Theoretical Chemistry I Quantum Mechanics

Axel Groß

16 October 2008

Preface

Theoretical Chemistry 1 Quantum Mechanics

Prof. Dr. Axel Groß
Phone: 50–22819
Room No.: O25/342
Email: axel.gross@uni-ulm.de
http://www.uni-ulm.de/theochem

Contents

1. Introduction – Wave Mechanics
2. Fundamental Concepts of Quantum Mechanics
3. Quantum Dynamics
4. Angular Momentum
5. Approximation Methods
6. Symmetry in Quantum Mechanics
7. Theory of chemical bonding
8. Scattering Theory
9. Relativistic Quantum Mechanics

Suggested Reading:

- J.J. Sakurai, *Modern Quantum Mechanics*, Benjamin/Cummings 1985
- G. Baym, *Lectures on Quantum Mechanics*, Benjamin/Cummings 1973
- F. Schwabl, *Quantum Mechanics*, Springer 1990

Criteria for getting the *Schein*:

- not specified yet

These lecture notes are based on the class “Theoretical Chemistry I – Quantum Mechanics” in the winter semester 2007 at the University of Ulm
(see http://www.uni-ulm.de/theochem/lehre/)

I am very grateful to Maximilian Lein who provided a LaTeX version of the original notes which have been the basis for this text; furthermore, he created many of the figures. Without his efforts this version of the lecture notes would not have been possible.

Ulm, October 2007 *Axel Groß*

Contents

1 Introduction - Wave mechanics

We will start by recalling some fundamental concepts of quantum wave mechanics based on the correspondence principle.

1.1 Postulates of Wave Mechanics

1. The state of a system is described by its wave function $\Psi(\mathbf{x},t)$ The *probability density* is defined as

   $$\rho(\mathbf{x},t) \equiv |\Psi(\mathbf{x},t)|^2 \tag{1.1}$$

   $|\Psi(\mathbf{x},t)|^2\, d^3x$ describes the probability to find the particle at time t in the volume element d^3x at $\mathbf{x}$.

2. Physical observables correspond to operators that act on the wave function. For example, the momentum $\mathbf{p}$ and the energy E are represented by the following derivatives

   $$\mathbf{p} \ \rightarrow \frac{\hbar}{i}\nabla \ , \tag{1.2}$$

   $$E \ \rightarrow i\hbar\frac{\partial}{\partial t} \ . \tag{1.3}$$

3. Starting from the Hamilton function H of classical mechanics, the time-dependent Schrödinger equation is given by

   $$E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}) \rightarrow i\hbar\frac{\partial}{\partial t}\ \Psi(\mathbf{x},t) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right)\Psi(\mathbf{x},t) \quad , \tag{1.4}$$

   i.e.

   $$i\hbar\frac{\partial}{\partial t}\ \Psi(\mathbf{x},t) = \ H\ \Psi(\mathbf{x},t) \tag{1.5}$$

   with the Hamiltonian

   $$H = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right) \tag{1.6}$$

4. Energy eigenstates are given by the time-independent Schrödinger equation

   $$(H - E)\ \Psi(\mathbf{x},t) = 0 \tag{1.7}$$

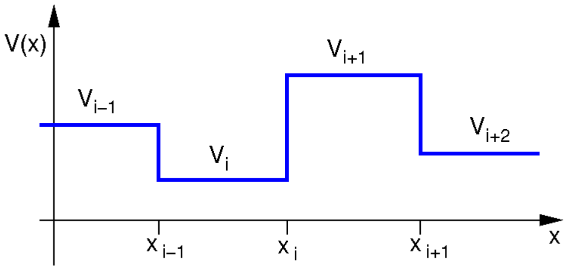


Figure 1.1: Illustration of a piecewise constant potential.

1.2 Simple problems

Although the time-dependent and the time-independent Schrödinger equations look rather simple, their solution is often not trivial. They correspond to second-order partial differential equations. Here we will consider some simple problems. For the sake of simplicity, we consider piecewise continuous potentials. Nevertheless, we will be show the variety of different solutions that are possible in quantum mechanics, such as bound states, scattering and tunneling. In spite of the fact that the chosen potentials are rather simple, we will see that the solutions of corresponding time-independent Schrödinger equation can be quite involved. Later we will learn other techniques such as algebraic formulations that make the solution much easier.

1.2.1 Piecewise constant potentials

Here we will derive the general form of the solution of the time-independent Schrödinger equation for a piecewise constant potential, i.e., potentials that have steps but that are otherwise flat. Such a potential is illustrated in Fig. 1.1. For every flat region of the potential labeled by the index i in Fig. 1.1, the time-independent Schrödinger equation (1.7) in one dimension can be rewritten as

$$\begin{aligned} \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V_i\right)\Psi_i(x) &= E\Psi_i(x), \qquad x_{i-1}<x\le x_i \\ \Leftrightarrow \frac{d^2\Psi_i(x)}{dx^2} &= -\frac{2m}{\hbar^2}(E-V_i)\Psi_i(x) \end{aligned} \tag{1.8}$$

We now assume that $E > V_i$. Then we can define the *wave vector*

$$k = \frac{\sqrt{2m(E-V_i)}}{\hbar}\ , \tag{1.9}$$

so that the Schrödinger equation simply becomes

$$\frac{d^2\Psi_i(x)}{dx^2} = -k^2\Psi(x)\ . \tag{1.10}$$

This is a well-known differential equation in physics and chemistry. Its solution just corresponds to a plane wave

$$\Psi_i(x) = e^{\pm ikx}\ , \tag{1.11}$$

where the $+$ sign presents a wave traveling to the right, i.e., in positive x-direction, and the $-$ sign a wave traveling to the left in negative x-direction.

For $E < V_i$, the *wave vector* (1.12) becomes purely imaginary, i.e. $k = i\kappa$ with

$$\kappa = \frac{\sqrt{2m(V_i - E)}}{\hbar} \; . \tag{1.12}$$

The Schrödinger equation is given by

$$\frac{d^2\Psi(x)}{dx^2} = \kappa^2\Psi(x) \; . \tag{1.13}$$

Now the solutions correspond to exponentially rising and exponentially vanishing functions

$$\Psi(x) = e^{\pm\kappa x} \; . \tag{1.14}$$

Finally we need to know how the solutions are connected at the steps. If the potential step is finite, then the second derivative Ψ'' makes a finite jump at x_i, so that both Ψ and the first derivative Ψ' are continuous at x_i. In other words,

$$\begin{aligned} \Psi_i(x_i) &= \Psi_{i+1}(x_i) \\ \Psi_i'(x_i) &= \Psi_{i+1}'(x_i) \end{aligned} \tag{1.15}$$

for $|V_i - V_{i+1}| < \infty$. For an infinite jump in the potential, i.e., for an infinitely high potential wall, Ψ' makes a finite jump, but Ψ is still continuous.

1.2.2 A simple atom: square well potential

Now we consider a particle in one dimension that can move freely for $|x| < a$, but is confined by infinite potential walls. This means that the potential is given by

$$V(x) = \begin{cases} 0 \; , & |x| < a \\ \infty \; , & |x| \geq a \end{cases} \; . \tag{1.16}$$

This potential is usually refered to as the "particle in a box". It can also be regarded as a very simple model for the electron of the hydrogen atom that is kept close to the proton. Since the potential is constant for $|x| < a$, the solution just corresponds to a superposition of plane waves (1.11). For $|x| > a$, the wave function has to vanish because of the infinite potential. At $|x| = a$, the first derivative of the wave function makes a jump, but the wave function itself is continuous, i.e.,

$$\Psi = \begin{cases} c_1 e^{+iqx} + c_2 e^{-iqx} & , |x| \leq a \\ 0 & , |x| > a \end{cases} \tag{1.17}$$

For $|x| \leq a$, the solutions correspond to a superposition of waves travelling to the right and to the left. Such a superpsoition leads to standing waves which can be written as sine and cosine functions, $\sin qx$ and $\cos qx$, respectively. On the otherhand, the fact that the wave functions have to vanish for $|x| > a$ means that $\Psi(x)$ has also to vanish

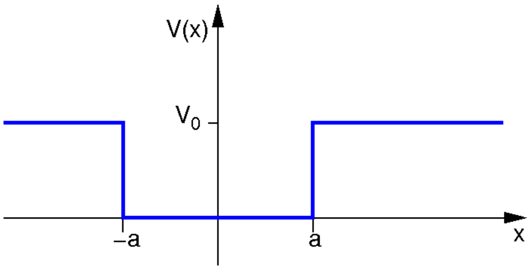


Figure 1.2: Square-well potential

at $|x| = a$, or in other words, they sine and cosine functions have to have a node there. This leads to the following conditions for the allowed wave vectors q:

$$\cos qa = 0 \;\Rightarrow qa = \frac{n\pi}{2} \;, \quad n = 2k+1 \tag{1.18}$$

$$\sin qa = 0 \;\Rightarrow qa = \frac{n\pi}{2} \;, \quad n = 2k \tag{1.19}$$

Hence the allowed eigenfunctions are given by

$$\Psi_n = \begin{cases} \cos \frac{n\pi}{2a} x & n = 2k+1 \\ \sin \frac{n\pi}{2a} x & n = 2k \end{cases} \tag{1.20}$$

The corresponding eigenenergies are

$$E_n = \frac{\hbar^2 q_n^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{n\pi}{2a}\right)^2 \tag{1.21}$$

The ground state function is given by cosine function that has no node between $-a$ and a. The further eigenstates alternate between sine and cosine function, and each higher eigenstate has one more node than the directly lower lying state. Note that the eigenenergies (1.21) grow with the quantum number n as $E_n \propto n^2$. This is different from the case of the hydrogen atom where the eigenenergies are proportional to $-1/n^2$.

Now we will make the system a little bit more realistic by considering walls of finite height. The square well potential is then given by

$$V(x) = V_0\,\theta(|x| - a) = \begin{cases} 0\,, & |x| < a \\ V_0\,, & |x| \geq a \end{cases} \qquad V_0 > 0 \text{ real} \quad \text{number}\,. \tag{1.22}$$

where $\theta(x)$ is the Heaviside step function which is defined as

$$\Theta(x) = \begin{cases} 0\,, & x < 0 \\ 1\,, & x \geq 0 \end{cases} \tag{1.23}$$

The resulting potential is shown in Fig. 1.2. Bound states exist for $0 < E < V_0$. Now we are confronted with a typical quantum phenomenon. Classically, particles can not penetrate into potential regions that are higher than their energy. However, quantum mechanically, there is a non-zero probability to find a particle in potential regions that are finite but larger than the energy of the particle. The particles can still move freely for $|x| \leq a$, but now they penetrate into the potential walls. The time-independent Schrödinger equation becomes

$$\Psi'' = -q^2\Psi \qquad q \equiv \frac{\sqrt{2mE}}{\hbar} \qquad |x| \leq a \tag{1.24}$$

$$\Psi'' = \kappa^2\Psi \qquad \kappa \equiv \frac{\sqrt{2m(V_0 - E)}}{\hbar} \qquad |x| > a \tag{1.25}$$

For $|x| \leq a$, the solutions are again oscillatory, i.e., the are plane waves that can be represented by exponential functions with purely imaginary exponent or as a combination of sine and cosine functions. In contrast, for $|x| > a$ the basic solutions are a linear combination of exponentially rising and vanishing functions

$$\Psi = \begin{cases} c_1 e^{+iqx} + c_2 e^{-iqx} & , |x| \leq a \\ c_+ e^{+\kappa x} + c_- e^{-\kappa x} \ , & |x| > a \end{cases} \tag{1.26}$$

The solutions have to vanish for $x \to \pm\infty$, or in other words

$$\int_{-\infty}^{\infty} |\Psi(x)|^2 dx < \infty \ . \tag{1.27}$$

Therefore one can only have the exponentially decreasing part for $x > a$ and the exponentially increasing component for $x < -a$. One can only say that $e^{+\kappa x}$ is not normalizable for $x > a$ and analogously $e^{-\kappa x}$ not for $x < -a$. Furthermore, since $V(x)$ is an even potential, the solutions can be characterized according to their symmetry, i.e., the solutions are either even or odd as we will see in later chapters. This means that they can be represented by cosine or sine functions, respectively. If we have even symmetry, the solution will be

$$\Psi(\mathbf{x}) = \begin{cases} A\cos qx & |x| \leq a \\ e^{-\kappa|x|} & |x| > a \end{cases} \tag{1.28}$$

For odd symmetry, we get

$$\Psi(\mathbf{x}) = \begin{cases} B\sin qx & |x| \leq a \\ \pm e^{-\kappa|x|} & |x| > a \end{cases} \tag{1.29}$$

It is interesting to note that even such a simple example as the square well potential is not that easy to solve. We will illustrate this in the following. Assume first that Ψ has even symmetry. Continuity at $x = a$ requires

$$A\cos qa = e^{-\kappa a} \tag{1.30}$$

Ψ' has to be continuous, too. From that, we get

$$= \tag{1.31}$$

If we devide (1.31) by (1.30), we obtain

$$\tan aq = \frac{\kappa}{q} \tag{1.32}$$

Now we introduce the dimensionless parameter

$$\lambda = a\frac{\sqrt{2mV_0}}{\hbar} \ , \tag{1.33}$$

so that Eq. (1.32) becomes

$$\begin{aligned} \tan aq = \frac{\kappa}{q} = \frac{\kappa a}{qa} &= \\ &= \\ &= \frac{\sqrt{\lambda^2 - (qa)^2}}{qa} \end{aligned} \tag{1.34}$$

This is a transcendental equation that cannot be solved analytically.

Now assume odd symmetry.

$$\begin{aligned} B \sin qa = & \qquad\qquad \mid Bq \cos qa = \\ \Rightarrow \ -\cot qa = \frac{\kappa}{q} &= \frac{\sqrt{\lambda^2 - (qa)^2}}{qa} \end{aligned} \tag{1.35}$$

Again a transcendental equation that can only be solved graphically. For the graphical solution we first not that κ/q behaves likes $1/q$ for small q. Furthermore:

$$\frac{\kappa}{q} = 0 \text{ for } qa = \frac{a}{\hbar}\sqrt{2mV_0} = \lambda \tag{1.36}$$

In Fig. 1.3 the graphical solution is sketched. For three different values of V_0 κ/q is plotted as a function of qa together with $\tan qa$ and $-\cot qa$. Every crossing point of the curves corresponds to a solution. Since κ/q diverges for $qa \to 0$, there is at least one crossing point with $\tan qa$. The lowest energy state is *always even*. When λ increases by $\pi/2$, there is another crossing point and hence one additional state. Hence the total number of states is given by

$$N_S = \left[\frac{2\lambda}{\pi}\right] = \left[\frac{2a\sqrt{2mV_0}}{\pi\hbar}\right] \tag{1.37}$$

with $[\alpha]$ nearest integer greater than α. Even and odd states alternate.

It is always instructive to look a limiting cases whose solution is exactly known. If the potential wall V_0 grow to infinity, we recover the particle-in-a-box problem. First of all, we see from (1.37) that $N_S \propto V_0$, i.e., for inifinite potential walls we also get infinitely many states, as is also reflected in Eq. (1.21). Furthermore, when V_0 increases, the curves for λ in Fig. 1.3 become higher and more flat. This means that they will cross $\tan qa$ and $-\cot qa$ at values closer and closer to $qa = n\pi/2$, which also corresponds to the case of the particle in a box.

1.2.3 Transmission-Reflection Problems

Transmission-reflection problems occur in chemistry when two particles meet each other. Here we treat such a problem as an one-dimensional potential step

$$V(x) = V_0\,\theta(x) \tag{1.38}$$

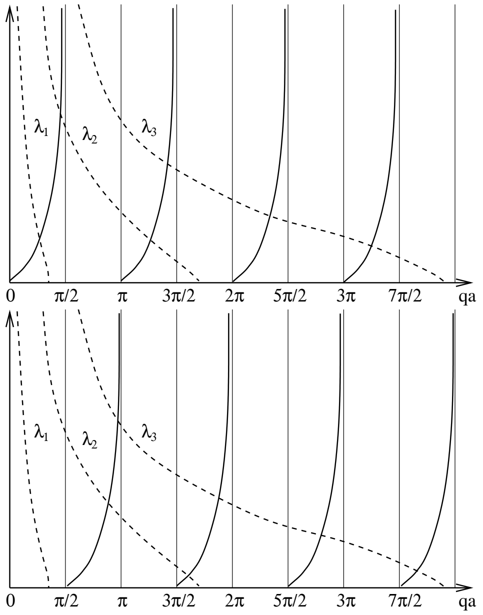


Figure 1.3: Graphical solution for the even and odd solutions of the square well problem

For $x < 0$, the potential is 0, for $x > 0$, the potential is V_0 (see Fig. 1.4). The Schrödinger equation for $x < 0$ and $x > 0$ is given by

$$\frac{d^2\Psi}{dx^2} = -\frac{2mE}{\hbar^2}\Psi = -k^2\Psi \qquad x < 0 \tag{1.39}$$

$$\frac{d^2\Psi}{dx^2} = -\frac{2m(E-V_0)}{\hbar^2}\Psi = -k'^2\Psi \qquad x > 0 \tag{1.40}$$

Let $E > V_0$. Suppose a particle is incident from the left.

$$\Psi_{\mathrm{I}}(x) = e^{+ikx} + re^{-ikx} \tag{1.41}$$

$$\Psi_{\mathrm{II}}(x) = te^{+ik'x} \tag{1.42}$$

r and t are the *reflection* and *transmission amplitudes*. We will now discuss the so-called probability flux which is given by

$$j(x) = \frac{\hbar}{m}\Im\big(\Psi^*\frac{d}{dx}\Psi\big) \tag{1.43}$$

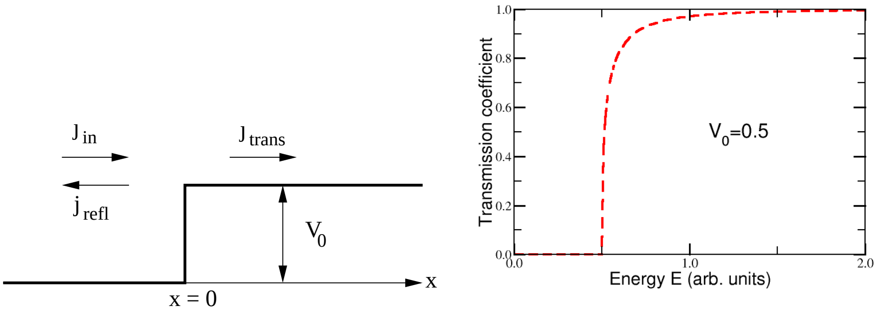


Figure 1.4: Potential step and transmission coefficient $T(E)$ for a potential step of height $V_0 = 0.5$.

For the particular problem of the potential step, we obtain

$$
\begin{aligned}
j_{\mathrm{I}}(x) &= \frac{\hbar}{m}\Im\big(\Psi_{\mathrm{I}}^{*}\frac{d}{dx}\Psi_{\mathrm{I}}\big) \\
&= \frac{\hbar}{m}\Im\big[(e^{-ikx}+r^{*}e^{ikx})ik(e^{ikx}-re^{-ikx})\big] \\
&= \frac{\hbar}{m}\Im\big[ik\big(1-|r|^2 \underbrace{-re^{-i2kx}+r^{*}e^{i2kx}}_{=\text{purely imaginary}}\big)\big] \\
&= \frac{\hbar k}{m}(1-|r|^2) \equiv j_{\mathrm{in}} - j_{\mathrm{refl}} \qquad (1.44) \\
j_{\mathrm{II}}(x) &= \frac{\hbar}{m}\Im\big[t^{*}e^{-ik'x}(ik')te^{ik'x}\big] \\
&= \frac{\hbar k'}{m}\,|t|^2 \equiv j_{\mathrm{trans}} \qquad (1.45)
\end{aligned}
$$

We will now define R and T, the *reflection* and *transmission coefficient*.

$$
R \equiv \frac{j_{\mathrm{ref}}}{j_{\mathrm{in}}} = |r|^2 \qquad (1.46)
$$

$$
T \equiv \frac{j_{\mathrm{trans}}}{j_{\mathrm{in}}} = \frac{k'}{k}\,|t|^2 \qquad (1.47)
$$

Due to particle conservation we have $j_{\mathrm{in}} - j_{\mathrm{refl}} = j_{\mathrm{out}}$, which leads to

$$
\begin{aligned}
&\frac{\hbar k}{m}(1-|r|^2) = \frac{\hbar k'}{m}\,|t|^2 \qquad j_{\mathrm{I}} = j_{\mathrm{II}} \\
&\Rightarrow R + T = 1 \qquad (1.48)
\end{aligned}
$$

If the potential is Hermitian, then the number of particles is conserved. If it is non-Hermitian, the potential must have an imaginary part not identically to zero. Imaginary potentials can describe the annihilation of particles.

It is left as an exercise to calculate the transmission coefficient for the potential step.

The result is given by

$$T(E) = \frac{4\sqrt{(E-V_0)E}}{\left(\sqrt{E} + \sqrt{(E-V_0)}\right)^2} \,, \tag{1.49}$$

which is also plotted in Fig. 1.4. There is another typical quantum phenomenon visible in Fig. 1.4: For energies slightly larger than the step height, the transmission coefficient is less than one. This means that there is *reflection above the barrier*. Although the energy higher than the potential step, a certain fraction of the incoming beam will still be reflected. Classically, there is no reflection above the barrier. It is true that the particles will be slowed down, but they will all propagate along the potential step.

1.2.4 A simple reaction: transmission through a potential barrier

The transmission through a potential barrier is a typical process in chemistry that occurs in any activated chemical reaction. However, the following considerations are also relevant for the understanding of the scanning tunneling microscope (STM), which has become one of the standard tools in surface chemistry. If the potential barrier has a finite width, then particles can be transmitted even with energies below the barrier height. This is again a typical quantum phenomenon called *tunneling* that is not possible in classical mechanics.

The potential we consider is again piecewise constant:

$$V(x) = V_0\,\theta(a - |x|) \tag{1.50}$$

This looks like the square well of Fig. 1.2, but just inverted.

The general solution for $0 < E < V_0$ can be written as

$$\Psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < -a \\ Ce^{-\kappa x} + De^{\kappa x} & -a \le x \le a \\ Fe^{ikx} + Ge^{-ikx} & x > a \end{cases} \tag{1.51}$$

where the quantities $k = \frac{1}{\hbar}\sqrt{2mE}$ and $\kappa = \frac{1}{\hbar}\sqrt{2m(V_0 - E)}$ are again the *wave numbers*.

The solution of the transmission problem is straightfoward but still tedious. We derive the explicit solution in the appendix of this chapter. Here we just consider the most important results.

