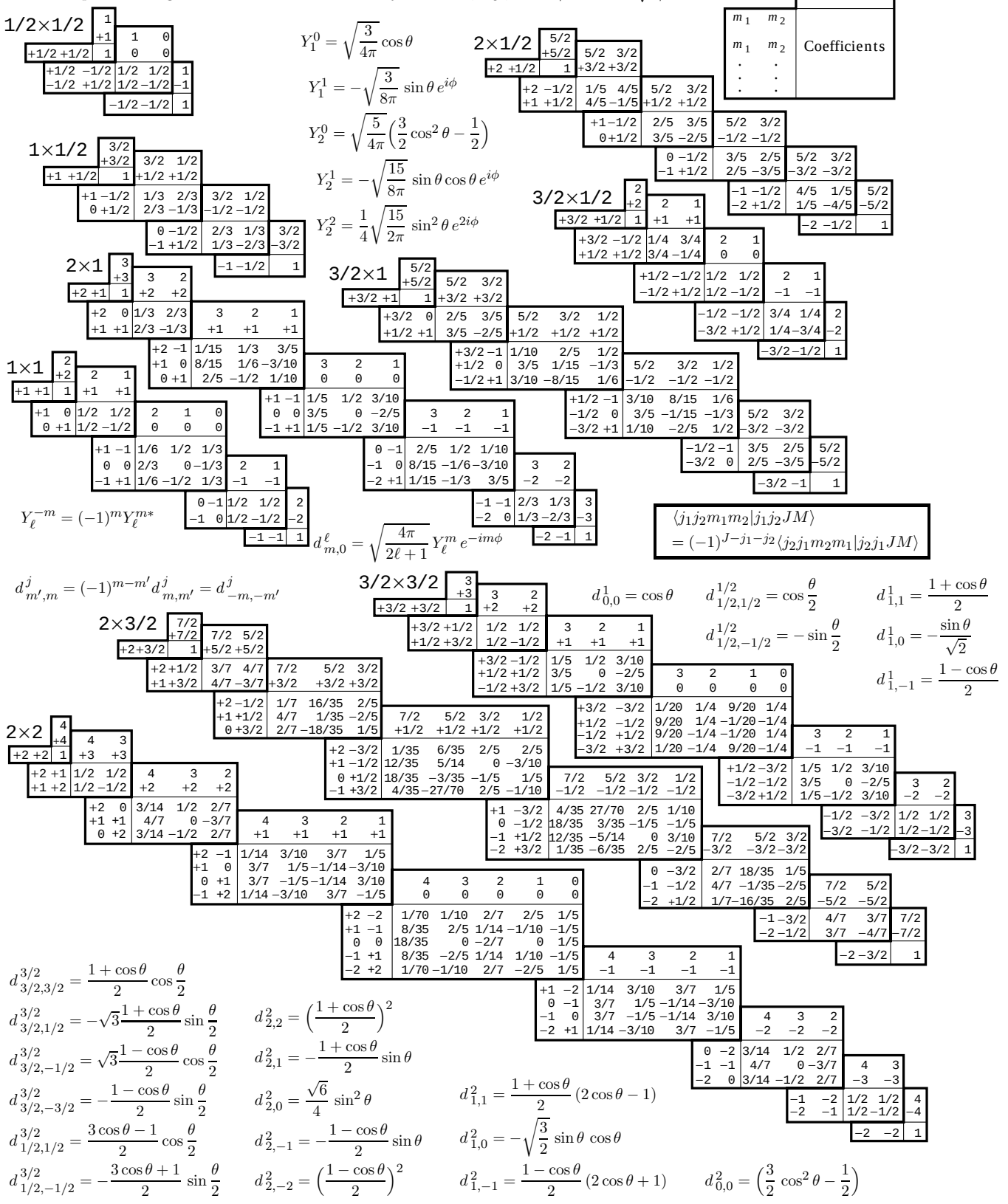


Figure 34.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974). The coefficients here have been calculated using computer programs written independently by Cohen and at LBNL.

Chapter 8

SYMMETRY

Physical systems often have symmetries. This is manifested by different observers obtaining the same results when they perform experiments under different conditions that are related by the symmetry. For example observer A and observer B that do experiments in laboratories at different locations obtain the same results because of translation symmetry. Similarly the results are the same when the laboratories are at different orientations because of rotation symmetry. These are general properties of the laws of Nature. In addition, complicated systems of many interacting particles sometimes have internal symmetries. In this chapter we will be concerned with space-time symmetries of non-relativistic particle systems, their classical and quantum formulation, their mathematical properties, and their usage in simplifying and solving problems. However, some of the discussion is more general and applies to more sophisticated theories in various branches of Physics.

8.1 Symmetry in classical physics

Observers A and B use their own coordinate systems to keep track of the particles. For the particle labelled by the index i let us define A's coordinates by $\mathbf{r}_i$ and B's coordinates by $\mathbf{r}'_i$. These are related to each other by coordinate transformations that involve several parameters. For example in the case of translations $\mathbf{r}'_i = \mathbf{r}_i + \mathbf{a}_i$, where $\mathbf{a}_i$ are the parameters. It is useful to consider nearby observers which are related to each other by infinitesimal coordinate transformations. If the infinitesimal parameters for N symmetries are $\epsilon_a,\ a = 1, 2, \cdots, N$, we may expand the relation between $\mathbf{r}_i$ and $\mathbf{r}'_i$ to first order in the ϵ_a's, and write

$$r_i^{'I} = r_i^I + \delta_\epsilon r_i^I \ , \qquad \delta_\epsilon r_i^I = \sum_a \epsilon_a f_i^{Ia}(\mathbf{r}, \dot{\mathbf{r}}) \tag{8.1}$$

where $I = 1, 2, 3$ denotes the vector index.

If the two observers A and B see identical physical phenomena and measure the same results, it must be that the equations that they use in terms of $\mathbf{r}_i$ and $\mathbf{r}'_i$ respectively have the *same form*. If one takes A's equations and substitutes $\mathbf{r}'_i$ instead of $\mathbf{r}_i$ the resulting equations are B's equations. Only if there is a symmetry, B's equations, rewritten in terms of $\mathbf{r}_i$, will yield A's equations in identical form, not otherwise.

Instead of discussing the symmetries of the equations of motion, it is more efficient to consider the action from which they are derived by a variational principle. The action S is constructed from a Lagrangian in the form $S(\mathbf{r}_i) = \int_1^2 dt\ L(\mathbf{r}_i(t), \dot{\mathbf{r}}_i(t))$. The Euler equations are then

$$\frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{\mathbf{r}}_i(t)} - \frac{\partial L}{\partial \mathbf{r}_i(t)} = 0\ . \tag{8.2}$$

There will be a symmetry provided, under the substitution $\mathbf{r}_i \to \mathbf{r}'_i$, the form of the action remains invariant up to a "constant"

$$S(\mathbf{r}) \to S(\mathbf{r}') = S(\mathbf{r}) + constant(1,2). \tag{8.3}$$

This equality will guarantee that the equations of motion derived for $\mathbf{r}'_i$ will have an identical form to those derived for $\mathbf{r}_i$. The "$constant(1,2)$" (which is zero in most cases) is allowed to depend on the coordinates or velocities at the end points $t = t_1, t_2$. Since the initial and final points are kept fixed ($\delta\mathbf{r}_i(t_1) = \delta\mathbf{r}_i(t_2) = 0$) in the variational principle that leads to the Euler equations (8.2) the "$constant(1,2)$" does not contribute to the equations of motion. Whenever the action has such a symmetry then two observers related to each other by the symmetry transformation must observe identical physics. Therefore, the mathematical formulation of symmetry at the classical level is reduced to the symmetries of the action under coordinate transformations as in (8.3). Now, substituting in (8.3) the infinitesimal transformation (8.1) valid up to first order in the parameters, and expanding $S(\mathbf{r}'_i) = S(\mathbf{r}_i + \delta_\epsilon \mathbf{r}_i) \approx S(\mathbf{r}_i) + \delta_\epsilon S$, one must get $\delta_\epsilon S = 0$ to first order because of the symmetry (8.3). Thus, symmetry implies

$$\delta_\epsilon S = \int_1^2 dt[\frac{\partial L}{\partial \dot{\mathbf{r}}_i(t)}\delta_\epsilon \dot{\mathbf{r}}_i(t) + \frac{\partial L}{\partial \mathbf{r}_i(t)}\delta_\epsilon \mathbf{r}_i(t)]\ =\ 0\ . \tag{8.4}$$

Here $\delta_\epsilon \dot{\mathbf{r}}_i(t)$ is obtained by taking a time derivative of (8.1). Let us give a few simple examples of symmetric actions.

(i) The free particle moving in one direction is a trivial but useful example. It is described by the Lagrangian $L(x) = \frac{1}{2}m\dot{x}^2$. This is invariant under translations $x' = x + \epsilon$ since $S(x) = S(x')$. Therefore two observers using coordinates whose origin differs by a translation must see the same motion of the free particle. Indeed the equation of motion used by A is $\ddot{x} = 0$, while the one used by B is $\ddot{x}' = 0$, which has the same form. These equations are derived from $S(x)$ and $S(x')$ respectively.

(ii) The free particle in 3 dimensions is described by the Lagrangian $L(\mathbf{r}) = \frac{1}{2}m\dot{\mathbf{r}}^2$. It has obvious rotation and translation symmetries. The infinitesimal transformations are $\delta\mathbf{r} = \boldsymbol{\omega} \times \mathbf{r} + \mathbf{a}$, and when substituted in (8.4) one gets zero.

(iii) The general central force problem for the two particle system, which is of great interest in many physical applications, has the Lagrangian

$$L(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}m_1^2\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2^2\dot{\mathbf{r}}_1^2 - V(|\mathbf{r}_1 - \mathbf{r}_2|). \tag{8.5}$$

This system is obviously symmetric under simultaneous translations and rotations of both coordinates as in example (ii), but in addition there is a symmetry under an extra internal rotation of the relative coordinate. The infinitesimal symmetry transformations are

$$\begin{array}{rcl} \delta\mathbf{r}_1 &=& \boldsymbol{\omega} \times \mathbf{r}_1 + \mathbf{a} + \frac{m_2}{m_1+m_2}\vec{\epsilon} \times (\mathbf{r}_1 - \mathbf{r}_2) \ , \\ \delta\mathbf{r}_2 &=& \boldsymbol{\omega} \times \mathbf{r}_2 + \mathbf{a} - \frac{m_1}{m_1+m_2}\vec{\epsilon} \times (\mathbf{r}_1 - \mathbf{r}_2). \end{array} \tag{8.6}$$

It is straightforward to check that the action remains invariant. That is, $S(\mathbf{r}_1', \mathbf{r}_2') = S(\mathbf{r}_1, \mathbf{r}_2) + \delta S$, with $\delta S = 0$ to first order in $\boldsymbol{\omega}$, $\mathbf{a}$, or $\vec{\epsilon}$ *(homework problem).*

The symmetries of example (iii) may be better understood by defining center of mass and relative coordinates, as in the chapter on the central force problem

$$\begin{array}{ll} \mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/M \ , & \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \ , \\ M = m_1 + m_2 \ , & \mu = m_1m_2/M \ . \end{array} \tag{8.7}$$

In terms of these the center of mass and relative motion decouple, and the Lagrangian in (8.5) splits into two independent pieces that do not communicate with each other

$$L = L_{CM} + L_{rel} \ , \qquad L_{CM} = \frac{1}{2}M\dot{\mathbf{R}}^2 \ , \qquad L_{rel} = \frac{1}{2}\mu\dot{\mathbf{r}}^2 - V(r). \tag{8.8}$$

Then from (8.6) one can derive the transformation rules of $\mathbf{R}$ and $\mathbf{r}$

$$\delta\mathbf{R} = \boldsymbol{\omega} \times \mathbf{R} + \mathbf{a} \ , \qquad \delta\mathbf{r} = (\boldsymbol{\omega} + \boldsymbol{\epsilon}) \times \mathbf{r} \ . \tag{8.9}$$

This allows one to interpret $(\boldsymbol{\omega}, \mathbf{a})$ as the symmetry parameters of the overall system, with the translations applied only to the center of mass, while $\vec{\epsilon}$ is clearly a symmetry parameter of an internal orbital rotation that leaves the Lagrangian L_{rel} invariant by itself.

8.2 Symmetry and classical conservation laws

The above examples illustrate some simple physical systems with symmetries. Now consider a general Lagrangian describing an arbitrary system of particles

located at $\mathbf{r}_i(t)$ at time t. Suppose the Lagrangian has a symmetry under the infinitesimal transformation of (8.1) with some specific functions $f_i^{Ia}(\mathbf{r}_j, \dot{\mathbf{r}}_j)$. According to Noether's theorem, that we will prove below, corresponding to every symmetry parameter ϵ_a there exists a conserved quantity $Q_a(\mathbf{r}_i, \dot{\mathbf{r}}_i)$ that is time independent. That is, even though the location and velocities of the particles may be changing with time, the conserved quantities Q_a, which are constructed from them, remain unchanged, *i.e.* $dQ_a/dt = 0$. The conservation of energy, momentum and angular momentum are some examples of consequences of symmetry. There are many more interesting cases in specific physical systems.

To construct the explicit form of $Q_a(\mathbf{r}_i, \dot{\mathbf{r}}_i)$ and prove Noether's theorem, first note that the symmetry of the action (8.3) is satisfied most generally provided the Lagrangian behaves as follows

$$L(\mathbf{r}_i'(t), \dot{\mathbf{r}}_i'(t)) = \frac{\partial t'}{\partial t} L(\mathbf{r}_i(t'), \dot{\mathbf{r}}_i(t')) + \frac{\partial}{\partial t}\alpha(\mathbf{r}_i(t), \dot{\mathbf{r}}_i(t)). \tag{8.10}$$

Here $t'(t, \epsilon)$ is a change of variables that generally may depend on the parameters of the symmetry transformation, and $\frac{\partial t'}{\partial t}$ is the Jacobian for the change of variables. α is some function of the dynamical variables and the parameters, which vanishes as $\epsilon_a \to 0$. The function α is zero in most cases, but not for every case, as will be seen in examples below. Also, in most cases $t'(t, \epsilon) = t$, otherwise the infinitesimal expansion gives $t' = t + \sum_a \epsilon_a \gamma^a(\mathbf{r}_i(t), \dot{\mathbf{r}}_i(t))$ with some functions γ^a. When equation (8.10) is integrated, the left side yields $\int_1^2 L(\mathbf{r}_i'(t), \dot{\mathbf{r}}_i'(t)) = S(\mathbf{r}_i')$, and the right side gives

$$\begin{array}{l} \int_1^2 dt \frac{\partial t'}{\partial t} L(\mathbf{r}_i(t'), \dot{\mathbf{r}}_i(t')) + \int_1^2 dt \frac{\partial}{\partial t}\Lambda(\mathbf{r}_i(t), \dot{\mathbf{r}}_i(t)) \\ = \int dt' L(\mathbf{r}_i(t'), \dot{\mathbf{r}}_i(t')) + \Lambda(1) - \Lambda(2) \\ = S(\mathbf{r}_i) + constant(1,2). \end{array} \tag{8.11}$$

Thus, the condition of symmetry (8.3) is equivalent to (8.10).

Now expand (8.10) to first order in ϵ_a on both sides. After dropping L from both sides it takes the form

$$\delta_\epsilon L = \frac{d\Lambda}{dt}, \tag{8.12}$$

where $\delta_\epsilon L = \frac{\partial L}{\partial r_i^I}\delta_\epsilon r_i^I + \frac{\partial L}{\partial \dot{r}_i^I}\delta_\epsilon \dot{r}_i^I$, and $\Lambda\left(\mathbf{r}_i, \dot{\mathbf{r}}_i\right) = \epsilon_a \gamma^a L + \epsilon_a \frac{\partial \alpha}{\partial \epsilon_a}$ is some function of the dynamical variables which is first order in the parameters ϵ^a. Λ is zero in most cases, but not always. That is, when one applies an infinitesimal symmetry transformation on the Lagrangian, the result is either zero, or at most a total time derivative. By contrast, if there is no symmetry, the result can be anything. Now use Euler's equation of motion $\frac{d}{dt}\frac{\partial L}{\partial \dot{r}_i^I} - \frac{\partial L}{\partial r_i^I} = 0$ (8.2), and recall that the canonical momentum is $p_i^I = \frac{\partial L}{\partial \dot{r}_i^I}$, to write (8.12) in the form

$$\frac{d}{dt}(\delta_\epsilon \mathbf{r}_i \cdot \mathbf{p}_i) = \frac{d\Lambda}{dt} \tag{8.13}$$

This shows that there exist a conserved quantity $dQ_a/dt = 0$ for every linearly independent infinitesimal parameter ϵ_a, and that it is given by

$$\epsilon^a Q_a = \delta_\epsilon \mathbf{r}_i \cdot \mathbf{p}_i - \Lambda. \tag{8.14}$$

Substituting the explicit form of the transformation (8.1), and differentiating with respect to ϵ^a on both sides one can write

$$Q_a(\mathbf{r}_i, \dot{\mathbf{r}}_i) = f_i^{aI}(\mathbf{r}_i, \dot{\mathbf{r}}_i)\, p_i^I - \frac{\partial}{\partial \epsilon_a}\Lambda(\mathbf{r}_i, \dot{\mathbf{r}}_i). \tag{8.15}$$

This is a generalized version of Noether's theorem that includes the Λ.

We have seen that a conserved charge is associated with every symmetry parameter. To compute it, all one needs is the explicit form of the infinitesimal symmetry transformation, i.e. the functions $f_i^{aI}(\mathbf{r}_i, \dot{\mathbf{r}}_i)$, and then plug it into the Lagrangian in order to compute Λ as in (8.12). Let us see how this works in the examples (i), (ii), (iii) above for which one finds $\Lambda = 0$:

(i) The free particle in one dimension has a translation invariance under $\delta x = a$. In this case $\delta L = m\dot{x}\dot{a} = 0$, since a is time independent. Hence $\Lambda = 0$. Then eq.(8.14) takes the form $aQ = ap$ where a has been used instead of ϵ. Eliminating a from both sides one identifies the conserved quantity as the momentum

$$Q = p. \tag{8.16}$$

So, the canonical momentum p is time independent. Indeed, the equation of motion is $\dot{p} = 0$, which confirms the expected conservation.

(ii) The free particle in three dimensions has a rotation invariance under $\delta\mathbf{r} = \boldsymbol{\omega} \times \mathbf{r}$, with $\Lambda = 0$. Then eq.(8.14) takes the form $\boldsymbol{\omega} \cdot \mathbf{Q} = (\boldsymbol{\omega} \times \mathbf{r}) \cdot \mathbf{p}$ where ω^a has been used instead of ϵ^a. Eliminating it from both sides one finds that the conserved quantity is angular momentum

$$\mathbf{Q} = \mathbf{r} \times \mathbf{p}. \tag{8.17}$$

In addition, there is a translation symmetry under $\delta\mathbf{r} = \mathbf{a}$ with $\Lambda = 0$, which yields another conserved quantity $\mathbf{p}$, which is the three dimensional momentum. Thus, for the free particle, it is dictated that momentum $\mathbf{p}$ and angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ are conserved because of translation and rotation symmetries respectively.

(iii) The central force problem has symmetries under the transformations given in (8.6). First, let us deal with translations $\delta\mathbf{r}_1 = \delta\mathbf{r}_2 = \mathbf{a}$, for which $\Lambda = 0$. Eq.(8.14) gives $\mathbf{a} \cdot \mathbf{Q} = \mathbf{a} \cdot \mathbf{p}_1 + \mathbf{a} \cdot \mathbf{p}_2$. Hence the conserved quantity is the total momentum of the system

$$\mathbf{Q} = \mathbf{p}_1 + \mathbf{p}_2 \equiv \mathbf{P}. \tag{8.18}$$

Next consider the rotations $\delta\mathbf{r}_i = \boldsymbol{\omega} \times \mathbf{r}_i$. Eq.(8.14) gives $\boldsymbol{\omega} \cdot \mathbf{Q} = (\boldsymbol{\omega} \times \mathbf{r}_1) \cdot \mathbf{p}_1 + (\boldsymbol{\omega} \times \mathbf{r}_2) \cdot \mathbf{p}_2$. Therefore the conserved quantity associated with rotations is the total angular momentum of the system

$$\mathbf{Q} = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2 \equiv \mathbf{L}_{tot}. \tag{8.19}$$

The total angular momentum may also be written in terms of the center of mass and relative coordinates defined in (8.7) and their canonical conjugate momenta (given in Chap.6) $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$ and $\mathbf{p} = (m_2\mathbf{p}_1 - m_1\mathbf{p}_2)/M$,

$$\mathbf{L}_{tot} = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2 = \mathbf{R} \times \mathbf{P} + \mathbf{r} \times \mathbf{p}. \tag{8.20}$$

Finally, consider the internal rotations associated with the parameter $\vec{\epsilon}$ for which $f_1^{Ia} = (m_2/M)\epsilon^{IaJ}(r_1^J - r_2^J)$ and $f_2^{Ia} = -(m_1/M)\epsilon^{IaJ}(r_1^J - r_2^J)$. The conserved quantity in this case reduces to the relative orbital angular momentum

$$\mathbf{L}_{rel} = \mathbf{r} \times \mathbf{p}. \tag{8.21}$$

Combining the above results one concludes that the center of mass and relative orbital angular momenta are independently conserved. This result is not surprising in view of the decupling of the center of mass and relative motions, as is evident from (8.8). Therefore, due to the symmetries associated with the 9 parameters $\mathbf{a}, \boldsymbol{\omega}, \vec{\epsilon}$ there are 9 conserved quantities: the total momentum $\mathbf{P}$, the total angular momentum $\mathbf{L}_{tot} = \mathbf{L}_{CM} + \mathbf{L}_{rel}$ and the relative orbital angular momentum $\mathbf{L}_{rel}$ respectively.

The more general cases involving non-trivial Λ may be illustrated through the following examples

(iv) Consider any Lagrangian in which the only time dependence comes through $\mathbf{r}_i(t)$. Then the transformation $\mathbf{r}_i'(t) = \mathbf{r}_i(t + \epsilon)$ that corresponds to a time translation $t'(t) = t + \epsilon$ is a symmetry in the sense of (8.3) with $(\partial t'/\partial t) = 1$, $\alpha = 0$. In this case $\delta\mathbf{r}_i = \epsilon\dot{\mathbf{r}}_i$ gives $\delta L = \epsilon\partial_t L$. Then eq.(8.14) becomes $\epsilon Q = \epsilon\mathbf{r}_i \cdot \mathbf{p}_i - \epsilon L$, and the conserved charge is recognized as the Hamiltonian

$$Q = H = \mathbf{r}_i \cdot \mathbf{p}_i - L.$$

Thus conservation of the total energy of a system is due to time translation symmetry.

(v) Consider the harmonic oscillator potential in the central force problem. It is sufficient to concentrate on the relative motion described in terms of the relative coordinates $L_{rel} = \frac{1}{2}\mu(\dot{\mathbf{r}}^2 - \omega^2\mathbf{r}^2)$, where ω is the frequency of oscillations. We have already seen in example (iii) that there is an internal rotational invariance that gives rise to the conservation of relative angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. In addition, this system has a symmetry under the transformation $\delta r^I = \epsilon^{IJ}\dot{r}^J$ where ϵ^{IJ} is a symmetric matrix of parameters. In this case one finds

$$\Lambda = (\mu/2)\epsilon^{IJ}(\dot{r}^I\dot{r}^J - \omega^2 r^I r^J). \tag{8.22}$$

Note that the index "a" on ϵ_a has been replaced by the double index (IJ) on ϵ^{IJ} and therefore the conserved quantity will be labelled by a symmetric tensor Q_{IJ}. Applying the general theorem one finds the conserved quantity from eq.(8.14)

$$\epsilon^{IJ}Q_{IJ} = \epsilon^{IJ}\dot{r}^J p^I - (\mu/2)\epsilon^{IJ}(\dot{r}^I\dot{r}^J - \omega^2 r^I r^J)$$

After replacing $p^I = \mu \dot{r}^I$ this becomes

$$Q_{IJ} = \frac{1}{2\mu} p^I p^J + \frac{\mu\omega^2}{2} r^I r^J \ . \tag{8.23}$$

Using the equations of motion for the harmonic oscillator one can explicitly show that indeed $dQ^{IJ}/dt = 0$ *(exercise).* The 6 conserved charges Q_{IJ} together with the three conserved angular momentum charges form the Lie algebra of SU(3) (see Chap.6, and problems at the end of this chapter). The symmetry of the Lagrangian applies also in any dimension $d = 2, 3, 4, \cdots$. In that case the symmetry is SU(d) (see below).

(vi) Consider the central force problem with the $1/r$ potential, as in the planetary problem, or as in the Hydrogen atom. Its Lagrangian that describes the relative motion is

$$L_{rel} = \frac{1}{2} m \dot{\mathbf{r}}^2 + \frac{Ze^2}{r} \tag{8.24}$$

This is evidently invariant under rotations, and therefore the relative angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is conserved as in example (iii). In addition, this system is symmetric under the transformation

$$\begin{aligned} \delta_\varepsilon \mathbf{r} &= \mathbf{r} \times (\boldsymbol{\varepsilon} \times \dot{\mathbf{r}}\,) + \boldsymbol{\varepsilon} \times (\mathbf{r} \times \dot{\mathbf{r}}) \\ \delta_\varepsilon r^I &= \varepsilon_a f^{Ia}(\mathbf{r}, \dot{\mathbf{r}}\,) = \varepsilon_a \left(\delta^{Ia} \mathbf{r} \cdot \dot{\mathbf{r}} + r^I \dot{r}^a - 2 r^a \dot{r}^I \right). \end{aligned} \tag{8.25}$$

The Lagrangian transforms to a total time derivative $\delta_\varepsilon L = \partial_t \Lambda$, where *(exercise)*

$$\Lambda = -m(\boldsymbol{\varepsilon} \times \dot{\mathbf{r}}\,) \cdot (\mathbf{r} \times \dot{\mathbf{r}}) - Ze^2 \frac{\boldsymbol{\varepsilon} \cdot \mathbf{r}}{r}. \tag{8.26}$$

Noëther's theorem gives (after using $\mathbf{p} = m\dot{\mathbf{r}}$)

$$m\mathbf{Q} = \mathbf{L} \times \mathbf{p} - mZe^2 \frac{\mathbf{r}}{r}.$$

This is known as the Runge-Lenz vector, and it is conserved *(exercise).* The three conserved charges $\mathbf{Q}$ together with three conserved angular momentum charges $\mathbf{L}$ form the Lie algebra of SO(4) (see below, and next chapter). The symmetry of the Lagrangian is also valid in any dimension $d = 2, 3, 4, \cdots$, provided one uses the second line in (8.25) that applies to any dimension. In the more general case the symmetry is SO($d+1$). This symmetry of the H-atom and its consequences will be discussed in more detail in the next chapter.

8.3 Symmetry in quantum mechanics

Let us consider the conserved quantities Q_a associated with a symmetry of the classical Lagrangian. In quantum mechanics these quantities are operators

expressed as functions of positions and momenta $Q_a(\mathbf{r}_i, \mathbf{p}_i)$. The time derivative of an operator is given by commuting it with the Hamiltonian $dQ_a/dt = i[H, Q_a]/\hbar$. Since the Q_a are time independent they must commute with the Hamiltonian

$$[Q_a, H] = 0 \ . \tag{8.27}$$

We can then conclude that every conserved "charge" Q_a, corresponding to a given symmetry of the theory, commutes with the Hamiltonian and can thus be simultaneously diagonalized with H itself.

In determining the Hilbert space on which H and Q^a operate, one can select those "charges" which commute among each other, say $Q^{a_1} \dots Q^{a_r}$, so that the states can be labelled by their eigenvalues

$$|state >= |E, q^{a_1}, \dots, q^{a_r} > \tag{8.28}$$

This implies that in the presence of a symmetry, there is *degeneracy* in the energy levels. An example is readily found in the case of rotational symmetry. A general state can be represented as $|E, l, m >$ where for the same energy level E there are many states for several values of the angular momentum eigenvalues l and m.

We have thus seen that Noether's theorem is very helpful in finding "good" quantum numbers as labels of the Hilbert space of a particular physical system. When the Hilbert space is labelled by symmetry quantum numbers, the quantum problem greatly simplifies and its interpretation becomes much more tractable. A major application of this fact is the angular momentum basis for rotationally symmetric systems.

To understand the mathematical structure we now turn our attention to another implication of Noether's theorem in Quantum Mechanics, namely that every symmetry operator Q^a is the infinitesimal generator of the corresponding transformation under which the Lagrangian and/or the Hamiltonian is invariant. This means that an infinitesimal symmetry transformation of any operator $A(\mathbf{r}_i, \mathbf{p}_i)$ in the quantum theory would be given by

$$\delta_\epsilon A = (i/\hbar)[\epsilon_a Q_a, A] \ . \tag{8.29}$$

One can easily prove that indeed the Q_a are the infinitesimal generators for the case $A = \mathbf{r}_i$ and when f_i^{Ia} are independent of velocities. To see this note that one can use the canonical commutation rules, $[p_j^J, r_i^I] = -i\hbar\delta^{IJ}\delta_{ij}$, to write

$$\delta r_i^I = (i/\hbar)[\epsilon_a Q_a, r_i^I] = (i/\hbar)\epsilon_a f_j^{Ja}[p_j^J, r_i^I] = \epsilon_a f_i^{Ia}, \tag{8.30}$$

thus reproducing the transformation law (8.1) through the quantum commutation rules. One can prove (8.29) for more general cases, but this will not be treated here (see problem on 3-dimensional harmonic oscillator).

The main lesson of the previous paragraph is that commutation with a symmetry operator Q_a has to be viewed as being closely related to applying an infinitesimal symmetry transformation. The result of commuting a symmetry operator with any other operator A is equivalent to applying an infinitesimal

symmetry transformation to the operator A. An important consequence of this is that the commutation with the Hamiltonian $[Q_a, H] = 0$ means that the Hamiltonian is invariant under infinitesimal symmetry transformations.

One may commute symmetry operators with each other $[Q_a, Q_b]$. This is equivalent to applying symmetry transformations to each other, since $\delta_\epsilon Q_b = (i/\hbar)[\epsilon \cdot Q, Q_b]$. By using Jacobi identities, one can show that the composite operator $[Q_a, Q_b]$ commutes again with the Hamiltonian

$$[[Q_a, Q_b], H] = [Q_a, [Q_b, H]] - [Q_b, [Q_a, H]] = 0, \tag{8.31}$$

since each term vanishes on the right hand side. It must be that $[Q_a, Q_b]$ is some linear combination of symmetry operators, since any operator that commutes with the Hamiltonian is a generator of symmetry transformations. Therefore, we expect the form

$$[Q_a, Q_b] = i\hbar f^c_{ab} Q_c. \tag{8.32}$$

This relation indeed holds for all the examples given in this chapter, with some set of constants f^c_{ab} *(exercise: find the constants for some of the examples).*

A relation such as (8.32), that holds for a *complete set* of operators that *close* into the same set under commutation, is called a Lie algebra. Lie algebras form the mathematical foundation of symmetry in quantum mechanics. Their study is extremely important both for the formulation of symmetries in fundamental theories and for performing computations to extract the physical information from them.

Up to now we have mainly discussed infinitesimal transformations. How about finite transformations? We will argue that for continuous groups the infinitesimal transformations contain all the information. By performing a series of infinitesimal transformations one can build up a finite transformation. For the operator A one may write the infinitesimal transformation in the form $A' = A + (i/\hbar)[\epsilon \cdot Q, A]$. To first order in ϵ this is the same as $A' = (1 + (i/\hbar)\epsilon \cdot Q)\ A\ (1 - (i/\hbar)\epsilon \cdot Q)$. For a finite transformation, one can divide it into N equal parts, apply it N times, and take the limit of large N. This amounts to

$$\begin{aligned} A' &= lim_{N\to\infty}(1 + (i/N\hbar)\epsilon \cdot Q)^N\ A\ (1 - (i/N\hbar)\epsilon \cdot Q)^N \\ &= exp(i\epsilon \cdot Q/\hbar)\ A\ exp(-i\epsilon \cdot Q/\hbar). \end{aligned} \tag{8.33}$$

Therefore, a finite transformation is applied through the exponentiation of the infinitesimal generators

$$U(\epsilon) = exp(i\epsilon \cdot Q/\hbar), \qquad A' = UAU^{-1}\ . \tag{8.34}$$

For Hermitian Q_a the U's are unitary, since upon Hermitian conjugation one finds $U^\dagger = U^{-1}$ (to prove it, think of U as an infinite series). Therefore, $UU^\dagger = 1 = U^\dagger U$. Finite transformations on states are applied according to the rule

$$|\psi >' = U|\psi > \ . \tag{8.35}$$

Under a unitary symmetry transformation the norm of the state does not change, since

$$`< \psi|\psi >' = < \psi|U^\dagger U|\psi > = < \psi|\psi > \quad . \tag{8.36}$$

From the above discussion we deduce the following important facts. They should never be forgotten:

1) The time translation generator for any quantum theory of the type given in example (iv), is the Hamiltonian of the total system, as shown in that example. A finite time translation by the amount Δt is given by $U = exp(-iH\Delta t/\hbar)$. Thus, a time translated state is $|\psi(t) >= \exp(-iH(t - t_0)/\hbar)|\psi(t_0) >$. Taking a derivative of this equation one arrives at the Schrödinger equation $i\hbar\frac{\partial}{\partial t}|\psi(t) >= H|\psi(t) >$. Thus, the time translated state is a formal solution of the Schrödinger equation.

2) The space translation generator is the *total* momentum operator of the system. For the general 2-body system this was derived in the context of example (ii). For a more complicated system, again it is the *total* momentum that generates translations on all canonical variables. A finite translation by the amount $\mathbf{a}$ on all operators and all states is given by $U = exp(i\mathbf{a} \cdot \mathbf{P}/\hbar)$.

3) The rotation generator is the *total* angular momentum operator of the system. For the two body system this was derived in example (iii), and the same result is true for a general system. When spin is included the total angular momentum $\mathbf{J}_{tot} = \mathbf{L}_{tot} + \mathbf{S}_{tot}$ is the generator. A finite rotation by the amount $\boldsymbol{\omega}$ on all operators and all states is given by $U = exp(i\boldsymbol{\omega} \cdot \mathbf{J}_{tot}/\hbar)$.

The above symmetry transformations, and other more specialized symmetry transformations that occur in specific systems, such as the Q_{IJ} of example (v), play an important role in the analysis and solution of the corresponding quantum mechanical systems.

The simplest symmetry is translations. Since the translation generator, *i.e.* total momentum, is simultaneously diagonalizable with the Hamiltonian, it is convenient to work in the basis of total momentum eigenstates.. This means plane waves for the center of mass variables. The total Hilbert space may be taken as the direct product of these plane waves with the Hilbert space for the relative motion. This result is a trivial consequence of translation symmetry which leads to the separation of center of mass variables.

The next simplest non-trivial symmetry is associated with rotations and angular momentum. For the 2-body central force problem, the *orbital* angular momentum basis for the *relative motion* is of great interest and this was the subject of the chapter on rotations. There, angular momentum was studied for an arbitrary system, including intrinsic spin.

In other chapters we also studied other examples of higher symmetry, including the cases of the two and three dimensional harmonic oscillators with

$SU(2)$ and $SU(3)$ symmetries. We will also study the case of the Coulomb force problem with $SO(4) = SU(2) * SU(2)$ symmetry. We will learn that the higher symmetry has important consequences on observable properties, for example on the degeneracy of the energy spectrum, transition rates, etc..

Finally we should mention discrete symmetries, such as parity, time reversal or charge conjugation invariances. These are distinguished from the above since for these there are no infinitesimal transformations. Rather, they are analogous to the finite transformations. The action is invariant just as in (8.3) or (8.10). Similarly, there are even more complicated discrete symmetries, such as the symmetries of crystals in condensed matter physics, which involve reflections, inversions, discrete translations and discrete rotations. Finite transformations of this type form discrete symmetry groups, while those admitting infinitesimal transformations form Lie groups.

The properties of Lie groups are deduced from the study of the corresponding Lie algebra of the type given in (8.32). All possible discrete groups and Lie groups are classified mathematically, and their properties extensively studied. The study of a Lie group and its representations is very much like solving a quantum mechanics problem in the sense of finding a complete set of eigenstates and eigenvalues that diagonalize simultaneous observables. The study of the rotation group or its Lie algebra is the study of angular momentum and spin, as was presented in a previous chapter (see also problems on $SL(2,R)$, $SU(3)$).

There are also physical theories involving symmetries between bosons and fermions. These are called supersymmetries.. The parameters of supersymmetry are bosonic (like real or complex numbers) or fermionic (Grassmann numbers).

8.4 Symmetry in time dependent systems

Up to this point we mainly discussed systems in which the time dependence appears only through the dynamical variables $x(\tau), p(\tau)$. However we also encounter physical systems that are described by a Lagrangian or Hamiltonian in which time may appear explicitly; how do we deal with symmetries for such systems? We will discuss this in the first order formalism by writing the Lagrangian in the form

$$L = \dot{x}^i p_i - H(x, p, \tau) \tag{8.37}$$

where $H(\vec{x}(\tau), \vec{p}(\tau), \tau)$ is assumed to depend on time explicitly. Consider the transformations generated by some operator $Q_\varepsilon(\vec{x}(\tau), \vec{p}(\tau), \tau)$ which may have an explicit τ dependence in addition to the implicit τ dependence through $\vec{x}(\tau)$ and $\vec{p}(\tau)$. The transformations of $\vec{x}, \vec{p}$ are obtained by computing the Poisson bracket at equal τ

$$\delta_\varepsilon x_i(\tau) = \{x_i(\tau), Q_\varepsilon(x(\tau), p(\tau), \tau)\}, \quad \delta_\varepsilon p_i(\tau) = \{p_i(\tau), Q_\varepsilon(x(\tau), p(\tau), \tau)\} \tag{8.38}$$

In evaluating the Poisson bracket τ is treated like a parameter. Then these give

$$\delta_\varepsilon x_i = \frac{\partial Q_\varepsilon}{\partial p_i}, \ \delta_\varepsilon p_i = -\frac{\partial Q_\varepsilon}{\partial x_i}. \tag{8.39}$$

Using these we compute the transformation of the Lagrangian $L = \dot{x}_i p_i - H(x,p,\tau)$ and obtain

$$\begin{aligned}
\delta_\varepsilon L &= \frac{d(\delta_\varepsilon x_i)}{d\tau}p_i + \frac{dx_i}{d\tau}\delta_\varepsilon p_i - \frac{\partial H}{\partial x_i}\delta_\varepsilon x_i - \frac{\partial H}{\partial p_i}\delta_\varepsilon p_i && (8.40)\\
&= \frac{d\left(\frac{\partial Q_\varepsilon}{\partial p_i}\right)}{d\tau}p_i - \frac{dx_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial Q_\varepsilon}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial Q_\varepsilon}{\partial x_i} && (8.41)\\
&= \frac{d}{d\tau}\left(p_i\frac{\partial Q_\varepsilon}{\partial p_i}\right) - \frac{dp_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial p_i} - \frac{dx_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial x_i} - \{H, Q_\varepsilon\} && (8.42)\\
&= \frac{d}{d\tau}\left(p_i\frac{\partial Q_\varepsilon}{\partial p_i} - Q_\varepsilon\right) + \frac{\partial Q_\varepsilon}{\partial \tau} + \{Q_\varepsilon, H\}. && (8.43)
\end{aligned}$$

In going from the third line to the fourth line we used the fact that the *total* time derivative of $Q_\varepsilon(\vec{x}(\tau),\vec{p}(\tau),\tau)$ indicated by $\frac{dQ_\varepsilon}{d\tau} = \frac{\partial Q}{\partial \tau} + \frac{dp_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial p_i} + \frac{dx_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial x_i}$ is obtained by differentiating with respect to the explicit τ dependence (partial derivative $\frac{\partial Q}{\partial \tau}$) as well as the implicit τ dependence (chain rule $\frac{dp_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial p_i} + \frac{dx_i}{d\tau}\frac{\partial Q_\varepsilon}{\partial x_i}$).

Since the Hamiltonian is the generator of infinitesimal time translations on the canonical variables, namely $\{x_i, H\} = \partial_\tau x_i$, $\{p_i, H\} = \partial_\tau p_i$, and $\{Q_\varepsilon, H\} = \frac{\partial Q_\varepsilon}{\partial x_i}\partial_\tau x_i + \frac{\partial Q_\varepsilon}{\partial p_i}\partial_\tau p_i$, the last two terms in Eq.(8.43) represent the total time derivative of Q_ε, namely $\frac{dQ_\varepsilon}{d\tau} = \frac{\partial Q_\varepsilon}{\partial \tau} + \{Q_\varepsilon, H\}$. If there is a symmetry, then Q_ε is conserved, which means its *total* time derivative $\frac{dQ_\varepsilon}{d\tau}$ must vanish *when the equations of motion are used*. However, to show that there is a symmetry, one must show that $\frac{\partial Q_\varepsilon}{\partial \tau} + \{Q_\varepsilon, H\}$ vanishes *without using the equations of motion* for $x(\tau), p(\tau)$. If that works, then we see from Eq.(8.43) that $\delta_\varepsilon L$ reduces to a total τ derivative $\delta_\varepsilon L = \frac{d\Lambda}{d\tau}$ with

$$\Lambda = p_i\frac{\partial Q_\varepsilon}{\partial p_i} - Q_\varepsilon, \quad \rightarrow \quad Q_\varepsilon = p_i\frac{\partial Q_\varepsilon}{\partial p_i} - \Lambda \tag{8.44}$$

Recalling that Q_ε was defined as $Q_\varepsilon = p_i\delta_\varepsilon x_i - \Lambda$, we check the consistency when the expressions above are inserted. Namely we can verify that $Q_\varepsilon = p_i\delta_\varepsilon x_i - \Lambda = p_i\left(\frac{\partial Q_\varepsilon}{\partial p_i}\right) - \left(p_i\frac{\partial Q_\varepsilon}{\partial p_i} - Q_\varepsilon\right) = Q_\varepsilon$ is consistent, as expected.

In conclusion, there is a symmetry whenever one can find a $Q_\varepsilon(x(\tau),p(\tau),\tau)$ whose equal τ Poisson bracket with $H(x(\tau),p(\tau),\tau)$ satisfies

$$\frac{\partial Q_\varepsilon}{\partial \tau} + \{Q_\varepsilon, H\} = 0. \tag{8.45}$$

without using the equations of motion. It should be emphasized that $\frac{\partial Q_\varepsilon}{\partial \tau}$ is the τ-derivative with respect to the explicit τ only. This characterization of a symmetry is valid whether H depends on τ explicitly or not, and can be taken as a more general statement of a symmetry. The generators of the symmetry Q_a are the coefficients of the τ-independent parameters ε^a in the expansion

$$Q_\varepsilon(x(\tau),p(\tau),\tau) = \varepsilon^a Q_a(x(\tau),p(\tau),\tau). \tag{8.46}$$

The $Q_a\left(x\left(\tau\right),p\left(\tau\right),\tau\right)$ should form a Lie algebra under the equal τ Poisson bracket at any τ in the classical theory, and after operator ordering, under the equal τ Lie bracket at any τ in the quantum theory.

8.5 A brief tour of Lie groups

Symmetry transformations form a group. What is a group? A group G is a set of elements $G=\{g\}$, together with an operation symbolically denoted by a dot $\{\cdot\}$, that satisfies the properties of (1) closure, (2) associativity, (3) identity, (4) inverse. Closure means that under the operation two elements in G give a third element in G, that is

$$g_1\cdot g_2=g_3\ ,\qquad and\qquad g_3\in G\ . \tag{8.47}$$

Associativity requires that, when three elements are combined under the operation, one obtains the same result if combined in either order according to the parentheses below

$$(g_1\cdot g_2)\cdot g_3=g_1\cdot(g_2\cdot g_3). \tag{8.48}$$

However, one is not allowed to change the order of the elements 1,2,3 , unless the operation happens to be commutative. The set must include the identity element, designated by **1**, whose property is

$$1\cdot g=g=g\cdot 1\ . \tag{8.49}$$

Finally, the inverse of each element must be included in the set G. When an element g is combined with its inverse g^{-1} they must produce the identity element:

$$g\cdot g^{-1}=1=g^{-1}\cdot g\ . \tag{8.50}$$

Now consider the set of all symmetry transformations that may be applied on a physical system. More precisely, consider the action of the system that remains invariant under the symmetries. To comprehend it better, it may help to keep the examples of translations and/or rotations in mind. It is physically evident that two symmetry transformations applied one after the other may be regarded as a single symmetry transformation that leaves the Lagrangian invariant. Therefore, symmetries close under this combination into another symmetry transformation. Applying the symmetry transformation backward is equivalent to the inverse element, and applying no transformation at all is equivalent to the identity element. Therefore, the inverse and identity elements are in the set of symmetries. Finally, one must check associativity of symmetry transformations to insure that one has a group. Certainly for translations and rotations this is intuitively evident, but more generally the product of operators (applied on states or operators) is an associative multiplication in quantum mechanics .

When the symmetries have continuous parameters, they can be studied in the neighborhood of the identity. These are called Lie groups. As we saw an

infinitesimal transformation is generated by the conserved charges Q_a, and we may write the symmetry transformation on any operator or state in terms of $U(\boldsymbol{\epsilon}) = 1 + i\boldsymbol{\epsilon}\cdot Q/\hbar$. The finite symmetry transformation is given by $U(\boldsymbol{\epsilon}) = exp(i\boldsymbol{\epsilon}\cdot Q/\hbar)$. Therefore these operators must satisfy the group properties when applied on states (as in eq.(8.35)) or operators (as in eq.(8.34))

$$\begin{array}{ll} 1- & U(\boldsymbol{\epsilon}_1)U(\boldsymbol{\epsilon}_2) = U(\boldsymbol{\epsilon}_3)\ , \\ 2- & [U(\boldsymbol{\epsilon}_1)U(\boldsymbol{\epsilon}_2)]U(\boldsymbol{\epsilon}_3) = U(\boldsymbol{\epsilon}_1)[U(\boldsymbol{\epsilon}_2)U(\boldsymbol{\epsilon}_3)], \\ 3- & U(0) = 1\ , \\ 4- & U(-\boldsymbol{\epsilon}) = U(\boldsymbol{\epsilon})^\dagger = U(\boldsymbol{\epsilon})^{-1};\ U(\boldsymbol{\epsilon})U(-\boldsymbol{\epsilon}) = 1\ . \end{array} \tag{8.51}$$

The properties of specific groups are hidden in the closure property (1), and by specifying the combination of two parameters into a third one with the functions $\boldsymbol{\epsilon}_3^a(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)$, one completely determines the group. By taking infinitesimal $(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)$ and expanding to lowest order one can give a general relation

$$\vec{\epsilon}_3^{\,a} = \vec{\epsilon}_1^{\,a} + \vec{\epsilon}_2^{\,a} + \frac{1}{2} f^a_{bc} \vec{\epsilon}_1^{\,b} \vec{\epsilon}_2^{\,c} + \cdots \tag{8.52}$$

in terms of the constants f^a_{bc}. These constants are called the structure constants of the Lie group, and they carry all the information about the group. To see their significance consider the product of two symmetry operators in the form $e^A e^B$. For small A and B one has

$$e^A e^B = 1 + A + B + AB + \frac{1}{2}A^2 + \frac{1}{2}B^2 + \cdots = exp(A + B + \frac{1}{2}[A,B] + \cdots). \tag{8.53}$$

Now use $A = i\boldsymbol{\epsilon}_1 \cdot Q/\hbar$, $B = i\boldsymbol{\epsilon}_2 \cdot Q/\hbar$ and compare the first line of (8.51) to (8.53). Using the definition of $\boldsymbol{\epsilon}_3$ in (8.52) one finds that the generators of symmetry transformations must satisfy the Lie algebra

$$[Q_a, Q_b] = i\hbar f^c_{ab} Q_c\ . \tag{8.54}$$

with the same structure constants given in (8.52).

The expansion in (8.53) can be continued to higher orders $exp(A)\ exp(B) = \exp(C)$, and it will be seen that the higher powers of A or B contained in C come always with commutators, e.g. $[A,[A,B]]$ or $[B,[B,A]]$ etc. *(exercise).* Therefore, the Lie algebra (8.54) contains all the information about the closure of the group; thus, knowing only the Lie algebra is sufficient to fully construct $\boldsymbol{\epsilon}_3(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)$ to all orders. For this reason, the group property is the same for any representation of the generators that satisfy the Lie algebra. In particular, the smallest matrix representation of the Lie group (given below) is useful for the explicit computation of $\boldsymbol{\epsilon}_3(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)$.

To impose associativity (second line of eq.(8.51)) one must consider three infinitesimal transformations. In quantum mechanics products of operators is an associative multiplication. Therefore, this requirement is automatically satisfied. Using this same associativity of products of operators one can prove the Jacobi identity

$$[Q_a,[Q_b,Q_c]+[Q_b,[Q_c,Q_a]+[Q_c,[Q_a,Q_b]=0 \tag{8.55}$$

by simply opening up the brackets. However, a second way of evaluating this identity is to use the Lie algebra 8.54 repeatedly.. Then we obtain the following condition on the structure constants

$$f^d_{ak}f^k_{bc}+f^d_{bk}f^k_{ca}+f^d_{ck}f^k_{ab}=0\ , \tag{8.56}$$

which is valid for any value of the indices (a,b,c,d). The two equations (8.54) and (8.56) define a Lie algebra.

All possible Lie algebras were classified by Cartan in the 19^{th} century. His work consists of finding all possible solutions to the Jacobi identities in (8.56) and providing explicit sets of constants f^c_{ab} that determine the Lie algebra. Therefore all possible continuous symmetry transformations are mathematically classified and their properties studied. Cartan identified a set of Lie algebras called "simple Lie algebras". These form the building blocks of all symmetries among bosons or among fermions. In addition there are superalgebras discovered during the 1970's that are relevant for supersymmetries between bosons and fermions. For supersymmetry there is also a set of "simple superalgebras". All remaining Lie algebras or superalgebras can be constructed by either taking direct sums of "simple" ones (then they are called "semi-simple"), or by taking certain limits called "contractions" (then they are called "solvable"). Here we provide the list of compact simple Lie algebras given by Cartan. Their non-compact versions (i.e. $SU(n,m)$, $SO(n,m)$ etc.) are obtained by analytic continuation:

algebra	**group**	**rank (N)**	**dimension (D)**
A_N	$SU(N+1)$	$1,2,3,\cdots,\infty$	$(N+1)^2-1$
B_N	$SO(2N+1)$	$1,2,3,\cdots,\infty$	$N(2N+1)$
C_N	$Sp(2N)$	$1,2,3,\cdots,\infty$	$N(2N+1)$
D_N	$SO(2N)$	$1,2,3,\cdots,\infty$	$N(2N-1)$
E_N	E_6,E_7,E_8	$6,7,8$	$78,133,248$
F_4	F_4	4	54
G_2	G_2	2	14

(8.57)

The value of N is the number of commuting generators, and it is called the "rank", while the value of D is the number of total generators in the Lie algebra and it is called its "dimension". One may consider functions of the generators that commute with each other as well as with all the generators. These are called Casimir operators. The number of independent Casimir operators also coincides with the rank N. In the above notation there are a few Lie algebras that are identical with each other. These identities are

$$\begin{aligned} &A_1=B_1=C_1,\\ &B_2=C_2,\\ &D_2=A_1\oplus A_1,\\ &A_3=D_3, \end{aligned} \tag{8.58}$$

while all others are independent.

As an example consider the familiar rotation group in 3-dimensions. Because of the equivalence, this corresponds to $SU(2) \sim SO(3) \sim Sp(2)$ or $A_1 = B_1 = C_1$. It has one commuting operator which may be chosen as J_3 and has one Casimir operator which is $\vec{J} \cdot \vec{J}$.

To make the classification of symmetries more concrete it is useful to consider the following analysis, due to Chevalley, that provides the smallest *representation* of most of the above groups and algebras in terms of matrices.

The rotation group $SO(3)$ can be characterized by saying that it leaves invariant dot products of any two vectors $\mathbf{r}_1 \cdot \mathbf{r}_2$. Similarly, the group $SO(d)$ can be characterized as the transformations that leaves invariant dot products of real vectors in d-dimensions. Consider the transformations on such a vector written in the form of a column $x' = Rx$, or in term of components

$$x'_I = R_{IJ} x_J, \qquad I = 1, 2, \cdots d \ . \tag{8.59}$$

The dot product may be written as a product of a row vector, which is the transpose vector, with a column vector, in the form $x^T y = x_1 y_1 + \cdots + x_d y_d$. Therefore the invariance requires that R is an orthogonal matrix:

$$x^T y = x'^T y' = x^T R^T R y \qquad \rightarrow R^T R = 1 \ . \tag{8.60}$$

Orthogonal matrices give again orthogonal matrices under matrix multiplication. Therefore, they close into the same set. Furthermore, matrix multiplication is associative, there is a unit matrix, and there is an inverse. Therefore orthogonal matrices in d-dimensions form a group, namely $SO(d)$.

If R is a transformation connected to the identity infinitesimally, then it must be possible to expand it in the form $R = 1 + A + \cdots$, with a matrix A_{IJ} that contains infinitesimal parameters. Imposing the orthogonality condition requires an antisymmetric A. Thus, for a finite transformation we can write

$$orthogonal \quad \rightarrow R = e^A, \qquad A^T = -A \ . \tag{8.61}$$

Now counting the number of independent parameters contained in A, one finds the dimension of the Lie algebra $SO(d)$, namely $D = d(d-1)/2$, which coincides with the numbers given in Cartan's table for $SO(2N+1)$ and $SO(2N)$. Two antisymmetric matrices close under commutation $[A_1, A_2]$ into another antisymmetric matrix *(exercise)*; therefore, the Lie algebra is closed. Associativity, and hence Jacobi identity is a property of matrix multiplication, so (8.56) is automatically satisfied.

The same argument may now be used for complex vectors in d dimensions, $z_I, \ I = 1, 2, \cdots d$. The transformation is $z' = Uz$. But now define the dot product by taking the hermitian conjugate $z^\dagger w = z_1^* w_1 + \cdots + z_d^* w_d$, and demand invariance. This requires a unitary matrix, $U^\dagger U = 1$. An infinitesimal expansion is given by $U = 1 + H + \cdots$, and unitarity requires an anti-Hermitian matrix H. Furthermore, if U has unit determinant, then H must be traceless. Unitary matrices of determinant one give again matrices in the same set under ordinary

matrix multiplication. The associativity, unity and inverse conditions are also satisfied *(exercise)*. Therefore such matrices form a group. Such unitary transformations are called $SU(d)$, or special unitary transformations in d-dimensions (special means that the determinant is 1). They are characterized by

$$unitary \quad \rightarrow U = e^H, \qquad H^\dagger = -H, \qquad tr(H) = 0 \ . \tag{8.62}$$

The anti-hermiticity condition requires purely imaginary entries on the diagonal $H_{aa} = iy_a$, and related complex entries above and below the diagonal $H_{12} = -H_{21}^* = x_{12} + iy_{12}$, etc.. Furthermore, the trace condition allows only $(d-1)$ independent diagonal entries. The number of independent *real* parameters in such an H is $(d-1) + 2 \times d(d-1)/2 = d^2 - 1$, and this agrees with Cartan's table for $d = N+1$. Two arbitrary anti-hermitian traceless matrices close under commutation $[H_1, H_2]$ into another anti-hermitian traceless matrix; therefore the Lie algebra (8.54) and Jacobi identities (8.56) are satisfied automatically.

Next consider quaternions instead of real or complex numbers of the previous cases. A quaternion is a generalization of a complex number; it has one real and three imaginary directions. It may be written in terms of 2×2 Pauli matrices $q = a + i\boldsymbol{\sigma} \cdot \mathbf{b}$. The quaternionic conjugate, which generalizes complex conjugate, is $\bar{q} = a - i\boldsymbol{\sigma} \cdot \mathbf{b}$. Using quaternionic vectors q_I in d dimensions, and taking the transformation $q' = Sq$ and a dot product in the form $\bar{q}^T q = \bar{q}_1 q_1 + \cdots + \bar{q}_d q_d$, we may repeat the analysis above. The result is $\bar{S}^T = S^{-1}$. The infinitesimal form of such a matrix $S = 1 + Q + \cdots$ with the condition that the matrix Q is odd under the combined transposition and quaternionic conjugation. This gives the symplectic group $Sp(2d)$:

$$symplectic \quad \rightarrow S = e^Q, \qquad \bar{Q}^T = -Q \ . \tag{8.63}$$

As in the previous cases such quaternionic matrices (S, Q) form a Lie group and algebra. The conditions on the matrix Q require that the diagonal is purely quaternion imaginary $Q_{aa} = i\boldsymbol{\sigma} \cdot \mathbf{y}_a$, and that the elements above and below the diagonal are related, $Q_{12} = -\bar{Q}_{21} = x_{12} + i\boldsymbol{\sigma} \cdot \mathbf{y}_{12}$, etc.. Therefore, the number of independent *real* parameters is $3d + 4 \times d(d-1)/2 = d(2d+1)$, in agreement with Cartan's table for $Sp(2d)$. Two arbitrary quaternionic matrices Q that are odd under the combined transposition and quaternionic conjugation close under commutation $[Q_1, Q_2]$ into the same set of matrices, therefore the Lie algebra (8.54) and Jacobi identities (8.56) are satisfied.

To understand the groups E_6, E_7, E_8, F_4, G_2 it is useful to consider octonions, which is a generalization of complex and quaternionic numbers. But the analysis is complicated (because octonionic multiplication is not associative), and will not be given here since it is not needed in this book. Another topic that will not be discussed further in this chapter is superalgebras, but we will come back to them in examples in later chapters.

The following subgroup and sub-algebra structures can be easily determined:

1– $SO(d-1)$ is a subgroup of $SO(d)$; $SU(d-1)$ is a subgroup of $SU(d)$; $Sp(2d-2)$ is a subgroup of $Sp(2d)$. This is evident since one can choose

to make transformations on the first $(d-1)$ components of a vector while leaving the last component unchanged.

2– $d\times d$ anti-symmetric matrices A are a subset of $d\times d$ anti-hermitian matrices H; therefore $SO(d)$ is a subgroup of $SU(d)$.

3– If quaternions are represented by 2×2 matrices, as above, the size of the matrices are really $2d\times 2d$, and transposition combined with quaternionic conjugation becomes equivalent to ordinary hermitian conjugation. Such quaternionic $2d\times 2d$ matrices Q form a subset of $2d\times 2d$ anti-hermitian matrices H; hence $Sp(2d)$ is a subgroup of $SU(2d)$.

4– $SO(2d)$ as well as $Sp(2d)$ have an $SU(d)$ subgroup.

There are further subgroup/sub-algebra structures that are harder to see or to explain, and they will not be dealt with here.

8.6 SL(2,R) and its representations

SU(2) is the simplest non-Abelian group. It can be interpreted as the rotation group in three dimensions SO(3) . The properties of SU(2) and its representations have been discussed in previous chapters. SL$(2,R)$ is equivalent to SU$(1,1)$ or to SO$(2,1)$ which has the interpretation of the Lorentz group in 2 space and 1 time dimensions. It can also be thought of as the conformal group in 0 space and one time dimensions. Just like SU(2) the group SL$(2,R)$ has a large number of applications in Physics, with interpretations which are not necessarily related to space-time transformations. The parameters of the rotation group are angles, and the group element of SU(2) , i.e. the $D^j_{mm'}(\boldsymbol{\omega})$, is periodic when the angles are changed by 2π when j is integer, or by 4π when j is half integer. Thus, SU(2) is a compact group since its parameters have a finite range. SL(2,R) is the simplest non-compact group. Two of its parameters have infinite range while the third one has a compact range. SL$(2,R)$ can be formally thought of as an analytic continuation of SU(2) , but its *unitary* representations are not analytic continuations of the representations of SU(2). We will discuss here the unitary representations of SL$(2,R)$ and we will show that in some representations SL(2,R) is closely connected to the Hydrogen atom.

The SL(2,R) Lie Algebra is given by

$$[J_0,J_1]=iJ_2 \quad , \quad [J_0,J_2]=-iJ_1 \quad , \quad [J_1,J_2]=-iJ_0 \tag{8.64}$$

This algebra is related to the SU(2) Lie algebra by the analytic continuation $J_{1,2}\rightarrow iJ_{1,2}$ and $J_3\rightarrow J_0$. Equivalently one can think that the parameters associated with these generators are analytically continued. J_0 is compact, the other two are non-compact. It is useful to think of J_0 as pointing in the time-like direction and of $J_{1,2}$ as pointing in two space-like directions.

One can define $J_\pm=J_1\pm iJ_2$ and rewrite the commutation rules in the form

$$[J_0,J_\pm]=\pm J_\pm, \quad [J_+,J_-]=-2J_0. \tag{8.65}$$

These differ from those of SU(2) only by the minus sign in the second equation. The representations of $SL(2,R)$ can be labelled by the eigenvalues of the quadratic Casimir operator $C_2 = J_0^2 - J_1^2 - J_2^2 = j\,(j+1)$ and those of the compact generator $J_0 = m$. These states labelled as $|jm>$ have properties closely related to those of SU(2). These representations were first worked out by Bargmann and they have been extensively studied in the literature. The problems at the end of this chapter provide a guide for deriving them by close analogy to SU(2) representations.

A particular realization of this algebra was given in terms of oscillators in the problem section of chapter 5. The reader is advised to review that exercise and understand that it corresponds to special value of j. A generalization of that oscillator representation is found in Chapter 9, section 9.4.1 for more general values of j that correspond to a subset of representations called the discrete series (see below).

It is also useful rewrite the commutation rules in terms of the lightcone type combinations

$$[J_0 + J_1\,,\, J_0 - J_1] = -2iJ_2 \quad , \quad [J_2\,,\, J_0 \pm J_1] = \pm i(\,J_0 \pm J_1). \tag{8.66}$$

In this section we will explore this approach in detail and obtain all representations of $SL(2,R)$ in a different and interesting basis that has a close connection to the H-atom.

8.6.1 A construction

Introduce a canonical set of variables $[q,p] = i$ and two constants s, σ. We now construct a representation of the generators in terms of these

$$\begin{aligned} J_0 + J_1 &= p \\ J_0 - J_1 &= qpq + 2qs + \frac{\sigma^2}{p} \\ J_2 &= \frac{1}{2}\,(qp+pq) + s \end{aligned} \tag{8.67}$$

Here q and p are Hermitian operators and s, σ are real parameters. Hence *the generators are Hermitian*. By explicit commutation we verify that the commutation rules for the currents are indeed satisfied

$$\begin{aligned} [J_0 + J_1\,,\, J_0 - J_1] &= \left[p\,,\, qpq + 2qs + \frac{\sigma^2}{p}\right] \\ &= -2i\left(\frac{1}{2}\,(qp+pq) + s\right) \\ &= -2iJ_2 \end{aligned} \tag{8.68}$$

and

$$\begin{aligned}
[J_2\,,\,J_0+J_1] &= \left[\frac{1}{2}\,(qp+pq)+s\,,\,p\right] = ip = i\,(J_0+J_1) \qquad (8.69)\\
[J_2\,,\,J_0+J_1] &= \left[\frac{1}{2}\,(qp+pq)+s\,,\,qpq+2qs+\frac{\sigma^2}{p}\right]\\
&= -i\left(qpq+2qs+\frac{\sigma^2}{p}\right) = -i\,(J_0-J_1)\,.
\end{aligned}$$

8.6.2 Casimir

The quadratic Casimir operator is computed as

$$\begin{aligned}
J_0^2-J_1^2-J_2^2 &= \frac{1}{2}\,(J_0+J_1)\,(J_0-J_1)+\frac{1}{2}\,(J_0-J_1)\,(J_0+J_1)-J_2^2\\
&= \frac{1}{2}p\left(qpq+2qs+\frac{\sigma^2}{p}\right)+\frac{1}{2}\left(qpq+2qs+\frac{\sigma^2}{p}\right)p-\\
&\qquad -\left(\frac{1}{2}\,(qp+pq)+s\right)^2\\
&= \frac{1}{2}(pqpq+qpqp)-\frac{1}{4}(pqpq+qpqp+qppq+pqqp)+\sigma^2-s^2\\
&= \sigma^2-s^2+\frac{1}{4}\,[q,p]\,[q,p]\\
&= \sigma^2-s^2-\frac{1}{4} \qquad (8.70)
\end{aligned}$$

Thus, we may identify

$$\begin{aligned}
j(j+1) &= \sigma^2-s^2-\frac{1}{4} \qquad (8.71)\\
j &= -\frac{1}{2}\pm\sqrt{\sigma^2-s^2}
\end{aligned}$$

We will see later that the unitary representations (with normalizable states) are uniquely identified as follows

$$\begin{aligned}
-\infty &< j(j+1) < -\frac{1}{4}, \;\rightarrow j=-\frac{1}{2}+i\sqrt{s^2-\sigma^2} && \text{principal series, } s^2-\sigma^2>0\\
-\frac{1}{4} &< j(j+1) < 0, \;\rightarrow j=-\frac{1}{2}+\sqrt{\sigma^2-s^2} && \text{supplementary or discrete series, } 0<\sigma^2-s^2<\tfrac{1}{4}\\
0 &< j(j+1) < \infty, \;\rightarrow j=-\frac{1}{2}+\sqrt{\sigma^2-s^2} && \text{discrete series } \sigma^2-s^2\geq\tfrac{1}{4} \qquad (8.72)
\end{aligned}$$

That is, considering all real values of $j(j+1)$, for $j(j+1)<-1/4$ one finds the principal series, for $-1/4<j(j+1)<0$ one finds the supplementary series

or the discrete series, and for $j(j+1) \geq 0$ there is only the discrete series. So, the discrete series occurs for all $\sigma^2 \geq s^2$ or $j \geq -1/2$. Note that there is a non-trivial representation for $j = 0$. The value of j is complex only for the principal series, although $j(j+1)$ is real in all cases, since it is the eigenvalue of a Hermitian operator. Weyl reflections that flip the sign of the square roots produce equivalent representations, therefore it is sufficient to choose one sign of the square root to identify the value of j, as a convention, as we have done above.

8.6.3 Wavefunction and unitarity

One may work in the vector space that diagonalizes momentum $|p>$. As in usual quantum mechanics this is a complete and orthonormal set of states

$$< p|p' >= \delta\left(p - p'\right), \qquad \int_{-\infty}^{+\infty} dp\, |p >< p| = 1 \tag{8.73}$$

The operator q acts as

$$< p|\, q = i\partial\ < p| \tag{8.74}$$

An arbitrary normalizable state in the Hilbert space may be expanded in this basis

$$|\psi >= \int_{-\infty}^{+\infty} dp\, |p >\ \psi(p), \qquad \psi(p) =< p|\psi > . \tag{8.75}$$

In particular one complete set of normalizable states is obtained by diagonalizing the compact generator $J_0 = \frac{1}{2}\left(p + qpq + 2qs + \frac{\sigma^2}{p}\right)$

$$J_0|jm >= m|jm > . \tag{8.76}$$

These states may be expanded in the momentum basis, with the normalization condition specified as

$$\begin{aligned} |jm \ > \ &= \int_{-\infty}^{+\infty} dp\, |p >\ \psi_m(p), \qquad \psi_m(p) =< p|jm > . \\ < \ jm|jm' >&= \delta_{mm'} = \int_{-\infty}^{+\infty} dp\, \psi_m^*(p)\psi_{m'}(p) \end{aligned} \tag{8.77}$$

The eigenvalue condition provides a differential equation for the wavefunction

$$-\partial\left(p\partial\psi_m\right) + 2is\partial\psi_m + \left(p + \frac{\sigma^2}{p} - 2m\right)\psi_m = 0. \tag{8.78}$$

This equation simplifies by defining

$$\begin{aligned} \psi_m(p) \ &= \ p^{is-1/2}\phi_m(p) \\ \delta_{mm'} \ &= \ \int_{-\infty}^{+\infty} \frac{dp}{|p|}\, \phi_m^*(p)\phi_{m'}(p) \end{aligned} \tag{8.79}$$

Then the equation takes the same form as the radial equation for the Hydrogen atom

$$\left(-\partial^2 + \frac{j(j+1)}{p^2} - \frac{2m}{p} + 1\right)\phi_m(p) = 0. \tag{8.80}$$

Near $p = 0$ the two independent solutions behave as p^{j+1} or p^{-j}. The real part of the power must be positive otherwise the wavefunction is not normalizable. With the definition of j given in eq.(8.72) the unique solution is picked as the one that behaves as p^{j+1} when $\sigma^2 - s^2 > 1/4$, while both solutions are admissible for all other cases.