Consider a particle incident from the left, i. e. $G = 0$. The incoming wave amplitude is then given by A, the reflected wave amplitude is given by B and the transmitted flux is given by F. For $E < V_0$, the transmission coefficient is given by

$$\begin{aligned} T(E) \equiv |t(E)|^2 = |F/A|^2 &= \frac{1}{1 + \left\{ \dfrac{\left(k^2 + \kappa^2\right)^2}{4k^2\kappa^2} \right\} \sinh^2 2\kappa a} \\ &= \frac{1}{1 + \left\{V_0^2/4E(V_0 - E)\right\} \sinh^2 \left(2a\sqrt{2m(V_0 - E)/\hbar^2}\right)} \end{aligned} \tag{1.52}$$

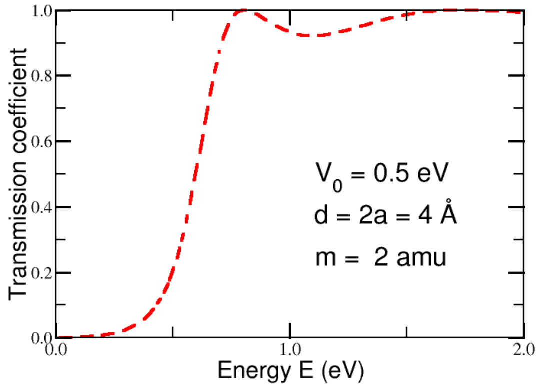


Figure 1.5: Transmission coefficient $T(E)$ for a square barrier of height $V_0 = 0.5\,\text{eV}$, width $d = 2a = 4\,\text{Å}$, and $m = 2\,\text{amu}$.

For energies larger than the barrier, i.e., $E > V_0$, the transmission probability is given by

$$T(E) = \frac{1}{1 + \left\{V_0^2/4E(E-V_0)\right\}\sin^2\left(2a\sqrt{2m(E-V_0)/\hbar^2}\right)} \tag{1.53}$$

Note that Eq. (1.53) also decribes the transmission across the square-well potential (1.22) with V_0 replaced by $-V_0$.

The transmission coefficient $T(E)$ for a square barrier of height $V_0 = 0.5\,\text{eV}$, width $d = 2a = 4\,\text{Å}$, and $m = 2\,\text{amu}$ corresponding to a H_2 molecule has been plotted in Fig. 1.5 using Eqs. (1.52) and (1.53). For $E < V_0 = 0.5\,\text{eV}$, there is already a significant transmission probability of up to 20%. For $E > V_0$, the transmission probability is in general less than unity. Note that there is an oscillatory structure in the transmission probability with resonances where the probability is one. An inspection of Eq. (1.53) shows that these resonances occur when

$$2a\sqrt{\frac{2m(E-V_0)}{\hbar^2}} = n\pi, \quad n = 1, 2, 3, \ldots \tag{1.54}$$

Now we consider the limiting case of a very high and wide barrier, e. g. $\kappa \cdot a \gg 1$ which also corresponds to $0 < E \ll V_0$. Under these conditions, we can expand the $\sinh 2\kappa a \approx \frac{1}{2}e^{2\kappa a} \gg 1$. Then the transmission coefficient is approximately equal to

$$\begin{aligned} T(E) &\approx \frac{16(\kappa k)^2}{(\kappa^2+k^2)^2} e^{-4\kappa a} \\ &= \frac{16E(V_0-E)}{V_0^2} \exp\left(-\frac{4}{\hbar}\sqrt{2m(V_0-E)}a\right) \\ \Rightarrow T(E) &\propto \exp\left(-\frac{4}{\hbar}\sqrt{2m(V_0-E)}a\right) \end{aligned} \tag{1.55}$$

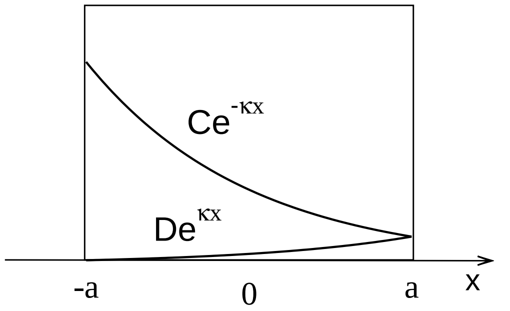


Figure 1.6: Wave function in the potential barrier

Thus, for large and high barriers, tunnelling is suppressed exponentially. The transmission probability decreases exponentially with the width a of the barrier and the square root of the difference between the energy and the energetic height of the barrier. This is a purely quantum mechanical process.

Let us have a look inside the barrier. From the boundary conditions, we can derive that

$$\begin{aligned} C &= F \cdot \frac{1}{2}\left(1 - i\frac{k}{\kappa}\right) e^{(\kappa+ik)a} \\ D &= F \cdot \frac{1}{2}\left(1 + i\frac{k}{\kappa}\right) e^{(-\kappa+ik)a} \end{aligned} \tag{1.56}$$

Now we use the fact that the coefficient F is proportional to the transmission amplitude $t(E)$ (see (1.52)). Then in the case of a high and wide barrier, i. e. $\kappa a \gg 1$, we get from Eq. (1.55)

$$F \propto \sqrt{T(E)} \propto e^{-2\kappa a} \tag{1.57}$$

and

$$\begin{aligned} C &\propto e^{-\kappa a + ika} \\ D &\propto e^{-3\kappa a + ika} \\ \Rightarrow\ Ce^{-\kappa x}\big|_{x=a} &\propto e^{-2\kappa a} \propto De^{\kappa x}\big|_{x=a} \end{aligned} \tag{1.58}$$

In the end, F consists of two parts – an exponentially decreasing part, $Ce^{-\kappa x}$ and an exponentially increasing part, $De^{\kappa x}$, which add up to F at $x = a$ (see Fig. 1.6).

Continuous Potential Barrier

If we have a continuous potential, then we approximate $V(x)$ by individual square barriers of width dx (see Fig. 1.7), which means that we replace the step width $2a$ in Eq. (1.55) by dx. In the limiting case of a high and wide barrier, the transmission probability can then be derived by multiplying the transmission probabilities of each

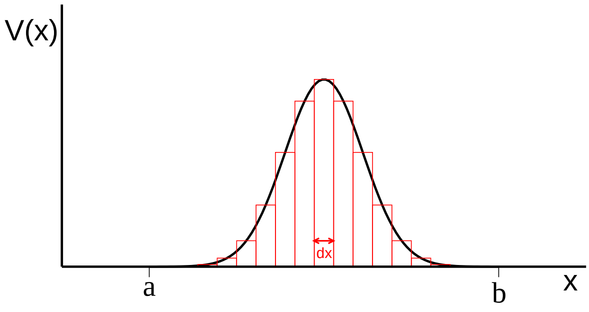


Figure 1.7: Decomposition of a continuous barrier into rectangular barriers

segment:

$$T(E) = \Pi_{i=1}^{n} e^{-\frac{2}{\hbar}\sqrt{2m(V(x_i)-E)}dx} = \exp\big(-\frac{2}{\hbar}\sum_{i=1}^{n}\sqrt{2m(V(x_i)-E)}dx\big)$$

$$\xrightarrow{n\to\infty} \exp\big(-\frac{2}{\hbar}\int_a^b \sqrt{2m(V(x)-E)}\,dx\big) \tag{1.59}$$

1.2.5 Appendix: Transmission through a potential barrier

Here we show the explicit solution of the transmission through a square potential barrier. For the potential barrier $V(x) = V_0\,\theta(a-|x|)$ of Eq. (1.50), the general solution for $0 < E < V_0$ can be written as

$$\Psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x < -a \\ Ce^{-\kappa x} + De^{\kappa x} & -a \le x \le a \\ Fe^{ikx} + Ge^{-ikx} & x > a \end{cases} \tag{1.60}$$

where the quantities $k = \frac{1}{\hbar}\sqrt{2mE}$ and $\kappa = \frac{1}{\hbar}\sqrt{2m(V_0 - E)}$ are again the *wave numbers*.

First of all, we write down the matching conditions at $x = -a$.

$$Ae^{-ika} + Be^{ika} = Ce^{\kappa a} + De^{-\kappa a} \tag{1.61}$$

$$ik\big[Ae^{-ika} - Be^{ika}\big] = -\kappa\big[Ce^{\kappa a} - De^{-\kappa a}\big] \tag{1.62}$$

In matrix notation, this is easier to solve.

$$\begin{pmatrix} e^{-ika} & e^{ika} \\ e^{-ika} & -e^{ika} \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} e^{\kappa a} & e^{-\kappa a} \\ \frac{i\kappa}{k}e^{\kappa a} & -\frac{i\kappa}{k}e^{-\kappa a} \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix}$$

$$M(a)\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} \tag{1.63}$$

Here $M(a)$ is given by

$$M(a) = \frac{1}{2}\begin{pmatrix} (1+\frac{i\kappa}{k})e^{\kappa a + ika} & (1-\frac{i\kappa}{k})e^{-\kappa a + ika} \\ (1-\frac{i\kappa}{k})e^{\kappa a - ika} & (1+\frac{i\kappa}{k})e^{-\kappa a - ika} \end{pmatrix} \tag{1.64}$$

The matching conditions at $x = a$ are similar.

$$M(-a)\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix} \tag{1.65}$$

$$M(a)M^{-1}(-a)\begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} \tag{1.66}$$

where $M^{-1}(-a)$ is given by

$$M^{-1}(-a) = \frac{1}{2}\begin{pmatrix} (1-\frac{ik}{\kappa})e^{\kappa a+ika} & (1+\frac{ik}{\kappa})e^{\kappa a-ika} \\ (1+\frac{ik}{\kappa})e^{-\kappa a+ika} & (1-\frac{ik}{\kappa})e^{-\kappa a-ika} \end{pmatrix} \tag{1.67}$$

The solution for the coefficients is

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \big(\cosh 2\kappa a + \frac{i\varepsilon}{2}\sinh 2\kappa a\big)e^{i2ka} & \frac{i\eta}{2}\sinh 2\kappa a \\ -\frac{i\eta}{2}\sinh 2\kappa a & \big(\cosh 2\kappa a - \frac{i\varepsilon}{2}\sinh 2\kappa a\big)e^{-i2ka} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix} \tag{1.68}$$

with $\varepsilon = \frac{\kappa}{k} - \frac{k}{\kappa}$ and $\eta = \frac{\kappa}{k} + \frac{k}{\kappa}$.

Consider a particle incident from the left, i. e. $G = 0$. The incoming wave amplitude is then given by A, the reflected wave amplitude is given by B and the transmitted flux is given by F. Thus we get from the matrix equation (1.68)

$$\begin{aligned} A &= F\left(\cosh 2\kappa a + \frac{i\varepsilon}{2}\sinh 2\kappa a\right)e^{i2ka} \\ B &= F\cdot\left(-\frac{i\eta}{2}\right)\sinh 2\kappa a \end{aligned} \tag{1.69}$$

The transmission amplitude is given by $t(E) \equiv \frac{F}{A}$.

$$\begin{aligned} t(E) = \frac{F}{A} &= \frac{F}{F\left(\cosh 2\kappa a + \frac{i\varepsilon}{2}\sinh 2\kappa a\right)e^{i2ka}} \\ &= \frac{e^{-i2ka}}{\cosh 2\kappa a + i\frac{\varepsilon}{2}\sinh 2\kappa a} \end{aligned} \tag{1.70}$$

Now we want to calculate the transmission coefficient.

$$\begin{aligned} T(E) \equiv |t(E)|^2 &= \frac{1}{\left(\cosh 2\kappa a + i\frac{\varepsilon}{2}\sinh 2\kappa a\right)\left(\cosh 2\kappa a - i\frac{\varepsilon}{2}\sinh 2\kappa a\right)} \\ &= \frac{1}{1+\left(1+\frac{\varepsilon^2}{4}\right)\sinh^2 2\kappa a} \end{aligned} \tag{1.71}$$

This can be rewritten to give

$$\begin{aligned} T(E) \equiv |t(E)|^2 = |F/A|^2 &= \frac{1}{1 + \left\{ \dfrac{\left(k^2+\kappa^2\right)^2}{4k^2\kappa^2} \right\} \sinh^2 2\kappa a} \\ &= \frac{1}{1 + \left\{ V_0^2/4E(V_0-E) \right\} \sinh^2 \left(2a\sqrt{2m(V_0-E)/\hbar^2} \right)} \end{aligned} \tag{1.72}$$

2 Fundamental Concepts of Quantum Mechanics

2.1 Introduction

Let us start with first discussing the *Stern-Gerlach experiment* performed in 1922.

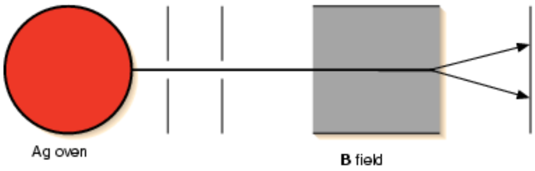


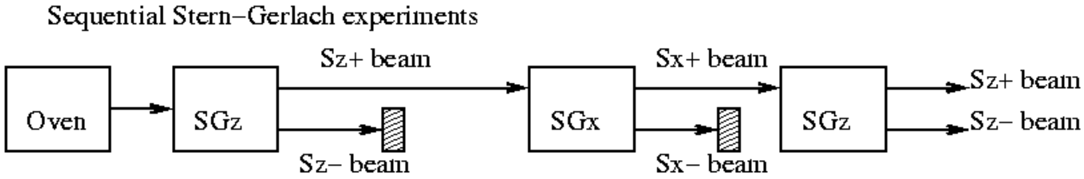
Figure 2.1: Diagram of the Stern-Gerlach-Experiment

The magnetic moment of the silver atoms is proportional to the magnetic moment of the $5s^1$ electron, the inner electron shells do not have a net magnetic moment. The Force in z-direction in an inhomogeneous magnetic field is given by

$$F_z = \frac{\partial}{\partial z}(\mu \cdot \mathbf{B}) \approx \mu_z \frac{\partial B_z}{\partial z} \tag{2.1}$$

We expect from classical mechanics that the atoms are distributed randomly with a peak in the middle. But we observe two different peaks; if we calculate the magnetic moment from the data obtained, we get that the magnetic moment is either $S = +\hbar/2 e_z$ or $S = -\hbar/2 e_z$ – the electron spin is *quantitized*.

Historically, more sophisticated experiments followed. Instead of using just one magnet, several magnets are used in series, so that sequential Stern-Gerlach experiments can be performed:



They show that the spin is quantized in every direction by the amount above, $\pm\hbar/2$. It also suggests that selecting the S_x+ component after a Stern-Gerlach experiment in x-direction completely destroys any previous information about S_z. There is in fact an analogon in classical mechanics – the transmission of polarized light through polarization filters.

The following correspondence can be made

$$\begin{aligned} S_z \pm atoms &\leftrightarrow x-, y - \text{polarized light} \\ S_x \pm atoms &\leftrightarrow x'-, y' - \text{polarized light} \end{aligned} \tag{2.2}$$

where x' and y'-axes are x and y axes rotated by $45°$.

Notation. We write the S_{z+} state as $|S_z;+\rangle$ or $|S_z;\uparrow\rangle$; similarly, the S_{z-} state corresponds to $|S_z;\downarrow\rangle$. We assume for S_x states *superposition* of S_z states

$$|S_x;\uparrow\rangle \overset{?}{=} \frac{1}{\sqrt{2}}|S_z;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_z;\downarrow\rangle \tag{2.3}$$

$$|S_x;\downarrow\rangle \overset{?}{=} -\frac{1}{\sqrt{2}}|S_z;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_z;\downarrow\rangle \tag{2.4}$$

This really is a *two-dimensional space*! What about the S_y states – it should be a linear combination of the S_z states, too. However, all possible combinations seem to be used up. The analogy is circular polarized light. Right circularly polarized light can be expressed as

$$\mathbf{E} = E_0 \left[\frac{1}{\sqrt{2}} e_x e^{i(kz-\omega t)} + \frac{i}{\sqrt{2}} e_y e^{i(kz-\omega t)} \right] \tag{2.5}$$

Can we use this analogy to define the S_y states?

$$|S_y;\uparrow\rangle \overset{?}{=} \frac{1}{\sqrt{2}}|S_z;\uparrow\rangle + \frac{i}{\sqrt{2}}|S_z;\downarrow\rangle \tag{2.6}$$

$$|S_y;\downarrow\rangle \overset{?}{=} -\frac{1}{\sqrt{2}}|S_z;\uparrow\rangle - \frac{i}{\sqrt{2}}|S_z;\downarrow\rangle \tag{2.7}$$

We already note here that only the "direction" in the vector space is of significance, not the "length" of the vectors.

2.2 Kets, Bras, and Operators

Consider a complex vector space of dimension d which is related to the nature of the physical system.

The space of a single electron spin is two-dimensional whereas for the description of a free particle a vector space of denumerably infinite dimension is needed.

2.2.1 Kets

The vector space is called Hilbert Space. The *physical state* is represented by a *state vector*. Following Dirac, a state vector is called *ket* and denoted by $|\alpha\rangle$.

They suffice the usual requirements for vector spaces (commutative, associative addition, existence of null ket and inverse ket, and scalar multiplication).

One important postulate is that $|\alpha\rangle$ and $c \cdot |\alpha\rangle$ with $c \neq 0$ correspond to the same physical state. Mathematically this means that we deal with rays rather than vectors

A *physical observable* can be represented by an *operator*. Operators act on kets from the left.

$$A(|\alpha\rangle) = A|\alpha\rangle \tag{2.8}$$

In general, applying an operator to a ket cannot be expressed as a scalar multiplication, i.e.,

$$A(|\alpha\rangle) \neq c\,|\alpha\rangle \;\text{ in general} \tag{2.9}$$

with c any complex number.

Analogously to eigenvectors, there are eigenkets

$$A(|\alpha\rangle) = a\,|\alpha\rangle \;\;,\;\; A(|\alpha'\rangle) = a'\,|\alpha'\rangle \;\;, \tag{2.10}$$

with eigenvalues $a, a', \ldots$.

Example

Spin 1/2 System. $S_z\,|S_z,\uparrow\rangle = +\frac{\hbar}{2}\,|S_z,\uparrow\rangle$, $S_z\,|S_z,\downarrow\rangle = -\frac{\hbar}{2}\,|S_z,\downarrow\rangle$

2.2.2 Bra space and inner product

The Bra space is the vector space dual to the ket space. It is spanned by the eigenbras $\{\langle a'|\}$ which correspond to the eigenkets $\{|a'\rangle\}$. There is an isomorphism that assigns each ket onto its bra.

$$\langle\alpha| \leftrightarrow |\alpha\rangle \qquad c_\alpha\,\langle\alpha| + c_\beta\,\langle\beta| \leftrightarrow c_\alpha^*\,|\alpha\rangle + c_\beta^*\,|\beta\rangle \tag{2.11}$$

Note the complex-conjugated coefficients.

Now we introduce the inner or scalar product[1].

$$\begin{aligned} \langle\Box|\Box\rangle : V^* \times V &\Rightarrow \mathbb{C} \\ (\langle\beta|\,,|\alpha\rangle) &\longmapsto \langle\beta|\alpha\rangle \end{aligned} \tag{2.12}$$

It is Hermitian and positive definite, i.e.,

$$\begin{aligned} \langle\beta|\alpha\rangle &= \langle\alpha|\beta\rangle^* \\ \langle\alpha|\alpha\rangle &\geq 0\;. \end{aligned} \tag{2.13}$$

Two kets $|\alpha\rangle$ and $|\beta\rangle$ are said to be orthogonal if

$$\langle\alpha|\beta\rangle = 0\;. \tag{2.14}$$

We can explicitly normalize a ket $|\tilde{\alpha}\rangle$ by

$$\tag{2.15}$$

$\sqrt{\langle\tilde{\alpha}|\tilde{\alpha}\rangle}$ is known as the norm of $|\tilde{\alpha}\rangle$.

---

[1] bra)(ket

2.3 Operators

X and Y are said to be equal, $X = Y$, if $X|\alpha\rangle = Y|\alpha\rangle$. for arbitrary $|\alpha\rangle$. Then we write $X = Y$. Operator addition is commutative and associative,

$$X + Y = Y + X \; ; \; X + (Y + Z) = (X + Y) + Z \; . \tag{2.16}$$

Operators are usually linear, that is,

$$X(a_1|\alpha_1\rangle + a_2|\alpha_2\rangle) = a_1 X|\alpha_1\rangle + a_2 X|\alpha_2\rangle \tag{2.17}$$

An exception is for example the time-reversal operator which is *antilinear*.

An operator acts on a bra from the *right* side.

Note. $X|\alpha\rangle$ and $\langle\alpha|X$ are in general *not dual* to each other. The correspondence is

$$X|\alpha\rangle \leftrightarrow \langle\alpha|X^\dagger \tag{2.18}$$

The operator $X^\dagger$ is called the *Hermitian adjoint*, or simply the adjoint of X. By definition, an operator is said to be *Hermitian*, if and only if $X^\dagger = X$.

2.3.1 Multiplication of operators

The multiplication of operators corresponds to the successive application of operators. It is in general *non-commutative*, i. e.

$$XY \neq YX \tag{2.19}$$

but associative.

$$X(YZ) = (XY)Z \tag{2.20}$$

Furthermore, the usual rules of a non-Abelian group apply. The Hermitian adjoint of a product of operators is given by

$$(XY)^\dagger \; = Y^\dagger X^\dagger \tag{2.21}$$

2.3.2 Outer Product

The outer product of $|\beta\rangle$ and $\langle\alpha|$ is defined as

$$(|\beta\rangle)(\langle\alpha|) = |\beta\rangle\langle\alpha| \tag{2.22}$$

It is to be regarded as an operator.

$$|\beta\rangle\langle\alpha|\underbrace{|\gamma\rangle}_{\text{ket}} = \underbrace{|\beta\rangle}_{\text{ket}}\underbrace{\langle\alpha|\gamma\rangle}_{\in\mathbb{C}} \tag{2.23}$$

Some more rules. Let X be an operator.

$$\big(|\beta\rangle\langle\alpha|\big)^\dagger = |\alpha\rangle\langle\beta| \tag{2.24}$$

$$\big(|\beta\rangle\big)\big(X\langle\alpha|\big) = \big(|\beta\rangle X\big)\big(\langle\alpha|\big) = |\beta\rangle X\langle\alpha| \tag{2.25}$$

$$\langle\beta|X|\alpha\rangle = \langle\alpha|X^\dagger|\beta\rangle^*$$

$$X \text{ Hermitian} \Leftrightarrow \langle\beta|X|\alpha\rangle = \langle\alpha|X|\beta\rangle^* \tag{2.26}$$

2.3.3 Base Kets and Matrix Representations

Theorem

The eigenvalues of a Hermitian operator A are real; the eigenkets of A corresponding to different eigenvalues are orthogonal.

Proof

□

Usually, we will assume that eigenkets are normalized, i. e. $\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$. Thus, the eigenkets form an orthogonal set.

2.3.4 Eigenkets as Base Kets

Normalized eigenkets of A form a complete orthonormal set, i. e. an arbitrary ket $|\beta\rangle$ can be expressed as a linear combination of eigenkets.

$$|\beta\rangle = \sum_k c_k |\alpha_k\rangle \tag{2.27}$$

If we multiply by $\langle\alpha_j|$ we get that $\langle\alpha_j|\beta\rangle = c_j$. Thus, we get

$$|\beta\rangle = \sum_k |\alpha_k\rangle \langle\alpha_k|\beta\rangle \tag{2.28}$$

$$\Rightarrow \sum_k |\alpha_k\rangle \langle\alpha_k| = \mathbf{1} \tag{2.29}$$

2.3.5 Resolution of the Identity, Completeness Relation, or Closure

(2.29) can be extremely useful.