In solving the equation we take advantage of the analogy to the H-atom. In comparison to the H-atom equation we have $mass = 1/2$, angular momentum= j, Coulomb potential with $Ze^2 = 2m$, and energy $E = -1$. However, there are also some important differences: (i) For the H-atom the radial coordinate is positive, while here the momentum takes both positive and negative values; (ii) Furthermore, the angular momentum term $j(j+1)$ is positive for the H-atom, but here it can take either sign.

Note that we are seeking normalizable states with negative "energy" $E = -1$. This can occur only if the effective potential $V(p) = \frac{j(j+1)}{p^2} - \frac{2m}{p}$ is attractive (i.e. negative) in a finite range of p. The possible cases are given in Fig.1.

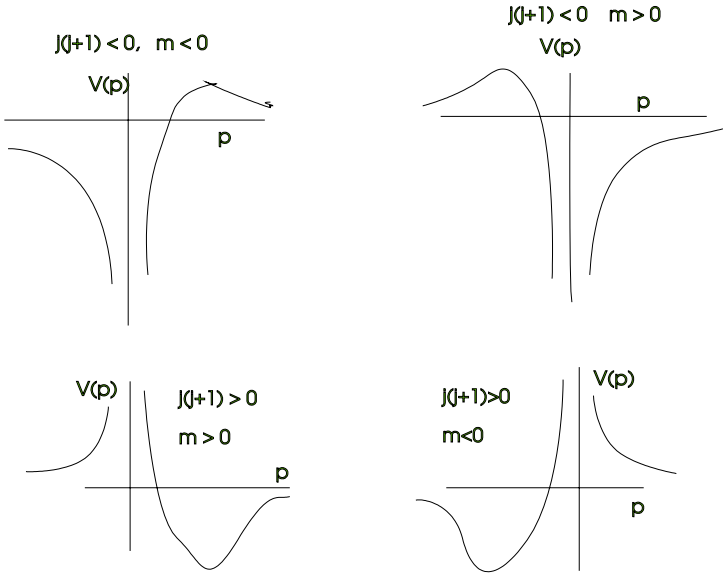


Fig.1 - Effective potential.

If $j(j+1) = \sigma^2 - s^2 - \frac{1}{4}$ is positive, then there are bound states only if $2m/p$ is positive. Therefore, if $m > 0$ the momentum must also be positive, and if $m < 0$ then the momentum must be negative. This is the case for the discrete series. This case is exactly analogous to the H-atom that has only positive radial coordinate. We can simply take over the H-atom solutions that are given

in terms of Laguerre polynomials. The bound state energy for the H-atom is

$$E_n = -\frac{mass}{2}\frac{(Ze^2)^2}{(j+n+1)^2} \tag{8.81}$$

where $n = 0, 1, 2, \cdots$ is the radial quantum number (and we used $\hbar = c = 1$). If we substitute the quantities in our equation we obtain

$$-1 = -\frac{1}{4}\frac{(2m)^2}{(j+n+1)^2}. \tag{8.82}$$

Therefore, the allowed quantum numbers are

$$|m| = j + 1 + n, \quad \text{with } j + 1 = 1/2 + \sqrt{\sigma^2 - s^2}. \tag{8.83}$$

This is in agreement with the well known result for the discrete series (both positive and negative).

The analysis of the H-atom is conducted for $j(j+1) > 0$. But it extends also to the case $-1/4 < j(j+1)$ provided one concentrates on the case of $sign(m) = sign(p)$ and keep only the solution that behaves like p^{j+1} near the origin. Then the form of the solution is identical to the one just given above, but now with $-1/2 < j$ or all values of $\sigma^2 > s^2$. These are all the discrete series representations. We emphasize that for all discrete series representations the wavefunction $\psi_m(p) =< p|m >$ is non-vanishing only when the sign of m and p are the same, and it is given in terms of the H-atom wavefunctions.

If $j(j+1) = \sigma^2 - s^2 - \frac{1}{4}$ is negative, then the angular momentum term is an attractive potential that swamps the Coulomb term near the origin. Then we can obtain bound states with negative energies for any sign of m as well as any sign of p by allowing both solutions that behave like p^{j+1} or p^{-j} near $p = 0$. This is the case for the principal and supplementary series. By studying the orthonormal properties of the states we discover that the allowed values of m are given by $m = |m_0| + n$ where n is any positive or negative integer, and m_0 may be chosen in the range $0 < |m_0| < 1$ except for the special values $j + 1$ or $-j$.

8.7 PROBLEMS

1. The Lorentz group in one-time and two-space dimensions has infinitesimal generators that are somewhat similar to rotations in three dimensions, but with some sign changes. To see this, introduce canonical conjugates $[x_\mu, p_\nu] = i\hbar g_{\mu\nu}$ where $g_{\mu\nu} = diag\,(-1, 1, 1)$ is the Minkowski metric. Then consider the Lorentz operators $L_{\mu\nu} = x_\mu p_\nu - x_\nu p_\mu$ that are analogous to orbital angular momentum except for the sign difference introduced through the metric $g_{\mu\nu}$ instead of $\delta_{\mu\nu}$. Show that the commutation rules of the $L_{\mu\nu}$ among themselves may be rewritten in the form

$$[J_+, J_-] = -2J_0, \quad [J_0, J_\pm] = \pm J_\pm. \tag{8.84}$$

and give the explicit forms of J_0 and $J_\pm = (J_1 \pm iJ_2)$ in terms of the x_μ, p_μ. Note that, as compared to the angular momentum algebra there is only one sign change in the first commutator. This is the Lie algebra of SL(2,R)=SU(1,1) instead of SU(2). Show that the quadratic Casimir operator that commutes with all the generators $J_{0,\pm}$ is

$$\begin{aligned} C &= J_0^2 - (J_+J_- + J_-J_+)/2 \\ &= J_0^2 - J_1^2 - J_2^2. \end{aligned} \tag{8.85}$$

2. The quantum mechanical states for the general SL(2,R) algebra in (8.84) may be found with the same methods that were used in the analysis of angular momentum. In this case the simultaneous eigenstates of C and J_0 on states labelled as $|jm>$ are

$$J_0|jm> = \hbar m|jm>, \quad C|jm> = \hbar^2 j(j+1)\,|jm>\,. \tag{8.86}$$

Note that C is not a positive definite operator, but is Hermitian. Find all possible values of j, m that are consistent with hermiticity. of J_1, J_2, J_0 and unitarity (i.e. states of positive norm). Note that there are three (or four, depending on how one counts) independent solutions. These are called the principal series (complex j =?), the supplementary series real (j =?) and the discrete series (two of them) (real j =?). What are the values of m in each case?

3. Show that the following set of operators form the SL(2,R) Lie algebra

$$\begin{aligned} J_0 - J_1 &= x \\ J_2 &= -\frac{1}{2}(xp + px) - s \\ J_0 + J_1 &= pxp + 2sx + \frac{\sigma^2}{x}, \end{aligned} \tag{8.87}$$

where s, σ are constants. Construct the Casimir operator C and find the value of j in terms of s, σ. Which series are reproduced when $\sigma = 0$ or when $s = 0$?

4. Consider the following Lagrangian in 3 dimensions

$$L = \frac{1}{2}\frac{\dot{\mathbf{r}}^2}{1+\mathbf{r}^2} - \frac{1}{2}\frac{(\dot{\mathbf{r}}\cdot\mathbf{r})^2}{(1+\mathbf{r}^2)^2} \tag{8.88}$$

Evidently this Lagrangian is symmetric under rotations $\delta_\omega \mathbf{r}_i = (\boldsymbol{\omega}\times\mathbf{r})_i$ since everything is written in terms of dot products. In addition, show that there is also a hidden symmetry under the transformation $\delta_\varepsilon \mathbf{r}_i = \boldsymbol{\varepsilon}_i + \boldsymbol{\varepsilon}\cdot\mathbf{r}\;\mathbf{r}_i$. Construct the conserved quantities due to both of these symmetries, call them $\mathbf{L}$ and $\mathbf{K}$ respectively, and express them in terms of the canonical momentum (note $\mathbf{p}$ is not just $\dot{\mathbf{r}}$). Show that they are indeed conserved when you use the equations of motion. Compute the Poisson brackets of these six generators and identify the Lie algebra that they satisfy.

5. Recall the Wigner-Eckhart theorem for a $j = 1/2$ tensor operator $K_n, n = +1/2, -1/2$, applied on an arbitrary state $K_n|jm >=?$. Write out explicitly the right hand side, including the explicit Clebsch-Gordan coefficients and the unknown reduced matrix elements. Thus, the right hand side has two unknown coefficients A and B that are independent of the magnetic quantum numbers n, m.

6. Consider the following collection of 8 operators:J_0, J_+, J_- (angular momentum or isospin), Y, K_n, Q_n, where the Y operator is a $j = 0$ operator, while the K_n and Q_n operators are both $j = 1/2$ tensor operators with $n = +1/2, -1/2$ as in problem 1. We will require that the K_n, Q_n operators are complex and that they are hermitian conjugates of each other by demanding that

$$K^\dagger_{+1/2} = -Q_{-1/2}, \quad K^\dagger_{-1/2} = +Q_{+1/2} \ .$$

The statement that Y, K_n, Q_n are tensor operators with the above specified values of $j = 0, 1/2$, is equivalent to the following commutation rules with the angular momentum operators

$$\begin{aligned}
[J_0, J_\pm] &= \pm J_\pm, \quad [J_+, J_-] = 2J_0, \quad [Y, J_{\pm,0}] = 0, \\
[J_0, K_{\pm 1/2}] &= \pm\frac{1}{2} K_{\pm 1/2}, \quad [J_+, K_{-1/2}] = K_{1/2}, \quad [J_-, K_{1/2}] = K_{-1/2}, \\
[J_0, Q_{\pm 1/2}] &= \pm\frac{1}{2} Q_{\pm 1/2}, \quad [J_+, Q_{-1/2}] = Q_{1/2}, \quad [J_-, Q_{1/2}] = Q_{-1/2}, \\
[J_\pm, K_{\pm 1/2}] &= 0, \quad [J_\pm, Q_{\pm 1/2}] = 0
\end{aligned} \tag{8.89}$$

Note that in the third line $\pm K_{\mp 1/2}$ may be substituted for $Q_{\pm 1/2}$ and the commutation rules would be consistent with hermitian conjugation (note a similar statement for commutators below). If these commutation rules are supplemented with

$$\begin{aligned}
[Y, K_{\pm 1/2}] &= K_{\pm 1/2}, \quad [Y, Q_{\pm 1/2}] = -Q_{\pm 1/2}, \\
[K_{\pm 1/2}, Q_{\mp 1/2}] &= -(J_0 \pm \frac{3}{2} Y), \quad [K_{\pm 1/2}, Q_{\pm 1/2}] = \pm J_\pm, \\
[K_n, K_m] &= 0 = [Q_n, Q_m].
\end{aligned} \tag{8.90}$$

then it is found that the Jacobi identities

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{8.91}$$

are satisfied with the above commutation rules when A, B, C are any of the 8 operators $\mathbf{J}, Y, K_n, Q_n$.

When a system of operators close (that is no new operators appear on the right hand side) under commutation rules that are consistent with the Jacobi identity, it is said that these operators form a Lie algebra. In the present case these 8 operators form the Lie algebra of $SU(3)$ and the 3 operators J_0, J_+, J_- form the Lie sub-algebra $SU(2)$. Verify that the Jacobi identity is satisfied for

(i) $A = J_+, B = J_-, C = J_0$,

(ii) $A = K_{\pm 1/2}, B = Q_{\mp 1/2}, C = J_0$ and $C = Y$

(iii) $A = K_{\pm 1/2}, B = Q_{\pm 1/2}, C = J_\pm$.

7. Consider the 3-dimensional harmonic oscillator and its 3 creation-annihilation operators. Recall that the eigenstates of the Hamiltonian are the number states

$$|n_1, n_2, n_3 >= \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} \frac{(a_3^\dagger)^{n_3}}{\sqrt{n_3!}} |0> . \tag{8.92}$$

In addition to the total number operator $\hat{N} = a_1^\dagger a_1 + a_2^\dagger a_2 + a_3^\dagger a_3$ define the following 8 operators that commute with it

$$\begin{array}{l} J_0 = \frac{1}{2}(a_1^\dagger a_1 - a_2^\dagger a_2), \quad J_+ = a_1^\dagger a_2, \quad J_- = a_2^\dagger a_1 \\ U_+ = a_1^\dagger a_3, \quad U_- = a_3^\dagger a_1, \quad V_+ = a_2^\dagger a_3, \quad V_- = a_3^\dagger a_2 \\ Y = \frac{1}{3}(a_1^\dagger a_1 + a_2^\dagger a_2 - 2\, a_3^\dagger a_3). \end{array} \tag{8.93}$$

Using the commutation rules for harmonic oscillators it is not difficult to show that these 8 operators close under commutation, hence they form a Lie algebra. The operators **J** constructed from the first two oscillators form an $SU(2)$ sub-algebra. Furthermore, by comparison to the previous problem we may identify

$$U_+ \equiv K_{1/2},\, V_+ \equiv K_{-1/2},\, U_- \equiv K_{1/2}^\dagger = -Q_{-1/2},\; V_- = K_{-1/2}^\dagger = Q_{1/2.} \tag{8.94}$$

Hence these 8 operators form the $SU(3)$ Lie algebra. By analogy to the $SU(2)$ states of the 2D harmonic oscillator the number states $|n_1, n_2, n_3 >$ may be rewritten in terms of the eigenvalues of the commuting operators $\hat{N}, J_0, Y$ with eigenvalues n, m, y respectively

$$|n, m, y >= \frac{(a_1^\dagger)^{n/3+y/2+m}}{\sqrt{(n/3+y/2+m)!}} \; \frac{(a_2^\dagger)^{n/3+y/2-m}}{\sqrt{(n/3+y/2-m)!}} \; \frac{(a_3^\dagger)^{n/3-y}}{\sqrt{(n/3-y)!}} \; |0> \tag{8.95}$$

where we have defined

$$n = n_1 + n_2 + n_3, \quad m = \frac{1}{2}(n_1 - n_2), \quad y = \frac{1}{3}(n_1 + n_2 - 2n_3).$$

Verify that indeed these operators have the stated eigenvalues. Furthermore, by concentrating only on the first two oscillators and comparing to the 2-dimensional case it is useful to define the quantum number j

$$j = \frac{1}{2}(n_1 + n_2) = n/3 + y/2. \tag{8.96}$$

So, the state may be labelled $|n, j, m, y >$ with the additional quantum number j, but keeping in mind that it is a function of the others. From

the positivity of the integers $n_I \geq 0$ one derives the allowed values and ranges of these quantum numbers

$$\begin{array}{rcl} n & = & 0, 1, 2, 3, \cdots . \\ (j, y) & = & (\frac{n}{2}, \frac{n}{3}), \cdots , \ (\frac{n-k}{2}, \frac{n-3k}{3}), \cdots , (0, -\frac{2n}{3}) \\ m & = & -j, \ -j+1, \cdots , \ j-1, \ j \end{array} \tag{8.97}$$

- Compute the action of the 3 creation and annihilation operators on the states, writing the resulting state in terms of the labels of the type $|n, j, m, y>$.
- Using this result now obtain the action of the 8 operators which do not change the value of n, and again give your result in terms of the states labelled by $|n, j, m, y>$. Since the value of n cannot change you have found irreducible representations of $SU(3)$ for each value of n. You will compare your result to the general solution of the $SU(3)$ problem given below and find out that it corresponds to a subclass of irreducible representations.

8. Consider the 8 operators of $SU(3)$ of the previous problem but now constructed with fermionic oscillators that satisfy $\left\{ b_\alpha , b^\dagger_\beta \right\} = \delta_{\alpha\beta}$. Verify that the Lie algebra closes with identical coefficients as the bosonic case. Show that the list of all possible values of the quantum numbers (n, j, y) that may be constructed with fermions is

$$\begin{array}{rcl} (n, j, y) & = & (0, 0, 0), \ (1, 0, -2/3), \ (1, 1/2, 1/3) \\ & & (3, 0, 0), \ (2, 0, 2/3), \ (2, 1/2, -1/3), \\ j & = & \frac{1}{2} \left[\left(\frac{2}{3} (n \, mod 3) + y \right) mod 2 \right] . \end{array} \tag{8.98}$$

Of course, for each value of j the remaining quantum number must take values $-j \leq m \leq j$. If we interpret j as isospin and y as hypercharge, then the three states with $n = 1$ have the quantum numbers of up, down, strange quarks, while the three states with $n = 2$ have the quantum numbers of antiquarks (also of di-quarks). The states with $n = 0, 3$ are $SU(3)$ singlets. There are other applications and/or interpretations of the mathematics. For example the three fermions $b^\dagger_\alpha$ with $\alpha = 1, 2, 3$ may represent the three colors of quarks. Then we see that we can make a color singlet state $(3, 0, 0)$ by putting three quarks of different colors together $b^\dagger_1 b^\dagger_2 b^\dagger_3 |0>$ since the 8 $SU(3)$ generators annihilate this state.

9. Let us return to the general formulation of $SU(3)$ in problem 2. For every Lie algebra one can find polynomials constructed from operators that commute with all operators in the Lie algebra. These are called Casimir operators. For $SU(2)$ the Casimir operator is $\mathbf{J} \cdot \mathbf{J}$ and it commutes with every $\mathbf{J}_i$. For $SU(3)$ there are two Casimir operators, one is quadratic and the other cubic. Verify that the following two Casimir operators C_2, C_3 commute with all 8 operators of the Lie algebra (if you find this too long

do it only for $J_0, Y, K_{+1/2}, Q_{+1/2}$)

$$\begin{array}{ll} C_2 = & \mathbf{J}^2 + \frac{3}{4}Y(Y+2) + \sum_n K_n^\dagger K_n \\ C_3 = & \mathbf{J}^2(Y+1) - \frac{1}{4}Y(Y+1)(Y+2) + \sum_n 2nK_n^\dagger J_0 K_n \\ & + K_{+1/2}^\dagger J_+ K_{-1/2} + K_{-1/2}^\dagger J_- K_{+1/2} - \frac{1}{2}\sum_n K_n^\dagger (Y+2) K_n \end{array} \tag{8.99}$$

You now have 5 operators that commute with each other: $C_2, C_3, \mathbf{J}^2, J_0, Y$, and therefore they are simultaneously diagonalizable. Just as it is convenient to parametrize the eigenvalues of $\mathbf{J}^2$ as $j(j+1)$, it is also convenient to parametrize the eigenvalues of C_2 and C_3 in terms of two integers p, q as follows

$$\begin{array}{l} C_2 = p + q + (p^2 + pq + q^2)/3, \\ C_3 = (p-q)(p+2q+3)(q+2p+3)/27. \end{array} \tag{8.100}$$

Thus, a complete labelling for an $SU(3)$ state is $|p, q; j, m, y>$ where y is the eigenvalue of the Y operator and the rest are as described above.

10. Using the harmonic oscillator construction of the $SU(3)$ generators show that the operators C_2, C_3 can be rewritten only in terms of the total number operator $\hat{N}$ (this is shown by rearranging the orders of the harmonic oscillators). Hence, for this special case the two Casimir operators are not independent from each other, and their eigenvalues q, p are not most general. Show that the only possible eigenvalues are $(p, q) = (n, 0)$ where n is the eigenvalue of $\hat{N}$ as it appeared above.

11. The general labelling of $SU(3)$ states allows us, by definition, to immediately write down the action of the 4 operators $J_{\pm,0}$ and Y on the states. Write them down. There remains to learn how to act with the remaining 4 operators $K_{\pm 1/2}$ and $Q_{\pm 1/2}$. They act like ladder operators that shift the eigenvalues j, m, y but cannot shift p, q. It is straightforward to consider the commutations $[Y, K_n]$ and $[Y, Q_n]$ to learn how y is shifted. In addition we can figure out immediately how j, m are shifted up to 4 unknown coefficients A, B, C, D from the knowledge that K_n, Q_n are $j = 1/2$ tensor operators and using the results of problem1

$$\begin{array}{rl} K_{\pm 1/2}|pqjmy> = & A_{jmy}\ |pq, j+\frac{1}{2}, m \pm \frac{1}{2}, y+1> \\ & \pm B_{\,jmy}\,|pq, j-\frac{1}{2}, m \pm \frac{1}{2}, y+1> \\ Q_{\pm 1/2}|pqjmy> = & C_{jmy}\ |pq, j+\frac{1}{2}, m \pm \frac{1}{2}, y-1> \\ & \pm D_{jmy}\ |pq, j-\frac{1}{2}, m \pm \frac{1}{2}, y-1>\,. \end{array} \tag{8.101}$$

In order to completely determine the remaining coefficients we need to impose the commutation rules for $[K_n, Q_m]$ and also insure that C_2, C_3 have the eigenvalues given above. Impose these conditions and verify that the 4 coefficients are given by

$$\begin{array}{ll} A_{jmy} = \sqrt{-\frac{G(j+y/2+1)\ (j\pm m+1)}{(2j+2)(2j+1)}}, & B_{jmy} = \sqrt{\frac{G(-j+y/2)\ (j\mp m)}{2j(2j+1)}}, \\ C_{jmy} = -\sqrt{\frac{G(-j+y/2-1)\ (j\pm m+1)}{(2j+2)(2j+1)}}, & D_{jmy} = \sqrt{-\frac{G(j+y/2)\ (j\mp m)}{2j(2j+1)}}. \end{array} \tag{8.102}$$

where

$$\begin{aligned} G(x) &= (x+\tfrac{p+2q+3}{3})(x+\tfrac{p-q}{3})(x-\tfrac{2p+q+3}{3}) \\ &= x^3-(C_2+1)x-C_3. \end{aligned} \tag{8.103}$$

Note that with these coefficients we find that, once p, q are fixed, there are a finite number of states that are shifted into each other since we must satisfy the inequalities

$$\begin{aligned} -G(j+y/2+1) \geq 0, \quad G(-j+y/2) \geq 0, \\ G(-j+y/2-1) \geq 0, \quad -G(j+y/2) \geq 0. \end{aligned} \tag{8.104}$$

in order to have real coefficients A, B, C, D (this is the condition of unitarity= positive definite norm). Then the allowed quantum numbers (j, y) are shown in the figure. The states are represented by the corners of the cells in this figure. Note the maximum and minimum values of (j, m) at the extremities of the lattice.

Verify that the allowed quantum numbers are the ones given in the figure, and that the commutation rules $[K_n, Q_m]$ are satisfied and C_2, C_3 are diagonal on these states.

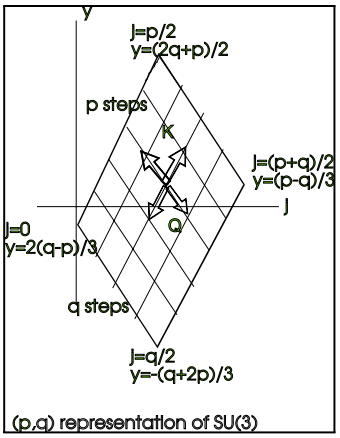


Fig. 8.1 – Basis of states for SU(3).

12. Using all of the above information, find the set of states for the cases

$$(p,q)=(1,0); \quad (p,q)=(2,0); \quad (p,q)=(1,1); \quad (p,q)=(3,0). \tag{8.105}$$

How many states are there in each case? Compare your result to the states $(n, 0)$ of the harmonic oscillator model.

36. SU(n) MULTIPLETS AND YOUNG DIAGRAMS

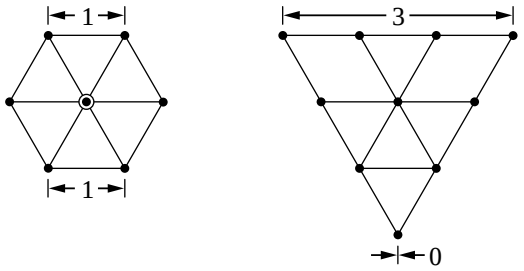
Written by C.G. Wohl (LBNL).

This note tells (1) how SU(n) particle multiplets are identified or labeled, (2) how to find the number of particles in a multiplet from its label, (3) how to draw the Young diagram for a multiplet, and (4) how to use Young diagrams to determine the overall multiplet structure of a composite system, such as a 3-quark or a meson-baryon system.

In much of the literature, the word "representation" is used where we use "multiplet," and "tableau" is used where we use "diagram."

36.1. Multiplet labels

An SU(n) multiplet is uniquely identified by a string of ($n-1$) nonnegative integers: $(\alpha, \beta, \gamma, \ldots)$. Any such set of integers specifies a multiplet. For an SU(2) multiplet such as an isospin multiplet, the single integer α is the number of *steps* from one end of the multiplet to the other (*i.e.*, it is one fewer than the number of particles in the multiplet). In SU(3), the two integers α and β are the numbers of steps across the top and bottom levels of the multiplet diagram. Thus the labels for the SU(3) octet and decuplet



are (1,1) and (3,0). For larger n, the interpretation of the integers in terms of the geometry of the multiplets, which exist in an ($n-1$)-dimensional space, is not so readily apparent.

The label for the SU(n) singlet is $(0, 0, \ldots, 0)$. In a flavor SU(n), the n quarks together form a $(1, 0, \ldots, 0)$ multiplet, and the n antiquarks belong to a $(0, \ldots, 0, 1)$ multiplet. These two multiplets are *conjugate* to one another, which means their labels are related by $(\alpha, \beta, \ldots) \leftrightarrow (\ldots, \beta, \alpha)$.

36.2. Number of particles

The number of particles in a multiplet, $N = N(\alpha, \beta, \ldots)$, is given as follows (note the pattern of the equations).

In SU(2), $N = N(\alpha)$ is

$$N = \frac{(\alpha+1)}{1} \ . \tag{36.1}$$

In SU(3), $N = N(\alpha, \beta)$ is

$$N = \frac{(\alpha+1)}{1} \cdot \frac{(\beta+1)}{1} \cdot \frac{(\alpha+\beta+2)}{2} \ . \tag{36.2}$$

In SU(4), $N = N(\alpha, \beta, \gamma)$ is

$$N = \frac{(\alpha+1)}{1} \cdot \frac{(\beta+1)}{1} \cdot \frac{(\gamma+1)}{1} \cdot \frac{(\alpha+\beta+2)}{2} \cdot \frac{(\beta+\gamma+2)}{2} \cdot \frac{(\alpha+\beta+\gamma+3)}{3} \ . \tag{36.3}$$

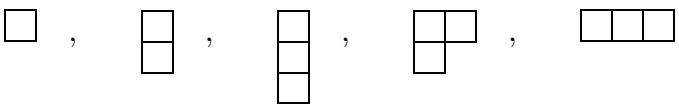
Note that in Eq. (36.3) there is no factor with $(\alpha+\gamma+2)$: only a *consecutive* sequence of the label integers appears in any factor. One more example should make the pattern clear for any SU(n). In SU(5), $N = N(\alpha, \beta, \gamma, \delta)$ is

$$\begin{aligned} N = {} & \frac{(\alpha+1)}{1} \cdot \frac{(\beta+1)}{1} \cdot \frac{(\gamma+1)}{1} \cdot \frac{(\delta+1)}{1} \cdot \frac{(\alpha+\beta+2)}{2} \cdot \frac{(\beta+\gamma+2)}{2} \\ & \times \frac{(\gamma+\delta+2)}{2} \cdot \frac{(\alpha+\beta+\gamma+3)}{3} \cdot \frac{(\beta+\gamma+\delta+3)}{3} \cdot \frac{(\alpha+\beta+\gamma+\delta+4)}{4} \ . \end{aligned} \tag{36.4}$$

From the symmetry of these equations, it is clear that multiplets that are conjugate to one another have the same number of particles, but so can other multiplets. For example, the SU(4) multiplets (3,0,0) and (1,1,0) each have 20 particles. Try the equations and see.

36.3. Young diagrams

A Young diagram consists of an array of boxes (or some other symbol) arranged in one or more *left-justified* rows, with each row being *at least as long* as the row beneath. The correspondence between a diagram and a multiplet label is: The top row juts out α boxes to the right past the end of the second row, the second row juts out β boxes to the right past the end of the third row, *etc.* A diagram in SU(n) has at most n rows. There can be any number of "completed" columns of n boxes buttressing the left of a diagram; these don't affect the label. Thus in SU(3) the diagrams

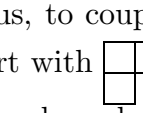


represent the multiplets (1,0), (0,1), (0,0), (1,1), and (3,0). In any SU(n), the quark multiplet is represented by a single box, the antiquark multiplet by a column of $(n-1)$ boxes, and a singlet by a completed column of n boxes.

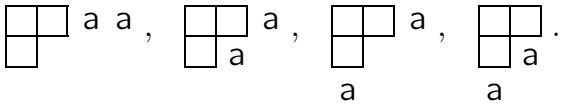
36.4. Coupling multiplets together

The following recipe tells how to find the multiplets that occur in coupling two multiplets together. To couple together more than two multiplets, first couple two, then couple a third with each of the multiplets obtained from the first two, *etc.*

First a definition: A sequence of the letters $a, b, c, \ldots$ is *admissible* if at any point in the sequence at least as many a's have occurred as b's, at least as many b's have occurred as c's, *etc.* Thus *abcd* and *aabcb* are admissible sequences and *abb* and *acb* are not. Now the recipe:

(a) Draw the Young diagrams for the two multiplets, but in one of the diagrams replace the boxes in the first row with a's, the boxes in the second row with b's, *etc.* Thus, to couple two SU(3) octets (such as the π-meson octet and the baryon octet), we start with  and  . The *unlettered* diagram forms the *upper left-hand corner* of all the enlarged diagrams constructed below.

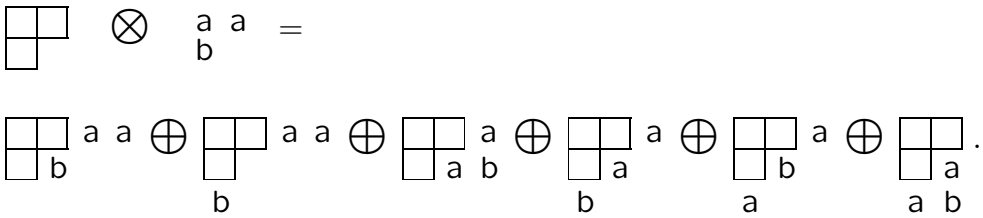
(b) Add the a's from the lettered diagram to the right-hand ends of the rows of the unlettered diagram to form all possible legitimate Young diagrams that have no more than one a per column. In general, there will be several distinct diagrams, and all the a's appear in each diagram. At this stage, for the coupling of the two SU(3) octets, we have:



(c) Use the b's to further enlarge the diagrams already obtained, subject to the same rules. Then throw away any diagram in which the full sequence of letters formed by reading *right to left* in the first row, then the second row, *etc.*, is not admissible.

(d) Proceed as in (c) with the c's (if any), *etc.*

The final result of the coupling of the two SU(3) octets is:



Here only the diagrams with admissible sequences of a's and b's and with fewer than four rows (since $n = 3$) have been kept. In terms of multiplet labels, the above may be written

$$(1,1) \otimes (1,1) = (2,2) \oplus (3,0) \oplus (0,3) \oplus (1,1) \oplus (1,1) \oplus (0,0) \ .$$

In terms of numbers of particles, it may be written

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{10} \oplus \overline{\mathbf{10}} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1} \ .$$

The product of the numbers on the left here is equal to the sum on the right, a useful check. (See also Sec. 13 on the Quark Model.)

Chapter 9

SOME APPLICATIONS OF SYMMETRY

In this chapter we will discuss exactly solvable quantum mechanical systems which have a symmetry structure. We will illustrate the power of group theory in solving the problems and shedding additional light on the structure of the system as compared to the more standard Schrödinger's differential equation approach. We will discuss one dimensional problems with special potentials such as the Morse potential and it's $SU(2)$ structure, two dimensional problems such as a particle in a magnetic field with its "magnetic translation group" structure, The hydrogen atom in d-dimensions with its $SO(d+1)$ structure for $d = 2, 3, \cdots$, and the Interacting Boson Model of large nuclei with its $SU(6)$ structure.

9.1 H-atom in d-dimensions and SO(d+1)

The Hydrogen atom in d-dimensions is described by the Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} - \frac{Ze^2}{|\mathbf{r}|}, \tag{9.1}$$

where $\mathbf{r}, \mathbf{p}$ are vectors in d-dimensions. The Schrödinger equation for the central force problem in $d-$dimensions was discussed in chapter 6. The complete wavefunction is

$$\psi(\mathbf{r}) = r^{-\frac{1}{2}(d-1)} f_{El_d}(r)\ T_{I_1 I_2 \cdots I_l}(\hat{\mathbf{r}}) \tag{9.2}$$

where $f_{El_d}(r)$ is the radial wavefunction and $T_{I_1 I_2 \cdots I_l}(\hat{\mathbf{r}})$ is the angular wavefunction that is analogous to the spherical harmonics $Y_{lm}(\hat{\mathbf{r}})$ in three dimensions. As discussed in Chapter 6 (see section (6.6) and problem (6.10)) $T_{I_1 I_2 \cdots I_l}(\hat{\mathbf{r}})$ can be represented as a completely symmetric traceless tensor constructed from the direct product of l powers of the unit vector $\hat{\mathbf{r}}_I$. The norm is

$$\int d^d\mathbf{r}\ |\psi(\mathbf{r})|^2 = \int_0^\infty dr\ |f_{El_d}(r)|^2 = 1. \tag{9.3}$$

It was shown in section (6.7) that the eigenvalue problem could be reduced to a radial equation that looked just like the usual radial equation in 3-dimensions

$$\left(-\partial_r^2 + \frac{l_d(l_d+1)}{r^2} - \frac{2mZe^2}{r} - \frac{2mE}{\hbar^2}\right) f_{El_d}(r) = 0, \tag{9.4}$$

where

$$l_d = l + \frac{d-3}{2}, \tag{9.5}$$

and $l = 0, 1, 2, \cdots$, which is a quantum number analogous to the orbital angular momentum, is the rank of the completely symmetric traceless tensor $T_{I_1 I_2 \cdots I_l}(\hat{\mathbf{r}})$. The number of independent components of this tensor for a fixed value of l in d dimensions is

$$N(l,d) = \frac{(l+d-3)!}{(d-2)!\ l!}\,(2l+d-2)\,. \tag{9.6}$$

In 3 dimensions this number reduces to $N(l,3) = 2l+1$, i.e. the familiar number of spherical harmonics $Y_{lm}(\hat{\mathbf{r}})$. In 2 dimensions it becomes $N(l,2) = 2$ for $l \neq 0$, and $N(l,2) = 1$ for $l = 0$, which is consistent with the number of angular momentum wavefunctions $\exp(im\phi)$, with $m = \pm l$.

The solution of the eigenvalue problem proceeds just like the 3-dimensional case, and we find that the energy is quantized as

$$E_n = -\frac{mc^2}{2}\frac{Z^2\alpha^2}{\left(n + \frac{d-3}{2}\right)^2}, \tag{9.7}$$

where $n = n_r + l + 1$, and $n_r = 0, 1, 2, \cdots$ is the radial quantum number. The ranges of the quantum numbers may be rewritten as

$$\begin{array}{l} n = 1, 2, 3, \cdots \\ l = 0, 1, \cdots (n-1). \end{array} \tag{9.8}$$

The degeneracy of the states for a fixed value of n is

$$D_n(d) = \sum_{l=0}^{n-1} N(l,d) = \frac{(n+d-3)!}{(d-1)!\ \ (n-1)!}\,(2n+d-3)\,. \tag{9.9}$$

This general result agrees with the degeneracy of the states computed in Chapter 6 in 2 and 3 dimensions

$$\begin{array}{lll} D_n(2) = 2n-1 & \text{in} & d=2 \\ D_n(3) = n^2 & \text{in} & d=3 \end{array}\ . \tag{9.10}$$

A plot of the levels in 2 and 3 dimensions is given in Fig.9.1. The quantum numbers l, n are shown on the horizontal and vertical axes respectively, and the degeneracy of each state is indicated.

How is this degeneracy explained? The Schrödinger equation approach outlined above has produced degenerate states but has not given a clue for why

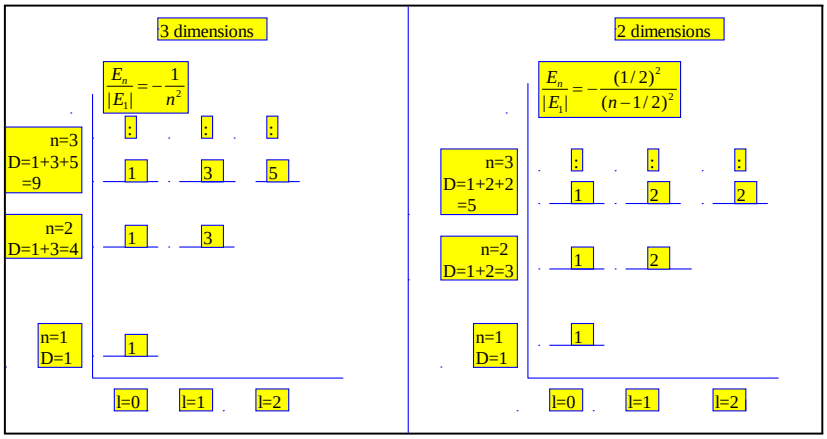


Figure 9.1: Fig.9.1– Degeneracies in d=2 and 3.

there are unexpected degeneracies. The Hamiltonian is manifestly invariant under rotations in d dimensions that form the group $SO(d)$. This explains the degeneracy for a given value of angular momentum l, but it does not explain why states with different values of angular momentum have the same energy. Therefore, we must seek a larger symmetry that is not manifest but nevertheless is present. As a first clue note that, in two dimensions the number of degenerate states, $D_n(2) = 2n-1$, coincides with the number of angular momentum states, $2j+1$, provided we identify $j = n-1$. This suggest that there may be an underlying $SO(3)$ symmetry for $d = 2$. Generalizing this clue to 3 dimensions one comes up with $SO(4)$ symmetry for $d = 3$. More generally, this observation suggests that in d dimensions we might expect $SO(d+1)$ symmetry.

This detective-like approach gets reinforced with a stronger clue. Using the definition of eq.(9.6) we may write the number of states computed in eq.(9.9) in the form

$$D_n(d) = N(n-1,\ d+1), \tag{9.11}$$

which indicates that the degeneracy coincides with the number of independent components in a traceless symmetric tensor of rank $(n-1)$ in $(d+1)$ dimensions. This remarkable coincidence strongly suggests that the H-atom has a Hilbert space isomorphic to the states of the rotation group $SO(d+1)$, as if there is one extra dimension!

9.2 The symmetry algebra

With these clues we are now ready to construct the symmetry. To begin with, the $SO(d)$ rotation symmetry has operators

$$L_{IJ} = r_I\, p_J - r_J\, p_I \tag{9.12}$$

that generate infinitesimal rotations in the (I, J) plane. For example, rotations in the 1-2 plane, $\delta_{12} r_1 = \varepsilon_{12}\, r_2, \delta_{12} r_2 = -\varepsilon_{12}\, r_1$, $\delta_{12} r_3 = \delta_{12} r_3 = \cdots = 0$, are generated as follows

$$\delta_{12} r_I = -\frac{i}{\hbar}\varepsilon_{12}[L_{12}, r_I] = \varepsilon_{12}(r_2\partial_1 - r_1\partial_2) r_I = \varepsilon_{12}\left(r_2\delta_{1I} - r_1\delta_{2I}\right). \tag{9.13}$$

In two dimensions there is only one generator L_{12}, while in three dimensions there are the familiar rotation operators $L_{23} \equiv L_1$, $L_{31} \equiv L_2$, $L_{12} \equiv L_3$. Their commutation rules follow from those of $(r_I,\, p_I)$ for any d

$$[L_{IJ}, L_{KL}] = i\hbar(\delta_{IK}\, L_{JL} + \delta_{JL}\, L_{IK} - \delta_{IL}\, L_{JK} - \delta_{JK}\, L_{IL}). \tag{9.14}$$

These operators commute with the Hamiltonian since H is constructed only from rotationally invariant dot products. This symmetry is responsible for the degenerate states with a fixed value of l. By the same token, the quantities L_{IJ} are all time independent since they commute with the Hamiltonian.

We must now look for additional operators that will give a higher symmetry beyond $SO(d)$. In three dimensions it was known in the classical mechanics of planetary motion (that has the same form of Hamiltonian) that in addition to the time independent angular momentum vector there is another constant vector, the Runge-Lenz vector given by

$$\mathbf{A} = \mathbf{L} \times \mathbf{p} - mZe^2\, \hat{\mathbf{r}}, \quad \dot{\mathbf{A}} = \mathbf{0}. \tag{9.15}$$

In quantum mechanics we expect this vector to commute with the Hamiltonian. In d dimensions the Runge-Lenz vector is generalized to

$$A_I = \frac{1}{2}\left(L_{IJ}\, p^J + p^J L_{IJ}\right) - mZe^2 \frac{r_I}{r}, \tag{9.16}$$

where the hermitian combination of L_{IJ} and p^J is taken. One can now check that this observable does indeed commute with the Hamiltonian (see problem (4))

$$[H, A_I] = 0. \tag{9.17}$$

Hence we can regard A_I as generators of a symmetry that leave the energy invariant. The infinitesimal transformations of r_I, p_I that give this invariance may be computed from the commutators (see problem (1)

$$\delta r_I = -\frac{i}{\hbar}\varepsilon^J [A_J, r_I], \quad \delta p_I = -\frac{i}{\hbar}\varepsilon^J [A_J, p_I]. \tag{9.18}$$

We now seek the Lie algebra of all the symmetry operators by commuting L_{IJ}, A_K among themselves. After a lot of algebra one finds (see problem (2))

$$[A_I, A_J] = i\hbar(-2mH)\; L_{IJ}, \tag{9.19}$$

and of course, that A_I rotates like a vector as indicated by the commutation rule

$$[L_{IJ}, A_K] = i\hbar\left(\delta_{IK}\, A_J - \delta_{JK}\, A_I\right). \tag{9.20}$$

We notice that if it were not for the operator H on the right hand side of (9.19), the set L_{IJ}, A_K would form a Lie algebra with constant coefficients (as opposed to operators). However, since H commutes with all the operators L_{IJ}, A_K, it effectively acts like a constant and therefore we may define a more convenient operator

$$L_{I0} = \frac{A_I}{\sqrt{-2mH}} \tag{9.21}$$

that has commutation rules with constant coefficients

$$\begin{aligned} [L_{IJ}, L_{K0}] &= i\hbar\,(\delta_{IK}\,L_{J0} - \delta_{JK}\,L_{I0}) \\ [L_{I0}, L_{J0}] &= i\hbar L_{IJ}. \end{aligned} \tag{9.22}$$

Hence we have a closed Lie algebra among the operators L_{IJ}, L_{K0}.

Which Lie algebra is this? To consult Cartan's table of Lie algebras we need the total number of symmetry generators. Counting the number of L_{IJ} and L_{I0}'s we get

$$\frac{1}{2}d(d-1) + d = \frac{1}{2}d(d+1), \tag{9.23}$$

which is the same number of generators as $SO(d+1)$. To finally show that the symmetry is really $SO(d+1)$ it is convenient to define the rotations as if we have $d+1$ coordinates labelled by $\mu = 0, I$, where $I = 1, 2, \cdots d$.

$$L_{\mu\nu} = \begin{cases} L_{IJ} & \text{for} \quad \mu = I,\ \nu = J \\ L_{I0} = -L_{0I} & \text{for} \quad \begin{cases} \mu=I,\ \nu=0 \\ \mu=0,\ \nu=I \end{cases} \\ L_{00} = 0 & \text{for} \quad \mu = \nu = 0 \end{cases} \tag{9.24}$$

Then the commutation rules (9.14,9.22) take the form of the Lie algebra of $SO(d+1)$

$$[L_{\mu\nu}, L_{\kappa\lambda}] = i\hbar(\delta_{\mu\kappa}\,L_{\nu\lambda} + \delta_{\nu\lambda}\,L_{\mu\kappa} - \delta_{\mu\lambda}\,L_{\nu\kappa} - \delta_{\nu\kappa}\,L_{\mu\lambda}). \tag{9.25}$$

Therefore the symmetry of the H-atom in d dimensions is $SO(d+1)$!

9.3 The dynamical symmetry structure of H-atom

Having established that the Hamiltonian commutes with all the generators of $SO(d+1)$ one may wonder whether H is a function of the Casimir operators of $SO(d+1)$ (this need not be the case in all instances). The quadratic Casimir operator which commutes with every $L_{\mu\nu}$ is given by

$$C_2 = \frac{1}{2}\sum_{\mu\neq\nu}(L_{\mu\nu})^2 = \frac{1}{2}\sum_{I\neq J}(L_{IJ})^2 + \sum_I \frac{(A_I)^2}{-2mH}. \tag{9.26}$$

Using the explicit expression (9.16) for the Runge-Lenz vector one can show that its square takes the form (see problem (3))

$$\sum_I (A_I)^2 = 2mH \left[\frac{1}{2}\sum_{I\neq J}(L_{IJ})^2 + \hbar^2\left(\frac{d-1}{2}\right)^2\right] + (mZe^2)^2. \tag{9.27}$$

The term proportional to $\hbar^2$ arises from the reordering of quantum operators; it is absent in the classical theory. Inserting this result into the expression for the Casimir operator we find a relation between the Hamiltonian and C_2

$$H = -\frac{mc^2(Z\alpha)^2}{2}\left[\frac{1}{\hbar^2}C_2 + \left(\frac{d-1}{2}\right)^2\right]^{-1}, \tag{9.28}$$

where we have used the expression for the fine structure constant $\alpha = e^2/\hbar c$.

Having established this result, we can now reverse the reasoning: Given a symmetry group $SO(d+1)$ one can define all the states on which the symmetry acts irreducibly. This means labelling the states with the eigenvalues of the Casimir operators and other mutually commuting operators. It is often convenient to choose the additional operators as the Casimirs of the next largest subgroup, then the Casimirs of next sub-subgroup, etc. until we reach the smallest subgroup $SU(2)$. The eigenvalues of all these Casimirs are known from the representation theory of groups. Then, the eigenvalues and the eigenstates of the Hamiltonian are completely solved by group theoretical methods.

9.3.1 2 dimensions

To see how this works, let's start with the two dimensional Hydrogen atom for which the symmetry is $SO(3)$. The generators of the symmetry are

$$\begin{aligned}
J_3 &= L_{12} = r_1p_2 - r_2p_1 \\
J_1 &= \frac{1}{\sqrt{-2mH}}\left[\frac{1}{2}\left(L_{12}\,p_2 + p_2L_{12}\right) - mZe^2\frac{r_1}{r}\right] \\
J_2 &= \frac{1}{\sqrt{-2mH}}\left[\frac{1}{2}\left(-L_{12}\,p_1 - p_1L_{12}\right) - mZe^2\frac{r_2}{r}\right] \\
C_2 &= J_1^2 + J_2^2 + J_3^3
\end{aligned} \tag{9.29}$$

The eigenstates and eigenvalues $|jm>$ are well known to students of quantum mechanics. We only need to decide which of these states belong to the H-atom. Since m is the eigenvalue of the orbital rotation generator $J_3 = L_{12}$, it can only take integer eigenvalues. Hence j must be integer. Now we compute the

eigenvalues of H

$$\begin{aligned} H\ |jm\ > &= -\frac{mc^2(Z\alpha)^2}{2}\left[\frac{1}{\hbar^2}C_2+\left(\frac{2-1}{2}\right)^2\right]^{-1}\ |jm> \\ &= -\frac{mc^2(Z\alpha)^2}{2}\left[j(j+1)+1/4\right]^{-1}|jm> \\ &= -\frac{mc^2(Z\alpha)^2}{2(j+1/2)^2}\ |jm>, \qquad j=0,1,2,\cdots \end{aligned} \tag{9.30}$$

Evidently we have recovered the same eigenstates and eigenvalues that were computed with the Schrödinger equation approach in (9.7), provided we identify $j=n-1$ as anticipated at the end of that section.

9.3.2 3 dimensions

The H-atom in 3 dimensions was first solved by Pauli. In this case there are the three angular momentum operators $\mathbf{L}=\mathbf{r}\times\mathbf{p}$ and the three Runge-Lenz vectors $\mathbf{K}=\mathbf{A}/\sqrt{-2mH}$. It is instructive to rewrite their commutation rules given in (9.25) in the form

$$\begin{aligned} [L_I,L_J] &= i\hbar\,\epsilon_{IJK}\,L_K \\ [L_I,K_J] &= i\hbar\,\epsilon_{IJK}\,K_K \\ [K_I,K_J] &= i\hbar\,\epsilon_{IJK}\,L_K. \end{aligned} \tag{9.31}$$

It is further useful to define the combinations

$$\mathbf{J}^{(\pm)}=\frac{1}{2}\left(\mathbf{L}\pm\mathbf{K}\right) \tag{9.32}$$

that have commutation rules

$$\begin{aligned} [J_I^{(+)},J_J^{(+)}] &= i\hbar\,\epsilon_{IJK}\ J_K^{(+)} \\ [J_I^{(-)},J_J^{(-)}] &= i\hbar\,\epsilon_{IJK}\ J_K^{(-)} \\ [J_I^{(+)},J_J^{(-)}] &= 0. \end{aligned} \tag{9.33}$$

So, $\mathbf{J}^{(\pm)}$ satisfy two independent $SU(2)$ Lie algebras that are equivalent to the $SO(4)$ Lie algebra

$$SO(4)\sim SU(2)_+\otimes SU(2)_-\ . \tag{9.34}$$

Therefore the states of $SO(4)$ may be rewritten as the states of $SU(2)_+\otimes SU(2)_-$

$$|j_+,m_+;j_-,m_->,\quad j_\pm=\text{half integers} \tag{9.35}$$

This is the complete solution of the quantum mechanics problem (eigenstates and eigenvalues) for $SO(4)$. We only need to decide which of these states belong to the Hydrogen atom. Furthermore, this notation brings us to familiar territory since we can now use the mathematics of addition of angular momentum.

Indeed, as seen from the definitions in (9.32), the orbital angular momentum of the hydrogen atom is the diagonal sum of the two commuting $SU(2)$ algebras

$$\mathbf{L} = \mathbf{J}_{(+)} + \mathbf{J}_{(-)}. \tag{9.36}$$

Therefore, we may also find it convenient to use the "total angular momentum" states that determine l in terms of $j_\pm$

$$|j_+, j_-; l, m_l >, \quad l = |j_+ - j_-|, \cdots, (j_+ + j_-) \tag{9.37}$$

The quadratic Casimir operator of $SO(4)$ given in (9.26) may be rewritten in terms of the Casimir operators of the two $SU(2)'s$

$$C_2 = \mathbf{L}^2 + \mathbf{K}^2 = 2\left(\mathbf{J}^2_{(+)} + \mathbf{J}^2_{(-)}\right) \tag{9.38}$$

Furthermore, in the present problem $\mathbf{L}$ and $\mathbf{K}$ are perpendicular $\mathbf{L}\cdot\mathbf{K}$ =0 (this is easy to see in vector notation in the classical theory as given in eq.(9.15), but is also true as operators). A consequence of this is $\mathbf{J}^2_{(+)} = \mathbf{J}^2_{(-)}$ so that their eigenvalues are forced to be the same

$$j_+ = j_- \equiv j. \tag{9.39}$$

This means that for the Hydrogen atom we can only admit the subset of $SO(4)$ states (9.37) of the form

$$|j, j; l, m_l >, \quad l = 0, 1, \cdots, 2j\ , \quad j = 0, 1/2, 1, 3/2, \cdots \quad . \tag{9.40}$$

On these states the $SO(4)$ Casimir (9.38) takes the values

$$C_2\,|j, j; l, m_l >= 4\hbar^2 j(j+1)\,|j, j; l, m_l >, \tag{9.41}$$

and the Hamiltonian has the eigenvalues

$$\begin{aligned} H\ |j, j; l, m_l \ > \ &= -\frac{mc^2(Z\alpha)^2}{2}\left[\frac{1}{\hbar^2}C_2 + \left(\frac{3-1}{2}\right)^2\right]^{-1} |j, j; l, m_l > \\ &= -\frac{mc^2(Z\alpha)^2}{2}\left[4j(j+1)+1\right]^{-1} |j, j; l, m_l > \\ &= -\frac{mc^2(Z\alpha)^2}{2(2j+1)^2}\ |j, j; l, m_l >, \qquad \begin{cases} j = 0, 1/2, 1, 3/2, \cdots \\ l = 0, 1, \cdots, 2j \end{cases} \end{aligned} \tag{9.42}$$

This result is in complete agreement with the Schrödinger equation approach, provided we identify the total quantum number as

$$n = 2j + 1. \tag{9.43}$$

Furthermore, the degeneracy is easy to understand. The number of states in the SO(4) multiplet $|j_+, m_+; j_-, m_- >$ is $(2j_+ + 1)(2j_- + 1)$. However, for the H-atom we specialize to $j_+ = j_- = j$, which leads to the degeneracy

$$D = (2j+1)^2 = n^2, \tag{9.44}$$

as found in Fig.9.1. Evidently, we have now gained a lot more insight into the multiplet and symmetry structure of the H-atom in three dimensions.

9.3.3 d dimensions

With the experience gained in $d = 2$ and $d = 3$ we are now prepared to identify the symmetry structure of the eigenstates and eigenvalues of the H-atom in higher dimensions. The general $SO(d+1)$ problem has a Hilbert space labelled by the $[(d+1)/2]$ Casimir operators. For the H-atom we need to identify the appropriate representations that correspond to a subset of these Casimir eigenvalues. In $d = 2, 3$ we saw that we had to restrict to a subset also.

The degeneracy computed in eq.(9.11) identifies the number of states that should be found in the $SO(d+1)$ multiplet as $D_n(d) = N(n-1,\, d+1)$. Namely, the correct multiplet corresponds to the traceless symmetric tensor of rank $(n-1)$ in $(d+1)$ dimensions. That is, we are invited to think of the Hilbert space of the H-atom as being isomorphic to the Hilbert space of angular momentum in $(d+1)$ dimensions. The quadratic Casimir operator is then easily evaluated: we just apply the same formulas derived in Chapter 6, eq.(6.90), but with the substitution $l \to (n-1)$ and $d \to (d+1)$, namely

$$C_2(n-1, d+1) = \hbar^2(n-1)[(n-1) + (d+1) - 2)]. \tag{9.45}$$

Therefore the Hamiltonian has the eigenvalue

$$\begin{aligned} E_n(d) &= -\frac{mc^2(Z\alpha)^2}{2}\left[\frac{1}{\hbar^2}C_2 + \left(\frac{d-1}{2}\right)^2\right]^{-1} \\ &= -\frac{mc^2(Z\alpha)^2}{2}\left[n + \frac{d-3}{2}\right]^{-2} \end{aligned} \tag{9.46}$$

which agrees completely with the Schrödinger equation approach.

The mathematical structure of the symmetry is now fully clarified. The symmetry is isomorphic to rotations in $d+1$ dimensions and the Hilbert space of the H-atom is isomorphic to angular momentum states in $d+1$ dimensions. We conclude that the H-atom in d dimensions is a quantum system isomorphic to a particle moving on the surface of the unit sphere in $d+1$ dimensions. The Hamiltonian of such a particle reduces precisely to C_2 when the constraint

$$x_\mu = (x_0, \mathbf{x}), \qquad x_0^2 + \mathbf{x}^2 = 1 \tag{9.47}$$

is taken into account

$$C_2 = \sum_{\mu=0}^{d} (p_\mu)^2 = \frac{1}{2}\sum_{\mu\neq\nu} (L_{\mu\nu})^2\,, \tag{9.48}$$

with

$$L_{\mu\nu} = -i\hbar(x_\mu\partial_\nu - x_\nu\partial_\mu). \tag{9.49}$$

The Hilbert space of the particle problem is precisely the traceless symmetric tensors $T_{\mu_1\mu_2\cdots\mu_{(n-1)}}(x)$ constructed from a unit vector in $d+1$ dimensions. The Hamiltonians of the particle problem in $d+1$ dimensions and of the H-atom are related to each other as given above.

9.4 Interacting oscillators and dynamical symmetries.

There are many problems in physics that can be formulated (at least approximately) in terms of harmonic oscillator creation-annihilation operators and their interactions.

We recall from Chapter () that any multiparticle Hamiltonian which is at the most a quadratic polynomial in the positions or momenta may always be diagonalized exactly in terms of normal modes that correspond to independent harmonic oscillators. For the purpose of the present section, these harmonic oscillators will be called "free", and the free Hamiltonian will be denoted by

$$H_0 = \hbar \sum_{i=1}^{N} \omega_i \, a_i^\dagger a_i \tag{9.50}$$

where ω_i is the frequency of the normal mode i.

In applications these creation-annihilation operators may represent the creation-annihilation of particles (electrons, protons, neutrons, etc.) or collective excitations of a crystal such as phonons, or collective bosons in nuclei that correspond to paired nucleons, etc.. Accordingly $a_i^\dagger, a_i$ may be fermions or bosons and therefore their quantization is specified with canonical anti-commutators or commutators respectively

$$\begin{aligned} fermions \quad &: \quad \left\{a_i, a_j^\dagger\right\} = \delta_{ij}, \quad \{a_i, a_j\} = 0 = \left\{a_i^\dagger, a_j^\dagger\right\} \\ bosons \quad &: \quad \left[a_i, a_j^\dagger\right] = \delta_{ij}, \qquad [a_i, a_j] = 0 = \left[a_i^\dagger, a_j^\dagger\right]. \end{aligned} \tag{9.51}$$

These particles are allowed to interact with each other. During the interaction the particle number may be conserved or violated. Interactions of these types may involve the creation of n particles and the annihilation of m ones, as represented by the operator

$$a_{i_1}^\dagger a_{i_2}^\dagger \cdots a_{i_n}^\dagger a_{j_1} a_{j_2} \cdots a_{j_m} \ . \tag{9.52}$$

It turns out that quite a few physical systems have an approximate description in terms of Hamiltonians that involve particle conserving quartic interactions of the form

$$H = \hbar \sum_{i=1}^{N} \omega_i \, a_i^\dagger a_i + \sum \lambda_{i_1 i_2 j_1 j_2} \, a_{i_1}^\dagger a_{i_2}^\dagger a_{j_1} a_{j_2} + \cdots \ . \tag{9.53}$$

There are, of course, also other systems for which there are particle non-conserving quartic interactions, as well as interactions that are described by other polynomials of creation-annihilation operators (including cubic, etc.) as implied by the dots $\cdots$.

In general it is not easy to find the eigenstates and eigenvalues of these systems. However, in some cases the coefficients $\lambda_{i_1 i_2 j_1 j_2}$ etc. have certain relations among them such that the overall Hamiltonian has a symmetry structure.

Then the symmetry permits an exact quantum mechanical solution of the system. In the following we give a few examples of this type. Some of them are purely mathematical structures that help to explain the idea, while some others are real applications that have been very useful in the description of complex physical systems (interacting boson model, etc.)

9.4.1 SL(2,R) and harmonic oscillator in d-dimensions

Consider the following construction of the generators for SL(2,R)

$$\begin{aligned} J_+ &= \frac{1}{2}\vec{a}^\dagger\cdot\vec{a}^\dagger, \quad J_- = \frac{1}{2}\vec{a}\cdot\vec{a} \\ J_0 &= \frac{1}{2}\vec{a}^\dagger\cdot\vec{a}+\frac{d}{4} \end{aligned} \tag{9.54}$$

where the creation-annihilation operators are vectors in $d-$dimensions $a_I, a_I^\dagger$, $I = 1, 2, \cdots d$. The commutation rules are

$$[J_+, J_-] = -2J_0, \quad [J_0, J_\pm] = \pm J_\pm \ . \tag{9.55}$$

These differ from SU(2) only by a minus sign, i.e. $-2J_0$ instead of $2J_0$. We have learned in problem () that for SL(2,R) we can simultaneously diagonalize J_0 and the quadratic Casimir operator given by

$$\begin{aligned} C_2 &= J_0^2 - \frac{1}{2}\left(J_+J_- + J_-J_+\right) \\ &= J_0(J_0-1) - J_+J_- \end{aligned} \tag{9.56}$$

The states for SL(2,R) are labelled $|j, m>$. In the general case j, m may take values in any of the four unitary representations: principal series, supplementary series, and the two discrete series. We will see that in the present case we obtain all of them, but only for special values of j in each series.

We may consider a Hamiltonian constructed from the oscillators $a_I, a_I^\dagger$, $I = 1, 2, \cdots d$ such that H is a function of only C_2. Let's construct the Hamiltonian (or C_2) in terms of oscillators by substituting the generators in terms of oscillators

$$\begin{aligned} C_2 &= \left(\frac{1}{2}\vec{a}^\dagger\cdot\vec{a}+\frac{d}{4}\right)^2 - \frac{1}{2}\vec{a}^\dagger\cdot\vec{a} - \frac{d}{4} - \frac{1}{4}\vec{a}^\dagger\cdot\vec{a}^\dagger\ \vec{a}\cdot\vec{a} \\ &= \frac{d^2}{16} - \frac{d}{4} + \frac{(d-2)}{4}\vec{a}^\dagger\cdot\vec{a} + \frac{1}{4}\left(\vec{a}^\dagger\cdot\vec{a}\ \vec{a}^\dagger\cdot\vec{a} - \vec{a}^\dagger\cdot\vec{a}^\dagger\ \vec{a}\cdot\vec{a}\right) \end{aligned} \tag{9.57}$$

So, this is a particle conserving type quartic interaction with some special coefficients. In this case we anticipate two symmetries. The first is $SO(d)$, the rotation symmetry in $d-$dimensions, due to the fact that only dot products of the oscillators appear. In addition, we also have an $SL(2,R)$ symmetry since the generators $J_{0,\pm}$ commute with C_2.

There is one further simplification we can make by noticing that the quartic interaction can be rewritten in terms of the Casimir operator of $SO(d)$. Recall from the study of the d-dimensional harmonic oscillator, that the angular momentum operator in $d-$dimensions is

$$L_{IJ} = i\hbar\left(a_I^\dagger a_J - a_J^\dagger a_I\right) \tag{9.58}$$

and the Casimir operator that commutes with all L_{IJ} is

$$\begin{aligned} C_2\left(SO(d)\right) &= \frac{1}{2\hbar^2}\sum_{I,J}\left(L_{IJ}\right)^2 \\ &= (d-2)\ \vec{a}^\dagger\cdot\vec{a} + \left(\vec{a}^\dagger\cdot\vec{a}\ \vec{a}^\dagger\cdot\vec{a} - \vec{a}^\dagger\cdot\vec{a}^\dagger\ \vec{a}\cdot\vec{a}\right) \end{aligned} \tag{9.59}$$

Furthermore, the eigenvalues of angular momentum for the harmonic oscillator in $d-$dimensions are completely determined at level $\vec{a}^\dagger\cdot\vec{a} = n$ as follows (as in eq.())

$$\begin{aligned} \frac{1}{2\hbar^2}\sum_{I,J}\left(L_{IJ}\right)^2 &= l(l+d-2) \\ l &= n, n-2, \cdots, \begin{cases} 0 & if\ n = even \\ 1 & if\ n = odd \end{cases} \end{aligned} \tag{9.60}$$

Putting these results together we can write C_2 for SL(2,R) in terms of the eigenvalues of the harmonic oscillator states

$$\begin{aligned} C_2 &= \frac{1}{4}\left(\frac{d^2}{4} - d + \frac{1}{2\hbar^2}\sum_{I,J}\left(L_{IJ}\right)^2\right) \\ j(j+1) &= \frac{1}{4}\left(\frac{(d-2)^2}{4} + l(l+d-2) - 1\right) \\ &= \frac{1}{4}\left(l_d + \frac{1}{2}\right)^2 - \frac{1}{4} \\ j &= -\frac{1}{2} \pm \frac{1}{2}\left(l + \frac{d-2}{2}\right) \\ j+1 &= \frac{1}{2}\left(l + \frac{d}{2}\right)\ or\quad \frac{1}{2}\left(-l - \frac{d}{2} + 1\right) \end{aligned} \tag{9.61}$$

9.5 Particle in a magnetic field

Consider a spinless particle moving in a time independent magnetic field $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, where $\mathbf{A}$ is the electromagnetic vector potential. The Hamiltonian is

$$\begin{aligned} H &= \frac{1}{2}m\mathbf{v}^2 \\ m\mathbf{v} &= \mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{r}) \end{aligned} \tag{9.62}$$

where $\mathbf{p}$ is the momentum canonically conjugate to $\mathbf{r}$. Note that $\mathbf{v}$ has the commutation rules

$$\begin{aligned}[(mv_I)\,,(mv_J)] &= i\hbar\frac{e}{c}(\nabla_I A_J - \nabla_J A_I) \\ &= i\hbar\frac{e}{c}\,\varepsilon_{IJK}\,B_K(\mathbf{r})\,. \end{aligned} \tag{9.63}$$

The Hamiltonian may be rewritten in the form

$$H = \frac{1}{2m}\left(\mathbf{p}^2 - 2\frac{e}{c}\mathbf{A}\cdot\mathbf{p} + i\frac{e}{c}\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{e^2}{c^2}\mathbf{A}^2\right), \tag{9.64}$$

where the $\boldsymbol{\nabla}\cdot\mathbf{A}$ arises from commuting $\mathbf{p}$ to the right hand side. We will consider a couple of examples of magnetic fields

1. A constant magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ pointing in the z-direction, where B is a constant.
2. A radial magnetic field of the form $\mathbf{B}(r) = (\frac{Br_0^2}{r^2})\,\hat{\mathbf{r}}$ due to a magnetic monopole.

Constant magnetic field

For a constant magnetic field $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ the vector potential is[1]

$$\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}. \tag{9.65}$$

[1] More generally the vector potential takes the form

$$\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r} + \boldsymbol{\nabla}\Lambda$$

where Λ is an arbitrary gauge function of x, y, z. A convenient gauge choice which we use in the text is $\Lambda = 0$, but other gauge choices are also convenient for various purposes. For example $\mathbf{B} = \hat{\mathbf{z}}B$ is also obtained if $\mathbf{A}$ is of the form ($A_1 = 0,\ A_2 = x_1 B,\ A_3 = 0$). This case may be written in the form $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r} + \boldsymbol{\nabla}(Bx_1x_2)$, or $\Lambda = Bx_1x_2$. The Hamiltonian is simply

$$H = \frac{1}{2m}\left(p_1^2 + (p_2 - eBx_1/c)^2 + p_3^2\right).$$

This Hamiltonian can be diagonized in a basis in which p_2 and p_3 are diagonal. Then the motion in the x_1 direction corresponds to a translated harmonic oscillator. The choice of gauge changes the explicit symmetry of the problem. In the present case there is explicit translation invariance in the $\hat{\mathbf{y}}, \hat{\mathbf{z}}$ directions. But the actual physical symmetry as well as the physical results are still the same as the ones presented in the text, as seen in the following discussion.

In the general gauge consider the modified momentum and angular momentum operators

$$\begin{aligned}\bar{\mathbf{p}} &= m\mathbf{v} + \frac{e}{2c}\mathbf{B}\times\mathbf{r} = \mathbf{p} - \frac{e}{c}\boldsymbol{\nabla}\Lambda \\ \mathbf{J} &= \mathbf{r}\times\bar{\mathbf{p}}\end{aligned}$$

The commutation rules of the $\bar{p}_I$ are the same as the canonical p_I, namely

$$\begin{aligned}[r_I, \bar{p}_J] &= i\hbar\,\delta_{IJ}, \\ [\bar{p}_I, \bar{p}_J] &= i\hbar\frac{e}{c}(\nabla_I\nabla_J\Lambda - \nabla_J\nabla_I\Lambda) = 0.\end{aligned}$$

The Hamiltonian takes the form

$$\begin{aligned} H &= \frac{p_3^2}{2m} + \frac{m}{2}\left(v_1^2 + v_2^2\right) \\ &= \frac{1}{2m}\left(p_1^2 + p_2^2 + p_3^2\right) + \frac{e^2B^2}{8mc^2}\left(x_1^2 + x_2^2\right) - \frac{eB}{2mc}L_3 \\ &= \frac{p_3^2}{2m} + \frac{1}{2m}\left(p_r^2 + \frac{L_3^2 - \hbar^2/4}{r^2}\right) + \frac{e^2B^2}{8mc^2}r^2 - \frac{eB}{2mc}L_3 \end{aligned} \tag{9.66}$$

where $r = (x_1^2 + x_2^2)^{1/2}$, $p_r = r^{-1/2}\left(-i\hbar\partial_r\right)r^{1/2}$ are canonical conjugate radial variables in 2-dimensions (see chapter on central force problem) and $L_3 = (x_1p_2 - x_2p_1)$ is the angular momentum in the z-direction. We recognize the 2D harmonic oscillator as part of the Hamiltonian.

This Hamiltonian is evidently symmetric under translations and rotations along the $\hat{\mathbf{z}}$-axis. The generators of these symmetries p_3 and L_3 are constants of motion. The classical motion of the particle can be easily inferred: to keep the constants of motion, the particle must move uniformly along the z-direction while rotating uniformly along the z-direction. So, this is a spiraling motion along the z-axis combined with oscillations perpendicular to the z-axis.

The quantum wavefunction is labelled with the simultaneous eigenvalues of energy $H = E$, angular momentum $L_3 = \hbar\mu$ and momentum $p_3 = \hbar k$. Then $H|E,\mu,k> = E|E,\mu,k>$ is rewritten in the form

$$\begin{aligned} &\left[\tfrac{1}{2m}\left(p_r^2 + \tfrac{\hbar^2(\mu^2-1/4)}{r^2}\right) + \tfrac{m\omega^2}{2}r^2\right]|E,\mu,k> \\ &\qquad = (E + \hbar\omega\mu - \tfrac{\hbar^2k^2}{2m})|E,\mu,k> \end{aligned} \tag{9.67}$$

where $\omega = \frac{eB}{2mc}$. The left hand side is just the 2-dimensional harmonic oscillator Hamiltonian whose eigenvalues are $\hbar\omega\left(n+1\right)$. Recall that the quantum number

---

This allows us to compute easily the commutators of $\mathbf{J} = \mathbf{r} \times \bar{\mathbf{p}}$ with $\mathbf{J}, \mathbf{r}, m\mathbf{v} = \bar{\mathbf{p}} - \frac{e}{2c}\mathbf{B} \times \mathbf{r}$, and $H = \frac{1}{2}m\mathbf{v}^2$ in any gauge as follows

$$\begin{aligned} [J_I, J_J] &= i\hbar\,\varepsilon_{IJK}\,J_K \\ [J_I, r_J] &= i\hbar\,\varepsilon_{IJK}\,r_K, \quad [J_I, \bar{p}_J] = i\hbar\,\varepsilon_{IJK}\,\bar{p}_K \\ [J_3, v_1] &= i\hbar\,v_2, \quad [J_3, v_2] = -i\hbar\,v_1 \\ [J_3, mv_3] &= 0, \quad [J_3, H] = 0, \quad [mv_3, H] = 0. \end{aligned}$$

This shows that in any gauge J_3 and $mv_3 = \bar{p}_3$ commute with the Hamiltonian as well as each other. Hence they generate symmetries, and their eigenvalues label the degenerate energy eigenstates.

In the gauge $\Lambda = 0$ the generators of the symmetries become $mv_3 = p_3$, $J_3 = L_3$, hence the symmetries are translations and rotations along the z-direction. However, even if $\Lambda \neq 0$ breaks the manifest symmetries, as in the example above, there still are "translation" and "rotation" symmetries that lead to the same degeneracies and classification of states. In the example above these are given by $\bar{p}_3 = p_3$ and the following form of J_3.

$$\begin{aligned} J_3 &= x_1\bar{p}_2 - x_2\bar{p}_1 \\ &= x_1p_2 - x_2p_1 + \frac{eB}{2c}(x_2^2 - x_1^2), \\ [J_3, H] &= 0. \end{aligned}$$

n is the sum of a radial quantum number n_r and angular momentum $l = |\mu|$. Thus identifying this eigenvalue with the right hand side we find

$$\begin{aligned} E &= \frac{eB\hbar}{2mc}(n_r + |\mu| - \mu + 1) + \frac{\hbar^2 k^2}{2m}, \\ n_r &= 0, 2, 4, 6 \cdots \\ \mu &= 0, \pm 1, \pm 2, \cdots \end{aligned} \tag{9.68}$$

Note that the vacuum state is infinitely degenerate since $n_r + |\mu| - \mu = 0$ occurs for an infinite number of values. In fact in order to exhibit correctly the degeneracy of each state we define

$$2N = n_r + |\mu| - \mu \ , \tag{9.69}$$

since $n_r + |\mu| - \mu$ is an even integer. Then we may rewrite

$$\begin{aligned} E &= \frac{eB\hbar}{mc}\left(N + \frac{1}{2}\right) + \frac{\hbar^2 k^2}{2m} \\ N &= 0, 1, 2, \cdots \\ \mu &= -N, -N+1, \cdots, 0, 1, 2, \cdots \end{aligned} \tag{9.70}$$

We now see the infinite degeneracy as labelled by the magnetic quantum number μ. A given energy level N may be achieved either by radial excitations (n_r) or by orbital excitations that spiral in the *clockwise* direction (negative μ), or a combination of the two. There is no energy gain for anti-clockwise orbital excitations (positive μ).

The wavefunction in position space $\psi_{n,\mu,k}(\mathbf{r}) =< \mathbf{r}|E, \mu, k >$ is given in terms of the 2D harmonic oscillator radial wavefunction $R_{n_r,\mu}(r)$

$$\begin{aligned} \psi_{N,\mu,k}(\mathbf{r}) &= Ce^{ikx_3}e^{i\mu\phi}R_{n_r,\mu}(r), \\ n_r &= 2N + \mu - |\mu| \end{aligned} \tag{9.71}$$

where the 2D harmonic oscillator radial quantum number n_r is related to the total quantum number N and degeneracy label μ of the current problem. This probability amplitude is consistent with the classical motion described above (spiraling in the z-direction combined with oscillating in the radial direction). For a given energy level N the lowest radial quantum number $n_r = 0$ is obtained for the most negative magnetic quantum number $\mu = -N$, corresponding to a particle that is spiraling in the direction opposite to the magnetic field.

Algebraic method

There is another quick algebraic solution that yields the same results. Since this gives more insight and provides an example of useful methods as well, we describe it here. From (9.63) we have $[v_1, v_2] = i\hbar eB/cm^2$. Except for the overall normalization, these commutation rules are isomorphic to the commutation rules

of position-momentum operators. So we may define one-dimensional harmonic oscillators and rewrite the Hamiltonian in terms of them

$$\begin{aligned} a &= \sqrt{\frac{cm^2}{2\hbar eB}}(v_1 + iv_2), \quad a^\dagger = \sqrt{\frac{cm^2}{2\hbar eB}}(v_1 - iv_2) \\ H &= \frac{p_3^2}{2m} + \frac{m}{2}\left(v_1^2 + v_2^2\right)^2 = \frac{p_3^2}{2m} + \frac{eB\hbar}{mc}\left(a^\dagger a + \frac{1}{2}\right) \end{aligned} \tag{9.72}$$

Therefore, the eigenstates are

$$\frac{1}{\sqrt{N!}}(a^\dagger)^N|0,k>, \tag{9.73}$$

and the spectrum is

$$\begin{aligned} E &= \frac{\hbar^2 k^2}{2m} + \frac{eB\hbar}{mc}\left(N + \frac{1}{2}\right) \\ N &= 0, 1, 2, \cdots \end{aligned} \tag{9.74}$$

This agrees with the spectrum obtained above.

Next we need to understand the role of angular momentum and explain the degeneracy as well. The Hamiltonian is axially symmetric, so it is simultaneously diagonalizable with angular momentum L_3. The commutation rules with the oscillators are

$$[L_3, a] = a, \quad \left[L_3, a^\dagger\right] = -a^\dagger \tag{9.75}$$

Therefore, each time a creation operator is applied on a state it lowers its angular momentum by one unit. So, we must identify the magnetic quantum number of the state as follows

$$|N, \mu, k> = \frac{1}{\sqrt{N!}}(a^\dagger)^N|0, \mu + N, k> . \tag{9.76}$$

However, we still need to find out the allowed values for μ. For this we need the properties of the wavefunction.

The wavefunction is separable as

$$\begin{aligned} \psi_{N,\mu,k}(\mathbf{r}) &= <\mathbf{r}|N, \mu, k> \\ &= \frac{e^{ikx_3}}{\sqrt{2\pi\hbar}}\, F_{N\mu}(x_1, x_2), \end{aligned} \tag{9.77}$$

where $F_{N\mu}$ is computed by exploring the properties of a and $a^\dagger$ as follows

$$\begin{aligned} < x_1, x_2|a =& < x_1, x_2| \left[\sqrt{\tfrac{c}{2\hbar eB}}(p_1 + ip_2) + \sqrt{\tfrac{eB}{8\hbar c}}(x_2 - ix_1)\right] \\ =& \left[\sqrt{\tfrac{\hbar c}{2eB}}(\partial_2 - i\partial_1) + \sqrt{\tfrac{eB}{8\hbar c}}(x_2 - ix_1)\right] < x_1, x_2| \end{aligned} \tag{9.78}$$

Defining

$$w = \sqrt{\frac{eB}{2\hbar c}}\,(x_2 + ix_1) = i\sqrt{\frac{eB}{2\hbar c}} r e^{-i\phi} \tag{9.79}$$

we can write

$$\begin{aligned} < \quad & x_1, x_2|a = \left(\frac{d}{dw} + \frac{w^*}{2}\right) < x_1, x_2| \\ < \quad & x_1, x_2|a^\dagger = \left(-\frac{d}{dw^*} + \frac{w}{2}\right) < x_1, x_2|. \end{aligned} \tag{9.80}$$

This leads to

$$\begin{aligned} F_{N\mu}(x_1, x_2) &= \frac{1}{\sqrt{N!}} < x_1, x_2|(a^\dagger)^N|0, \mu + N > \\ &= \frac{1}{\sqrt{N!}}\left(-\frac{d}{dw^*} + \frac{w}{2}\right)^N F_{0,\mu+N}(x_1, x_2) \\ F_{0,\mu+N}(x_1, x_2) &\equiv < x_1, x_2|0, \mu + N > \end{aligned} \tag{9.81}$$

The ground state wavefunction $F_{0,\mu+N}$ satisfies a differential equation derived by acting with a to the right or left as follows

$$0 =< x_1, x_2|a|0, \mu + N >= \left(\frac{d}{dw} + \frac{w^*}{2}\right) F_{0,\mu+N}. \tag{9.82}$$

The general solution of this first order equation is $F_{0,\mu+N}\,(w, w^*) = f_{\mu+N}(w^*)\exp(-ww^*/2)$. But we also know that the angular momentum of the state is $L_3 = \mu+N$, which demands that the wavefunction be proportional to $\exp(i\phi(\mu + N))$. This fixes $f_{\mu+N}(w^*) = C\,(w^*)^{\mu+N}$. So, we have derived the wavefunction except for the overall normalization

$$\begin{aligned} F_{0,\mu+N}(x_1, x_2) &= C_{\mu+N}\ (w^*)^{\mu+N}\ e^{-\frac{1}{2}ww^*} \\ F_{N\mu}(x_1, x_2) &= \frac{C_{\mu+N}}{\sqrt{N!}}\left(-\frac{d}{dw^*} + \frac{w}{2}\right)^N \left[(w^*)^{\mu+N}\ e^{-\frac{1}{2}ww^*}\right] \\ &= \frac{C_{\mu+N}}{\sqrt{N!}} e^{-\frac{1}{2}ww^*}\left(-\frac{d}{dw^*} + w\right)^N (w^*)^{\mu+N} \end{aligned} \tag{9.83}$$

The vacuum state must have positive values of angular momentum so that its wavefunction is well behaved near the origin $x_1 = x_2 = 0$. Hence $\mu + N \geq 0$, which implies that the degeneracy of the state at energy N is given by

$$\mu = -N, -N + 1, \cdots, 0, 1, 2, \cdots . \tag{9.84}$$

This is the same result obtained above.

Particle in a magnetic monopole field

Consider a particle moving in a radial magnetic field of the form

$$\mathbf{B}(r) = \frac{g}{4\pi r^2}\,\hat{\mathbf{r}} = -\boldsymbol{\nabla}\frac{g}{4\pi r}. \tag{9.85}$$

The value of the magnetic field at $r = r_0$ is normalized to the constant $B_0 \equiv B(r_0) = \frac{g}{4\pi r_0^2}$. The divergence of this field is

$$\boldsymbol{\nabla} \cdot \mathbf{B} = -\boldsymbol{\nabla}^2 \frac{g}{4\pi r} = g\delta^3(\mathbf{r}), \tag{9.86}$$

which indicates that the field is due to a magnetic monopole at the origin whose magnetic charge is equal to the flux passing through the spherical shell at $r = r_0$,

$$g = 4\pi r_0^2 B_0.$$

The vector potential that produces this magnetic field through $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$ is the Dirac monopole field. In spherical coordinates (r, θ, ϕ) it is given as

$$\mathbf{A} = -\frac{g}{4\pi} \frac{\cot\theta}{r} \hat{\boldsymbol{\phi}} + \boldsymbol{\nabla}\Lambda. \tag{9.87}$$

where $\hat{\boldsymbol{\phi}}$ is the unit vector in the ϕ-direction, and Λ is a gauge function that we will fix to $\Lambda = 0$ in our discussion. To verify this result recall that

$$\begin{aligned} \boldsymbol{\nabla} &= \hat{\mathbf{r}}\partial_r + \hat{\boldsymbol{\theta}} \frac{1}{r}\partial_\theta + \hat{\boldsymbol{\phi}} \frac{1}{r\sin\theta}\partial_\phi \\ \partial_\theta \hat{\boldsymbol{\phi}} &= 0, \quad \partial_\phi \hat{\boldsymbol{\phi}} = -(\hat{\boldsymbol{\theta}}\cos\theta + \hat{\mathbf{r}}\sin\theta). \end{aligned} \tag{9.88}$$

The Hamiltonian (9.62) may be rewritten in the form

$$\begin{aligned} H &= \frac{m}{2}\left(v_r^2 + \frac{1}{r^2}(\mathbf{r}\times\mathbf{v})^2\right) \\ v_r &\equiv \frac{1}{2}(\hat{\mathbf{r}}\cdot\mathbf{v} + \mathbf{v}\cdot\hat{\mathbf{r}}) = \frac{1}{m}p_r \end{aligned} \tag{9.89}$$

where $p_r = \frac{1}{r}\left(-i\hbar\partial_r\right) r$ is the radial momentum that is canonically conjugate to r in three dimensions, as seen in previous chapters. This has the familiar form of spherical decomposition. The term $(\mathbf{r}\times m\mathbf{v})^2$ looks like the square of angular momentum . However, since the components of $\mathbf{v}$ do not commute with each other (see (9.63)) the expression for angular momentum needs modification so that it forms the correct Lie algebra. Thus one finds that there is a modified expression that works as follows

$$\begin{aligned} \mathbf{J} &= \mathbf{r}\times(m\mathbf{v}) + \hbar s\hat{\mathbf{r}} \\ [J_I, J_J] &= i\hbar\varepsilon_{IJK}\, J_K \\ s &= -eg/4\pi\hbar c. \end{aligned} \tag{9.90}$$

where s is dimensionless. It is useful to represent $\mathbf{J}$ purely in terms of angles

$$\begin{aligned} \mathbf{J} &= \hbar s\ \hat{\mathbf{r}} + \hat{\boldsymbol{\theta}}\ (\hbar s\cot\theta + \frac{i\hbar}{\sin\theta}\partial_\phi) - \hat{\boldsymbol{\phi}}\ i\hbar\partial_\theta \\ J_\pm &= \hbar e^{\pm i\phi}(\pm\partial_\theta + \cot\theta\,\partial_\phi + \frac{s}{\sin\theta}), \quad J_0 = -i\hbar\partial_\phi \\ (\mathbf{r}\times m\mathbf{v})^2 &= (\mathbf{J} - \hbar s\hat{\mathbf{r}})^2 = \mathbf{J}^2 - \hbar^2 s^2 = \hbar^2[j(j+1) - s^2]. \end{aligned} \tag{9.91}$$

The terms proportional to s are due to the presence of the magnetic field. Thus, the Hamiltonian can be written in terms of the modified angular momentum

$$H = \frac{1}{2m}[p_r^2 + \frac{1}{r^2}(\mathbf{J}^2 - \hbar^2 s^2)]. \tag{9.92}$$

Therefore, H commutes $\mathbf{J}$, which implies that H is invariant under the modified "rotations" generated by $\mathbf{J}$. Note the problem with the same symmetries can be generalized by including also an arbitrary potential $V(r)$

$$H' = \frac{1}{2m}[p_r^2 + \frac{1}{r^2}(\mathbf{J}^2 - \hbar^2 s^2)] + V(r). \tag{9.93}$$

The states may now be labelled as $|E, j, m>$ and we need to find out the allowed values of j, E. To do so we must consider the wavefunction in position space

$$\psi_{Ejm}(\mathbf{r}) = <\mathbf{r}|E, j, m> = R_{Ej}(r)\, g_{jm}(\theta, \phi) \tag{9.94}$$

and impose the structure of $\mathbf{J}$ in terms of angles. Thus, by sandwiching the $J_0, J_\pm, \mathbf{J}^2$ operators between states we derive the following constraints

$$\begin{array}{l} -i\partial_\phi g_{jm}(\theta,\phi) = m\, g_{jm}(\theta,\phi) \\ e^{\pm i\phi}(\pm\partial_\theta + \cot\theta\,\partial_\phi + \frac{s}{\sin\theta}) g_{jm}(\theta,\phi) = \pm\sqrt{j(j+1) - m(m\pm 1)}\, g_{j,m\pm 1}(\theta,\phi) \\ \frac{-1}{\sin^2\theta}[(\sin\theta\,\partial_\theta)^2 + (\partial_\phi - is\cos\theta)^2] g_{jm}(\theta,\phi) = j(j+1) g_{jm}(\theta,\phi) \end{array}$$

The solution is the rotation matrix up to a normalization constant since it satisfies the same conditions (e.g. compare the third equation to eq.() in chapter 7),

$$g_{jm}(\theta,\phi) = C_{jms} D^j_{ms}(\phi,\theta,0) = C_{jms} e^{im\phi} d^j_{ms}(\theta). \tag{9.95}$$

Furthermore, the radial wavefunction satisfies the equation

$$\begin{array}{rcl} [p_r^2 + \frac{\hbar^2 q(q+1)}{r^2} + v(r)]\, R_{Ej}(r) & = & (2mE)\; R_{Ej}(r) \\ j(j+1) - s^2 & \equiv & q(q+1) \end{array} \tag{9.96}$$

which is identical to the radial equation of the free particle, as studied in chapter 6, if $v(r) = 0$. In this case, its solutions are given in terms of spherical Bessel functions $R_{Ej}(r) = j_q(kr)$, and the energy $E = \hbar^2 k^2/2m$ is continuous.

We see that, as compared to the free particle angular solution $Y_{lm}(\theta,\phi) \sim e^{im\phi} d^j_{m0}(\theta)$, the angular solution in the present case involves $d^j_{ms}(\theta)$, where the non-zero s is due to the magnetic field. However, for consistency with angular momentum quantization, it must be that s is quantized! Since the allowed quantized values of j are either half integer or integer, so the allowed quantized values of s are also half integers or integers. Therefore

$$|s| = \left|\frac{eg}{4\pi\hbar c}\right| = \text{integer or } \left(\text{integer} + \frac{1}{2}\right) \tag{9.97}$$

This quantization seems surprising at first, but it is consistent with other arguments that the charges of monopoles must be quantized. Therefore, its magnetic

field is also quantized. The value of $|s|$ is a-priori determined by the magnetic monopole at the origin. The minimum non-trivial value of $|s|$ is 1/2.

Now we are prepared to give the allowed values of j. Since s plays the role of third component of angular momentum in $d^j_{ms}(\theta) =< jm|\exp(iJ_2\theta)|js >$, then $j \geq |s|$. Thus the solution is completed by giving the following list of the allowed values for the quantum numbers

$$\begin{aligned} j &= |s|\,,\ |s|+1,\ |s|+2,\ |s|+3,\cdots \\ q &= [(j+1/2)^2 - s^2]^{1/2} \\ \psi_{Ejm}(\mathbf{r}) &= C_{jms}\ j_q(kr)\, d^j_{ms}(\theta)\, e^{im\phi}. \end{aligned} \tag{9.98}$$

We see that there is a minimum value of q

$$q_{\min}(s) = -1/2 + \sqrt{|s| + 1/4}. \tag{9.99}$$

Hence, for non-zero s the wavefunction must vanish at the origin, which means that the particle has little probability to sit on top of the monopole at the origin.

9.6 PROBLEMS

1. Compute $[A_I, H]$ and show that it is zero. It is useful to use as much as possible that L_{IJ} commutes with scalars. For some of the computation, it is also useful to recall that $\mathbf{p}^2$ can be written in terms of the radial momentum and angular momenta

$$\mathbf{p}^2 = p_r^2 + \frac{1}{2r^2} L_{IJ}^2, \tag{9.100}$$

and that the commutator between angles and the radial momentum vanishes $[\hat{r}_I, p_r^2] = 0$.

2. Compute $[A_I, r_J]$ that appeared in eq.(9.18). It is useful to use as much as possible the properties of $[L_{IJ}, r_K]$. Give your answer in terms of the quantities $D = \frac{1}{2}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r})$, the antisymmetric L_{IJ}, and the symmetric traceless tensor $Q_{IJ} = r_I p_J + r_J p_I - \frac{2}{d} D$. Recall that the set (L_{IJ}, Q_{IJ}) form together the $SU(d)$ Lie algebra and D, which is the dimension operator, commutes with these dimensionless quantities.

   The infinitesimal transformation that leaves the *classical* action for the H-atom invariant is obtained from $= -\frac{i\varepsilon_J}{\hbar}[A_I, r_J]$ as a linear combination of these operators (find it). Now, replacing $\mathbf{p} = m\dot{\mathbf{r}}$ and ignoring orders of operators classically show that it reduces to

$$\begin{aligned} \delta_\varepsilon \mathbf{r} &= \boldsymbol{\varepsilon}\ \mathbf{r}\cdot\dot{\mathbf{r}} - \mathbf{2}\boldsymbol{\varepsilon}\cdot\mathbf{r}\ \ \dot{\mathbf{r}} + \boldsymbol{\varepsilon}\cdot\dot{\mathbf{r}}\ \ \mathbf{r} \\ &= \mathbf{r}\times(\boldsymbol{\varepsilon}\times\dot{\mathbf{r}}) + \boldsymbol{\varepsilon}\times(\mathbf{r}\times\dot{\mathbf{r}}). \end{aligned} \tag{9.101}$$

   (the first line is in any dimension, the second one is valid only in three dimensions). Next show that the classical Lagrangian $L = \frac{1}{2}m\dot{\mathbf{r}}^2 + \frac{Ze^2}{r}$

transforms to a total time derivative, which is consistent with the symmetry of the action (see chapter on symmetry)

$$\begin{aligned} \delta\varepsilon L &= \partial_t \Lambda_\varepsilon \\ \Lambda_\varepsilon &= -m(\boldsymbol{\varepsilon} \times \dot{\mathbf{r}}\,) \cdot (\mathbf{r} \times \dot{\mathbf{r}}) - Ze^2 \frac{\boldsymbol{\varepsilon} \cdot \mathbf{r}}{r}. \end{aligned} \tag{9.102}$$

3. Compute the commutator $[A_I, A_J]$ that appeared in eq.(9.19). It helps to first compute $[A_I, r_J]$ and $[A_I, p_J]$ as in problem 2.

4. Compute $\sum_I A_I^2$ that appeared in eq.(9.27) and verify the result given in the text.

5. Consider the Hydrogen atom perturbed by external fields or forces such that

$$\begin{aligned} H &= H_0 + H_1 \\ H_0 &= \frac{\mathbf{p}^2}{2m} - \frac{Ze^2}{r} \\ H_1 &= \varepsilon\, A_3 \end{aligned} \tag{9.103}$$

where $\mathbf{A} = \frac{1}{2}(\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}) - mZe^2\hat{\mathbf{r}}$ is the Runge-Lenz vector, and ε is a positive parameter. Using the $SU(2)_+ \otimes SU(2)_-$ symmetry structure of the problem, as given in the attached class notes, solve *exactly* for the eigenvalues and eigenstates of the total Hamiltonian.

Then plot the 4 states for the $n = 2$ level, indicating their relative energies, together with their quantum numbers.

9.8 Particle on spherical shell in monopole field

The problem we have just solved may be modified by putting a constraint that the particle be restricted to move on the surface of the spherical shell at $r = r_0$. This implies that the radial momentum vanishes on the wavefunction $p_r\psi = 0$. Therefore, the Hamiltonian in (9.89,9.92) reduces to

$$\begin{aligned} H &= \frac{m}{2r_0^2}(\mathbf{r}\times\mathbf{v})^2 \\ &= \frac{1}{2mr_0^2}(\mathbf{J}^2 - \hbar^2 s^2) \\ &= \frac{-\hbar^2}{2mr_0^2}\frac{1}{\sin^2\theta}[(\sin\theta\,\partial_\theta)^2 + (\partial_\phi - is\cos\theta)^2] \end{aligned} \tag{9.116}$$

We have already solved this problem above, and found the solution listed in (9.98), but in this case the energy is quantized

$$\begin{aligned} E(j) &= \frac{\hbar^2\left[j(j+1) - s^2\right]}{2mr_0^2}, \\ \psi_{jm}(\hat{\mathbf{r}}) &= C_{jms}\; d^j_{ms}(\theta)\, e^{im\phi} \end{aligned} \tag{9.117}$$

Let us now take $r_0 = 1$ in some units, and consider the stereographic projection of the spherical shell on the plane that passes through the equatorial plane. The mapping is done by drawing a straight line from the north pole to a point on the plane. The line cuts the sphere at a point parametrized by the spherical coordinates (θ, ϕ), and cuts the plane at a point parametrized by the cylindrical coordinates (ρ, ϕ). Note that we have the same coordinate ϕ in both cases, as can be seen by drawing a picture. Graphically it is also easy to see that

$$\rho = \cot\frac{\theta}{2}. \tag{9.118}$$

The north pole ($\theta = 0$) is mapped to infinity ($\rho = \infty$) and the south pole ($\theta = \pi$) is mapped to the origin ($\rho = 0$). We next make one more change of variables to a set of generalized coordinates ($q_1 = \ln\rho,\ q_2 = \phi$) whose ranges are $-\infty < q_1 < \infty$, and $0 \le q_2 \le 2\pi$, corresponding to the infinite strip. The relation between the spherical angles and the variables on the strip are now

$$e^{q_1 + iq_2} = \cot\frac{\theta}{2}e^{i\phi}. \tag{9.119}$$

The Hamiltonian takes a simple form in terms of the variables on the strip by noting that

$$\begin{aligned} \sin\theta\;\partial_\theta &= \partial_{q_1}, \\ &= \frac{-1}{\sin^2\theta}[(\sin\theta\,\partial_\theta)^2 + (\partial_\phi - is\cos\theta)^2] \end{aligned}$$

Chapter 10

VARIATIONAL METHOD

With this chapter we begin the study of approximation methods in quantum mechanics. We will study several methods that yield approximate results of different kinds. The variational method, the WKB approximation, and perturbation theory, both time independent and time dependent, are the major approximation techniques. The variational method is a "quick and dirty" approach to obtain an approximate description of the ground state energy and wavefunction. The accuracy of the results depend mostly on the physical intuition that one has about the system. The WKB approximation yields the leading orders in an expansion of $\hbar$. Thus, it is a semi-classical approximation and works well for states with large quantum numbers. Finally perturbation theory is a systematic expansion in a small parameter that represents a small perturbation of an exactly solvable quantum mechanical system. In principle, with perturbation theory one may compute systematically to all orders, but in practice it is difficult to compute beyond a few orders. Low order perturbation theory works well when the parameter of expansion is indeed small.

The variational method may be applied to simple as well as complicated systems. It is worth trying this approach for a quick estimate of the properties of the system in its ground state. One may also obtain less reliable information on the next few excited states. The method requires a guess for the form of the ground state. The closer the guess is to the actual ground state, the better the results are. But, in the absence of an exact computation, there is no way to know how good the guess is. This is why a good physical intuition about the physical system is of great importance.

10.1 Variational theorems

The following two theorems are at the basis of the variational method. Consider any Hamiltonian H, and any quantum state $|\psi>$, and define the expectation

value for the energy in this state

$$E(\psi) = \frac{<\psi|H|\psi>}{<\psi|\psi>} .$$

Theorem 1 *When $|\psi>$ is varied $|\psi> \rightarrow |\psi + \delta\psi>$, then the energy $E(\psi)$ is stationary when $|\psi>$ is in the vicinity of an eigenstate of the Hamiltonian $|\psi> \sim |E_n>$.*

Theorem 2 *The energy $E(\psi)$ is always larger than the true ground state energy E_0, i.e. $E(\psi) \geq E_0$.*

To prove these theorems consider the variation of $E(\psi)$

$$\delta E(\psi) = \frac{1}{<\psi|\psi>} \ (<\delta\psi|H|\psi> - E(\psi) <\delta\psi|\psi>) + h.c.$$

where *h.c.* stands for hermitian conjugate. If $|\psi>$ is an eigenstate of energy $|E_n>$, then using $H|E_n> = E_n|E_n>$ and $E(\psi) = E_n$, it is evident that $\delta E(\psi) = 0$. This proves the first theorem.

Next consider $E(\psi) - E_0$, multiply the second term with identity in the form $1 = <\psi|\psi> / <\psi|\psi>$, and then introduce the identity operator $\mathbf{1} = \sum_m |E_m><E_m|$ as follows

$$\begin{aligned} E(\psi) - E_0 &= \tfrac{1}{<\psi|\psi>} (<\psi|H|\psi> - E_0 <\psi|\psi>) \\ &= \tfrac{1}{<\psi|\psi>} \textstyle\sum_m (<\psi|H|E_m><E_m|\psi> - E_0 <\psi|E_m><E_m|\psi>) \\ &= \tfrac{1}{<\psi|\psi>} \textstyle\sum_m (E_m - E_0) \ |<\psi|E_m>|^2 \geq 0 \end{aligned} \tag{10.1}$$

The last inequality follows since every term in the sum is positive thanks to the fact that E_0 is, by definition, the lowest energy. This proves the second theorem.

As a result of these theorems it is possible to obtain an estimate for the ground state energy by making an educated guess for the ground state wavefunction. One uses a trial wavefunction $\psi(x, \lambda_i)$ which includes some parameters λ_i that relate to the physical properties of the system. As a function of the parameters one is really considering a family of trial wavefunctions. By minimizing $E(\psi) = E(\lambda_i)$ with respect to these parameters, i.e. $\delta E(\psi) = 0$, or

$$\frac{\partial E}{\partial \lambda_i} = 0, \quad \rightarrow \ \lambda_i = \lambda_i^0$$

one obtains an estimate for the value of $E_0 \sim E(\lambda_i^0)$ (note that E_0 is a lower bound, since $E(\lambda_i) \geq E_0$). Thus the "best" ground state wavefunction within the trial family, $\psi_0(x) \sim \psi(x, \lambda_i^0)$, has the value of the parameters λ_i^0 that minimize the energy. One may improve the estimate of the ground state energy E_0, by including more parameters in the family of wavefunctions.

One may consider excited states as well. The first theorem holds. In the proof of the second theorem, one may substitute E_1, the next excited energy

level. If one insures that the state $|\psi>$ is orthogonal to the true ground state, $<\psi|E_0>=0$, then every term in the sum is positive, and therefore for such $|\psi>$ one gets $E(\psi) \geq E_1$. To consider higher excited states one must choose $|\psi>$ orthogonal to all lower energy eigenfunctions. The difficulty is that the true eigenfunctions are unknown. At best, one may compute at each step an estimate for the energy eigenfunctions, such as $\psi_0(x) \sim \psi(x, \lambda_i^0)$, etc.. Therefore, errors accumulate as one considers higher energy levels, so that the estimates become less and less reliable. The situation improves if there are additional conserved quantum numbers, such as angular momentum, that help define orthogonal eigenfunctions. Then one may obtain a good estimate for the lowest energy level with a fixed angular momentum $l = 0, 1, \cdots$,etc.

10.2 Examples

EXAMPLE 1.– Consider the Hamiltonian for a 1-dimensional infinite well

$$H = -\frac{\hbar^2}{2m}\partial_x^2 + V(x), \qquad \text{where} \quad V(x) = \begin{cases} 0 & |x| \leq a \\ \infty & |x| > a \end{cases} . \tag{10.2}$$

as given in the figure

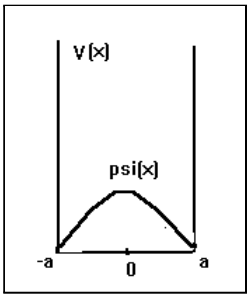


Fig.10.1.—Square well

The true (exact) solution for the ground state is known,

$$\psi_0(x) = \frac{1}{\sqrt{a}} \cos(\frac{\pi x}{2a})\, \theta(a - |x|), \qquad E_0 = \frac{\hbar^2 \pi^2}{8ma^2}.$$

Let us consider the features that are expected on intuitive grounds, and incorporate them in the trial wavefunction. For the ground state one expects an even wavefunction, with the greatest probability near the center of the well, and zero probability on the outside. As a first guess one may choose an non-normalized trial wavefunction without any parameters

$$\psi = \begin{cases} ((a^2 - x^2) & |x| \leq a \\ 0 & |x| > a \end{cases} \tag{10.3}$$

The expectation value of the Hamiltonian is

$$E(\psi) = \frac{-\frac{\hbar^2}{2m}\int_{-a}^{a} dx\ (a^2 - x^2)\frac{d^2}{dx^2}(a^2 - x^2)}{\int_{-a}^{a}(a^2 - x^2)^2 dx} = \frac{10}{\pi^2}\frac{\hbar^2\pi^2}{8ma^2}$$

which is already pretty close to the exact ground state energy E_0

$$E(\psi) = \frac{10}{\pi^2} E_0 = 1.013\ E_0.$$

One can do better by choosing a wave function which depends on a parameter λ. For example, let

$$\psi = \begin{cases} a^\lambda - |x|^\lambda & |x| \le a \\ 0 & |x| > a \end{cases} . \tag{10.4}$$

The

$$E(\psi) = \frac{(\lambda+1)(2\lambda+1)}{(2\lambda-1)} \frac{\hbar^2}{4ma^2}$$

Imposing now $\frac{dE}{d\lambda} = 0$ and solving for λ, one gets

$$\lambda_0 = \frac{1+\sqrt{6}}{2} \simeq 1.72$$

So that

$$E(\lambda_0) = \frac{5+2\sqrt{6}}{\pi^2} \frac{\pi^2\hbar^2}{8ma^2} \simeq 1.00298\ E_0$$

The approximate answer agrees with the exact one within 0.3%, and it is a great improvement over the no-parameter trial wavefunction chosen originally.

The next excited state can be considered by taking a trial wavefunction that is odd. Recall that in this problem there is a parity symmetry. An odd wavefunction is automatically orthogonal to the ground state wavefunction which is even. Therefore, the variational method can be applied in a straightforward manner to get an estimate of the first excited state. This is left as an exercise for the student (see problem ()).

EXAMPLE 2.—Suppose we now want to solve the problem of the H-atom by means of the variational method. The Hamiltonian is

$$H = \frac{\mathbf{p}^2}{2m} - \frac{Ze^2}{r}.$$

We need to make an educated guess for ψ. By studying just the asymptotic behavior of the Schrödinger equation, one expects an exponential fall-off. Therefore, one may take the trial wavefunction for the ground state

$$\psi = e^{-r/a}$$

where a is a variational parameter. The norm is

$$< \psi|\psi > \quad = \int e^{-\frac{2r}{a}} d^3r = \pi a^3 \tag{10.5}$$

and the expectation for the energy is

$$\begin{aligned} E(a) &= \tfrac{1}{<\psi|\psi>} < \psi|H|\psi > \\ &= \tfrac{1}{\pi a^3} \int e^{-\frac{r}{a}} \left(-\tfrac{\hbar^2 \boldsymbol{\nabla}^2}{2m} - \tfrac{Ze^2}{r} \right) e^{-\frac{r}{a}} d^3r \\ &= \tfrac{\hbar^2}{2ma^2} - \tfrac{Ze^2}{a} \end{aligned} \tag{10.6}$$

For later purposes it is useful to note that the first term is the expectation value of the kinetic energy, and the second is the expectation value of the potential energy. Minimizing with respect to a yields

$$\frac{dE(a)}{da} = 0 \quad \rightarrow \quad -\frac{\hbar^2}{ma^3} + \frac{Ze^2}{a^2} = 0 \ \rightarrow a = \frac{\hbar^2}{Zme^2} = \frac{a_0}{Z}$$

where a_0 is the Bohr radius. So the minimized energy is

$$\begin{array}{l} E(a) = -\frac{Ze^2}{2a_0} = E_0 \\ \psi_0 = e^{-\frac{Zr}{a_0}} \quad . \end{array} \tag{10.7}$$

This is actually identical to the exact result. The exactness of our result followed from the lucky choice of the trial wave function.

EXAMPLE 3.—Problem: Estimate the first excited state and energy of the H-atom, by considering a state with angular momentum $l = 1$, e.g. $\psi(\mathbf{r}) = R(r)\, Y_{10}(\hat{\mathbf{r}}) = R(r)$ $\cos\theta$, where $R(r)$ is a variational wavefunction.

10.3 Helium atom

Neglecting the motion of the nucleus, but taking into consideration the electron-electron repulsion, the Hamiltonian of the Helium atom is written as follows

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{2e^2}{|\mathbf{r}_1|} - \frac{2e^2}{|\mathbf{r}_2|} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}.$$

If it were not for the repulsive term, this Hamiltonian would be equivalent to the sum of two Hamiltonians for two hydrogen-like systems. In that case the total eigenstate would be the direct product of two hydrogen-like eigenstates

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{n_1 \ell_1 m_1}(\mathbf{r}_1)\ \psi_{n_2 \ell_2 m_2}(\mathbf{r}_2)\ . \tag{10.8}$$

This form neglects the intrinsic spin of the electron, to which we will return later. To take into account the repulsion as well, we make the educated guess that each electron will see an effective potential consisting of the Coulomb attraction of the nucleus, but partially shielded by the other electron. In that case, the trial wavefunction can be taken in the same form as 10.8, but with an effective charge Z replacing the nuclear charge. This effective charge plays the role of the variational parameter. Actually, for more accuracy, we may introduce two parameters Z_1, Z_2 , one for each electron.

Before applying the variational method, let's estimate the energy by adding naively the energy of each electron in hydrogen-like orbits $|\psi > \sim |n_1 l_1 m_1 > |n_2 l_2 m_2 >$. We would then make the following naive guess for the ground state energy

$$E_{n_1 n_2} = E_{n_1}^{(1)} + E_{n_2}^{(2)} = -\left(\frac{Z_1^2}{n_1^2} + \frac{Z_2^2}{n_2^2}\right)\ Ry$$

where $Ry \simeq$ 13.6 eV. For the ground state, both electrons are in hydrogen-like $1S$ states, $|\psi> \sim |1S> |1S>$. If we take naively $Z_1 = Z_2 = 2$, and set $n_1^2 = n_2^2 = 1$, we get

$$E_0 = -8\ Ry$$

For the state $|1S> \otimes |2S>$ or higher ones $|1S> \otimes |n_2 S>$, we need to take into account the screening effect of the $1S$-electron to obtain a fair estimate of the energy the state. In general, for such excited states we have the naive estimate $Z_1 = 2,\ Z_2 = 1$, so that

$$E = -\left[\left(\frac{2}{1}\right)^2 + \left(\frac{1}{n_2}\right)^2\right] Ry., \quad n_2 = 2, 3, 4, \cdots, \infty$$

The ionization energy is defined as the energy corresponding to the escape of the second electron, i.e. $n_2 = \infty$.

$$E_{ion} \simeq -4\ Ry.$$

Consider also the case of two electrons excited to the second energy level, with $Z_1 = \ Z_2 = 2$

$$E(2S, 2S) = -\left[\frac{4}{4} + \frac{4}{4}\right] Ry = -2\ Ry.$$

This energy is roughly twice as high as the ionization energy of one electron.

Collecting these estimates we compare the naive theory to experiment.

state	naive theory	experiment
|1S,1S>	-8 Ry	-5.8 Ry
|1S,2S>	-4.25 Ry	- ... Ry
⋮	⋮	⋮
|1S,∞S>	-4 Ry	–4 Ry
|2S,2S>	-2 Ry	———

We see roughly that it is more favorable to excite one of the electrons to arbitrary levels rather than exciting both electrons to the second level.

The way at which we have looked at our problem up to this point is very inaccurate indeed. We can do better if we introduce the parameters $Z_1,\ Z_2$ as variational parameters and minimize the energy. To find the energy of the $|1S> \otimes |1S>$ state we can start from the trial wave function with $Z_1 = \ Z_2 = Z_{eff}$ and using $a_0 = \hbar^2/me^2$

$$\begin{array}{ll} \psi_0(\mathbf{r}_1, \mathbf{r}_2) & = e^{-Z_{eff}(\frac{r_1}{a_0} + \frac{r_2}{a_0})} \quad \rightarrow \\ E(\psi) = \frac{<\psi|H|\psi>}{<\psi|\psi>} & = \frac{\int\int d^3\mathbf{r}_1\ d^3\mathbf{r}_2\ \psi(\mathbf{r}_1,\mathbf{r}_2)\ H\ \psi(\mathbf{r}_1,\mathbf{r}_2)}{\int\int d^3\mathbf{r}_1\ d^3\mathbf{r}_2\ \psi(\mathbf{r}_1,\mathbf{r}_2)\ \psi(\mathbf{r}_1,\mathbf{r}_2)} \end{array} \tag{10.9}$$

Let us start from the norm in the denominator, which involves two independent integrals, each one the same integral as (10.5) except for substituting

$a = a_0/Z_{eff}$,

$$< \psi|\psi >= \left(\frac{\pi a_0^3}{Z_{eff}^3}\right)^2 \tag{10.10}$$

For the numerator, the expectations of the kinetic and $1/r$ potential energy terms also involve decoupled integrals that can be performed with the same computations as (10.6) except for substituting $a = a_0/Z_{eff}$,

$$< \psi|\frac{\mathbf{p}_1^2}{2m}|\psi >= < \psi|\frac{\mathbf{p}_2^2}{2m}|\psi >= \left(\frac{e^2 Z_{eff}^2}{2a_0}\right)\left(\frac{\pi a_0^3}{Z_{eff}^3}\right)^2$$
$$< \psi| - \frac{2e^2}{|\mathbf{r}_1|}|\psi >=< \psi| - \frac{2e^2}{|\mathbf{r}_2|}|\psi >= \left(-\frac{2e^2 Z_{eff}}{a_0}\right)\left(\frac{\pi a_0^3}{Z_{eff}^3}\right)^2$$

where we have used $a_0 = \hbar^2/me^2$. The only new integral to be evaluated is the expectation of the repulsive term

$$\int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 e^{-Z_{eff}(\frac{r_1}{a_0}+\frac{r_2}{a_0})}\frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}e^{-Z_{eff}(\frac{r_1}{a_0}+\frac{r_2}{a_0})} \tag{10.11}$$

The above integral is of the type

$$\int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 \frac{e^{-Ar_1}e^{-Br_2}}{|\mathbf{r}_1 - \mathbf{r}_2|} = 2(4\pi)^2\frac{A^2 + B^2 + 3AB}{A^2B^2(A+B)^3} \tag{10.12}$$

Here $A = B = \frac{2Z_{eff}}{a_0}$, so that it becomes

$$e^2 \times 2(4\pi)^2\frac{5A^2}{8A^7} = \frac{5(4\pi)^2e^2}{4A^5} = \left(\frac{5e^2}{8}\frac{Z_{eff}}{a_0}\right)\left(\frac{\pi a_0^3}{Z_{eff}^3}\right)^2. \tag{10.13}$$

Putting everything together

$$\begin{aligned} E(Z_{eff}) &= 2\times\left(\tfrac{e^2 Z_{eff}^2}{2a_0}\right) + 2\times\left(-\tfrac{2e^2 Z_{eff}}{a_0}\right) + \left(\tfrac{5e^2}{8}\tfrac{Z_{eff}}{a_0}\right) \\ &= \tfrac{e^2}{a_0}\left(Z_{eff}^2 - \tfrac{27}{8}Z_{eff}\right). \end{aligned} \tag{10.14}$$

Minimizing the energy $\frac{\partial E}{\partial Z_{eff}} = 0$ yields

$$Z_{eff} = \frac{27}{16} = 1.6875. \tag{10.15}$$

We see that the effective charge seen by each electron is less than $Z = 2$, indicating that even when they are in the same orbit, they shield the nucleus from each other to a certain extent. So the best estimate for the ground state energy is

$$E_0(Z_{eff}) = \frac{e^2}{2a_0}2\times\left(\left(\frac{27}{16}\right)^2 - \frac{27}{8}\left(\frac{27}{16}\right)\right) \tag{10.16}$$

$$\approx -5.7\frac{e^2}{2a_0} = -5.7\ Ry. \tag{10.17}$$

Comparing this value with the experimental result of $E_0(expt) = -5.8\,\mathrm{Ry}$, we see that the agreement is not bad, and certainly better than our first guess of $-8\,Ry$.

Similarly, the estimates for the higher excited states can be improved by using two effective charges Z_1, Z_2 as variational parameters, and insuring that the variational state is orthogonal to the ground state that we have just computed above. The orthogonality may be imposed by putting the second electron in an $l = 1$ state. So, the variational state may be taken as

$$\begin{aligned} |1S, 2P > &\rightarrow \psi_{1S2P}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1, Z_1)\ \psi_{210}(\mathbf{r}_2, Z_2) \qquad (10.18) \\ \psi_{100}(\mathbf{r}_1, Z_1) &= e^{-Z_1 r_1/a_0}, \quad \psi_{210}(\mathbf{r}_2, Z_2) = e^{-Z_2 r_2/2a_0}\, r_2\, Y_{10}(\hat{\mathbf{r}}_2). \end{aligned}$$

where the second wavefunction is the $|2P >$ state of a hydrogen-like atom with an effective nuclear charge Z_2. The presence of $Y_{10}(\hat{\mathbf{r}}_2) = \cos\theta_2/\sqrt{4\pi}$ guaranties orthogonality to the ground state independently of its detailed r-dependence. It is also possible to take a $|1S, 2S >$ variational state, however since we do not know the *exact* r-dependence of the ground state wavefunction there is no way to choose the variational wavefunction orthogonal to it. The best we can do is to make $\psi_{1S2S}(\mathbf{r}_1, \mathbf{r}_2)$ orthogonal to the approximate ground state wavefunction (10.9) that we estimated above. Therefore it is expected that the errors in the ground state estimate will propagate to the excited state. So one may take a variational state of the form

$$\begin{aligned} &\psi_{1S2S}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1, Z_1)\ \psi_{200}(\mathbf{r}_2, Z_2) \\ &\psi_{100}(\mathbf{r}_1, Z_1) = e^{-Z_1 r_1/a_0}, \\ &\psi_{200}(\mathbf{r}_2, Z_2, c) = e^{-Z_2 r_2/2a_0}\ \left(\tfrac{Z_2 r_2}{2a_0} - c\,(Z_2)\right), \end{aligned} \qquad (10.19)$$

where the constant $c\,(Z_2) = 24Z_2/(8Z_2 + 27)$ is chosen so that the $|2S >$ wavefunction with an unknown Z_2 is orthogonal to the ground state $|1S >$ that has $Z_{eff} = 27/16$ (note that if Z_2 were equal to Z_{eff} then $c = 1$ just like the exact $|2S >$ Hydrogen wavefunction). Then both Z_1 and Z_2 can be used as variational parameters. According to physical intuition based on the screening effects, we expect to find $Z_1 \approx 2$ and $Z_2 \approx 1$ for both variational wavefunctions suggested above. The computations for these variational wavefunctions are left as exercises for the student (see problem (2)). However, there are additional important physical effects that must be taken into account before attempting to improve the variational computation. These have to do with the exclusion principle and the spin states of the two electrons, as discussed below.

Nevertheless, if we ignore the exclusion principle, we find the following variational results (see problem 2). For the $|1s2s >$ configuration the variational calculation gives $Z_1 = 1.9934$, and $Z_2 = 1.232$ and

$$\begin{aligned} E_{1s2s} &= E_1 + E_2 + I \\ &= -4 - 0.81586 + 0.50736 \qquad (10.20) \\ &= -4.3087 \end{aligned}$$

For the $|1s2p>$ configuration the variational calculation gives $Z_1 = 1.99828$, and $Z_2 = 1.0179$ and

$$E_{1s2p} = -4.252 \tag{10.21}$$

Effectively, the $2s$ electron is not as well shielded as the $2p$ electron (compare Z_2 for the two cases). This is reasonable since the $2s$ electron can penetrate more easily the shielding cloud of the $1s$ electron and "see" more of the nucleus. Because of this the $2s$ electron is more attracted to the nucleus relative to the $2p$ electron and its energy is lower $E_{1s2s} < E_{1s2p}$. The more careful analysis, including the exclusion principle, is consistent with this picture as seen below.

10.4 Identical particles

Let us consider a two body wavefunction $\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2)$ that describes the probability amplitude of two identical particles, where a_1, a_2 stand for a collection of quantum numbers such as spin, charge, etc. quantum numbers. The probability that these particles are found within some volume element at locations $\mathbf{r}_1, \mathbf{r}_2$ is $|\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2$. If the particles are identical, an observer will measure the same probability if the two particles are interchanged

$$|\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 = |\psi_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2.$$

This implies that under the permutation of the two particles the wavefunctions can differ at the most by a phase

$$\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2) = e^{i\phi(\psi)} \psi_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1).$$

We now argue that the phase $e^{i\phi(\psi)}$ should be the same for all wavefunctions ψ : The general state of the two particles can always be written in terms of some basis states $\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2) = \sum A_n \, \psi^{(n)}_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2)$. Under the interchange of the two particles each basis state can differ at the most by an overall phase $\exp(i\phi_n)$. Thus, one must have

$$\begin{aligned} \psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2) &= e^{i\phi(\psi)} \psi_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1) \\ \sum A_n \, \psi^{(n)}_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2) &= \sum A_n \, e^{i\phi_n} \, \psi^{(n)}_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1) \\ &= e^{i\phi(\psi)} \sum A_n \, \psi^{(n)}_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1) \end{aligned}$$

If the phases are different for each basis function it is impossible to satisfy these relations. Therefore, the phase must be universal for all wavefunctions $e^{i\phi(\psi)} = e^{i\phi_n} = e^{i\phi}$. Furthermore, by permuting the two particles twice we must return to the same wavefunction

$$\psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2) = e^{i\phi} \psi_{a_2a_1}(\mathbf{r}_2, \mathbf{r}_1) = e^{i2\phi} \psi_{a_1a_2}(\mathbf{r}_1, \mathbf{r}_2)$$

Therefore we must have

$$e^{i2\phi} = 1, \quad \text{or} \quad e^{i\phi} = \pm 1.$$

Let us formally write this result by defining a permutation operator Π which just interchanges *all* the attributes of the two particles (i.e. position, spin, etc.)

$$\begin{aligned} \Pi\psi_{a_1a_2}(\mathbf{r}_1,\mathbf{r}_2) &= \psi_{a_2a_1}(\mathbf{r}_2,\mathbf{r}_1) \\ &= \pm\psi_{a_1a_2}(\mathbf{r}_1,\mathbf{r}_2). \end{aligned}$$

This implies that the operator is diagonal on the two particle wavefunction. Furthermore, note that the Hamiltonian for two identical particles is symmetric under their interchange

$$\Pi H(1,2)\Pi^{-1} = H(2,1) = H(1,2),$$

By multiplying both sides of this equation by Π from the right, we find that the permutation operator commutes with the Hamiltonian,

$$\Pi H = H\Pi \quad \rightarrow \quad [\Pi, H] = 0.$$

This implies that Π is a constant of motion, and that it can be diagonalized simultaneously with the Hamiltonian. Thus, if we pick either the $+1$ or the -1 eigenvalue, it will remain fixed for all times.

It turns out that Nature has chosen a particular sign as follows:

- If the identical particles have integer spins, i.e. they are bosons, then the wavefunction must be symmetric, i.e. Π must have the $+1$ eigenvalue.
- If the identical particles have half-integral spins, i.e. they are fermions, then the wavefunction must be anti-symmetric, i.e. Π must have the (-1) eigenvalue.

Within non-relativistic quantum mechanics there is no fundamental explanation of this fact. But in relativistic quantum field theory this rule is derived from fundamental facts such as relativistic invariance, locality and causality. It turns out that all bosons must satisfy Klein-Gordon type equations with the field quantized with commutators, and all fermions must satisfy Dirac type equations with the field quantized with anticommutators. This explains simultaneously the symmetry or antisymmetry of the wavefunctions as well as the fact that bosons satisfy Bose-Einstein statistics and fermions satisfy Fermi-Dirac statistics. All this goes under the name of the "spin-statistics theorem" in relativistic quantum field theory.

Returning now to the non-relativistic wavefunction, let us assume that we have two identical fermions. Furthermore suppose that their motion is uncorrelated, so that the two particle wavefunction is to be constructed from the product of single fermion wavefunctions $\psi_{a_1}(\mathbf{r}_1)$, $\phi_{a_2}(\mathbf{r}_2)$. However, we must impose antisymmetry, therefore

$$\psi_{a_1a_2}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}\left(\psi_{a_1}(\mathbf{r}_1)\ \phi_{a_2}(\mathbf{r}_2) - \psi_{a_2}(\mathbf{r}_2)\ \phi_{a_1}(\mathbf{r}_1)\right).$$

Even if there is no interaction between the particles, the antisymmetry creates a correlation whose effects will be clearly demonstrated below. The clearest sign of a correlation is the fact that if the quantum numbers are the same then the identical particles cannot be found at the same point $\mathbf{r}_1 = \mathbf{r}_2$ since then the wavefunction vanishes. This is Pauli's famous exclusion principle.

10.5 Helium and the exclusion principle

Let us now return to the variational wavefunctions for Helium. Even though the Hamiltonian is independent of the spin of the electrons, the wavefunctions must be labelled by their spin quantum numbers in addition to the orbital quantum numbers $\psi_a(\mathbf{r}) \to \psi^{nlm_l,}_{m_s}(\mathbf{r})$ with $m_s = \pm\frac{1}{2}$ for each electron. In the present case the space dependence of the wavefunction is the same whether the electron spins up or down, and we may write it in the form of a 2-dimensional spinor $\psi_{nlm_l,m_s}(\mathbf{r}) = \psi_{nlm_l}(\mathbf{r})\chi_{m_s}$, where χ_{m_s} is the spinor $\binom{1}{0}$ for spin up and $\binom{0}{1}$ for spin down. The antisymmetric variational state is

$$\begin{gathered}\psi^{n_1l_1m_{l_1}}_{m_{s_1}}(\mathbf{r}_1)\ \psi^{n_2l_2m_{l_2}}_{m_{s_2}}(\mathbf{r}_2) - \psi^{n_1l_1m_{l_1}}_{m_{s_2}}(\mathbf{r}_2)\ \psi^{n_2l_2m_{l_2}}_{m_{s_1}}(\mathbf{r}_1) = \\ \psi^{n_1l_1m_{l_1}}(\mathbf{r}_1)\ \chi^{(1)}_{m_{s_1}}\ \psi^{n_2l_2m_{l_2}}(\mathbf{r}_2)\,\chi^{(2)}_{m_{s_2}} - \psi^{n_1l_1m_{l_1}}(\mathbf{r}_2)\ \chi^{(1)}_{m_{s_2}}\ \psi^{n_2l_2m_{l_2}}(\mathbf{r}_1)\,\chi^{(2)}_{m_{s_1}}\,.\end{gathered} \tag{10.22}$$

When the two electrons are in the ground state $(n_1l_1m_{l_1}) = (n_2l_2m_{l_2}) = (100)$, this structure is necessarily symmetric under the interchange of the coordinates and antisymmetric in the spin indices m_{s_1}, m_{s_2}

$$\psi_{1S1S}(\mathbf{r}_1,\mathbf{r}_2) \equiv \psi_{100}(\mathbf{r}_1, Z_{eff})\ \psi_{100}(\mathbf{r}_2, Z_{eff})\ \chi^{S=0}_{m_{s_1}m_{s_2}}. \tag{10.23}$$

The spin part is the antisymmetric 2×2 matrix

$$\chi^{S=0}_{m_{s_1}m_{s_2}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_{m_{s_1}m_{s_2}} \tag{10.24}$$

which has only one component, implying that the spins of the two electrons must couple to total spin $S = 0$. The space part is the same as the one assumed in a previous section, so that the estimates obtained for $Z_{eff} = 27/16$ and the energy $E_0 = -5.7$ Ry, are valid after taking into account the exclusion principle.

When one of the electrons is in the ground orbital and the other one is in an excited orbital then it is possible to construct two independent antisymmetric variational states from (10.22) as follows

$$\begin{aligned}\psi^{+}_{1Snl} &= \frac{\chi^{S=0}_{m_{s_1}m_{s_2}}}{\sqrt{2}}\left(\psi_{100}(\mathbf{r}_1,Z_1)\ \psi^{(+)}_{nlm}(\mathbf{r}_2,Z_2)\ + \psi_{100}(\mathbf{r}_2,Z_1)\ \psi^{(+)}_{nlm}(\mathbf{r}_1,Z_2)\right) \\ \psi^{-}_{1Snl} &= \frac{\chi^{S=1}_{m_{s_1}m_{s_2}}}{\sqrt{2}}\left(\psi_{100}(\mathbf{r}_1,Z_1)\ \psi^{(-)}_{nlm}(\mathbf{r}_2,Z_2)\ - \psi^{(-)}_{100}(\mathbf{r}_2,Z_1)\ \psi_{nlm}(\mathbf{r}_1,Z_2)\right)\end{aligned} \tag{10.25}$$

In the first one the spins are coupled to $S = 0$, therefore the orbital part must be symmetrized. In the second one the spins are coupled to a symmetric matrix, of which there are three possibilities that may be written in terms of Pauli

matrices as $\chi^{S=1}_{m_{s_1}m_{s_2}} \sim (\sigma_2\vec{\sigma})_{m_{s_1}m_{s_2}}$, implying that $S=1$. Therefore, the orbital wavefunctions must be anti-symmetrized when $S=1$.

We can now make a brief digression on the spectroscopic notation for multi-electron atoms. The notation goes as follows; consider S_{TOT}, J_{TOT} and L_{TOT}, and write the spectroscopic state, for the system as follows

$$^{(2S_{TOT}+1)}\left[L_{TOT}\right]_{J_{TOT}} \ ,$$

using the letters

$$L_{TOT} = S, P, D, F, G, \cdots$$

instead of the numbers $L_{TOT} = 0, 1, 2, 3, 4, \cdots$. We notice that, in our case, $S_{TOT} = 0, 1$; the He atom thus has its electrons paired either in a singlet state $(2S+1=1)$ or in a triplet state $(2S+1=3)$. Experimentally, the spectroscopic lines for these two spin states are different. The spectroscopic lines belonging to the triplet spin state are called *orthohelium*, and the ones with singlet spin state are called *parahelium*.

Another digression for a technical remark is necessary. In order to satisfy the second variational theorem for excited states, the variational wavefunctions must be chosen orthogonal to all the lower energy level wavefunctions. Since we are trying to keep Z_1, Z_2 as unknown variational parameters, orthogonality to the lower energy orbitals, that have a different value of Z_1, Z_2, is not guaranteed by the standard form of hydrogenic wavefunctions $R_{nl}(r,Z)\,Y_{lm}(\hat{\mathbf{r}})$, and therefore appropriate modifications must be made to $R_{nl}(r,Z)$. For example, for the $|1S, 2S>$ case we must choose

$$\begin{array}{c} \psi_{200}(\mathbf{r},Z_2,c_\pm) = e^{-Z_2 r_2/2a_0}\left(\frac{Z_2 r_2}{2a_0} - c_\pm\right) \\ c_+ = 24Z_2/(8Z_2+27), \quad c_- = 3Z_2/(2Z_1+Z_2) \end{array} \tag{10.26}$$

similar to eq.(10.19), and then insert it in eq.(10.25). However, by virtue of the orthogonality of spherical harmonics, for each new orbital, the highest angular momentum state $l = n-1$, i.e.e $\psi_{n,n-1,0}(\mathbf{r}) = R_{n,n-1}(r,Z)\,Y_{nn}(\hat{\mathbf{r}})$, is automatically guarantied to be orthogonal to all lower energy orbitals, with no modifications to $R_{n,n-1}(r,Z)$. An example of this is the $\psi_{210}((\mathbf{r},Z_2)$ function used in eq.(10.18).

Using these methods an estimate of the excited states can be given. The quantitative details are left as an exercise to the student (see problem (3)). On physical grounds we are expecting that Z_1 should be fixed by the variation around the value $Z_1 \sim 2$ and $Z_2 \sim 1$ (see problem). Here we discuss the qualitative features for the variational states $|1S2S>$ and $|1S2P>$ of eq.(10.25).

The expectation value of the energy takes the form

$$\begin{aligned} E_\pm(Z_1,Z_2) \quad = \quad & \left[\varepsilon_1(Z_1) + \varepsilon_{nl}(Z_2) + I_{nl}(Z_1,Z_2)\right] N^\pm_{nl}(Z_1,Z_2) \\ & \pm \left[2\Delta\varepsilon_{nl}(Z_1,Z_2) + J_{nl}(Z_1,Z_2)\right] N^\pm_{nl}(Z_1,Z_2). \end{aligned}$$

The norm factors of the state are defined by

$$\begin{gathered}\left\langle \psi^{\pm}_{1Snl}|\psi^{\pm}_{1Snl}\right\rangle = N_1(Z_1)\,N_n(Z_2) \pm |N_{nl}(Z_1,Z_2)|^2 \\ N_1(Z_1) =< 100,Z_1|100,Z_1 > \\ N_n(Z_2) =< nlm,Z_2|nlm,Z_2 > \\ N_{nl}(Z_1,Z_2) =< 100,Z_1|nlm,Z_2 >= \delta_{l0}N_{n0}(Z_1,Z_2) \\ N^{\pm}_{nl}(Z_1,Z_2) = \frac{N_1(Z_1)\,N_n(Z_2)}{N_1(Z_1)\,N_n(Z_2)\pm\delta_{l0}|N_{n0}(Z_1,Z_2)|^2}\end{gathered}$$

Note that when $l \neq 0$ the overall norm factor simplifies to $N^{\pm}_{nl}(Z_1,Z_2) = 1$. The energies are

$$\begin{gathered}\varepsilon_1(Z_1) =< 100,Z_1|H_0|100,Z_1 > /N_1(Z_1) \\ \varepsilon_{nl}(Z_2) =< nlm,Z_2|H_0|nlm,Z_2 > /N_n(Z_2) \\ \Delta_{nl}(Z_1,Z_2) = \delta_{l0}\,\frac{<100,Z_1|H_0|n00,Z_2>N_{n0}(Z_1,Z_2)}{N_1(Z_1)\,N_n(Z_2)} \\ I_{nl} \pm J_{nl} = \frac{\left\langle \psi^{\pm}_{1Snl}\left|\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\right|\psi^{\pm}_{1Snl}\right\rangle}{N_1(Z_1)\,N_n(Z_2)}\end{gathered}$$

H_0 is an H-like atom Hamiltonian with a nucleus charge of 2, $H_0 = \mathbf{p}^2/2m - 2e^2/r$. Again, for $l \neq 0$ the cross term is absent $\Delta_{nl}(Z_1,Z_2) = 0$. The $I_{nl}(Z_1,Z_2)$ and $J_{nl}(Z_1,Z_2)$ are called the direct and exchange integrals. They are given by

$$\begin{gathered}I_{nl} = e^2\int\int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \frac{|\psi_{100}(\mathbf{r}_1,Z_1)|^2|\psi_{nlm}(\mathbf{r}_2,Z_2)|^2}{N_1(Z_1)\,N_n(Z_2)\;\;|\mathbf{r}_1-\mathbf{r}_2|} \\ J_{nl} = e^2\int\int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \frac{\psi_{100}(\mathbf{r}_1,Z_1)\psi_{nlm}(\mathbf{r}_1,Z_2)\psi^*_{100}(\mathbf{r}_2,Z_1)\psi^*_{nlm}(\mathbf{r}_2,Z_2)}{N_1(Z_1)\,N_n(Z_2)\;\;|\mathbf{r}_1-\mathbf{r}_2|}.\end{gathered}$$

These integrals can be performed by using the formula (10.12) and its derivatives with respect to A or B.

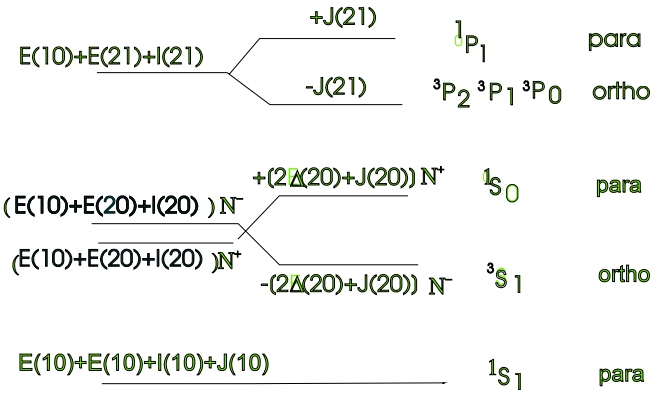


Fig. (10.2) - Qualitative plot of energy levels for helium.

It can be argued that $(\varepsilon_{1+}\varepsilon_n \pm 2\Delta\varepsilon_{n0}\delta_{l0})$, I_{nlm}, J_{nlm}, $N^{\pm}_n$, are all positive quantities, therefore qualitatively we expect the splitting of the energy levels given in Fig.(10.2). Using the procedure outlined in problem 3 we find the variational results summarized below. The experimental measurements for the excited levels of Helium E_{xpt} are included for comparison.

$$
\begin{array}{ll}
\textbf{Parahelium } (\mathbf{S=0}) & \textbf{Orthohelium } (\mathbf{S=1}). \\
\vdots & \vdots \\
2\ {}^1P_1 \left\{ \begin{array}{l} Z_1 = 1.99 \\ Z_2 = 1.0893 \\ E = -4.26138\ Ry \\ E_{xpt} = -4.24\ \ Ry \end{array} \right. & 2\ {}^3P_{2,1,0} \left\{ \begin{array}{l} Z_1 = \\ Z_2 = \\ E = \ Ry \\ E_{xpt} = -4.25\ \ Ry \end{array} \right. \\
2\ {}^1S_0 \left\{ \begin{array}{l} Z_1 = 1.99517 \\ Z_2 = 1.1137 \\ E = -4.2909\ Ry \\ E_{xpt} = -4.290\ \ Ry \end{array} \right. & 2\ {}^3S_1 : \left\{ \begin{array}{l} Z_1 = 2.0085 \\ Z_2 = 1.3891 \\ E = -4.3439\ Ry \\ E_{xpt} = -4.345\ \ Ry \end{array} \right. \\
1\ {}^1S_1 \left\{ \begin{array}{l} Z_1 = Z_2 = 27/16 \\ E = -5.6953\ Ry \\ E_{xpt} = -5.809\ \ Ry \end{array} \right. &
\end{array}
\tag{10.27}
$$

The quantitative estimates provided by the variational scheme above come pretty close to the measured values.