(2.30)

If $\langle\beta|$ is normalized, then $\sum_k |c_k|^2 = 1$. Each summand $|\alpha_k\rangle\langle\alpha_k|$ selects the portion of $|\beta\rangle$ parallel to $|\alpha_k\rangle$. Thus, it is a *projection* operator; it is denoted by $\Lambda_k = |\alpha_k\rangle\langle\alpha_k|$.

Therefore, every operator can be *represented* in a matrix via $X = \mathbf{1}X\mathbf{1}$; the bra index is the row index, the ket index is the column index.

$$
\begin{aligned}
X &= \sum_{k,j} \left(|\alpha_k\rangle \underbrace{\langle\alpha_k|}_{\text{row}} X \underbrace{|\alpha_j\rangle}_{\text{column}} \langle\alpha_j| \right) \\
&\hat{=} \begin{pmatrix} \langle\alpha_1|X|\alpha_1\rangle & \langle\alpha_1|X|\alpha_2\rangle & \dots \\ \langle\alpha_2|X|\alpha_1\rangle & \langle\alpha_2|X|\alpha_2\rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}
\end{aligned}
\tag{2.31}
$$

The *Hermitian adjoint* of an operator corresponds to the complex conjugate transposed matrix.

The successive application of operators corresponds to matrix multiplication.

Proof

□

Thus, state kets are represented by column vectors and bras by row vectors.

$$
|\beta\rangle \hat{=} \begin{pmatrix} \langle\alpha_1|\beta\rangle \\ \langle\alpha_2|\beta\rangle \\ \langle\alpha_3|\beta\rangle \\ \vdots \end{pmatrix}, \qquad
\begin{aligned}
\langle\gamma| &\hat{=} \left(\langle\gamma|\alpha_1\rangle, \langle\gamma|\alpha_2\rangle, \langle\gamma|\alpha_3\rangle, \dots \right) \\
&= \left(\langle\alpha_1|\gamma\rangle^*, \langle\alpha_2|\gamma\rangle^*, \langle\alpha_3|\gamma\rangle^*, \dots \right)
\end{aligned}
\tag{2.32}
$$

2.4 Spin 1/2 System

As base kets $|S_z;+\rangle = |S_z;\uparrow\rangle = |\uparrow\rangle$ and $|S_z;-\rangle = |\downarrow\rangle$ are used. Since nothing is special about the z-axis, this just corresponds to a convention. The base kets can also be written as a column matrix

$$
|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.33}
$$

The simplest operators is the identity operator $\mathbf{1}$. In the basis given above, the operator S_z is diagonal.

$$
\mathbf{1} = |\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.34}
$$

$$
S_x = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right) \equiv \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.35}
$$

$$
S_y = \frac{i\hbar}{2}\left(-|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right) \equiv \frac{i\hbar}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{2.36}
$$

$$
S_z = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right) \equiv \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.37}
$$

Nearly all operators we deal with are Hermitian operators. However, there are some important non-Hermitian operators, e. g. the so-called lowering and raising operator, S_- and S_+ respectively.

$$S_+ = \hbar \left|\uparrow\right\rangle \left\langle\downarrow\right| \hat{=} \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{2.38}$$

$$S_- = \hbar \left|\downarrow\right\rangle \left\langle\uparrow\right| \hat{=} \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{2.39}$$

$S_\pm$ raise and lower the spin by one unit of $\hbar$, respectively. As we will see later, they can be expressed as

$$S_\pm = S_x \pm i S_y \tag{2.40}$$

2.5 Measurements, Observables And The Uncertainty Relation

Consider a state $|\alpha\rangle = \sum_k c_k |\alpha_k\rangle = \sum_k |\alpha_k\rangle \langle\alpha_k|\alpha\rangle$. According to the *quantum theory of measurement*, after a measurement of the observable A the system is 'thrown' into an eigenstate of A, for example $|\alpha_k\rangle$. We measure A to be α_k.

The Result of a measurement yields one of the eigenvalues of A.

Theorem Postulate

The probability for a state α_k to be measured is

$$P(\alpha_k) = |\langle\alpha_k|\alpha\rangle|^2 \tag{2.41}$$

provided that $|\alpha\rangle$ normalized.

These probabilities $P(\alpha_k)$ can be determined with a large number of experiments performed on an *ensemble* of *identically prepared physical systems*, a so-called *pure ensemble*.

If the system already is in an eigenstate α_k, then the probability to measure α_k is 1.

The *expectation value* of an operator A with respect to state α is

$$\langle A\rangle \equiv \langle A\rangle_\alpha \equiv \langle\alpha|A|\alpha\rangle \tag{2.42}$$

This corresponds to the average measured value which can be derived from

$$\langle A\rangle = \tag{2.43}$$

which corresponds to a sum over the measured values a_j which are weighted by the probabilities $P(\alpha_j) = |\langle\alpha_j|\alpha\rangle|^2$.

Note. In general, expectation values do not coincide with eigenvalues, e.g. in the spin $1/2$ system, the expectation value is a number that 'arbitrarily' lies between $\pm\hbar/2$.

2.5.1 Compatible Observables

We define the *commutator* and the *anti-commutator* as follows:

$$[A, B] \equiv AB - BA \qquad \{A, B\} \equiv AB + BA \tag{2.44}$$

Note that the commutator and the anti-commutator also represent oeprators. If A and B are Hermitian, then the commutator $[A, B]$ is anti-Hermitian:

$$\begin{aligned} ([A, B])^\dagger &= (AB - BA)^\dagger = (B^\dagger A^\dagger - A^\dagger B^\dagger) \\ &= (BA - AB) = -[A, B]\ . \end{aligned} \tag{2.45}$$

Equivalently, if A and B are Hermitian, then $\{A, B\}$ is also Hermitian.

Definition Compatible Observables

Observables A and B are defined to be compatible, if the corresponding operators commute, i. e.

$$[A, B] = 0 \tag{2.46}$$

and incompatible, if $[A, B] \neq 0$.

If the observables A and B are compatible, then A measurements and B measurements do not interfere, as we will see below.

An important example for incompatible observables are S_x and S_y. In detail, the spin operators obey the commutation relations

$$[S_i, S_j] = i\hbar \sum_k \varepsilon_{ijk} S_k\ , \tag{2.47}$$

and the anti-commutation relations

$$\{S_i, S_j\} = \frac{\hbar^2}{2}\delta_{ij}\ , \tag{2.48}$$

where the total anti-symmetric Levi-Civita tensor ε_{ijk} is defined via the unit basis-vectors $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ of a right-handed Cartesian coordinate system as

$$\vec{e}_i \times \vec{e}_j = \sum_k \varepsilon_{ijk}\vec{e}_k. \tag{2.49}$$

Note that S_z and $\mathbf{S}^2 \equiv \sum_k S_k^2$ are in fact compatible.

Theorem Representation of Compatible Observables

Suppose that A and B are compatible observables and the eigenvalues of A are nondegenerate, i. e. $a_i \neq a_j \ \forall i \neq j$, then the matrix elements $\langle \alpha_i | B | \alpha_j \rangle$ are all diagonal.

Thus, both operators have a common set of eigenkets, their corresponding matrix representations can be diagonalized simultaneously.

Proof

Simple.

□

Immediately we see that A and B are diagonalized simultaneously. Suppose B acts on an eigenket of A.

$$\begin{aligned} B\,|\alpha_i\rangle &= \sum_k |\alpha_k\rangle \langle \alpha_k | B | \alpha_k \rangle \langle \alpha_k | \alpha_i \rangle \\ &= \left(\langle \alpha_i | B | \alpha_i \rangle \right) |\alpha_i\rangle \end{aligned} \tag{2.52}$$

This is just an eigenvalue equation for the operator B with eigenvalue

$$\beta_i = \langle \alpha_i | B | \alpha_i \rangle \tag{2.53}$$

The ket $|\alpha_i\rangle$ is therefore a simultaneous eigenket of A and B which might be characterized by $|\alpha_i, \beta_i\rangle$.

Remark. The theorem also holds for degenerate eigenvalues.

Example

Orbital Angular Momentum.

Assume that $A, B, C, \ldots$ with

$$[A, B] = [A, C] = [B, C] = \ldots = 0 \tag{2.54}$$

form a *maximal set* of compatible observables which means that we cannot add any more observables to our list without violating (2.54). Then the (collective) index $K_i = (a_i, b_i, c_i, \ldots)$ uniquely specifies the eigenket

$$|K_i\rangle = |a_i, b_i, c_i, \ldots\rangle \tag{2.55}$$

The completenes relation implies that

$$\sum_i |K_i\rangle \langle K_i| = \sum_{a_i} \sum_{b_i} \sum_{c_i} \ldots |a_i, b_i, c_i, \ldots\rangle \langle a_i, b_i, c_i, \ldots| = \mathbf{1} \,. \tag{2.56}$$

As an example, we consider again the angular momentum

$$\mathbf{1} = \sum_l \sum_{|m| \leq l} |lm\rangle \langle lm| \tag{2.57}$$

What does it mean when two operators are compatible or not?

Consider a successive measurement of compatible observables.

$$|\alpha\rangle \xrightarrow{A} |a_i, b_i\rangle \xrightarrow{B} |a_i, b_i\rangle \xrightarrow{A} |a_i, b_i\rangle \tag{2.58}$$

Thus, A and B measurements do not interfere, if A and B are compatible observables.

Now, imagine an experiment with a sequential selective measurement of *incompatible observables*.

$$|\alpha\rangle \xrightarrow{A} |a_i\rangle \xrightarrow{B} |b_j\rangle \xrightarrow{C} |c_k\rangle \tag{2.59}$$

The probability to find $|c_k\rangle$ (provided that $|c_k\rangle$ is normalized) is given by

$$P^{b_j}(c_k) = |\langle c_k|b_j\rangle|^2 \cdot |\langle b_j|a_i\rangle|^2 \tag{2.60}$$

We sum over all b_j to get the total probability for going through all possible b_j routes.

$$P(c_k) = \sum_j |\langle c_k|b_j\rangle|^2 \, |\langle b_j|a_i\rangle|^2 = \sum_j \langle c_k|b_j\rangle \langle b_j|a_i\rangle \langle a_i|b_j\rangle \langle b_j|c_k\rangle \tag{2.61}$$

What happens when the filter B is switched off, i.e., when the measurement of the observable B is not performed?

$$\begin{aligned} P'(c_k) &= |\langle c_k|a_i\rangle|^2 = \left|\sum_j \langle c_k|b_j\rangle \langle b_j|a_i\rangle\right|^2 \\ &= \sum_j \sum_l \langle c_k|b_j\rangle \langle b_j|a_i\rangle \langle a_i|b_l\rangle \langle b_l|c_k\rangle \end{aligned} \tag{2.62}$$

$P(c_k)$ and $P'(c_k)$ are *different* (double sum vs. single sum)! The important result is that it matters whether or not B is switched on.

2.5.2 Uncertainty Relation

First, we derive the more general uncertainty relation, then – later on – Heisenberg's Uncertainty-Relation as a special case of this uncertainty relation.

Given an observable A, we define ΔA as

$$\Delta A \equiv A - \langle A \rangle \tag{2.63}$$

Then we can define the *dispersion*

$$\left\langle (\Delta A)^2 \right\rangle = \left\langle \left(A^2 - 2A\langle A \rangle + \langle A \rangle^2\right) \right\rangle = \langle A^2 \rangle - \langle A \rangle^2 \tag{2.64}$$

which is sometimes also called *variance* or mean square deviation in accordance with probabilistic theory.

Remark. Sometimes ΔA denotes

$$\Delta A \equiv \sqrt{\langle (A - \langle A \rangle)^2 \rangle} \tag{2.65}$$

Example

Theorem Uncertainty Relation

$$\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \geq \frac{1}{4} \left| \langle [A, B] \rangle \right|^2 \tag{2.68}$$

If the observables do not commute, then there is some inherent "fuzziness" in the measurements.

For the proof, we need two lemmas.

Theorem Schwarz Inequality

$$\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq \left| \langle \alpha | \beta \rangle \right|^2 \tag{2.69}$$

Proof

□

Theorem

The expectation value of a Hermitian operator is purely real. The expectation value of an anti-Hermitian operator is purely imaginary.

The proofs are trivial.

Now we are in a position to prove the Uncertainty Relation.

Proof

We use the Schwarz Inequality.

$$|\alpha\rangle = \Delta A\,|\Box\rangle \qquad |\beta\rangle = \Delta B\,|\Box\rangle \tag{2.72}$$

In conjunction with the Hermiticity of A and B, we get

$$\left\langle(\Delta A)^2\right\rangle\left\langle(\Delta B)^2\right\rangle \geq |\langle\Delta A\,\Delta B\rangle|^2 \tag{2.73}$$

We can decompose $\Delta A \Delta B$ as the sum of half the commutator plus half of the anti-commutator of both operators[2], i.e. an anti-Hermitian part and a Hermitian part.

$$\Delta A\,\Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\} \tag{2.74}$$

$$\Rightarrow\ \langle\Delta A\,\Delta B\rangle = \frac{1}{2}\ \langle[\Delta A, \Delta B]\rangle + \frac{1}{2}\ \langle\{\Delta A, \Delta B\}\rangle \tag{2.75}$$

$$= \frac{1}{2}\ \underbrace{\langle[A, B]\rangle}_{\text{purely imaginary}} + \frac{1}{2}\underbrace{\langle\{\Delta A, \Delta B\}\rangle}_{\text{purely real}} \tag{2.76}$$

---

[2] similarly, we can decompose any real function in an odd and an even part

In order to get the last line, we have used that $[\Delta A, \Delta B] = [A, B]$. Now let us take the square modulus

$$\begin{aligned} |\langle \Delta A\, \Delta B \rangle|^2 &= \frac{1}{4} |\langle [A, B] \rangle|^2 + \frac{1}{4} |\langle \{\Delta A, \Delta B\} \rangle|^2 \\ &\geq \frac{1}{4} |\langle [A, B] \rangle|^2 \end{aligned} \tag{2.77}$$

That leads to the uncertainty relation. □

In the next section (sect. 2.6.3) we will see that the Heisenberg Uncertainty Principle is just a special case of the uncertainty relation (2.68) for the position and the momentum operator.

2.5.3 Change of basis

Consider two incompatible observables A and B (if both are compatible, they have a common spectrum of eigenkets!). The ket space may either be spanned by the eigenkets $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$.

How are these basis related? We want to find the operator that connects the representation of A with the representation of B.

Theorem

Given two sets of base kets, both satisfying orthonormality (and completeness), there exists an *unitary operator* U such that

$$|b_i\rangle = U\,|a_i\rangle \qquad \forall i \tag{2.78}$$

Unitary means that the Hermitian of the operator is the inverse.

$$U^\dagger U = 1 = U U^\dagger \tag{2.79}$$

As a proof, we will explicitly construct this operator.

Proof

We prove this theorem by explicit construction. Let us assume that U is the operator defined by

$$U \equiv \sum_i |b_i\rangle \langle a_i| \tag{2.80}$$

(2.81)

Now we still have to prove the unitarity.

(2.82)

The other direction $UU^\dagger = 1$ can be shown analogously.

□

The matrix representation of U in the old $\{|a_i\rangle\}$ basis is

$$\langle a_i|U|a_j\rangle = \langle a_i|b_j\rangle \ , \tag{2.83}$$

i.e. it corresponds to the inner product of old base bras and new base kets.

Let us consider an arbitrary ket $|\alpha\rangle$.

$$|\alpha\rangle = \sum_k |a_k\rangle \langle a_k|\alpha\rangle \tag{2.84}$$

The coefficients in the new basis are given by

$$\begin{aligned} \langle b_l|\alpha\rangle &= \sum_k \langle b_l|a_k\rangle \langle a_k|\alpha\rangle \\ &= \sum_k \langle a_l|U^\dagger|a_k\rangle \langle a_k|\alpha\rangle \ , \end{aligned} \tag{2.85}$$

i.e., the column matrix for $|\alpha\rangle$ in the new basis can be obtained by applying the square matrix $U^\dagger$ to the column matrix in the old basis:

$$(\text{new}) = U^\dagger(\text{old}) \tag{2.86}$$

In order to understand why $U^\dagger$ appears in eq. (2.86), one should note that if for example the coordinate system is rotated in real space, then the coordinates of vectors can be obtained by effectively rotating the vector by the same angle in the opposite direction which is represented by the transposed rotation matrix.

It is equally simple to show that the matrix elements of an operator X transforms according to

$$X' = U^\dagger X U \tag{2.87}$$

Quantities that are invariant under transformations are of special importance in physics. The *trace* of the operator is such an invariant quantity.

$$\text{tr}\, X \equiv \sum_i \langle a_i|X|a_i\rangle \ . \tag{2.88}$$

Proof

$$\begin{aligned}\mathrm{tr}\,X &\equiv \sum_i \langle a_i|X|a_i\rangle \\ &= \\ &= \\ &= \\ &= \sum_k \langle b_k|X|b_k\rangle \qquad (2.89)\end{aligned}$$

□

Further properties of the trace:

$$\mathrm{tr}\,(XY) = \mathrm{tr}\,(YX) \qquad (2.90)$$

$$\mathrm{tr}\,(U^\dagger X U) = \mathrm{tr}\,X \qquad (2.91)$$

$$\mathrm{tr}\,(|a_i\rangle\,\langle a_j|) = \delta_{ij}\ , \qquad \mathrm{tr}\,(|b_i\rangle\,\langle a_j|) = \langle a_j|b_i\rangle \qquad (2.92)$$

The *determinant* is also such an invariant quantity.

$$\det X := \ldots = \Pi_i \lambda_i\ , \qquad (2.93)$$

where λ_i are the eigenvalues of the operator.

2.5.4 Diagonalization

A matrix in its eigenbasis is diagonalized. Finding the eigenvalues and the eigenkets of an operator B is equivalent to finding the unitary matrix U that diagonalizes B. In fact, U consists of the eigenkets of B, as we will show explicitly.

Consider the eigenvalue equation

$$B\,|b_i\rangle = b_i\,|b_i\rangle\ . \qquad (2.94)$$

This can be rewritten as

$$\langle a_j|B|b_i\rangle = \sum_k \langle a_j|B|a_k\rangle\,\langle a_k|b_i\rangle = b_i\,\langle a_j|b_i\rangle\,. \qquad (2.95)$$

$|b_i\rangle$ in (2.94) stands for the ith eigenket of operator B. Thus we can write (2.95) in matrix notation using the $\{|a_i\rangle\}$ basis as

$$\begin{pmatrix} B_{11} & B_{12} & \ldots \\ B_{21} & B_{22} & \ldots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1^{(i)} \\ c_2^{(i)} \\ \vdots \end{pmatrix} = \qquad (2.96)$$

with

$$B_{nm} = \langle a_n|B|b_m\rangle \ , \tag{2.97}$$

and

$$c_n^{(i)} = \langle a_n|b_i\rangle \ . \tag{2.98}$$

From linear algebra one knows that nontrivial solutions for $c_n^{(i)}$ are only possible if the characteristic equation

$$\det(B - \lambda 1) = 0 \tag{2.99}$$

is satisified. The roots of (2.99) can be identified with the eigenvalues b_i. The corresponding eigenvectors with the coefficients $c_n^{(i)}$ can then be determined up to an overall constant that can be derived from the normalization condition. Comparing (2.96) with (2.83) we see that the elements of the unitary matrix U performing the transformation are just given by the coefficients $c_n^{(i)}$, i.e., U can be constructed from the column vectors formed by the eigenvectors of B:

$$U = \begin{pmatrix} U_{11} & U_{12} & U_{13} & \dots \\ U_{21} & U_{22} & U_{23} & \dots \\ U_{31} & U_{32} & U_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \left(\begin{array}{c|c|c|c} c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \dots \\ c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \dots \\ c_3^{(1)} & c_3^{(2)} & c_3^{(3)} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{array}\right) \tag{2.100}$$

2.6 Position, Momentum, and Translation

The observables so far have been assumed to exhibit discrete eigenvalue spectra. However, in quantum mechanics observable can have a continuous spectrum (i.e. a continuous set of eigenvalues). An example is p_z, the z component of the momentum.

Consider the eigenvalue equation with a continuous spectrum.

$$W\,|w\rangle = w\,|w\rangle \ , \tag{2.101}$$

where W is an operator and w is simply a number. What is the analogy to discrete spectra? A vector space spanned by eigenkets with a continuous spectrum has an infinite dimension. Many of the results derived for a finite-dimensional vector space with discrete eigenvalues can be generalized. Replace the Kronecker symbol δ_{ij} by the Dirac function $\delta(w)$. Replace the sum over the eigenvalues by an integral over a continuous variable.

$$\langle a_i|a_j\rangle = \delta_{ij} \iff \langle w'|w''\rangle = \delta(w' - w'') \tag{2.102}$$

$$\sum_i |a_i\rangle\,\langle a_i| = \mathbf{1} \iff \int |w'\rangle\,\langle w'|\; dw' = \mathbf{1} \tag{2.103}$$

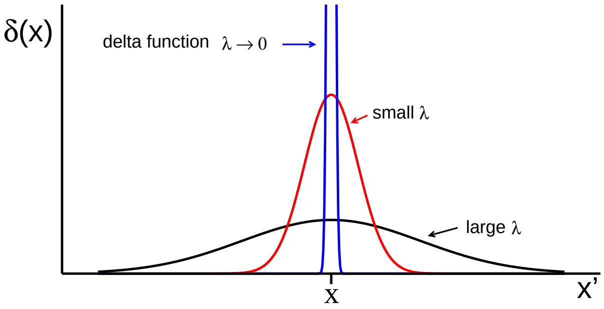


Figure 2.2: Delta function as a limit of Gaussian functions (2.114) for $\lambda \to 0$.

There are further analogies:

$$|\alpha\rangle = \sum_i |a_i\rangle \langle a_i|\alpha\rangle \iff |\alpha\rangle = \int |\omega\rangle \langle\omega|\alpha\rangle \; d\omega \tag{2.104}$$

$$\sum_i |\langle a_i|\alpha\rangle|^2 = 1 \iff \int \langle\alpha|\omega\rangle \langle\omega|\alpha\rangle \; d\omega = \int |\langle\alpha|\omega\rangle|^2 \; d\omega = 1 \tag{2.105}$$

$$\langle\beta|\alpha\rangle = \sum_i \langle\beta|a_i\rangle \langle a_i|\alpha\rangle \iff \langle\beta|\alpha\rangle = \int \langle\beta|\omega\rangle \langle\omega|\alpha\rangle \; d\omega \tag{2.106}$$

$$\langle a_i|A|a_j\rangle = a_i\delta_{ij} \iff \langle\omega'|W|\omega''\rangle = \omega'\delta(\omega' - \omega'') \tag{2.107}$$

2.6.1 Digression on the Dirac Delta function

The Dirac δ function is in fact a *distribution* with the properties

$$\delta(x - x') = 0 \qquad x \neq x' \tag{2.108}$$

$$\int_{\mathbb{R}} \delta(x - x') f(x') \, dx' = f(x) \qquad f \text{ arbitrary function} \tag{2.109}$$

$$\int_{\mathbb{R}} \delta(x - x') \, dx = 1 \tag{2.110}$$

The dimension of δ is given by the inverse dimension of its argument, so if for example x is a length, then $\delta(x)$ has the dimension 1/length. Furthermore, by doing partial integration, we get an expression for the derivative.

$$\int_{\mathbb{R}} f(x')\delta'(x - x') \, dx' = -\int_{\mathbb{R}} f'(x')\delta(x - x') \, dx' = -f'(x) \tag{2.111}$$

$$\delta(ax) = \frac{1}{|a|}\delta(x) \tag{2.112}$$

$$\delta(g(x)) = \sum_{\substack{g(a)=0 \\ g'(a)\neq 0}} \frac{\delta(x - a)}{|g'(x)|} \tag{2.113}$$

The δ function can be expressed as a limit of a real (in the sense of existing) function.

For example, the Gaussian function (see Fig. 2.2)

$$f_\lambda(x') \;=\; \frac{1}{\lambda\sqrt{2\pi}} \exp\left(-\frac{(x'-x)^2}{2\lambda^2}\right) \; . \tag{2.114}$$

is peaked around $x' = x$ with a width $\approx \lambda$ and unit area under the curve. The delta function can be regarded as the limit of the Gaussion function for $\lambda \to 0$.

2.6.2 Position and momentum eigenkets

We will now consider specific operators with a continuous eigen spectrum, namely the position and the momentum operators. To avoid confusion, we will denote operators with a hat, e. g. $\hat{x}$. We will first derive the properties of the operators in one dimension and then later generalize the results for three dimensions.

The eigenkets $|x'\rangle$ of the position operator $\hat{x}$ satisfy

$$\hat{x}\,|x'\rangle = x'\,|x'\rangle \tag{2.115}$$

The eigenkets $|x'\rangle$ are postulated to form a complete set which means that any physical state can be decomposed into an integral over the eigenstates.

$$|\alpha\rangle = \int_{\mathbb{R}} |x\rangle\,\langle x|\alpha\rangle \; dx \tag{2.116}$$

Now, the probability to find the particle in an intervall dx is

$$|\langle x|\alpha\rangle|^2 \; dx = \langle x|\alpha\rangle\,\langle\alpha|x\rangle \; dx \tag{2.117}$$

In this formalism, the wave function is identified with the inner product.

$$\langle x|\alpha\rangle \equiv \psi_\alpha(x) \tag{2.118}$$

This is the function $\psi_\alpha(x)$ of the state $|\alpha\rangle$; in this formalism based on Dirac it can be regarded as an expansion coefficient.

The comparison of (2.118) with the orthogonality relation (2.102) for the position eigenkets,

$$\langle x'|x''\rangle = \delta(x'-x'') \; , \tag{2.119}$$

leads to the conclusion, that the δ function can be regarded as the position operator eigenfunctions.

The inner product $\langle\beta|\alpha\rangle$ now becomes the overlap between two wave functions; equivalently, it can be interpreted as the probability amplitude to find $|\beta\rangle$ in $|\alpha\rangle$.

$$\langle\beta|\alpha\rangle = \int_{\mathbb{R}} \langle\beta|x\rangle\,\langle x|\alpha\rangle \; dx = \int_{\mathbb{R}} \psi^*_\beta(x)\psi_\alpha(x)\,dx \tag{2.120}$$

Now we introduce the momentum operator $\hat{p}$.

$$\hat{p}\,|p'\rangle = p'\,|p'\rangle \qquad \langle p'|p''\rangle = \delta(p'-p'') \tag{2.121}$$

$\{|p\rangle\}$ span the ket space in much the same way as $\{|x\rangle\}$:

$$|\alpha\rangle = \int |p'\rangle\,\langle p'|\alpha\rangle \; dp' \tag{2.122}$$

with the momentum space function

$$\langle p'|\alpha\rangle = \phi_\alpha(p') \ . \tag{2.123}$$

In fact, momentum can be regarded as a generator for translations. We define the infinitesimal translation operator $\mathcal{T}$.

$$\mathcal{T}(dx)\,|x\rangle = |x+dx\rangle \tag{2.124}$$

$$\mathcal{T}(dx) \equiv 1 - i\frac{\mathbf{p}\cdot d\mathbf{x}}{\hbar} \tag{2.125}$$

Here we have in fact already used the generalisation to three dimensions. For a finite translation Δx, we can subsequently apply infinitesimal translations. The result is given by

$$\begin{aligned}\mathcal{T}(\Delta x) &= \lim_{n\to\infty} \mathcal{T}^n(\Delta x/n)\\ &= \\ \\ &= \exp\left(-\frac{i}{\hbar}\mathbf{p}\cdot\Delta\mathbf{x}\right)\end{aligned} \tag{2.126}$$

2.6.3 Canonical Commutation Relations

Recall the *Poisson brackets* in classical mechanics:

$$\big\{A(q,p),B(q,p)\big\} = \sum_s \frac{\partial A}{\partial q_s}\frac{\partial B}{\partial p_s} - \frac{\partial B}{\partial q_s}\frac{\partial A}{\partial p_s} \tag{2.127}$$

where q_s and p_s are the generalized space and momentum coordinates of a s-dimensional configuration space. Note thate the dimension of the Poisson bracket is the dimension of A times B devided by the dimension of action.

The fundamental Poisson brackets for canonical coordinates (i. e. coordinates that satisfy the following relations) in $\mathbb{R}^3$ are

$$\big\{x_i,x_j\big\} = 0 \qquad \big\{p_i,p_j\big\} = 0 \qquad \big\{x_i,p_j\big\} = \delta_{ij} \tag{2.128}$$

In quantum mechanics the Poisson brackets become the commutator between operators.

$$\big[\hat{x}_i,\hat{x}_j\big] = 0 \qquad \big[\hat{p}_i,\hat{p}_j\big] = 0 \qquad \big[\hat{x}_i,\hat{p}_j\big] = i\hbar\delta_{ij} \tag{2.129}$$

Using the Uncertainty Relation, the Heisenberg Uncertainty Principle follows:

$$\big\langle(\Delta x)^2\big\rangle\,\big\langle(\Delta p_x)^2\big\rangle \geq \frac{1}{4}\,|\langle[x,p_x]\rangle|^2 = \frac{\hbar^2}{4} \tag{2.130}$$

It is often written as

$$\Delta x\,\Delta p_x \geq \frac{\hbar}{2} \tag{2.131}$$

Either 2.129 or the Uncertainty Relation has to be postulated to get the uncertainty principle – one follows from the other.

In general

$$
\begin{aligned}
\langle\beta|A|\alpha\rangle &= \int\int \langle\beta|x\rangle\,\langle x|A|x'\rangle\,\langle x'|\alpha\rangle\;dx\,dx' \\
&= \int\int \psi_\beta^*(x)\,\langle x|A|x'\rangle\,\psi_\alpha(x')\,dx\,dx'
\end{aligned}
\tag{2.132}
$$

Consider for example $A = \hat{x}^2$:

$$
\begin{aligned}
\left\langle\beta|\hat{x}^2|\alpha\right\rangle &= \\
\\
\\
&=
\end{aligned}
\tag{2.133}
$$

For an arbitrary function $f(\hat{x})$ of the operator $\hat{x}$, we have

$$
\langle\beta|f(\hat{x})|\alpha\rangle = \int \psi_\beta^*(x) f(x) \psi_\alpha(x)\,dx
\tag{2.134}
$$

As far as the momentum operator is concerned, it follows that the matrix element $\langle\beta|f(\hat{p})|\alpha\rangle$ in the position representation is expressed as

$$
\langle\beta|f(\hat{p})|\alpha\rangle = \int \phi_\beta^*(p) f(p) \phi_\alpha(p)\,dp
\tag{2.135}
$$

However, how does the momentum operator, i.e. $\langle\beta|\hat{p}|\alpha\rangle$, look like in the position representation?

In order to answer this question, we will first consider $\langle x'|\hat{p}|x''\rangle$. By using the fundamental commutation relation $[\hat{x},\hat{p}] = i\hbar$, it can be shown, as derived in the appendix on page 37, that

$$
\langle x''|\hat{p}|x'\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x''}\delta(x''-x')
\tag{2.136}
$$

It immediately follows that

$$
\begin{aligned}
\langle x|\hat{p}|\alpha\rangle &= \int \langle x|\hat{p}|x'\rangle\,\langle x'|\alpha\rangle\,dx' \\
&= \frac{\hbar}{i} \\
&= \frac{\hbar}{i} \\
&= \frac{\hbar}{i}\frac{\partial}{\partial x}\psi_\alpha(x)
\end{aligned}
\tag{2.137}
$$

This shows that the momentum operator in the position space representation corresponds to performing the partial derivative with respect to x and multiplication by $-i\hbar$.

We list some matrix elements of the momentum operator in the position representation:

$$\Rightarrow \ \langle x|\hat{p}|\alpha\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\psi_\alpha(x) \tag{2.138}$$

$$\Rightarrow \ \langle \beta|\hat{p}|\alpha\rangle = \int \psi^*_\beta(x')\frac{\hbar}{i}\frac{\partial}{\partial x'}\psi_\alpha(x')\, dx' \tag{2.139}$$

$$\Rightarrow \ \langle x|\hat{p}^n|\alpha\rangle = \left(\frac{\hbar}{i}\right)^n \frac{\partial^n}{\partial x^n}\psi_\alpha(x) \tag{2.140}$$

$$\Rightarrow \ \langle \beta|\hat{p}^n|\alpha\rangle = \int \psi^*_\beta(x')\left(\frac{\hbar}{i}\right)^n \frac{\partial^n}{\partial x'^n}\psi_\alpha(x')\, dx' \tag{2.141}$$

This means that in this derivation we do not postulate the correspondence $\mathbf{p} \rightarrow \frac{\hbar}{i}\nabla$, we simply require that the fundamental canonical commutation relations hold. We still obtain the same representation of the momentum operator.

2.7 Momentum-Space Wave Function

Instead of the position representation we can work equally well in the momentum representation. Haw are the wave functions in both representations related? To make a long story short, the answer is that both are the Fourier transform of each other.

Let's start by considering the basis of position space. The δ functions are the eigenfunctions of the position operator in position space. There must also be an eigenfunction for the position operator.

$\langle x'|p'\rangle$ induces a transformation function from the x representation to the p representation.

$$\begin{aligned}\langle x'|\hat{p}|p'\rangle &= \frac{\hbar}{i}\frac{\partial}{\partial x'}\langle x'|p'\rangle \\ &= p'\langle x'|p'\rangle \qquad\qquad |p'\rangle \text{ is eigenfunction}\end{aligned} \tag{2.142}$$

This is a differential equation for p. The solution is an exponential function.

$$\langle x'|p'\rangle = Ce^{\frac{i}{\hbar}p'x'}\ , \tag{2.143}$$

where C is a normalization constant.

$$\begin{aligned}\delta(x'-x'') &= \\ \\ &= \end{aligned} \tag{2.144}$$

$$\Rightarrow \ |C|^2 = \frac{1}{2\pi\hbar} \tag{2.145}$$

C can be a complex number; but we can choose any phase that we want, so we set C to be a real positive number.

$$\begin{aligned} C &= \frac{1}{\sqrt{2\pi\hbar}} \\ \Rightarrow \ \langle x'|p'\rangle &= \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p'x'}\end{aligned} \tag{2.146}$$

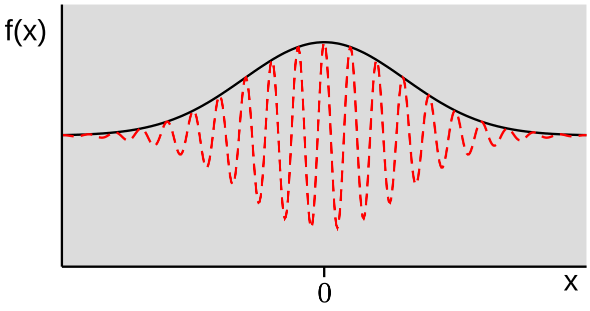


Figure 2.3: A Gaussian wave packet as described by eq. (2.149).

This is just a plane wave.

We will now consider the consequences of this particular form of the momentum eigenfunction for the representation of the wave functions. The wave function $\psi_\alpha(x')$ in position space can be rewritten as

$$\begin{aligned}\psi_\alpha(x') = \langle x'|\alpha\rangle &= \int \langle x'|p'\rangle \langle p'|\alpha\rangle \; dp' \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int e^{\frac{i}{\hbar}p'x'} \phi_\alpha(p') \, dp' \end{aligned} \tag{2.147}$$

Equivalently, the momentum space representation of $|\alpha\rangle$ can be decomposed as

$$\begin{aligned}\phi_\alpha(p') = \langle p'|\alpha\rangle &= \int \langle p'|x'\rangle \langle x'|\alpha\rangle \; dx' \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int e^{-\frac{i}{\hbar}p'x'} \psi_\alpha(x') \, dx' \end{aligned} \tag{2.148}$$

This shows that representations ψ_α and ϕ_α are just related by a *Fourier transformation*!

A very important kind of wave functions are the *Gaussian wave packets*.

2.7.1 Gaussian Wave Packets

A very important kind of wave functions are the *Gaussian wave packets*. They can be regarded as a plane wave characterized by a wave number $k = 2\pi/\lambda$ that is enveloped by a Gaussian (see Fig. 2.3).

$$\langle x'|\alpha\rangle = \frac{1}{\pi^{1/4}\sqrt{d}} e^{ikx - \frac{x'^2}{2d^2}} \tag{2.149}$$

The expectation value

$$\langle x\rangle = \int_{\mathbb{R}} \langle \alpha|x'\rangle \, x' \, \langle x'|\alpha\rangle \; dx' = \int_{\mathbb{R}} |\langle x'|\alpha\rangle|^2 \, x' dx' = 0 \; . \tag{2.150}$$

vanishes for symmetry reasons since $\langle x'|\alpha\rangle$ is an even function. Furthermore,

$$\langle x^2\rangle = \int_{\mathbb{R}} |\langle x'|\alpha\rangle|^2 \, x'^2 \, dx' = \frac{1}{\sqrt{\pi}d} \int_{\mathbb{R}} x'^2 e^{-\frac{x'^2}{d^2}} \, dx' = \frac{d^2}{2} \; . \tag{2.151}$$

Now we can determine the variance of the wave packet.

$$\left\langle(\Delta x)^2\right\rangle = \left\langle x^2\right\rangle - \langle x\rangle^2 = \frac{d^2}{2} \tag{2.152}$$

It is left to the student as an exercise to confirm the result that

$$\left\langle p^2\right\rangle = \tag{2.153}$$

$$\left\langle(\Delta p)^2\right\rangle = \left\langle p^2\right\rangle - \langle p\rangle^2 = \tag{2.154}$$

The uncertainty relation now reads as follows:

$$\left\langle(\Delta x)^2\right\rangle\left\langle(\Delta p)^2\right\rangle = \frac{\hbar^2}{4} \tag{2.155}$$

This shows that the Gaussian wave packet is a *minimum uncertainty wave packet*, i.e., it has the smallest possible uncertainty.

Let us now take a look at the Gaussian wave packet in momentum space:

$$\begin{aligned}\langle p'|\alpha\rangle &= \frac{1}{\sqrt{2\pi\hbar}}\frac{1}{\pi^{1/4}\sqrt{d}}\int_{\mathbb{R}} e^{-\frac{i}{\hbar}p'x'+ikx'-\frac{x'^2}{2d^2}}\,dx' \\ &= \frac{1}{\sqrt{2\pi\hbar}}\frac{1}{\pi^{\frac{1}{4}}\sqrt{d}}\int_{\mathbb{R}} e^{-\frac{1}{2}\left(\frac{x'}{d}+id\frac{p'-\hbar k}{\hbar}\right)^2-\frac{(p'-\hbar k)^2d^2}{2\hbar^2}}\,dx' \\ &= \sqrt{\frac{d}{\hbar\sqrt{\pi}}}e^{-\frac{(p'-\hbar k)^2d^2}{2\hbar^2}}\end{aligned} \tag{2.156}$$

It follows that the Fourier transformation of the Gaussian function in real space is a Gaussian in momentum space! It is centered at $\hbar k$ which is the mean momentum of the wave packet. Furthermore, the widths of Gaussian wave packets in real space and momentum space are inversely proportional.

2.7.2 Generalization to three dimensions

The position and momentum operator can be generalized to three dimensions as follows.

$$\hat{\mathbf{x}}\,|\mathbf{x}'\rangle = \mathbf{x}'\,|\mathbf{x}'\rangle \qquad \hat{\mathbf{p}}\,|\mathbf{p}'\rangle = \mathbf{p}'\,|\mathbf{p}'\rangle \tag{2.157}$$

$$\langle\mathbf{x}'|\mathbf{x}''\rangle = \delta(\mathbf{x}'-\mathbf{x}'') \tag{2.158}$$

$$\langle\beta|\hat{\mathbf{p}}|\alpha\rangle = \int \psi_\beta^*(\mathbf{x}')\frac{\hbar}{i}\nabla\psi_\alpha(\mathbf{x}')\,d^3x' \tag{2.159}$$

$$\langle\mathbf{x}'|\mathbf{p}'\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{\frac{i}{\hbar}\mathbf{p}'\cdot\mathbf{x}'} \tag{2.160}$$

2.7.3 Appendix: Position representation of the momentum operator

Here we will derive Eq. (2.162) which is important to determine the momentum operator, i.e. $\langle\beta|\hat{p}|\alpha\rangle$, in the position representation. We will first consider $\langle x'|\hat{p}|x''\rangle$ and use $[\hat{x}, \hat{p}] = i\hbar$.

$$\begin{aligned} \langle x''|[\hat{x},\hat{p}]|x'\rangle &= \langle x''|i\hbar|x'\rangle = i\hbar\delta(x'-x'') \\ &= \langle x''|\hat{x}\hat{p}-\hat{p}\hat{x}|x'\rangle = x''\,\langle x''|\hat{p}|x'\rangle - \langle x''|\hat{p}|x'\rangle\, x' \end{aligned} \tag{2.161}$$

This shows that $\langle x''|[\hat{x},\hat{p}]|x'\rangle$ vanishes for $x'' \neq x'$. Note that $\hat{p}$ still acts on both, $|x'\rangle$ and x'. A possible solution is

$$\langle x''|\hat{p}|x'\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x''}\delta(x''-x') \tag{2.162}$$

Proof

From (2.161) follows that

$$x''\,\langle x''|\hat{p}|x'\rangle - \langle x''|\hat{p}|x'\rangle\, x' = i\hbar\delta(x'-x'')\ , \tag{2.163}$$

i.e. in order to proove (2.162) we have to show that

$$-x''\frac{\partial}{\partial x''}\delta(x''-x') + \frac{\partial}{\partial x''}\delta(x''-x')x'' = \delta(x'-x'')\ . \tag{2.164}$$

Note that this looks like the product rule of differentation:

$$\frac{\partial}{\partial x''}\delta(x''-x')x'' = x''\frac{\partial}{\partial x''}\delta(x''-x') + \delta(x'-x'')\ . \tag{2.165}$$

However, since we have not proved yet that the usual differentiation rules also apply to Dirac's Delta function, we will now introduce an arbitrary function $f(x'')$ that is supposed to be differentiable in x''. We will then insert (2.164) and show that this leads to a correct result:

$$\begin{aligned} \Rightarrow f(x') &= \int f(x'')\delta(x'-x'')dx'' \\ &\overset{(2.164)}{=} \int f(x'')\left[-x''\frac{\partial}{\partial x''}\delta(x''-x') + \frac{\partial}{\partial x''}\delta(x''-x')x''\right]dx'' \\ &\overset{\text{part. int.}}{=} \int \left[\left(\frac{\partial}{\partial x''}f(x'')x''\right)\delta(x''-x') - \left(\frac{\partial}{\partial x''}f(x'')\right)\delta(x''-x')x''\right]dx'' \\ &= \left(\frac{\partial}{\partial x'}f(x')x'\right) - \left(\frac{\partial}{\partial x'}f(x')\right)x' \overset{\text{product rule}}{=} f(x') \end{aligned} \tag{2.166}$$

Since $f(x')$ is arbitrary, this means that (2.164) is correct. □

Note that

$$\langle x''|\hat{p}|x'\rangle = \frac{\hbar}{i}\left(\frac{\partial}{\partial x''} + f(x'')\right)\delta(x''-x') \tag{2.167}$$

is also a possible solution of 2.161. But one can show that this only causes a different overall phase for the wave function which still leads to the same state.

3 Quantum Dynamics

3.1 Time Evolution and the Schrödinger Equation

3.1.1 Time Evolution Operator

We first specify the notation. Assume that a physical system is in state α at a point in time t_0. Then the time evolution of the state is denoted by

$$|\alpha, t_0\rangle = |\alpha\rangle \xrightarrow{\text{time evolution}} |\alpha, t_0; t\rangle \tag{3.1}$$

The change of the system can be described by the time evolution operator $U(t, t_0)$.

$$|\alpha, t_0; t\rangle = U(t, t_0)\,|\alpha, t_0\rangle \tag{3.2}$$

We expand α into state kets.

$$\begin{aligned} |\alpha, t_0\rangle &= \sum_i c_i(t_0)\,|a_i\rangle \\ |\alpha, t\rangle &= \sum_i c_i(t)\,|a_i\rangle \end{aligned} \tag{3.3}$$

However, we require that the state $|\alpha, t_0; t\rangle$ remains normalized at all times. Thus,

$$\begin{aligned} \langle \alpha, t_0 | \alpha, t_0\rangle = 1 &= \sum_i |c_i(t_0)|^2 \\ &= \sum_i |c_i(t)|^2 = \langle \alpha, t_0; t | \alpha, t_0; t\rangle = \left\langle \alpha, t_0 | U^\dagger U | \alpha, t_0 \right\rangle \end{aligned} \tag{3.4}$$

This can only be satisfied if U is a *unitary operator* ($U(t, t_0)^\dagger U(t, t_0) = \mathbf{1}$), i.e., we require U to be unitary.

Furthermore, we require the so-called *composition property*:

$$U(t_2, t_0) = U(t_2, t_1) U(t_1, t_0) \qquad\qquad t_2 > t_1 > t_0 \tag{3.5}$$

Consider first the infinitesimal time evolution operator:

$$|\alpha, t_0; t_0 + dt\rangle = U(t_0 + dt, t_0)\,|\alpha, t_0\rangle \tag{3.6}$$

This operator must reduce to the unity operator as dt goes to zero; furthermore, it should be of first order in dt. These requirements are fulfilled by

$$U(t_0 + dt, t_0) := \mathbf{1} - i\Omega dt\ , \tag{3.7}$$

where Ω is a Hermitian operator.

Let's take a look at the composition property.

$$\begin{aligned} U(t_0+dt_1+dt_2,t_0+dt_1)U(t_0+dt_1,t_0) &= \big(1-i\Omega dt_2\big)\big(1-i\Omega dt_1\big) \\ &= \\ &\approx \\ &= U(t_0+dt_1+dt_2,t_0) \end{aligned} \tag{3.8}$$

Since we deal with *infinitesimal* time evolutions dt_i, terms of order $(dt)^2$ can be ignored. If Ω depends explicitly on time, it must be evaluated at t_0.