The physical reason for the lower energy in the spin triplet $S = 1$ states (either the $\psi^-_{1S2S} = {}^3S_1$or the $\psi^-_{1S2P} = {}^3P_{2,1,0}$) as compared to the corresponding $S = 0$ states (either the $\psi^+_{1S2S} = {}^1S_0$ or the $\psi^+_{1S2P} = {}^1P_1$) is the antisymmetry versus the symmetry of the wavefunction, as dictated by the Pauli exclusion principle. In the antisymmetric state the probability amplitude vanishes when the electrons get together $\psi^-(\ \mathbf{r}_1 = \mathbf{r}_2) = 0$. This means that the contribution from the positive repulsive energy to the integral $< \psi^-|\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}|\psi^- >$ is less important near $r_1 = r_2 = r$ as compared to the integral $< \psi^+|\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}|\psi^+ >$. Therefore, the antisymmetric state has a lower energy.

Similarly, there is a physical reason for the $|1S2S >$ state to be lower as compared to the $|1S2P >$. One explanation was provided at the end of section (10.3). Further physical arguments can be given as follows: The $2S$ electron is on the average farther from the nucleus as compared to the $1S$ electron. Due to the repulsion between the two electrons, the $1S$ electron is pushed a little closer to the nucleus, so that its negative energy increases. This effect also happens in the $|1S2P >$ state, however it is more important in the $2S$ case, because a $2P$ electron remains further away from the nucleus due to its angular momentum, and thus it cannot push the $1S$ electron toward the nucleus as much as the $2S$ electron can. These physical effects are evident from the fact that the effective charge Z_1 in the $|1S2S >$ case is larger than the $|1S2P >$ case.

10.6 Multi-electron atoms

10.6.1 Hund's rules and their applications

The variational approach and the Pauli exclusion principles can be applied to multi-electron atoms to get some quick estimates. The variational wavefunction

is taken to be a product of H-like wavefunctions for each electron. The Pauli exclusion principle, including spin, requires that each electron uccupies a different state. Thus, in the 1s level there can be at the most two electrons, at the 2s level also two, at the 2p level 6 electrons, an so on. The experimental ordering and maximum occupation numbers of the levels are

$$\begin{aligned}&(1s)^2\,(2s)^2\,(2p)^6\,(3s)^2\,(3p)^6\left[(4s)^2\,(3d)^{10}\right](4p)^6\left[(5s)^2\,(4d)^{10}\right]5(p\\&\left[(6s)^2\,(4f)^{14}\,(5d)^{10}\right](6p)^6\left[(7s)^2\,(5f)^{14}\,(6d)^{10}\right]\cdots\end{aligned}\tag{10.28}$$

where the exponents are the maximum number of electrons allowed at that level. Brackets are placed around those levels that are out of numerical or alphabetical order. For example $4s$ is lower than $3d$, etc. This ordering of levels can be easily remembered through the following grid

$$\begin{array}{ccccccc}1s & \swarrow & \swarrow & & & & \\ 2s & 2p & \swarrow & \swarrow & & & \\ 3s & 3p & 3d & \swarrow & \swarrow & & \\ 4s & 4p & 4d & 4f & \swarrow & \swarrow & \\ 5s & 5p & 5d & 5f & 5g & \swarrow & \swarrow \\ 6s & 6p & 6d & 6f & 6g & 6h & \swarrow \\ 7s & 7p & 7d & 7f & 7g & 7h & 7i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\end{array}\tag{10.29}$$

By reading the grid along the *minor* diagonals, in the direction indicated by the arrows, the correct physical ordering of levels emerge. The significance of this observation is not understood. In this notation the electronic configuration of various atoms are illustrated by the following examples

$$\begin{aligned}H &= (1s)^1, \quad He=(1s)^2, \quad Li=(1s)^2\,(2s)^1, \quad Be=(1s)^2\,(2s)^2,\\ B &= (1s)^2\,(2s)^2\,(2p)^1, \quad C=(1s)^2\,(2s)^2\,(2p)^2, \quad N=(1s)^2\,(2s)^2\,(2p)^3,\\ O &= (1s)^2\,(2s)^2\,(2p)^4, \quad F=(1s)^2\,(2s)^2\,(2p)^5, \quad Ne=(1s)^2\,(2s)^2\,(2p)^6.\end{aligned}\tag{10.30}$$

In these examples He, Be, Ne correspond to completely filled levels. Filled levels always have total angular momentum zero. Because of this their interactions with other atoms is weaker. This is related to the fact that in chemical reactions they are less active than other atoms. To understand this fact as well as the properties of other atoms, we consider a crude approximation: completely filled levels will be considered as part of an effective core that behaves like an effective nucleus of a given charge, and then try to understand the behavior of the remaining outer electrons in unfilled levels. The wavefunctions of each outer electrons $\psi_{\mu m}$ are labelled with intrinsic spin μ and orbital spin m labels, such that $\mu=\pm$ and $m=-l,-l+1,\cdots,l-1,l$. The product of the electronic wavefunctions $\psi_{\mu_1 m_1}\psi_{\mu_2 m_2}\psi_{\mu_3 m_3}\cdots$ have to be totally antisymmetrized due to Pauli's exclusion principle. For example if it is totally symmetric in the μ quantum numbers it must be totally anti-symmetric in the m quantum numbers, and

vice versa. More generally the symmetry could be mixed for μ, then m must have exactly the opposite mixed symmetry. Young tableaux can be used as a mathematical device to keep track of the symmetry. The symmetry properties in μ and m determine the total spin S and total orbital angular momentum L of the state. The total angular momentum J is obtained by using addition of angular momentum, thus $|L-S| \leq J \leq (L+S)$. The values of S, L, J determine also the energy level of the state $E = E(S, L, J)$. Under normal circumstances the atom prefers to be in the lowest energy state. These quantum numbers of the ground state determine the chemical properties of the atom.

Hund discovered that some general rules apply in trying to determine the spin, orbital and total angular momentum quantum numbers of the ground state. Hund's rules are:

1. Combine the intrinsic spins of electrons such as to get the largest possible total spin S_{tot} which is also consistent with the Pauli principle for the overall atom.

2. Coose the maximum value of L consistent with the Pauli principle and with rule #1.

3. If the shell is less than half full, choose the minimum $J = |L-S|$, and if the shell is more than half full choose the maximum value of $J = L+S$.

We now apply Hund's rules to find the ground states of various atoms, and discuss some of their properties.

Helium: $(1s)^2$, $L = 0$ since both electrons are in s states. This is symmetric in L, therefore in spin it must be antisymmetric. Thus we must combine the two spin 1/2 electrons to S=0. Then only $J = \dot{0}$ is possible. The spectroscopic notation for this configuration is ${}^1S_0$ where the superscript corresponds to the number of spin spin states $2S+1 = 1$, the subscript corresponds to total angular monentum $J = 0$, and the letter S corresponds to $L = 0$. The ground state energy was computed earlier in this chapter as $E_{(1s)^2} = (-Ry) \times 2 \times \frac{Z^{*2}}{1^2} \simeq -79\,eV$ with $Z^* = 27/16$. Consider ionizing one electron, then the remaining electron has energy $E = (-Ry) \times \frac{2^2}{1^2} = -54.4\,eV$. Therefore the ionization energy is $\Delta E = (79 - 54.4) = 24.6\,eV$.

Lithium: $(1s)^2\,2s$, S=1/2, L=0, therefore J=1/2. Spectroscopic notation: ${}^2S_{1/2}$. Last electron $\Delta E = (-Ry) \times \frac{Z^{*2}}{2^2}$. Naively $Z^* = 3 - 2 = 1$, but the 2s electron can penetrate the cloud, therefore really $Z^* \simeq 1.26$. Therefore, ionization energy is $\Delta E \simeq 5.4eV$.

Beryllium: $(1s)^2\,(2s)^2$, S=0 antisymmetric, space symmetric L=0, therefore J=0. The spectroscopic notation is ${}^1S_0$. The last two electrons contribute $\Delta E = (-Ry) \times \frac{Z^{*2}}{2^2} \times 2$ to the ground state energy. To estimate the effective charge Z^* one must take into account the fact that the $(2s)^2$ electrons can penetrate the cloud of the $(1s)^2$ electrons, so that the effective charge of the core is $Z = 4-2+\varepsilon$ where $\varepsilon > 0$. However, the $(2s)^2$ electrons also screen each other from the nucleus, almost cancelling ε. The net effect is an effective charge $Z^* \approx 2$ (for a

more careful estimate see problems 4 &5). Therefore $\Delta E\,(Be) \approx -13.6\times\frac{2^2}{2^2}\times 2 \approx -27.2$ eV. If one electron is ionized the remaining 2s electron sees the effective charge of the core which is estimated $2+\varepsilon \approx 2.3$ in problems 4&5. The remaining single electron has energy $-13.6 \times \frac{(2.3)^2}{4} eV \approx -17.9\ eV$. Then the estimate for the ionization energy is

$$27.2\,eV - 17.9\,eV = 9.3\,eV, \tag{10.31}$$

which experimentally correct. In fact, this number is used as an input to justify $\varepsilon \approx 0.3$ above.

Boron: $(1s)^2\,(2s)^2\,2p$, S=1/2, L=1, therefore J=1/2,3/2. The ground state has J=1/2 (Hund2). The spectroscopic notation is ${}^2P_{1/2}$. The ionization energy is 8.3 eV. It is smaller than that of Be since the 2p electron is somewhat higher than the 2s electron (for an estimate compare He excited states 1s2s to 1s2p, the average difference is about 1eV).

Carbon: $(1s)^2\,(2s)^2\,(2p)^2$, S=1 (Hund1) symmetric, L=1 antisymmetric, J=0 (Hund2). Spectroscopic notation: ${}^3P_0$. To estimate the ionization energy we may give a quick and dirty argument similar to that of Beryllium. However, since the mutual screening of the $(2p)^2$ electrons is not as important as that of the $(2s)^2$ electrons of Beryllium, we must use a Z^* somewhat larger than 2. Therefore the ionization energy is higher as compared to Beryllium. The experimental number is approximately 11.3 eV.

Nitrogen: $(1s)^2\,(2s)^2\,(2p)^3$, S=3/2 (Hund1) symmetric, L=1×1×1→0 antisymmetric, J=3/2. Spectroscopic notation: ${}^4S_{3/2}$. Ionization 14.5eV, Z^* increased because repulsion is less important.

Oxygen: $(1s)^2\,(2s)^2\,(2p)^4$, think 2 holes $(2p)^2$, S=1 (Hund1) symmetric, L=1 antisymmetric, J=2 (Hund2, more than half full), Spectroscopic notation: ${}^3P_2$. The fourth electron occupies m_l value already occupied, therefore repulsion is more important, so ionization 13.6eV.

Fluorine: $(1s)^2\,(2s)^2\,(2p)^5$, think 1 hole $(2p)^1$, S=1/2, L=1, J=3/2 (Hund, more than half), Spectroscopic notation: ${}^2P_{3/2}$. Ionization 17.4eV

Neon: $(1s)^2\,(2s)^2\,(2p)^6$, closed shell. S=0, L=0, J=0: ${}^1S_0$. Ionization 21.6eV.

The ionization keeps increasing since Z^* keeps increasing (more electrons, but shielding of nucleus not perfect). It was not monotonic from Beryllium to Boron because the 2p electrons are higher than the 2s electrons. Also not monotonic from Nitrogen to Oxygen because of the mixed symmetry.

Estimating some excited levels and their spectroscopic configurations.

10.7 Problems

1. Using the variational approach, estimate the energy and wavefunction of the *first excited state* for a particle trapped in the one dimensional infinite square well. Note that a good wavefunction would have the properties that it will vanish at the walls and be an odd function so that it will be orthogonal to the ground state.

2. Consider two electrons moving around a core of charge Z (which may be different than 2). The Hamiltonian is

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{Ze^2}{|\mathbf{r}_1|} - \frac{Ze^2}{|\mathbf{r}_2|} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{10.32}$$

   Use the variational principleto find the ground state energy of the system, assuming that both electrons are in the $(1s)$ Hydrogen-like orbitals. This system could describe the $(1s)^2$ core of the Li atom or the Be atom, etc.. What is the effective charge Z^* seen by each electron as a function Z? (recall that for He $Z = 2$, and the effective charge is $Z^* = 27/16$).

3. If the second electron is ionized what is the ground state energy of the remaining $(1s)$ electron? By taking the difference of the energies find the ionization energy of the second electron. Apply your results to the He atom ($Z = 2$) and find the ionization energy (experiment is about 24.6 eV).

4. Repeat problem 2 by assuming that the two electrons are in $(2s)$ states. This could describe the $2s$ electrons in the atom Be ($(1s)^2 (2s)^2$) with an effective core of Z (estimate for Be $Z = 4 - 2 + \varepsilon$, see next problem), or some other system. What is the binding energy of the electrons and what is the effective charge Z^* as a function of Z?

5. Repeat problem 3 by ionizing a $(2s)$ electron. This should provide an estimate for the ionization energy of Be, with $Z = 2+\varepsilon$ (the $(2s)^2$ electrons can penetrate the cloud of the $(1s)^2$ electrons, so the screening of the nucleus from charge 4 to charge 2 is not perfect, therefore $\varepsilon > 0$). Given that the experimental result is 9.3 eV, what is $Z =$?.

6. Apply the variational approach to solve for the ground state energies and wavefunctions of the systems described in the problems at the end of Chapter 1.

7. Using the $|1S, 2S>$ and the $|1S, 2P>$ variational wavefunctions given in the text in eqs.(10.19, 10.18), and neglecting for the moment the effects of the Pauli exclusion principle, estimate the energy of the first excited state. Compare your results to eq.(10.20,10.21).

8. Using the variational wavefunctions $|1S, 2S>$ and $|1S, 2P>$, and taking into account the Pauli exclusion principle, as in eqs.(10.25), compute

the necessary quantities defined in the text as functions of Z_1, Z_2. Then minimize the energies $E^{\pm}(1S2S)$ and $E^{\pm}(1S2P)$ to obtain variational estimates of the energy levels of parahelium and orthohelium. Verify the values of the effective charges and energies given in eq.(10.27)? Amplify the qualitative physical arguments in the text by supporting them with your quantitative computations.

9. Consider two distinguishable particles of spin $\frac{1}{2}$ (e.g. electron and positron) interacting via a spin dependent Hamiltonian of the form (with A, B positive)

$$H = A\vec{S}^{(1)} \cdot \vec{S}^{(2)} + B(S_0^{(1)} - S_0^{(2)}). \tag{10.33}$$

Assume that they are spinning in opposite directions so that the total third component of spin $S_0^{(1)} + S_0^{(2)} = 0$, but the total intrinsic spin $\vec{S}$ may be anything allowed.

   (a) What are the exact states and energies that you expect in the limits $A \to 0$ or $B \to 0$? Give your answer in either the direct product basis $|j_1 m_1 j_2 m_2>$ or total angular momentum basis $|j_1 j_2 j m>$ whichever is most convenient. List all eigenstates, compute their energies and indicate which one is the *ground state* in the appropriate limit.

   (b) Find the estimated ground state energy and wavefuction as a function of A, B using a variational approach. Show that your energy and wavefunction tend to the expected limits as either A or B vanish.

10. Consider the atom Ti that has 22 electrons. Recall that atomic shells are ordered as $1s, 2s, 2p, 3s, 3p,\ 4s, 3d, 4p, 5s, 4d, \cdots$,

   (a) Give the electron configuration for this atom (e.g. $He = (1s)^2$, $Li = (1s)^2(2s)^1$, etc.). Counting both spin and orbital angular momentum states, how many states are available to the outer shell electrons after you take into account the Pauli exclusion principle? (for example $(2p)^2$ has 6 states for each electron and $\frac{6\times 5}{2} = 15$ states for both electrons because of antisymmetrization). Among these states show that the ground state has the spectroscopic configuration ${}^3F_2$ and justify it with appropriate arguments.

   (b) What are the S, L, J quantum numbers of all the other states available to the outer shell electrons? If their energies were proportional only to $J(J+1) - L(L+1) - S(S+1)$ construct a diagram for the energy levels, and give the number of states at each level, together with their spectroscopic notation (the total number of states should match the number of states in part (a)). You may use the table provided for symmetric (S) and antisymmetric (A) products of orbital angular momentum wavefunctions.

11. The deuteron consists of a proton and a neutron which have masses 938 MeV/c^2 and 940 MeV/c^2, respectively. It has a binding energy of 2.23

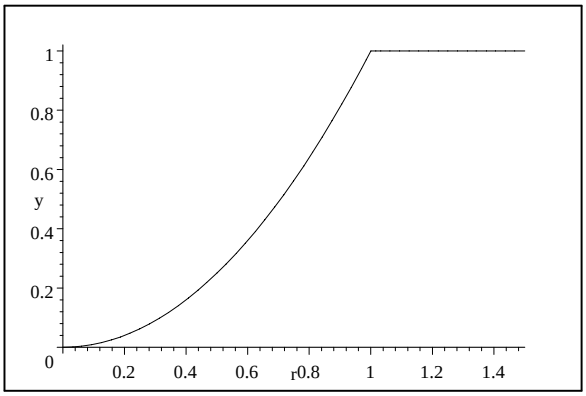
MeV/c^2. In order to obtain an estimate of its structure we can neglect the spins, and treat the system non-relativistically. The Hamiltonian, with the Yukawa potential is given by $H = p^2/2m - V_0 e^{-\mu r}/\mu r$,where m is the reduced mass, $\mu = m_\pi c/\hbar$ with the pion mass $m_\pi = 139\, MeV/c^2$. Consider its ground state with zero angular momentum. The radial wavefunction $f(r) = rR(r)$ must behave asymptotically as $f(r) \to C\exp(-\kappa r)$. As a variational state that is also well behaved at the origin we will take $f(r) = C\exp(-\kappa r)[1 - \exp(-\lambda r)]$, with $\lambda > 0$.

a) Compute κ^{-1} by analyzing the asymptotic behavior of the radial differential equation.

b) Express both κ^{-1} and μ^{-1} in centimeters and by comparing them decide whether the deuteron is a tightly or loosely bound system.

c) Using λ as a variational parameter estimate the ground state energy in terms of the unknown V_0 and then give the value of V_0 that agrees with the observed binding energy. From this value estimate the ratio of the strong interactions to the electromagnetic interactions between two protons at nuclear distances.

12. A particle of mass m is trapped in a 3-dimensional harmonic oscillator type potential with a cutoff at $r = b$, namely $V(\mathbf{r}) = \frac{k\mathbf{r}^2}{2}\theta(b - r) + \frac{kb^2}{2}\theta(r - b)$ as plotted in the figure



V(r)/($\frac{1}{2}$kb^2) plotted versus r/b

The special degeneracies of the 3-dimensional harmonic oscillator are not expected to survive when $b < \infty$. We would like to compute the shifts of the energy levels relative to the $b \to \infty$ case. Recall that in the $b \to \infty$ limit, $V(\mathbf{r}) = \frac{k\mathbf{r}^2}{2}$, the energies are $E_n = \hbar\omega(n + 3/2)$ where $\omega = \sqrt{k/m}$, with the angular momentum taking values $l = n, n-2, \cdots, (0 \text{ or } 1)$.
Use the variational approach to estimate the **ground** state energy. Use some analysis of the Schrödinger equation to guide you in making

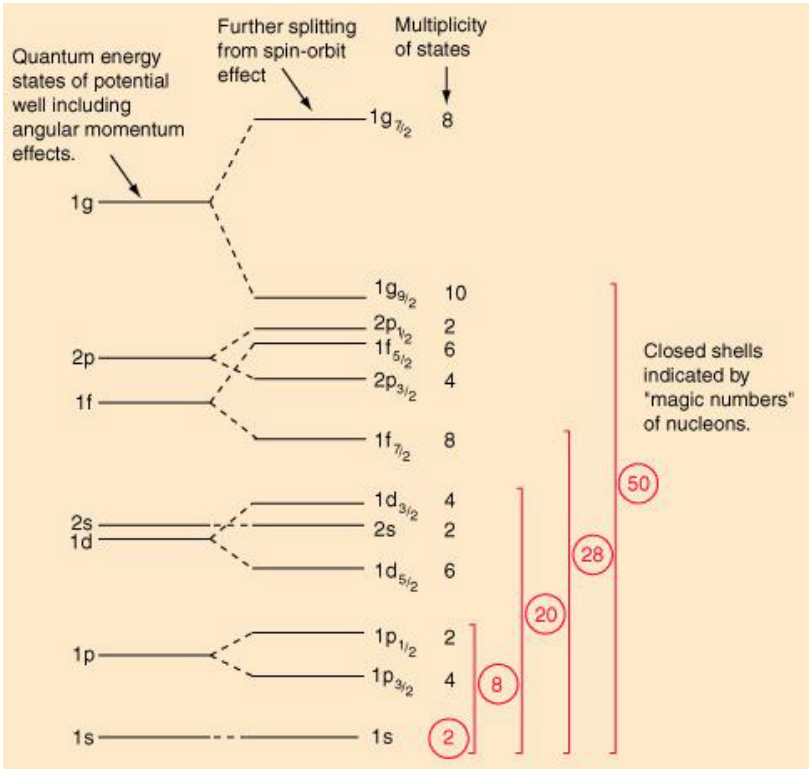
an educated guess for the wavefunction. Propose your best educated guesses for **any** angular momentum l, using **one** variational parameter, and using **two** variational parameters. Explain your reasoning for making those proposals. For the computations that follow use the guess with a single parameter and concentrate only on the $l = 0$ ground state. You can use Mathematica or other programs to compute, plot and do numerical estimates. Give your answer in the form $E = \hbar\omega\varepsilon$, where you estimate ε for several values of b (it is better to use the dimensionless combination $\alpha = \frac{m\omega b^2}{\hbar}$). Provide a table of results for $\varepsilon(\alpha)$ for a range of values of α. Is there a range of α (or b) without any bound states? Test the validity of your estimate by taking the limit $b \to \infty$, and comparing to the known exact result. How close are you?

13. The Hamiltonian for a spinning top and a perturbation is given by

$$H = \frac{1}{2I}\mathbf{J}^2 + aJ_z + bJ_x, \tag{10.35}$$

where $\mathbf{J}$ are angular momentum operators. If $b = 0$ it is evident that the states $|jm>$ are exact energy eigenstates. Use a variational approach with a states $|\theta, jm> = \exp(-i\theta J_y)|jm>$, where θ is a variational parameter, to estimate the new energy levels. Is this method correctly applicable for all values of j, m?

14. The Nuclear Shell Model is based on the 3-dimensional harmonic oscillator hamiltonian H_0 with a perturbation of the form $H_1 = -\overrightarrow{L} \cdot \overrightarrow{S} f(r)$ where f is positive, and $\overrightarrow{S}$ is the spin of the proton or neutron moving in the average harmonic oscillator potential. Use addition of angular momentum, as we did for the H-atom, to give a spectroscopic labelling of the states (nL_j) and show the splitting of the states (watch the sign of the interaction). Now consider putting identical protons and/or neutrons together in these levels to construct an estimate of the ground state of a nucleus (as in the variational approach, but with nothing to vary). Observing the Pauli exclusion principle, find how many identical particles can go on each level? Count the *total* number of particles that give you filled shells. These numbers are the magic numbers of nuclear physics; when a nucleus has that many protons or neutrons it is more stable. Show your reasoning for how you obtain them.



Chapter 11

WKB APPROXIMATION

The Wentzel-Kramers-Brillouin (WKB) approximation is a semiclassical approximation based on an expansion in powers of $\hbar$. Therefore it is expected to work best for states with large quantum numbers (such as energy, angular momentum, etc.) in which one can evoke the correspondance principle between quantum and classical physics. However, in certain cases the WKB method works better than naively expected, and sometimes it may even yield the exact result (such as the harmonic oscillator energy levels). In any case, one may trust the WKB approximation for highly excited states and then try to push its limits toward the low lying states with less confidence. In this sense it is complementary to the variational method that works well near the ground state. Together, these two techniques give quick but not very precise information for low lying states and highly excited states.

11.1 Semiclassical expansion

We will discuss only the problem of a particle in a potential, or the interacting two particle problem that can be reduced to the one particle problem in the center of mass. Similar methods would apply to more complicated systems as well. Thus, consider the Schrödinger equation with potential energy $V(\mathbf{r})$

$$i\hbar\partial_t\psi(\mathbf{r},t) = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right)\psi(\mathbf{r},t). \tag{11.1}$$

Without loss of generality one may write

$$\psi(\mathbf{r},t) = A\exp\left(iW(\mathbf{r},t)/\hbar\right), \tag{11.2}$$

as long as $W(\mathbf{r},t)$ is a complex function. The normalization constant A could be absorbed into W, but it is convenient to keep it separate, so that W may be defined up to an additive complex constant. Substituting this form into the Schrödinger equation one gets

$$\partial_t W + \frac{1}{2m}\left(\boldsymbol{\nabla}W\right)^2 + V - \frac{i\hbar}{2m}\boldsymbol{\nabla}^2 W = 0. \tag{11.3}$$

First consider the limit $\hbar = 0$

$$\partial_t W_0 + \frac{1}{2m} \left(\boldsymbol{\nabla} W_0\right)^2 + V = 0, \tag{11.4}$$

and compare this form to the Hamilton-Jacobi equation for the principal function in classical mechanics

$$\partial_t W_0 + H(\mathbf{r}, \mathbf{p}) = 0. \tag{11.5}$$

It is evident that these two equations are the same for $H = \mathbf{p}^2/2m + V(\mathbf{r})$, provided we identify

$$\mathbf{p}(t) = \boldsymbol{\nabla} W_0(\mathbf{r}, t). \tag{11.6}$$

Thus, the solution of the Hamilton-Jacobi *classical* equations yield the function $W_0(\mathbf{r}, t)$ from which one gets the classical particle trajectories $\mathbf{p}(t)$ as a function of time. Another way of writing this classical result is the familiar form for the current in quantum mechanics

$$\mathbf{p}(t) = -\frac{i\hbar}{2} \left(\psi_0^* \overleftrightarrow{\boldsymbol{\nabla}} \psi_0\right), \tag{11.7}$$

where $\psi_0 = \exp(iW_0(\mathbf{r}, t)/\hbar)$, with real W_0. This means that, in the $\hbar = 0$ limit, the probability amplitude ψ_0 contains all the information about the classical motion of the system. To go beyond classical mechanics consider an expansion in powers of $\hbar$

$$W = W_0 + \hbar W_1 + \hbar^2 W_2 + \hbar^3 W_3 + \cdots . \tag{11.8}$$

Replacing it in (11.3) and collecting powers of $\hbar$, one gets a set of equations that must be satisfied for each order of $\hbar$

$$\begin{array}{l} \partial_t W_0 + \frac{1}{2m} \left(\boldsymbol{\nabla} W_0\right)^2 + V = 0 \\ \partial_t W_1 + \frac{1}{m} \boldsymbol{\nabla} W_0 \cdot \boldsymbol{\nabla} W_1 - \frac{i}{2m} \boldsymbol{\nabla}^2 W_0 = 0 \\ \partial_t W_2 + \frac{1}{m} \boldsymbol{\nabla} W_0 \cdot \boldsymbol{\nabla} W_2 + \frac{1}{2m} \left(\boldsymbol{\nabla} W_1\right)^2 - \frac{i}{2m} \boldsymbol{\nabla}^2 W_1 = 0 \\ \quad \vdots \end{array} \tag{11.9}$$

These equations may be solved in principle by feeding the solutions of the lower orders into the higher orders. It is evident that W is a complex function. For an energy eigenstate the exact eigenfunction has the form

$$\begin{aligned} \psi(\mathbf{r}, t) &= \psi_E(\mathbf{r}) e^{-iEt/\hbar} = A \exp\left[\frac{i}{\hbar}(S(\mathbf{r}) - Et)\right] \\ W &= S(\mathbf{r}) - Et, \qquad S = S_0 + \hbar S_1 + \hbar^2 S_2 + \cdots \end{aligned} \tag{11.10}$$

This implies that all the *explicit* time dependence is in $W_0 = S_0(\mathbf{r}) - Et$, while all higher orders $W_n = S_n(\mathbf{r})$ have vanishing *partial* time derivatives $\partial_t W_1 = \partial_t W_2 = \cdots = 0$. Therefore the equations simplify to

$$\begin{array}{l} \frac{1}{2m} \left(\boldsymbol{\nabla} S_0\right)^2 + V = E \\ \boldsymbol{\nabla} S_0 \cdot \boldsymbol{\nabla} S_1 - \frac{i}{2} \boldsymbol{\nabla}^2 S_0 = 0 \\ \boldsymbol{\nabla} S_0 \cdot \boldsymbol{\nabla} S_2 + \frac{1}{2} \left(\boldsymbol{\nabla} S_1\right)^2 - \frac{i}{2} \boldsymbol{\nabla}^2 S_1 = 0 \\ \quad \vdots \end{array} \tag{11.11}$$

It is not easy to solve these equations in general, however they are quite simple in one dimension. Therefore, consider either a genuine one dimensional problem, or a higher dimensional problem that can be reduced to an effective one dimensional Schrödinger equation because of some symmetry. For example, the central force problem in $d-$dimensions with a rotationally invariant potential $V(r)$ is reduced to the radial Schrödinger equation, which is equivalent to a one dimensional problem in the effective potential (see chapter on the central force problem)

$$\begin{aligned} V_{eff}(r) &= V(r) + \frac{\hbar^2 l_d(l_d+1)}{r^2} \\ l_d &= l + \frac{1}{2}(d-3), \quad l = 0, 1, 2, \cdots . \end{aligned} \tag{11.12}$$

For any such one dimensional problem the Schrödinger wavefunction $\psi_E(x) = A \exp iS(x)/\hbar$ can be reconstructed from the solutions of the simplified equations in one variable

$$\begin{array}{l} \frac{1}{2m}\left(S_0'\right)^2 + V = E \\ S_0' S_1' - \frac{i}{2} S_0'' = 0 \\ S_0' S_2' + \frac{1}{2}\left(S_1'\right)^2 - \frac{i}{2} S_1'' = 0 \\ \qquad \vdots \end{array} \tag{11.13}$$

where the prime is the derivative $S_n' \equiv \partial_x S_n$, and x may be used as the symbol for the radial variable r. It will be useful to consider separately two types of domains of x, called "region I" or "region II", as in Figs.11.1 to 11.5. In these regions we define

$$\begin{cases} E > V(x) & \text{region I}: \quad p(x) \equiv \sqrt{2m\left(E - V(x)\right)} \\ E < V(x) & \text{region II}: \quad \tilde{p}(x) \equiv \sqrt{2m\left(V(x) - E\right)} \end{cases} \tag{11.14}$$

The physical meaning of $E - V(x)$ is the kinetic energy of the particle at the point x. Therefore, one may interpret $p(x)$ as the momentum of the particle at the point x. Then

$$\begin{aligned} S_0' &= \pm p, \quad S_1' = \frac{ip'}{2p}, \\ S_2' &= \pm\left(\frac{3}{8}\frac{(p')^2}{p^3} - \frac{1}{4}\frac{p''}{p^2}\right) \\ &\vdots \end{aligned} \tag{11.15}$$

The solution of the first two equations in the corresponding regions are

$$S_0(x) = \begin{cases} \pm \int^x dx'\ p(x') \\ \pm i \int^x dx'\ \tilde{p}(x') \end{cases}, \quad S_1(x) = \begin{cases} \frac{i}{2}\ln p(x) \\ \frac{i}{2}\ln \tilde{p}(x) \end{cases}. \tag{11.16}$$

The third equation is solved by

$$S_2(x) = \begin{cases} \mp\frac{1}{4}\frac{p'}{p^2} \mp \frac{1}{8}\int^x \frac{(p')^2}{p^3} \\ \pm\frac{i}{4}\frac{\tilde{p}'}{\tilde{p}^2} \pm \frac{i}{8}\int^x \frac{(\tilde{p}')^2}{\tilde{p}^3} \end{cases}. \tag{11.17}$$

Therefore, to first order in $\hbar$ the wavefunction takes the form

$$\begin{aligned} \psi_E(x) &= A \exp\left(iS_0/\hbar + iS_1 + O(\hbar)\right) \\ \psi_E(x) &= \begin{cases} \frac{A_I^{\pm}}{\sqrt{p(x)}} e^{\pm \frac{i}{\hbar}\int^x dx'\ p(x')}\ [1+O(\hbar)] \\ \frac{A_{II}^{\pm}}{\sqrt{\tilde{p}(x)}} e^{\pm \frac{1}{\hbar}\int^x dx'\ \tilde{p}(x')}\ [1+O(\hbar)]\,. \end{cases} \end{aligned} \tag{11.18}$$

The contributions from $S_2, S_3, \cdots$ vanish as $\hbar \to 0$. Physical boundary conditions must be considered in order to decide which combination of solutions are valid in the different regions of type I , II. Such considerations also determine the constants $A_I^{\pm}, A_{II}^{\pm}$ in the appropriate regions.

11.2 Extrapolation through turning points

Before proceeding further, let us investigate the conditions under which the approximation is valid. The obvious criterion is that all neglected terms must be small. Formally, the higher order terms in $\hbar$ are small, however one must take care that the coefficients that multiply these powers are not large. Therefore, the correct criterion is

$$\hbar\,|S_2| << |S_1|\,. \tag{11.19}$$

It is clear from (11.17) that this condition cannot be met in the vicinity of $x \sim x_i(E)$ where the local momentum vanishes $p(x_i) = \tilde{p}(x_i) = 0$, or

$$V(x_i) = E. \tag{11.20}$$

At these points the classical momentum changes sign, thus such points are the classical "turning points". Hence the expression for the wavefunction given above is certainly not valid near the classical turning points. By examining the expressions for S_1, S_2 one notices that, as long as $p(x), \tilde{p}(x)$ are not near zero, and they are slowly varying functions, the condition is met in the domains of x where

$$\frac{\hbar\,|p'|}{p^2} << 1\ , \tag{11.21}$$

and similarly for $\tilde{p}(x)$. The physical meaning of this criterion is understood by defining the local deBroglie wavelength $\lambda(x) = \hbar/p(x)$ and writing

$$\lambda(x)\left|\frac{p'(x)}{p(x)}\right| << 1. \tag{11.22}$$

The local rate of change of momentum is $|p'\delta x/p| = |\delta p/p|$. The rate of change of momentum for one wavelength is approximately obtained by substituting $\delta x \sim \lambda(x)$. Therefore, as long as the rate of change of momentum in the interval of one wavelength is small, the approximation is good in that domain of x. These conditions are met for a slowly varying potential $V(x)$ in the domains that are sufficiently far away from turning points. In such domains the expression for the

wavefunction given above is a good approximation. Since there may be several regions of type I, II for a given potential $V(x)$ one must connect the solutions in different regions to one another. To do so more information near the turning points are needed. So, consider again the exact Schrödinger equation in the form (no expansion in $\hbar$)

$$\left[\hbar^2\partial_x^2 + p^2(x)\right]\psi_E(x) = 0 \tag{11.23}$$

and solve it near a classical turning point, for which

$$p^2(x) \approx \alpha(x-x_i)^n\left[1 + O\left(x-x_i\right)\right]. \tag{11.24}$$

Normally $n = 1$, but it will be kept arbitrary in the present analysis. The solution near the turning point $x \sim x_i$ can be written in the form

$$\psi_E(x) = \begin{cases} A_i^{\pm}\sqrt{\frac{\int_{x_i}^{x} p(x')\,dx'}{p(x)}}\; J_{\pm\frac{1}{n+2}}\left(\frac{1}{\hbar}\int_{x_i}^{x} p(x')\,dx'\right) \\ \mp A_i^{\pm}\sqrt{\frac{\int_{x}^{x_i} \tilde{p}(x')\,dx'}{\tilde{p}(x)}}\; I_{\pm\frac{1}{n+2}}\left(\frac{1}{\hbar}\int_{x}^{x_i} \tilde{p}(x')\,dx'\right), \end{cases} \tag{11.25}$$

where $J_{\pm\nu}(z), I_{\pm\nu}(\tilde{z})$ are the standard Bessel functions with index $\nu = \frac{1}{n+2}$. The two forms of the solution are analytic continuations of each other as $(x-x_i)$ changes sign, hence the solution is valid on both sides of the turning point, with the same overall constants $A_i^{\pm}$. Actually, (11.25) is valid for all values of x strictly only for $p^2(x) = \alpha(x-x_i)^n$. Now consider (11.25) in terms of the general $p^2(x)$ that is not strictly equal to $\alpha(x-x_i)^n$ but which merely behaves like $\alpha(x-x_i)^n$ near the turning point. Then, in an expansion near the turning point, only the leading terms of the form above would be valid. The Schrödinger equation is not expected to be satisfied by (11.25) away from the turning point. What is then the advantage of writing the solution in terms of the general $p(x)$? Here comes the crucial observation: remarkably, the asymptotic forms of $J_\nu(z), I_\nu(\tilde{z})$ for *real* $z, \tilde{z}$

$$\begin{aligned} &J_{\pm\nu}(z) \to \sqrt{\tfrac{2}{\pi z}}\cos\left(z \mp \tfrac{\nu\pi}{2} - \tfrac{\pi}{4}\right)\left[1 + O(\tfrac{1}{z})\right] \\ &I_{\pm\nu}(\tilde{z}) \to \sqrt{\tfrac{1}{2\pi z}}\left\{e^{\tilde{z}}\;\left[1 + O(\tfrac{1}{\tilde{z}})\right] + e^{-i\pi(\frac{1}{2}\pm\nu)}e^{-\tilde{z}}\;\left[1 + O(\tfrac{1}{\tilde{z}})\right]\right\} \\ &I_{-\nu}(\tilde{z}) - I_\nu(\tilde{z}) \to \sin(\nu\pi)\sqrt{\tfrac{2}{\pi\tilde{z}}}e^{-\tilde{z}}\;\left[1 + O(\tfrac{1}{\tilde{z}})\right] \end{aligned} \tag{11.26}$$

are such that, far away from the turning point, i.e. for large $z = \frac{1}{\hbar}\int_{x_i}^{x} p$ or $\tilde{z} = \frac{1}{\hbar}\int_{x}^{x_i}\tilde{p}(x')$, appropriate linear combinations of (11.25) agree with the asymptotic forms (11.18). Therefore, even though one may not trust (11.25) at intermediate points, it is reliable near the turning points as well as in regions asymptotically far away from them. This allows one to use (11.25) to interpolate between regions of type I, II and find the relation between the constants $A_I^{\pm}, A_{II}^{\pm}$ in all asymptotic regions, for any given potential $V(x)$.

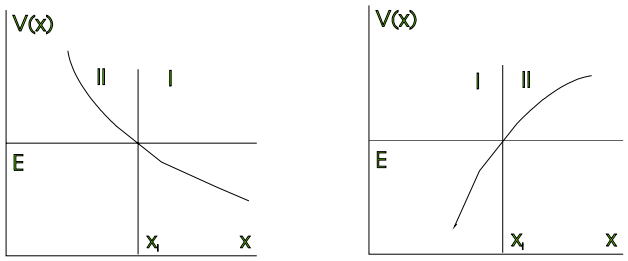


Fig. (11.1) - Regions I and II.

The result of this analysis is the following asymptotic connection formulas: For a potential $V(x)$ that is decreasing through the turning point, with region II on the left of region I, as in Fig. (11.1a), the useful connection formulas are (see problem 1)

$$\begin{aligned} &\frac{\exp\left(-\frac{1}{\hbar}\int_x^{x_i} dx'\,\tilde{p}(x')\right)}{\sqrt{\tilde{p}(x)}} \quad \underset{II\to I}{\rightarrow} \quad \frac{\cos\left(\frac{1}{\hbar}\int_{x_i}^{x} dx'\,p(x') - \frac{\pi}{4}\right)}{\sqrt{p(x)}\ \sin\frac{\pi}{4+2n}} \\ &\frac{\sin\phi\ \exp\left(\frac{1}{\hbar}\int_x^{x_i} dx'\,\tilde{p}(x')\right)}{2\sqrt{\tilde{p}(x)}\ \sin\frac{\pi}{4+2n}} \quad \underset{II\leftarrow I}{\leftarrow} \quad \frac{\cos\left(\int_{x_i}^{x}\frac{1}{\hbar}dx'\,p(x') - \frac{\pi}{4} + \phi\right)}{\sqrt{p(x)}}, \end{aligned} \tag{11.27}$$

where ϕ is an arbitrary phase that gives a linear combination of sines and cosines in the asymptotic region I. Similarly, when the potential $V(x)$ is increasing through the turning point, with region I on the left of region II, as in Fig. (11.1b), the connection formulas are

$$\begin{aligned} &\frac{\cos\left(\frac{1}{\hbar}\int_x^{x_i} dx'\,p(x') - \frac{\pi}{4}\right)}{\sqrt{p(x)}\ \sin\frac{\pi}{4+2n}} \quad \underset{I\leftarrow II}{\leftarrow} \quad \frac{\exp\left(-\frac{1}{\hbar}\int_{x_i}^{x} dx'\,\tilde{p}(x')\right)}{\sqrt{\tilde{p}(x)}} \\ &\frac{\cos\left(\frac{1}{\hbar}\int_x^{x_i} dx'\,p(x') - \frac{\pi}{4} + \phi\right)}{\sqrt{p(x)}} \quad \underset{I\to II}{\rightarrow} \quad \frac{\sin\phi\ \ \exp\left(\frac{1}{\hbar}\int_{x_i}^{x} dx'\,\tilde{p}(x')\right)}{2\sqrt{\tilde{p}(x)}\ \sin\frac{\pi}{4+2n}}\ . \end{aligned} \tag{11.28}$$

In the connection formula $I \to II$ the negligible decaying exponential $\exp(-\tilde{z})$ is dropped. Note that for $n = 1$, which is the usual case, one has $\sin\frac{\pi}{4+2n} = \frac{1}{2}$. For some potentials, such as square wells or barriers, there are special boundary conditions at the turning points, since one cannot write $p^2(x) \sim (x - x_i)^n$. Then the above analysis should be modified. The asymptotic forms (11.18) are still valid, but the interpolating wavefunctions (11.25) must be changed according to the problem at hand. For example, for the infinite square well of the type shown in Fig. (11.2) the wavefunction in region I must vanish at the wall. In this case the connection formula is simply

$$0 \quad \underset{II\leftrightarrow I}{\leftrightarrow} \quad \frac{\sin\left(\frac{1}{\hbar}\int_{x_i}^{x} dx'\,p(x')\right)}{\sqrt{p(x)}}. \tag{11.29}$$

More generally, for a finite square well or barrier, one may use a linear combination of the asymptotic forms, extrapolate it naively all the way to the barrier from both sides, and fix the relative coefficients by requiring continuity of the wavefunction and its derivative at the turning points.

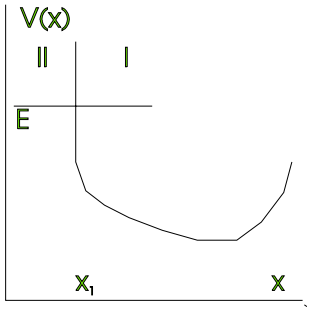


Fig. (11.2) - At infinite wall $\psi(x_1) = 0$.

11.3 Applications

11.3.1 Bound states

Consider the potential well of Fig.(11.3). What are the eigenvalues and eigenstates? Of course, for an arbitrary function $V(x)$ it may not be possible to obtain an exact answer. But the WKB method gives a quick approximate answer that is accurate for sufficiently excited levels.

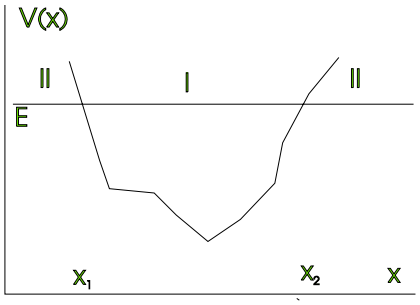


Fig. (11.3) - Potential well.

There are two regions of type II. The physical boundary conditions require that the wavefunction should decrease asymptotically. Therefore the asymptotic solutions in the left and right regions are selected from those given in (11.18) as

$$\psi_L^{II}(x) = \frac{A_L}{\sqrt{\tilde{p}(x)}} \exp\left(-\frac{1}{\hbar}\int_x^{x_1} dx'\,\tilde{p}(x')\right) \tag{11.30}$$

$$\psi_R^{II}(x) = \frac{A_R}{\sqrt{\tilde{p}(x)}} \exp\left(-\frac{1}{\hbar}\int_{x_2}^{x} dx'\,\tilde{p}(x')\right)$$

The connection formulas can be applied from left as well as right in order to obtain the solution in region I far away from the turning points. The two expressions are

$$\begin{aligned} \psi^I(x) &= A_L \frac{\cos\left(\frac{1}{\hbar}\int_{x_1}^{x} dx'\, p(x') - \frac{\pi}{4}\right)}{\sqrt{p(x)}\ \sin\frac{\pi}{4+2n}} \\ &= A_R \frac{\cos\left(\frac{1}{\hbar}\int_{x}^{x_2} dx'\, p(x') - \frac{\pi}{4}\right)}{\sqrt{p(x)}\ \sin\frac{\pi}{4+2n}}. \end{aligned} \tag{11.31}$$

These should be the same function in the domain I. Using $\int_{x_1}^{x} dx'\, p(x') = \int_{x_1}^{x_2} dx'\, p(x') - \int_{x}^{x_2} dx'\, p(x')$, and $\cos(\theta) = (-1)^N \cos(N\pi - \theta)$ one sees that these two expressions are indeed the same provided $A_L = (-1)^N A_R$ and $\int_{x_1}^{x_2} p(x')\, dx' = \hbar\pi\left(N + \frac{1}{2}\right)$. Therefore there is a bound state provided the energy is quantized as follows

$$\int_{x_1}^{x_2} \sqrt{2m(E - V(x))}dx = \hbar\pi\left(N + \frac{1}{2}\right). \tag{11.32}$$

Note that $x_1(E)$ and $x_2(E)$ also contain energy dependence. Solving for $E(N)$ one obtains the quantized energy levels.

As an example consider the harmonic oscillator $V(x) = \frac{m\omega^2}{2}x^2$, with $x_1(E) = -\frac{1}{\omega}\sqrt{2E/m}$ and $x_2(E) = \frac{1}{\omega}\sqrt{2E/m}$. The integral gives

$$\int_{-\frac{1}{\omega}\sqrt{2E/m}}^{\frac{1}{\omega}\sqrt{2E/m}} \sqrt{(2mE - m^2\omega^2x^2)}dx = E\frac{\pi}{\omega} = \hbar\pi\left(N + \frac{1}{2}\right) \tag{11.33}$$

Therefore $E(N) = \hbar\omega(N + 1/2)$. Although we only had the right to expect an approximate result for the excited states, we have obtained an exact result which works even for the ground state. This kind of exact result is an accident (which can be explained!). For additional examples see the problems at the end of the chapter.

Let us now consider a spherical potential in the $l = 0$ angular momentum state, that has an infinite wall at $r = 0$, as depicted in Fig. (11.4).

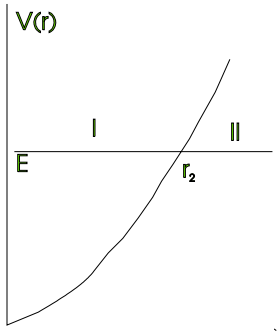


Fig. (11.4) - $V_{eff}(r) = V(r)$ for $l = 0$.

What are the eigenvalues? Near $r = 0$ the wavefunction behaves as r^{l_d+1}, and it has the following asymptotic form in region II on the right

$$\psi_R^{II}(r) = \frac{A_R}{\sqrt{\tilde{p}(r)}} \exp\left(-\frac{1}{\hbar}\int_{r_2}^{r} dr'\, \tilde{p}(r')\right). \tag{11.34}$$

The connection formula gives the wavefunction in region I away from the turning point

$$\psi^I(r) = A_R \frac{\cos\left(\frac{1}{\hbar}\int_r^{r_2} dr'\, p(r') - \frac{\pi}{4}\right)}{\sqrt{p(r)}\ \sin\frac{\pi}{4+2n}} \tag{11.35}$$

According to (11.29) this wavefunction must vanish at the origin. This requires the argument of the cosine to be $(N + 1/2)\pi$. Therefore

$$\int_0^{r_2(E)} dr \sqrt{2m(E - V_{eff}(r))} = \hbar\pi\left(N + \frac{3}{4}\right). \tag{11.36}$$

For example, for the 3-dimensional harmonic oscillator we have $V_{eff}(r) = \frac{m\omega^2}{2}r^2$ in the $l = 0$ state. Doing the integrals we obtain

$$E_{WKB} = \hbar\omega\left(2N + \frac{3}{2}\right) \quad \text{for} \quad l = 0. \tag{11.37}$$

Recall that the exact energy eigenvalues of the 3-dimensional harmonic oscillator are $E_n = \hbar\omega(n+3/2)$, and the angular momentum is $l = n, (n-2), (n-4), \cdots, 0$ or 1. Zero angular momentum is possible only if $n = 2N = even$. Therefore, the exact *zero* angular momentum states are

$$E_{exact} = \hbar\omega\left(2N + \frac{3}{2}\right) \quad \text{for} \quad l = 0. \tag{11.38}$$

We see that the WKB approximation gives the correct behavior in the quantum number N, as well as the correct value of the ground state. Thus, as expected, the WKB approximation gives the correct answer for highly excited states, but the fact that it is also exact for low lying states is an accident for the harmonic oscillator. This latter feature is not expected to be true in general, and for a better approximation for low lying states one should use the variational approach.

As a further example consider the Hydrogen atom for any dimension, and compute the energy eigenvalues using the WKB method. Using the effective potential in the radial equation for any dimension (see Chapter 6), the WKB quantization condition becomes

$$\begin{aligned} I &= \int_{r_1}^{r_2} dr \left(2mE + \frac{2mZe^2}{r} - \frac{\hbar^2 l_d(l_d+1)}{r^2}\right)^{1/2} \\ &= \hbar\pi\left(N + \frac{1}{2}\right) \end{aligned} \tag{11.39}$$

where $l_d = l + (d-3)/2$. The solution for the turning points is (with E = negative):

$$\begin{aligned} r_1 &= \frac{e^2 Z}{2|E|} - \sqrt{\left(\frac{e^2 Z}{2|E|}\right)^2 - \frac{\hbar^2}{2m|E|} l_d(l_d+1)} \\ r_2 &= \frac{e^2 Z}{2|E|} + \sqrt{\left(\frac{e^2 Z}{2|E|}\right)^2 - \frac{\hbar^2}{2m|E|} l_d(l_d+1)} \end{aligned} \tag{11.40}$$

To do the integral it is useful to make a change of variables

$$\begin{aligned} &u = \ln(r), \quad du = \tfrac{dr}{r} \\ &I = \sqrt{2m\,|E|}\int_{\ln r_1}^{\ln r_2} \sqrt{(e^u - r_1)\,(r_2 - e^u)}du \\ &I = \sqrt{2m\,|E|}\ \tfrac{\pi}{2}\left(r_1 + r_2 - 2\sqrt{r_1 r_2}\right) \\ &I = \sqrt{2m\,|E|}\ \tfrac{\pi}{2}\left(\tfrac{e^2 Z}{|E|} - 2\sqrt{\tfrac{\hbar^2}{2m|E|} l_d(l_d+1)}\right) \\ &I = \tfrac{\pi}{2}\left(e^2 Z\sqrt{\tfrac{2m}{-E}} - 2\hbar\sqrt{l_d(l_d+1)}\right) = \hbar\pi\left(N + \tfrac{1}{2}\right) \end{aligned} \tag{11.41}$$

Solving for the energy E one obtains the WKB approximation for the eigenvalues

$$\begin{aligned} E_{WKB} &= -\frac{mc^2 Z^2 \alpha^2}{2}\left[\sqrt{l_d(l_d+1)} + N + \frac{1}{2}\right]^{-2} \\ l_d &= l + \frac{d-3}{2} \qquad N, l = 0, 1, 2, \cdots \end{aligned} \tag{11.42}$$

Comparing to the exact results of chapter 6

$$E_{exact} = -\frac{mc^2 Z^2 \alpha^2}{2}\left[(l + N + 1) + \frac{d-3}{2}\right]^{-2} \tag{11.43}$$

one sees that the WKB approximation is good for large quantum numbers N, l, d, but not for small ones, as expected. However, note that there would be exact agreement if one replaces $\sqrt{l_d(l_d+1)} \to \sqrt{l_d(l_d+1) + \frac{1}{4}} = l_d + 1/2$.

Repeating the same approach for the harmonic oscillator in d-dimensions one finds (see problem 1)

$$\begin{aligned} E_{WKB} &= \hbar\omega\left(\sqrt{l_d(l_d+1)} + 2N + 1\right) \\ E_{exact} &= \hbar\omega\left(l + 2N + \frac{d}{2}\right) \end{aligned} \tag{11.44}$$

Again, the result is good for large quantum numbers N, l, d, but not for small ones, as expected. But there would be exact agreement if one replaces $\sqrt{l_d(l_d+1)} \to \sqrt{l_d(l_d+1) + \frac{1}{4}} = l_d + 1/2$.

11.3.2 Tunneling

Consider α-decay, the process of a nucleus of charge $(Z+2)$ splitting up into an α particle of charge 2 and a smaller nucleus of charge Z. At small distances within

the nucleus the α particle is attracted back to the center by the nuclear force that is larger than the electrostatic repulsion between the charge Z fragment and the α. But its kinetic energy overcomes the attraction and it escapes. Once it is outside of the range of the nuclear force $R \sim 10^{-12} cm$, there remains only the electrostatic repulsion and the angular momentum barrier. The effective potential energy

$$V_{eff} = V_{nuclear}(r) + \frac{2Ze^2}{r} + \frac{\hbar^2 l(l+1)}{r^2} \tag{11.45}$$

has the shape given in Fig.(11.5) for zero angular momentum.

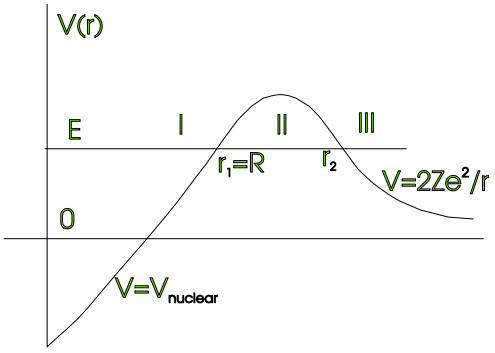


Fig. (11.5) - Barrier for $\alpha-$decay.

At large distances in region III the $1/r$ Coulomb repulsion

$$V_{eff}(r) \underset{r \gg R}{\rightarrow} \frac{2Ze^2}{r} \tag{11.46}$$

dominates all other terms . What is the lifetime of the nucleus τ =? It is inversely proportional to the decay probability P, which is given by the ratio of the probabilities for finding the α-particle outside versus inside the nucleus.

$$\tau = \tau_0 / P, \quad P = \frac{|\psi(out)|^2}{|\psi(in)|^2} \tag{11.47}$$

Since we want to avoid the details of the nuclear potential we will try to compute P up to an overall constant and absorb some of the details of the nuclear potential in a redefinition of τ_0. The asymptotic form of the wavefunction outside of the nucleus must be an outgoing wave of the WKB form

$$\psi_{WKB}(out) = \frac{A_{III}}{\sqrt{p(r)}} \exp \frac{i}{\hbar} \left[\int_{r_2}^{r} dr' \, p(r') - \frac{\pi}{4} + \phi \right]. \tag{11.48}$$

The WKB connection formulas (for $n = 1$) gives the asymptotic expressions for region II when r is far away from both r_1 and r_2

$$\psi^{II}_{WKB}(r) = A_{III} \; e^{i(\phi - \pi/2)} \; \frac{\exp\left(\frac{1}{\hbar} \int_r^{r_2} dr' \, \tilde{p}(r')\right)}{\sqrt{\tilde{p}(r)}} \tag{11.49}$$

Inside the nucleus $\psi(in)$ must vanish at the origin $r=0$. Feeding this information to region II by using (11.25) gives (i.e. first write it in terms of $J_{\pm 1/3}$ and then analytically continue through r_1 to region II)

$$\psi^{II}(r) = A_I \sqrt{\frac{\int_{r_1}^{r} \tilde{p}(r')}{\tilde{p}(r)}} \left[I_{-1/3}\left(\int_{r_1}^{r} \frac{\tilde{p}(r')}{\hbar} \right) - I_{1/3}\left(\int_{r_1}^{r} \frac{\tilde{p}(r')}{\hbar} \right) \right] \tag{11.50}$$

This form is valid near $r=r_1$ and far away from it, but is not valid near r_2. The asymptotic form of this expression away from r_1 is

$$\psi^{II}(r) \underset{r \gg r_1}{\rightarrow} \psi^{II}_{WKB}(r) = \frac{A_I \, \exp\left(-\frac{1}{\hbar} \int_{r_1}^{r} dr' \, \tilde{p}(r') \right)}{\sqrt{\tilde{p}(r)}}. \tag{11.51}$$

This is connected to the asymptotic WKB form inside the nucleus

$$\psi_{WKB}(in) = \frac{A_I}{\sqrt{p(r)}} \sin\left[\frac{1}{\hbar} \int_0^r dr' \, p(r') \right]. \tag{11.52}$$

One sees that A_I is a measure of the maximum amplitude inside the nucleus and A_{III} is a similar measure for outside the nucleus. Therefore the decay probability is proportional to the ratio $\left| A_{III}/A_I \right|^2$. Matching the two forms (11.49,11.51) of $\psi^{II}_{WKB}(r)$ at any r in the region $r_1 \ll r \ll r_2$ one finds the ratio A_{III}/A_I

$$\frac{A_{III}}{A_I} = e^{-i(\phi - \pi/2)} \exp\left(-\frac{1}{\hbar} \int_{r_1}^{r_2} dr' \, \tilde{p}(r') \right). \tag{11.53}$$

Squaring it gives the decay probability up to an overall factor C that may be considered a constant for slowly varying potentials

$$P = C \exp\left(-\frac{2}{\hbar} \int_{r_1(E)}^{r_2(E)} dr' \, \tilde{p}(r') \right). \tag{11.54}$$

The overall normalization is fixed to $C=1$ by requiring $P=1$ at energies for which $r_1(E) = r_2(E)$. Here $r_1(E), r_2(E)$ are the turning points that solve the equation $V_{eff}(r) = E$, as shown in Fig.(11.5). A discussion of the details of the nuclear potential is avoided by estimating that r_1 is of the order of the nuclear radius

$$r_1 \approx R \approx 10^{-12} cm. \tag{11.55}$$

Similarly, $r_2(E)$ is estimated by using the asymptotic form of the effective potential in (11.46) which is purely Coulomb, $r_2 \sim \frac{2Ze^2}{E}$. Furthermore, the energy E is equal to the kinetic energy of the free α particle outside of the nucleus $E \approx \frac{1}{2} m_\alpha v^2$, where the velocity v is measured by an observer. Here one must use the reduced mass that appears in the radial equation

$$m_\alpha = \frac{M_\alpha M_Z}{M_\alpha + M_Z} \tag{11.56}$$

where M_α, M_Z are the actual masses of the α and the remaining nuclear fragment. Thus, both r_1 and r_2 are pinned down approximately by experimental quantities

$$r_1 \approx R, \quad r_2 \approx \frac{4Ze^2}{m_\alpha v^2}. \tag{11.57}$$

The decay probability is then

$$P = \exp\left(-\frac{2}{\hbar}\int_R^{\frac{4Ze^2}{m_\alpha v^2}} dr\ \sqrt{2m_\alpha V_{eff}(r) - m_\alpha^2 v^2}\right). \tag{11.58}$$

In the region of integration one can ignore the short range nuclear potential. Furthermore, specializing to the zero angular momentum state $l = 0$, the integral can be easily performed, leading to the decay rate

$$\begin{aligned} \tau_{(Z+2)\to Z+\alpha} &= \tau_0 \exp\left[\frac{2}{\hbar}\int_R^{\frac{4Ze^2}{m_\alpha v^2}} dr\sqrt{\frac{4Ze^2 m_\alpha}{r} - m_\alpha^2 v^2}\right] \\ &= \tau_0\ \exp\left[\frac{4Ze^2}{\hbar v}\left\{\pi - \gamma - \cos\gamma\right\}\right] \end{aligned} \tag{11.59}$$

where $\gamma < \pi/2$ is defined by

$$\sin\gamma = \sqrt{\frac{m_\alpha v^2 R}{4Ze^2}} = \sqrt{\frac{R}{r_2}} < 1, \tag{11.60}$$

and $4Ze^2/\hbar v = m_\alpha v r_2/\hbar$ is an angular momentum just outside the nucleus, in dimensionless units.

The decay rate computed above is a function of Z and v. It can be plotted against these quantities and compared to experiment. The characteristic lifetime is of order $\tau_0 \sim 10^{-21}$ sec. The exponential can be large. The fit is a reasonably good one over a large range of lifetimes.

11.4 Problems

1. Compute the energy eigenvalues and eigenstates for the harmonic oscillator in d-dimensions in the WKB approximation.

2. Compute the energy levels for the 1-dimensional potential

$$V(x) = \left(\frac{|x|}{a} - 1\right)\ V_0\,\theta\left(a - |x|\right). \tag{11.61}$$

   in the WKB approximation. If $\frac{2mV_0a^2}{\hbar^2\pi^2} = 9$, how many bound states are there?

3. Compute the energy levels for $H = \mathbf{p}^2/2m + \gamma r$ in the WKB approximation.

4. Derive the connection formulas (11.27,11.28) by using the properties of Bessel functions.

5. Consider examples of one-dimensional potentials, of the type discussed in the text and the problems in Chapter 4, for which the transmission coefficient can be computed exactly. Then apply the WKB method to the same problem to compute the probability of tunneling. Compare the exact and WKB results.

Chapter 12

PERTURBATION THEORY

Perturbation theory is a systematic expansion of the physical quantities of a system in terms of a small parameter in the Hamiltonian. It provides accurate results to the extent that the expansion parameter is small. Typically it arises when the Hamiltonian has the form

$$H = H_0 + H', \tag{12.1}$$

where H' is small as compared to H_0, within certain validity criteria that will be discussed below. The underlying assumption is that the physical system described by H_0 is exactly solved, and that the system described by $H = H_0 + H'$ is difficult to solve exactly. Then perturbation theory provides a series expansion of all physical quantities (energies, states, matrix elements) in powers of H'. In principle the expansion provides answers to any desired accuracy, but in practice the approach is useful if one does not need to go beyond a few terms, usually first order or second order perturbation theory.

In this chapter only time independent perturbation theory is discussed. The time dependent cases will be covered in the next chapter. To begin, the exact *formal* solution of the eigenstates and eigenvalues of the total system described by $H = H_0 + H'$ are given. This closed form may be computed exactly, if possible for certain cases. Otherwise, it has a form that is convenient for an expansion in powers of H', so that perturbation theory appears as an ordinary series expansion.

12.1 Diagonalization of H

One assumes that the quantum problem for H_0 is already solved. That is, H_0 has been diagonalized, and if there are any additional operators $\{A_0^i,\ i = 1, 2, \cdots\}$ that commute with H_0, the complete set has been found and simultaneously

diagonalized together with H_0. These eigenvalues and eigenstates define the complete Hilbert space that we will refer to as the "zeroth order" or "perturbative" Hilbert space. The eigenvalues of the operators may be used to label the zeroth order states as $|E_n^0, a_{0m}^i\rangle$:

$$\begin{aligned} H_0|E_n^0, a_{0m}^i\rangle &= E_n^0|E_n^0, a_{0m}^i\rangle \\ A_0^i|E_n^0, a_{0m}^i\rangle &= a_{0m}^i|E_n^0, a_{0m}^i\rangle \end{aligned} \tag{12.2}$$

Similarly, in principle, there is an "exact" Hilbert space which is defined by the eigenvalues of the total H and a complete set of operators $\{A^i,\ i = 1, 2, \cdots\}$ that commute with it. The states of the exact Hilbert space $|E_n, a_m^i\rangle$ satisfy

$$\begin{aligned} H|E_n, a_m^i\rangle &= E_n|E_n, a_m^i\rangle \\ A^i|E_n, a_m^i\rangle &= a_m^i|E_n, a_m^i\rangle \end{aligned} \tag{12.3}$$

In general the set $\{A^i,\ i = 1, 2, \cdots\}$ is different than the set $\{A_0^i,\ i = 1, 2, \cdots\}$, although, depending on the problem, some of the operators may overlap. Having mentioned the existence of the operators A_0^i and A^i, they will be suppressed from the discussion from now on, and we will concentrate only on the eigenvalues E_n^0, E_n as if they are the only labels of the states $|E_n^0\rangle, |E_n\rangle$. Both sets are complete and orthonormal

$$\begin{array}{ll} \sum_n |E_n^0\rangle\langle E_n^0| = \mathbf{1}, & \langle E_n^0|E_m^0\rangle = \delta_{nm} \\ \sum_n |E_n\rangle\langle E_n| = \mathbf{1}, & \langle E_n|E_m\rangle = \delta_{nm} \end{array} \tag{12.4}$$

In the completeness relation the sum would be replaced by an integral for the eigenvalues that are continuous. Similarly, in the orthogonality relations the Kronecker delta δ_{nm} would be replaced by the Dirac delta function for continuous eigenvalues. Keeping this in mind, for the sake of simplicity, we will continue to use orthogonality and completeness relations as if all eigenvalues are discrete. It is easy to make the appropriate modification when they are not discrete.

One may expand any quantum state $|\psi\rangle$ in terms of either set. In particular the exact eigenstates may be expressed as a linear combination of the zeroth order states by multiplying with the identity operator

$$\begin{aligned} |E_n\rangle &= \sum_{n'} |E_{n'}^0\rangle\langle E_{n'}^0|E_n\rangle \\ &= \sum_{n'} |E_{n'}^0\rangle U_{n'n}^\dagger \\ U_{n'n}^\dagger &\equiv \langle E_{n'}^0|E_n\rangle \end{aligned} \tag{12.5}$$

The matrix U is unitary as a consequence of the completeness and orthogonality conditions

$$\begin{aligned} \left(UU^\dagger\right)_{nn'} &= \textstyle\sum_{n''} U_{nn''}U_{n''n'}^\dagger \\ &= \textstyle\sum_{n''} \langle E_n|E_{n''}^0\rangle\langle E_{n''}^0|E_{n'}\rangle \\ &= \langle E_n|E_{n'}\rangle = \delta_{nn'}. \end{aligned} \tag{12.6}$$

One may construct a unitary operator $\hat{U}$ that maps the perturbative Hilbert space to the exact Hilbert space and establishes a one-to-one correspondance between the states

$$\hat{U} = \sum_k |E_k^0\rangle\langle E_k|, \quad \hat{U}^\dagger = \sum_k |E_k\rangle\langle E_k^0|, \tag{12.7}$$

$$|E_n^0\rangle = U|E_n\rangle, \quad |E_n\rangle = U^\dagger|E_n^0\rangle, \tag{12.8}$$

where the second line follows from orthogonality. The matrix elements of $\hat{U}^\dagger$ evaluated in the perturbative Hilbert space coincide with $U^\dagger_{n'n}$ as seen below

$$\langle E_{n'}^0|\hat{U}^\dagger|E_n^0\rangle = \sum_k \langle E_{n'}^0|E_k\rangle\langle E_k^0|E_n^0\rangle \tag{12.9}$$

$$= \langle E_{n'}^0|E_n\rangle = U^\dagger_{n'n} \tag{12.10}$$

Here we have used orthogonality and the definition of $U^\dagger_{n'n}$ given in (12.5).