Now we check the unitarity:

$$U^\dagger(t_0+dt,t_0)U(t_0+dt,t_0) = \big(\mathbf{1}+i\Omega^\dagger dt\big)\big(\mathbf{1}-i\Omega dt\big) = \mathbf{1}+\Omega^2 dt^2 \approx \mathbf{1} \tag{3.9}$$

In order to further specify Ω we refer to classical mechanics. The Hamiltonian is the generator of time evolution. So, we will use the ansatz

$$\Omega = \frac{H}{\hbar} \tag{3.10}$$

$$U(t_0+dt,t_0) = 1-\frac{iHdt}{\hbar} \tag{3.11}$$

3.1.2 Derivation of the Schrödinger Equation

Use the composition property (3.5).

$$U(t+dt,t_0) = U(t+dt,t)U(t,t_0) = \left(\mathbf{1}-\frac{iHdt}{\hbar}\right)U(t,t_0) \tag{3.12}$$

$$U(t+dt,t_0)-U(t,t_0) = -\frac{iHdt}{\hbar}U(t,t_0)$$

$$\implies i\hbar\frac{U(t+dt,t_0)-U(t,t_0)}{dt} = i\hbar\frac{\partial}{\partial t}U(t,t_0) = HU(t,t_0) \tag{3.13}$$

This is the Schrödinger Equation for the time evolution operator. If we multiply (3.13) by $|\alpha,t_0\rangle$, we obtain

$$i\hbar\frac{\partial}{\partial t}U(t,t_0)\,|\alpha,t_0\rangle = HU(t,t_0)\,|\alpha,t_0\rangle \tag{3.14}$$

Since $|\alpha,t_0\rangle$ does not depend on t, this can be rewritten (using (3.2)) as

$$i\hbar\frac{\partial}{\partial t}\,|\alpha,t_0;t\rangle = H\,|\alpha,t_0;t\rangle \tag{3.15}$$

This is the *Schrödinger Equation* for a state ket. Any time evolution of a state ket is given by the time evolution operator $U(t,t_0)$. In order to determine $U(t,t_0)$, we have to first consider the formal solutions of the Schrödinger equation for the time evolution operator (3.13).

3.1.3 Formal Solution for $U(t,t_0)$

There are three cases.

Case 1 H is independent of time. Then U is

$$U(t,t_0) = e^{-\frac{i}{\hbar}H(t-t_0)} \tag{3.16}$$

Proof

Note that we have not introduced yet what the meaning of a function of an operator is. We simply use the series expansion of the function where only products of the operator appear. For example,

$$\exp A \;= \tag{3.17}$$

Thus we also expand the exponential function in eq. (3.16).

Alternative proof: We regard U as a compound of infinitesimal time-evolution operators.

$$\lim_{n\to\infty}\left[1-\frac{(iH/\hbar)(t-t_0)}{n}\right]^n = e^{-\frac{i}{\hbar}H(t-t_0)} \tag{3.20}$$

□

Case 2 The Hamiltonian is time-dependent, but $[H(t_1),H(t_2)] = 0$ for all t_1 and t_2.

Now the time evolution operator becomes

$$U(t,t_0) = e^{-i/\hbar\int_{t_0}^{t} H(t)\,dt} \tag{3.21}$$

The proof is similar to case 1, simply replace $H(t-t_0)$ by $\int_{t_0}^{t} H(t)\,dt$.

Case 3 We have a time-dependent evolution and the Hamiltonians do not commute at different points in time. Now we are facing the problem that we cannot simply use the integral $\int_{t_0}^{t} H(t)\,dt$ in the exponential since in general $\exp(A)\exp(B) \neq \exp(A+B)$ if A and B do not commute. In fact, one can show that if the commutator of the two operators A and B commutes with these operators, i.e. $[[A,B],A] = [[A,B],B] = 0$, then

$$e^{A+B} \;=\; e^A e^B e^{-[A,B]/2}\;. \tag{3.22}$$

For the time-evoluation operator, the general solution, if $[H(t_1), H(t_2)] \neq 0$ for $t_1 \neq t_2$, is given by

$$U(t,t_0) = 1 + \sum_{n\in\mathbb{N}} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t_1} \int_{t_1}^{t_2} \cdots \int_{t_{n-1}}^{t_n} H(t_1)H(t_2)\cdots H(t_n)\, dt_1\, dt_2 \cdots dt_n \tag{3.23}$$

This is the so-called *Dyson series*; the proof will be provided later.

Energy Eigenkets

In classical mechanics, the Hamiltonian H corresponds to the total energy of a system. Consequently, the eigenvalues of the Hamiltonian can be regarded as *energy eigenvalues*. Consider now the base kets of an observable A with

$$[A, H] = 0 \tag{3.24}$$

Then the eigenkets of A are also the *energy eigenkets* of H.

$$H |a_i\rangle = E_i |a_i\rangle \tag{3.25}$$

Expand the time-evolution operator.

$$e^{-\frac{i}{\hbar}Ht} = \sum_{i,j} |a_j\rangle \langle a_j| e^{-iHt/\hbar} |a_i\rangle \langle a_i| = \sum_i |a_i\rangle e^{-iE_it/\hbar} \langle a_i| \tag{3.26}$$

We can decompose an arbitrary ket $|\alpha, t_0 = 0\rangle$ into eigenkets,

$$|\alpha, t_0 = 0\rangle = \sum_i c_i |a_i\rangle \ , \tag{3.27}$$

with $c_i = \langle a_i|\alpha\rangle$. When applying the time evolution operator, we get

$$\begin{aligned} |\alpha, t_0 = 0; t\rangle = e^{-iHt/\hbar} |\alpha, t_0 = 0\rangle &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-iHt}{\hbar}\right)^n |\alpha, t_0 = 0\rangle \\ &= \sum_i \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-iHt}{\hbar}\right)^n |a_i\rangle \langle a_i|\alpha\rangle \\ &= \sum_i c_i \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-iE_it}{\hbar}\right)^n |a_i\rangle = \qquad (3.28) \\ \Rightarrow \ c_i(t) = c_i(t=0) e^{-iE_it/\hbar} & \qquad (3.29) \end{aligned}$$

Thus, the expansion coefficients change in time with their modulus unchanged.

Now let us assume that the initial state is an eigenstate of A:

$$|\alpha, t_0 = 0\rangle = |a_j\rangle \tag{3.30}$$

Then the time evolution is simply given by

$$|\alpha, t_0 = 0; t\rangle = e^{-iE_jt/\hbar} |a_j\rangle \tag{3.31}$$

Thus, the system remains in the state $|a_j\rangle$ at all times, just the phase is modulated. If an observable is compatible with the Hamiltonian H, then it is a constant of motion.

Time Dependence of Expectation Values

It is easy to show that the expectation value of any observable with respect to *an eigenket* does not change under time evolution:

$$\begin{aligned} \langle B \rangle &= \langle \alpha, t_0 = 0; t | B | \alpha, t_0 = 0; t \rangle \\ &= \left\langle a_j | e^{iE_j t/\hbar} B e^{-iE_j t/\hbar} | a_j \right\rangle \\ &= \langle a_j | B | a_j \rangle \end{aligned} \tag{3.32}$$

Therefore eigenkets of H are also called *stationary kets*.

Consider now the decomposition of a ket into stationary kets.

$$\begin{aligned} |\alpha, t_0 = 0\rangle &= \sum_j c_j \, |a_j\rangle \\ \Rightarrow \langle B \rangle &= \Big[\sum_i c_i^* \, \langle a_i| \, e^{+iE_i t/\hbar}\Big] B \Big[\sum_j c_j \, |a_j\rangle \, e^{-iE_j t/\hbar}\Big] \\ &= \sum_{i,j} c_i^* c_j \, \langle a_i | B | a_j \rangle \, e^{-i(E_j - E_i)t/\hbar} \end{aligned} \tag{3.33}$$

This is a sum over oscillatory terms with frequency

$$\omega_{ij} = \frac{E_j - E_i}{\hbar} \tag{3.34}$$

Example: Spin Precession

Consider a spin-1/2 system in an external magnetic field.

$$H = -\frac{e}{m_e c} \mathbf{S} \cdot \mathbf{B} \tag{3.35}$$

Here, $e < 0$ for the electron, $\mathbf{B} = B\mathbf{e}_3$.

$$H = -\frac{eB}{m_e c} S_z \tag{3.36}$$

As we can directly see, S_z commutes with H, they have simultaneous eigenkets.

$$E_\pm = \mp \frac{e\hbar B}{2m_e c} \tag{3.37}$$

Define the cyclotron frequency $\omega = \frac{|e|B}{m_e c}$. From that, the Hamiltonian simplifies to $H = \omega S_z$. This means that the time-evolution operator is simply given by

$$U(t, 0) = \exp\left(\frac{-i\omega S_z t}{\hbar}\right) \ . \tag{3.38}$$

Any general state can now be decomposed into the two spin eigenkets. How does it change under time-evolution?

$$\begin{aligned} |\alpha\rangle &= c_+ \, |+\rangle + c_- \, |-\rangle \\ \Rightarrow |\alpha, t_0 = 0; t\rangle &= c_+ e^{-i\omega t/2} \, |+\rangle + c_- e^{+i\omega t/2} \, |-\rangle \end{aligned} \tag{3.39}$$

Let $|\alpha\rangle$ be the $|S_x,+\rangle$ state, i.e.,

$$|\alpha\rangle = |S_x,+\rangle = \tfrac{1}{\sqrt{2}}\big(|+\rangle + |-\rangle\big) \tag{3.40}$$

Then the probability to find the $S_x\pm$ states as a function of time can be obtained via

$$\begin{aligned}|\langle S_x,\pm|\alpha,t_0=0;t\rangle|^2 &= \left|\tfrac{1}{\sqrt{2}}\big(\langle+|\pm\langle-|\big)\tfrac{1}{\sqrt{2}}\big(e^{-i\omega t/2}\,|+\rangle + e^{+i\omega t/2}\,|-\rangle\big)\right|^2\\ &= \left|\tfrac{1}{2}\big(e^{-i\omega t/2}\pm e^{+i\omega t/2}\big)\right|^2 = \begin{cases}\cos^2\frac{\omega}{2}t & \text{for } S_x+\\ \sin^2\frac{\omega}{2}t & \text{for } S_x-\end{cases}\end{aligned} \tag{3.41}$$

According to Eq. (2.43) ($\langle A\rangle = \sum_j a_j\,|\langle\alpha_j|\alpha\rangle|^2$), the expectation value of S_x as a function of time can be expressed as

$$\begin{aligned}\langle S_x\rangle &= \frac{\hbar}{2}\,|\langle S_x,+|\alpha,t_0=0;t\rangle|^2 - \frac{\hbar}{2}\,|\langle S_x,-|\alpha,t_0=0;t\rangle|^2\\ &= \frac{\hbar}{2}\big(\cos^2\frac{\omega}{2}t - \sin^2\frac{\omega}{2}t\big) = \frac{\hbar}{2}\cos\omega t\end{aligned} \tag{3.42}$$

Analogously, we get for the other two expectation values

$$\langle S_y\rangle = \frac{\hbar}{2}\sin\omega t \tag{3.43}$$

$$\langle S_z\rangle = 0 \tag{3.44}$$

Thus the spin precesses in the xy plane with an angular frequency ω.

3.1.4 Schrödinger versus Heisenberg Picture

Recall that the time evolution operator is unitary. Thus it preserves the inner product although the state kets change in time.

$$\begin{aligned}\langle\beta,t_0=0;t|\alpha,t_0=0;t\rangle &= \big\langle\beta|U^\dagger(t,0)U(t,0)|\alpha\big\rangle\\ &= \langle\beta|\alpha\rangle\end{aligned} \tag{3.45}$$

However, operators are not affected by these unitary transformations. Since the state kets change in time, $\langle\beta|X|\alpha\rangle$ will also change in time:

$$\langle\beta,t_0=0;t|X|\alpha,t_0=0;t\rangle = \big\langle\beta|U^\dagger(t,0)\,X\,U(t,0)|\alpha\big\rangle \tag{3.46}$$

This suggests that the unitary transformations can be regarded in two different ways: first, the *Schrödinger picture*: the state kets change, but the operators are unchanged.

$$|\alpha\rangle \longrightarrow U(t,0)\,|\alpha\rangle \tag{3.47}$$

In the second one, the *Heisenberg picture*, the operators evolve in time but the state kets remain unchanged:

$$X \longrightarrow U^\dagger(t,0)\,X\,U(t,0) \tag{3.48}$$

This second approach is in fact closer to classical mechanics where we do not introduce state kets but consider observables such as the position of the particle that changes with time.

In the following, a Heisenberg picture observable is

$$A^{(\mathrm{H})}(t) = U^\dagger(t)\, A^{(\mathrm{S})}\, U(t) \tag{3.49}$$

As we can see, Heisenberg and Schrödinger observables coincide at $t = 0$. The corresponding states in the two pictures are given by

$$\begin{aligned} |\alpha, t_0 = 0; t\rangle_H &= |\alpha, t_0 = 0\rangle \\ |\alpha, t_0 = 0; t\rangle_S &= U(t)\, |\alpha, t_0 = 0\rangle \end{aligned} \tag{3.50}$$

The expectation values do not depend on the picture used as can be shown very easily:

$$\begin{aligned} {}_S\left\langle \alpha, t_0 = 0; t | A^{(\mathrm{S})} | \alpha, t_0 = 0; t \right\rangle_S &= \left\langle \alpha, t_0 = 0 | U^\dagger(t)\, A^{(\mathrm{S})}\, U(t) | \alpha, t_0 = 0 \right\rangle \\ &= {}_H\left\langle \alpha, t_0 = 0; t | A^{(\mathrm{H})}(t) | \alpha, t_0 = 0; t \right\rangle_H \end{aligned} \tag{3.51}$$

Heisenberg Equation of Motion

As we have seen, in the Heisenberg picture the operators evolve in time. Now we would like to derive an equation of motion for the operators. To do so, we start with Schrödinger equation of motion for U:

$$i\hbar \frac{\partial}{\partial t} U = HU \quad \Leftrightarrow \quad -i\hbar \frac{\partial}{\partial t} U^\dagger = U^\dagger H \tag{3.52}$$

$$\begin{aligned} \Rightarrow \frac{dA^{(\mathrm{H})}}{dt} &= \frac{\partial U^\dagger}{\partial t} A^{(\mathrm{S})} U + U^\dagger A \frac{\partial U}{\partial t} \\ &= \\ &= \frac{1}{i\hbar}\left[A^{(\mathrm{H})}, U^\dagger H U\right] \end{aligned} \tag{3.53}$$

For a time-independent Hamiltonian $H(t) \equiv H$: $U^\dagger H U = H$.

$$\Rightarrow \frac{dA^{(\mathrm{H})}}{dt} = \frac{1}{i\hbar}\left[A^{(\mathrm{H})}, H\right] \tag{3.54}$$

This is called the *Heisenberg equation of motion*.

Compare the commutator in quantum mechanics with the Poisson bracket!

$$\frac{dA}{dt} = \left\{A, H\right\}_{\mathrm{cl}} \Leftrightarrow \frac{dA^{(\mathrm{H})}}{dt} = \frac{1}{i\hbar}\left[A^{(\mathrm{H})}, H\right] \tag{3.55}$$

$$\frac{\left[\ldots, \ldots\right]_{\mathrm{qm}}}{i\hbar} \Leftrightarrow \left\{\ldots, \ldots\right\} \tag{3.56}$$

So far, we have introduced the Hamiltonian as the generator of time evolution, but we have not explicitly specified how the Hamiltonian looks like in quantum mechanics. Hence we now address how to construct the appropriate Hamiltonian operator. We simply assume that the classical x and p coordinate can be replaced by their corresponding operators in quantum mechanics. Furthermore, we require the Hamiltonian to be Hermitian. Therefore we perform the replacement

$$xp \longrightarrow \frac{1}{2}\left(\hat{x}\hat{p} + \hat{p}\hat{x}\right) \tag{3.57}$$

Furthermore note that which it is rather easy to show

$$\big[x_i, F(\mathbf{p})\big] = +i\hbar\frac{\partial F}{\partial p_i} \tag{3.58}$$

$$\big[p_i, G(\mathbf{x})\big] = -i\hbar\frac{\partial G}{\partial x_i}\ , \tag{3.59}$$

if F and G can be expanded in powers of p and x, respectively.

With that, we can construct the Hamiltonian. For a free particle, we will convert the Hamiltonian from classical mechanics to an operator in quantum mechanics, and then solve the Heisenberg equation of motion.

$$H = \frac{\mathbf{p}^2}{2m} \tag{3.60}$$

$$\frac{d\hat{p}_i}{dt} = \frac{1}{i\hbar}\big[\hat{p}_i, H\big] =$$

$$\frac{d\hat{x}_i}{dt} = \frac{1}{i\hbar}\big[\hat{x}_i, H\big] = \tag{3.61}$$

$$\Rightarrow \hat{x}_i(t) = \hat{x}_i(0) + \frac{\hat{p}_i(0)}{m}t \tag{3.62}$$

So the solution is like a classical trajectory.

For $t = 0$, the fundamental commutation relations hold. But for $t \neq 0$, they do not commute anymore.

$$\big[\hat{x}_i(0), \hat{x}_j(0)\big] = 0 \tag{3.63}$$

$$\big[\hat{x}_i(t), \hat{x}_i(0)\big] = \left[\hat{x}_i(0) + \tfrac{\hat{p}_i(0)}{m}t, \hat{x}_i(0)\right] = -\frac{i\hbar}{m}t \tag{3.64}$$

The uncertainty relation for the measurements of the position (and analogously for the momentum) is

$$\left\langle(\Delta x_i)^2\right\rangle_t \left\langle(\Delta x_i)^2\right\rangle_{t=0} \geq \frac{\hbar^2}{4m^2}\cdot t^2 \tag{3.65}$$

Thus, the position of the particle becomes more and more uncertain with time if it was well-localized at time $t = 0$.

Now, consider additionally a potential $V(\mathbf{x})$.

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

$$\frac{d\hat{p}_i}{dt} = \frac{1}{i\hbar}\big[\hat{p}_i, V(\hat{\mathbf{x}})\big] = -\frac{\partial}{\partial x_i}V(\hat{\mathbf{x}}) \tag{3.66}$$

$$\frac{d\hat{x}_i}{dt} = \frac{\hat{p}_i}{m} \tag{3.67}$$

The last equation still holds, because $[\hat{x}_i, V(\hat{\mathbf{x}})] = 0$. Now we take the second derivative with respect to time by applying the Heisenberg equation of motion to the first derivative:

$$\frac{d^2\hat{x}_i}{dt^2} = \frac{1}{i\hbar}\left[\frac{d\hat{x}_i}{dt}, H\right] = \frac{1}{i\hbar}\left[\frac{\hat{p}_i}{m}, H\right] = \frac{1}{m}\frac{d\hat{p}_i}{dt} = \frac{-1}{m}\frac{\partial}{\partial x_i}V(\hat{\mathbf{x}}) \tag{3.68}$$

Finally we obtain in vectorial form

$$\Rightarrow \ m\frac{d^2\hat{\mathbf{x}}}{dt^2} = -\nabla V(\hat{\mathbf{x}}) \tag{3.69}$$

Equation 3.69 is the quantum mechanical analogue of *Newton's second law of motion* for the Heisenberg picture operators. By taking the expectation value of (3.69) with respect to a Heisenberg state ket, we obtain

$$\Rightarrow \ m\frac{d^2\left\langle\hat{\mathbf{x}}\right\rangle}{dt^2} = \frac{d\left\langle\hat{\mathbf{p}}\right\rangle}{dt} = -\left\langle\nabla V(\hat{\mathbf{x}})\right\rangle \tag{3.70}$$

This is the so-called *Ehrenfest theorem* which is in fact independent of the picture since expectation values are the same in both pictures:

Theorem Ehrenfest

The center of a wave packet (i.e. the expectation value of its position) moves like a classical particle subjected to $V(\mathbf{x})$.

3.1.5 Base Kets and Transition Amplitudes

Let's take a look at the Schrödinger picture where we have an observable A acting on a state ket.

$$A\left|a_i\right\rangle = a_i\left|a_i\right\rangle \tag{3.71}$$

In the Heisenberg picture, the operator changes with time.

$$A^{(\mathrm{H})}(t) = U^\dagger A(0) U \tag{3.72}$$

$$t=0: \qquad A(0)\left|a_i\right\rangle = a_i\left|a_i\right\rangle = A(0)\cdot\underbrace{UU^\dagger}_{=\mathbf{1}}\left|a_i\right\rangle$$

$$\Rightarrow \underbrace{U^\dagger A(0) U}_{=A^{(\mathrm{H})}(t)}\cdot U^\dagger\left|a_i\right\rangle = a_i U^\dagger\left|a_i\right\rangle \tag{3.73}$$

This is an eigenvalue equation for $A^{(\mathrm{H})}$.

$$A^{(\mathrm{H})}(t)\big(U^\dagger\left|a_i\right\rangle\big) = a_i\big(U^\dagger\left|a_i\right\rangle\big) \tag{3.74}$$

Thus, $\left\{U^\dagger\left|a_i\right\rangle\right\}$ form the base kets in the Heisenberg picture which in fact move in time

$$\left|a_i,t\right\rangle_{\mathrm{H}} = U^\dagger\left|a_i\right\rangle \tag{3.75}$$

The base kets satisfy the wrong-sign Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}\left|a_i,t\right\rangle = -H\left|a_i,t\right\rangle \tag{3.76}$$

The Heisenberg picture base kets rotate oppositely compared to Schrödinger picture state kets.

That corresponds to the two points of view in a coordinate transformation.

We expand $A^{(\mathrm{H})}(t)$.

$$\begin{aligned} A^{(\mathrm{H})}(t) &= \sum_i |a_i, t\rangle\, a_i\, \langle a_i, t| = \sum_i U^\dagger |a_i\rangle\, a_i\, \langle a_i| U \\ &= U^\dagger A^{(\mathrm{S})} U \end{aligned} \tag{3.77}$$

This is consistent with our previous discussion.

What are the expansion coefficients in the Schrödinger picture (*left*) and in the Heisenberg picture (*right*)?

$$c_i(t) = \underbrace{\langle a_i|}_{\text{base bra}}\ \underbrace{U|\alpha, t_0 = 0\rangle}_{\text{state ket}} = \underbrace{\langle a_i| U}_{\text{base bra}} \underbrace{|\alpha, t_0 = 0\rangle}_{\text{state ket}} \tag{3.78}$$

The expansion coefficients are the same, as are the *transition amplitudes*.

$$b_{ij}(t) = \underbrace{\langle b_j|}_{\text{base bra}}\ \underbrace{U|\alpha_i\rangle}_{\text{state ket}} = \underbrace{\langle b_j| U}_{\text{base bra}}\ \underbrace{|\alpha_i\rangle}_{\text{state ket}} \tag{3.79}$$

3.1.6 Summary

	Schrödinger Picture	**Heisenberg Picture**
state ket	moving	stationary
observable	stationary	moving
base kets	stationary	moving oppositely

3.2 Schrödinger's Wave Equation

In the first chapter we have introduced the *wave equation approach*. We will show now that it is absolutely equivalent to the approach presented here.