The aim is to find the eigenstates $|E_n\rangle$, eigenvalues E_n, and the operator $\hat{U}$ as a function of H'. In the zeroth order Hilbert space $|E_n^0\rangle$ it is assumed that the matrix element of any operator may be computed. In particular one assumes that the following computations have already been performed

$$\langle E_m^0|\ H\ |E_n^0\rangle \equiv H_{mn} = E_n^0\,\delta_{mn} + H'_{mn}. \tag{12.11}$$

Then the Hamiltonian operator may be expressed in either basis by using the different forms of the identity operator

$$\begin{aligned} H &= \mathbf{1}H\mathbf{1} \\ &= \sum_{nn'} |E_n^0\rangle\langle E_n^0|\ H\ |E_{n'}^0\rangle\langle E_{n'}^0| \\ &= \sum_{nn'} |E_n^0\rangle \left[E_n^0\,\delta_{mn} + H'_{mn}\right] \langle E_{n'}^0| \\ &= \sum_{mm'} |E_m\rangle\langle E_m|\ H\ |E_{m'}\rangle\langle E_{m'}| \\ &= \sum_{mm'} |E_m\rangle E_m \delta_{mm'} \langle E_{m'}| \end{aligned} \tag{12.12}$$

By using the relation $|E_n^0\rangle = \sum_m |E_m\rangle U_{mn}$ between the two complete Hilbert spaces the third line may be rewritten in a form that is comparable to the last line

$$H = \sum_{mm'} |E_m\rangle \left[\sum_{nn'} U_{mn}\left(E_0^n\delta_{nn'} + H'_{nn'}\right) U^\dagger_{n'm'}\right] \langle E_{m'}| \tag{12.13}$$

This shows that the matrix U which provides the map between the two Hilbert spaces must be identified with the matrix that performs the diagonalization of the Hamiltonian in matrix form

$$U^\dagger_{mn}\left(E_0^n\delta_{nn'} + H'_{nn'}\right) U_{n'm'} = E_m\delta_{mm'}. \tag{12.14}$$

So, in principle this is a method for computing both the eigenvalues and eigenstates.

The aim of this chapter is to develop an approximation technique for computing $E_n, |E_n\rangle, U_{mn}$, but before doing so we will first obtain some exact relations that follow from the definitions given above. Furthermore, there are sufficiently simple cases for which the diagonalization procedure can be carried out exactly. It is useful to discuss these exact expressions before considering the perturbative expansion.

12.2 Two level Hamiltonian

It is useful to solve the problem exactly in a simple case and use the result as a guide for the perturbative expansion. The 2-level system is an important problem worth discussing exactly, since it has many practical physical applications (lasers, magnetic resonance, $K^0 - \bar{K}^0$ system in particle physics, etc.). This case also serves as a simplified laboratory to illustrate the methods of approximation. The Hamiltonian has the form

$$H = H_0 + H' = \begin{pmatrix} E_1^0 & 0 \\ 0 & E_2^0 \end{pmatrix} + \begin{pmatrix} k_1 & h \\ h^* & k_2 \end{pmatrix} \tag{12.15}$$

where we have used the basis

$$|E_1^0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |E_2^0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.16}$$

We seek a transformation U that diagonalizes the matrix and identify the eigenstates. This problem was solved in Chapter 3, section 3.8, where the Hamiltonian was written as

$$H = \begin{pmatrix} x+y & y\,\tan\theta\, e^{i\phi} \\ y\,\tan\theta\, e^{-i\phi} & x-y \end{pmatrix}, \tag{12.17}$$

with the definitions

$$\begin{aligned} x &= \frac{1}{2}(E_1^0 + E_2^0) + \frac{1}{2}(k_1 + k_2) \\ y &= \frac{1}{2}(E_1^0 - E_2^0) + \frac{1}{2}(k_1 - k_2) \\ \tan\theta &= \frac{|h|}{y}, \quad e^{i2\phi} = \frac{h}{h^*}. \end{aligned} \tag{12.18}$$

Then, the exact energy eigenvalues and eigenfunctions are conveniently given by

$$\begin{aligned} E_1 &= x + \frac{y}{\cos\theta} \qquad , \qquad E_2 = x - \frac{y}{\cos\theta} \\ |E_1\rangle &= \begin{pmatrix} \cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2}\, e^{-i\phi} \end{pmatrix}, \quad |E_2\rangle = \begin{pmatrix} \sin\frac{\theta}{2}\, e^{i\phi} \\ \cos\frac{\theta}{2} \end{pmatrix} \\ U &= \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2}\, e^{i\phi} \\ -\sin\frac{\theta}{2}\, e^{-i\phi} & \cos\frac{\theta}{2} \end{pmatrix} \end{aligned} \tag{12.19}$$

The reader can verify directly that these expressions satisfy $H|E_i\rangle = E_i|E_i\rangle$, and $UHU^\dagger = E$.

Perturbation theory corresponds to expanding in the small parameters k_1, k_2, h. For the purpose of comparing to perturbation theory it will be useful to rewrite this result in terms of the original parameters

$$E_1 = \frac{1}{2}\left(E_1^0 + E_2^0 + k_1 + k_2\right) + \frac{1}{2}\left[\left(E_1^0 - E_2^0 + k_1 - k_2\right)^2 + 4\left|h\right|^2\right]^{1/2} \quad (12.20)$$

$$E_2 = \frac{1}{2}\left(E_1^0 + E_2^0 + k_1 + k_2\right) - \frac{1}{2}\left[\left(E_1^0 - E_2^0 + k_1 - k_2\right)^2 + 4\left|h\right|^2\right]^{1/2} \quad (12.21)$$

$$\cos\frac{\theta}{2} = \frac{1}{\sqrt{2}}\left[1 + \left(1 + \frac{4\left|h\right|^2}{\left(E_1^0 - E_2^0 + k_1 - k_2\right)^2}\right)^{-1/2}\right]^{1/2} \quad (12.22)$$

$$\sin\frac{\theta}{2} = \frac{\sqrt{2}\left|h\right|\left[1 + \frac{4|h|^2}{\left(E_1^0 - E_2^0 + k_1 - k_2\right)^2} + \left(1 + \frac{4|h|^2}{\left(E_1^0 - E_2^0 + k_1 - k_2\right)^2}\right)^{1/2}\right]^{-1/2}}{\left(E_1^0 - E_2^0 + k_1 - k_2\right)} \quad (12.23)$$

A consistent expansion is obtained by rescaling the parameters by an overall factor $\lambda k_1, \lambda k_2, \lambda h$, and expanding in powers of λ. The result is

$$E_1 = E_1^0 + \lambda k_1 + \lambda^2 \frac{\left|h\right|^2}{E_1^0 - E_2^0} + \lambda^3 \frac{\left(-k_1 + k_2\right)\left|h\right|^2}{\left(E_1^0 - E_2^0\right)^2} + \lambda^4 \frac{\left|h\right|^2\left(\left(-k_1 + k_2\right)^2 - \left|h\right|^2\right)}{\left(E_1^0 - E_2^0\right)^3} + \cdots$$

$$E_2 = E_2^0 + \lambda k_2 - \lambda^2 \frac{\left|h\right|^2}{E_1^0 - E_2^0} - \lambda^3 \frac{\left(-k_1 + k_2\right)\left|h\right|^2}{\left(E_1^0 - E_2^0\right)^2} - \lambda^4 \frac{\left|h\right|^2\left(\left(-k_1 + k_2\right)^2 - \left|h\right|^2\right)}{\left(E_1^0 - E_2^0\right)^3} - \cdots$$

$$\sin\frac{\theta}{2} = \lambda \frac{\left|h\right|}{E_1^0 - E_2^0} - \lambda^2 \frac{\left(k_1 - k_2\right)\left|h\right|}{\left(E_1^0 - E_2^0\right)^2} - \lambda^3 \left|h\right| \frac{3\left|h\right|^2 + 2\left(k_1^2 - k_2^2\right)}{2\left(E_1^0 - E_2^0\right)^3} + \cdots$$

$$\cos\frac{\theta}{2} = 1 - \lambda^2 \frac{\left|h\right|^2}{2\left(E_1^0 - E_2^0\right)^2} + \lambda^3 \frac{\left(k_1 - k_2\right)\left|h\right|^2}{\left(E_1^0 - E_2^0\right)^3} + \cdots \quad (12.24)$$

At the end of this computation we set $\lambda = 1$. This expansion should be compared to the computation using perturbation theory in the following section.

12.3 Formal exact solution

We will now construct the exact solution in the general case. This will be a formal solution in the sense that more work will be needed for a complete answer. However, the form of the formal solution is useful because it will lend itself to a systematic perturbative expansion in powers of H'.

Consider the exact eigenvalue problem

$$(H - E_n)|E_n\rangle = 0, \tag{12.25}$$

and rewrite it by using $H = H_0 + H'$ and $E_n = E_n^0 + \Delta_n$, where Δ_n is the energy difference between the exact and the perturbative state. The equation above becomes

$$(H_0 - E_n^0)|E_n\rangle = (\Delta_n - H')\,|E_n\rangle. \tag{12.26}$$

If one takes a dot product with the bra $\langle E_n^0|$ the left hand side vanishes. This shows that the state $(\Delta_n - H')\,|E_n\rangle$ is orthogonal to $|E_n^0\rangle$

$$\langle E_n^0|\,(\Delta_n - H')\,|E_n\rangle = 0. \tag{12.27}$$

From this equation we find a formula for the energy difference Δ_n

$$(E_n - E_n^0) = \frac{\langle E_n^0|H'|E_n\rangle}{\langle E_n^0|E_n\rangle} = \Delta_n \tag{12.28}$$

Since the right hand side of (12.26) does not contain the state $|E_n^0\rangle$ nothing changes if it is multiplied by the projection operator ($Q_n^2 = Q_n$)

$$Q_n = 1 - |E_n^0\rangle\langle E_n^0| = \sum_{k\neq n} |E_k^0\rangle\langle E_k^0|. \tag{12.29}$$

Next eq.(12.26) is rewritten by multiplying it with the inverse $(H_0 - E_n^0)^{-1}$ on both sides

$$|E_n\rangle = (H_0 - E_n^0)^{-1} Q_n\,(\Delta_n - H')\,|E_n\rangle + Z_n^{-1/2}|E_n^0\rangle. \tag{12.30}$$

The inverse $(H_0 - E_n^0)^{-1}$ is well defined on the right hand side since the state $|E_n^0\rangle$ is absent in the first term, and this is emphasized by the insertion of the projector Q_n. Hence $(H_0 - E_n^0)^{-1}Q_n$ must be understood as the inverse of $H_0 - E_n^0$ in the subspace of the Hilbert space excluding the state $|E_n^0\rangle$. On the right hand side of (12.30) one is free to add an arbitrary term proportional to $|E_n^0\rangle$ with proportionality constant $Z_n^{-1/2}$, because if one applies $(H_0 - E_n^0)$ on both sides of (12.30) the original equation (12.26) is recovered for any $Z^{-1/2}$. Thus the two equations (12.26) and (12.30) are equivalent.

The exact state $|E_n\rangle$ can now be obtained in terms of the perturbative state by solving (12.30) formally

$$|E_n\rangle = Z_n^{-1/2}\left[1 - (H_0 - E_n^0)^{-1} Q_n\,(\Delta_n - H')\right]^{-1}|E_n^0\rangle. \tag{12.31}$$

The meaning of $Z^{-1/2}$ is now understood as a normalization constant. It can be computed in two ways. The first is by dotting eq.(12.30) with $\langle E_n^0|$ on both sides and using the norm $\langle E_n^0|E_n^0\rangle = 1$

$$Z_n^{-1/2} = \langle E_n^0|E_n\rangle. \tag{12.32}$$

The second is by imposing $\langle E_n|E_n\rangle = 1$ for the $|E_n\rangle$ given in Eq.(12.31). We can now construct the operator $\hat{U}$ through its definition (12.7)

$$\hat{U}^\dagger = \sum_n Z_n^{-1/2} \left[1 - (H_0 - E_n^0)^{-1} Q_n \left(\Delta_n - H'\right)\right]^{-1} |E_n^0\rangle\langle E_n^0| \,. \tag{12.33}$$

Its matrix elements are given by

$$U^\dagger_{mn} = \langle E_m^0| \left[1 - (H_0 - E_n^0)^{-1} Q_n \left(\Delta_n - H'\right)\right]^{-1} |E_n^0\rangle Z_n^{-1/2}. \tag{12.34}$$

This exact expression for $U^\dagger_{mn}$ involves only matrix elements in the perturbative Hilbert space $|E_n^0\rangle$. Note that since $\langle E_n^0|Q_n = 0$, the diagonal entries $U^\dagger_{mn}$ are given by the normalization factor $U^\dagger_{nn} = Z_n^{-1/2}$. The exact Hilbert space $|E_n\rangle$ is obtained either in the form of Eq.(12.30) or by inserting the $U^\dagger_{mn}$ above the sum of Eq.(12.5)

$$|E_n\rangle = \sum_m |E_m^0\rangle U^\dagger_{mn}. \tag{12.35}$$

In summary, the equations (12.28,12.31,**??**,12.33) provide exact expressions for $E_n, |E_n\rangle, \hat{U}^\dagger$ in terms of the original Hilbert space $|E_n^0\rangle$ and the matrix elements of the operators H_0, H'. This is a formal solution because Δ_n appears non-linearly in $|E_n\rangle$ and Z_n, and satisfies the non-linear equation (12.28), from which it remains to be solved. This non-linear equation for Δ_n is equivalent to the secular equation for the eigenvalues

$$\det\left(H - E_n^0 - \Delta_n\right) = 0. \tag{12.36}$$

If the secular equation is solved, the above formulas provide the eigenstates $|E_n\rangle$ and the unitary operator U. When it cannot be solved exactly we need the perturbative expansion discussed in the next section.

12.3.1 2-level problem revisited

In order to appreciate the general exact formulas of this section it is useful to apply them to the exactly solvable 2-level problem of the previous section, and rederive the same exact results using the general formalism given above. So, if we take $E_1^0 + k_1 = x + y$, $E_2^0 + k_2 = x - y$, with

$$|E_1^0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |E_2^0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{12.37}$$

we have the setup

$$H_0 = \begin{pmatrix} x+y & 0 \\ 0 & x-y \end{pmatrix}, \quad H' = \begin{pmatrix} 0 & y\tan\theta\, e^{i\phi} \\ y\tan\theta\, e^{-i\phi} & 0 \end{pmatrix}. \tag{12.38}$$

The projection operators are

$$Q_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

We then find

$$(H_0 - E_1^0)^{-1} Q_1 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{-2y} \end{pmatrix}, \quad (H_0 - E_2^0)^{-1} Q_2 = \begin{pmatrix} \frac{1}{2y} & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\begin{aligned} &\left[1 - (H_0 - E_1^0)^{-1} Q_1 \left(\Delta_1 - H'\right)\right]^{-1} \\ &= \left[1 - \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{-2y} \end{pmatrix} \begin{pmatrix} \Delta_1 & y \tan\theta \, e^{i\phi} \\ y \tan\theta \, e^{-i\phi} & \Delta_1 \end{pmatrix}\right]^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} \tan\theta \, e^{-i\phi} & 1 + \frac{1}{2y}\Delta_1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ \frac{y \tan\theta}{2y+\Delta_1} e^{-i\phi} & \frac{2y}{2y+\Delta_1} \end{pmatrix}. \end{aligned}$$

Therefore, the eigenstate follows from (12.31)

$$|E_1\rangle = Z_1^{-1/2} \begin{pmatrix} 1 & 0 \\ \frac{y \tan\theta}{2y+\Delta_1} e^{-i\phi} & 2\frac{y}{2y+\Delta_1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = Z_1^{-1/2} \begin{pmatrix} 1 \\ \frac{y \tan\theta \, e^{-i\phi}}{2y+\Delta_1} \end{pmatrix}.$$

The normalization follows from $\langle E_1 | E_1 \rangle = 1$,

$$Z_1 = 1 + \left(\frac{y \tan\theta}{2y + \Delta_1}\right)^2,$$

and the equation for the energy difference Δ_1 from (12.28)

$$\begin{aligned} \Delta_1 &= Z_1^{1/2} \langle E_1^0 | H' | E_1 \rangle \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & y \tan\theta \, e^{i\phi} \\ y \tan\theta \, e^{-i\phi} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{y \tan\theta \, e^{-i\phi}}{2y+\Delta_1} \end{pmatrix} \\ &= \frac{y^2 \tan^2\theta}{\Delta_1 + 2y}. \end{aligned} \tag{12.39}$$

Similarly, for the second state one gets

$$\left[1 - (H_0 - E_2^0)^{-1} Q_2 \left(\Delta_2 - H'\right)\right]^{-1} = \begin{pmatrix} -\frac{2y}{\Delta_2 - 2y} & \frac{y \tan\theta}{\Delta_2 - 2y} e^{i\phi} \\ 0 & 1 \end{pmatrix},$$

and

$$|E_2\rangle = Z_2^{-1/2} \begin{pmatrix} \frac{y \tan\theta \, e^{i\phi}}{\Delta_2 - 2y} \\ 1 \end{pmatrix}, \; Z_2 = 1 + \frac{y^2 \tan^2\theta}{(2y - \Delta_2)^2}, \; \Delta_2 = \frac{y^2 \tan^2\theta}{\Delta_2 - 2y}. \tag{12.40}$$

The solutions of the quadratic equations for Δ_1, Δ_2 (12.39,12.40) are (the choice of roots is consistent with the limit $E_{1,2} \to E_{1,2}^0$ while $\theta \to 0$)

$$\Delta_1 = -y + \frac{y}{\cos\theta}, \quad \Delta_2 = y - \frac{y}{\cos\theta}.$$

This leads to

$$Z_1^{-1/2} = Z_2^{-1/2} = \cos\frac{\theta}{2}\,.$$

This exact result is in full agreement with the exact solution of the previous section, using $E_n = E_n^0 + \Delta_n$

$$E_1 = x + \frac{y}{\cos\theta}, \quad E_2 = x - \frac{y}{\cos\theta}. \tag{12.41}$$

Furthermore, the exact states $|E_n\rangle$ are also in agreement

$$|E_1\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2}\,e^{-i\phi} \end{pmatrix}, \quad |E_2\rangle = \begin{pmatrix} \sin\frac{\theta}{2}\,e^{i\phi} \\ \cos\frac{\theta}{2} \end{pmatrix}. \tag{12.42}$$

12.4 Perturbative expansion (non-degenerate)

If H' is small compared to H_0 we can approximate the exact solution by using an expansion in powers of H'. In order to do this systematically we multiply H' with the parameter λ, and expand in powers of λ. For this purpose it is useful to work with the unnormalized state $\overline{|E_n\rangle} \equiv Z_n^{1/2}|E_n\rangle$ and write the exact results of the previous section as follows

$$\overline{|E_n,\lambda\rangle} = \left[1 - (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\}\right]^{-1} |E_n^0\rangle \tag{12.43}$$

$$\Delta_n(\lambda) = \langle E_n^0|\lambda H' \left[1 - (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\}\right]^{-1} |E_n^0\rangle \tag{12.44}$$

$$Z_n(\lambda) = \overline{\langle E_n,\lambda||E_n,\lambda\rangle} \tag{12.45}$$

$$|E_n,\lambda\rangle = \overline{|E_n,\lambda\rangle} Z_n^{-1/2}(\lambda)\,. \tag{12.46}$$

The systematic expansion in λ is

$$\overline{|E_n,\lambda\rangle} = |E_n^0\rangle + \lambda\,|E_n\rangle^{(1)} + \lambda^2\,|E_n\rangle^{(2)} + \lambda^3\,|E_n\rangle^{(3)} + \cdots \tag{12.47}$$

$$\Delta_n(\lambda) = 0 + \lambda\Delta_n^{(1)} + \lambda^2\Delta_n^{(2)} + \lambda^3\Delta_n^{(3)} + \lambda^4\Delta_n^{(4)} + \cdots \tag{12.48}$$

$$Z_n(\lambda) = 1 + \lambda Z_n^{(1)} + \lambda^2 Z_n^{(2)} + \lambda^3 Z_n^{(3)} + \lambda^4 Z_n^{(4)} + \cdots \tag{12.49}$$

This is plugged into the equations above and powers of λ^n are compared on both sides. It is useful to do the expansion in two steps. In the first step one expands the inverse $[1 - \cdots]^{-1}$ formally as a power series

$$\left[1 - (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\}\right]^{-1} = 1 + (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\}$$
$$+ (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\} (H_0 - E_n^0)^{-1} Q_n \left\{\Delta_n(\lambda) - \lambda H'\right\} + \cdots$$

The terms that involve the $\Delta_n(\lambda)$ in bold can be dropped since the whole expression is applied on $|E_n^0\rangle$ and we obtain $\Delta_n(\lambda)\,Q_n|E_n^0\rangle = 0$ for each such term of the expansion. In the second step substitute $\Delta_n(\lambda) = 0 + \lambda\Delta_n^{(1)} + \lambda^2\Delta_n^{(2)} + \cdots$ for the remaining $\Delta_n(\lambda)$'s and re-expand in powers of λ. By collecting powers of λ consistently the various terms in the expansion are obtained (the details

are srtaightforward and left as an exercise for the student). Next use the projector Q_n in the form $Q_n = \sum_{k\neq n} |E_k^0\rangle\langle E_k^0|$ as given in (12.29) and evaluate the operator H_0 on its own eigenstates wherever it occurs. At the end of this procedure one sets $\lambda = 1$.

The results are as follows. The un-normalized state $\overline{|E_n\rangle}$ up to second power in H' is

$$\overline{|E_n\rangle} = |E_n^0\rangle + \sum_{k\neq n} |E_k^0\rangle \frac{H'_{kn}}{E_n^0 - E_k^0} - \sum_{k\neq n} |E_k^0\rangle \frac{H'_{kn}H'_{nn}}{\left(E_n^0 - E_k^0\right)^2} + \sum_{k\neq n, l\neq n} |E_k^0\rangle \frac{H'_{kl}H'_{ln}}{\left(E_n^0 - E_k^0\right)\left(E_n^0 - E_l^0\right)} + \cdots \tag{12.50}$$

The energy eigenvalue up to third power in H' is

$$E_n = E_n^0 + H'_{nn} + \sum_{k\neq n} \frac{H'_{nk}H'_{kn}}{E_n^0 - E_k^0} - H'_{nn} \sum_{k\neq n} \frac{H'_{nk}H'_{kn}}{\left(E_n^0 - E_k^0\right)^2} + \sum_{k\neq n,\, l\neq n} \frac{H'_{nk}H'_{kl}H'_{ln}}{\left(E_n^0 - E_k^0\right)\left(E_n^0 - E_k^0\right)} + \cdots \tag{12.51}$$

The normalization can be shown to satisfy $Z_n^{-1} = \partial E_n/\partial E_n^0$. We need the lowest terms in the expansion for $Z^{-1/2}$, which is obtained by taking the square root and then expanding in powers of H'

$$Z_n^{-1/2} = 1 - \frac{1}{2} \sum_{k\neq n} \frac{\left|H'_{kn}\right|^2}{\left(E_n^0 - E_k^0\right)^2} + \cdots \tag{12.52}$$

Finally, the normalized state is obtained by multiplying $Z_n^{-1/2}\overline{|E_n\rangle}$ and re-expanding in powers of H'

$$|E_n\rangle = |E_n^0\rangle \left(1 - \frac{1}{2} \sum_{k\neq n} \frac{\left|H'_{kn}\right|^2}{\left(E_n^0 - E_k^0\right)^2}\right) + \sum_{k\neq n} |E_k^0\rangle \frac{H'_{kn}}{E_n^0 - E_k^0} - \sum_{k\neq n} |E_k^0\rangle \frac{H'_{kn}H'_{nn}}{\left(E_n^0 - E_k^0\right)^2} + \sum_{k\neq n, l\neq n} |E_k^0\rangle \frac{H'_{kl}H'_{ln}}{\left(E_n^0 - E_k^0\right)\left(E_n^0 - E_l^0\right)} + \cdots \tag{12.53}$$

The $U_{mn}^\dagger = \langle E_m^0|E_n\rangle$ are then, up to second order given by

$$U_{mn}^\dagger = \delta_{mn} \left(1 - \frac{1}{2} \sum_{k\neq n} \frac{\left|H'_{kn}\right|^2}{\left(E_n^0 - E_k^0\right)^2}\right) + \tag{12.54}$$

$$+ (1 - \delta_{mn}) \left\{ \frac{H'_{mn}}{E_n^0 - E_m^0} - \frac{H'_{mn}H'_{nn}}{\left(E_n^0 - E_m^0\right)^2} + \sum_{l\neq n} \frac{H'_{ml}H'_{ln}}{\left(E_n^0 - E_m^0\right)\left(E_n^0 - E_l^0\right)} \right\} + \cdots$$

In these formulas $E_n^0 - E_k^0$ appear in the denominator. The expansion is valid *for level* n provided the energy differences $\left|E_n^0 - E_k^0\right|$ with all other energy levels $k \neq n$ are not small. Typically, the approximation is valid for a given level n if for that level

$$|H'_{kn}| \ll \left|E_n^0 - E_k^0\right| \tag{12.55}$$

is satisfied. If this is the case, this expansion is called non-degenerate perturbation theory for level n. If the contrary is true for some levels, either because of degenerate or nearly degenerate states for which $E_n^0 - E_k^0 = 0$ or small, one needs to reconsider the expansion carefully for those levels and apply the methods of degenerate perturbation theory that is discussed next.

12.5 Degenerate perturbation theory

In the previous discussion of perturbation theory we assumed that there was a one to one correspondence between the set of eigenvectors $\{|E_n^0\rangle\}$ and the corresponding set of eigenvalues $\{E_n^0\}$; that is, we have postulated no degeneracy in our energy eigenvalues. We also mentioned that if such degeneracy is present we cannot use the non-degenerate perturbation theory because of the appearance of zeroes in the denominator of the expressions of the E_n's or $|E_n\rangle$ in eqs.(12.51,12.53).

There would also be a problem even if no degeneracy is present, but the condition (12.57) is violated. Then the ratio

$$\frac{|H'_{nk}|}{|E_n^0 - E_k^0|} \tag{12.56}$$

is not small for some of the terms in the sums in eqs.(12.51,12.53), and the perturbative expansion is no longer a good approximation. It must be emphasized that these problems will depend on the specific eigenvalue E_n^0. The pertubative expansion continues to hold for the states labelled by E_n^0 that satisfy

$$|H'_{nk}| \ll \left|E_n^0 - E_k^0\right| \quad all \quad k \neq n \tag{12.57}$$

The problem occurs only for the states E_n which violate this condition.

The cure to the problem is to perform an exact (or almost exact) *partial diagonalization* of the matrix $H_{mn} = H_{mn}^0 + H'_{mn}$ in the blocks of degenerate or almost degenerate states. If necessary, one can relabel the states so that the degenerate or almost degenerate states appear in blocks in the form

$$H = \begin{pmatrix} E_1^0 & a & \# & \# & \# & \# \\ a^* & E_1^0 + \varepsilon_1 & \# & \# & \# & \# \\ \# & \# & E_2^0 & b & c & \# \\ \# & \# & b^* & E_2^0 + \varepsilon_2 & d & \# \\ \# & \# & c^* & d^{\#} & E_2^0 + \varepsilon_2' & \# \\ \# & \# & \# & \# & \# & \ddots \end{pmatrix}. \tag{12.58}$$

The entries indicated by # are assumed not to create a problem. In the first 2×2 block there would be a problem if $|a| \geq \varepsilon_1$, including the degenerate limit $\varepsilon_1 = 0$. Similarly in the 3×3 block there would be a problem if $|b| \geq \varepsilon_2$, or $|c| \geq \varepsilon_2'$ or $|d| \geq |\varepsilon_2 - \varepsilon_2'|$, including the degenerate limit $\varepsilon_2 = \varepsilon_2' = 0$. One must then diagonalize these blocks exactly or approximately "by hand", or without the perturbative expansion, as discussed in the previous sections. This defines a new basis $\left\{|\tilde{E}_n\rangle\right\}$ and new eigenvalues through the transformation U_{mn} that performs the partial diagonalization. U has a block diagonal form

$$U = \begin{pmatrix} U_{11}^{(1)} & U_{12}^{(1)} & 0 & 0 & 0 & 0 \\ U_{21}^{(1)} & U_{22}^{(1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & U_{11}^{(2)} & U_{12}^{(2)} & U_{13}^{(2)} & 0 \\ 0 & 0 & U_{21}^{(2)} & U_{22}^{(2)} & U_{23}^{(2)} & 0 \\ 0 & 0 & U_{31}^{(2)} & U_{32}^{(2)} & U_{33}^{(2)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots \end{pmatrix} \tag{12.59}$$

Each block is determined as the matrix that diagonalizes the corresponding block of the Hamiltonian, i.e.

$$U^\dagger H U = \tilde{H} = \begin{pmatrix} \tilde{E}_1 & \text{“0”} & \tilde{\#} & \tilde{\#} & \tilde{\#} & \tilde{\#} \\ \text{“0”} & \tilde{E}_1 + \tilde{\varepsilon}_1 & \tilde{\#} & \tilde{\#} & \tilde{\#} & \tilde{\#} \\ \tilde{\#} & \tilde{\#} & \tilde{E}_2 & \text{“0”} & \text{“0”} & \tilde{\#} \\ \tilde{\#} & \tilde{\#} & \text{“0”} & \tilde{E}_2 + \tilde{\varepsilon}_2 & \text{“0”} & \tilde{\#} \\ \tilde{\#} & \tilde{\#} & \text{“0”} & \text{“0”} & \tilde{E}_2 + \tilde{\varepsilon}_2' & \tilde{\#} \\ \tilde{\#} & \tilde{\#} & \tilde{\#} & \tilde{\#} & \tilde{\#} & \ddots \end{pmatrix} \tag{12.60}$$

The entries labelled by "0" are 0, unless U is chosen differently as discussed below. The off diagonal blocks are also transformed to new values labelled by $\tilde{\#}$, and the states are transformed to the new basis

$$|\tilde{E}_n\rangle = \sum_k |E_k^0\rangle U_{kn}^\dagger. \tag{12.61}$$

Now, the matrix $\tilde{H}$ corresponds to the original Hamiltonian computed in the new basis, and it may again be split into two parts

$$\tilde{H}_{mn} = \langle \tilde{E}_m | H | \tilde{E}_n \rangle = \tilde{H}_{mn}^0 + \tilde{H}_{mn}'. \tag{12.62}$$

As long as the condition

$$\left|\tilde{H}_{nk}'\right| \ll \left|\tilde{E}_n - \tilde{E}_k\right| \quad all \quad k \neq n \tag{12.63}$$

is satisfied in the new basis (for a given $\tilde{E}_n$), one may apply perturbation theory from this point on for the level $\tilde{E}_n$, as in the previous section.

If the partial diagonalization is actually accomplished, the entries labelled by "0" are 0. For small enough matrices, of course one can perform the diagonalization exactly. If this step is too difficult technically, one may resort to finding a U that gives small enough numbers for the entries labelled by "0", because then perturbation theory would also be applicable.

12.5.1 More on degeneracy

There is one more step to discuss if the eigenvalues in the same block are still exactly degenerate after the partial diagonalization. For example, suppose in the first block $\tilde{\varepsilon}_1 = 0$, while "0"=0. This degeneracy continues to create a problem as follows. For simplicity consider a 3-level problem whose Hamiltonian matrix has the form

$$H = \begin{pmatrix} A & 0 & C \\ 0 & A & D \\ C^* & D^* & B \end{pmatrix} \tag{12.64}$$

So that we identify

$$H_0 = \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & B \end{pmatrix}, \quad H' = \begin{pmatrix} 0 & 0 & C \\ 0 & 0 & D \\ C^* & D^* & 0 \end{pmatrix} \tag{12.65}$$

The basis is given by

$$\langle E_1^0| = (1\ 0\ 0)\,,\ \langle E_2^0| = (0\ 1\ 0)\,,\ \langle E_3^0| = (0\ 0\ 1) \tag{12.66}$$

and the initial energies are $E_1^0 = A$, $E_2^0 = A$, $E_3^0 = B$. There is a problem in the perturbation series due to the degeneracy $E_1^0 = E_2^0 = A$ if C or D is non-zero, as follows. The expansion of (12.51) up to second order is

$$E_1 = E_1^0 + \frac{H'_{12}H_{21}}{E_1^0 - E_2^0} + \frac{H'_{13}H'_{31}}{E_1^0 - E_3^0} + \cdots = A + \frac{0}{0} + \frac{|C|^2}{A - B} + \cdots \tag{12.67}$$

$$E_2 = E_2^0 + \frac{H'_{21}H'_{12}}{E_2^0 - E_1^0} + \frac{H'_{23}H'_{32}}{E_2^0 - E_3^0} + \cdots = A + \frac{0}{0} + \frac{|D|^2}{A - B} + \cdots \tag{12.68}$$

$$E_3 = E_3^0 + \frac{H'_{31}H'_{13}}{E_3^0 - E_1^0} + \frac{H'_{32}H'_{23}}{E_3^0 - E_2^0} + \cdots = B + \frac{|C|^2}{B - A} + \frac{|D|^2}{B - A} + \cdots \tag{12.69}$$

The $\frac{0}{0}$ is undetermined, and one must find its meaning. To do so, we apply a unitary transformation in such a way as to create a fully isolated 1×1 block with all off diagonals zeros

$$U^\dagger H U = \begin{pmatrix} \# & 0 & 0 \\ 0 & \# & \# \\ 0 & \#^* & \# \end{pmatrix} \tag{12.70}$$

This is easy to accomplish because any unitary transformation of the 2×2 block form

$$U = \begin{pmatrix} \# & \# & 0 \\ \# & \# & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{12.71}$$

leaves H_0 unchanged $U^\dagger H_0 U = H_0$ since the 2×2 block of H_0 is proportional to the matrix 1. Then U can be chosen to rotate the row vector $\begin{pmatrix} C^* & , D^* \end{pmatrix}$ and the corresponding column vector to point in only the second direction. Since a unitary transformation cannot change the length of the vector the result takes the form $\begin{pmatrix} 0 & , \sqrt{|C|^2 + |D|^2} \end{pmatrix}$. The transformation is easy to find

$$U^\dagger H U = \begin{pmatrix} A & 0 & 0 \\ 0 & A & \sqrt{|C|^2 + |D|^2} \\ 0 & \sqrt{|C|^2 + |D|^2} & B \end{pmatrix} \equiv \tilde{H} \tag{12.72}$$

$$U^\dagger = \begin{pmatrix} \frac{D}{\sqrt{|C|^2+|D|^2}} & \frac{-C}{\sqrt{|C|^2+|D|^2}} & 0 \\ \frac{C^*}{\sqrt{|C|^2+|D|^2}} & \frac{D^*}{\sqrt{|C|^2+|D|^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12.73}$$

The new basis is given by $\langle \tilde{E}_n | = \sum_k \langle E_k^0 | U_{kn}^\dagger$

$$\langle \tilde{E}_1 | = \begin{pmatrix} \frac{D}{\sqrt{|C|^2+|D|^2}}, & \frac{-C}{\sqrt{|C|^2+|D|^2}}, & 0 \end{pmatrix} \tag{12.74}$$

$$\langle \tilde{E}_2 | = \begin{pmatrix} \frac{C^*}{\sqrt{|C|^2+|D|^2}}, & \frac{D^*}{\sqrt{|C|^2+|D|^2}}, & 0 \end{pmatrix} \tag{12.75}$$

$$\langle \tilde{E}_3 | = \begin{pmatrix} 0, & 0, & 1 \end{pmatrix} \tag{12.76}$$

In this new basis we have

$$\tilde{H}_0 = \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & B \end{pmatrix}, \quad \tilde{H}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \sqrt{|C|^2 + |D|^2} \\ 0 & \sqrt{|C|^2 + |D|^2} & 0 \end{pmatrix} \tag{12.77}$$

Since the first 1×1 block is fully isolated in the form (12.70), the exact eigenvalue is $E_1 = A$, and the exact eigenstate is $\langle E_1 | = \langle \tilde{E}_1 |$ as given above. For the remaining 2×2 block perturbation theory may be applied without any problem to get the eigenvalues

$$E_1 = A \tag{12.78}$$

$$E_2 = \tilde{E}_2 + \frac{\tilde{H}'_{23}\tilde{H}'_{32}}{\tilde{E}_2 - \tilde{E}_3} + \cdots = A + \frac{|C|^2 + |D|^2}{A - B} + \cdots \tag{12.79}$$

$$E_2 = \tilde{E}_3 + \frac{\tilde{H}'_{32}\tilde{H}'_{23}}{\tilde{E}_3 - \tilde{E}_2} + \cdots = B + \frac{|C|^2 + |D|^2}{B - A} + \cdots \tag{12.80}$$

We see that the 0/0 problem of eqs.(12.67,12.68) is resolved, and the correct answer obtained.

We can verify that this is the correct answer by computing the exact eigenvalues for this problem. The secular equation for the exact diagonalization of the original Hamiltonian is

$$\det \begin{pmatrix} A-\lambda & 0 & C \\ 0 & A-\lambda & D \\ C^* & D^* & B-\lambda \end{pmatrix} = 0 \tag{12.81}$$

This gives the cubic equation

$$(A-\lambda)\left(\lambda^2 - \lambda\,(A+B) + AB - |C|^2 - |D|^2\right) = 0 \tag{12.82}$$

which has three solutions

$$\lambda = E_1 = A \tag{12.83}$$

$$\lambda = E_2 = \frac{1}{2}\,(A+B) - \frac{1}{2}\sqrt{(A+B)^2 + 4\left(|C|^2 + |D|^2\right)} \tag{12.84}$$

$$\lambda = E_3 = \frac{1}{2}\,(A+B) + \frac{1}{2}\sqrt{(A+B)^2 + 4\left(|C|^2 + |D|^2\right)} \tag{12.85}$$

Assuming $(A+B)^2 \gg \left(|C|^2 + |D|^2\right)$ one may expand the square root and compare the result to the perturbative computation above, and verify that it is the same expansion.

In summary, in the degenerate or nearly degenerate case, one must use other means to resolve the degeneracy so that for the remaining problem perturbation theory would be applicable. In some cases one may use the symmetries of the problem to find a "good" basis that would be appropriate for the perturbative computations.

12.6 Fine structure of Hydrogen

In a previous chapter, the Hydrogen-like atom with Hamiltonian $H_0 = \frac{p^2}{2\mu} - \frac{Ze^2}{r}$ was discussed, and its eigenvalues and eigenstates were completely determined. In this section we will consider two corrections to H_0, namely the relativistic correction due to the fast motion of the electron, and the spin-orbit coupling correction due to the spin of the electron. These two corrections are of the same order of magnitude and therefore they must be treated simultaneously. Together they correspond to the "fine structure" of the Hydrogen atom for $Z = 1$.

12.6.1 Relativistic correction

The energy of a relativistic free particle is given by $E = \sqrt{c^2\mathbf{p}^2 + m^2c^4}$. The kinetic energy K is obtained after subtracting the rest energy $E_0 = mc^2$, thus

$$K = \sqrt{c^2\mathbf{p}^2 + m^2c^4} - mc^2. \tag{12.86}$$

The velocity of the particle is given by

$$\frac{\mathbf{v}}{c} = \frac{\mathbf{p}c}{E} = \frac{\mathbf{p}c}{\sqrt{c^2\mathbf{p}^2 + m^2c^4}}, \tag{12.87}$$

and it cannot exceed the velocity of light as seen from the formula. Small velocities occur when the mass term is much larger than the momentum term. In this case one may use the expansion of the square root $\sqrt{1+x^2} = 1 + \frac{1}{2}x^2 - \frac{1}{8}x^4 + \cdots$, with $x^2 = \left(c^2\mathbf{p}^2\right) / \left(m^2c^4\right)$, to write the kinetic energy in the form

$$K = mc^2\left(\sqrt{1+x^2} - 1\right) = \frac{\mathbf{p}^2}{2m} - \frac{1}{8}\frac{\left(\mathbf{p}^2\right)^2}{m^3c^2} + \cdots \tag{12.88}$$

The first term is the non-relativistic kinetic energy already included in H_0, and the second term is the relativistic correction which takes the form

$$H_{rel} = -\frac{1}{2mc^2}\left(\frac{\mathbf{p}^2}{2m}\right)^2. \tag{12.89}$$

For later convenience it will be useful to write $\frac{\mathbf{p}^2}{2m} = H_0 + Ze^2/r$ and insert it in H_{rel}

$$H_{rel} = -\frac{1}{2mc^2}\left[H_0^2 + Ze^2\left(H_0\frac{1}{r} + \frac{1}{r}H_0\right) + \frac{Z^2e^2}{r^2}\right] \tag{12.90}$$

where the orders of operators are respected.

We may now compare the size of H_{rel} to that of H_0 for a generic state of the atom, with $Z = 1$. We had computed before $\langle\frac{\mathbf{p}^2}{2m}\rangle \sim \langle\frac{e^2}{2r}\rangle \sim \langle H_0\rangle$ with $\langle H_0\rangle \sim \frac{1}{2}mc^2\alpha^2$, where $\alpha = e^2/\hbar c \approx 1/137$. Using these for orders of magnitude we may estimate

$$\frac{\langle H_{rel}\rangle}{\langle H_0\rangle} \approx \frac{\langle H_0\rangle^2}{2mc^2\langle H_0\rangle} = \frac{1}{4}\alpha^2 \approx 1.3\times10^{-5}. \tag{12.91}$$

This shows that the relativistic correction is fairly small, and that perturbation theory at first order would provide an adequate approximation.

12.6.2 Spin-orbit coupling

We begin with a classical argument to motivate the additional energy due to the spin-orbit coupling. From the point of view of the nucleus the electron goes around it, say counterclockwise. However, from the point of view of the electron the nucleus goes around it in a circle in the opposite direction, say clockwise. The moving charged nucleus creates an electric current, which in turn produces a magnetic field $\mathbf{B}$ at the center of the circle where the electron is located. The magnetic moment of the spinning electron $\boldsymbol{\mu}$ interacts with the magnetic field and this produces an additional interaction term in the Hamiltonian

$$H_{L\cdot S} = -\frac{1}{2}\boldsymbol{\mu}\cdot\mathbf{B} \tag{12.92}$$

The extra factor of $\frac{1}{2}$ is explained by the "Thomas precession". A more direct explanation is provided by the Dirac equation in the non-relativistic approximation. This additional energy must be added to the total Hamiltonian of the Hydrogen atom. As we will see, it is proportional to the dot product of the orbital angular momentum and the spin of the electron $L \cdot S$, and this is why it is called the "spin-orbit coupling".

Every spinning particle has a magnetic moment proportional to its spin and inversely proportional to its mass

$$\boldsymbol{\mu} = \frac{ge\mathbf{S}}{2mc}, \tag{12.93}$$

where g is the gyro-magnetic ratio. For the electron we have $g = 2$ as explained by the Dirac equation. The induced magnetic field at the center of the circle is given in terms of the velocity tangential to the circle $\mathbf{v} = \mathbf{p}/m$, and the electric field $\mathbf{E}(r)$

$$\mathbf{B} = -\frac{1}{c}(\mathbf{v} \times \mathbf{E}).$$

The electric field is computed in terms of the scalar potential, which is the Coulomb potential (there is no time dependent vector potential in this problem)

$$\mathbf{E} = -\boldsymbol{\nabla}\Phi = -\boldsymbol{\nabla}\left(\frac{-Ze}{r}\right) = \frac{-Ze}{r^3}\mathbf{r}$$

Therefore

$$\mathbf{B} = \frac{Ze}{cr^3}(\frac{\mathbf{p}}{m} \times \mathbf{r}) = -\frac{Ze}{mcr^3}\mathbf{L}$$

where $\mathbf{L}$ is the angular momentum. So now we have the energy

$$H_{L\cdot S} = -\frac{1}{2}\left(\frac{2e\mathbf{S}}{2mc}\right) \cdot \left(-\frac{Ze}{mcr^3}\mathbf{L}\right) = \frac{1}{2}\frac{Ze^2}{m^2c^2r^3}\mathbf{S} \cdot \mathbf{L} = \frac{e^2\alpha^2}{2a_0}\frac{a_0^3}{r^3}\frac{\mathbf{S} \cdot \mathbf{L}}{\hbar^2} \tag{12.94}$$

where we have used the relations for the fine structure constant $\alpha = e^2/\hbar c$ and the Bohr radius $a_0 = \hbar^2/\left(me^2\right)$ that we had learned in the study of the H-atom. We may again estimate the size of this correction for $Z = 1$ by using $\langle H_0 \rangle \sim \frac{1}{2}mc^2\alpha^2$. Then, we find

$$\frac{\langle H_{L\cdot S} \rangle}{\langle H_0 \rangle} \sim \alpha^2 \tag{12.95}$$

which is of similar order of magnitude to the relativistic correction. Therefore $H_{L\cdot S}$ and H_{rel} must be taken into account simultaneously as corrections of the same order.

12.6.3 First order pertubation

The total Hamiltonian is

$$H = H_0 + H' \tag{12.96}$$

$$H_0 = \frac{p^2}{2m} - \frac{Ze^2}{r}, \quad H' = H_{rel} + H_{L\cdot S} \tag{12.97}$$

Since we have argued that H' is about 10^{-4} smaller than H_0 we may apply first order perturbation theory to obtain the fine structure correction.

The basis of states must now include the information about the spin and the angular momentum of the electron. Therefore, the unperturbed states may be labelled as $|E_n^0\rangle \sim |nlm_l; sm_s\rangle$. This means that every state of the H-atom we had discussed before without spin, is now doubled due to the two possible states of the spin $m_s = \pm\frac{1}{2}$. However, due to the spin-orbit coupling it is more convenient to work in the total angular momentum basis $|E_n^0\rangle = |nls; jm_j\rangle$, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and the eigenvalues $|jm_j\rangle$ correspond to the operators $\mathbf{J}^2 \to \hbar^2 j(j+1)$ and $\mathbf{J}_3 \to \hbar m_j$ when applied on the states. With this choice of basis vectors, H_0 and H' are diagonal as far as spin is concerned, since we may express $\mathbf{L}\cdot\mathbf{S}$ as follows :

$$\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\left[(\mathbf{L}+\mathbf{S})^2 - \mathbf{L}^2 - \mathbf{S}^2\right] = \frac{1}{2}\left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right)$$
$$\to \frac{\hbar^2}{2}\left(j(j+1) - l(l+1) - \frac{3}{4}\right) \text{ on states.}$$

When j takes the values $j = l \pm \frac{1}{2}$ the quantity in paranhesis becomes l or $-(l+1)$ respectively.

The lowest lying H-atom levels are labelled as follows. For $n = 1$ the orbital angular momentum is $l = 0$ and the electron spin is $s = \frac{1}{2}$, therefore the total spin is $j = \frac{1}{2}$ and $m_j = \pm\frac{1}{2}$. In atomic physicists's notation these two states are labelled as $|1s_{1/2}\rangle$ where the m_j label is suppressed. Before H' is included, the two $n = 1$ levels are degenerate and have energy $E_1^0 = -\frac{1}{2}mc^2\alpha^2 = -\frac{e^2}{2a_0} = -13.6$ eV. Similarly for $n = 2$, the orbital angular momentum is $l = 0, 1$ and the electron spin is $s = \frac{1}{2}$, therefore the total spin is $j = \frac{1}{2}$ ($l = 0$) or $j = \frac{1}{2}, \frac{3}{2}$ ($l = 1$). In atomic physicists's notation these eight states are labelled as $|2s_{1/2}\rangle$ or $|2p_{1/2}\rangle, |2p_{3/2}\rangle$, where again the m_j labels are suppressed. Before H' is included the eight level n=2 states are degenerate and have energy $E_2^0 = -\frac{e^2}{4a_0} = -3.4$ eV. We can go on in this way to define to define the zeroth order basis, and then construct the matrix elements of $H_0 + H'$ in this basis as follows (the m_j labels are suppressed)

$$(H_0+H')_{nm} = \begin{array}{c|cccc} & 1s_{1/2} & 2s_{1/2} & 2p_{1/2} & 2p_{3/2}\ \cdots \\ \hline 1s_{1/2} & E_1^0 + k_{1s} & h_{12} & 0 & 0\ \cdots \\ 2s_{1/2} & h_{12}^* & E_2^0 + k_{2s} & 0 & 0\ \cdots \\ 2p_{1/2} & 0 & 0 & E_2^0 + k_{2p_1} & 0\ \cdots \\ 2p_{3/2} & 0 & 0 & 0 & E_2^0 + k_{2p_3}\cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \tag{12.98}$$

Let us first explain the zero entries. These are due to angular momentum conservation. Since $H_0 + H'$ is invariant under rotations $\mathbf{J}$ (including rotation of the spin $\mathbf{S}$) it commutes with the total generator of rotations $\mathbf{J}$. Since $H_0 + H'$ is a scalar under rotations its the matrix elements between different values of j, m_j vanish. Furthermore $H_0 + H'$ commutes also with $\mathbf{L}^2$, and therefore its matrix

elements between different values of l also must vanish. The matrix element h_{12} does not vanish since the total spin j and orbital quantum number l of the states $1s_{1/2}, 2s_{1/2}$ are the same. However, off diagonal elements, such as h_{12}, contribute only to second order in perturbation theory (see Eqs.(12.51,12.53)). Therefore the effects of off diagonal terms will be negligible compared to the diagonal elements $k_{1s}, k_{2s}, k_{2p_1}, k_{2p_3}$ that contribute in first order.

The general matrix element of H' is computed in the zeroth order basis as follows

$$H'_{nl,n'l} = \langle nlj|H'|n'lj\rangle \tag{12.99}$$

where $|nlj\rangle$ are the H_0 eigenstates with the quantum numbers $s = \frac{1}{2}$ and m_j suppressed. Note the same values of l, j in the bra and ket because of the conservation laws mentioned above. The matrix elements of H_{rel} and H_{LS} given in Eqs.(12.90,12.94)) are computed as follows, where we use $H_0|nlj\rangle = E_n^0|nlj\rangle$, with $E_n^0 = -\frac{1}{2n^2}mc^2\alpha^2 = -\frac{e^2}{2n^2a_0} - \frac{13.6}{n^2}$ eV

$$\langle nlj|H_{rel}|n'lj\rangle = -\frac{1}{2mc^2}\left(\left(E_n^0\right)^2\delta_{nn'} + \left(E_n^0 + E_{n'}^0\right)\frac{e^2}{a_0}\langle nl|\frac{a_0}{r}|n'l\rangle + \frac{e^4}{a_0^2}\langle nl|\frac{a_0^2}{r^2}|n'l\rangle\right)$$
$$\langle nlj|H_{LS}|n'lj\rangle = \frac{e^2\alpha^2}{2a_0}\frac{1}{2}\left(j\left(j+1\right) - l\left(l+1\right) - \frac{3}{4}\right)\langle nl|\frac{a_0^3}{r^3}|n'l\rangle$$

The matrix elements of powers of r are given by

$$\langle nl|r^{-k}|n'l\rangle = \int_0^\infty dr\ r^{2-k}\langle nl|r\rangle\langle r|n'l\rangle = \int_0^\infty dr\ r^{2-k}R_{nl}\left(r\right)R_{n'l}\left(r\right),$$

where $R_{nl}\left(r\right)$ are the radial wavefunctions studied earlier in the chapter on the H-atom. Using these functions we evaluate the integrals and obtain

$$\langle nl|\frac{1}{r}|nl\rangle = \frac{1}{a_0}\frac{1}{n^2},\quad \langle nl|\frac{1}{r^2}|nl\rangle = \frac{1}{a_0^2}\frac{2}{n^3\left(2l+1\right)},$$
$$\langle nl|\frac{1}{r^3}|nl\rangle = \frac{1}{a_0^3}\frac{2}{n^3\left(2l+1\right)\left(l+1\right)l}\ ,\ \cdots$$

Combining these results we find the diagonal elements of H' which we define as $k_{nlj} \equiv \langle nlj|H'|nlj\rangle$

$$\begin{aligned} k_{nlj} &= E_n^0\alpha^2\left[\left(-\frac{3}{4n^2} + \frac{1}{n\left(l+1/2\right)}\right) - \frac{j\left(j+1\right) - l\left(l+1\right) - \frac{3}{4}}{n\left(2l+1\right)\left(l+1\right)l}\right] \\ &= E_n^0\alpha^2\left(-\frac{3}{4n^2} + \frac{1}{n\left(j+1/2\right)}\right) \end{aligned} \tag{12.100}$$

where the equivalence of the the two expressions can be verified for $j = l \pm \frac{1}{2}$. Note that the result depends only on j rather than j and l independently. Using these results for $n = 1, 2$ and $l = 0, 1$ we get the matrix elements of $\left(H_0 + H'\right)_{nm}$

in the lowest states of the H-atom. These come out as follows

$$E_1^0 + k_{1s} = -\frac{e^2}{2a_0}\left(1 + \frac{1}{4}\alpha^2\right), \quad E_2^0 + k_{2s} = -\frac{e^2}{2a_0}\frac{1}{2^2}\left(1 + \frac{5}{16}\alpha^2\right)$$
$$E_2^0 + k_{2p_1} = -\frac{e^2}{2a_0}\frac{1}{2^2}\left(1 + \frac{5}{16}\alpha^2\right), \quad E_2^0 + k_{2p_3} = -\frac{e^2}{2a_0}\frac{1}{2^2}\left(1 + \frac{1}{16}\alpha^2\right)$$

The terms of order α^2, namely the k's, are the corrections to the energy levels in first order perturbation theory, as see from Eqs.(**??**,**??**).

The off diagonal term h_{12} can be computed in a similar way, and it is found to be of similar order of magnitude to the k's above. h_{12} contributes to the correction of the energy levels or the states in second order perturbation theory through Eqs.(12.51,12.53). Since the term $|h_{12}|^2 / \left(E_1^0 - E_2^0\right)$ is much smaller compared to the k's by a a factor of $\alpha^2 \sim 10^{-4}$, the second order perturbation correction can safely be neglected in this level of approximation to the H-atom. Note that the $2s_{1/2}$ and $2p_{1/2}$ states are still degenerate at this level of accuracy. In the real world there is a tiny splitting called the Lamb shift, whose measurement by Lamb was the source of a Nobel prize. The shifts of the energy levels due to all corrections are shown in the diagram.

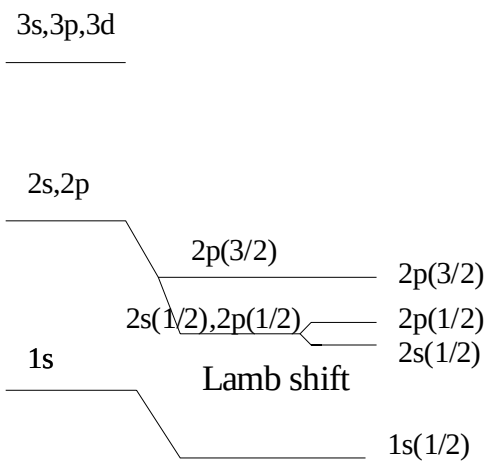


Fig.12.1. Hyperfine structure and Lamb shift.

There are further small corrections due to additional physical effects. One of them is the vacuum quantum fluctuations of the electromagnetic field and pair creation of electrons and positrons out of the vacuum. This is understood in detail in quantum electrodynamics, and explains the Lamb shift to 12 decimal places in agreement with experiment. Another small physical effect is the the non-zero size of the nucleus which leads to a potential energy $V(r)$ for a charge distribution in the vicinity of the nucleus as opposed to the point charge

approximation of $V=-e^2/r$. These corrections are treated in the problems at the end of this chapter.

12.7 H-atom in an external magnetic field

When an electron of charge $-e$ is placed in an external electromagnetic field, its kinetic energy must be modified by replacing the momentum with the covariant momentum and by adding the energy due to the interaction with the external scalar potential. Furthermore the magnetic moment of the electron interacts with the magnetic field and produces additional energy. Hence, including the interaction with a potential $V(r)$, the total Hamiltonian is

$$H=\frac{1}{2m}\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)^2-eA_0+\boldsymbol{\mu}\cdot\mathbf{B}+\mathbf{V}(r) \tag{12.101}$$

The magnetic moment is proportional to the spin of the particle $\boldsymbol{\mu}=g\mu_0\mathbf{S}$ where μ_0 is called the Bohr magneton, and g is gyromagnetic ratio. For the electron, the Dirac equation explains that $g=2$ (there are also extremely small quantum corrections) and μ_0 has the value

$$\mu_0=\frac{e\hbar}{2mc}\simeq 0.6\times 10^{-8}\ Gauss. \tag{12.102}$$

In quantum mechanics, with $\mathbf{p}\rightarrow-i\hbar\boldsymbol{\nabla}$, we have

$$\frac{1}{2m}\left(i\hbar\boldsymbol{\nabla}+\frac{e}{c}\mathbf{A}\right)^2\psi(\mathbf{r})=\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2+i\frac{\hbar e}{mc}\mathbf{A}\cdot\boldsymbol{\nabla}\right)\psi(\mathbf{r})+\left(i\frac{\hbar e}{2mc}\boldsymbol{\nabla}\cdot\mathbf{A}+\frac{e^2}{2mc^2}\mathbf{A}^2\right)\psi(\mathbf{r})$$

where the first parenthesis is a differential operator while the last one can be regarded an additional contribution to the potential. In the Coulomb gauge one may take $\boldsymbol{\nabla}\cdot\mathbf{A}=0$.

Now consider a constant external magnetic field $\mathbf{B}$, and no external electric field. This implies $A_0=0$ and $\mathbf{A}=\frac{1}{2}\mathbf{r}\times\mathbf{B}$, which satisfies $\boldsymbol{\nabla}\times\mathbf{A}=\mathbf{B}$. Then we note that the term

$$i\frac{\hbar e}{2mc}\mathbf{A}\cdot\boldsymbol{\nabla}=i\frac{\hbar e}{2mc}(\mathbf{r}\times\mathbf{B})\cdot\boldsymbol{\nabla}=\frac{e}{2mc}\mathbf{B}\cdot(\mathbf{r}\times\mathbf{p})=\mu_0\mathbf{B}\cdot\mathbf{L} \tag{12.103}$$

produces an interaction between the magnetic field and the orbital angular momentum proportional to the Bohr magneton μ_0. Combined with the magnetic moment interaction these two terms give the energy $\mu_0\mathbf{B}\cdot(\mathbf{L}+g\mathbf{S})$.

Now consider the Hydrogen atom with the potential $V(r)$ that includes the relativistic and spin-orbit corrections, and then place it in an external magnetic field. The Hamiltonian is $H=H_0+H'$ where

$$H_0=\frac{p^2}{2m}-\frac{Ze^2}{r}+H_{rel}+H_{L\cdot S},\quad H'=\mu_0B(J_3+S_3). \tag{12.104}$$

where we assume that $\mathbf{B}$ points in the z direction, and used $L_3 + 2S_3 = J_3 + S_3$. We recall $H_{rel} + H_{L\cdot S}$ are of order $10^{-4} eV$. Using the value of μ_0 we estimate H' is of order $\left(10^{-8} eV\right) \times \frac{B}{Gauss}$. Therefore H' is small compared to $H_{rel} + H_{L\cdot S}$ for $B < 10^4$ $Gauss$ which is an extremely large field.

The energy levels of H_0 were computed in the previous section in the basis $|nljm\rangle$. We recall the diagonal elements

$$\langle nljm|H_0|nljm\rangle = E_{nj} = -\frac{e^2}{2a_0}\frac{1}{n^2}\left(1 + \alpha^2\left(-\frac{3}{4n^2} + \frac{1}{n\,(j+1/2)}\right)\right)$$

We now compute the matrix elements of H' in the same basis

$$\langle nljm|H'|n'l'j'm'\rangle = \mu_0 B\left(m\delta_{jj'} + (S_3)_{jj'}\right)\delta_{mm'}\delta_{ll'}\delta_{mm'} \tag{12.105}$$

where

$$(S_3)_{jj'} = \sum_{l,m_l,m_s} \langle jm|lm_l\frac{1}{2}m_s\rangle m_s \langle lm_l\frac{1}{2}m_s|j'm\rangle \tag{12.106}$$

Here $m_l = m \mp m_s$, $m_s = \pm 1/2$ and $j = l \pm 1/2$. Using the Clebsch-Gordan coefficients, we compute

$$for\ \ j = j' = l \pm 1/2:\ (S_3)_{jj} = \pm\frac{m}{2l+1}, \tag{12.107}$$

$$for\ \ j = l \pm 1/2\ \ and\ \ j' = l \mp 1/2:\ (S_3)_{jj'} = -\frac{\sqrt{(l+1/2)^2 - m^2}}{2l+1}. \tag{12.108}$$

With this information we can setup the matrix for the total Hamiltonian

$1s_{1/2}$	$2s_{1/2}$	$2p_{1/2}$	$2p_{3/2}$ $\cdots$
$-\frac{e^2}{2a_0}\left(1+\frac{1}{4}\alpha^2\right)$ $+2m\mu_0 B$	h_{12}	0	0 $\cdots$
h_{12}^*	$-\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right)$ $+2m\mu_0 B$	0	0 $\cdots$
0	0	$-\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right)$ $+\frac{2}{3}m\mu_0 B$	$-\frac{\mu_0 B}{3}\sqrt{\frac{9}{4}-m^2}\,\delta_{\lvert m\rvert,\frac{1}{2}}$
0 $\vdots$	0 $\vdots$	$-\frac{\mu_0 B}{3}\sqrt{\frac{9}{4}-m^2}\,\delta_{\lvert m\rvert,\frac{1}{2}}$	$-\frac{e^2}{8a_0}\left(1+\frac{1}{16}\alpha^2\right)$ $+\frac{4}{3}m\mu_0 B$ $\vdots$

This matrix is easily diagonalized. In the first block $h_{12} \sim \frac{e^2\alpha^2}{2a_0}$ is neglected since its contribution to the eigenvalue in second order perturbation theory is of order $\left(|h_{12}|^2 / \frac{e^2}{2a_0}\right) \sim \frac{e^2}{2a_0}\alpha^4$. Therefore, the eigenvalues in the first block are approximately the diagonal entries in the matrix above. Similarly, for the $2p_{3/2}$, $m = \pm 3/2$ states, the matrix is already diagonal. For the second block that describes the $2p_{1/2}, 2p_{3/2}$ states with $m = \pm 1/2$, the exact eigenvalues are

$$E_{2p}^{\pm}(m) = -\frac{e^2}{8a_0}\left(1+\frac{3}{16}\alpha^2\right) + m\mu_0 B \pm \left[\left(\frac{e^2}{8a_0}\frac{2}{16}\alpha^2 + \frac{1}{3}m\mu_0 B\right)^2 + (\mu_0 B)^2\left(\frac{1}{4} - \frac{m^2}{9}\right)\delta_{|m|,\frac{1}{2}}\right]^{1/2}$$

For non-zero magnetic field all degeneracies have been broken, and every state has a different eigenvalue. We examine the splitting of the states for various values of m as the value of B increases from zero toward large values.

The behavior of the energy levels as a function of B is shown in Figure (12.2) which is not to scale, and exagerated, to illustrate the effects. The vertical axis is energy and the horizontal axis is the magnetic field. Before the hyperfine splitting the two lowest levels of the H-atom are shown (at zero B). The bottom one is the $1s_{1/2}$ and the top one corresponds to the degenerate $2s_{1/2}, 2p_{1/2}$,$2\mathrm{p}_{3/2}$ states. After the hyperfine splitting one sees three levels (still at zero B). The bottom one is the $1s_{1/2}$, the middle one is the degenerate states $2\mathrm{s}_{1/2}, 2p_{1/2}$ (except for the Lamb shift, not shown) and the top one is the $2p_{3/2}$. In the presence of the magnetic field all remaining degeneracies due to rotational symmetry are broken as shown (non-zero B). As B increases the splitting gets larger, as shown by the diverging lines. Note that there are critical values of the magnetic field at which there is level crossing which can occur for many levels. These are experimentally observed and analysed, and are found to be in agreement with the computations.

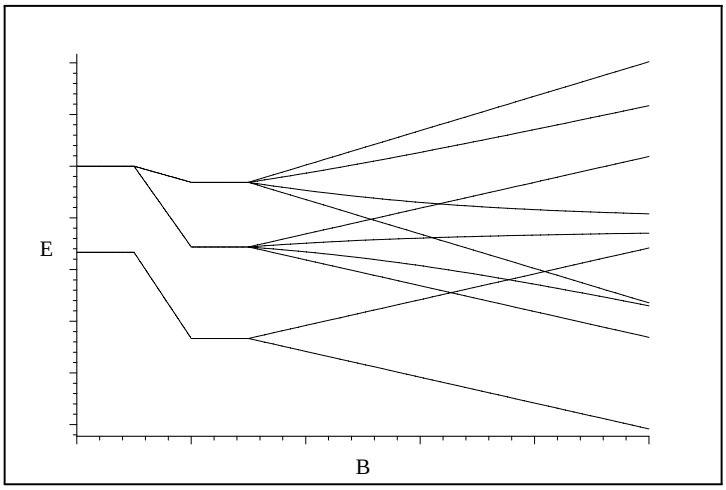


Fig.12.2 - bottom $1s_{\frac{1}{2}}$, middle $2s_{\frac{1}{2}}, 2p_{\frac{1}{2}}$, top $2p_{\frac{3}{2}}$

For small of $\mu_0 B$ first and second order perturbation theory is sufficient to describe the splitting. This splitting of levels is called the Zeeman effect. This amounts to the expansion of the expressions above up to the second power in $(\mu_0 B)$. Of course, for the $\left(1s_{1/2}, m = \pm 1/2\right)$, $\left(2s_{1/2}, m = \pm 1/2\right)$ and $\left(2p_{3/2}, m = \pm 3/2\right)$ there is only the first power, so that for various values of m's the splitting is linear in $(\mu_0 B)$. For the $E_{2p}^{\pm}(m), m = \pm 1/2$ eigenvalues the expansion includes the second and higher powers as well which follows from the

Taylor expansion

$$E_{2p}^{-}=-\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right)+\frac{1}{3}m\mu_0 B-16\frac{\left(\mu_0 B\right)^2}{\alpha}\frac{2a_0}{e^2}\left(\frac{1}{4}-\frac{m^2}{9}\right)\delta_{|m|,\frac{1}{2}}+\cdots$$
$$E_{2p}^{+}=-\frac{e^2}{8a_0}\left(1+\frac{1}{16}\alpha^2\right)+\frac{1}{3}m\mu_0 B+16\frac{\left(\mu_0 B\right)^2}{\alpha}\frac{2a_0}{e^2}\left(\frac{1}{4}-\frac{m^2}{9}\right)\delta_{|m|,\frac{1}{2}}+\cdots$$

This amounts to second order perturbation theory. From first order perturbation theory we see that the splitting is proportional to $m\left(\mu_0 B\right)$, therefore for positive m the level goes up and for negative values it goes down. Thus the splitting of the four levels of $2p_{3/2}$ is such that they appear in the order $m=3/2,1/2,-1/2,-3/2$ from top to bottom. For the degenerate $2s_{1/2},2p_{1/2}$ levels the slope of the splitting is smaller for the $2p_{1/2}$ states. Therefore the inner two lines that emanate from the middle level in the figure are the $m=\pm 1/2$ levels of the $2p_{1/2}$ states, while the outer lines are the $m=\pm 1/2$ levels of the $2s_{1/2}$ states.

For large values of B the splitting of the levels increases linearly only for some every levels, while it tends to some constant values for others. This is called the Paschen-Bach effect. The linear nature is evident for the $\left(1s_{1/2},m=\pm 1/2\right)$, $\left(2s_{1/2},m=\pm 1/2\right)$ and $\left(2p_{3/2},m=\pm 3/2\right)$, while for the $2p_{1/2},2p_{3/2}$ states with $m=\pm 1/2$, the behavior is seen by computing the asymptotic expansion of $E_{2p}^{\pm}\left(m\right)$. We find

$$\begin{aligned}
E_{2p}^{-}\left(\frac{1}{2}\right)&=-\frac{e^2}{8a_0}\left(1+\frac{19}{48}\alpha^2\right)+\cdots,\\
E_{2p}^{+}\left(-\frac{1}{2}\right)&=-\frac{e^2}{8a_0}\left(1+\frac{17}{48}\alpha^2\right)+\cdots\\
E_{2p}^{+}\left(\frac{1}{2}\right)&=-\frac{e^2}{8a_0}\left(1+\frac{17}{48}\alpha^2\right)+\left(\mu_0 B\right)+\cdots,\\
E_{2p}^{-}\left(-\frac{1}{2}\right)&=-\frac{e^2}{8a_0}\left(1+\frac{19}{48}\alpha^2\right)-\left(\mu_0 B\right)+\cdots
\end{aligned}$$

The neglected terms vanish as $B\to\infty$, and have the form $\pm\left(\frac{1}{4}-\frac{m^2}{9}\right)\left(\frac{e^2}{2a_0}\right)^2\alpha^4\times\left(\left(32\right)^2\left(\mu_0 B\right)\right)^{-1}$. As seen in the figure, the energies of the $2p_{1/2},m=1/2$ state and $2p_{3/2},m=-1/2$ state approach the constant values $E_{2p}^{+}\left(-\frac{1}{2}\right),E_{2p}^{-}\left(\frac{1}{2}\right)$ respectively for large values of B as computed above.