Keep in mind that $\mathbf{x}$ is an *operator* and $\mathbf{x}'$ is the eigenvalue to the state $|\mathbf{x}'\rangle$ with respect to the position operator.

A wave function is defined as the inner product of a state and an operator – $\psi_\alpha(\mathbf{x}', t) \equiv \langle \mathbf{x}'|\alpha, t_0; t\rangle$. The Hamiltonian operator is given by

$$H \equiv H(\mathbf{p}, \mathbf{x}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}) \tag{3.80}$$

$V(\mathbf{x})$ is a *local* operator

$$\langle \mathbf{x}''|V(\mathbf{x})|\mathbf{x}'\rangle = V(\mathbf{x}')\, \delta^3(\mathbf{x}' - \mathbf{x}'')\ , \tag{3.81}$$

where $V(\mathbf{x}')$ is a function of the eigenvalue, not of the operator. The Schrödinger equation for a state ket in the $\mathbf{x}$-representation is given by

$$i\hbar \frac{\partial}{\partial t} \langle \mathbf{x}'|\alpha, t_0; t\rangle = \langle \mathbf{x}'|H|\alpha, t_0; t\rangle \tag{3.82}$$

The kinetic energy and the potential energy can also be described by inner products.

$$\left\langle \mathbf{x}'|\tfrac{\mathbf{p^2}}{2m}|\alpha, t_0; t\right\rangle = -\frac{\hbar^2}{2m}\nabla^2 \langle \mathbf{x}'|\alpha, t_0; t\rangle \tag{3.83}$$

$$\langle \mathbf{x}'|\, V(\mathbf{x})\, |\alpha, t_0; t\rangle = V(\mathbf{x}') \langle \mathbf{x}'|\alpha, t_0; t\rangle \tag{3.84}$$

The result is the *time-dependent Schrödinger equation of motion* in wave mechanics.

$$i\hbar\frac{\partial}{\partial t} \langle \mathbf{x}'|\alpha, t_0; t\rangle = -\frac{\hbar^2}{2m}\nabla^2 \langle \mathbf{x}'|\alpha, t_0; t\rangle + V(\mathbf{x}') \langle \mathbf{x}'|\alpha, t_0; t\rangle$$

$$i\hbar\frac{\partial}{\partial t}\psi_\alpha(\mathbf{x}', t) = -\frac{\hbar^2}{2m}\nabla^2\psi_\alpha(\mathbf{x}', t) + V(\mathbf{x}')\psi_\alpha(\mathbf{x}', t) \tag{3.85}$$

Suppose $|a_i, t_0; t\rangle$ is an energy eigenfunction.

$$\langle \mathbf{x}'|a_i, t_0; t\rangle = \langle \mathbf{x}'|a_i\rangle\, e^{-\frac{i}{\hbar}E_i t}$$

$$-\frac{\hbar^2}{2m}\nabla^2 \langle \mathbf{x}'|a_i\rangle + V(\mathbf{x}') \langle \mathbf{x}'|a_i\rangle = E_i \langle \mathbf{x}'|a_i\rangle \tag{3.86}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{x}') + V(\mathbf{x}')\psi_i(\mathbf{x}') = E_i\psi_i(\mathbf{x}') \tag{3.87}$$

This is the time-independent Schrödinger equation (for stationary states). This shows that the approach using kets and bras is entirely equivalent to the wave mechanics.

Let us consider solutions of (3.87) in general. For bound states we have

$$E < \lim_{\|x\|\to\infty} V(x) \;\Rightarrow\; \lim_{\|x\|\to\infty} \psi(\mathbf{x}') = 0 \tag{3.88}$$

From the theory of partial differential operators, we know that we have a discrete eigenvalue spectrum.

The interpretation of the wave function is not that the particle is 'smeared out', but it is rather a probability distribution. We introduce the *probability density*.

$$\rho(\mathbf{x}, t) := |\psi(\mathbf{x}, t)|^2 = |\langle \mathbf{x}|\alpha, t_0; t\rangle|^2 \tag{3.89}$$

Now we take the time derivative of this expression.

$$\frac{\partial}{\partial t}\rho(\mathbf{x}, t) = \frac{\partial}{\partial t}\big[\psi^*(\mathbf{x}, t)\psi(\mathbf{x}, t)\big] = \dot{\psi}^*\psi + \psi^*\dot{\psi}$$

$$= \frac{1}{-i\hbar}(H\psi^*)\psi + \frac{1}{i\hbar}\psi^*(H\psi) \tag{3.90}$$

Terms with the potential $V(\mathbf{x})$ vanish:

$$\Rightarrow\; \frac{\partial}{\partial t}\rho(\mathbf{x}, t) = \frac{\hbar}{2im}\big[(\nabla^2\psi^*)\psi - \psi^*(\nabla^2\psi)\big] \tag{3.91}$$

Now we define the *probability flux*.

$$\mathbf{j}(\mathbf{x}, t) = \frac{\hbar}{2im}\big[\psi^*\nabla\psi^* - (\nabla\psi)\psi\big] = \frac{\hbar}{m}\Im\big[\psi^*(\nabla\psi)\big] \tag{3.92}$$

$$\Rightarrow \frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 \tag{3.93}$$

This is a *continuity equation*!

We rewrite the wave equation in terms of the probability density.

$$\psi(\mathbf{x},t) = \sqrt{\rho(\mathbf{x},t)} \cdot e^{\frac{i}{\hbar}S(\mathbf{x},t)} \; , \tag{3.94}$$

where ρ and S are real functions, $\rho > 0$. ρ has already been defined. In order to derive the meaning of S, we note that

$$\begin{aligned} \psi^* \nabla \psi &= \sqrt{\rho} e^{-\frac{i}{\hbar}S} \cdot \big(\nabla \sqrt{\rho} + \frac{i}{\hbar}\sqrt{\rho}\nabla S\big) e^{\frac{i}{\hbar}S} \\ &= \sqrt{\rho}\,\nabla\big(\sqrt{\rho}\big) + \frac{i}{\hbar}\rho \nabla S \end{aligned} \tag{3.95}$$

Now the the probability flux can be simplified to

$$\mathbf{j} = \frac{\rho \nabla S}{m} \tag{3.96}$$

The spatial variation of the phase characterizes the probability flux. For instance, for a plane wave (a momentum eigenfunction) we have

$$\psi(\mathbf{x},t) \propto e^{\frac{i}{\hbar}(\mathbf{px}-Et)} \tag{3.97}$$

$$\Rightarrow \;\; \nabla S = \nabla(\mathbf{px}) = \mathbf{p} \tag{3.98}$$

Now we will take the classical limit: insert 3.97 into the time-dependent Schrödinger equation.

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}\psi(\mathbf{x},t) &= -\frac{\hbar^2}{2m}\nabla^2 \psi(\mathbf{x},t) + V(\mathbf{x})\psi(\mathbf{x},t) \\ \Rightarrow \;\; i\hbar \left(\frac{\partial}{\partial t}\sqrt{\rho} + \frac{i}{\hbar}\sqrt{\rho}\frac{\partial}{\partial t}S \right) e^{\frac{i}{\hbar}S} &= -\frac{\hbar^2}{2m}\nabla \cdot \big[(\nabla\sqrt{\rho} + \frac{i}{\hbar}\sqrt{\rho}\,\nabla S) e^{\frac{i}{\hbar}S} \big] + \sqrt{\rho} V e^{\frac{i}{\hbar}S} \\ &= \end{aligned} \tag{3.99}$$

Taking the classical limit (usually) means $\hbar \to 0$. In this case, we neglect all terms containing $\hbar$. The following terms survive:

$$\frac{1}{2m}\left|\nabla S(\mathbf{x},t)\right|^2 + V(\mathbf{x}) + \frac{\partial S}{\partial t} = 0 \tag{3.100}$$

This is the analogue to the *Hamilton-Jacobi equation* in classical mechanics! Here, $S(\mathbf{x},t)$ is Hamilton's principle function (i.e. action). A stationary state has the time dependence $e^{-\frac{i}{\hbar}Et}$, i.e. S is separable. It follows that

$$S(\mathbf{x},t) = W(\mathbf{x}) - Et \; , \tag{3.101}$$

where W is Hamilton's characteristic function. As time goes on, a surface of constant S advances.

3.3 Harmonic Oscillator

The *harmonic oscillator* is very important from both a pedagogical as well as a practical point of view. One important point is that it can be solved analytically. Furthermore, every system behaves like a harmonic oscillator for small displacements or vibrations about the energy minimum positions.

The Hamiltonian of a harmonic oscillator in one dimension is given by

$$H \equiv H(p,x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{3.102}$$

First, some terminology. We define the *annihilation operator* a as

$$a \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{i}{m\omega}p\right) \tag{3.103}$$

and the *creation operator* $a^\dagger$ as

$$a^\dagger \equiv \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{i}{m\omega}p\right) \tag{3.104}$$

The operators a and $a^\dagger$ are both *non-Hermitian*, which is clear since the are the Hermitian adjoint of each other and different. Their specific names will soon become clear. These operators have a simple commutation relation:

$$\begin{aligned}[] [a, a^\dagger] &= \frac{m\omega}{2\hbar}\left[\left(x + \frac{i}{m\omega}p\right), \left(x - \frac{i}{m\omega}p\right)\right] \\ &= \frac{1}{2\hbar}\big(-i[x,p] + i[p,x]\big) = \frac{1}{2\hbar}(\hbar + \hbar) = \mathbf{1} \end{aligned} \tag{3.105}$$

We next define the *number operator* N

$$N \equiv a^\dagger a \tag{3.106}$$

$$\begin{aligned} N = a^\dagger a &= \\ \\ &= \frac{H}{\hbar\omega} - \frac{1}{2} \end{aligned} \tag{3.107}$$

Thus we can re-write the Hamiltonian of the harmonic oscillator as

$$H = \hbar\omega\big(N + \frac{1}{2}\big) \tag{3.108}$$

N is a Hermition operator. We will call its eigenkets $|n\rangle$ which satisfy the eigenvalue equation $N|n\rangle = n|n\rangle$. Because of (3.108), the eigenkets of the number operator are also eigenkets of the Hamiltonian.

$$H|n\rangle = \hbar\omega\big(n + \frac{1}{2}\big)|n\rangle \tag{3.109}$$

The energy eigenvalues are $E_n = \hbar\omega\big(n + \frac{1}{2}\big)$. We will show below that n must be a nonnegative integer.

Properties of a. Let us start with the commutator of N and a and $a^\dagger$ respectively.

$$\left[N,a\right] = \left[a^\dagger a, a\right] = a^\dagger \underbrace{[a,a]}_{=0} + \underbrace{[a^\dagger, a]}_{=-\mathbf{1}} a = -a \tag{3.110}$$

Equivalently,

$$\left[N,a^\dagger\right] = \ldots = +a^\dagger \tag{3.111}$$

In order to find out why $a^\dagger$ is called creation operator we let $Na^\dagger$ act on eigenkets of N.

$$Na^\dagger \left|n\right\rangle = \big(\underbrace{[N,a^\dagger]}_{=a^\dagger} + a^\dagger N\big) \left|n\right\rangle = (n+1) a^\dagger \left|n\right\rangle \tag{3.112}$$

Thus, $a^\dagger \left|n\right\rangle$ is an eigenket of N with eigenvalue $(n+1)$. This means that acting with $a^\dagger$ on a state $\left|n\right\rangle$ leads to an higher excited state $\left|n+1\right\rangle$, i.e., we have *created* a state with an additional quantum unit of energy $\hbar\omega$.

Likewise, $a\left|n\right\rangle$ is an eigenket of N with eigenvalue $(n-1)$.

$$Na\left|n\right\rangle = \ldots = (n-1)a\left|n\right\rangle, \tag{3.113}$$

which corresponds to a lower excited state, i.e., a quantum unit of energy $\hbar\omega$ has been annihilated. This is the reason why $a^\dagger$ and a are called creation and annihilation operators, respectively.

Eq. (3.113) implies that there is a constant c with

$$a\left|n\right\rangle \overset{!}{=} c\left|n-1\right\rangle \tag{3.114}$$

By convention, we require that c is real and positive.

$$\begin{aligned} \left\langle n\middle|a^\dagger a\middle|n\right\rangle = |c|^2 = \langle n|N|n\rangle = n \\ \Rightarrow\ c = \sqrt{n} \end{aligned} \tag{3.115}$$

Analogously for $a^\dagger$. Thus, the result can be summarized as follows.

$$a\left|n\right\rangle = \sqrt{n}\left|n-1\right\rangle \tag{3.116}$$

$$a^\dagger\left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle \tag{3.117}$$

Successive application of a gives us

$$\begin{aligned} a^2\left|n\right\rangle &= \sqrt{n(n-1)}\left|n-2\right\rangle \\ a^3\left|n\right\rangle &= \sqrt{n(n-1)(n-2)}\left|n-3\right\rangle \end{aligned}$$

If n is an integer, the squence terminates at a certain point.

Assume that n is not an integer.

$$\langle n|N|n\rangle = n = \big(\langle n|\, a^\dagger\big)\big(a\left|n\right\rangle\big) \geq 0 \tag{3.118}$$

Thus, n must be positive integer so that the sequence terminates. Since the eigenvalues of N are integers, it is called number operator. The smallest value of n is 0.

The smallest possible value of n, zero, corresponds to the ground state $|0\rangle$[1], which is the lowest possible eigenstate with energy $E_0 = \frac{1}{2}\hbar\omega$. A very important consequence is that the ground energy is *not* 0! The non-zero ground state energy, the so-called zero-point energy, is a consequence of the uncertainty principle.

Excited states can be obtained by applying the creation operator to the ground state.

$$|1\rangle = a^\dagger |0\rangle \qquad |2\rangle = \frac{a^\dagger}{\sqrt{2}}|1\rangle = \frac{\left(a^\dagger\right)^2}{\sqrt{2!}}|0\rangle \ , \qquad |n\rangle = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}}|0\rangle \tag{3.119}$$

These are the eigenkets of H and N with energy eigenvalues

$$E_n = \hbar\omega\big(n + \frac{1}{2}\big) \tag{3.120}$$

The matrix elements with respect to a and $a^\dagger$ are

$$\begin{aligned} \langle m|a|n\rangle &= \sqrt{n}\delta_{m,n-1} \\ \left\langle m|a^\dagger|n\right\rangle &= \sqrt{n+1}\delta_{m,n+1} \end{aligned} \tag{3.121}$$

Now the matrix elements of the x and p operators follow:

$$\langle m|x|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\big(\sqrt{n}\delta_{m,n-1} + \sqrt{n+1}\delta_{m,n+1}\big) \tag{3.122}$$

$$\langle m|p|n\rangle = i\sqrt{\frac{m\hbar\omega}{2}}\big(-\sqrt{n}\delta_{m,n-1} + \sqrt{n+1}\delta_{m,n+1}\big) \tag{3.123}$$

where

$$x = \sqrt{\frac{\hbar}{2m\omega}}\big(a + a^\dagger\big) \tag{3.124}$$

$$p = i\sqrt{\frac{m\hbar\omega}{2}}\big(-a + a^\dagger\big) \tag{3.125}$$

How do the eigenfunctions look like? We know that if we apply the annihilation operator to the ground state, we get the zero ket, $a|0\rangle = 0$.

$$\langle x'|a|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\left\langle x'\middle|x + i\frac{p}{m\omega}\middle|0\right\rangle = 0 \tag{3.126}$$

Introducing $x_0 := \sqrt{\frac{\hbar}{m\omega}}$ and using $\langle x'|p|\alpha\rangle = -i\hbar\frac{d}{dx}\langle x'|\alpha\rangle$ leads to a differential equation:

$$\left(x' + x_0^2\frac{d}{dx'}\right)\langle x'|0\rangle = 0 \tag{3.127}$$

The (normalized) solution is a Gaussian function.

$$\langle x'|0\rangle =: \Psi_0(x') = \frac{1}{\pi^{\frac{1}{4}}\sqrt{x_0}}e^{-\frac{1}{2}\left(\frac{x'}{x_0}\right)^2} \tag{3.128}$$

---

[1] it is *not* the zero-ket, but rather the ground state by convention

The excited states are also Gaussian functions multiplied with polynomials of x'.

$$\langle x'|1\rangle = \langle x'|a^\dagger|0\rangle = \frac{1}{\sqrt{2}x_0}\left(x' - x_0^2\frac{d}{dx'}\right)\langle x'|0\rangle$$
$$\langle x'|2\rangle = \frac{1}{\sqrt{2}}\langle x'|(a^\dagger)^2|0\rangle = \frac{1}{\sqrt{2!}}\left(\frac{1}{\sqrt{2}x_0}\right)^2\left(x' - x_0^2\frac{d}{dx'}\right)^2\langle x'|0\rangle \tag{3.129}$$

In general, we have

$$\langle x'|n\rangle = \frac{1}{\pi^{\frac{1}{4}}\sqrt{2^n n!}x_0^{n+\frac{1}{2}}}\left(x' - x_0^2\frac{d}{dx'}\right)^n e^{-\frac{1}{2}\left(\frac{x'}{x_0}\right)^2} \tag{3.130}$$

These solutions are still normalized.

Expectation Values of x^2 and p^2 in the ground state.

$$x^2 = \frac{\hbar}{2m\omega}\left(a^2 + a^{\dagger^2} + a^\dagger a + aa^\dagger\right) \tag{3.131}$$
$$\langle x^2\rangle_0 = \langle 0|x^2|0\rangle = \frac{\hbar}{2m\omega} = \frac{x_0^2}{2} \tag{3.132}$$

Just the last term in the bracket survives, due to the orthonormality.

$$p^2 = \frac{m\hbar\omega}{2}\left(-a^2 - a^{\dagger^2} + a^\dagger a + aa^\dagger\right) \tag{3.133}$$
$$\langle p^2\rangle_0 = \frac{m\hbar\omega}{2} \tag{3.134}$$

Thus we obtain the expectation values of the kinetic and the potential energy.

$$\left\langle \frac{p^2}{2m}\right\rangle_0 = \frac{\hbar\omega}{4} = \frac{1}{2}\langle H\rangle$$
$$\left\langle \frac{m\omega^2x^2}{2}\right\rangle_0 = \frac{\hbar\omega}{4} = \frac{1}{2}\langle H\rangle \tag{3.135}$$

This complies with the result in classical mechanics (*virial theorem* or *equipartition theorem*).

Furthermore, because of symmetry we have

$$\langle x\rangle = 0 = \langle p\rangle \tag{3.136}$$

Plugged into the uncertainty relation, we see that the ground state is a minimum uncertainty product.

$$\langle(\Delta x)^2\rangle_0\langle(\Delta p)^2\rangle_0 = \langle x^2\rangle_0\langle p^2\rangle_0 = \frac{\hbar}{2m\omega}\cdot\frac{m\hbar\omega}{2} = \frac{\hbar^2}{4} \tag{3.137}$$

Analogously, excited states fulfill

$$\langle(\Delta x)^2\rangle_n\langle(\Delta p)^2\rangle_n = \left(n + \frac{1}{2}\right)^2\hbar^2 \geq \frac{\hbar^2}{4} \tag{3.138}$$

Note: in a classical harmonic oscillator, we encounter the particle most probably at the turning points. In quantum mechanics, we will most probably find it in the middle!

3.3.1 Time Development of the Harmonic Oscillator

Heisenberg Picture

According to the Heisenberg equation of motion, we get

$$\frac{dp}{dt} = \frac{1}{i\hbar}\big[p, H\big] = -\frac{\partial H}{\partial x} = -m\omega^2 x \tag{3.139}$$

$$\frac{dx}{dt} = \frac{1}{i\hbar}\big[x, H\big] = \frac{\partial H}{\partial p} = \frac{p}{m} \tag{3.140}$$

Instead of the usual way in classical mechanics of solving that problem, we take a different approach here. By writing the equations of motion in terms of a and $a^\dagger$, we obtain two uncoupled differential equations:

$$\Longrightarrow \frac{d}{dt}a^\dagger = i\omega a^\dagger \qquad \frac{d}{dt}a = -i\omega a \tag{3.141}$$

The solutions are simply exponentials.

$$a(t) = a_0 e^{-i\omega t} \qquad a^\dagger(t) = a_0^\dagger e^{i\omega t} \tag{3.142}$$

Note that $N = a^\dagger a$ and $H = \hbar\omega(N + \frac{1}{2})$ are time-independent in the Heisenberg picture, as they should be, since they commute with the Hamiltonian so that they have to constant according to the Heisenberg equation of motion.

We rewrite the last formula and express a and $a^\dagger$ in terms of x and p.

$$x(t) + \frac{i}{m\omega}p(t) = x(0)e^{-i\omega t} + \frac{i}{m\omega}p(0)e^{-i\omega t} \tag{3.143}$$

$$x(t) - \frac{i}{m\omega}p(t) = x(0)e^{i\omega t} - \frac{i}{m\omega}p(0)e^{+i\omega t} \tag{3.144}$$

We equate the Hermitian and the anti-Hermitian part of this equation – they have to be equal (corresponds to real and imaginary part of a number, but we are dealing with operators here)!

$$x(t) = x(0)\cos\omega t + \frac{p(0)}{m\omega}\sin\omega t \tag{3.145}$$

$$p(t) = -m\omega x(0)\sin\omega t + p(0)\cos\omega t \tag{3.146}$$

These equations for the operators look like the classical equations of motion. But although $x(t)$ and $p(t)$ oscillate, their expectation values still vanish, $\langle x(t)\rangle_n = 0 = \langle p(t)\rangle_n$, with respect to any eigenstate $|n\rangle$ because both $x(0)$ and $p(0)$ change n by ± 1, and $|n\rangle$ and $|n \pm 1\rangle$ are orthogonal.

In order to obtain oscillations, we must in fact consider a superposition of two eigenstates, say $|0\rangle$ and $|1\rangle$:

$$|\alpha\rangle = c_0\,|0\rangle + c_1\,|1\rangle \tag{3.147}$$

$$\Rightarrow\ \langle x(t)\rangle_\alpha = \langle x(t)\rangle_\alpha(t) \tag{3.148}$$

This motivates the question whether or not there are states that imitate classical oscillators. It turns out that the so-called *coherent states* behave like a classical oscillator. They can be expressed as eigenstates of the non-Hermitian annihilation operator a; they obey the eigenvalue equation, $a\left|\lambda\right\rangle = \lambda\left|\lambda\right\rangle$, $\lambda \in \mathbb{C}$.

Properties.

1. $\left|\lambda\right\rangle = \sum\limits_{n=0}^{\infty} f(n)\left|n\right\rangle$ with

$$|f(n)|^2 = \tfrac{\bar{n}^n}{n!}e^{-\bar{n}} \geq 0 \tag{3.149}$$

   where $\bar{n} = \langle N\rangle_\lambda = \sum\limits_{n\in\mathbb{N}_0} |f(n)|^2 n$ which is the mean value of the occupation number. The distribution of $|f(n)|^2$ with respect to n corresponds to a *Poisson distribution*.