12.8 H-atom in an external electric field

We recall the Hamiltonian in the presence of an electromagnetic field

$$H=\frac{1}{2m}\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)^2-eA_0-\boldsymbol{\mu}\cdot\mathbf{B}+\mathbf{V}\left(r\right)$$

For a constant external electric field the gauge potential is given by $\mathbf{A}$ =0 and $A_0 = -\mathbf{E}\cdot\mathbf{r}$, so that $\mathbf{E} = -\boldsymbol{\nabla} A_0$ is verified. We choose the electric field in the z direction, hence $A_0 = -Ez = Er\cos\theta$. Therefore the Hamiltonian takes the form

$$H = H_0 + eEr\cos\theta \tag{12.109}$$

where H_0 is the Hydrogen atom Hamiltonian. We compute the matrix elements in $|nljm_j\rangle$ basis. Of course H_0 is diagonal as in the previous sections. The perturbation has the following matrix elements

$$\langle nljm|eEr\cos\theta|n'l'j'm'\rangle = eE\delta_{mm'}\langle nl|r|n'l'\rangle\langle ljm|\cos\theta|l'j'm\rangle$$

We will concentrate on a given level n, such as $n = n' = 2$ as an illustration. The orbital angular momentum l, l' must differ by one unit because the operator $\cos\theta$ or $\mathbf{r}$ has odd parity while the states have parity $(-1)^l$. Therefore we take the values $l = 0$ and $l' = 1$. Thus we need to calculate $\langle 2s_{1/2}|eEr\cos\theta|2p_{1/2,3/2}\rangle$

$$\langle 2s_{1/2}, m|eEr\cos\theta|2p_{1/2,3/2}, m\rangle = eE\langle 2s|r|2p\rangle\ \langle s_{1/2}, \pm\frac{1}{2}|\cos\theta|p_{1/2,3/2}, \pm\frac{1}{2}\rangle$$

First we compute

$$\langle 2s|r|2p\rangle = \int_0^\infty drr^3 R_{20}(r)\, R_{21}(r) = -3\sqrt{3}\frac{a_0}{Z}$$

Next we compute the matrix elements of $\cos\theta$ by transforming into the $|lm_l, sm_s\rangle$ states by using Clebsch-Gordan coefficients $\langle j, m|lm_l, sm_s\rangle$. Furthermore we write $\cos\theta$ in terms of the spherical harmonics $\cos\theta = \sqrt{4\pi/3}Y_{10}(\theta)$

$$\langle s_{1/2}, m|\cos\theta|p_{1/2,3/2}, m\rangle = \langle\frac{1}{2}, m|0, 0, \frac{1}{2}, m\rangle\ \langle 0, 0|\cos\theta|1, 0\rangle\ \langle 1, 0, \frac{1}{2}, m|p_{1/2,3/2}, m\rangle$$

On the left side only $l = 0$ contributes since we start with the $l = 0$ state $s_{1/2}$. This also fixes $m_l = 0$. Similarly on the right hand side we must have $l = 1$. Since $\cos\theta = \sqrt{4\pi/3}Y_{10}$ also has $m = 0$ on the right hand side we must also have $m_l = 0$. The Clebsch-Gordan are obtained from a table $\langle 1, 0, \frac{1}{2}, \pm\frac{1}{2}|\frac{1}{2}, \pm\frac{1}{2}\rangle = \mp\sqrt{1/3}$, and $\langle 1, 0, \frac{1}{2}, \pm\frac{1}{2}|\frac{3}{2}, \pm\frac{1}{2}\rangle = \sqrt{2/3}$, while

$$\langle 0, 0|\cos\theta|1, 0\rangle = \sqrt{4\pi/3}\int d\Omega Y_{00}Y_{10}Y_{10} = \sqrt{1/3}$$

Therefore we have the matrix elements

$$\langle s_{1/2}, \pm\frac{1}{2}|eEr\cos\theta|p_{1/2}, \pm\frac{1}{2}\rangle = eE\left(-3\sqrt{3}\frac{a_0}{Z}\right)\left(\mp\sqrt{1/3}\sqrt{1/3}\right) = \pm\frac{\sqrt{3}}{Z}eEa_0$$
$$\langle s_{1/2}, \pm\frac{1}{2}|eEr\cos\theta|p_{3/2}, \pm\frac{1}{2}\rangle = eE\left(-3\sqrt{3}\frac{a_0}{Z}\right)\left(\sqrt{2/3}\sqrt{1/3}\right) = -\frac{\sqrt{6}}{Z}eEa_0.$$

This provides the necessary information to construct the matrix H

	$2s_{1/2}$	$2p_{1/2}$	$2p_{3/2}$ $\cdots$
$2s_{1/2}$	$-\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right)-\frac{\Delta}{2}$	$\pm\sqrt{3}eEa_0$	$-\sqrt{6}eEa_0$ $\cdots$
$2p_{1/2}$	$\pm\sqrt{3}eEa_0$	$-\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right)+\frac{\Delta}{2}$	0 $\cdots$
$2p_{3/2}$	$-\sqrt{6}eEa_0$	0	$-\frac{e^2}{8a_0}\left(1+\frac{1}{16}\alpha^2\right)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

(12.110)

The Lamb shift Δ is also included.

If we diagonalize the 2×2 sector of the $2s_{1/2}, 2p_{1/2}$ states exactly, we find the eigenvalues

$$E_\pm = -\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right) \pm \frac{1}{2}\sqrt{\Delta^2+12e^2E^2a_0^2} \tag{12.111}$$

For small electric field this has the expansion

$$eEa_0 \ll \Delta:\ E_\pm = -\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right) \pm \frac{\Delta}{2} \pm \frac{3e^2E^2a_0^2}{\Delta}+\cdots \tag{12.112}$$

This is second order perturbation theory. For large electric field we have another expansion

$$eEa_0 \gg \Delta:\ E_\pm = -\frac{e^2}{8a_0}\left(1+\frac{5}{16}\alpha^2\right) \pm \sqrt{3}eEa_0 \pm \frac{\sqrt{3}\Delta^2}{24eEa_0}+\cdots \tag{12.113}$$

In the new basis the difference between eigenvalues $E_+ - E_- \approx 2\sqrt{3}eEa_0$ is comparable to the off diagonal elements that mix with the $2p_{3/2}$ states. Therefore, one must apply the nearly degenerate perturbation theory methods to include the effects of the remaining off diagonal terms. The situation is somewhat similar to the one discussed in Eqs.(12.64-12.85), now with a matrix of the form

$$\begin{pmatrix} A-a & 0 & a \\ 0 & A+a & a \\ a & a & B \end{pmatrix} \tag{12.114}$$

The solution of the complete problem is left as an exercise for the student in a homework problem.

12.9 Problems

1. Apply the general formulas for the formal exact solution for $E_n, |E_n\rangle, Z_n, U^\dagger_{mn}$ to the 2-level problem, and show that you obtain the same exact results as the standard matrix diagonalization methods.

2. Consider two distinguishable particles of spin $\frac{1}{2}$ (e.g. electron and positron) interacting via a spin dependent Hamiltonian of the form

$$H = A\vec{S}^{(1)} \cdot \vec{S}^{(2)} + B(S_z^{(1)} - S_x^{(2)}).$$

   Assuming that $B \ll A$, compute the the energy to second order and the normalized wavefunction to first order? (note that this is a 4×4 matrix problem).

3. Finite size nucleus.

The usual non-relativistic treatment of the hydrogen atom makes use of many physical idealizations. One of these consists of assuming that the proton is a point charge. In fact, the charge of the proton is distributed over a very small spatial region and is described by a spherically symmetrical charge density $\rho(r)$.

a) Compute the change in the potential from a pure Coulomb potential

$$\Delta V(r) = \begin{matrix} V^{(\rho)}(r) & \text{[distributed charge]} \\ -V(r) & \text{[point charge]} \end{matrix}$$

which arises from an arbitrary charge distribution $\rho(r)$, with, of course, the total charge fixed.

b) Compute approximately the first order energy shift in an S-state of hydrogen which results from this perturbation. That is, compute the first order energy shift to lowest order in $Z\alpha$. [Hint: ΔV is concentrated in a very small spatial region so that, to lowest order in $Z\alpha$, the wave function can be replaced by its value at the origin,

$$\psi_{ns}(0) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{na_0}\right)^{3/2}.$$

Your answer should involve only the mean square radius of the charge distribution.]

c) What is the order of magnitude in eV of this energy shift? What is the order of magnitude of the shift for P states? Compare the fine structure of H. How big are these effects for a muon bound to a proton?

4. Lamb shift.

The Lamb Shift - A crude estimate of the Lamb shift can be obtained through the following considerations: The dynamics of a normal mode of the electromagnetic field is equivalent to that of a harmonic oscillator (see class notes). Therefore, upon quantization each mode acquires a zero point energy of $\hbar\omega/2$ (vacuum energy). It can be checked that although the average E.M. field strength is zero in vacuum, the mean-square value of the field is non-zero. (This follows from $\langle 0|a|0\rangle = 0$, but $\langle 0|aa^{\dagger}|0\rangle \neq 0$.) This vacuum fluctuation of the E.M. field gives an extra interaction with any charged particle. Now, in the H-atom this causes a jiggling of the electron about its average position. This mean square fluctuation can be estimated to be (for a nucleus of charge Z)

$$\langle (\delta r)^2 \rangle = \frac{2\alpha}{\pi} \left(\frac{\hbar}{mc}\right)^2 \ln \frac{1}{Z\alpha} \quad .$$

Therefore, the electron "sees" a different E.M. potential at its various points about the average position. This leads then to an additional average energy

$$\langle V(\underset{\sim}{r}+\delta\underset{\sim}{r}) - V(\underset{\sim}{r}) \rangle = \langle \delta\underset{\sim}{r} \cdot \overset{0}{\underset{\sim}{\nabla}} V + \frac{1}{2}\delta r_i \delta r_j \frac{\partial^2 V}{\partial r_i \partial r_j} + \ldots \rangle$$

$$= \frac{1}{6}\langle (\delta r)^2 \rangle \nabla^2 V + \cdots \qquad \text{use: } \vec{\nabla}^2 \frac{1}{r} = -4\pi\delta^3(\vec{r})$$

where V is the Coulomb potential.

Calculate in 1st order perturbation theory the energy shift due to this interaction for any state $|n\ell\ldots\rangle$. In particular calculate the energy splitting for the $2s_{1/2}, 2p_{1/2}$ states (i.e. the Lamb Shift). How does your result compare with experiment? (experiment: 1057 MHz)

5. Apply the methods of Eqs.(12.64-12.85), or any of the other methods discussed in class, to find the eigenvalues and eigenstates of the Hamiltonian in Eq.(12.110)

Chapter 13

TIME DEPENDENT PROBLEMS

We now turn our attention to time dependent problems. We consider a Hamiltonian of the form $H = H_0 + H'(t)$ where we assume that all the eigenvalues and eigenstates of H_0 are known, and that the complete set of states are labelled as $|E_n\rangle$. When $H'(t)$ is zero, the time dependence of these states is simply $e^{-itH_0/\hbar}|E_n\rangle = |E_n\rangle e^{-itE_n/\hbar}$, implying that the energy eigenstate changes at most by a phase, and does not make a transition to another state. However, when $H'(t)$ is not zero, eigenstates of H_0 are no longer stationary in time. A common question is: what is the probability amplitude that the system will make a transition from some initial energy eigenstate $|E_n\rangle$ to some other final energy eigenstate $|E_m\rangle$ when the time dependent interaction is turned on?

More generally the system may be in some general initial state $|i\rangle$ which may or may not be an energy eigenstate of H_0, and we would like to compute the probability amplitude for making a transition to some general final state $|f\rangle$. To answer the question we need to compute the time development of the state $|i\rangle$ which is governed by the Schrödinger equation. As long as $H'(t)$ is zero we have already learned in past chapters that the time development of any state will be given by the time translation operator $|\psi, t\rangle = e^{-itH_0/\hbar}|\psi\rangle$. Now we will imagine that $H'(t)$ remains zero up to some time t_0, or that $H'(t)$ it is turned on very slowly so that it starts to become appreciable around the time $t = t_0$, and compute the subsequent time development of the state. This will be given by the time translation operator which we will call $U(t, t_0)$, so that

$$|i, t\rangle^+ = U(t, t_0)\,|i, t_0\rangle. \tag{13.1}$$

The superscript + is to remind us that the evolution of this state is determined by the full Hamiltonian $H = H_0 + H'(t)$. Our first task is to compute the operator $U(t, t_0)$ and then compute the transition amplitude.

While the initial state of the system develops in time, our measuring apparatus also develops in time. We will assume that our measuring apparatus is not

subject to the interactions described by $H'(t)$ and that its time development is governed by H_0. To find out if the system described by state $|i\rangle$ will make a transition to some final state $|f\rangle$ we must prepare a detector described by the state $|f\rangle$. Up to the time $t = t_0$, the time development of both states $|i\rangle, |f\rangle$ are governed by H_0,

$$|i, t_0\rangle = e^{-it_0 H_0/\hbar}|i\rangle, \quad |f, t_0\rangle = e^{-it_0 H_0/\hbar}|f\rangle. \tag{13.2}$$

But for $t > t_0$ the detector is still governed by H_0 while the system is governed by the total Hamiltonian $H = H_0 + H'(t)$. Therefore for $t > t_0$ the time dependence of the detector is given by

$$|f, t\rangle = e^{-i(t-t_0)H_0/\hbar}|f, t_0\rangle. \tag{13.3}$$

The transition amplitude $A_{fi}(t)$ measured at time t is simply the overlap between the system and the detector at time t, and is given by

$$A_{fi}(t, t_0) = \langle f, t|i, t\rangle^+ = \langle f, t_0|e^{i(t-t_0)H_0/\hbar} U(t, t_0)\, |i, t_0\rangle. \tag{13.4}$$

By substituting $|i, t_0\rangle, |f, t_0\rangle$ this can be further developed to the form

$$A_{fi}(t, t_0) = \langle f|e^{itH_0/\hbar} U(t, t_0)\, e^{-it_0 H_0/\hbar}|i\rangle, \tag{13.5}$$

where the states $|i\rangle, |f\rangle$ are the initial and final states well before the interaction was turned on. Note that on the left t appears while on the right t_0 appears. The time dependence of the transition amplitude will be determined by understanding the time translation operator $U(t, t_0)$ in more detail.

There are many possible physical circumstances in Nature to which this setup is applied. In some cases, such as the turning on of some external electromagnetic field at time t_0 is controlled by the experimentalist in the laboratory, then t_0 is some fixed time and t is some subsequent time. In other cases, such as the scattering of two free particles by short range forces, the interaction $H'(t)$ is turned on very slowly as the particles approach each other as a function of the distance between them; it becomes appreciable only for a very short time period, and it becomes negligible again as the particles fly apart after scattering. For such circumstances we will be interested in taking the limits $t_0 \to -\infty$ and $t \to \infty$, and then $|i\rangle, |f\rangle$ are interpreted as the initial and final states that describe free particles. In the latter case $A_{fi}(\infty, -\infty)$ is called the S-matrix, or the scattering matrix, which we will study in more detail in the next chapter.

Before we deal with the problem of a general time dependent $H'(t)$, let us consider the case of a perturbation H' which turns on at $t = t_0$, but remains constant as a function of time after that point. Then we already know the time translation operator

$$U(t, t_0) = \exp\left(-\frac{i}{\hbar}(t - t_0)(H_0 + H')\right), \text{ iff } \partial_t H' = 0. \tag{13.6}$$

which leads to the transition amplitude

$$A_{fi}(t, t_0) = \langle f|e^{itH_0/\hbar} e^{-i(t-t_0)\left(H_0+H'\right)/\hbar} e^{-it_0 H_0/\hbar}|i\rangle, \text{ iff } \partial_t H' = 0. \tag{13.7}$$

In this expression the time dependence of the process has been made quite explicit. Note that in general H_0, H' are operators that do not commute with each other, and therefore the exponentials cannot be combined naively. Beyond this point the computation depends on the details of the operators H_0, H'. We will return to this expression later when we consider some examples.

13.1 Interaction picture

To compute the time development operator for the general case $H'(t)$, we will distinguish between different formalisms that deal with time dependent problems in quantum mechanics. These are called the Schrödinger picture, the interaction picture and the Heisenberg picture.

The Schrödinger picture is the one we discussed above. In this case operators such as $\mathbf{r}, \mathbf{p}$ or their functions $O(\mathbf{r}, \mathbf{p})$ are all taken at time zero in the Hamiltonian formalism. The time dependence is all in the states and is governed by the Schrödinger equation

$$i\hbar\partial_t|\psi, t\rangle = H|\psi, t\rangle. \tag{13.8}$$

The Heisenberg picture is the opposite, where operators such as $\mathbf{r}(t), \mathbf{p}(t)$ or their functions $O(\mathbf{r}(t), \mathbf{p}(t))$ are taken as dynamical observables that develop in time while the states are all time independent. To distinguish the Heisenberg states from the Schrödinger states we will append an extra H as a subscript. When the Hamiltonian is independent of time the Heisenberg states are related to the Schrödinger states, which have no subscript, as

$$|\psi, t\rangle = e^{-itH/\hbar}|\psi\rangle_H \;\leftrightarrow |\psi\rangle_H = e^{itH/\hbar}|\psi, t\rangle, \tag{13.9}$$

while the Heisenberg operators O_H are related to the Schrödinger operators O by

$$O_H(t) = e^{itH/\hbar}Oe^{-itH/\hbar}. \tag{13.10}$$

The interaction picture is somewhere in between the Heisenberg and Schrödinger pictures. It is defined by

$$|\psi, t\rangle = e^{-itH_0/\hbar}|\psi, t\rangle_I \;\leftrightarrow |\psi, t\rangle_I = e^{itH_0/\hbar}|\psi, t\rangle \tag{13.11}$$

where, compared to the Heisenberg picture, $e^{itH_0/\hbar}$ appears instead of $e^{itH/\hbar}$. The equation of motion for the interaction picture state $|\psi, t\rangle_I$ is derived by computing its time derivative as follows

$$\begin{aligned} &i\hbar\partial_t|\psi, t\rangle_I = i\hbar\partial_t\left(e^{itH_0/\hbar}|\psi, t\rangle\right) = e^{itH_0/\hbar}\left(i\hbar\partial_t|\psi, t\rangle\right) - H_0e^{itH_0/\hbar}|\psi, t\rangle \\ &= e^{iH_0t/\hbar}\left(H_0 + H'(t)\right)|\psi, t\rangle - H_0|\psi, t\rangle_I = \left(e^{itH_0/\hbar}H'(t)e^{-itH_0/\hbar}\right)|\psi, t\rangle_I. \end{aligned}$$

Define now the interaction picture Hamiltonian

$$H_I(t) = e^{itH_0/\hbar}H'(t)e^{-itH_0/\hbar}. \tag{13.12}$$

Then we see that the time development of the interaction picture state is governed just by $H_I(t)$

$$i\hbar\partial_t|\psi,t\rangle_I = H_I(t)|\psi,t\rangle_I. \tag{13.13}$$

The solution of this equation may be written as

$$|\psi,t\rangle_I = U_I\left(t,t_1\right)|\psi,t_1\rangle_I \tag{13.14}$$

where $U_I\left(t,t_1\right)$ is the time translation operator in the interaction picture, and t_1 is some intial time. If the original perturbation $H'\left(t\right)$ were zero, then the interaction picture Hamiltonian would be zero, and the inteaction picture state $|\psi,t\rangle_I$ would be time independent. Indeed, in the interaction picture the initial and final states up to the point $t=t_0$ are just the original states $|i\rangle,|f\rangle$ as seen below

$$t\leq t_0:\ |i,t\rangle_I = e^{itH_0/\hbar}|i,t\rangle = e^{itH_0/\hbar}e^{-itH_0/\hbar}|i\rangle = |i\rangle, \tag{13.15}$$

and similarly for $|f,t\rangle_I = |f\rangle$. This makes sense since during the time period $t\leq t_0$ both $H'\left(t\right)$ and $H'_I\left(t\right)$ vanish. Therefore, in the interaction picture $|i,t_0\rangle_I = |i\rangle$ and the time development of the initial state at times later than t_0 is given by

$$t\geq t_0:\ |i,t\rangle_I = U_I\left(t,t_0\right)|i\rangle. \tag{13.16}$$

We compare this to the result in Eq.(13.1) obtained in the Schrödinger picture by using Eq.(13.11) which relates the two pictures

$$\begin{aligned} |i,t\rangle^+ &= U\left(t,t_0\right)|i,t_0\rangle = U\left(t,t_0\right)e^{-it_0H_0/\hbar}|i\rangle &(13.17)\\ &= e^{-itH_0/\hbar}|\psi,t\rangle_I = e^{itH_0/\hbar}U_I\left(t,t_0\right)|i\rangle &(13.18)\end{aligned}$$

From the last expressions in the two lines we extract the relation between the time translation operators in the two pictures. Noting that the relation is independent of the state $|i\rangle$, or that it must be true for all states, we can write the following equation that relates the time translation operators in the two pictures

$$U\left(t,t_0\right) = e^{-itH_0/\hbar}U_I\left(t,t_0\right)e^{it_0H_0/\hbar},\ \ U_I\left(t,t_0\right) = e^{itH_0/\hbar}U\left(t,t_0\right)e^{-it_0H_0/\hbar} \tag{13.19}$$

Inserting this in the transition amplitude given in Eq.(13.5) we find

$$A_{fi}\left(t,t_0\right) = \langle f|U_I\left(t,t_0\right)|i\rangle. \tag{13.20}$$

So the transition amplitude can be computed directly in the interaction picture as the matrix element of the time translation operator $U_I\left(t,t_0\right)$ sandwiched between the time independent states $|i\rangle,|f\rangle$. All the time dependence is now in $U_I\left(t,t_0\right)$ which remains to be computed.

13.2 Integral equations

We have already seen that the time translation operator in the interaction picture $U_I(t,t_0)$ satisfies an integral equation that represents the solution to the Schrödiger equation and incorporates also the initial condition $U_I(t_0t_0)=1$

$$U_I(t,t_0)=1-\frac{i}{\hbar}\int_{t_0}^{t}dt'\ H_I(t')U_I(t',t_0). \tag{13.21}$$

The formal solution was given in the previous section in the form of the Dyson series.

This integral equation implies also an integral equation for the time translation operator in the Schrödinger picture. By inserting Eq.(13.19) in Eq.(13.21) we obtain

$$e^{itH_0/\hbar}U(t,t_0)e^{-it_0H_0/\hbar}=1-\frac{i}{\hbar}\int_{t_0}^{t}dt'\ H_I(t')e^{it'H_0/\hbar}U(t',t_0)e^{-it_0H_0/\hbar}$$

Furthermore by replacing $H_I(t')=e^{it'H_0/\hbar}H'(t')\,e^{-it'H_0/\hbar}$ and re-arranging the factors we obtain

$$U(t,t_0)=e^{-i(t-t_0)H_0/\hbar}-\frac{i}{\hbar}\int_{t_0}^{t}dt'\ e^{-i\left(t-t'\right)H_0/\hbar}H'(t')U(t',t_0). \tag{13.22}$$

By applying both sides of this equation on the state $|i,t_0\rangle$ and using Eq.(13.1) we derive an integral equation for the state $|i,t\rangle^+$

$$|i,t\rangle^+=|i,t\rangle-\frac{i}{\hbar}\int_{t_0}^{t}dt'\ e^{-i\left(t-t'\right)H_0/\hbar}H'(t')|i,t'\rangle^+. \tag{13.23}$$

The state $|i,t\rangle$ without the superscript + implies that its time evolution is determined by H_0, not by the full H. The integral equation is a formal solution to the full Schrödinger equation $i\hbar\partial_t|i,t\rangle^+=\left(H_0+H'(t)\right)|i,t\rangle^+$ and satisfies the boundary condition. We will return to the investigation of this integral equation in later section.

13.3 Dyson series

Now we compute $U_I(t,t_0)$ formally for the general time dependent Hamiltonian $H'(t)$. First we derive a differential equation for $U_I(t,t_0)$ by taking the derivative of Eq.(13.16) and using Eq.(13.13)

$$i\hbar\partial_t|\psi,t\rangle_I=H'_I(t)\,|\psi,t\rangle_I\ \ \rightarrow\ \ i\hbar\partial_t\left(U_I(t,t_0)\,|\psi\rangle\right)=H'_I(t)\,U_I(t,t_0)\,|\psi\rangle$$

Since this equation must be satisfied for every $|\psi\rangle$ it must be true as an operator equation

$$i\hbar\partial_tU_I(t,t_0)=H'_I(t)\,U_I(t,t_0). \tag{13.24}$$

This can be converted to the following integral equation that satisfies not only the differential equation, but also the initial condition $U_I(t_0 t_0) = 1$

$$U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt'\ H_I(t')U_I(t',t_0). \tag{13.25}$$

A solution for $U_I(t,t_0)$ can now be obtained by recursively substituting the expression for $U_I(t,t_0)$ into the right hand side, so that

$$\begin{aligned} U_I(t,t_0) &= 1 - \tfrac{i}{\hbar}\int_{t_0}^{t} dt_1 H_I(t_1)\left[1 - \tfrac{i}{\hbar}\int_{t_0}^{t} dt_2 H_I(t_2)U_I(t_2 t_0)\right] \\ &= 1 - 1/\hbar\int_{t_0}^{t} dt_1 H_I(t_1) + \left(-\tfrac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt_1 H_1^{'}(t_1)\int_{t_0}^{t_1} dt_2\ H_I(t_2)U_I(t_2 t_0) \\ &= 1 - \tfrac{i}{\hbar}\int_{t_0}^{t} H_I(t_1)dt_1 + \left(-\tfrac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt_1 H_1^{'}(t_1)\int_{t_0}^{t_1} dt_2 H_I(t_2) \\ &\quad + \left(-\tfrac{i}{\hbar}\right)^3 \int_{t_0}^{t} dt_1 H_1^{'}(t_1)\int_{t_0}^{t_1} dt_2 H_I(t_2)\int_{t_0}^{t_2} dt_3 H_I(t_3) + \cdots \end{aligned}$$

In this way we obtain the infinite series

$$U_I(t,t_0) = \sum_{n=0}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \left[H_I(t_1)\cdots H_I(t_n)\right]. \tag{13.26}$$

Note that the $H_I(t_k)$ are time ordered, with the later times to the left side. This order could not be altered since the $H_I(t_k)$ gererally do not commute with each other at different times. Following this observation we define a time ordering operator T which is applied on a product of operators that depends on various instances of time. It instructs to write the operator with the larger time to the left side of the operator with the smaller time. It is given by the definition

$$T\left\{A(t_1)B(t_2)\right\} = A(t_1)B(t_2)\theta(t_1 - t_2) + B(t_2)A(t_1)\theta(t_2 - t_1). \tag{13.27}$$

Inside the time ordering operator T the order of the operators could be switched, even if the operators $A(t_1), B(t_2)$ do not commute, since the result is still the same after the definition of T is used

$$T\left\{A(t_1)B(t_2)\right\} = T\left\{B(t_2)A(t_1)\right\}. \tag{13.28}$$

Therefore, inside the T ordering operator any set of operators $A(t), B(t')$ commute at arbitrary times. Extending the definition of time ordering to the product of any number of operators, we can write the time ordered product

$$H_I(t_1)\cdots H_I(t_n) = T\left(H_I(t_1)\cdots H_I(t_n)\right), \tag{13.29}$$

and inside the time ordering operator T the orders of the $H_I(t_k)$ can be switched arbitrarily. Thus inside the time ordering, the products $(H_I(t_1)\cdots H_I(t_n))$ can be regarded as a completely symmetric function of the times $t_1, t_2, \cdots, t_n$ as if the $H_I(t_k)$ commute at any t_k.

It can be argued that for a symmetric function $f(t_1,\ldots,t_n)$, nested integrals can be rewritten as integrations in the full range as follows

$$\int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n f(t_1,\cdots t_n) = \frac{1}{n!}\int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n f(t_1,\cdots t_n).$$

This is easily verified for $n = 2$. First one can see that by renaming $t_1 \leftrightarrow t_2$ and using the symmetry $f(t_1, t_2) = f(t_2, t_1)$, it follows that $\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \; f(t_1, t_2)$ $= \int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 f(t_1, t_2)$. This says that integrating $f(t_1, t_2)$ in the two different regions that are indicated gives the same result. Then one can show, just by looking at a diagram of the integration region in the (t_1, t_2) plane, that the integral in the second region can also be written in a different order of nesting the integrals $\int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 f(t_1, t_2) = \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \; f(t_1, t_2)$. Now one can write the original nested integral as half the sum of the integrations in the two equivalent regions $\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \; f(t_1, t_2)$ and $\int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \; f(t_1, t_2)$. But this is precisely $\frac{1}{2!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \; f(t_1, t_2)$. The general proof for any n can be given with a similar reasoning.

Having established this result, we can write $U_I(t, t_0)$ in the following way by pulling out the time ordering operator in front of the sum

$$\begin{aligned} U_I(t, t_0) &= \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar} \right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n T \left(H_I(t_1) \cdots H_I(t_n) \right) \\ &= T \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar} \right)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n \left(H_I(t_1) \cdots H_I(t_n) \right) \\ &= T \exp \left(-\frac{i}{\hbar} \int_{t_0}^{t} H_I(t') dt' \right) \end{aligned} \tag{13.30}$$

In the last step the series is formally summed to what is called a "time ordered exponential". The last expression is a definition. It cannot be computed directly; its meaning and computation follows from the Dyson series form that appears in the second line or the first line of Eq.(13.30). Note that if the $H_I(t_k)$ commute at different times then there is no need for time ordering and the result becomes a usual exponential that solves the differential equation or the integral equation in Eqs.(13.24,13.25). In any case, the Dyson series is very useful for perturbative approximations to time dependent problems since the series is given in powers of the perturbation for any $H'(t)$.

13.3.1 Time dependent perturbation theory

Time dependent perturbation theory is now well defined. The perturbative approximation to the transition amplitude $A_{fi}(t, t_0) = \langle f | U_I(t, t_0) | i \rangle$ is given by computing the matrix elements

$$\begin{aligned} A_{fi}(t, t_0) &= \langle f | i \rangle - \frac{i}{\hbar} \int_{t_0}^{t} dt_1 \langle f | H_I(t_1) | i \rangle \\ &+ \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \langle f | H_I(t_1) H_I(t_2) | i \rangle + \cdots \end{aligned} \tag{13.31}$$

Recall that $H_I(t) = e^{itH_0/\hbar} H_I(t) e^{-itH_0/\hbar}$. Inserting the energy eigenstates as a complete set of states $1 = \sum_n |E_n^0\rangle\langle E_n^0|$, and assuming also that $|i\rangle, |f\rangle$ are

energy eigenstates (not so in general), we obtain

$$A_{fi}(t,t_0) = \delta_{fi} - \frac{i}{\hbar}\int_{t_0}^{t} dt_1 e^{it_1(E_f-E_i)/\hbar} H'_{fi}(t_1) \tag{13.32}$$
$$+\left(-\frac{i}{\hbar}\right)^2 \sum_n \int_{t_0}^{t} dt_1 e^{it_1(E_f-E_n)/\hbar} H'_{fn}(t_1) \int_{t_0}^{t_1} dt_2 e^{it_2(E_n-E_i)/\hbar} H'_{ni}(t_2) + \cdots$$

The time integrals cannot be done until the time dependence of $H'(t)$ is specified.

Of course, if $H'(t)$ is time independent, then the integrations are straightforward. In that case we expect agreement with the computation of Eq.(13.7) which should correspond to the full sum of the Dyson series

$$\begin{aligned} A_{fi}(t,t_0) &= \langle f|e^{itH_0/\hbar} e^{-\frac{i}{\hbar}(t-t_0)\left(H_0+H'\right)} e^{-it_0H_0/\hbar}|i\rangle, \text{ iff } \partial_t H' = 0 \\ &= e^{itE_f/\hbar} e^{it_0E_i/\hbar} \langle f|e^{-i(t-t_0)\left(H_0+H'\right)/\hbar}|i\rangle, \end{aligned} \tag{13.33}$$

After performing the integrals in Eq.(13.32) the series in powers of H' should match at each order. Therefore, the sum of the series should become the expression given above. The direct proof by performming the integrals is left as an exercise in problem 2. A more clever proof is the general discussion in the next section.

13.3.2 Summing up the Dyson series

There are a few special cases in which it is actually possible to sum up the infinite Dyson series. In this section we will discuss the case of a time independent H', and the case of a time dependence of the form

$$H_0 + H'(t) = e^{itH_1/\hbar} H_2 e^{-itH_1/\hbar}, \tag{13.34}$$

where both H_1, H_2 are time independent arbitrary operators. We will also discuss a general property of the Dyson series that is useful for transformations from one picture to another, including Schrödinger, interaction and more general pictures, and then give the most general form of $H'(t)$ for which the Dyson series can be summed up.

First we deal with the case of a time independent H'. For this special case the time translation operator in the Schrödinger representation $U(t,t_0)$ is already computed in Eq.(13.6). The time translation operator in the interaction representation follows from Eq.(13.19) and can now be written in the form

$$U_I(t,t_0) = T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} H_I(t')dt'\right) = e^{itH_0/\hbar} U(t,t_0)\, e^{-it_0H_0/\hbar}. \tag{13.35}$$

This allows us to rewrite the relation above as

$$\begin{aligned} &T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\; e^{it'H_0/\hbar} H' e^{-it'H_0/\hbar}\right) \\ &= e^{itH_0/\hbar} \exp\left(-\frac{i}{\hbar}(t-t_0)\left(H_0+H'\right)\right) e^{-it_0H_0/\hbar}, \;\text{ iff }\; \partial_t H' = 0. \end{aligned} \tag{13.36}$$

Next we turn to the highly interesting less trivial case in Eq.(13.34) which has several very important applications discussed in the next section. For this case it is useful to define a third picture which we will call the "stationary picture". We define the stationary picture states $|\psi,t\rangle_s$ as

$$|\psi,t\rangle_s = e^{-itH_1/\hbar}|\psi,t\rangle^+ = e^{-itH_1/\hbar}e^{itH_0/\hbar}|\psi,t\rangle_I \tag{13.37}$$

in analogy to the interaction picture or the Heisenberg picture. This form is inspired by the fact that the original Schrödinger wavefunction $|\psi,t\rangle^+$ satisfies the equation

$$i\hbar\partial_t|\psi,t\rangle^+ = \left(e^{itH_1/\hbar}H_2e^{-itH_1/\hbar}\right)|\psi,t\rangle^+, \tag{13.38}$$

Using this equation to compute the derivative of $|\psi,t\rangle_s$ one finds the simple equation

$$i\hbar\partial_t|\psi,t\rangle_s = (H_1+H_2)\,|\psi,t\rangle_s \tag{13.39}$$

which can be integrated

$$|\psi,t\rangle_s = e^{-i(H_1+H_2)(t-t_0)/\hbar}|\psi,t_0\rangle_s \tag{13.40}$$

From this we obtain the solution in the Schrödinger picture

$$|\psi,t\rangle^+ = e^{itH_1/\hbar}e^{-i(H_1+H_2)(t-t_0)/\hbar}e^{-it_0H_1/\hbar}|\psi,t_0\rangle \tag{13.41}$$

This gives the time translation operators in the Schrödinger picture and the interaction picture as follows

$$U(t,t_0) = e^{itH_1/\hbar}e^{-i(H_1+H_2)(t-t_0)/\hbar}e^{-it_0H_1/\hbar} \tag{13.42}$$

$$U_I(t,t_0) = e^{itH_0/\hbar}e^{itH_1/\hbar}e^{-i(H_1+H_2)(t-t_0)/\hbar}e^{-it_0H_1/\hbar}e^{-it_0H_0/\hbar} \tag{13.43}$$

The combination of the time translations $e^{itH_0/\hbar}e^{itH_1/\hbar}$ is equivalent to a single time translation $V(t)$ given by

$$V(t) = e^{itH_0/\hbar}e^{itH_1/\hbar}. \tag{13.44}$$

$V(t) = e^{iQ(t)}$ through the Baker-Campbell-Housedorff theorem discussed in Eq.(close1),

$$Q(t) = t(H_0+H_1) + \frac{t^2}{2}[H_0,H_1] + \frac{t^3}{12}[H_0,[H_0,H_1]] + \frac{t^3}{12}[H_1,[H_1,H_0]] + \cdots. \tag{13.45}$$

In some cases the series terminates, while in some other cases it is possible to use properties of Lie groups to compute the product of the exponentials directly and obtain $Q(t)$ fully. In particular, this shows how to sum up the Dyson series $T\exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt' H_I'(t')\right) = U_I(t,t_0)$ when $H_I'(t)$ has the form $H_I'(t) = e^{itH_0/\hbar}\left(e^{itH_1/\hbar}H_2e^{-itH_1/\hbar} - H_0\right)e^{-itH_0/\hbar}$.

A more general property of the time ordered product is

$$T\exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt'\ A(t')\right) = V(t)\,T\exp\left(-\frac{i}{\hbar}\int_{t_0}^t dt'\ \tilde{A}(t')\right)V^{-1}(t_0) \tag{13.46}$$

where

$$\tilde{A}(t') = V^{-1}(t') A(t') V(t') - V^{-1}(t') i\hbar\partial_{t'} V(t'), \tag{13.47}$$

This relation is satisfied for any $A(t)$ and any $V(t)$ that has a well defined inverse. In particular, if A is hermitian and V is unitary then $\tilde{A}$ is also hermitian. One can verify Eq.(13.46) by noting that it is correct at $t = t_0$, and that the time derivatives ∂_t of both sides are equal. The rule for taking the time derivative of a time ordered exponential is given in Eq.(13.24). This relation can be used to make transformations to arbitrary pictures. Thus, if $A(t')$ is the Hamiltonian in the interaction picture, and one makes a transformation from the interaction picture to another pictute by applying the operator $V(t)$, then the Hamiltonian in the new picture is $\tilde{A}(t')$. We see that the transformations discussed above $V(t) = e^{itH_0/\hbar}$ between the interaction and Schrödinger pictures is a special case of this more general transformation with a general $V(t)$. The transformation $V(t)$ is analogous to a local gauge transformation is a gauge theory[1].

Observe that if one can find a $V(t')$ that gives a constant $\tilde{A}$ for some given $A(t')$, then the integral on the right hand side can be performed $T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\ \tilde{A}\right) = e^{-i(t-t_0)\tilde{A}/\hbar}$, and hence the Dyson series can be summed up explicitly

$$T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\ A(t')\right) = V(t)\exp\left(-\frac{i}{\hbar}(t-t_0)\tilde{A}\right)V^{-1}(t_0). \tag{13.48}$$

Turning this observation around we can show that any $A(t)$ that has this property must be of the form

$$A(t) = V(t)\tilde{A}V^{-1}(t) - V(t) i\hbar\partial_t V^{-1}(t) \tag{13.49}$$

where $\tilde{A}$ is time independent and $V(t)$ is arbitrary.

The special forms of the Hamiltonians discussed above for which the Dyson series was summed up are special applications of this observation since they correspond to some specific form of $V(t)$ and $A(t)$. We can now state that the most general Hamiltonian for which we can sum up the Dyson series is of the form $H = H_0 + H'(t)$ with

$$H_0 + H'(t) = V(t)\tilde{H}V^{-1}(t) - V(t) i\hbar\partial_t V^{-1}(t). \tag{13.50}$$

where $\tilde{H}$ is any time independent operator and $V(t)$ is any unitary operator with any time dependence. This can be verified by retracing the discussion above backwards, and is left as an exercise for the student.

13.4 Two level system

Consider a physical system that, under certain circumstances and for certain phenomena, can be approximated by two quantum levels. It turns out that

[1] This relation can be recognized in the context of gauge theories as a local gauge transformation U(t) of a gauge potential $A(t)$ and its effect on the corresponding Wilson line integral constructed from $A(t)$.

this kind of approximation is indeed valid for several important phenomena in Nature, including spin or nuclear magnetic resonance, masers, and the $\mathrm{K}^0-\bar{\mathrm{K}}^0$ elementary particle system.

In the stable configuration the system is described by the two level Hamiltonian H_0 given by

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} = E_0 \times 1 + \hbar\omega_{12}\frac{\sigma_3}{2} \tag{13.51}$$

where we parametrized $E_0 = \frac{1}{2}(E_1 + E_2)$ and $\omega_{12} = \frac{1}{\hbar}(E_1 - E_2)$ and introduced the Pauli matrix σ_3. Now consider disturbing the system by a time dependent perturbation of the form

$$H'(t) = \begin{pmatrix} 0 & \hbar\gamma e^{-i\omega t} \\ \hbar\gamma^* e^{i\omega t} & 0 \end{pmatrix} \tag{13.52}$$

In a physical setting this could represent the turning on of an external electromagnetic field. Note that $H_0 + H'(t)$ is of the form of Eq.(13.34) which we studied above. Explicitly we have

$$H_0 + H'(t) = \begin{pmatrix} e^{-i\omega t/2} & 0 \\ 0 & e^{i\omega t/2} \end{pmatrix} \begin{pmatrix} E_1 & \hbar\gamma \\ \hbar\gamma^* & E_2 \end{pmatrix} \begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2} \end{pmatrix} \tag{13.53}$$

which identifies

$$H_1 = \begin{pmatrix} -\frac{\hbar\omega}{2} & 0 \\ 0 & \frac{\hbar\omega}{2} \end{pmatrix} = -\hbar\omega\frac{\sigma_3}{2}, \tag{13.54}$$

$$H_2 = \begin{pmatrix} E_1 & \hbar\gamma \\ \hbar\gamma^* & E_2 \end{pmatrix} = E_0 \times 1 + \hbar\omega_{12}\frac{\sigma_3}{2} + \hbar\gamma_1\sigma_1 + \hbar\gamma_2\sigma_2 \tag{13.55}$$

where $\gamma = \gamma_1 - i\gamma_2$. Therefore we expect to solve this problem exactly by using the methods of the previous section.

The interaction picture Hamiltonian $H_I(t) = e^{itH_0/\hbar}H'(t)e^{-itH_0/\hbar}$ becomes

$$H_I(t) = \begin{pmatrix} 0 & \gamma e^{-i(\omega-\omega_{12})t} \\ \gamma^* e^{i(\omega-\omega_{12})t} & 0 \end{pmatrix}. \tag{13.56}$$

The Schrödinger equation in the interaction picture $i\hbar\partial_t|\psi,t\rangle_I = H_I(t)|\psi,t\rangle_I$ is now a matrix differential equation with two unknown functions ${}_I\langle\psi,t| = (\psi_1^*(t), \psi_2^*(t))$. This can be solved by coupled differential equation methods. But we can also apply our general solution of the previous section to this matrix problem, and give directly the solution in the form of Eq.(13.43) with $e^{itH_0/\hbar}e^{itH_1/\hbar} = e^{iQ(t)}$. In the present case the Baker-Campbell-Housedorff formula Eq.(13.45) is simple because $[H_0, H_1] = 0$ and we have $Q(t) = (H_0 + H_1)t$. The solution $|\psi,t\rangle_I = e^{it(H_0+H_1)/\hbar}e^{-i(t-t_0)(H_1+H_2)/\hbar}e^{-it_0(H_0+H_1)/\hbar}|\psi\rangle$ can be written in matrix form. The part of H_0 and H_2 proportional to E_0 cancels out since it is proportional to the identity, and we obtain

$$|\psi,t\rangle_I = e^{it(\omega_{12}-\omega)\frac{\sigma_3}{2}}e^{-i(t-t_0)\left((\omega_{12}-\omega)\frac{\sigma_3}{2}+\gamma_1\sigma_1+\gamma_2\sigma_2\right)}e^{-it_0(\omega_{12}-\omega)\frac{\sigma_3}{2}}|\psi\rangle \tag{13.57}$$

where $|\psi\rangle$ is a general initial state. The exponential of the Pauli matrices is computed with the methods we discussed before when we studied rotations $\exp(i\mathbf{v}\cdot\boldsymbol{\sigma}) = \cos v + i\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}\sin v$. In this case the components of the vector $\mathbf{v}$ are $v_1 = -(t-t_0)\gamma_1$, $v_2 = -(t-t_0)\gamma_2$ and $v_3 = -(t-t_0)(\omega_{12}-\omega)/2$, and the length of the vector is

$$v = \Omega(t-t_0),\ \Omega \equiv \left(|\gamma|^2 + \frac{1}{4}(\omega_{12}-\omega)^2\right)^{1/2}, \tag{13.58}$$

while the quantity $\hat{\mathbf{v}}\cdot\boldsymbol{\sigma} = \boldsymbol{\sigma}\cdot\mathbf{v}/v$ is

$$\hat{\mathbf{v}}\cdot\boldsymbol{\sigma} = \frac{1}{\Omega}\left((\omega_{12}-\omega)\frac{\sigma_3}{2} + \gamma_1\sigma_1 + \gamma_2\sigma_2\right). \tag{13.59}$$

We can now compute transition amplitudes. If the initial state is $|i\rangle$ is the first energy eigenstate $|i\rangle = |E_1\rangle$ and the final state is the second energy eigenstate $|f\rangle = |E_2\rangle$ then the transition amplitude is

$$A_{21}(t) = \langle E_2| e^{it(\omega_{12}-\omega)\frac{\sigma_3}{2}} e^{-i(t-t_0)\left((\omega_{12}-\omega)\frac{\sigma_3}{2}+\gamma_1\sigma_1+\gamma_2\sigma_2\right)} e^{-it_0(\omega_{12}-\omega)\frac{\sigma_3}{2}} |E_1\rangle \tag{13.60}$$

$$= e^{-\frac{i}{2}t(\omega_{12}-\omega)}\left(e^{-i(t-t_0)\left((\omega_{12}-\omega)\frac{\sigma_3}{2}+\gamma_1\sigma_1+\gamma_2\sigma_2\right)}\right)_{21} e^{-\frac{i}{2}t_0(\omega_{12}-\omega)} \tag{13.61}$$

$$= e^{-\frac{i}{2}t(\omega_{12}-\omega)}\left(\cos v + i\hat{\mathbf{v}}\cdot\boldsymbol{\sigma}\sin v\right)_{21} e^{-\frac{i}{2}t_0(\omega_{12}-\omega)} \tag{13.62}$$

$$= i\frac{\gamma}{\Omega}\sin(\Omega(t-t_0))\ e^{-\frac{i}{2}(t+t_0)(\omega_{12}-\omega)} \tag{13.63}$$

By squaring it we obtain the probability for making this transition

$$|A_{21}(t)|^2 = \frac{|\gamma|^2}{\Omega^2}\sin^2(\Omega(t-t_0)). \tag{13.64}$$

Similarly we can compute the probablity for staying in the same state, which is $|A_{11}(t)|^2$. Since the probabilty for all transitions must sum up to 1, we must have $|A_{11}(t)|^2 + |A_{21}(t)|^2 = 1$. Indeed we can prove generally that the sum of all probabilities for transitions to all states starting from some initial state must sum up to 1 as follows

$$\sum_n |A_{ni}|^2 = \sum_n A^*_{ni}A_{ni} = \sum_n \langle i|\left(U_I(t,t_0)\right)^\dagger |n\rangle\langle n|U_I(t,t_0)|i\rangle = \langle i|i\rangle = 1 \tag{13.65}$$

Note that the sum over n includes the initial state i. Therefore, the probability for staying in the same state $|A_{11}(t)|^2$ is

$$|A_{11}(t)|^2 = 1 - \frac{|\gamma|^2}{\Omega^2}\sin^2(\Omega(t-t_0)). \tag{13.66}$$

If we have a more general initial and final states, they are generally normalized states of the form $|i\rangle = \cos\theta_i|E_1\rangle + \sin\theta_i|E_2\rangle$ and $|f\rangle = -\sin\theta_f|E_1\rangle + \cos\theta_f|E_2\rangle$. The transition amplitude is then

$$A_{fi}(t) = -\sin\theta_f\cos\theta_i\ A_{11}(t) + \cos\theta_f\cos\theta_i\ A_{21}(t) - \sin\theta_f\sin\theta_i A_{12}(t) + \cos\theta_f\sin\theta_i\ A_{22}(t).$$

Let us plot the probabilities $|A_{21}(t)|^2$ and $|A_{11}(t)|^2$ as a function of time as in Fig.13.1.

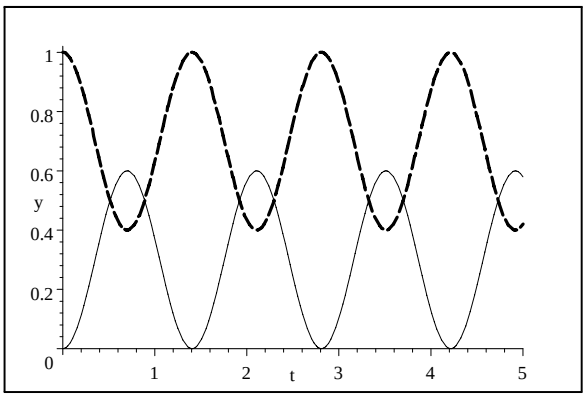


Fig. 13.1 - $|A_{11}(t)|^2$ dash, $|A_{11}(t)|^2$ solid

We see that the system keeps making transitions between the two states. At time $t = t_0 = 0$ the system starts out 100% in state 1 as seen from the figure (dashed curve at 1, solid curve at 0). As time goes on the state 1 gets depleted (decreasing dashed curve) while state 2 gets populated (increasing solid curve). But the state 2 does not get populated 100%; instead it reaches a maximum amplitude given by $\frac{|\gamma|^2}{\Omega^2} < 1$. At this point it begins to lose its population while state 1 begins to absorb it. This process of emission and absorbtion oscillates over time and goes on indefinitely as shown in the figure. The maximum transition probability is $\frac{|\gamma|^2}{\Omega^2}$ which is reached periodically at certain times. This quantity is given by

$$\frac{|\gamma|^2}{\Omega^2} = \frac{|\gamma|^2}{|\gamma|^2 + \frac{1}{4}(\omega_{12} - \omega)^2} \tag{13.67}$$

In Fig.13.2 we plot it as a function of the frequency of the external field ω

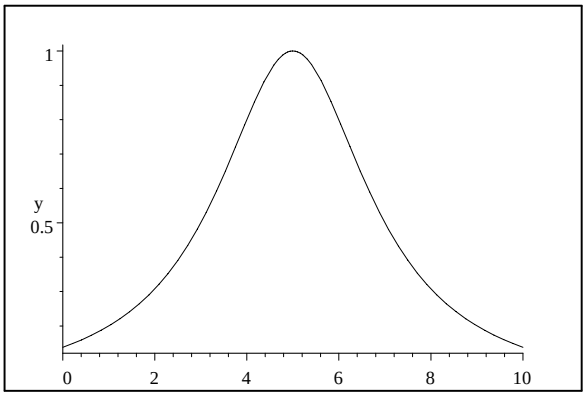


Fig.13.2 - Amplitude peaks at $\omega = \omega_{12}$.

It a bell-shaped curve that peaks at $\omega = \omega_{12}$, at which point $\frac{|\gamma|^2}{\Omega^2} = 1$. Thus, when the external field is adjusted to the critical frequency ω_{12} then state 2 can get populated 100% and the system can make a full transition from state 1 to state 2. This frequency is called the resonance frequency since the effect is the maximum possible. At the resonance frequency both curves at Fig.13.1 have the maximum amplitude and oscillate as a function of time between 0 and 1.

One can perform experiments in which the system is observed and the curves in Figures 13.1 and 13.2 are plotted. From the location of the peak of the curve one reads off $\omega = \omega_{12}$ and from this learns the energy difference between the two levels

$$E_1 - E_2 = \hbar\omega_{12}. \tag{13.68}$$

Then one measures the width of the curve when the amplitude is half of its maximum $\frac{|\gamma|^2}{\Omega^2} = \frac{1}{2}$. Lets call the measured width q. The half width is given by a value of the frquency ω_c that satisfies $\omega_c - \omega_{12} = q/2$. Inserting this in the relation $\frac{|\gamma|^2}{\Omega^2} = \frac{1}{2}$ we can extract the quantity $|\gamma|$

$$\frac{|\gamma|^2}{|\gamma|^2 + \frac{1}{4}\left(\frac{q}{2}\right)^2} = \frac{1}{2} \;\rightarrow\; |\gamma| = \frac{q}{4}. \tag{13.69}$$

This value of $|\gamma|$ is a measure of the strength of the interaction between the external field and the system. This measurement is used to extracts some relevant property of the system as we will see in the next section.

13.4.1 Spin magnetic resonance

An example of the two-level problem discussed before is given by the spin magnetic resonance of a set of nuclei at rest. In fact, if we apply to such system a constant magnetic field $\vec{B}_0$ in the z direction $\mathbf{B}_0 = B_0\,\hat{z}$, and a "small" oscillating field in the $x-y$ plane, $B_x(t) = b_0\cos\omega t$ and $B_y(t) = b_0\sin\omega t$. We can write the Hamiltonian

$$H(t) = -\boldsymbol{\mu}\cdot\mathbf{B}(t). \tag{13.70}$$

The magnetic moment is proportinal to the spin operator $\boldsymbol{\mu} = \mu\mathbf{S}$ of the nucleus , then

$$H(t) = -\mu(B_0 S_z + b_0\cos\omega t\; S_x + b_0\sin\omega t\; S_y) \tag{13.71}$$

or

$$H = H_0 + H'(t) \tag{13.72}$$

with

$$H_0 = -\mu B_0 S_z, \quad H'(t) = -\mu\left(b_0\cos\omega t\; S_x + b_0\sin\omega t\; S_y\right) \tag{13.73}$$

We can write

$$H'(t) = e^{-i\omega t S_z/\hbar}\left(-\mu b_0 S_x\right)e^{i\omega t S_z/\hbar} \tag{13.74}$$

Therefore, this is of the form $H'(t) = e^{itH_1/\hbar}H_2 e^{-itH_1/\hbar}$ which we can solve exactly, with $H_1 = \omega S_z$ and $H_2 = -\mu b_0 S_x$.

The states of H_0 are angular momentum states $|jm\rangle$ for a fixed j, and has eigenvalues $E_m = -\mu B_0 m$. The transitions will be between different rotation states and can be computed for any j in terms of the D-functions for rotation matrices. The general case is left as an exercise for the student. Here we discuss the case of spin $j = 1/2$. For that case the spin operator is represented by Pauli matrices $\mathbf{S} = \hbar\boldsymbol{\sigma}/2$. Therefore the Hamiltonian takes the 2×2 matrix form

$$H = H_0 + H'(t) = \begin{pmatrix} -\frac{1}{2}\hbar\mu B_0 & 0 \\ 0 & \frac{1}{2}\hbar\mu B_0 \end{pmatrix} + \begin{pmatrix} 0 & \hbar\gamma e^{-i\omega t} \\ \hbar\gamma^* e^{+i\omega t} & 0 \end{pmatrix}, \quad (13.75)$$

with

$$|\gamma| = \frac{1}{2}\mu b_0. \quad (13.76)$$

This is precisely the problem we studied in the previous section. By making the measurments of the resonance frequency ω and of the width of the curve q indicated in the previous section we learn

$$\mu B_0 = \omega, \quad \frac{1}{2}\mu b_0 = \frac{q}{4}. \quad (13.77)$$

Since an experimentalist has control of the external fields B_0, b_0 and can vary them, these measurements allow us to extract the value of the magnetic moment.

The setup can also be used in reverse to make instruments for which μ is known, and use it to measure electromagnetic properties of systems.

13.4.2 Maser

An Ammonia molecule (NH_3) has two opposite parity eigenstates $|A\rangle$ and $|S\rangle$. The energy of the state $|A\rangle$ is slightly higher than the energy of the state $|S\rangle$. Suppose now we apply an electric field in a given direction oscillating at a frequency

$$\omega \simeq (E_A - E_S)/\hbar \quad . \quad (13.78)$$

In this case we reproduce the same situation as we have discussed before, where the transition probability $|A_{21}(t)|^2$ has a maximum peak of value 1. Suppose now we construct the following apparatus.

The molecular beam of NH_3 containing both $|A\rangle$ and $|S\rangle$ states goes through a filter. The emerging beam now contains only the $|A\rangle$ state. Then the beam goes through a microwave cavity that has an oscillating electric field. The excited $|A\rangle$ state now makes transitions to the lower $|S\rangle$ state and the beam oscillates between the two states as it continues its journey through the cavity. The amount of time it takes for one cycle of full transition from $|A\rangle$ to $|S\rangle$ at

resonance frequency, when $\Omega = |\gamma|$, is $\Delta t = \frac{\pi}{2|\gamma|}$. If one now chooses the length of the cavity L in such a way that the outgoing beam of NH_3 is completely in the $|S\rangle$ state just when it exits the cavity, then one expects that the energy inside the cavity would rise, as it happens in practice. This excess energy is given up to the time dependent potential inside the cavity, so that the radiation field grows in magnitude. This is the MASER = Microwave Amplification by Stimulated Emission of Radiation.

13.5 Time independent interaction.

Consider once again the probability amplitude $A_{fi}(t) = \langle ft|it\rangle^+$ and recall the integral equation 13.23

$$|it\rangle^+ = |it\rangle - i/\hbar \int_{t_0}^{t} dt' e^{-iH_0/\hbar(t-t')} H'|it'\rangle^+ \tag{13.79}$$

We recall that the above expression for $|it\rangle^+$ is a formal solution for a state at any time t. So now

$$\langle ft|it\rangle^+ = \langle ft|it\rangle - i/\hbar \int_{t_0}^{t} dt' e^{-iE_f/\hbar(t-t')} \langle ft'|H'(t')|it'\rangle^+ \tag{13.80}$$

We see the structure

$$\langle ft|it\rangle^+ = \delta_{fi} - \frac{i}{\hbar}\tilde{T}_{fi} \tag{13.81}$$

where we have defined

$$\tilde{T}_{fi} = \int_{t_0}^{t} dt' \langle ft'|H'(t')|it'\rangle^+ = \int_{t_0}^{t} dt' e^{+i/\hbar E_f(t'-t_0)} \langle ft_0|H'(t')|it'\rangle^+ \tag{13.82}$$

Suppose now H' does not depend on time and that we are looking for a perturbative approximation by inserting $|it\rangle^+ = |it\rangle$ on the right hand side, then

$$\begin{aligned} \tilde{T}_{fi} &= \textstyle\int_{t_0}^{t} dt' e^{i/\hbar E_f(t'-t_0)} \langle ft_0|H' e^{-i/\hbar E_i(t'-t_0)}|it_0\rangle \\ &= \textstyle\int_{t_0}^{t} dt' \langle ft_0|H'|it_0\rangle e^{-i/\hbar(E_i-E_f)(t'-t_0)} \\ &= \langle ft_0|H'|it_0\rangle 2\pi\delta\left(\left(E_i - E_f\right)/\hbar\right) \end{aligned} \tag{13.83}$$

where in the last step we took the limit $t_0 \to -\infty$ and $t \to \infty$. The result shows that when the Hamiltonian does not depend on the time explicitly the energy is conserved. The integral equation can be solved by continuing the iteration and one finds that at each order of the perturbative expansion the delta function appears. We can thus write generally the following form for the transition amplitude when H' is time independent

$$A_{fi}(t) = \langle ft|it\rangle^+ = \delta_{fi} - i2\pi\delta(E_f - Ei)T_{fi} \tag{13.84}$$

Because of energy conservation we are now forced to choose eigenstates $|it\rangle^+$ and $|it\rangle$ such that $H|i,t\rangle^+ = E_i|i,t\rangle^+$, and $H_0|i,t\rangle = E_i|i,t\rangle$. This matching of

the energies is easily satisfied when both H_0, H have continuous energy eigenvalues. Because of this, the time dependence of these states can be made explicit $|i,t\rangle^+ = e^{-i/\hbar E_i t}|E_i\rangle^+$, $|i,t\rangle = e^{-i/\hbar E_i t}|E_i\rangle$. Then the integral equation becomes

$$|E_i\rangle^+ = |E_i\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt' e^{-i/\hbar(H_0-E_i)(t-t')}H'|E_i\rangle^+ \tag{13.85}$$

Similarly the transition amplitude becomes

$$A_{fi} = \langle E_f|E_i\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt' e^{i/\hbar(E_f-E_i)t'}\langle E_f|H'|E_i\rangle^+ \tag{13.86}$$

Taking the limits $t_0 \to -\infty$ and $t \to \infty$ we recover the delta function. A_{fi} in this limit is called the S-matrix, so we can write the form

$$S_{fi} = \delta_{fi} - 2\pi i\delta\left(E_f - E_i\right)T_{fi} \tag{13.87}$$

where we identify

$$T_{fi} \ = \langle E_f|H'|E_i\rangle^+. \tag{13.88}$$

This is the transition amplitude when $f \neq i$ which we would like to compute. In the limits considered it is called the T-matrix.

We would now like to take the limit $t_0 \to \infty$ more carefully. Before we do this we must first give a prescription; namely, first replace E_i with $E_i + i\varepsilon$ (ε infinitesimal but positive) and then set $t_0 = -\infty$. Since H' is time-independent, we can readily perform the integration on t', to obtain

$$|E_i\rangle^+ = |E_i\rangle + \left(\frac{1 - e^{+i/\hbar(E_i+i\varepsilon-H_o)(t-t_0)}}{E_i - H_0 + i\varepsilon}\right)H'|E_i\rangle^+ \tag{13.89}$$

We see that the "$i\varepsilon$-presciption" is related to the boundary condition that $|E_i\rangle^0$ is an eigenstate in the remote past. It also represents a way to define a green's function for the operator $(E_i - H_0)$. If we now take the limit $t_0 \to -\infty$, we have

$$|E_i\rangle^+ = |E_i\rangle + \frac{1}{(E_i - H_0 + i\varepsilon)}H'|E_i\rangle^+ \tag{13.90}$$

This equation is called *Lipmann-Schwinger* equation. We could arrive at the same equation by a time independent approach as follows

$$H|E_i\rangle^+ = E_i|E_i\rangle^+, \ \to \ (H_0 + H')|E_i\rangle^+ = E_i|E_i\rangle^+, \ \to \ (E_i - H_0)|E_i\rangle^+ = H'|E_i\rangle^+ \tag{13.91}$$

The solution of the last relation is exactly given by the Lippmann-Schwinger equation. In fact, we can verify that the Lippmann-Schwinger equation satisfies the last equation

$$(E_i - H_0)|E_i\rangle^+ = (E_i - H_0)|E_i\rangle^0 + \frac{E_i - H_0}{E_i - H_0 + i\varepsilon}H'|E_i\rangle^+ \underset{\varepsilon\to 0}{=} 0 + H'|E_i\rangle^+ \tag{13.92}$$

We see that in the time independent derivation the inverse of $(E_i - H_0)$ needs to be given some meaning. As we saw above the $i\varepsilon$ prescription derived through the time dependent formalism provides the meaning and it implies causality.

The Lippmann Schwinger equation can be solved formally for $|E_i\rangle^+$ in terms of $|E_i\rangle$ as follows

$$|E_i\rangle^+ = \left[1 - \frac{1}{(E_i - H_0 + i\varepsilon)} H'\right]^{-1} |E_i\rangle, \tag{13.93}$$

By inserting this in the expression above we obtain the following expression for the transition amplitude

$$T_{fi} = \langle E_f | H' | E_i \rangle^+ = \langle E_f | T(E_i) | E_i \rangle \tag{13.94}$$

$$T(E_i) = H' \left[1 - (E_i - H_0 + i\varepsilon)^{-1} H'\right]^{-1} \tag{13.95}$$

$$= \left[1 - H'(E_i - H_0 + i\varepsilon)^{-1}\right]^{-1} H'. \tag{13.96}$$

If we expand $T(E_i)$ in powers as a geometric series and insert the identity in the form of intermediate states we obtain the series form

$$T_{fi} = \sum_{n=1}^{\infty} \sum_{k_1 \cdots k_{n-1}} (-1)^{n-1} \frac{(H')_{fk_1} (H')_{k_1 k_2} \cdots (H')_{k_{n-1} i}}{(E_f - E_{k_1} + i\varepsilon)(E_f - E_{k_2} + i\varepsilon) \cdots (E_{f_{n-1}} - E_{k_{n-1}} + i\varepsilon)} \tag{13.97}$$

We will return to study this expression in more detail.

13.6 Cross section for scattering

We will now compute the cross section of a scattering process which is proportional to $|S_{fi}|^2$. In fact, writing

$$S_{fi} = \delta_{fi} - i2\pi\delta(E_f - E_i) T_{fi} \tag{13.98}$$

for $f \neq i \quad |S_{fi}|^2 = 2\pi\delta(0) 2\pi\delta(E_f - E_i) |T_{fi}|^2$ writing

$$\int_{-\infty}^{\infty} dt' e^{-i/\hbar (E_f - E_i) t'} = 2\pi\hbar\delta(E_f - E_i) \tag{13.99}$$

we see that

$$2\pi\hbar\delta(0) = \int_{-\infty}^{\infty} dt' = \lim_{t \to \infty} t \tag{13.100}$$

Define now the transition probability *per unit time* by dividing out this infinite time factor. Indeed in an measurement an observation takes a certain amount time, not an infinite amount of time, and one may report the average probability per unit time by dividing by the length of time it takes for the observation. Therefore we obtain

$$\dot{P}_{fi} = \frac{\Delta |S_{fi}|^2}{\Delta t} = \frac{2\pi}{\hbar} \delta(E_f - E_i) |T_{fi}|^2. \tag{13.101}$$

Next we take into account the limitation of instruments that, rather than measuring the transition to a single energy eigenstate, they actually measure the transition to a collection of states in the same energy neighborhood as E_f within some resolution dE_f. So the transition rate must be multiplied by the number of states in the energy interval determined by the resolution of the instrument. We represent the number of states as $\rho(E_f)dE_f$ where $\rho(E_f)$ is the number density per unit energy interval, per unit volume (in space), in the vicinity of E_f. We will later see how to compute $\rho\,(E)$ in some examples. Thus, the measured probability per unit time of the transition from the initial state to a group of final states is dW_{fi}

$$dW_{fi} = \dot{P}_{fi}\ \rho(E_f)dE_f \tag{13.102}$$

for the energy range $E_f \pm \frac{1}{2}dE_f$. For the complete energy range, we have the integral

$$W_{fi} = \int \dot{P}_{fi}\ \rho(E_f)dE_f = \frac{2\pi}{\hbar}\int \delta(E_f - E_i)|T_{fi}|^2\rho(E_f)dE_f \tag{13.103}$$

$$= \frac{2\pi}{\hbar}|T_{fi}|^2\rho(E_f)\Big|_{E_f=E_i}. \tag{13.104}$$

13.6.1 Fermi's golden rule

If we evaluate $|T_{fi}|^2$ in the first approximation, then we use

$$T_{fi}^{(1)} = (H')_{fi} \tag{13.105}$$

$$W_{fi}^{(1)} = \frac{2\pi}{\hbar}|(H')_{fi}|^2\rho(E_f)\Big|_{E_f=E_i} \tag{13.106}$$

This is such a simple and useful formula that it has acquired the name "Fermi's Golden Rule".