2. It can be obtained by translating the oscillator ground state.

3. It satisfies the minimum uncertainty product relation at all times[2].

3.4 Harmonic Oscillator using Wave Mechanics

We will revisit the harmonic oscillator. The eigenstates of the harmonic oscillators are

$$\psi_n(\mathbf{x}) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{2^n n!}}\frac{1}{x_0^{n+\frac{1}{2}}}\left(x - x_0^2\frac{d}{dx}\right)^n e^{-\frac{1}{2}\left(\frac{x}{x_0}\right)^2} \tag{3.150}$$

We introduce the dimensionless variable $y := \frac{x}{x_0}$.

$$\psi_n(y) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{2^n n!}}\left(y - \frac{d}{dy}\right)^n e^{-\frac{y^2}{2}} \tag{3.151}$$

Consider the following operator identity.

$$\begin{aligned} e^{-\frac{y^2}{2}}\left(y - \frac{d}{dy}\right)e^{\frac{y^2}{2}} &= -\frac{d}{dy} \\ e^{-\frac{y^2}{2}}\left(y - \frac{d}{dy}\right)^n e^{\frac{y^2}{2}} &= (-1)^n\frac{d^n}{dy^n} \end{aligned} \tag{3.152}$$

Now the equations simplify to

$$\begin{aligned} \left(y - \frac{d}{dy}\right)^n e^{-\frac{y^2}{2}} &= e^{-\frac{y^2}{2}}e^{y^2}e^{-\frac{y^2}{2}}\left(y - \frac{d}{dy}\right)^n e^{\frac{y^2}{2}}e^{-y^2} \\ &= e^{-\frac{y^2}{2}}\underbrace{e^{y^2}(-1)^n\frac{d^n}{dy^n}e^{-y^2}}_{=:H_n(y)} = e^{-\frac{y^2}{2}}H_n(y) \end{aligned} \tag{3.153}$$

---

[2] the proof is left as an excercise

$$\Rightarrow \psi_n(y) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{2^n n!}} e^{-\frac{y^2}{2}} H_n(y) \tag{3.154}$$

The functions H_n are called *Hermite polynomials*. They satisfy the following equations

$$\int_{\mathbb{R}} e^{-y^2} H_n(y) H_m(y)\, dy = \sqrt{\pi} 2^n n! \delta_{mn} \tag{3.155}$$

$$\left[\frac{d^2}{dy^2} - 2y\frac{d}{dy} + 2n\right] H_n(y) = 0 \tag{3.156}$$

The first Hermite polynomials are:

$$H_0(y) = 1, \quad H_1(y) = 2y, \quad H_2(y) = 4y^2 - 2, \quad H_3(y) = 8y^3 - 12y, \quad \dots \tag{3.157}$$

3.4.1 Symmetry of the Wave Function

We define the parity operator π.

$$\pi f(x) \equiv f(-x) \tag{3.158}$$

π is an involution, i. e. $\pi^2 = \mathrm{id}$. Even and odd functions are eigenfunctions of π with eigenvalues ± 1.

$$f(-x) = \begin{cases} +f(x) & \text{even parity} \\ -f(x) & \text{odd parity} \end{cases} \tag{3.159}$$

Let $V(x)$ be an even potential, $\pi V(x) = V(x)$. The kinetic energy operator is always even, since it only contains second derivatives. Hence the Hamiltonian with an even potential is even. Consequently, $[H, \pi] = 0$.

Theorem

Suppose H commutes with the parity operator, i. e. $[H, \pi] = 0$, and $|n\rangle$ is a nondegenerate eigenket of H with eigenvalue E_n

$$H|n\rangle = E_n|n\rangle \tag{3.160}$$

Then $|n\rangle$ is also a parity eigenket.

Proof

The two operators commute and E_n is a nondegenerate eigenvalue. We have already proved that they have a mutual set of eigenkets.

Let us specifically consider the parity operator.

1. We first show that $\frac{1}{2}(1 \pm \pi)\,|n\rangle$ is a parity eigenket:
$$\pi\big(\frac{1}{2}(1 \pm \pi)\big)\,|n\rangle = \pm\frac{1}{2}(1 \pm \pi)\,|n\rangle \tag{3.161}$$
2. Furthermore, $\frac{1}{2}(1 \pm \pi)\,|n\rangle$ is an energy eigenket with eigenvalue E_n because, since H and π commute, we get
$$H\big(\frac{1}{2}(1 \pm \pi)\big)\,|n\rangle = \frac{1}{2}(E_n \pm \pi E_n)\,|n\rangle = E_n \cdot \frac{1}{2}(1 \pm \pi)\,|n\rangle \tag{3.162}$$
   Since E_n is nondegenerate, $|n\rangle$ and $\frac{1}{2}(1 \pm \pi)\,|n\rangle$ can only differ by a multiplicative constant $\Rightarrow$ $|n\rangle$ is a parity eigenket.

□

An example is the harmonic oscillator: its parity is $(-1)^n$.

3.5 Ensembles and the density operator

So far we discussed measurements where we have performed the same experiment on identically prepared particles which can all be specified by the same state ket $|\alpha\rangle$. We denote such a system as a *pure ensemble* in the following. How do we describe a beam where for example 50% of the particles are in a particular state $|a_1\rangle$, the other 50% in a state $|a_2\rangle$? Is the correct description $|\alpha\rangle = (|a_1\rangle + |a_2\rangle)/\sqrt{2}$?

To answer, let us consider a spin state
$$|\alpha\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}\,. \tag{3.163}$$

As we have already shown, this state does certainly not correspond to a 50-50 mixture of spin up and spin down states but rather to a spin eigenstate of the S_x operator with eigenvalue $\hbar/2$. In the linear superposition (3.163), there is crucial phase relation between $|\uparrow\rangle$ and $|\downarrow\rangle$ states determining the spin orientation. In contrast, in a mixture there is in general not specific phase relation between the different parts; therefore, we will call it more precisely an *incoherent mixture*.

In order to deal with such a situation, we introduce the *fractional populations* or *probability weights* w_i that give the population of a particular state $|a_i\rangle$. For a 50-50 mixture, we have
$$w_1 \;=\; 0.5\,, \quad w_2 \;=\; 0.5\,. \tag{3.164}$$

The w_i are real numbers. These fractional populations satisfy the typical normalization condition of probability weights, namely
$$\sum_i w_i \;=\; 1\,. \tag{3.165}$$

If only one w_i is non-zero, i.e., our mixture only contains one state, then we call it a *pure ensemble*. However, if more than one w_i are non-zero, then we have a *mixed* *ensemble*. If we measure some observable B on a mixed ensemble, the average value of an observable B is given by the so-called *ensemble average*

$$[B] \equiv \sum_i w_i \ \langle a_i|B|a_i\rangle = \sum_i \sum_j w_i |\langle b_j|a_i\rangle|^2 b_j \ , \tag{3.166}$$

where $|b_j\rangle$ is an eigenket of B.

Using a general basis $\{c_j\}$, the ensemble average can be written as

$$\begin{aligned}[B] &= \sum_i w_i \sum_j \sum_k \langle a_i|c_j\rangle \langle c_j|B|c_k\rangle \langle c_k|a_i\rangle \\ &= \sum_j \sum_k \left(\sum_i w_i \langle c_k|a_i\rangle \langle a_i|c_j\rangle \right) \langle c_j|B|c_k\rangle \ . \end{aligned} \tag{3.167}$$

The expression in the brackets in eq. (3.167) does not depend on the specific observable B but only on properties of the ensemble. This information can be expressed in form of the *density operator* ρ

$$\rho = \sum_i w_i \ |a_i\rangle \langle a_i| \ . \tag{3.168}$$

The elements of the corresponding *density matrix* are given by

$$\langle c_k|\rho|c_j\rangle = \sum_i w_i \langle c_k|a_i\rangle \langle a_i|c_j\rangle \ . \tag{3.169}$$

Comparing (3.169) with (3.167), we see that the ensemble average can be written in terms of the density operator as

$$\begin{aligned}[B] &= \sum_j \sum_k \langle c_k|\rho|c_j\rangle \langle c_j|B|c_k\rangle \ = \sum_k \langle c_k|\rho B|c_k\rangle \\ &= \mathrm{tr}(\rho B) \end{aligned} \tag{3.170}$$

Since the trace is independent of the representation, any convenient basis can be used to evaluate $\mathrm{tr}(\rho B)$.

There are important properties of the density operator. First of all, it is Hermitian, which follows from its definition in (3.168). Second, it satisfies the normalization condition

$$\begin{aligned}\mathrm{tr}(\rho) &= \sum_i \sum_j w_i \langle c_j|a_i\rangle \langle a_i|c_j\rangle \ = \\ &= \sum_i w_i \langle a_i|a_i\rangle \ = \sum_i w_i \ = \ 1 \ . \end{aligned} \tag{3.171}$$

A pure ensemble is characterized by $w_i = 1$ for some state $|a_i\rangle$ and $w_j = 0$ for all other kets $|a_j\rangle$. Hence the density operator for a pure ensemble can ber written just as

$$\rho = \ |a_i\rangle \langle a_i| \ . \tag{3.172}$$

Thus the density operator for a pure ensemble just corresponds to a projection operator. Operating twice with the same projection operator on a state does not change anything compared to performing the projection just once. Consequently, a projection operator and thus also the density operator is idempotent which means that

$$\rho^2 = \rho \,, \tag{3.173}$$

and thus also $tr(\rho^2) = 1$. Therefore one eigenvalue of the density operator for a pure ensemble is one, the rest is zero. When diagonalized, the density matrix for a pure ensemble has one eigenvalue one on the diagonal, all other matrix elements are zero.

Let us look at some examples of spin $\frac{1}{2}$ systems. If the system is completely polarized in the $S_z \downarrow$ state, the density matrix is given by

$$\begin{aligned} \rho = |\downarrow\rangle\langle\downarrow| &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{3.174}$$

A completely polarized beam in the $S_x \uparrow$ state:

$$\begin{aligned} \rho = |S_x \uparrow\rangle\langle S_x \uparrow| &= \left(\frac{1}{\sqrt{2}}\right) (|\uparrow\rangle + |\downarrow\rangle) \left(\frac{1}{\sqrt{2}}\right) (\langle\uparrow| + \langle\downarrow|) \\ &= \end{aligned} \tag{3.175}$$

An unpolarized beam can be regarded as a fifty-fifty mixture of a spin-up ensemble and a spin-down ensemble:

$$\begin{aligned} \rho &= \frac{1}{2} |\uparrow\rangle\langle\uparrow| + \frac{1}{2} |\downarrow\rangle\langle\downarrow| \\ &= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \end{aligned} \tag{3.176}$$

It is an easy exercise to show that the same matrix results if we consider the unpolarized beam to be made up of either $S_x \uparrow$ and $S_x \downarrow$ or $S_y \uparrow$ and $S_y \downarrow$ states, as it should be. Thus ρ for an unpolarized beam is just the identity operator divided by 2 which is the dimensionality of the problem. Therefore,

$$\mathrm{tr}(\rho S_x) = \mathrm{tr}(\rho S_y) = \mathrm{tr}(\rho S_z) = 0 \,, \tag{3.177}$$

since the S_i are traceless. And hence the ensemble average of $\mathbf{S}$ vanishes for a random ensemble,

$$[\mathbf{S}] = 0 \,. \tag{3.178}$$

which means that there is no preferred direction for the spin. Finally, we will consider an example for a partially polarized beam, for example a 75-25 mixture of a pure $S_z \uparrow$ and a pure $S_x \uparrow$ ensemble, i.e. $w(S_z \uparrow) = 0.75$ and $w(S_x \uparrow) = 0.25$. The corresponding density matrix can be written as

$$\rho = \tag{3.179}$$

Using this density matrix, the ensemble averages for the three spin components are given by

$$[S_x] = \frac{\hbar}{8} \,, \quad [S_y] = 0 \,, \quad [S_z] = \frac{3\hbar}{8} \,. \tag{3.180}$$

It is of course desirable to be able to specify the degree of randomness of a given ensemble. To this end, we introduce the quantity σ by

$$\sigma \; = \; -\mathrm{tr}(\rho \ln \rho) \; . \tag{3.181}$$

It might not be immediately obvious how to obtain the logarithm of a matrix. Most easily it is done in the basis in which ρ is diagonal because then

$$\sigma \; = \; -\sum_j \rho_{jj}^{\mathrm{diag}} \ln \rho_{jj}^{\mathrm{diag}} \; . \tag{3.182}$$

For a pure ensemble such as (3.174), $\sigma = 0$ because of either $\rho_{jj}^{\mathrm{diag}} = 0$ or $\ln \rho_{jj}^{\mathrm{diag}} = 0$ for each term in the sum (3.182). For a completely random ensemble such as (3.176), the density matrix consists in any representation of diagonal elements that are all $\rho_{jj} = 1/N$, where N is the dimension of the system, whereas all non-diagonal elements vanish. Hence σ is given by

$$\sigma \; = \; -\sum_{j=1}^{N} \frac{1}{N} \ln \left(\frac{1}{N} \right) \; = \; \ln N \; . \tag{3.183}$$

It can in fact be shown that $\ln N$ is the maximum value for σ under the normalization condition (3.171). It is apparent that σ is a quantitative measure of disorder. For an ordered (pure) system, σ vanishes while for a disordered (random) system it is maximal. From thermodynamics we learn that the *entropy* is a measure for disorder. In fact, the quantity σ is directly related to the entropy per particle s via

$$s \; = \; k_{\mathrm{B}} \sigma \; , \tag{3.184}$$

where k_{B} is the Boltzmann constant. Thus eq. (3.184) can be regarded as the definition of the entropy in quantum statistical mechanics.

Finally we want to address the questions how quantum ensembles evolve in time. Let us assume that at the specific time $t = 0$ the density operator is given by

$$\rho(t=0) = \sum_i w_i \; |a_i\rangle \langle a_i| \; . \tag{3.185}$$

If the system is left undisturbed, then the fractional populations will not be changed. Hence it is only the state kets that are evolving in time. Thus we have

$$\begin{aligned} \frac{\partial}{\partial t}\rho(t) &= \sum_i w_i \; \left[\left(\frac{\partial}{\partial t} |a_i, t\rangle \right) \langle a_i, t| + |a_i, t\rangle \left(\frac{\partial}{\partial t} \langle a_i, t| \right) \right] \\ &= \frac{1}{i\hbar} \sum_i w_i \; \left[H |a_i, t\rangle \langle a_i, t| - |a_i, t\rangle \langle a_i, t| H \right] \; , \end{aligned} \tag{3.186}$$

where we have used the fact that the state kets evolve according to the Schrödinger equation. Consequently, we have

$$i\hbar \frac{\partial}{\partial t} \rho(t) = \; -[\rho, H] \; . \tag{3.187}$$

This so-called *Liouville-von Neumann* equation looks like the Heisenberg equation of motion except for the opposite sign. This should be not disturbing since ρ is not a dynamical observable in the Heisenberg picture. It is rather constructed from state kets and bras in the Schrödinger picture which evolve in time according to the Schrödinger equation.

4 Angular Momentum

In this chapter, we will discuss the quantum description of the angular momentum. We will see that are angular momentum operators that have no analogue in classical mechanics, namely the spin operators. We will present, how the energy levels of the hydrogen atoms can be determined. The addition of spin and orbital angular momentum presented at the end of the chapter is essentiell for the characterization of electronic orbitals in chemistry.

4.1 Rotations and Angular Momentum

In analogy to the treatment of translations, we will here use the fact that angular momentum is the generator of rotations.

Rotations are described by *orthogonal operators*, $R \in \mathcal{O}$. This is a group of real operators (that can be represented by real matrices). For rotation around the e_3 axis in $\mathbb{R}^3$ we get

$$R_3(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.1}$$

Note. Here we regard *active rotations*, i. e. the coordinate axes are kept fixed and the physical system is rotated.

Let us specifically consider an infinitesimal rotation by $\delta\varphi$ where terms of higher order are neglected.

$$R_3(\delta\varphi) = \begin{pmatrix} 1 - {\delta\varphi^2}/{2} & -\delta\varphi & 0 \\ \delta\varphi & 1 - {\delta\varphi^2}/{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

Likewise, by cyclic permutation, we get the appropriate matrices for rotations around e_1 and e_2.

$$R_1(\delta\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\delta\varphi^2}{2} & -\delta\varphi \\ 0 & \delta\varphi & 1 - \frac{\delta\varphi^2}{2} \end{pmatrix} \tag{4.3}$$

$$R_2(\delta\varphi) = \begin{pmatrix} 1 - \frac{\delta\varphi^2}{2} & 0 & \delta\varphi \\ 0 & 1 & 0 \\ -\delta\varphi & 0 & 1 - \frac{\delta\varphi^2}{2} \end{pmatrix} \tag{4.4}$$

This may be verified by the cyclic permutation $x \to y, y \to z, z \to x$.

Successive *infinitesimal* rotations commute to first order but not to second order:

$$R_1(\delta\varphi)R_2(\delta\varphi) - R_2(\delta\varphi)R_1(\delta\varphi) = \begin{pmatrix} 0 & -\delta\varphi^2 & 0 \\ \delta\varphi^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.5}$$

$$= R_3(\delta\varphi^2) - \mathbf{1}\ , \tag{4.6}$$

where terms of order higher than $\delta\varphi^2$ have been ignored.

Rotations in Quantum Mechanics

Given a rotation operator R, we associate an operator $D(R)$ in the appropriate ket space such that $|\alpha\rangle_R = D(R)\,|\alpha\rangle$.

Note. Let $R \in \mathcal{O}(3)$; $D(R)$ acts on state vectors in ket space.

The matrix representation of $D(R)$ depends on the dimensionality of the particular ket space. We will now construct $D(R)$.

Construction of $D(R)$

Recall the properties of infinitesimal operators:

$$U_\varepsilon = 1 - iG\varepsilon \tag{4.7}$$

For infinitesimal translations, $G = p/\hbar$, $\varepsilon \to dx$. For time evolution $G = H/\hbar$, $\varepsilon \to dt$.

In classical mechanics, angular momentum is the generator of rotation. In quantum mechanics, we define the *angular momentum operator* J_k for infinitesimal rotation around the k-th axis by

$$G \to \frac{J_k}{\hbar} \qquad \varepsilon \to d\varphi \tag{4.8}$$

Up to now we still do not know what J_k actually *is*. In general

$$D(\hat{n}, d\varphi) = 1 - i\frac{\mathbf{J}\cdot\hat{n}}{\hbar}d\varphi \tag{4.9}$$

It is very important to note that the angular momentum is *not defined as* $\mathbf{x} \times \mathbf{p}$. For instance, spin angular momentum is not related to $\mathbf{x}$ and $\mathbf{p}$ at all!

Again, we obtain finite rotations by exponentiation.

$$\begin{aligned} D_3(\varphi) &= \lim_{n\to\infty}\left[1 - i\left(\frac{J_z\varphi}{n\hbar}\right)\right]^n \\ &= e^{-\frac{i}{\hbar}J_z\varphi} = 1 - \frac{iJ_z\varphi}{\hbar} - \frac{J_z^2\varphi^2}{2\hbar^2} + \dots \end{aligned} \tag{4.10}$$

We need just one more concept in order to uniquely specify the angular momentum operator. We propose that

Theorem Postulate: $D(R)$ has the *same group properties* as R

In group theory one says that $D(R)$ is a representation of R. In detail, this means:

1. *Identity.* $R \cdot \mathbf{1} = R \Rightarrow D(R) \cdot D(\mathbf{1}) = D(R)$
2. *Closure.* $R_1 R_2 = R_3 \Rightarrow D(R_1)D(R_2) = D(R_3) = D(R_1 R_2)$
3. *Inverse.* $R^{-1}R = \mathbf{1}$, $RR^{-1} = \mathbf{1}$
   $\Rightarrow D(R)^{-1}D(R) = D(R^{-1})D(R) = \mathbf{1} = D(R)D(R)^{-1} = D(R)D(R^{-1})$
4. *Associativity.* $(R_1R_2)R_3 = R_1(R_2R_3) = R_1R_2R_3$
   $\Rightarrow [D(R_1)D(R_2)]D(R_3) = D(R_1)[D(R_2)D(R_3)] = D(R_1)D(R_2)D(R_3)$

Now we will write down the rotation operator analogue of (4.6):

$$\begin{aligned}&\left(1 - \tfrac{i}{\hbar}J_x\delta\varphi - \tfrac{1}{2\hbar^2}J_x^2\delta\varphi^2\right)\left(1 - \tfrac{i}{\hbar}J_y\delta\varphi - \tfrac{1}{2\hbar^2}J_y^2\delta\varphi^2\right)\\ &\quad - \left(1 - \tfrac{i}{\hbar}J_y\delta\varphi - \tfrac{1}{2\hbar^2}J_y^2\delta\varphi^2\right)\left(1 - \tfrac{i}{\hbar}J_x\delta\varphi - \tfrac{1}{2\hbar^2}J_x^2\delta\varphi^2\right) = \mathbf{1} - \tfrac{i}{\hbar}J_z\delta\varphi^2 - \mathbf{1}\end{aligned}$$

This gives the *fundamental commutation relations for the angular momentum operator* by equating terms of same order in $\delta\varphi^2$,

$$\left[J_x, J_y\right]\frac{\delta\varphi^2}{\hbar^2} = i\frac{J_z\delta\varphi^2}{\hbar} + \mathcal{O}(\delta\varphi^3) \tag{4.11}$$

$$\left[J_i, J_j\right] = i\hbar sum_k\varepsilon_{ijk}J_k \tag{4.12}$$

Note that in contrast to the translations, the generators of infinitesimal rotations do not commute, i.e., they form a *non-Abelian* group.

4.2 Spin $\frac{1}{2}$ Systems and Finite Rotations

The spin operators are

$$\begin{aligned}S_1 &= \tfrac{\hbar}{2}\left(\left|\uparrow\right\rangle\left\langle\downarrow\right| + \left|\downarrow\right\rangle\left\langle\uparrow\right|\right)\\ S_2 &= i\tfrac{\hbar}{2}\left(-\left|\uparrow\right\rangle\left\langle\downarrow\right| + \left|\downarrow\right\rangle\left\langle\uparrow\right|\right)\\ S_3 &= \tfrac{\hbar}{2}\left(\left|\uparrow\right\rangle\left\langle\uparrow\right| - \left|\downarrow\right\rangle\left\langle\downarrow\right|\right)\end{aligned} \tag{4.13}$$

It is easy to show that the S_i satisfy the fundamental commutation relations

$$\left[S_i, S_j\right] = i\hbar\sum_k \varepsilon_{ijk}S_k \tag{4.14}$$

Since the S_i satisfy the fundamental commutation relations of angular momentum, we propose that a rotation by φ about the z-axis can be expressed as

$$\begin{aligned}\left|\alpha\right\rangle_R &= D_3(\varphi)\left|\alpha\right\rangle\\ D_3(\varphi) &= e^{-\frac{i}{\hbar}S_3\varphi}\end{aligned} \tag{4.15}$$

Does this really make sense? Let's consider the expectation value, e. g. $\left\langle S_1\right\rangle_\alpha$.

$$\left\langle S_1\right\rangle_\alpha \longrightarrow \left\langle\alpha|S_1|\alpha\right\rangle_R = \left\langle\alpha|D_3^\dagger(\varphi)S_1D_3(\varphi)|\alpha\right\rangle \tag{4.16}$$

Does this really correspond to a rotation?

$$\begin{aligned} D_3^\dagger(\varphi) S_1 D_3(\varphi) &= \\ &= \\ &= \\ &= S_1 \cos\varphi - S_2 \sin\varphi \end{aligned} \tag{4.17}$$

Before interpreting this result, we will present an alternative way to derive this result which will use the fundamental commutation relation instead of being restricted just to the spin $\frac{1}{2}$ systems.