13.6.2 First order, time dependent perturbation

$$A_{fi}(+) = \langle ft|it\rangle^+ = \langle f,t|U(t,t_0|i,t_0\rangle_I \tag{13.107}$$

For $U_I(t,t_0) = 1 - (i/\hbar)\int_{t_0}^{t} H_I(t')dt' + O(H_I'^2)$, we have

$$\begin{aligned} A_{fi}(t) &= \delta_{fi} - i/\hbar\int_{t_0}^{t} dt'\ \langle f|H_I|i\rangle \\ &= \delta_{fi} - i/\hbar\int_{t_0}^{t} dt'\ \langle ft_0|e^{-iH_0t_0/\hbar}e^{+iH_0t\prime/\hbar}H'e^{-iH_0t'/\hbar}e^{+iH_0t_0/\hbar}|i,t_0\rangle^+ \\ &= \delta_{fi} - i/\hbar\int_{t_0}^{t} dt' e^{i/\hbar(E_f-E_i)(t'-t_0)}\langle f,t_0|H'|i,t_0\rangle \end{aligned} \tag{13.108}$$

Consider now the perturbation due to a radiation field (photons)

$$H'(t) = Ae^{i\omega t} + A^\dagger e^{-i\omega t}. \tag{13.109}$$

where A is some operator. Then

$$A_{fi}(t) = \delta_{fi} - \frac{i}{\hbar} e^{i(E_f - E_i)t_0/\hbar} \left[\begin{array}{c} \langle f, t_0|A|i, t_0\rangle \int_{t_0}^{t} e^{i[(E_f + E_i) + \hbar\omega]t'/\hbar} dt' \\ + \langle f, t_0|A^\dagger|i, t_0\rangle \int_{t_0}^{t} e^{i[(E_f - E_i) - \hbar\omega]t'/\hbar} dt' \end{array} \right] \tag{13.110}$$

If we send $t_0 \to -\infty$ and $t \to \infty$, we obtain delta functions $\delta(E_f - E_i \pm \hbar\omega)$, which indicate energy emission or absorbtion, and certainly not energy conservation. This is called stimulated emission or absorbtion under the influence of an external radiation field. Evidently one of the terms must vanish if the other does not. In computing the probability $|A_{fi}|^2$ we apply the same reasoning as the previous section, and divide out the by the factor $2\pi\hbar\delta(0)$. Then for $f \neq i$, we obtain the transition probability per unit time in the form

$$d|A_{fi}(t)|^2/dt = \frac{2\pi}{\hbar} \left\{ \begin{array}{c} \delta(E_f - E_i + \hbar\omega)|\langle f|A|i\rangle|^2 \\ + \delta(E_f - E_i - \hbar\omega)|\langle f|A^\dagger|i\rangle|^2 \end{array} \right\} \tag{13.111}$$

This must again be multiplied by the density of final states as discussed in the previous section. A transition is possible only if $E_f - E_i = \pm\hbar\omega$.

13.6.3 Ionization of a Hydrogen atom by a radiation field

Consider the H-atom in a time dependent oscillating electric field which is spacially constant. The scalar potential is $A_0 = -\mathbf{r} \cdot \mathbf{E}(t)$ and the electron interacts with the electric field via the term $-eA_0$ in the Hamiltonian. Therefore, for an electric fiels in the z direction the perturbation is

$$H'(t) = e\mathbf{r} \cdot \mathbf{E}(t) = eE(t)\ \mathbf{r} \cdot \hat{\mathbf{E}}. \tag{13.112}$$

We will take $E(t) = E_0\left(e^{i\omega t} + e^{-i\omega t}\right)$, and choose later a convenient direction $\hat{\mathbf{E}}$ for the electric field. Then

$$H'(t) = Ae^{i\omega t} + A^\dagger e^{-i\omega t},\ A = eE_0\mathbf{r} \cdot \hat{\mathbf{E}} \tag{13.113}$$

Following the computation in the previous section we have

$$\frac{d}{dt}|A_{fi}(t)|^2 = \frac{2\pi}{\hbar} \left\{ \begin{array}{c} \delta(E_f - E_i + \hbar\omega)|\langle f|A|i\rangle|^2 \\ + \delta(E_f - E_i - \hbar\omega)|\langle f|A^\dagger|i\rangle|^2 \end{array} \right\}. \tag{13.114}$$

For the ionization of Hydrogen, we take the initial state to be the ground state, while the final state is a free particle state with mometum $\mathbf{p}$. Therefore we need to compute

$$\langle f|A|i\rangle = eE_0\langle \mathbf{p}|\mathbf{r} \cdot \hat{\mathbf{E}}|100\rangle\mathbf{r}' \tag{13.115}$$

$$= eE_0 \int d^3\mathbf{r}'\ \langle \mathbf{p}|\mathbf{r}'\rangle\ r'\hat{\mathbf{r}}' \cdot \hat{\mathbf{E}}\ \langle \mathbf{r}'|100\rangle \tag{13.116}$$

$$= eE_0 \int d^3r\ \frac{1}{(L)^{3/2}} \exp\left(-\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}'\right)\ r'\ \hat{\mathbf{r}}' \cdot \hat{\mathbf{E}}\ \frac{e^{-r'/a_0}}{\sqrt{\pi a_0^3}} \tag{13.117}$$

$$= \frac{eE_0}{\sqrt{\pi}(La_0)^{3/2}} \int d^3\mathbf{r}'\ \exp\left(-\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}'\right)\ r'\ \hat{\mathbf{r}}' \cdot \hat{\mathbf{E}}\ e^{-r'/a_0} \tag{13.118}$$

We have used the plane wave expression $\langle \mathbf{p}|\mathbf{r}'\rangle = \frac{1}{(L)^{3/2}}\exp\left(-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}'\right)$ for a particle in a box of size L. It is now convenient to choose $\mathbf{p}$ in the z direction, and let $\hat{\mathbf{E}}$ be in the z-x plane, with an angle θ relative to the z plane. The vector $\mathbf{r}'$ points in a general direction with angles $\left(\theta', \phi'\right)$ and makes the angle θ'' with the direction $\hat{\mathbf{E}}$. Thus

$$\hat{\mathbf{r}}' = \left(\sin\theta'\cos\phi',\ \sin\theta'\sin\phi', \cos\theta'\right) \tag{13.119}$$

$$\hat{\mathbf{p}} = (0,0,1)\,,\quad \hat{\mathbf{E}} = (\sin\theta, 0, \cos\theta) \tag{13.120}$$

So we can derive the relation between the angles

$$\hat{\mathbf{E}}\cdot\hat{\mathbf{r}}' = \cos\theta'\cos\theta + \sin\theta\sin\theta'\cos\phi',\quad \hat{\mathbf{p}}\cdot\hat{\mathbf{r}}' = \cos\theta' \tag{13.121}$$

The integrals can now be performed

$$\langle f|A|i\rangle = \frac{32\pi e E_0\,(p/\hbar)\,a_0^5\cos\theta}{\sqrt{\pi L^3 a_0^3}\left(1+\frac{1}{\hbar^2}p^2a_0^2\right)^3}. \tag{13.122}$$

The transition probability per unit time to a group of final states with approximately the same energy E_f is given by

$$dW_{fi} = \int |\langle f|A|i\rangle|^2 \times \frac{2\pi}{\hbar}\delta(E_f - E_i + \hbar\omega) \times \rho(E_f)dE_f \times L^3 \tag{13.123}$$

where the number of states available to the emitted particle in the box is $\rho(E_f)dE_f \times L^3$. Thus, we also need to discuss the density of final states. For a free particle in a box the momentum is quantized $p_i = 2\pi\hbar n_i/L$. The number of states in the range $dp_1dp_2dp_3$ is $dn_1dn_2dn_3 = (L/2\pi\hbar)^3\,dp_1dp_2dp_3$. The density of states per unit volume is obtained by dividing by the volume $(dn_1dn_2dn_3)/L^3 = d^3p/(2\pi\hbar)^3$. This is identified with the number of states in some energy interval $\rho(E_k)dE_k$ where $\rho(E_k)$ is the density of states per unit energy. Therefore the density of states for the free particle emitted by the atom is given by

$$\rho(E_k)dE_k = \frac{d^3p}{(2\pi\hbar)^3}. \tag{13.124}$$

Writing $d^3p = d^2\Omega p^2dp$ and using $E = p^2/2m$ to convert $dE = \frac{1}{m}pdp$, we find

$$\rho(E)dE = \frac{1}{(2\pi\hbar)^3}d^2\Omega p^2dp = \frac{pmd\Omega}{(2\pi\hbar)^3}dE. \tag{13.125}$$

The result is

$$dW_{fi} = \frac{256e^2E_0^2a_0^7mp^3}{\pi\hbar^6(1+\frac{1}{\hbar^2}p^2a_0^2)^6}\cos^2\theta d^2\Omega. \tag{13.126}$$

From conservation of energy we have

$$p^2/2m = -me^4/2\hbar^2 + \hbar\omega \tag{13.127}$$

We solve for p and substitute to get a final expression in terms of the frequency of the external field ω.

13.7 Problems

1. Prove that for a symmetric function $f(t_1, \ldots, t_n)$, ordered integrals can be rewritten as integrations in the full range as follows.

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n f(t_1, t_2 \cdots t_n) = \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n f(t_1, t_2, \cdots t_n).$$

2. Perform the integrals in Eq.(13.32) for order n, and show that the series sums up to the expression in Eq.(13.7).

3. Prove that the Dyson series satisfies the relation

$$T \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} dt' \; A(t')\right) = V(t) \exp\left(-\frac{i}{\hbar}(t - t_0)\tilde{A}\right) V^{-1}(t_0)$$

with $\tilde{A}(t') = V^{-1}(t') A(t') V(t') - V^{-1}(t') i\hbar \partial_{t'} V(t')$, for any $A(t')$ and any $V(t)$. This can be shown that proving that both sides satisfy the same boundary condition and that their derivatives are equal.

4. Consider a time dependent Hamiltonian $H = H_0 + H'(t)$ of the form

$$H_0 + H'(t) = V(t) \tilde{H} V^{\dagger}(t) - V(t) i\hbar \partial_t V^{\dagger}(t).$$

where $\tilde{H}$ is any time independent operator and $V(t)$ is any unitary operator with any time dependence. Obtain the solution of the Schrödinger equation $|i, t\rangle^{+}$ for $t > t_0$ if the initial state at very early times is a state $|i\rangle$ in the Hilbert space owhose time dependence is governed by H_0.

5. Consider now the general spin magnetic resonance problem for arbitrary spin j. The Hamiltonian is $H = H_0 + H'$, with $H_0 = \tilde{H}_0 - \Omega J_0$ and $H' = -\lambda\Omega(J_1 \cos\omega t - J_2 \sin\omega t)$. The part $\tilde{H}_0$ has eigenvalues $E_0(j)$ that are rotationally symmetric. $\Omega = \mu B$ is the Larmor frequency. Suppose you have an initial state at $t = t_0$ which is an eigenstate of H_0 and labelled as $|jm\rangle$. The interaction $H'(t)$ is turned on at $t = t_0$. Find the probability amplitude for observing a final state of H_0, $|jm'\rangle$ at times $t > t_0$? You can solve this problem easily by using the methods discussed in the text for transforming to a new basis in which the Hamiltonian becomes time independent. Then you should be able to give an exact result in terms of rotation D-functions. Compare your result to the one obtained in the text when you specialize to $j = 1/2$.

6. Consider a system that is described by H_1 for $t < t_0$ and by H_2 for $t > t_0$. Both H_1 and H_2 are time independent. Let us label the respective Hilbert spaces with indices as $|*, t\rangle_1, |*, t\rangle_2$ during the appropriate time intervals.

   a) What are the time development operators before and after t_0 ? If we have an initial state labelled as $|i, t_1\rangle_1$ for $t_1 < t_0$ give its time dependence $|i, t\rangle^{+}$ for $t > t_0$ in terms of these time development operators.

b) What is the probability amplitude for making a transition from an initial eigenstate of H_1, $|E_1\rangle_1$, to a final eigenstate of H_2, $|E_2\rangle_2$? Give the time dependence of the transition amplidude, if any.

7. Consider a H-like atom whose nuclear charge undergoes a sudden change $Z \to Z+1$ (example via beta decay). If the atom is initially in its ground state, what is the probability that it will be excited to the 2s level ? Give an exact answer using the known Hydrogen-like atom eigenstates (with arbitrary nuclear charge). Hint: use the formalism of the previous problem.

8. Consider again the physical process of the previous problem. Formulate the problem as a small change of the potential and apply perturbation theory when Z is large. How does your result compare to the result of the previous problem?

9. Consider a Hamiltonian $H = H_0 + H'_S$, where H'_S is time independent. This was discussed in class. Now, you will complete some of the steps as an exercise, starting with the S-matrix the form

$$S_{fi} = \langle f|U_I(\infty,-\infty)|i\rangle, \;\; U_I(\infty,-\infty) = T\left[\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty} dt' H_I(t')\right)\right].$$

Perform all the time ordered integrals for each term of the series expansion by using the $i\varepsilon$ prescription, and show that you obtain

$$\begin{aligned} S_{fi} &= \delta_{fi} - i2\pi\delta(E_f - E_i)T_{fi} \\ T_{fi} &= \sum_{n=0}^{\infty}\langle f|H'_S\frac{1}{E_i - H_0 + i\varepsilon}H'_S \;\cdots\; \frac{1}{E_i - H_0 + i\varepsilon}H'_S|i\rangle \\ &= \langle f|\left[1 - H'_S\frac{1}{E_i - H_0 + i\varepsilon}\right]^{-1} H'_S|i\rangle \end{aligned}$$

in agreement with the methods presented in class.

10. Consider a l-dimensional harmonic oscillator described by

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

where the frequency $\omega(t)$ is time dependent

$$\omega(t) = \omega_0 + \omega_1\cos(bt).$$

and $\omega_1 \ll \omega_0$. Use perturbation theory to calculate the probability that a transition occurs from the ground state, as a function of time, given that the system is initially at the ground state. To which state(s) is the main transition ?

Chapter 14

SCATTERING THEORY

14.1 Lippmann-Schwinger equation

We are now ready to attack a typical scattering problem of an incoming free particle of momentum $\mathbf{p} = \hbar\mathbf{k}$ interacting with a time independent potential $V(\mathbf{r})$. This process is time dependent even though the Hamiltonian does not explicitly depend on time. The Hamiltonian is

$$H = H_0 + H^{'} = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}).$$

The goal is to determine the probability amplitude $\Psi_{\mathbf{k}}(\mathbf{r})$ for finding an outgoing particle at a position $\mathbf{r}$ after the scattering of the incoming particle. This information can be obtained by solving the Schrödinger equation with the appropriate initial conditions. One approach is to derive an integral equation, called the Lippmann-Schwinger equation, that is equivalent to the Schrödinger equation and which incorporates the boundary conditions.

Because of the time independent Hamiltonian, energy is conserved during the scattering process, so we must match the energy eigenvalues of H_0 and H as we learned in the previous chapter. Since H_0 governs the incoming state $|E\rangle$, while H governs the state $|E\rangle^+$ during the scattering process, we write the eigenvalue equations with the same E

$$H|E\rangle^+ = E|E\rangle^+, \;\; H_0|E\rangle = E|E\rangle.$$

As we have learned in the previous chapter, the Lippmann-Schwinger equation describes the evolution of the initial state $|E\rangle$ into the state $|E\rangle^+$ as a result of the interaction

$$|E\rangle^+ = |E\rangle + (E - H_0 + i\varepsilon)^{-1} H^{'} |E\rangle^+.$$

Recall that the $+i\varepsilon$ prescription takes into account causality in the time evolution of the system.

The energy eigenstate of $H_0 = \mathbf{p}^2/2m$ is the free particle state $|E\rangle = |\mathbf{k}\rangle$ with definite momentum $\mathbf{p}|\mathbf{k}\rangle = \hbar\mathbf{k}|\mathbf{k}\rangle$, and $E = \frac{\hbar^2 k^2}{2m}$. We can define the energy

eigenstate of the total H as $|E\rangle^+ \equiv |\mathbf{k}\rangle^+$ where the superscript + indicates that the state corresponds the evolution of the initial free particle state $|\mathbf{k}\rangle$. The direction of the vector $\mathbf{k}$ still indicates the direction of the incoming particle. So we can write the Schrödinger equation with the total Hamiltonian as

$$\left(\frac{\mathbf{p}^2}{2m} + V(\mathbf{r})\right)|\mathbf{k}\rangle^+ = \frac{\hbar^2 k^2}{2m}|\mathbf{k}\rangle^+.$$

Rescaling with the overall factor $2m/\hbar^2$ and re-arranging terms we have

$$\left(k^2 - \mathbf{p}^2/\hbar^2\right)|\mathbf{k}\rangle^+ = v\,(r)\,|\mathbf{k}\rangle^+, \;\; v\,(r) \equiv \frac{2m}{\hbar^2}V(r).$$

The Lippmann-Schwinger equation takes the following form in our case

$$|\mathbf{k}\rangle^+ = |\mathbf{k}\rangle + \left(k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon\right)^{-1} v\,(r)\,|\mathbf{k}\rangle^+.$$

We can now dot this expression on the left by the position bra $\langle\mathbf{r}|$ and insert identity in the form $1 = \int d^3\mathbf{r}'|\mathbf{r}'\rangle\langle\mathbf{r}'|$ to obtain

$$\langle\mathbf{r}|\mathbf{k}\rangle^+ = \langle\mathbf{r}|\mathbf{k}\rangle + \int d^3\mathbf{r}'\,\langle\mathbf{r}|\left(k^2 - p^2/\hbar^2 + i\varepsilon\right)^{-1}|\mathbf{r}'\rangle\langle\mathbf{r}'|v\,(r)\,|\mathbf{k}\rangle^+.$$

The quantity $\langle\mathbf{r}|\mathbf{k}\rangle^+$ is the wavefunction $\Psi_{\mathbf{k}}(\mathbf{r}) \equiv \langle\mathbf{r}|\mathbf{k}\rangle^+$ which represents the probability amplitude that an incoming particle of momentum $\hbar\mathbf{k}$ will be found at position $\mathbf{r}$ after scattering. Similarly, $\langle\mathbf{r}|\mathbf{k}\rangle = e^{+i\mathbf{k}\cdot\mathbf{r}}/(2\pi)^{3/2}$ is the plane wave for the incoming free particle. The quantity

$$G_k^+(\mathbf{r},\mathbf{r}') = \langle\mathbf{r}|\left(k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon\right)^{-1}|\mathbf{r}'\rangle$$

is the Green function for inverting the operator $\left(k^2 - \mathbf{p}^2/\hbar^2\right)$ with appropriate boundary conditions imposed at $t \to -\infty$ (remote past). So now, we can write the Lippmann-Schwinger equation in the form

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{+i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}} + \int d^3\mathbf{r}'\,G_k^{(+)}(\mathbf{r},\mathbf{r}')v(\mathbf{r}')\Psi_{\mathbf{k}}(\mathbf{r}').$$

It is easy to see that the Green function satisfies the differential equation $(\nabla_{\mathbf{r}}^2 + k^2)G_k^{(+)}(\mathbf{r},\mathbf{r}') = \delta^{(3)}(\mathbf{r} - \mathbf{r}')$, namely

$$\begin{aligned}(\nabla_{\mathbf{r}}^2 + k^2)G_k^{(+)}(\mathbf{r},\mathbf{r}') &= \int \frac{d^3\mathbf{q}}{(2\pi)^3}(-q^2 + k^2)(k^2 + q^2 + i\varepsilon)^{-1}e^{+i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} \\ &= \int \frac{d^3\mathbf{q}}{(2\pi)^3}e^{+i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} = \delta^{(3)}(\mathbf{r}-\mathbf{r}').\end{aligned}$$

We can compute the Green function by inserting identity in momentum space

$1 = \int d^3\mathbf{q}\ |\mathbf{q}\rangle\ \langle\mathbf{q}|$, with $\mathbf{p}|\mathbf{q}\rangle = |\mathbf{q}\rangle\hbar\mathbf{q}$, and using the plane wave for $\langle\mathbf{r}|\mathbf{q}\rangle$

$$\begin{aligned} G_k^{(+)}(\mathbf{r},\mathbf{r}') &= \int d^3\mathbf{q}\ \langle\mathbf{r}|\mathbf{q}\rangle\ \langle\mathbf{q}|\left(k^2-\mathbf{p}^2/\hbar^2+i\varepsilon\right)^{-1}|\mathbf{r}'\rangle \\ &= \int d^3\mathbf{q}\frac{e^{+i\mathbf{q}\cdot\mathbf{r}}}{(2\pi)^{3/2}}(k^2-q^2+i\varepsilon)^{-1}\frac{e^{-i\mathbf{q}\cdot\mathbf{r}'}}{(2\pi)^{3/2}} \\ &= \frac{1}{(2\pi)^3}\int d^3\mathbf{q}\ e^{+i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\ (k^2-q^2+i\varepsilon)^{-1} \end{aligned}$$

The integration $\int d^3\mathbf{q}$ is performed as follows

$$\begin{aligned} G_k^{(+)} &= \frac{1}{(2\pi)^3}\int_0^\infty q^2dq\ \left(k^2-q^2+i\varepsilon\right)^{-1}\int_0^{2\pi}d\phi\int_{-1}^{1}d(\cos\theta)e^{+iq|\mathbf{r}-\mathbf{r}'|\cos\theta} \\ &= \frac{1}{(2\pi)^3}\int_0^\infty dq\ q^2(k^2-q^2+i\varepsilon)^{-1}\frac{2\pi}{iq|\mathbf{r}-\mathbf{r}'|}\left(e^{+iq|\mathbf{r}-\mathbf{r}'|}-e^{-iq|\mathbf{r}-\mathbf{r}'|}\right) \\ &= \frac{i}{(2\pi)^2|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty}dq\ \frac{q}{k^2-q^2+i\varepsilon}e^{-iq|\mathbf{r}-\mathbf{r}'|} \end{aligned}$$

The remaining integral can be performed by complex integration. We can write $k^2-q^2+i\varepsilon = (k-q+i\varepsilon)(k+q+i\varepsilon)$ since k is positive and $\varepsilon\to 0$. This shows that there are two poles in the complex q plane at $q = \pm(k+i\varepsilon)$. Since $|\mathbf{r}-\mathbf{r}'|$ is always positive we can close the contour clockwise in the lower complex q plane, and pick up the residue of the pole at $q = -k-i\varepsilon$. Therefore

$$\begin{aligned} G_k^{(+)} &= \frac{i\ (-2\pi i)\ Res\{q=-k-i\varepsilon\}}{(2\pi)^2|\mathbf{r}-\mathbf{r}'|} = \frac{i(\pi i)e^{ik|\mathbf{r}-\mathbf{r}'|}}{(2\pi)^2|\mathbf{r}-\mathbf{r}'|} \\ &= -\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}. \end{aligned}$$

Finally we can write

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{+i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}} - \int d^3\mathbf{r}'\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}v(r')\Psi_{\mathbf{k}}(\mathbf{r}') \tag{14.1}$$

This integral equation determines the probability amplitude $\Psi_{\mathbf{k}}(\mathbf{r})$ of finding the outgoing particle at a position $\mathbf{r}$ after the scattering of the incoming free particle of momentum $\mathbf{p} = \hbar\mathbf{k}$.

14.2 Scattering amplitude, S-Matrix, T-matrix

In general, the observation point $\mathbf{r}$ is located at a distance much greater than the scattering center. So it is appropriate to take the limiting value of the above integral equation for $|\mathbf{r}|\to\infty$. The first term on the right hand side will not be affected by the limit, since it represents an oscillating plane wave. The

integral in the second term has a term $e^{ik|\mathbf{r}-\mathbf{r}'|}|\mathbf{r}-\mathbf{r}'|^{-1}$. Assuming a reasonably short range potential, the $\mathbf{r}'$ integration will get its main contribution for small values of $\mathbf{r}'$ close to the center. Therefore the denominator of this term will be approximated by $|\mathbf{r}-\mathbf{r}'| \sim |\mathbf{r}|$ for large values $|\mathbf{r}|$. However, in the exponential we perform the following expansion

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 - 2\mathbf{r}\cdot\mathbf{r}' + r'^2} \simeq r - \hat{\mathbf{r}}\cdot\mathbf{r}' + O(r'^2/r)$$

so that for large r we have

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{+i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}} - \frac{e^{ikr}}{4\pi r}\int d^3\mathbf{r}'\, e^{-ik\hat{\mathbf{r}}\cdot\mathbf{r}'} v(r')\Psi_{\mathbf{k}}(\mathbf{r}').$$

We define

$$\mathbf{k}' = k\hat{\mathbf{r}}$$

and interpret this vector as the momentum of the outgoing particle. This is justified since $\mathbf{r}$ is the location of a detector where the scattered particle is detected, and therefore $\hat{\mathbf{r}}$ is the direction of the scattered particle. Furthermore, since the particle is a free particle after the scattering takes place, and since energy is conserved, its momentum must have absolute value $\hbar k$ so that the energy is $E = \hbar^2 k^2/2m$. Therefore, we have now an expression for the wavefunction that depends on the incoming and outgoing momenta $(\mathbf{k}, \mathbf{k}')$

$$\begin{aligned}\Psi_{\mathbf{k}}(\mathbf{r}) &\underset{r\to\infty}{\to} \frac{e^{+i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}} - \frac{e^{ikr}}{4\pi r}\int d^3\mathbf{r}'\, e^{-i\mathbf{k}'\cdot\mathbf{r}'} v(r')\Psi_{\mathbf{k}}(\mathbf{r}') \\ &= \frac{1}{(2\pi)^{3/2}}\left[e^{+i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{ikr}}{r} f(\mathbf{k},\mathbf{k}')\right] \end{aligned} \tag{14.2}$$

where

$$\begin{aligned} f(\mathbf{k},\mathbf{k}') &= -\frac{(2\pi)^3}{4\pi}\int d^3\mathbf{r}' \frac{e^{-i\mathbf{k}'\cdot\mathbf{r}'}}{(2\pi)^{3/2}} v(r')\Psi_k(\mathbf{r}') \\ &= -2\pi^2 \langle \mathbf{k}'|v(r)|\mathbf{k}\rangle^+ \end{aligned} \tag{14.3}$$

The last line is verified by inserting identity in the form $1 = \int d^3\mathbf{r}'|\mathbf{r}'\rangle\langle\mathbf{r}'|$. We may notice that the term $\frac{e^{ikr}}{(2\pi)^{3/2}r}$ is a spherical wave that originates at the scattering center, weighted by a function $f(\mathbf{k},\mathbf{k}')$. Therefore, $f(\mathbf{k},\mathbf{k}')$ represents the probability amplitude for an incoming particle of momentum $\mathbf{k}$ to be scattered into the direction $\hat{\mathbf{r}}$ with momentum $\mathbf{k}' = k\hat{\mathbf{r}}$.

Recall that that $\langle \mathbf{k}'|\mathbf{k}\rangle^+$ is the transition amplitude $A_{fi}(\infty, -\infty)$, which is the S-matrix. Analysing this quantity directly through the Lippmann-Schwinger equation, $|\mathbf{k}\rangle^+ = |\mathbf{k}\rangle + (k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon)^{-1} v(r)|\mathbf{k}\rangle^+$, we obtain

$$\langle \mathbf{k}'|\mathbf{k}\rangle^+ = \delta^{(3)}(\mathbf{k}'-\mathbf{k}) + \langle \mathbf{k}'|v(r)|\mathbf{k}\rangle^+ \left(k^2 - k'^2 + i\varepsilon\right)^{-1}$$

As long as $\mathbf{k}'$ is not in the forward direction we may drop the delta function, and the remaining term $\langle \mathbf{k}'|v|\mathbf{k}\rangle^+ \left(k^2 - k'^2 + i\varepsilon\right)^{-1}$ is identified as the T-matrix. From this expression we recognize the scattering amplitude $f(\mathbf{k}, \mathbf{k}')$ as the residue of the pole in the momentum wave function $\langle \mathbf{k}'|\mathbf{k}\rangle^+$, multiplied by the factor $(-2\pi^2)$, as given in Eq.(14.3).

Thus there are two ways of computing the scattering amplitude $f(\mathbf{k}, \mathbf{k}')$. The first is by examining the asymptotic behavior of the wavefunction $\Psi_{\mathbf{k}}(\mathbf{r})$, where $\Psi_{\mathbf{k}}(\mathbf{r})$ is the scattering solution of the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right)\Psi_{\mathbf{k}}(\mathbf{r}) = \frac{\hbar^2 k^2}{2m}\Psi_{\mathbf{k}}(\mathbf{r}).$$

Since the asymptotic behavior has the form of Eq.(14.2), $f(\mathbf{k}, \mathbf{k}')$ is obtained by taking the ratio of the coefficients for the incoming plane wave $e^{+i\mathbf{k}\cdot\mathbf{r}}$ and the outgoing spherical wave $\frac{e^{ikr}}{r}$. A second approach for computing $f(\mathbf{k}, \mathbf{k}')$ is to examine the T-matrix in momentum space. The residue of the pole $\left(k^2 - k'^2 + i\varepsilon\right)^{-1}$ times the factor $(-2\pi^2)$ is the scattering amplitude $f(\mathbf{k}, \mathbf{k}')$.

The discussion above was given in three space dimensions, but it can be repeated for any dimension. In particular the transmission and reflection coefficients that one learns to compute by solving the Schrödinger equation in one dimension correspond to the scattering amplitude. The student should develop experience for computing $f(\mathbf{k}, \mathbf{k}')$ by solving some of the problems at the end of the chapter.

14.3 Differential and total cross section

The differential cross section is defined as follows

$$d\sigma = \frac{\text{number of particles scattered into solid angle } d\Omega \text{ per unit time}}{\text{number of incident particles per unit time per unit area}}$$

Thus an experimentalist simply counts particles in the incident beam and in the detector that captures the scattered particles, and reports the ratio described above. The mathematical expression for what he sees is computed as follows. The denominator is just the probability current of the incident particles $\frac{dN_{in}}{dt(dA)} = |\hat{\mathbf{k}} \cdot \mathbf{J}_{in}|$, where the current is to be computed from the incident wavefunction far away from the interaction region. The numerator can be expressed as the area of an infinitesimal detector times the number of radially scattered particles per unit area per unit time. This amounts to the probablity current of the outgoing particles $\frac{dN_{out}}{dt(dA)} = |\hat{\mathbf{r}} \cdot \mathbf{J}_{out}|$ times the area element $r^2 d\Omega$ that represents an infinitesimal detector located at some solid angle. The current is to be computed from the outgoing wavefunction far away from the interaction region. Therefore one should compare the experimentalist measurements to the expression

$$d\sigma = \frac{|\hat{\mathbf{r}} \cdot \mathbf{J}_{out}|\ r^2 d\Omega}{|\hat{\mathbf{k}} \cdot \mathbf{J}_{in}|}$$

Recall that the probability current for any wavefunction Ψ is given by

$$\mathbf{J} = -i\frac{\hbar}{2m}(\Psi^*\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*) = \mathrm{Re}\left(\frac{1}{m}\Psi^*\left(-i\hbar\boldsymbol{\nabla}\right)\Psi\right)$$

We can thus compute $\mathbf{J}_{in}$ and $\mathbf{J}_{out}$ by using the asymptotic form of the wavefunction given in Eq.(14.2)

$$\Psi_{in}(\mathbf{r}) = e^{+i\mathbf{k}\cdot\mathbf{r}}, \quad \Psi_{out}(\mathbf{r}) = \frac{e^{ikr}}{r}f(\mathbf{k},\mathbf{k}').$$

We do not need to worry about the overall normalization since we are interested only in the ratio. Hence for the incident current we have $|\hat{\mathbf{k}}\cdot\mathbf{J}_{in}| = \frac{\hbar k}{m}$, and for the scattered current we have

$$|\hat{\mathbf{r}}\cdot\mathbf{J}_{out}| = \left|\mathrm{Re}\left(\frac{-i\hbar}{m}\Psi^*\frac{\partial}{\partial r}\Psi\right)\right| \underset{r\to\infty}{=} \frac{\hbar k}{mr^2}|f(\mathbf{k},\mathbf{k}')|^2$$

Therefore the differential cross section is

$$d\sigma = |f(\mathbf{k},\mathbf{k}')|^2 d\Omega_{k'}.$$

In real life the detector has a finite size and the experimentalist counts all the particles that reach the finite size detector. His measurement correspond to integrating the differential cross section over the size of the detector

$$\sigma = \int_{\mathrm{det}\,ector} |f(\mathbf{k},\mathbf{k}')|^2 d\Omega_{k'} \tag{14.4}$$

The total cross section corresponds experimentally to setting up detectors everywhere around the scattering center and capturing all scattered particles. This amounts to integrating the differentail cross section over all angles

$$\sigma_{tot} = \int_{tot} |f(\mathbf{k},\mathbf{k}')|^2 d\Omega_{k'}$$

The cross section has unit of area. It can be pictured as a small effective disk of area σ which deflects the beam from its initial path. The larger/smaller the cross section the larger/smaller the disk.

14.4 Optical theorem

We will first discuss the implications of the conservation of probability in scattering theory. Consider the Schrödinger equation for any potential energy $V(r)$

$$i\hbar\partial_t\Psi = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right)\Psi.$$

The divergence of the probability current $\mathbf{J} = -i\,(\hbar/2m)\,(\Psi^*\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*)$ is computed by substituting $\boldsymbol{\nabla}^2\Psi$ from the Schrödinger equation. The potential terms cancel and one finds

$$-i\frac{\hbar}{2m}\boldsymbol{\nabla}\cdot(\Psi^*\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Psi^*) = -\partial_t|\Psi|^2.$$

The right hand side is the time derivative of the probability density $\rho = |\Psi|^2$. Hence this equation is nothing more than the continuity equation for the probability current

$$\boldsymbol{\nabla}\cdot\mathbf{J}+\frac{\partial\rho}{\partial t} = 0.$$

For energy eigenstates $\Psi_E(\mathbf{r},t) = e^{-iEt/\hbar}\Psi(\mathbf{r})$ the probability density is time independent $\frac{\partial}{\partial t}\left[\Psi_E(\mathbf{r},t)\Psi_E^*(\mathbf{r},t)\right] = \frac{\partial}{\partial t}\left[\psi(\mathbf{r})\psi^*(\mathbf{r})\right] = 0$, and therefore the current has vanishing divergence $\boldsymbol{\nabla}\cdot\mathbf{J}$ =0. Its integral on a spherical volume reduces to a surface integral on the spherical shell by using Stoke's theorem

$$0 = \int_{sphere} d^3\mathbf{r}\ \boldsymbol{\nabla}\cdot\mathbf{J} = \int_{shell} d\Omega r^2(\hat{\mathbf{r}}\cdot\mathbf{J})$$

We let the volume go to infinity, $r\to\infty$, and compute this integral for the asymptotic form of the scattering wavefunction $\Psi_{\mathbf{k}}(\mathbf{r},t) \to (2\pi)^{-3/2}\left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\mathbf{k},\mathbf{k}')e^{ikr}/r\right]$ as follows. We compute

$$r^2(\hat{\mathbf{r}}\cdot\mathbf{J}) = \frac{\hbar}{m(2\pi)^3}\left\{r^2\hat{\mathbf{r}}\cdot\mathbf{k} + k\left|f(\mathbf{k},\mathbf{k}')\right|^2 + r\,\mathrm{Im}\left[i\,(k+\hat{\mathbf{r}}\cdot\mathbf{k})\,e^{ikr-i\mathbf{k}\cdot\mathbf{r}}f(\mathbf{k},\mathbf{k}')\right]\right\}$$

where we used that $\hat{\mathbf{r}}\cdot\boldsymbol{\nabla}$ is just the radial derivative $\partial/\partial r$, and dropped non-leading terms that become negligible as $r\to\infty$. Then the integral of the first term vanishes $\int d\Omega\ \hat{\mathbf{r}}\cdot\mathbf{k}$ =0, and the rest give

$$0 = \int d\Omega\left[k\left|f(\mathbf{k},\mathbf{k}')\right|^2 + kr\,\mathrm{Im}\left\{f(\mathbf{k},\mathbf{k}')\left(i+i\hat{\mathbf{r}}\cdot\hat{\mathbf{k}}\right)e^{ikr-i\mathbf{k}\cdot\mathbf{r}}\right\}\right] = k\sigma_{tot} + I \tag{14.5}$$

The first term is proportional to the total cross section as defined in the previous section $\int d\Omega k\left|f(\mathbf{k},\mathbf{k}')\right|^2 = k\sigma_{tot}\,(k)$. In the second integral I we recall that the direction of $\mathbf{k}'$ is the same as the direction of $\mathbf{r}$, therefore we write $f(\mathbf{k},\mathbf{k}') = f\,(k,\theta)$ to indicate the angle dependence of the scattering amplitude. The integration over ϕ in the solid angle integral $d\Omega = d\phi d\,(\cos\theta)$ is performed trivially $\int d\phi = 2\pi$. The remaining integral over θ

$$I = \lim_{r\to\infty} 2\pi kr\,\mathrm{Im}\left\{\int_{-1}^{1} d\,(\cos\theta)\ \ e^{ikr(1-\cos\theta)}\,(i+i\cos\theta)\,f\,(k,\theta)\right\}$$

is to be performed for $r\to\infty$. Due to the rapid fluctuations of the phase, we use the saddle point method. The phase is stationary at $\theta = 0$, since this is the solution of $\partial_\theta\,(1-\cos\theta) = \sin\theta = 0$. Expanding the integrand around

$\theta = 0$, and then rotating the contour along the steepest descent direction in the complex plane $\theta = \sqrt{\frac{2ix}{kr}}$, we obtain

$$\begin{aligned} I &= \lim_{r\to\infty} 2\pi kr \,\mathrm{Im}\left\{2if\left(k,0\right)\int e^{ikr\theta^2/2}d\theta^2/2\right\} \\ &= \mathrm{Im}\left\{-4\pi f\left(k,0\right)\int_0^\infty e^{-x}dx\right\} = -4\pi\,\mathrm{Im}\, f\left(k,0\right) \end{aligned}$$

Therefore Eq.(14.5) reduces to the "optical theorem"

$$\mathrm{Im}\left(f\left(k,0\right)\right) = \frac{k}{4\pi}\sigma_{tot}\left(k\right). \tag{14.6}$$

This theorem relates the total cross section to the imaginary part of the scattering amplitude in the forward direction, and is a direct consequence of the fact that the total probability density is conserved.

14.4.1 Generalized optical theorem

There is a more general statement that applies to the scattering amplitude $f(\mathbf{k},\mathbf{k}')$ for any potential $V\left(r\right)$. Consider two solutions of the Schrödinger equation $\Psi_E\left(\mathbf{r},t\right) = e^{-Et/\hbar}\Psi_{\mathbf{k}_1}\left(\mathbf{r}\right)$, and $\Psi_E\left(\mathbf{r},t\right) = e^{-Et/\hbar}\Psi_{\mathbf{k}_2}\left(\mathbf{r}\right)$ with the same energy but with the momenta $\mathbf{k}_1, \mathbf{k}_2$ pointing in two arbitrary directions although they have the same magnitude

$$|\mathbf{k}_1| = |\mathbf{k}_2| = k, \tag{14.7}$$

and therefore the same energy $E = \hbar^2k^2/2m$. Using the Schrödinger equation with any potential $V\left(r\right)$ one can easily prove the identity

$$\Psi^*_{\mathbf{k}_1}\left(\mathbf{r}\right)\boldsymbol{\nabla}^2\Psi_{\mathbf{k}_2}\left(\mathbf{r}\right) - \Psi^*_{\mathbf{k}_2}\left(\mathbf{r}\right)\boldsymbol{\nabla}^2\Psi_{\mathbf{k}_1}\left(\mathbf{r}\right) = 0.$$

This can be written in the form of a divergenceless current

$$\boldsymbol{\nabla}\cdot\mathbf{J} = 0,\ \mathbf{J} \equiv \Psi^*_{\mathbf{k}_1}\left(\mathbf{r}\right)\boldsymbol{\nabla}\Psi_{\mathbf{k}_2}\left(\mathbf{r}\right) - \Psi^*_{\mathbf{k}_2}\left(\mathbf{r}\right)\boldsymbol{\nabla}\Psi_{\mathbf{k}_1}\left(\mathbf{r}\right).$$

Integrating it over a large sphere, and applying Stoke's theorem we derive an equation similar to Eq.(14.5)

$$\begin{aligned} 0 &= \int_{sphere} d^3\mathbf{r}\ \boldsymbol{\nabla}\cdot\mathbf{J} = \int_{shell} d\Omega r^2(\hat{\mathbf{r}}\cdot\mathbf{J}) \\ &= \lim_{r\to\infty}\int d\Omega r^2\left\{\Psi^*_{\mathbf{k}_1}\left(\mathbf{r}\right)\frac{\partial}{\partial r}\Psi_{\mathbf{k}_2}\left(\mathbf{r}\right) - \Psi^*_{\mathbf{k}_2}\left(\mathbf{r}\right)\frac{\partial}{\partial r}\Psi_{\mathbf{k}_1}\left(\mathbf{r}\right)\right\}, \end{aligned}$$

The wavefunctions have the form $\Psi_{\mathbf{k}_{1,2}}(\mathbf{r}) \to (2\pi)^{-3/2}\left[e^{i\mathbf{r}\cdot\mathbf{k}_{1,2}} + f(\mathbf{k}_{1,2},\mathbf{k}')e^{ikr}/r\right]$ at $r\to\infty$, with $\mathbf{k}' = k\hat{\mathbf{r}}$. To evaluate the surface integral we expand the plane

waves $e^{i\mathbf{r}\cdot\mathbf{k}_{1,2}}$ in terms of outgoing and incoming spherical waves $e^{\pm ikr}/r$ for large r,

$$\exp\left(i\mathbf{r}\cdot\mathbf{k}_{1,2}\right) \underset{r\to\infty}{=} 2\pi\left[\frac{e^{+ikr}}{ikr}\delta\left(\Omega_{k_{1,2}}-\Omega_r\right)-\frac{e^{-ikr}}{ikr}\delta\left(\Omega_{k_{1,2}}-\Omega_r\right)\right] \tag{14.8}$$

where we have used Eq.(14.7). This relation is derived as outlined in a problem at the end of this chapter. It is now straightforward to take the radial derivatives and evaluate the surface integral. After a rearrangement of the terms the result can be written as

$$\frac{1}{2i}\left(f(\mathbf{k}_1,\mathbf{k}_2)-f^*\left(\mathbf{k}_2,\mathbf{k}_1\right)\right)=\frac{k}{4\pi}\int d\Omega_{k'}f^*\left(\mathbf{k}_2,\mathbf{k}'\right)f(\mathbf{k}_1,\mathbf{k}'). \tag{14.9}$$

The optical theorem discussed in the previous section is a special limit of this more general relation. Indeed for the case of $\mathbf{k}_1=\mathbf{k}_2=\mathbf{k}$, Eq.(14.9) reduces to Eq.(14.6).

14.5 Scattering of identical particles

Consider two identical particles that are scattered by a finite range potential $V(\mathbf{r})$. The total wavefunction $\Psi(\mathbf{r}_1,\mathbf{r}_2)$ can be written in terms of the center of mass and relative coordinates $\mathbf{R}$ and $\mathbf{r}$ respectively. In this basis the wavefunction is separable

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) \;=\; \Psi(\mathbf{R},\mathbf{r}) \;=\; \Psi_{cm}(\mathbf{R})\Psi_{rel}(\mathbf{r})$$

Since $\Psi_{cm}(\mathbf{R})$ is trivial we concentrate on $\Psi_{rel}(\mathbf{r})$ and for simplicity rename it $\Psi\left(\mathbf{r}\right)$.

Allowing for spin degrees of freedom, each wavefunction has an additional spin index to represent an appropriate spinning wavefunction. Therefore the two particle wavefunction has a pair of spin indices, and is denoted as $\Psi_{\sigma_1,\sigma_2}(\mathbf{r})$. Under the interchange of two identical particles this wavefunction must satisfy quantum statistics.

First we consider spinless bosons. There are no spin indices and the wavefunction must be even under the interchange of the positions of the two particles. For the relative wavefunction this amounts to

$$\text{Spinless bosons:}\;\; \Psi(\mathbf{r})=\Psi(-\mathbf{r})$$

Hence the wave function for scattered particles must have the form

$$\begin{aligned}\Psi_{\mathbf{k}}(\mathbf{r}) &= \frac{C}{(2\pi)^{3/2}}\left[(e^{i\mathbf{k}\cdot\mathbf{r}}+e^{-i\mathbf{k}\cdot\mathbf{r}})+\frac{e^{ikr}}{r}\left(f(\mathbf{k},\hat{\mathbf{r}})+f(\mathbf{k},-\hat{\mathbf{r}})\right)\right]\\ &= \frac{C}{(2\pi)^{3/2}}\left[(e^{i\mathbf{k}\cdot\mathbf{r}}+e^{-i\mathbf{k}\cdot\mathbf{r}})+\frac{e^{ikr}}{r}\left(f(E,\theta)+f(E,\pi-\theta)\right)\right]\end{aligned}$$

where we have used $f(\mathbf{k},\hat{\mathbf{r}})=f(E,\theta)$.

Next we consider spin 1/2 fermions. The wavefunction must be odd under the interchange of the positions and of the spins of the two particles. We may rewrite the wavefunction in terms of total spin 0,1 components by using the Clebsch-Gordan coefficients $\chi^0_{\sigma_1\sigma_2} \equiv \langle \frac{1}{2}, \sigma_1; \frac{1}{2}, \sigma_2 | 0, 0\rangle$ and $\chi^s_{\sigma_1\sigma_2} \equiv \langle \frac{1}{2}, \sigma_1; \frac{1}{2}\sigma_2 | 1, s\rangle$ where $s = 0, \pm 1$

$$\Psi_{\sigma_1\sigma_2}(\mathbf{r}) = \chi^0_{\sigma_1\sigma_2}\Psi^0_{\mathbf{k}}(\mathbf{r}) + \chi^s_{\sigma_1\sigma_2}\Psi^s_{\mathbf{k}}(\mathbf{r}).$$

Under the interchange of the two spins the spin 0 channel is antisymmetric $\chi^0_{\sigma_1\sigma_2} = -\chi^0_{\sigma_2\sigma_1}$, while the spin 1 channel is symmetric $\chi^s_{\sigma_1\sigma_2} = +\chi^s_{\sigma_2\sigma_1}$. Therefore $\Psi^0_{\mathbf{k}}(\mathbf{r})$ must be symmetric while $\Psi^s_{\mathbf{k}}(\mathbf{r})$ must be antisymmetric so that the overall wavefunction is antisymmetric $\Psi_{\sigma_1\sigma_2}(\mathbf{r}) = -\Psi_{\sigma_2\sigma_1}(-\mathbf{r})$

$$\text{Spin 1/2 fermions: } \Psi^0_{\mathbf{k}}(\mathbf{r}) = \Psi^0_{\mathbf{k}}(-\mathbf{r}), \quad \Psi^s_{\mathbf{k}}(\mathbf{r}) = -\Psi^s_{\mathbf{k}}(-\mathbf{r}).$$

Thus we can write the following forms

$$\Psi^0_{\mathbf{k}}(\mathbf{r}) = \frac{C_0}{(2\pi)^{3/2}}\left[(e^{i\mathbf{k}\cdot\mathbf{r}} + e^{-i\mathbf{k}\cdot\mathbf{r}}) + \frac{e^{ikr}}{r}\left(f_0(E,\theta) + f_0(E,\pi-\theta)\right)\right]$$
$$\Psi^s_{\mathbf{k}}(\mathbf{r}) = \frac{C_s}{(2\pi)^{3/2}}\left[(e^{i\mathbf{k}\cdot\mathbf{r}} - e^{-i\mathbf{k}\cdot\mathbf{r}}) + \frac{e^{ikr}}{r}\left(f_1(E,\theta) - f_1(E,\pi-\theta)\right)\right]$$

From these expressions we derive some general observable consequences of quantum statistics, independent of the details of the forces, as follows. Consider the differential cross section observed by an experimentalist. In the case of spinless bosons we have

$$\text{Spinless bosons: } \left(\frac{d\sigma}{d\Omega}\right)_B = |f(\theta) + f(\pi-\theta)|^2. \tag{14.10}$$

In the case of spin 1/2 fermions, if the final spin is observed to be spin zero we have

$$\text{Spin 1/2 fermions, total spin 0: } \left(\frac{d\sigma}{d\Omega}\right)_0 = \frac{1}{4}|f_0(\theta) + f_0(\pi-\theta)|^2$$

where the factor of 1/4 comes from spin averaging over the initial spins of the fermions, with a factor of 1/2 for each fermions $(1/2)^2 = 1/4$. If the final total spin is observed to be 1 but the orientation is not observed, we get

$$\text{Spin 1/2 fermions, total spin 1: } \left(\frac{d\sigma}{d\Omega}\right)_1 = \frac{3}{4}|f_1(\theta) - f_1(\pi-\theta)|^2$$

where the factor of 1/4 is explained above, while the factor of 3 is due to the fact that there are 3 spin states but the probablities are summed over the 3 states. In general, if the initial or final spins are not observed, then the differential cross section includes factors that correspond to averaging over the initial spins and summing over the final spins.

Note that at $\theta = \pi/2$ the differential cross section $\left(\frac{d\sigma}{d\Omega}\right)_1$ vanishes, which means that in the spin 1 channel an experimentalist should see no particles

scattered at 90^0 from the direction of the beam. This is in clear distinction to both $\left(\frac{d\sigma}{d\Omega}\right)_B$ and $\left(\frac{d\sigma}{d\Omega}\right)_0$ which do not vanish. This is a rather general property that follows only from quantum statistics, independent of any details of the scattering system.

There is an even easier test under further assumptions. If the total spin of the scattered fermions is not observed either, then the differential cross section is is obtained by summing over all spin components in the final state. Assuming that the scattering force is independent of spin we use $f_0(\theta) = f_1(\theta) = f(\theta)$ and obtain

$$\text{Fermions: } \left(\frac{d\sigma}{d\Omega}\right)_F = \frac{1}{4}|f(\theta) + f(\pi-\theta)|^2 + \frac{3}{4}|f(\theta) - f(\pi-\theta)|^2. \quad (14.11)$$

Comparing boson versus fermion scattering at $\theta = \pi/2$, we have

$$\left(\frac{d\sigma}{d\Omega}\right)_B \to_{\theta\to\pi/2} 4|f(\pi/2)|^2, \quad \left(\frac{d\sigma}{d\Omega}\right)_B \to_{\theta\to\pi/2} |f(\pi/2)|^2 \quad (14.12)$$

We notice an enhancement for bosons relative to fermions. This is due to the constructive interference for bosons at $\theta = \pi/2$, and the destructive interference for fermions in the spin 1 channel.

14.6 Born approximation

We have seen that the differential cross section $d\sigma$ is directly proportional to $|f(\mathbf{k},\mathbf{k}')|^2$, which, in turn, was expressed as $f(\mathbf{k},\mathbf{k}') = -2\pi^2\langle\mathbf{k}'|v(r)|\mathbf{k}\rangle^+$. Scattering experiments then give us a possibility to discover the matrix elements of $V(r)$ in momentum space thus revealing the properties of the interaction $V(r)$.

This property is particularly evident in the *First Born Approximation* denoted as $f^{(1)}(\mathbf{k},\mathbf{k}')$, which corresponds to approximating the state $|\mathbf{k}\rangle^+$ with the free particle state $|\mathbf{k}\rangle$. We will later investigate the conditions under which the approximation is valid. In this case indeed $f^{(1)}(\mathbf{k},\mathbf{k}')$ is proportional to the matrix elements of $V(r)$ in the free particle Hilbert space $\langle\mathbf{k}'|V(r)|\mathbf{k}\rangle$ which is given by representing the states with plane waves as in Eq.(14.3)

$$f^{(1)}(\mathbf{k}',\mathbf{k}) = -2\pi^2 \int d^3\mathbf{r} \frac{e^{-i\mathbf{k}'\cdot\mathbf{r}}}{(2\pi)^{3/2}} \frac{2m}{\hbar^2} V(r) \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}}.$$

This is just a Fourier transform of the potential using the momentum transfer $\hbar(\mathbf{k}'-\mathbf{k})$

$$f^{(1)}(\mathbf{k}'-\mathbf{k}) = -\frac{m}{2\pi\hbar^2}\int d^3\mathbf{r} V(r) e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} \quad (14.13)$$

For a rotationally invariant potential this is a function of only $|\mathbf{k}'-\mathbf{k}| = 2k\sin(\theta/2)$. Thus, from the scattering amplitude one can extract the details of the potential $V(r)$ through the inverse Fourier transform. One can thus probe the scatterer by bombarding it with a beam of known particles and extract the nature of the

forces between beam and target by measuring the energy and angle dependence of the differential cross section.

To justify the approximation we start from the general expression of the scattering amplitude $f(\mathbf{k},\mathbf{k}') = -2\pi^2\langle\mathbf{k}'|v\,(r)\,|\mathbf{k}\rangle^+$, and the Lippmann-Schwinger equation $|\mathbf{k}\rangle^+ = |\mathbf{k}\rangle + Gv|\mathbf{k}\rangle^+$, where $G = \left(k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon\right)^{-1}$ is the operator version of the Green function. We solve formally for $|\mathbf{k}\rangle^+$ and substitute it in $f(\mathbf{k},\mathbf{k}')$ to get the exact expressions

$$|\mathbf{k}\rangle^+ = (1-Gv)^{-1}|\mathbf{k}\rangle,\ f(\mathbf{k},\mathbf{k}') = -2\pi^2\langle\mathbf{k}'|v\,(r)\,(1-Gv)^{-1}|\mathbf{k}\rangle. \tag{14.14}$$

If the incoming energy $E = \hbar^2k^2/2m$ is sufficiently large compared to the potential energy $V\,(r)$, then Gv is roughly of order $v/k^2 = V\,(r)\,/E$ and is small. This will be examined in more detail below. Under such circumstances we can approximate the expressions for $|\mathbf{k}\rangle^+$ or $f(\mathbf{k},\mathbf{k}')$ by expanding in a power series and obtain

$$f(\mathbf{k}',\mathbf{k}) = -2\pi^2\ \langle\mathbf{k}'|v + vGv + vGvGv + \cdots|\mathbf{k}\rangle. \tag{14.15}$$

The n^{th} Born approximation is defined by keeping the first n terms in this expansion. Thus, the first Born approximation is $f^{(1)}(\mathbf{k}',\mathbf{k})$ as given above, while the second Born approximation is given by $f(\mathbf{k}',\mathbf{k}) = f^{(1)}(\mathbf{k}',\mathbf{k}) + f^{(2)}(\mathbf{k}',\mathbf{k})$ with $f^{(2)}(\mathbf{k}',\mathbf{k})$ given by

$$f^{(2)}(\mathbf{k}',\mathbf{k}) = -2\pi^2\langle\mathbf{k}'|vGv|\mathbf{k}\rangle,$$

and so on.

14.6.1 Scattering from the Yukawa Potential

As an example consider the Yukawa potential

$$V(r) = V_0\frac{e^{-r/r_0}}{r/r_0},$$

where r_0 is the distance scale that determines the range of the potential. Assuming a sufficiently high incoming energy compared to the potential strength V_0 we can approximate the scattering amplitude by using the first Born approximation. This amounts to the Fourier transform of the Yukawa potential as given in Eq.(14.13)

$$\begin{aligned} f^{(1)}(\mathbf{k},\mathbf{k}') &= -\frac{2m}{4\pi\hbar^2}\int_0^\infty r^2dr\ V_0\frac{e^{-r/r_0}}{r/r_0}\ \int_0^{2\pi}d\phi\int_{-1}^{1}d(\cos\theta)e^{+i|\mathbf{k}-\mathbf{k}'|(\cos\theta)r} \\ &= -\frac{m}{\hbar^2}\int_0^\infty r^2dr\ V_0\frac{e^{-r/r_0}}{r/r_0}\frac{e^{+i|\mathbf{k}-\mathbf{k}'|r} - e^{-i|\mathbf{k}-\mathbf{k}'|r}}{i|\mathbf{k}-\mathbf{k}'|r} \end{aligned}$$

After cancelling the r^2 factors in the numerator and denominator, the remaining integral is easily performed and we obtain

$$f^{(1)}(\mathbf{k},\mathbf{k}') = -\frac{\left(2mV_0r_0^3/\hbar^2\right)}{(\mathbf{k}'-\mathbf{k})^2r_0^2+1} = -\frac{\left(2mV_0r_0^3/\hbar^2\right)}{4k^2r_0^2\sin^2(\theta/2)+1}. \tag{14.16}$$

From this expression we obtain the differential cross section due to the Yukawa potential as a function of incoming energy $E = \hbar^2 k^2/2m$ and scattering angle θ

$$\frac{d\sigma}{d\Omega} = \left(\frac{\left(2mV_0 r_0^3/\hbar^2\right)}{4k^2 r_0^2 \sin^2(\theta/2) + 1} \right)^2 .$$

The results of experiments can be plotted as a function of k and θ for different ranges of these experimentally adjustable parameters and compared to the calculated expression above.

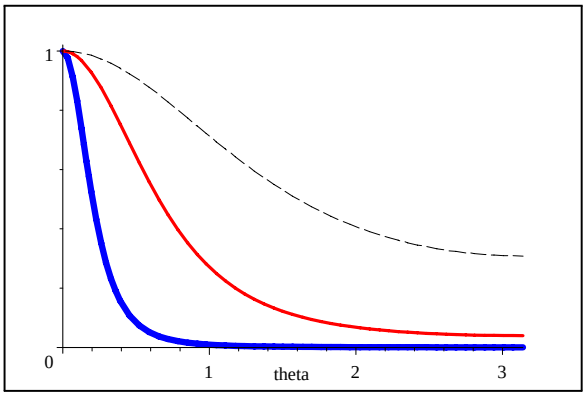


Fig.14.1 - Plots of $d\sigma/d\Omega$ versus θ

In Fig.14.1 the theoretical differential cross section is plotted for $0 \leq \theta \leq \pi$ for three different values of the energy. The vertical axis is normalized by dividing by the value of the forward scattering cross section, $(d\sigma/d\Omega)(\theta = 0) = \left(2mV_0 r_0^3/\hbar^2\right)^2$, which is independent of both k, θ. The dotted line is at some low energy, the thin line is at a higher energy, and the thick line is at an even higher energy. We see that the forward peak (small θ) is sharper for higher energies, which indicates that more and more particles tend to be scattered in the forward direction as the incoming energy is increased. It is clear that by obtaining such curves experimentally and comparing to the predicted shapes one can extract the range of the potential r_0 and the strength of the potential V_0, thus reconstructing the potential $V(r)$.

It is useful to compare this result to the equivalent computation via a Feynman graph in first order perturbation theory in relativistic quantum field theory.

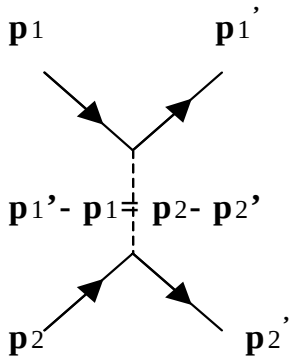


Fig.14.2 - Particle exchange

The Feynman diagram in Fig.14.2 describes two particles of initial 4-momenta p_1^μ and p_2^μ interacting by exchanging another particle (dotted line) and coming out with 4-momenta $p_1'^\mu$ and $p_2'^\mu$. The strength of the interaction at each vertex is represented by a constant g. The exchanged particle propagates from one interaction point in space-time to the other by carrying the 4-momentum transfer

$$p^\mu = p_1'^\mu - p_1^\mu = p_2^\mu - p_2'^\mu.$$

This is written in two forms due to momentum and energy conservation $p_1^\mu + p_2^\mu = p_1'^\mu + p_2'^\mu$. The scattering amplitude in lowest order perturbation theory for this process is given by the mathematical expression for this Feynman graph, which is

$$f^{(1)} = gG\left(p\right)g.$$

Here $G\left(p\right)$ is the Feynman propagator $G\left(p\right) = \left(p^2c^2 - m^2c^4 + i\varepsilon\right)^{-1}$, where m is the mass of the exchanged particle and $p^2 = p\cdot p = \left(p_0\right)^2 - \mathbf{p}^2$ is the relativistic dot product which takes the form $p^2 = \left(p_1' - p_1\right)^2 = \left(E_1' - E_1\right)^2 - \left(\mathbf{p}_1' - \mathbf{p}_1\right)^2$. The $i\varepsilon$ in the propagator can be set to zero as long the Feynman graph has no loops.

In the case of elastic scattering off a heavy (static) target, as in our case of scattering off a potential $V\left(r\right)$, the energy is conserved at each vertex $E_2' = E_2$ or $E_1' = E_1$. Therefore the Feynman propagator depends only on the 3-momentum transfer $G\left(p\right) = -\left(\left(\mathbf{p}_1' - \mathbf{p}_1\right)^2 c^2 + m^2c^4\right)^{-1}$. To compare to our notation in potential scattering, we take $\mathbf{p}_1 = \hbar\mathbf{k}$ and $\mathbf{p}_1' = \hbar\mathbf{k}'$. Then we find that the Feynman amplitude $gG\left(p\right)g$ takes an identical form to the first Born approximation for the Yukawa potential as in Eq.(14.16)

$$f^{(1)}(\mathbf{k},\mathbf{k}') = \frac{-g^2}{\left(\hbar\mathbf{k}' - \hbar\mathbf{k}\right)^2 c^2 + m^2c^4} = -\frac{\left(g/mc^2\right)^2}{\left(\hbar/mc\right)^2\left(\mathbf{k}'-\mathbf{k}\right)^2 + 1}.$$

By comparing these expressions for $f^{(1)}(\mathbf{k},\mathbf{k}')$ we also learn that the range r_0 and strength V_0 of the Yukawa potential are determined by the mass m and the

coupling constant g of the exchanged particle respectively

$$r_0 = \frac{\hbar}{mc}, \;\; V_0 = \frac{g^2}{2\hbar c}. \tag{14.17}$$

This expression for r_0 is just the Compton wavelength of the exchanged particle.

This point was the crucial observation in Yukawa's proposal for the pion. He suggested that pion exchange was the underlying process that gives rise to the strong interactions among nucleons. By using $\hbar c = 197$ MeV fm, and the known range of the nuclear interaction, $r_0 \simeq 1$ to 2 fm, he predicted the existence and the mass of the pion $m_\pi c^2 = 140$ MeV, although at that time he thought he was describing the muon.

As we know it today, the underlying force in strong interactions is the color interactions among quarks via gluon exchange, and these give rise to an effective interaction that is described mainly as pion exchange among nucleons. The effective interaction description is valid as long as the energy of the scattering is not much larger than a few hundred MeV, as is the case in the nucleus and in low energy particle physics. But when the scattering energies go well above a few GeV the details of the quark-gluon structure become evident through the form of the scattering cross section in its dependence on energy and angle.

The weak force is mediated by the exchange of the $W^\pm$ and Z^0 particles whose masses are $m_W c^2 = 80.4\ GeV$ and $m_Z c^2 = 91.2\ GeV$. These large masses explain the very short range of the weak interaction

$$(r_0)_{weak} \approx \frac{\hbar c}{m_W c^2} = \frac{197\ MeV \times fm}{80.4\ GeV} \simeq 2.5 \times 10^{-18}\ m \tag{14.18}$$

Similarly, the electromagnetic force is mediated by the exchange of massless photons, therefore the range of the electromagnetic force is infinite.

14.6.2 Scattering from the Coulomb Potential

The Coulomb potential is obtained from the Yukawa potential in the large r_0 limit, provided that $r_0 V_0 = q_1 q_2$ is held fixed as $r_0 \to \infty$. Then the Yukawa potential reduces to the Coulomb potential $V(r) = q_1 q_2 / r$. We can apply the same reasoning to every step of the computation of the scattering amplitude and cross section for the Yukawa potential and therefore obtain the corresponding results for the Coulomb potential by the limiting procedure. Thus, from Eq.(14.16) we obtain the first Born approximation for Coulomb scattering

$$\frac{d\sigma}{d\Omega} = \left(\frac{q_1 q_2}{4E \sin^2 \theta/2} \right)^2,$$

where we have replaced $E = \hbar^2 k^2 / 2m$. Since $\hbar$ does not appear in this expression we may conclude that this must be the same as the classical result for the differential cross section. Indeed this should be expected on the basis of the correspondance principle at high energies, which is when the Born approximation is valid.

This formula was derived by Rutherford to describe the scattering of α-particles from thin foils. From this expression that fitted his experiments he concluded that the α particles were scattered by a charged pointlike heavy center which resided inside the atoms that made up the thin foil. This followed from the fact that this formula was derived under the assumption of a pointlike positively charged heavy target as described by the Coulomb potential. If this assumption is replaced by another one in which the positive charge is distributed throughout the atom, then the energy and angular dependence of the differential cross section is quite different. These observations led Rutherford to propose the structure of the atom, namely that the atom must contain a small point-like heavy charged nucleus, much smaller than the size of the atom.

14.6.3 Validity of the Born approximation

The exact scattering amplitude which was given in Eq.(14.3) can be written as

$$f(\mathbf{k},\mathbf{k}')=-\frac{1}{4\pi}\int d^3\mathbf{r}\ e^{-i\mathbf{k}'\cdot\mathbf{r}}\ v(r)\ (2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}).$$

where $\Psi_{\mathbf{k}}(\mathbf{r})$ satisfies the integral equation in Eq.(14.1)

$$(2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}) = e^{+i\mathbf{k}\cdot\mathbf{r}} - \int d^3\mathbf{r}'\,\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}v(r')\ (2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}')$$

The first Born approximation $f(\mathbf{k},\mathbf{k}') \simeq f^{(1)}(\mathbf{k},\mathbf{k}')$ was obtained by keeping only the first term in $(2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}+\Delta_{\mathbf{k}}(\mathbf{r})$ where

$$\Delta_{\mathbf{k}}(\mathbf{r}) = -\frac{1}{4\pi}\int d^3\mathbf{r}'\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}v(\mathbf{r}')\ (2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}').$$

Hence in $f(\mathbf{k},\mathbf{k}')$ we have neglected a term of the form

$$\Delta f(\mathbf{k}',\mathbf{k}) = -\frac{1}{4\pi}\int d^3\mathbf{r}\ e^{-i\mathbf{k}'\cdot\mathbf{r}}\ v(r)\ \Delta_{\mathbf{k}}(\mathbf{r}).$$

We would like to estimate the error we make by neglecting the term $\Delta f(\mathbf{k},\mathbf{k}')$. Since the integrals in these equations involve a potential $v(\mathbf{r})$ with a finite range, we may approximate $\Delta_{\mathbf{k}}(\mathbf{r}) \approx \Delta_{\mathbf{k}}(\mathbf{r}=0)$ in the formula for $\Delta f(\mathbf{k},\mathbf{k}')$. Thus, $\Delta f(\mathbf{k}',\mathbf{k})$ is proportional to $\Delta_{\mathbf{k}}(\mathbf{r}=0)$. This will be negligible compared to the the first Born approximation $f^{(1)}(\mathbf{k},\mathbf{k}')$, as long as near $\mathbf{r}=0$ the $\Delta_{\mathbf{k}}(0)$ part of the wavefunction $(2\pi)^{3/2}\Psi_{\mathbf{k}}(0) = 1+\Delta_{\mathbf{k}}(0)$ is negligible compared to 1. Thus we must check under what conditions we obtain

$$\text{validity:}\quad |\Delta_{\mathbf{k}}(0)| \ll 1 \tag{14.19}$$

This condition takes the following form when we approximate $(2\pi)^{3/2}\Psi_{\mathbf{k}}(\mathbf{r}')$ by the plane wave consistently with the Born approximation

$$|\Delta_{\mathbf{k}}(0)| \simeq \left|-\frac{1}{4\pi}\int d^3\mathbf{r}'\frac{e^{ikr'}}{r'}v(r')e^{i\mathbf{k}\cdot\mathbf{r}'}\right| \ll 1 \tag{14.20}$$

The angular integrals are easily performed for a rotationally invariant potential and the condition becomes

$$\Delta_{\mathbf{k}}(0) \simeq \frac{1}{k}\int_0^\infty dr \sin(kr) e^{ikr} v(r) = \frac{C}{k} \tag{14.21}$$

where C is a constant determined by the potential. This shows that $\Delta_{\mathbf{k}}(0)$ decreases as the incoming energy increases. Therefore for sufficiently large incoming energies as compared to the strength of the potential we see that the Born approximation is valid.

However, the criterion $|\Delta_{\mathbf{k}}(0)| \ll 1$ could be satisfied for some potentials even if the incoming energies are not very high. In particular, consider the low incoming energy limit given by $k \to 0$. In this limit $\Delta_{\mathbf{k}}(0)$ becomes

$$\lim_{k\to 0} \Delta_{\mathbf{k}}(0) = \int_0^\infty dr' r' v(r') \ll 1. \tag{14.22}$$

As long as this quantity is much smaller than 1, the Born approximation can be applied even at low energies.

Validity for the spherical square well

We now check the criterion to the spherical square well $V(r) = -V_0 \theta(a - r)$. For any energy we compute

$$\Delta_{\mathbf{k}}(0) \simeq \frac{1}{k}\frac{2mV_0}{\hbar^2}\int_0^a \sin(kr) e^{ikr} dr = i\frac{mV_0 a^2}{\hbar^2 (ka)}\left(1 - \frac{\sin ak}{ak} e^{iak}\right)$$

For large k this vanishes as an inverse power, $\Delta_{\mathbf{k}}(0) \simeq \frac{mV_0 a^2}{\hbar^2 (ak)}$, and therefore $|\Delta_{\mathbf{k}}(0)| \ll 1$ can always be satisfied for sufficiently large incoming energies.

For validity of the First Born approximation at low energies we require

$$|\Delta_{\mathbf{k}}(0)| \ll 1 \quad \to \frac{2m|V_0|}{\hbar^2}\int_0^a dr' r' = \frac{m|V_0| a^2}{\hbar^2} \ll 1. \tag{14.23}$$

This relation can also be derived by an intuitive type of reasoning as follows. For sufficiently low energies the particle could be trapped in the spherical well. If that occurs, then the particle is located within a distance $\Delta x \sim a$, therefore by the uncertainty principle it has momentum $p \sim \hbar/a$. Its kinetic energy is estimated as $E_{kin} \approx p^2/2m = \frac{\hbar^2}{2a^2 m}$. For scattering to happen, we must require that the particle is not trapped in the well, which would be the case when the kinetic energy is larger than the depth of the well

$$\text{no binding:} \quad |V_0| \ll \frac{\hbar^2}{2a^2 m}$$

Except for the factor of 2 this is the same as the validity criterion derived above. This amounts to saying that the first Born approximation is valid whenever the particle has sufficient energy so that it is not trapped in the well.

While this intuitive discussion is instructive, it should not be taken too far. For example, if we have a potential barrier instead of a potential well there will never be a bound state, but still there is a criterion of validity for the first Born approximation which is the same as Eq.(14.23).

Validity for the Yukawa Potential

We now analyze the criterion for the Yukawa potential

$$\begin{aligned}\Delta_{\mathbf{k}}(0)&=\frac{1}{k}\frac{2mV_0}{\hbar^2}\int_0^\infty \sin(kr)e^{ikr}\frac{e^{-r/r_0}}{r/r_0}dr \\ &= \frac{mV_0r_0^2}{\hbar^2\left(kr_0\right)}\left[\tan^{-1}\left(2kr_0\right)+\frac{i}{2}\ln\left(1+4k^2{r_0}^2\right)\right]\end{aligned}$$

This expression is similar to the one above for the square well, which is not surprizing since the square well can be thought of as a deformed version of the Yukawa potential. If we analyze the above expression for low and high energies, we find

$$kr_0 \gg 1:\ \Delta_{\mathbf{k}}(0) \rightarrow \frac{mV_0r_0^2}{\hbar^2\left(kr_0\right)}\left(\frac{\pi}{2}+i\ln\left(2kr_0\right)+O\left(\frac{1}{kr_0}\right)\right) \tag{14.24}$$

$$kr_0 \ll 1:\ \Delta_{\mathbf{k}}(0) \rightarrow 2\frac{mV_0r_0^2}{\hbar^2}\left(1+O\left(kr_0\right)\right) \tag{14.25}$$

We again find that the first Born approximation is always valid for sufficiently high energies. Furthermore, when

$$2mV_0r_0^2/\hbar^2 \ll 1$$

it is also valid at low energies. As expected the critrion as well as the meaning of these equations is similar to the case of the spherical square well.

Validity for the Coulomb Potential

One could try to analyze the validity criterion for the Coulomb potential by taking a limit of the Yukawa potential. Then we find

$$\Delta_{\mathbf{k}}(0)=\frac{mq_1q_2}{\hbar^2k}\lim_{r_0\rightarrow\infty}\left[\tan^{-1}\left(2kr_0\right)+\frac{i}{2}\ln\left(1+4k^2{r_0}^2\right)\right] \tag{14.26}$$

$$=\frac{mq_1q_2}{\hbar^2k}\left(\frac{\pi}{2}+\frac{i}{2}\ln\left(\infty\right)\right) \tag{14.27}$$

Because of the diverging term we could suspect that the first Born approximation would not be valid in this case. However, it turns out that the 1^{st} Born approximation is valid also for the Coulomb potential, the reason being that any divergence can be absorbed in the phase of $f(\mathbf{k},\mathbf{k}')$, thus producing a finite differential cross section.

We discussed a few examples in which the Born approximation works quite well. However, not every potential behaves so nicely with respect to the first Born approximation. While it is always possible to obtain a good approximation at high energies, for some potentials it is not always possible to use the Born approximation at at low energies. This happens when the criterion

$$\left| \int_0^\infty r' v(r') dr' \right| \ll 1$$

is not satisfied. If suffices to consider the simple example of $V = V_0 \left(\frac{r_0}{r}\right)^3$, for which the divergence at $r = 0$ is too strong, making the above condition invalid.

14.6.4 Optical theorem and the Born approximation

We discussed above that the optical theorem is satisfied independent of the details of any potential $V(r)$. Consider the expression for the scattering amplitude in the first Born approximation $f^{(1)}(k, \theta)$ as given in Eq.(14.13). It is straightforward to see that this expression is real since the integral does not change when $\mathbf{k} - \mathbf{k}'$ is replaced by $-(\mathbf{k} - \mathbf{k}')$. The optical theorem appears to be violated by the the first Born approximation since $\mathrm{Im}\left\{f^{(1)}(k, 0)\right\} = 0$ while the cross section as computed in the first Born approximation $\sigma(k) = \int d\Omega \left|f^{(1)}(k, \theta)\right|^2$ obviously is nonzero. Is there a problem?

There is no problem since $\mathrm{Im}\left\{f^{(1)}(k, 0')\right\}$ is first order, while $\int d\Omega \left|f^{(1)}(k, \theta)\right|^2$ is second order, in the expansion of the exact formulas in powers of the potential as in Eq.(14.15). If we insert an artificial parameter λ to keep track of the order of the perturbative expansion, then the optical theorem reads

$$\mathrm{Im}\left\{\sum_{n=1}^{\infty} \lambda^n f^{(n)}(k, 0)\right\} = \int d\Omega \left|\sum_{m=1}^{\infty} \lambda^m f^{(m)}(k, \theta)\right|^2 .$$

The eqution is satisfied for each power of λ independently. For $n = 1$, we have $\mathrm{Im}\left\{f^{(1)}(\mathbf{k} = \mathbf{k}')\right\} = 0$, while, for $n = 2$, we obtain

$$\mathrm{Im}\left\{f^{(2)}(k, 0)\right\} = \int d\Omega \left|f^{(1)}(k, \theta)\right|^2 .$$

So, the optical theorem predics that indeed $\mathrm{Im}\left\{f^{(1)}(k, 0)\right\} = 0$, and that the total cross section determines the imaginary part of the second Born approximation in the forward direction. The generalized optical theorem of Eq.(14.9) makes an even stronger predictions about similar relations of terms in of successive Born approximations.

14.7 The eikonal approximation

When $V(r)$ is a slowly varying function in r, and the incoming energy is sufficiently large so as to satisfy $\hbar^2 k^2 / 2m \gg V(r)$, we can find an approximation

based on semi-classical methods. This is the basis of the eikonal approximation discussd in this section. The wavefunction may be written in the form

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} e^{iS(\mathbf{k},\mathbf{r})/\hbar}$$

for a complex $S(\mathbf{k},\mathbf{r})$. Then the Schrödinger equation reduces to an equation for $S(\mathbf{k},\mathbf{r})$

$$(\boldsymbol{\nabla} S)^2 - i\hbar \boldsymbol{\nabla}^2 S = 2m(E - V(r))$$

Since $V(r)$ is slowly varying with r, one expects that $S(\mathbf{k},\mathbf{r})$ is also a slowly varying function in r, and neglect second order derivatives of $S(\mathbf{k},\mathbf{r})$. This can also be justified on the basis of large momenta which is represented by the first derivative term $(\boldsymbol{\nabla} S)^2$. Under this assumption the equation reduces to the classical Hamilton-Jacobi equation

$$(\boldsymbol{\nabla} S)^2 = 2m(E - V(r)).$$

This is like the WKB approximation, but now it is argued on the basis of the large incoming energy, rather than a small $\hbar$. The solution to this equation will be taken as the approximation for $\Psi_{\mathbf{k}}(\mathbf{r})$ for the purpose of computing the scattering amplitude

$$\begin{aligned} f(\mathbf{k},\mathbf{k}') &= -\frac{1}{4\pi} \int d^3\mathbf{r}\ e^{-i\mathbf{k}'\cdot\mathbf{r}}\ v(r)\ (2\pi)^{3/2} \Psi_{\mathbf{k}}(\mathbf{r}) \\ &= -\frac{1}{4\pi} \int d^3\mathbf{r}\ e^{-i\mathbf{k}'\cdot\mathbf{r}}\ v(r)\ e^{iS(\mathbf{k},\mathbf{r})/\hbar} \end{aligned}$$

To solve the eikonal equation, consider the sketch below

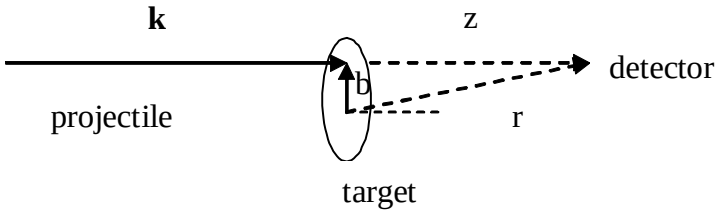


Fig.14.3 - Eikonal approximation

The projectile hits the target at some distance b from its center. Due to its high energy it does not change course very much, so it roughly continues its path along the incident direction, and then it is observed at some distance z from the target, at position $\mathbf{r}$ from the center. The small angle that the vector $\mathbf{r}$ makes with the axis of the target is the scattering angle θ. The vector $\mathbf{r}$ may be written as

$$\mathbf{r} = \mathbf{b} + z\hat{\mathbf{k}},\ r = \sqrt{b^2 + z^2}$$

where $\hat{\mathbf{k}}$ is the direction of the incident particle. We compute $(\boldsymbol{\nabla} S)^2 \left(\mathbf{b} + z\hat{\mathbf{k}}\right)$ in this parametrization. Since $(\boldsymbol{\nabla} S)$ represents the semi-classical momentum we

expect that the derivatives with respect to $\mathbf{b}$ are negligibly small as compared to derivatives with respect to z. Therefore we can write $\left(\boldsymbol{\nabla} S\right)^2 \simeq \left(\partial S/\partial z\right)^2$ and obtain the equation

$$\left(\frac{\partial S}{\partial z}\right)^2 = 2m\left(E - V\left(\sqrt{b^2+z^2}\right)\right)$$

The solution is

$$S(z,b) = \int_{-\infty}^{z} \sqrt{2m(E - V\left(\sqrt{b^2+z'^2}\right)}dz' + const.$$

The constant has to be chosen in such a way that the wavefunction $\Psi_{\mathbf{k}}(\mathbf{r})$ reduces to the the free particle plane wave when the potential vanishes. Therefore we require $\frac{1}{\hbar}S(z) \to kz$ when $V \to 0$, which gives

$$\frac{1}{\hbar}S(z) = kz + \int_{-\infty}^{z}\left[\sqrt{k^2 - v\left(\sqrt{b^2+z'^2}\right)} - k\right]dz'$$

For high energies, $k^2 \gg v(r)$, we can approximate $\sqrt{k^2 - v} \simeq k - \frac{v}{2k}$, so that

$$\frac{1}{\hbar}S(z) \simeq kz - \frac{1}{2k}\int_{-\infty}^{z} v\left(\sqrt{b^2+z'^2}\right)dz'.$$

Hence the wavefunction becomes

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \Psi_{\mathbf{k}}(\mathbf{b} + \hat{\mathbf{k}}z) = \frac{e^{ikz}}{(2\pi)^{3/2}}\ \exp\left(-\frac{i}{2k}\int_{-\infty}^{z} v\left(\sqrt{b^2+z'^2}\right)dz'\right)$$

We are now ready to compute the scattering amplitude using the above outgoing wave-function. Defining

$$\mathbf{k}' = k\hat{\mathbf{r}}, \quad \mathbf{r}' = \mathbf{b} + \hat{\mathbf{k}}z', \quad \hat{\mathbf{r}}\cdot\hat{\mathbf{k}} = \cos\theta,$$

where $\mathbf{r}$ is the location of the detector, and $\mathbf{r}'$ is some other point, we have

$$f(\mathbf{k}',\mathbf{k}) = -\frac{2\pi^2}{(2\pi)^3}\int d^3\mathbf{r}'\ e^{-i\mathbf{k}'\cdot\mathbf{r}'}e^{i\mathbf{k}\cdot\mathbf{r}'}v\left(\sqrt{b^2+z'^2}\right)\exp\left(-\frac{i}{2k}\int_{-\infty}^{z'} v\left(\sqrt{b^2+z''^2}\right)dz''\right) \tag{14.28}$$

where we substituted $e^{ikz'} = e^{i\mathbf{k}\cdot\mathbf{r}'}$. We replace $\mathbf{r} = \mathbf{b} + \hat{\mathbf{k}}z'$ and $d^3\mathbf{r}'$=$d^2\mathbf{b}dz'$, furthermore note that the factor of v in the integrand may be rewritten by taking a derivative of the exponential, as follows

$$f(\mathbf{k}',\mathbf{k}) = -\frac{1}{4\pi}\int d^2\mathbf{b}e^{-i\left(\mathbf{k}'-\mathbf{k}\right)\cdot\mathbf{b}}\int_{-\infty}^{\infty}dz\, e^{ikz(1-\hat{\mathbf{k}}'\cdot\hat{\mathbf{k}})}\left(\frac{2k}{-i}\right)\frac{\partial}{\partial z'}\exp\left(-\frac{i}{2k}\int_{-\infty}^{z'} v\left(\sqrt{b^2+z''^2}\right)dz''\right).$$

For small scattering angle θ we can approximate $e^{ikz(1-\hat{\mathbf{k}}'\cdot\hat{\mathbf{k}})} \approx e^{i0} = 1$. Then the integrand over z becomes a total derivative, so the integration is trivial and we get contributions from the boundaries at $z = \pm\infty$ in the form $\left[\exp\left(-\frac{i}{2k}\int_{-\infty}^{z'} v\left(\sqrt{b^2+z''^2}\right)dz''\right)\right]_{-\infty}^{\infty}$. Defining the phase

$$\Delta(b) = -\frac{1}{4k}\int_{-\infty}^{\infty} dz' v\left(\sqrt{b^2+z'^2}\right) = -\frac{m}{2k\hbar^2}\int_{-\infty}^{\infty} dz' V\left(\sqrt{b^2+z'^2}\right),$$

we can write the expression

$$f(\mathbf{k}',\mathbf{k}) = \frac{k}{2\pi i}\int d^2\mathbf{b}\; e^{-i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{b}}\left[e^{2i\Delta(b)}-1\right].$$

Next we write $(\mathbf{k}'-\mathbf{k})\cdot\mathbf{b} = b\,|\mathbf{k}'-\mathbf{k}|\cos\phi$, $d^2\mathbf{b} = bdbd\phi$, and $|\mathbf{k}'-\mathbf{k}| = 2k\sin\theta/2$, and perform the integral over ϕ by using

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-b|\mathbf{k}'-\mathbf{k}|\cos\phi}d\phi = J_0(2kb\sin\frac{\theta}{2}).$$

So we obtain

$$f(E,\theta) = -ik\int_0^{\infty} bdbJ_0(2kb\sin\frac{\theta}{2})\left[e^{2i\Delta(b)}-1\right] \tag{14.29}$$

Here b is called the impact parameter. The integration over b gets its main contribution up to a distance r_0, which represents the range of the potential $V(r)$, because $\Delta(b)$ vanishes for $b \gg r_0$.

This result was derived under the assumption of a small scattering angle θ, therefore the argument of the Bessel function should be $2kb\sin\frac{\theta}{2} \approx kb\theta$. However, it is possible to reconsider the derivation more carefully and show that the final expression given above is valid also for large values of the angle. Hence the expression above for any angle is the scattering amplitude in the eikonal approximation.

At high energies we can relate the eikonal and Born approximations by examining Eq.(14.28). In the large k limit we can drop the last phase factor

$$\exp\left(-\frac{i}{2k}\int_{-\infty}^{z'} v\left(\sqrt{b^2+z''^2}\right)dz''\right) \to 1. \tag{14.30}$$

Then Eq.(14.28) becomes precisely the first Born approximation. We can thus conclude that the eikonal approximation and the 1^{st} Born approximation are equivalent at high energies.

We will later see that the eikonal approximation may also be related to the partial wave scattering amplitudes in the limit of high energy scattering. From this relation, and also from a direct computation, one can verify that the Eikonal approximation satisfies the optical theorem

14.8 Partial waves

Let us consider scattering from a rotationaly invariant potential $V = V(r)$. In our previous discussion the beam of incoming particles was represented by the momentum ket $|\mathbf{k}\rangle$. Since the Hamiltonian $H_0 = \mathbf{p}^2/2m$ commutes with $\mathbf{L}^2$ and L_z, the free particle can also be represented in the angular momentum and energy basis as $|klm\rangle$. One basis has an axpansion in terms of the other

$$|\mathbf{k}\rangle = \sum_{lm} |klm\rangle\langle lm|\hat{\mathbf{k}}\rangle = \sum_{lm} |klm\rangle Y_{lm}\left(\hat{\mathbf{k}}\right).$$

where $E = \hbar^2 k^2/2m$. The total Hamiltonian, including the potential is also invariant under rotations. By the same arguments, the exact eigenstate $|\mathbf{k}\rangle^+$ can also be expressed in the angular momentum basis $|klm\rangle^+$, and have the expansion

$$|\mathbf{k}\rangle^+ = \sum_{lm} \quad |klm\rangle^+\langle lm|\hat{\mathbf{k}}\rangle = \sum_{lm} |klm\rangle^+ Y_{lm}\left(\hat{\mathbf{k}}\right).$$

Since $\mathbf{L}$ commutes with the total Hamiltonian, angular momentum is conseerved during the scattering process.

Let us see what is the advantage of using the angular momentum basis. Consider the following sketch of a beam of particles hitting the target of a size determined by a radius a. Semi-classically, the radius a represents the range of the interaction potential.

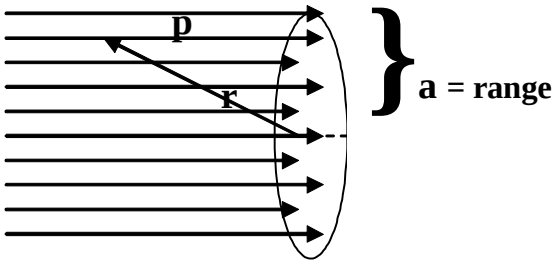


Fig.14.4 - Angular momentum of incoming particle.

The angular momentum of the particle is $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. The length of this vector is $L = rp\sin\theta = pb$, where $b = r\sin\theta$ is the impact parameter. When the impact parameter exceeds the range a the particle cannot be scattered. Therefore, the maximum angular momentum in the scattering process is determined semi-classically as $L_{\max} \sim pa = \hbar ka = \hbar l_{\max}$. This leads us to expect that, in the expansion in the angular momentum basis, mainly the modes of angular momentum $l \leq l_{\max}$ will contribute to the scattering amplitude. Hence, for small energies (k small) and/or small ranges of the potential, ka is a small

number, and thus only few "partial waves" are expected to play an important role in the expansion.

To obtain quantitative results, start from the Lippmann-Schwinger equation $|\mathbf{k}\rangle^+ = |\mathbf{k}\rangle + G_k v|\mathbf{k}\rangle^+$. After substituting the angular momentum expansion for each state, and matching the coefficients of $Y_{lm}\left(\hat{\mathbf{k}}\right)$ on both sides, this equation reduces to

$$|klm\rangle^+ = |klm\rangle + G_k v|klm\rangle^+,\ G_k = \left(k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon\right)^{-1}$$

since the operators G, v are rotationally invariant. The bra $\langle\mathbf{r}|$ can also be expanded in angular momentum basis $\langle\mathbf{r}| = \sum_{lm}\langle\hat{\mathbf{r}}|lm\rangle\langle rlm| = \sum_{lm} Y^*_{lm}(\hat{\mathbf{r}})\langle rlm|$. Taking a dot product with the bra $\langle rlm|$ the Lippmann-Schwinger equation becomes

$$\langle rlm|klm\rangle^+ = \langle rlm|klm\rangle + \langle rlm|G_k v(r)|klm\rangle^+$$

The values of l, m must be the same for the kets and bra since otherwise we get zero due to orthogonality. Next we insert the identity operator $1 = \int_0^\infty r'^2 dr' \sum_{l'm'} |r'l'm'\rangle\langle r'l'm'|$ and again use conservation of angular momentum for rotationally invariant uperators to argue that only $l' = l$ and $m' = m$ can contribute. In fact every term is independent of m, and the equation takes the form

$$\langle rl|kl\rangle^+ = \langle rl|kl\rangle + \int_0^\infty r'^2 dr' \langle rl|G_k|r'l\rangle v(r')\langle r'l|kl\rangle^+.$$

Thus the Lippmann-Schwinger equation becomes

$$A_l(k,r) = j_l(kr) + \int_0^\infty r'^2 dr' G^l_k(r,r') v(r') A_l(k,r') \tag{14.31}$$

where $G^l_k(r,r') \equiv \langle rl|G_k|r'l\rangle$ is the Green function, and the quantities $\langle rl|kl\rangle$ and $\langle rl|kl\rangle^+$ are just the radial wavefunctions for the free particle and the interacting particle respectively. For the free particle we consult section (6.8) to remind ourselves that $\langle rl|kl\rangle$ is written in terms of the spherical Bessel function as

$$\langle rl|kl\rangle = \sqrt{\pi/2} j_l(kr).$$

It satisfies the radial equation

$$\left(-\frac{1}{r}\partial_r^2 r + \frac{l(l+1)}{r^2} - k^2\right) j_l(kr) = 0.$$

Similarly, for the interacting particle we define $\langle rl|kl\rangle^+ \equiv \sqrt{2/\pi} A_l(kr)$, where $A_l(k,r)$ satisfies the radial equation including the potential

$$\left[-\frac{1}{r}\partial_r^2 r + \frac{l(l+1)}{r^2} - v(r) - k^2\right] A_l(k,r) = 0.$$

Next we compute the Green function. We insert identity in energy and angular momentum space $1 = \int_0^\infty k'^2 dk' \sum_{l'm'} |k'l'm'\rangle\langle k'l'm'|$ and again argue

that only $l' = l$ and $m' = m$ can contribute due to the rotation invariance of the operator G_k. Therefore

$$G_k^l(r,r') \equiv \langle rl|G_k|r'l\rangle = \int_0^\infty k'^2 dk' \ \langle rl| \left(k^2 - \mathbf{p}^2/\hbar^2 + i\varepsilon\right)^{-1} |k'l\rangle\langle k'l|r'l\rangle$$

The operator $\mathbf{p}^2$ is diagonal and gives the eigenvalue $\hbar^2 k'^2$, and we get

$$G_k^l(r,r') = \int_0^\infty k'^2 dk' \frac{\langle rl|k'l\rangle\langle k'l|r'l\rangle}{k^2 - k'^2 + i\varepsilon} = \frac{2}{\pi}\int_0^\infty k'^2 dk' \frac{j_l(k'r)j_l(k'r')}{k^2 - k'^2 + i\varepsilon}.$$

It can be checked that this expression for the Green function is a solution to the equation

$$\left(\frac{1}{r}\partial_r^2 r - \frac{l(l+1)}{r^2} + k^2\right) G_k^l(r,r') = \frac{1}{r^2}\delta(r-r').$$

To obtain an explicit expression for $G_k^l(r,r')$, we need to compute the integral. Since the integrand is even we can write

$$G_k^l(r,r') = \frac{1}{\pi}\int_{-\infty}^\infty q^2 dq \frac{j_l(qr_>)j_l(qr_<)}{k^2 - q^2 + i\varepsilon},$$

where $r_<$ is a symbol to indicate the smaller among r and r' and similarly $r_>$ indicates the larger among r and r'. We will perform the integral by using complex integration techniques in the complex q plane. The function $j_l(z)$ can be written in terms of the spherical Bessel functions $h_l(z), h_l^*(z)$ which have the asymptotic behavior in the complex z plane indicated below

$$j_l(z) = \frac{1}{2}\left(h_l(z) + h_l^*(z)\right) \underset{|z|\to\infty}{\to} \frac{e^{iz}}{2iz} - \frac{e^{-iz}}{2iz}$$

This behavior can be used to perform the integral as follows. We substitute $j_l(qr_>) = \frac{1}{2}\left(h_l(qr_>) + h_l^*(qr_>)\right)$. For the term involving $h_l(qr_>)$ the contour $-\infty < q < \infty$ can be closed in the upper half plane, and similarly for the term involving $h_l^*(qr_>)$ the contour can be closed in the lower half plane. The asymptotic behavior of $j_l(qr_<)$ in each of these regions is overwhelmed by $j_l(qr_>)$ therefore it can be ignored in the argument for the choice of countours at the infinite circles. The upper half plane contour encloses the ploe at $q = k + i\varepsilon$ in the clockwise direction, while the lower half plane contour encloses the ploe $q = -k - i\varepsilon$ in the counterclockwise direction. Then the result of the integral can be written as as the sum of the residues

$$\begin{aligned}
G_k^l(r,r') &= 2\pi i \ \mathrm{Res}\left(k + i\varepsilon\right) - 2\pi i \ \mathrm{Res}\left(-k - i\varepsilon\right) \\
\mathrm{Res}\left(k + i\varepsilon\right) &= \frac{1}{2}\frac{1}{\pi}\frac{k^2 j_l(kr_<)h_l(kr_>)}{-2k}, \\
\mathrm{Res}\left(-k - i\varepsilon\right) &= \frac{1}{2}\frac{1}{\pi}\frac{k^2 j_l(-kr_<)h_l^*(-kr_>)}{2k}.
\end{aligned}$$

After using properties of the Bessel functions one finds that the second term is equal to the first term, so we obtain finally

$$G_k^l(r,r') = -ik\ j_l(kr_<)\ h_l(kr_>).$$

The Lippmann-Schwinger equation in Eq.(14.31) for $A_l(k,r)$ now takes the form

$$\begin{aligned} A_l(k,r) &= j_l(kr) - ik\int_0^\infty dr' r'^2 j_l(kr_<)h_l(kr_>)v(r')A_l(kr') \\ &= j_l(kr) - ikh_l(kr)\int_0^r dr' r'^2 j_l(kr')v(r')A_l(kr') \\ &- ikj_l(kr)\int_r^\infty dr' r'^2 h_l(kr')v(r')A_l(kr') \end{aligned}$$

We can now compute the asymptotic limit of this expression for large r by using the asymptotic behavior. The second term vanishes so, after substituting $j_l(kr) = \frac{1}{2}\left(h_l(kr) + h_l^*(kr)\right)$, the wavefunction $A_l(k,r)$ takes the form $A_l(k,r) \underset{r\to\infty}{\to} \frac{1}{2}h_l^*(kr) + \frac{1}{2}h_l(kr)S_l(k)$ where

$$S_l(k) = 1 - 2ik\int_0^\infty dr' r'^2 j_l(kr')v(r')A_l(k,r'). \tag{14.32}$$

Thus, using the asymptotic limit of the Bessel functions $h_l(kr)$ and $h_l^*(kr)$ given above we have the asymptotic behavior

$$A_l(k,r) \underset{r\to\infty}{\to} -\frac{1}{2ikr}\left[e^{-i(kr-l\pi/2)} - e^{i(kr-l\pi/2)}S_l(k)\right]. \tag{14.33}$$

The general form is $A_l(k,r) \to a_l e^{-ikr}/kr + b_l e^{ikr}/kr$, where the first/second term is proportional to an incoming/outgoing spherical wave. Therefore, $S_l(k)$ is identified as the ratio of the coefficients of the outgoing and incoming spherical waves modulo the extra signs $(-1)^{l+1}$ that arise from the l dependent phases, that is $S_l(k) = (-1)^{l+1}\, b_l/a_l$.

We will now prove that $S_l(k)$ is a pure phase. From the arguments leading to the optical theorem we know that the probability current density satisfies $r^2\int \hat{\mathbf{r}}\cdot\mathbf{J}d\Omega = 0$. The radial current is computed as $\hat{\mathbf{r}}\cdot\mathbf{J} \propto i(A_l^*(k,r)\partial_r A_l(k,r) - A_l(k,r)\partial_r A_l^*(k,r))$ where we insert the asymptotic form of $A_l\,(k,r)$ given above and keep only the leading terms for large r. Then we obtain

$$0 = r^2\int \hat{\mathbf{r}}\cdot\mathbf{J}d\Omega \propto 2ik(|S_l|^2 - 1).$$

From this we conclude that S_l is a pure phase. It must be emphasized that this result follows from conservation of probability. Thus we can write

$$S_l(k) = e^{2i\delta_l(k)} = 1 - 2ik\int_0^\infty dr r^2 j_l(kr)v(r)A_l(k,r) \tag{14.34}$$

where $\delta_l(k)$ is real and is called the phase shift. If the potential $v\left(r\right)$ vanishes the phase shift also vanishes. Therefore the phase shift in the outgoing spherical wave relative to the incoming spherical wave is a measure of the interaction. Note that this expression corresponds to the matrix element below

$$S_l(k) = 1 - \frac{ik}{\pi}\langle kl|v\left(r\right)|kl\rangle^+ = e^{2i\delta_l(k)}. \tag{14.35}$$

We now turn to the computation of the scattering amplitude. Recall that

$$f(k,\theta) = -2\pi^2\langle \mathbf{k}'|v|\mathbf{k}\rangle^+.$$

We insert identity in position space and use the definitions of the various dot products and other quantities defined above

$$\begin{aligned} f(k,\theta) &= -2\pi^2\sum_{lm}\int_0^\infty r^2 dr\langle \mathbf{k}'|rlm\rangle\langle rlm|v(r)|\mathbf{k}\rangle^+ \\ &= -2\pi^2\frac{2}{\pi}\sum_l\int_0^\infty r^2 dr j_l(kr)v\left(r\right)A_l(kr)\sum_m\langle\hat{\mathbf{k}}'|lm\rangle\langle lm|\hat{\mathbf{k}}\rangle \\ &= -4\pi\sum_l\frac{1-S_l(k)}{2ik}\sum_m Y_{lm}^*(\hat{\mathbf{k}}')Y_{lm}(\hat{\mathbf{k}}) \\ &= \sum_l\frac{S_l(k)-1}{2ik}P_l(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')(2l+1) \end{aligned}$$

Hence the angular momentum expansion of the scattering amplitude is obtained in terms of the Legendre polynomials $P_l(\cos\theta)$

$$f(k,\theta) = \sum_l(2l+1)f_l(k)P_l(\cos\theta) \tag{14.36}$$

with $f_l(k)$ written in various forms in terms of the phase shift

$$f_l(k) = \frac{e^{2i\delta_l(k)}-1}{2ik} = \frac{1}{k}e^{i\delta_l(k)}\sin\delta_l\left(k\right) = \frac{1}{k}\frac{1}{\cot\delta_l\left(k\right)-i}. \tag{14.37}$$

From these expressions we learn that all possible values of $kf_l(k)$ must lie on a circle centered at $i/2$ in the complex plane, as seen from the form

$$kf_l(k) = \frac{i}{2} + \frac{1}{2}e^{2i\delta_l(k)-i\pi/2}$$

and as shown in the figure below which is called an Argand diagram

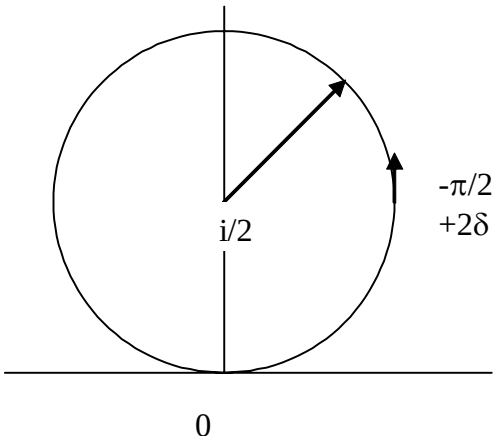


Fig.14.5 - Argand diagram.

The point on the circle at the tip of the arrow is the value of the complex number $\frac{i}{2} + \frac{1}{2}\exp(2i\delta_l(k) - i\pi/2)$. The arc segment from the point zero to the tip of the arrow spans the angle $2\delta_l(k)$. The tip of the arrow changes as a function of k but always remains on the circle for elastic scattering for which probability is conserved. If probability is not conserved, as in inelastic scattering (discussed later), then $kf_l(k)$ is no longer on the circle, but moves into the interior of the circle. By doing experimental measurements it is possible to obtain the trajectory of the point on the Argand diagram as function of the energy for either elastic or inelastic scattering.

14.8.1 Differential cross section

The differential cross section was given before as $\frac{d\sigma}{d\Omega} = |f(k,\theta)|^2$. In terms of the partial waves it becomes

$$\frac{d\sigma}{d\Omega} = \frac{1}{k^2}\sum_{ll'}(2l+1)(2l'+1)e^{i(\delta_l - \delta_{l'})} \times \sin\delta_l \sin\delta_{l'} P_l(\cos\theta) P_{l'}(\cos\theta)$$

where $f_l(k) = \frac{1}{k}e^{i\delta_l}\sin\delta_l$ has been used. This expression becomes useful when there are a few partial waves which dominate the sum, as in the case of low energy scattering for which $l_{\max} = ka$ could be small, as argued at the beginning of section (14.8).

The total cross section $\sigma_{tot} = \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta)\left(\frac{d\sigma}{d\Omega}\right)$ simplifies thanks to the orthogonality of the Legendre polynomials

$$\int_{-1}^{1} P_l(x)P_{l'}(x)dx = \frac{2}{2l+1}\delta_{ll'}.$$

Hence,

$$\sigma_{tot}(k) = \frac{4\pi}{k^2}\sum_l (2l+1)\sin^2\delta_l(k).$$

Again, this is a useful expression when a few partial waves dominate the sum.

We will now prove that the optional theorem is satisfied by the expression for $f(k,\theta)$ in Eq.(14.36). In the forward direction we have $P_l(\cos 0) = 1$, so that

$$f(k,0) = \frac{1}{k}\sum_l (2l+1)e^{i\delta_l}\sin\delta_l.$$

The imaginary part of $f(k,0)$

$$\mathrm{Im}\left(f(k,0)\right) = \frac{1}{k}\sum_l (2l+1)\sin^2\delta_l$$

is clearly proportional to the total cross section given above, and indeed satisfies the optical theorem

$$\sigma_{tot} = \frac{4\pi}{k}\mathrm{Im}\left(f(k,0)\right).$$

14.8.2 Phase shift δ_l

We have seen that the asymptotic behavior of the radial wavefunction takes the form

$$A_l(kr) \underset{r\to\infty}{\to} \frac{-1}{2ikr}\left\{e^{-i(kr-l\pi/2)} - e^{i(kr-l\pi/2+2\delta_l)}\right\}$$

where the outgoing spherical wave is phase shifted by the additional phase $2\delta_l$ compared to the free particle. Thus δ_l is a measure of the interaction. Here we want to prove that the phase shift $\delta_l(k)$ is a decreasing function of l. This point is intuitively understood through Fig.14.4, which suggests that there is little or no scattering when $l > l_{\max} = ka$. Hence at a fixed k, as l increases the phase shift must decrease since we approach and then surpass $l_{\max}$. To see this point more formally consider the radial differential equation for $A_l(kr)$

$$\left[\frac{1}{r}\partial_r^2 r - \left(\frac{l(l+1)}{r^2} + v(r)\right) + k^2\right] A_l(k,r) = 0.$$

As l increases the effective potential $\frac{l(l+1)}{r^2} + v(r)$ gets more and more dominated by the $\frac{l(l+1)}{r^2}$ term. Therefore the radial equation at large l becomes more and more like the one for the free particle, which implies that for large l the radial wavefunction $A_l(k,r)$ approaches the free particle wavefunction $j_l(kr)$. Thus, at large l we may approximate Eq.(14.34), which determines the phase shift, by the following expression

$$e^{2i\delta_l} \simeq (1 + 2i\delta_l + \cdots) \simeq 1 - 2ik\int_0^\infty dr r^2 j_l(kr)v(r)j_l(kr)$$

If the range of $v(r)$ is small, the contributions to the integral come from the vicinity of $r \sim 0$, where we can use $j_l(kr) \to (kr)^l/(2l+1)!!$. Therefore for large l we obtain the expression

$$\delta_l \simeq -\frac{k^{2l+1}}{\left[(2l+1)!!\right]^2}\int_0^\infty r^{(2l+2)}v(r)dr.$$

Now suppose $V(r)$ is approximated by a constant V_0 within the range $r < a$. Then we obtain

$$\delta_l \simeq -\frac{2mV_0a^2}{\hbar^2}\frac{(ka)^{2l+1}}{(2l+3)\left[(2l+1)!!\right]^2}.$$

This expression decreases rapidly for $l \ll ka$, showing that indeed there is a maximum angular momentum $l_{\max} \sim ka$ beyond which the phase shift vanishes.

14.8.3 Low energy scattering and partial waves

Consider the low energy scattering where very few partial waves contribute to the total cross section, as described above. In particular suppose $ka < 1$ so that only the $l = 0$ partial wave contributes. Then the total cross section is determined just by $\delta_0(k)$

$$for\ k \to 0: \quad \frac{d\sigma}{d\Omega} \simeq \frac{1}{k^2}\sin^2\delta_0, \quad \sigma_{tot} \simeq \frac{4\pi}{k^2}\sin^2\delta_0. \tag{14.38}$$

This differential cross section is totally independent of the direction of observation, thus the scattering at low energies is isotropic. As the energy increases the angular dependence becomes more and more significant.

14.8.4 Connection with the eikonal approximation

We now examine the scattering amplitude

$$f(k,\theta) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\left[e^{2i\delta_l(k)}-1\right]$$

for large energies. Since $(ka) \sim l_{\max}$ is large, we can approximate the sum over l by a continuous integral over a continuous variable l; therefore,

$$f(k,\theta) = \frac{1}{2ik}\int_0^\infty dl(2l+1)P_l(\cos\theta)\left[e^{2i\delta_k(k)}-1\right]$$

This integral my be written in terms of the impact parameter by writing $l = kb$ as explained in the discussion related to Fig.14.4. Then $dl = kdb$ and $(2l+1) \approx 2kb$, and $P_l(\cos\theta) = P_{kb}(\cos\theta)$. Now, for large k and small θ, in the regime $kb\theta =$ finite, the Legengre polynomial becomes the zeroth Bessel function

$$P_{kb}(\cos\theta) \underset{k=big,\theta=small}{\to} J_0(kb\theta).$$

Furthermore we can write $\delta_l(k) = \delta_{kb}(k) = \Delta(k,b)$ and obtain

$$f(k,\theta) = -ik\int_0^\infty bdbJ_0(kb\theta)\left[e^{2i\Delta(k,b)}-1\right]$$

This expression is the scattering amplitude in the *eikonal approximation* when $\delta_l(r)$ is replaced by $\Delta_l(b)$.

Now, since the optical theorem is satisfied in the partial wave method we can conclude that it is also satisfied in the eikonal approximation.

14.8.5 Finite range potential

Consider the radial differential equation

$$\left(\frac{1}{r}\partial_r^2 r - \frac{l(l+1)}{r^2} - v(r) + k^2\right)A_l(k,r) = 0$$

The boundary conditions to the above equations are

$$\begin{aligned} r\to 0 \qquad & A_l(k,r) \;\to\; r^l \\ r\to\infty \qquad & A_l(k,r) \to a_l\frac{e^{-ikr}}{r} + b_l\frac{e^{ikr}}{r} \end{aligned}$$

The coefficients $a_l(k), b_l(k)$ are computed in principle by solving the differential equation. As we have learned their ratio gives the phase shift

$$\exp\left(2i\delta_l(k)\right) = (-1)^{l+1}\frac{b_l}{a_l}.$$

Consider a finite range potential that vanishes outside of a range a. The vanishing could be smooth, but for simplicity we will assume an abrubt drop to zero potential. An example is the spherically symmetric well or barrier. Then, the radial differential equation takes the form

$$\begin{aligned} \left(\tfrac{1}{r}\partial_r^2 r - \tfrac{l(l+1)}{r^2} + k^2\right)A_l^{(out)}(k,r) = 0 \qquad & r > a \\ \left(\tfrac{1}{r}\partial_r^2 r - \tfrac{l(l+1)}{r^2} - v(r) + k^2\right)A_l^{(in)}(kr) = 0 \qquad & r < a \end{aligned}$$

The solutions are given as follows. For $r > a$ the outside solution is written in terms of the free particle spherical Bessel functions

$$r > a : A_l^{out}(kr) = c_l\left(h_l^*(kr) + e^{2i\delta_l(k)}h_l(kr)\right)$$

where we have already identified the phase shift as one of the ratio of the two terms as explained before. This can also be written in the form

$$A_l^{out}(r) = c_l e^{i\delta_l}\left[\cos\delta_l j_l(kr) - \sin\delta_l n_l(kr)\right]$$

Similarly there is some solution $A_l^{in}(r)$ for $r < a$, which is assumed to satisfy the correct boundary condition at $r = 0$. The logarithmic derivatives of the inside and outside solutions must match at $r = a$.

$$\left(\frac{1}{A_l^{out}} r\partial_r A_l^{out}\right)_{r=a} = \left(\frac{1}{A_l^{in}} r\partial_r A_l^{in}\right)_{r=a} \equiv \gamma_l^{in}$$

The left side of this equation gives

$$\left(\frac{1}{A_l^{out}} r\partial_r A_l^{out}\right)_{r=a} = ka\frac{\cos\delta_l \; j_l'(ka) - \sin\delta_l \; n_l'(ka)}{\cos\delta_l \; j_l(ka) - \sin\delta_l \; n_l(ka)}$$

where $j_l'(z) = \partial_z j_l(z)$ and $n_l'(z) = \partial_z n_l(z)$. From the matching condition we can solve for $\delta_l(k)$ and obtain

$$\tan\delta_l = \frac{(ka)\, j_l'(ka) - \gamma_l^{in} j_l(ka)}{(ka)\, n_l'(ka) - \gamma_l^{in} n_l(ka)}.$$

Once γ_l^{in} is computed, this equation determines the phase shift for every l and every k.

14.8.6 Hard Sphere Scattering

Consider the potential

$$V(r) = \begin{cases} 0 & for \quad r > a \\ \infty & for \quad r < a \end{cases}$$

This represents an impenetrable sphere. Therefore the solution inside is zero $A_l^{in}(k, r) = 0$. The matching condition requires $A_l^{out}(k, a) = 0$ at $r = a$. Using the solution obtained above gives

$$\tan\delta_l(k) = \frac{j_l(ka)}{n_l(ka)}, \text{ or } e^{2i\delta_l(k)} = -\frac{h_l^*(ka)}{h_l(ka)}.$$

The phase shifts for this potential are now known for any l. For example, for $l = 0$ we have

$$\tan\delta_0 = \frac{j_0(ka)}{n_0(ka)} = \frac{\sin(ka)/ka}{-\cos(ka)/ka} = -\tan(ka)$$

which gives

$$\delta_0 = -ka.$$

One can show generally that the sign of the phase shift is the opposite sign as the potential. That is, for a repulsive potential the sign is negative and for an attractive potential the sign is positive. In the present case the hard sphere is a repulsive potential, and we see explicitly that the sign computed above is consistent with the expectation.

The total cross section is

$$\begin{aligned}\sigma_{tot} &= \frac{4\pi}{k^2}\sum_l (2l+1)\sin^2\delta_l \\ &= \frac{4\pi}{k^2}\sin^2\delta_0 + \cdots \\ &= \frac{4\pi}{k^2}\sin^2(ka) + \cdots\end{aligned}$$

where the dots $\cdots$ represent the contribution from higher partial waves.

At low energies the $l=0$ term dominates. Therefore we get the low energy limit of the cross section for small ka as

$$\sigma_{tot} \simeq 4\pi a^2.$$

This value is 4 times the classical geometric cross section of a disk πa^2. We see that the quantum effect is very significant at low energies.

At high energies we compute the phase shift $\exp(2i\delta_l(k)) = -h_l^*(ka)/h_l(ka)$ by using the asymptotic forms of the Bessel functions. These differ according to whether $l < ka$ or $l > ka$. When $l > ka$ the asymptotic limit of $h_l(ka)$ is real and therefore $\delta_l(k) \to 0$. On the other hand for $l < ka$ we find the asymptotic limit $h_l(ka) \to \exp(ika - i(l+1)\pi/2)/ka$, and this gives $\exp(2i\delta_l(k)) \underset{k\to\infty}{\to} \exp\left(2i\left(-ka + l\frac{\pi}{2}\right)\right)$, or

$$\delta_l(k) \underset{k\to\infty}{\to} -ka + l\frac{\pi}{2} \text{ for } l < ka.$$

This gives the high energy cross section

$$\sigma_{tot} = \frac{4\pi}{k^2}\sum_{l=0}^{ka}(2l+1)\sin^2\left(ka - l\frac{\pi}{2}\right).$$

To get an estimate of the sum we may replace $\sin^2\left(ka - l\frac{\pi}{2}\right)$ by the average value of the $\sin^2$ function, which is $1/2$, and replace the sum by an integral. Then we obtain

$$\sigma_{tot} = \frac{4\pi}{k^2}\frac{1}{2}\int_0^{ka} 2l\ dl \simeq \frac{4\pi}{k^2}\frac{(ka)^2}{2} = 2\pi a^2.$$

This value of the high energy quantum cross section is still twice as big as the classical geometrical cross section.

To understand the origin of the factor of two consider the scattering amplitude at high energies

$$f(k,\theta) = \frac{1}{2ik}\sum_{l=0}^{ka}(2l+1)\left((-1)^l\exp(-2ika) - 1\right)P_l(\cos\theta)$$

The differential cross section $d\sigma/d\Omega = |f(k,\theta)|^2$ for this expression is plotted in Fig.14.6. We see a very sharp peak in the forward direction at small values of θ as shown in the figure, followed by a uniform smooth curve for most of the range from small θ up to $\theta = \pi$. The hight of the forward peak is $[d\sigma/d\Omega]_{\theta=0} = ka^3/2$, while the average hight of the smooth part is $a^2/4$

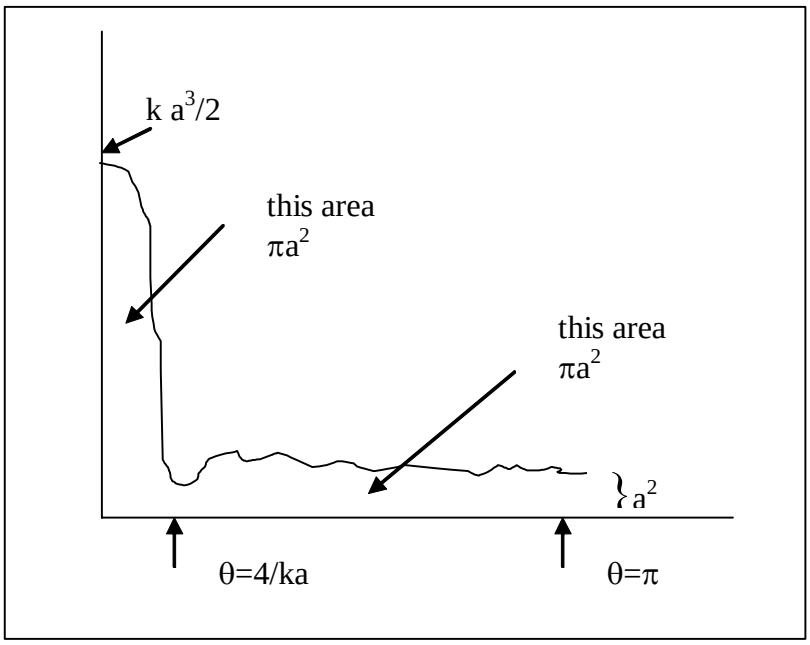


Fig. 14.6 - Forward peak.

The width of the forward peak vanishes as $\Delta\theta \sim 4/ka$, but the area under the forward peak is πa^2 and does not vanish as k increases. The area under the curve for most of the full range is also πa^2. Thus the total cross section of $2\pi a^2$ is accounted for by the forward peak and the rest of the differential cross section. The classical cross section of πa^2 misses the forward peak.

14.9 Inelastic scattering

14.9.1 Unitarity

14.9.2 Black holes

Inelastic scattering

If the target is not just a point particle, but has some structure, such as a proton made of quarks, or a nucleus made of nucleons, etc., then, besides elastic scattering, there are additional phenomena that happen in a collision. The target may get excited to a new state, the initial particle may be absorbed and a new particle may come out, or many particles may emerge in the final state

in a high energy collision, etc.. When all phenomena of both target and projectiles are taken into account the overall probability or unitarity is conserved, however the probability of one type of species of particles appears not to be conserved by itself. In potential scattering theory (which is the formalism used here)phenomena such as particle creation and annihilation cannot be described since one deals with the wavefunction $\psi(\mathbf{r},t)$ of a single species of particles. Since one can keep track only of one type of particle,probability or unitarity will appear to be violated when one attempts to describe inelastic scattering in this formalism. Effectively, one can try to describe the violation of unitarity as being due to a complex potential $V(r)$.

Elastic and inelastic cross sections

Here we will assume that spherical symmetry is valid, so one can analyze the scattering in terms of partial waves. We will also assume that outside the target region the potential decreases to zero, so that at fixed l the outgoing wave is a spherical wave for a free particle. The scattering amplitude $f(k,\theta)$ has both an elastic and an inelastic part. The elastic part is the one that corresponds to the one species of particles that is present in the in-going beam and then detected by an observer after the scattering. The wavefunction $\psi(\mathbf{r},t)$ that describes the probability amplitude of this particle satisfies the Schrödinger equation with a complex potential $V(r)$. Therefore we may apply the same reasoning as before, and derive that the wavefunction in spherical coordinates

$$\psi_E(\mathbf{r},t) = \sqrt{\frac{2}{\pi}} \sum_{lm} A_l^{(elastic)}(k,r)\, Y_{lm}(\theta,\phi) \tag{14.39}$$

has the asymptotic behavior that was derived in the previous formalism

$$A_l^{(elastic)}(kr) \underset{r\to\infty}{\to} \frac{-1}{2ikr}\left\{e^{-i(kr-l\pi/2)} - S_l(k)\, e^{i(kr-l\pi/2)}\right\}. \tag{14.40}$$

The only difference is that when there is inelastic scattering the amplitude of the outgoing wave is less than one

$$|S_l(k)| < 1. \tag{14.41}$$

The elastic scattering amplitude is given in terms of $S_l(k)$ as before

$$f^{el}(k,\theta) = \frac{1}{2ik} \sum_l (2l+1)\, P_l(\cos\theta)\, (S_l(k)-1) \tag{14.42}$$

and the elastic cross section is

$$\begin{aligned} \sigma_{el}(k) &= \int d\Omega \left| f^{el}(k,\theta)\right|^2 \\ &= \frac{\pi}{k^2} \sum_l (2l+1)\, \left|S_l(k)-1\right|^2 \end{aligned} \tag{14.43}$$

On the other hand we can obtain the total cross section by following the same steps as in the derivation of the optical theorem. We had seen that the number of particles taken out of the initial beam per second is given in terms of the scattering amplitude in the forward direction. In the present case the same reasoning gives

$$\left(\begin{array}{c} number\ of\ particles\ taken \\ out\ of\ the\ initial\ beam \end{array} \right) = \frac{v_k}{(2\pi)^{3/2}} \frac{4\pi}{k}\ \mathrm{Im}\left[f^{el}(k,0)\right] \tag{14.44}$$

The initial flux $v_k/(2\pi)^{3/2}$ is also given as before in terms of the velocity $v_k = \hbar k/m$. The ratio of these quantities is the total cross section, therefore

$$\begin{aligned} \sigma_{tot}(k) &= \frac{4\pi}{k}\ \mathrm{Im}\left[f^{el}(k,0)\right] \\ &= \frac{2\pi}{k^2}\sum_l (2l+1)\ \left[1 - \mathrm{Re}\left(S_l(k)\right)\right]\ . \end{aligned} \tag{14.45}$$

Finally, using the fact that the total cross section is the sum of the elastic and inelastic cross sections

$$\sigma_{tot} = \sigma_{el} + \sigma_{inel} \tag{14.46}$$

we derive the inelastic cross section

$$\sigma_{inel}(k) = \frac{\pi}{k^2}\sum_l (2l+1)\ \left[1 - |S_l(k)|^2\right]\ . \tag{14.47}$$

If the phase shift $\delta_l(k)$ defined by

$$S_l(k) = e^{2i\delta_l(k)} = 1 - 2ik\int_0^\infty dr\ r^2\, j_l(kr)\, v(r)\, A_l(k,r) \tag{14.48}$$

were real, then $|S_l(k)|$ would have been 1 and the inelastic cross section would have vanished. This is certainly the case when the potential is real, as we have proven before. But for a complex potential $V(r)$ the phase shift is complex and the inelastic cross section is non-zero.

Black sphere scattering

As an example of inelastic scattering consider an ideal black sphere. It has the property of absorbing any particle that touches it. Particles that do not touch it continue their travel unaffected. This is similar to a black hole that swallows all matter that come within its horizon. Thus, the potential $V(r)$ is such that

$$V(r) = \left\{ \begin{array}{cc} r < a: & complete\ \ absorbtion \\ r \rangle a: & V = 0 \end{array} \right. \tag{14.49}$$

To take this property into account consider angular momentum semi-classically

$$\begin{aligned} \vec{L} &= \mathbf{r} \times \mathbf{p} \\ \hbar l &\approx b \times \hbar k \end{aligned} \tag{14.50}$$

where b is the impact parameter $b = r\left|\hat{\mathbf{r}} \times \hat{\mathbf{p}}\right|$. If the impact parameter is smaller than the radius a of the black sphere then the particle will be completely absorbed since it will touch it, and if it is larger than a there will be no scattering at all. Then the properties of the potential imply the following properties of $S_l(k)$

$$\begin{aligned} l < ka &: \quad S_l(k) = 0 \quad (no \;\; outgoing \;\; wave) \\ l \geq ka &: \quad S_l(k) = 1 \quad (zero \;\; phase \;\; shift) \end{aligned} \tag{14.51}$$

With these we may compute the elastic and inelastic cross sections by plugging the values of $S_l(k)$ in the expressions given above. We obtain

$$\begin{aligned} \sigma_{el} &= \frac{\pi}{k^2}\sum_{l=0}^{ka-1}(2l+1)\;\left|0-1\right|^2 + \frac{\pi}{k^2}\sum_{l=ka}^{\infty}(2l+1)\;\left|1-1\right|^2 \\ &= \frac{\pi}{k^2}\left\{2\frac{(ka-1)\,ka}{2} + ka\right\} + 0 \\ &= \pi a^2 \end{aligned} \tag{14.52}$$

and similarly $\sigma_{inel} = \pi a^2,\ \sigma_{tot} = 2\pi a^2$. Also, the elastic differential cross section is

$$\frac{d\sigma_{el}}{d\Omega} = \frac{1}{4k^2}\left|\sum_{l=0}^{ka-1}(2l+1)\;P_l(\cos\theta)\right|^2 \tag{14.53}$$

Black hole radiation

The classical "black" properties of a black hole derive from the fact that the gravitational attraction is extremely large within its horizon. The horizon is the surface of a sphere centered at the black hole and whose radius is given by the Schwarzchild radius

$$a = \frac{GM}{c^2} \tag{14.54}$$

where G is Newton's gravitational constant, M is the mass of the black hole and c is the velocity of light. Any matter that falls inside its horizon is completely absorbed. Light or other signals cannot be emitted by a black hole because their paths bend under the gravitational attraction so strongly that they cannot escape the black hole. This is why the black hole is "black". However Hawking showed that in quantum mechanics black holes do radiate energy just like a black body, contrary to the expectation in classical mechanics. To understand this quantum mechanical effect we make an analogy between the black sphere scattering described above and the black hole radiation. Classically, the black sphere of the previous section has very similar properties to a black hole. Namely, all matter that comes within the radius a is completely absorbed, and the classical cross section for emission is zero. However, in the quantum calculation we found out that the black sphere does radiate with a

cross section

$$\frac{d\sigma_{el}}{d\Omega} = \frac{1}{4k^2}\left|\sum_{l=0}^{ka-1}(2l+1)\ P_l(\cos\theta)\right|^2 \tag{14.55}$$
$$\sigma_{el} = \pi a^2.$$

This is the analog of the Hawking radiation. This is simultaneously coupled with a violation of unitarity, since $|S_l(k)|$ differs from 1 for $l < ka$. Hawking's computation has raised many puzzles that remain unsolved today. These include the violation of unitarity or the "information puzzle" as well as issues revolving around "black hole thermodynamics". If the black hole radiates like a black body one wonders about its final state. Does it completely disappear or is there a remnant? If it completely disappears then all the information that went in comes out only in the form of thermal radiation, which is white noise devoid of any detailed quantum numbers except for the mass, spin and charge that characterizes a black hole. On the other hand the matter that went in carries all sorts of other quantum numbers. Hence there is loss of information. This is called the"information puzzle" and it is coupled to the violation of unitarity.

There has been a lot of inconclusive debate on what really goes on at the fundamental level. Some, including Hawking, believe that the laws of Quantum Mechanics break down and that they should be changed, others believe that the calculation as well as the physics is incomplete but have not managed to do it right either so far. Here is our point of view at the current time (1996). In the case of inelastic scattering we know that unitarity is re-established by taking into account all microstates of the projectiles and of the target. This requires a change of formalism to describe correctly the physics at the quantum level for the systems that are involved (see below for an example). The analogous steps are required to understand correctly black hole phenomena. However, in the case of the black hole the physical theory that describes its structure is not yet fully available. In the first place, one needs a correct theory for quantum gravity, and we know that a naive quantization of Einstein's General Relativity (which predicts black holes classically, and Hawking's radiation semi-classically) is not the correct quantum theory. Superstring theory is the only theory known to have a consistent description of quantum gravity. Its generalizations that includes "D-branes", which is under development in 1996, seems to have the correct ingredients for solving the puzzles in terms of the microstates of a black hole. These questions are of fundamental importance because it is believed that their resolution would lead us to the fundamental unified theory for all physical phenomena.

14.9.3 Inelastic electron scattering

14.10 Analytic properties of the scattering amplitude

14.10.1 Jost function

14.10.2 Unitarity

14.10.3 Bound states

14.10.4 Regge poles

14.11 Problems

1. Consider the Lippmann-Schwinger formalism for scattering $|\psi\rangle^+ = |\psi\rangle + \frac{1}{k^2-p^2/\hbar^2+i\varepsilon}\frac{2mV}{\hbar^2}|\psi\rangle^+$, and apply it *in one dimension* as follows:

   (a) Derive the x-space version of this equation by computing the Green's function $G(x,x') = \langle x|\frac{1}{k^2-p^2/\hbar^2+i\varepsilon}|x'>$.

   (b) Apply the formalism to the one dimensional problem with an attractive delta function potential $V = -\frac{\hbar^2\gamma}{2m}\delta(x)$, and compute fully the wavefuntion for all x.

   (c) Consider $x > 0$ or $x < 0$ and then give the scattering amplitude for the transmitted and reflected waves (i.e. forward and backward scattering).

   (d) What can you tell about the bound state spectrum by analysing the scattering amplitude that you have derived?

2. Prove the identity in Eq.(14.8) for the expansion of a plane wave in terms of spherical waves at asymptotic values of r. Hint: use the expansion of the free particle wavefunction in an angular momentum basis as given in Eq.(6.115,**??**)

$$\exp(i\mathbf{k}\cdot\mathbf{r}) = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\mathbf{\Omega}_r)Y^*_{lm}(\mathbf{\Omega}_k)\; i^l\; j_l(kr)$$

   and use the asymptotic form of the Bessel functions $j_l(kr)$. Then perform the sum over l, m by using completeness properties of spherical harmonics to obtain the dot products $\langle\hat{\mathbf{r}}| \pm \hat{\mathbf{k}}\rangle = \delta(\Omega_r \mp \Omega_k)$.

3. A particle of mass m is scattered in a 3-dimensional harmonic oscillator potential with a cutoff $V(\mathbf{r}) = \frac{1}{2}m\omega^2\left(\mathbf{r}^2 - a^2\right)\theta(a-r)$ (the form of V is analogous to the finite well potential).

   (a) What are the two possible high energy approximations to the scattering amplitude? Write the expressions for both high energy approximations with all the details (limits, factors, etc.). Perform any integrals that are not hard and simplify your expressions as much as possible.

   (b) What are the two low energy approximations? Briefly mention the underlying reason for why this is a good approximation. What is the range of validity for either one in terms of the parameters of the given potential?

   (c) Compute the $l = 0$ partial wave phase shift by solving the radial Schrödinger equation with the correct boundary conditions.

4. Prove
$$\sigma_{tot} = \frac{m^2}{\pi\hbar^4}\int d^3\mathbf{r}\, d^3\mathbf{r}' V(r)V(r')\frac{\sin^2 k\,|\mathbf{r}-\mathbf{r}'|}{k^2\,|\mathbf{r}-\mathbf{r}'|^2}$$
in each of the following ways

   (a) By integrating the differential cross section computed using the 1^{st} Born approximation.

   (b) By applying the optical theorem, including the 2^{nd} Born approximation to the forward scattering amplitude (since $f^{(1)}(k,0)$ is real in the 1^{st} Born approximation).

5. Consider the delta-shell potential
$$V(r) = V_0\delta(\frac{|\mathbf{r}|}{r_0} - 1).$$

   (a) Compute the scattering amplitude in the first Born approximation. Under what conditions is the approximation valid at high or low energies? What is the width of the resulting high energy diffraction peak?

   (b) Calculate the scattering amplitude in the Eikonal approximation giving explicitly the phase $\Delta(k,b)$ and *carefully* specifying the range of integration for the impact parameter b. What is the range of validity of this approximation ?

   (c) Using the integral equation approach compute the radial wavefuction $A_l(k,r)$ exactly. Furthermore, derive an exact expression for the phase shift $\exp(2i\delta_l(k))$. Then analyze the low energy limit of the scattering amplitude and compare to your results above; is there agreement? What is the low energy limit, and the angular dependence, of the differential cross section?

6. Consider the scattering of a particle from a set of N identical potentials which are centered at points $\mathbf{r}_i,\ i = 1, 2, \cdots, N$, that is $V(\mathbf{r}) = \sum_{i=1}^{N} v(\mathbf{r}-\mathbf{r}_i)$. Show that the differential cross section in lowest Born approximation is given by
$$\frac{d\sigma}{d\Omega} = \frac{d\sigma_0}{d\Omega}\left|\sum_{i=1}^{N} e^{i\mathbf{q}\cdot\mathbf{r}_i}\right|^2,$$
where $\hbar\mathbf{q}$ is the momentum transfer in the collision, and $\frac{d\sigma_0}{d\Omega}$ is the differential cross section for the scattering off one potential. Assume now that these scattering potentials are equally spaced on a line and discuss the angular dependence of the cross section. What is the physical situation that is roughly simulated by this model?

7. Consider the spherical potential well $V(r) = V_0\theta(a-r)$ with either sign of V_0. By solving the differential equation with the correct boundary conditions obtain

(a) an explicit expression for the $l = 0$ phase shift, involving only trigonometric functions.

(b) an expression for arbitrary l, involving spherical bessel functions. In the limit $V_0 \to +\infty$ does your result agree with the one discussed in class for the hard sphere?

(c) How does the sign of the phase shift depend on the sign of V_0 (the sign of the phase shift is directly correlated with the attractive/repulsive nature of the potential, if the potential does not change signs).

8. Consider the delta-shell potential

$$V(r) = V_0 \delta \left(\frac{|\mathbf{r}|}{r_0} - 1 \right).$$

(a) Using either the differential or integral equation approach (recommended) calculate the phase shifts $e^{2i\delta_l}$ for any angular momentum l . Then show that $\delta_l(k) \to \Delta(k, b)$ in the range of validity (use $l \sim kb$, thus identifying $\delta_l(k) \sim \Delta(b, k)$).

(b) Compare the cross sections for 1^{st} Born and partial wave approaches.

9. It can be shown that for the potential $V = -\frac{2ba^2e^{-ar}}{[be^{-ar}+1]^2}$ the solution of the $l = 0$ Schrödinger equation which behaves like an incoming spherical wave is

$$\psi^{(-)}_{l=0}(k, r) = \frac{e^{-ikr}}{r} \frac{2k[be^{-ar} + 1] + ia[be^{-ar} - 1]}{[be^{-ar} + 1](2k - ia)}$$

while the one that behaves like an outgoing spherical wave is $\psi^{(+)}_{l=0}(k, r) = \psi^{(-)}_{l=0}(-k, r) = \left(\psi^{(-)}_{l=0}(k, r)\right)^*$.

(a) Find the partial wave scattering amplitude $S_{l=0}(k)$.

(b) In general the poles of $S_l(k)$ on the positive imaginary axis at $k = ia_n$, with $a_n > 0$, correspond to bound states with energy $E_n = -\hbar^2 a_n^2/2m$. From the analyticity properties of $S_{l=0}(k)$ computed in part (a), obtain the number of $l = 0$ bound states for various regions of the parameter space a, b (consider all signs and ranges from $-\infty$ to $+\infty$) and give the binding energies.

(c) Does a plot of the potential $V(r)$ for the various signs and regions of a, b that are relevant for part (b) agree qualitatively with the bound state structure you find? Give the plots of the potential energy and present your reasoning..

10. Recall that the general form of the scattering amplitude is given by $f = -2\pi^2 \langle \mathbf{k}' | (2m/\hbar^2) V(r) | \mathbf{k} \rangle^+$. Consider a potential V whose matrix elements are given by

$$\langle \mathbf{r}' | V | \mathbf{r} \rangle = \delta^3(\mathbf{r}' - \mathbf{r}) V_0(r) + \lambda \, u(r') u(r)$$

with $V_0(r)$ describing a hard sphere, i.e. $V_0(r) = \infty$ for $r < a$, and $V_0(r) = 0$ for $r \geq a$, while $u(r)$ is a Yukawa-like function $u(r) = exp(-r/b)/r$.

(a) What is a high energy approximation to the scattering amplitude?

(b) If you drop entirely the hard sphere part, solve exactly for the wavefunction and scattering amplitude. Which partial waves contribute to this scattering amplitude ? (Hint: use the integral equation approach and note that you can find a relation and solve for the unknown integral by integrating once more on both sides of the equation).

(c) Keeping all parts of the potential, show that the $l \neq 0$ phase shift is independent of $u\,(r)$, while the $l = 0$ phase shift satisfies

$$2kbd\cot(\delta_0 + ka) = (k^2b^2 + 1)^2 + (k^2b^2 - 1)\ d,$$

where $d = \frac{4\pi m}{\hbar^2}\lambda b^3 \exp(-2a/b)$ is a dimensionless combination of the constants.

11. Consider the *elastic* scattering of a *fast* electron by a hydrogen atom in the ground state. We will treat the system as a two-body system of two electrons in the presence of a static nucleus. The Hamiltonian is $H_0 + H'$ where $H_0 = H(\text{atom bound electron}) + H(\text{free electron kinetic energy})$ and H' is the potential energy due to the interaction

(a) Write the potential energy H' of the fast electron taking into account its interaction with the nucleus and the bound electron in the H-atom.

(b) The initial state $|i\rangle$ describes the bound electron in the ground state plus the free electron with some momentum. The final state $|f\rangle$ is the bound electron in the ground state and the free electron with a scattered momentum. In an appropriate approximation give the expression for the scattering amplitude without performing any integrals but carefully labelling all integration variables. When is the approximation valid ?

(c) Compute the differential cross section using the integrals $\int d^3r \frac{e^{i\vec{q}\cdot\vec{r}}}{r} = \frac{4\pi}{q^2}$ and $\langle 0|e^{i\vec{q}\cdot\vec{r}}|0 \rangle = \frac{16}{(4+q^2a_0^2)^2}$, where $|0\rangle$ is the ground state of H-atom and a_0 is the Bohr radius.

Chapter 15

Feynman's Path integral formalism

Chapter 16

Relativistic Quantum Mechanics

Chapter 17

Radiation field and photons