For this alternative derivation, we need the *Baker-Hausdorff lemma*: let G be Hermitian and λ a real number. Then we have the following identity:

$$\begin{aligned} e^{iG\lambda} A e^{-iG\lambda} &= A + i\lambda\big[G, A\big] + \frac{i^2\lambda^2}{2!}\big[G,[G,A]\big] + \dots \\ &+ \frac{i^n\lambda^n}{n!}[G,[G,[G,\dots[G,A]]]\dots] + \dots \end{aligned} \tag{4.18}$$

The proof is left as an exercise. Using this relation, we obtain

$$\begin{aligned} D_3^\dagger(\varphi) S_1 D_3(\varphi) &= \\ &= \\ &= S_1 \cos\varphi - S_2 \sin\varphi \end{aligned} \tag{4.19}$$

The expectation values of S_1 now becomes

$$\langle S_1\rangle_\alpha \longrightarrow \langle\alpha|S_1|\alpha\rangle_R = \langle S_1\rangle_\alpha \cos\varphi - \langle S_2\rangle_\alpha \sin\varphi \tag{4.20}$$

Analogously, for S_2

$$\langle S_2\rangle_\alpha \longrightarrow \langle\alpha|S_2|\alpha\rangle_R = \langle S_1\rangle_\alpha \sin\varphi + \langle S_2\rangle_\alpha \cos\varphi \tag{4.21}$$

Since S_3 commnutes with $D_3(\varphi)$, we have

$$\langle S_3\rangle_\alpha \longrightarrow \langle\alpha|S_3|\alpha\rangle_R = \langle S_3\rangle_\alpha \tag{4.22}$$

In general, we get that

$$\langle J_k\rangle \longrightarrow \sum_l R_{kl} \langle J_l\rangle \tag{4.23}$$

where $R = (R_{kl})$ is the rotation matrix. So, the expectation values behave like classical systems under rotation!

Now consider the rotation of a general ket, $D_3(\varphi)\,|\alpha\rangle$. We first decompose $|\alpha\rangle$ into base kets and apply $D_3(\varphi)$ to it:

$$\begin{aligned} |\alpha\rangle &= |{\uparrow}\rangle \langle{\uparrow}|\alpha\rangle + |{\downarrow}\rangle\langle{\downarrow}|\alpha\rangle \\ D_3(\varphi)\,|\alpha\rangle &= e^{-\frac{i}{\hbar}S_3\varphi}\,|\alpha\rangle = e^{-\frac{i}{2}\varphi}\,|{\uparrow}\rangle\langle{\uparrow}|\alpha\rangle + e^{\frac{i}{2}\varphi}\,|{\downarrow}\rangle\langle{\downarrow}|\alpha\rangle \end{aligned} \tag{4.24}$$

If we rotate by 2π, we get $-\left|\alpha\right\rangle$!

$$\left|\alpha\right\rangle_{R_3(2\pi)} \longrightarrow -\left|\alpha\right\rangle \tag{4.25}$$

This fact has been verified in neutron interferometry experiments.

Pauli Two-Component Formalism.

Here we introduce the tow-component spinor formalism introduced by Wolfgang Pauli, which allows a convenient treatment of spin $\frac{1}{2}$ systems. In general, spin states are represented by spinors:

$$\begin{aligned}
\left|\uparrow\right\rangle &= \begin{pmatrix}1\\0\end{pmatrix} = \chi_\uparrow \qquad \left|\downarrow\right\rangle = \begin{pmatrix}0\\1\end{pmatrix} = \chi_\downarrow \\
\left\langle\uparrow\right| &= (1,0) = \chi_\uparrow^\dagger \qquad \left\langle\downarrow\right| = (0,1) = \chi_\downarrow^\dagger \qquad (4.26)\\
\left|\alpha\right\rangle &= \begin{pmatrix}\left\langle\uparrow\right|\alpha\rangle\\ \left\langle\downarrow\right|\alpha\rangle\end{pmatrix} = \begin{pmatrix}c_\uparrow\\c_\downarrow\end{pmatrix} = \chi \qquad (4.27)
\end{aligned}$$

Spinors are usually denoted by χ. The S_j operators are represented by the so-called *Pauli matrices* σ_j.

$$\sigma_1 = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \qquad \sigma_2 = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix} \qquad \sigma_3 = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \tag{4.28}$$

Some properties are:

$$\begin{aligned}
\{\sigma_i,\sigma_j\} &= 2\delta_{ij} \ ,\\
[\sigma_i,\sigma_j] &= 2i\sum_k \varepsilon_{ijk}\sigma_k \ ,\\
\sigma_j^\dagger &= \sigma_j \ ,\\
\det(\sigma_j) &= -1 \ ,\\
\operatorname{tr}(\sigma_j) &= 0 \ . \qquad (4.29)
\end{aligned}$$

Let $\mathbf{a}$ be a vector in three dimensions. Then $\sigma\cdot\mathbf{a}$ is a (2×2) matrix given by

$$\sigma\cdot\mathbf{a} = \sum_k a_k\sigma_k = \begin{pmatrix}a_3 & a_1 - ia_2\\ a_1 + ia_2 & -a_3\end{pmatrix} \ . \tag{4.30}$$

Note that

$$\begin{aligned}
(\sigma\cdot\mathbf{a})(\sigma\cdot\mathbf{b}) &= \big(\sum_j a_j\sigma_j\big)\big(\sum_k b_k\sigma_k\big)\\
&= \sum_j\sum_k\big(\frac{1}{2}\{\sigma_j,\sigma_k\} + \frac{1}{2}[\sigma_j,\sigma_k]\big)a_jb_k\\
&= \mathbf{a}\cdot\mathbf{b}\mathbf{1} + i\sigma\cdot(\mathbf{a}\times\mathbf{b}) \qquad (4.31)
\end{aligned}$$

Let a_j be real.

$$\begin{aligned}
\Rightarrow\ \ (\sigma\cdot\mathbf{a})^2 &= |\mathbf{a}|^2\,\mathbf{1}\\
(\sigma\cdot\hat{n})^n &= \begin{cases}\mathbf{1} & n\in 2\mathbb{Z}\\ \sigma\cdot\hat{n} & n\in 2\mathbb{Z}+1\end{cases} \qquad (4.32)
\end{aligned}$$

Now we apply this to rotations $D(\hat{n}, \varphi)$:

$$D(\hat{n}, \varphi) = e^{-\frac{i}{\hbar}\mathbf{S}\cdot\hat{n}\varphi} = e^{-\frac{i}{2}\sigma\cdot\hat{n}\varphi} = \ldots = \mathbf{1}\cos\frac{\varphi}{2} - i\sigma\cdot\hat{n}\sin\frac{\varphi}{2} \tag{4.33}$$

A spinor transforms under rotation according to

$$\chi \longrightarrow e^{-\frac{i}{2}\sigma\cdot\hat{n}\varphi}\chi \tag{4.34}$$

The σ_ks are to remain unchanged under rotation. Thus σ is *not* to be regarded as a vector. Rather

$$\chi^\dagger\sigma_k\chi \longrightarrow \sum_l R_{kl}\chi^\dagger\sigma_k\chi \tag{4.35}$$

transforms as a vector.

As an application of the spinor formalism we repeat the problem of determining the eigenspinor of $\sigma\cdot\hat{n}$ with eigenvalue $+1$. $\hat{n}$ is characterized by a *polar angle* β and an *azimuthal angle* α. Here we achieve this by applying two rotations to $\chi_\uparrow$:

$$\begin{aligned} \chi_{\sigma\cdot\hat{n}} &= D_3(\alpha)D_2(\beta)\begin{pmatrix}1\\0\end{pmatrix} && (4.36)\\ &= \\ &= \\ &= \begin{pmatrix}\cos\frac{\beta}{2}e^{-i\frac{\alpha}{2}}\\ \sin\frac{\beta}{2}e^{i\frac{\alpha}{2}}\end{pmatrix} && (4.37)\end{aligned}$$

In one of the problems, the eigenket $|\mathbf{S}\cdot\hat{n};\uparrow\rangle$ has already been determined. The result was

$$|\mathbf{S}\cdot\hat{n};\uparrow\rangle = \cos\frac{\beta}{2}|\uparrow\rangle + \sin\frac{\beta}{2}e^{i\alpha}|\downarrow\rangle \tag{4.38}$$

This corresponds to the same state as $\chi_{\sigma\cdot\hat{n}}$, because the two representations only differ by an overall phase factor (here $e^{i\frac{\alpha}{2}}$) that does not change the ket.

4.3 Eigenvalues and Eigenstates of Angular Momentum

Instead of just concentrating on spin $\frac{1}{2}$ systems, we will now consider general angular momentum states. We define the operator $\mathbf{J}^2$ as

$$\mathbf{J}^2 \equiv J_xJ_x + J_yJ_y + J_zJ_z \tag{4.39}$$

It commutes with all the J_k operators:

$$\left[\mathbf{J}^2, J_k\right] = 0 \tag{4.40}$$

Proof

To prove this we check the case $[\mathbf{J}^2, J_z] = 0$.

$$\begin{aligned}\left[\mathbf{J}^2, J_z\right] &= \left[J_xJ_x + J_yJ_y + J_zJ_z, J_z\right] \\ &= J_x[J_x, J_z] + [J_x, J_z]J_x + J_y[J_y, J_z] + [J_y, J_z]J_y \\ &= J_x(-i\hbar J_y) - i\hbar J_yJ_x + J_y i\hbar J_x + i\hbar J_xJ_y \equiv 0 \end{aligned} \tag{4.41}$$

For the other two cases, the proof goes analogously. □

To be specific, we look for *simultaneous eigenkets of* $\mathbf{J}^2$ *and* J_z.

$$\mathbf{J}^2 |a, b\rangle = a |a, b\rangle \qquad J_z |a, b\rangle = b |a, b\rangle \tag{4.42}$$

The choice of J_z is just convention, we could have equally well chosen J_x or J_y. It is now convenient to define the *ladder operators*:

$$J_\pm = J_x \pm iJ_y \tag{4.43}$$

As we will see, they play a similar role as the annihilation and creation operator in the case of the harmonic oscillator. The commutation relations for the $J_\pm$ are

$$\left[J_+, J_-\right] = 2\hbar J_z \qquad \left[J_z, J_\pm\right] = \pm\hbar J_\pm \qquad \left[\mathbf{J}^2, J_\pm\right] = 0 \tag{4.44}$$

What is the effect of $J_\pm$ on $|a, b\rangle$?

$$\begin{aligned} J_z\big(J_\pm |a, b\rangle\big) &= \big(\underbrace{[J_z, J_\pm]}_{=\pm\hbar J_\pm} + J_\pm J_z\big) |a, b\rangle \\ &= (b \pm \hbar)J_\pm |a, b\rangle \end{aligned} \tag{4.45}$$

Thus, $J_\pm |a, b\rangle$ is an eigenvector of J_z with eigenvalues $b \pm \hbar$. (Recall the effect of $a^\dagger$ and a on the harmonic oscillator eigenstates.)

On other hand, $J_\pm$ does not change the eigenvalue of $\mathbf{J}^2$:

$$\mathbf{J}^2\big(J_\pm |a, b\rangle\big) = J_\pm \mathbf{J}^2 |a, b\rangle = aJ_\pm |a, b\rangle \tag{4.46}$$

In summary, we obtain

$$J_\pm |a, b\rangle = c_\pm |a, b \pm \hbar\rangle \tag{4.47}$$

with the proportionality constant $c_\pm$ yet unspecified. Now we rewrite $\mathbf{J}^2 - J_zJ_z$:

$$\mathbf{J}^2 - J_zJ_z = \frac{1}{2}\big(J_+J_- + J_-J_+\big) = \frac{1}{2}\big(J_+J_+^\dagger + J_+^\dagger J_+\big) \tag{4.48}$$

$J_+J_+^\dagger$ and $J_+^\dagger J_+$ both have nonnegative expectation values because of

$$J_+^\dagger |a, b\rangle \longleftrightarrow \langle a, b| J_+ \qquad J_+ |a, b\rangle \longleftrightarrow \langle a, b| J_+^\dagger \ . \tag{4.49}$$

Therefore

$$\big\langle a, b|\mathbf{J}^2 - J_zJ_z|a, b\big\rangle \geq 0 \Rightarrow (a - b^2) \geq 0 \ . \tag{4.50}$$

Consequently, $a > b^2$. Thus, there must be a $b_{\max}$ such that

$$J_+ |a, b_{\max}\rangle = 0 \; , \tag{4.51}$$

which also implies that

$$\Rightarrow \; J_- J_+ |a, b_{\max}\rangle = 0 \tag{4.52}$$

Again, we will rewrite the product

$$J_- J_+ = J_x J_x + J_y J_y - i(\underbrace{J_y J_x - J_x J_y}_{=[J_y,J_x]=-i\hbar J_z}) = \mathbf{J}^2 - J_z J_z - \hbar J_z \tag{4.53}$$

$$J_- J_+ |a, b_{\max}\rangle = \left(\mathbf{J}^2 - J_z J_z - \hbar J_z\right) |a, b_{\max}\rangle = 0 \tag{4.54}$$

$$\Rightarrow \; a - b_{\max}^2 - \hbar b_{\max} = 0 \; \Leftrightarrow \; a = b_{\max}(b_{\max} + \hbar) \tag{4.55}$$

Similarly, we can argue that there is a minimal value of $b = b_{\min}$ with

$$J_- |a, b_{\min}\rangle = 0$$

$$J_+ J_- = \mathbf{J}^2 - J_z J_z + \hbar J_z$$

$$J_+ J_- |a, b_{\min}\rangle = (a - b_{\min}^2 + \hbar b_{\min}) |a, b_{\min}\rangle \tag{4.56}$$

$$\Rightarrow \; a = b_{\min}(b_{\min} - \hbar) \tag{4.57}$$

Hence, $b_{\max} = -b_{\min}$; all allowed values of b lie in between $-b_{\max} \leq b \leq b_{\max}$.

What is $b_{\max} - b_{\min}$? $|a, b_{\max}\rangle$ must be reached from $|a, b_{\min}\rangle$ by applying J_+ a finite number of times. Therefore we have

$$|a, b_{\max}\rangle = c J_+^n |a, b_{\min}\rangle \tag{4.58}$$

$$\Rightarrow \; b_{\max} = b_{\min} + n\hbar = \tfrac{n\hbar}{2} \tag{4.59}$$

with n some integer. Usually one prefers to work with quantum *numbers* j and m that are defined by

$$j \equiv \tfrac{b_{\max}}{\hbar} = \tfrac{n}{2} \qquad a \equiv \hbar^2 j(j+1) \qquad m \equiv \tfrac{b}{\hbar} \tag{4.60}$$

The allowed m values range from $-j$ to j, i.e. are $2j+1$ allowed m values. If j is an integer, all m are also integers, while if j is a half-integer, all m values are half-integer.

Simultaneous eigenkets of $\mathbf{J}^2$ and J_z are usually denoted by $|j, m\rangle$ which obey the eigenvalue equations

$$\mathbf{J}^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle \tag{4.61}$$

$$J_z |j, m\rangle = m\hbar |j, m\rangle \tag{4.62}$$

j is called the rotational quantum number while m is the azimuthal quantum number.

4.3.1 Matrix Elements of Angular Momentum Operators

Since $|j, m\rangle$ is an eigenstate of both $\mathbf{J}^2$ and J_z, we obviously have

$$\begin{aligned} \langle j', m' | \mathbf{J}^2 | j, m \rangle &= j(j+1)\hbar^2 \delta_{j,j'} \delta_{m,m'} \\ \langle j', m' | J_z | j, m \rangle &= m\hbar \delta_{j,j'} \delta_{m,m'} \end{aligned} \tag{4.63}$$

What are the matrix elements of $J_\pm$? We first consider

$$\begin{aligned}\langle j', m'|J_+^\dagger J_+|j,m\rangle &= \langle j', m'|\mathbf{J}^2 - J_z J_z - \hbar J_z|j,m\rangle \\ &= \hbar^2\big(j(j+1) - m^2 - m\big)\delta_{j,j'}\delta_{m,m'}\end{aligned} \tag{4.64}$$

Now $J_+ |j,m\rangle = c_{jm}^+ |j, m+1\rangle$. A comparison with (4.64) yields

$$\left|c_{jm}^+\right| = \hbar^2\left[j(j+1) - m(m+1)\right] = \hbar^2(j-m)(j+m+1) \tag{4.65}$$

We demand that c_{jm}^+ is real and positive by convention, thus

$$J_+ |j,m\rangle = \sqrt{(j-m)(j+m+1)}\, |j, m+1\rangle \tag{4.66}$$

Analogously, we get for J_-

$$J_- |j,m\rangle = \sqrt{(j+m)(j-m+1)}\, |j, m-1\rangle \tag{4.67}$$

$$\Rightarrow\ \langle j', m'|J_\pm|j,m\rangle = \sqrt{(j \mp m)(j \pm m + 1)}\hbar\delta_{j',j}\delta_{m',m\pm1}\ . \tag{4.68}$$

4.3.2 Representations of the Rotation Operator

Now we are looking for the matrix elements of the rotation operator:

$$D_{m,m'}^{(j)}(R) = \left\langle j, m|e^{-\frac{i}{\hbar}\mathbf{J}\cdot\hat{n}\varphi}|j,m'\right\rangle \tag{4.69}$$

Since $\mathbf{J}^2$ commutes with any function of the J_k, $D(R)\,|j,m\rangle$ is still an eigenket of $\mathbf{J}^2$, as can be seen from

$$\mathbf{J}^2 D(R)\,|j,m\rangle = D(R)\mathbf{J}^2\,|j,m\rangle = j(j+1)\hbar^2 D(R)\,|j,m\rangle \tag{4.70}$$

Hence matrix elements between states with different j-values vanish, as already incorporated in (**??**). This means in other words that rotations do not change the j-value.

Consequently, $D_{m,m'}^{(j)}(R)$ is a $(2j+1)\times(2j+1)$ matrix. It represents an *irreducible representation of the rotation operator*, i. e. it cannot be diagonalized any further. The rotation matrices characterized by a specific j form a group. For $\varphi = 0$ we have the identity, and a rotation around an arbitrary axis in the opposite direction to the one considered is the inverse. Furthermore, the product of any two members is also a member of the group, i.e.

$$D_{m,m''}^{(j)}(R_1 R_2) = \sum_{m'} D_{m,m'}^{(j)}(R_1) D_{m',m''}^{(j)}(R_2)\ . \tag{4.71}$$

Thus, the structure of $SO(3)$ is carried over to $D(R)$, i. e. the group of all the $D(R)$ has to be *unitary*.

$$D_{m,m'}^{(j)}(R^{-1}) = D_{m,m'}^{(j)\dagger}(R) = D_{m',m}^{(j)*}(R) \tag{4.72}$$

Be careful, the order of m and m' has changed!

In order to understand the meaning of the matrix elements of the rotation operator, we rotate a state $|j,m\rangle$, i.e. it is transformed $|j,m\rangle \to D(R)\,|j,m\rangle$. Now we use the fact that $D(R)$ does not change the angular momentum j of a ket, so that for the insertion of the identity operator expressed as a sum over angular momentum eigenstates we only have to consider states with the same j:

$$D(R)\,|j,m\rangle = \sum_{m'} |j,m'\rangle \, \langle j,m'|\, D(R)\,|j,m\rangle = \sum_{m'} |j,m'\rangle \, D^{(j)}_{m',m}(R) \tag{4.73}$$

Hence we see that $D^{(j)}_{m',m}(R)$ is simply the amplitude for the rotated state to be in $|j,m'\rangle$ when the unrotated state was $|j,m\rangle$.

For spin $\frac{1}{2}$ systems, we have in fact already determined the rotation matrices in terms of the Pauli matrices (see (4.33)). For a general rotation, we have

$$D^{\frac{1}{2}}_{m,m'}(\hat{n},\varphi) = \begin{pmatrix} \cos\frac{\varphi}{2} - in_3 \sin\frac{\varphi}{2} & -(in_1 + n_2)\sin\frac{\varphi}{2} \\ (-in_1 + n_2)\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} + in_3\sin\frac{\varphi}{2} \end{pmatrix} \tag{4.74}$$

4.4 Orbital Angular Momentum

In order to derive the eigenstates, eigenvalues and matrix elements of the angular momentum operators, we only used the fact that the angular momentum is the generator of rotation. In classical mechanics, the angular momentum is defined as

$$\mathbf{L} \equiv \mathbf{x} \times \mathbf{p}\ . \tag{4.75}$$

In quantum mechanics, there is also the spin-angular momentum that has no classical equivalent. In order to distinguish (4.75) from the spin-angular momentum, it is refered to as the *orbital angular momentum* $\mathbf{L}$. If the spin-angular momentum is 0 or can be ignored, then the angular momentum $\mathbf{J}$ for a single particle is the same as the orbital angular momentum $\mathbf{L}$.

It is easy to show that $\mathbf{L}$ satisfies the fundamental commutation relations

$$\big[L_i, L_j\big] = i\hbar \sum_k \varepsilon_{ijk} L_k \tag{4.76}$$

$$\big[\mathbf{L}^2, L_i\big] = 0 \tag{4.77}$$

Now the question is: *Does* $\mathbf{L}$ *indeed generate rotation?*

In order to see this we let $1 - \frac{i}{\hbar}\delta\varphi L_z$ act on a position eigenket $|\mathbf{x}\rangle$. Using $[x_1, p_2] = 0$ and that $\mathbf{p}$ generates translations, we obtain

$$\begin{aligned}\big[1 - \frac{i}{\hbar}\delta\varphi L_z\big]\,|\mathbf{x}\rangle &= \big[1 - \frac{i}{\hbar}\delta\varphi(x_1 p_2 - x_2 p_1)\big]\,|\mathbf{x}\rangle = \big[1 - \frac{i}{\hbar} p_2 \delta\varphi x_1 + \frac{i}{\hbar} p_1 \delta\varphi x_2\big]\,|\mathbf{x}\rangle \\ &= |(x_1 - \delta\varphi x_2, x_2 + x_1\delta\varphi, x_3)\rangle \end{aligned} \tag{4.78}$$

This corresponds in fact to an infinitesimal rotation about the x_3 axis.

Now we will apply this operator to an arbitrary state ket $|\alpha\rangle$.

$$\left\langle \mathbf{x}|1 - \frac{i}{\hbar}\delta\varphi L_3|\alpha \right\rangle = \langle (x_1 + x_2\delta\varphi, x_2 - x_1\delta\varphi, x_3)|\alpha\rangle \tag{4.79}$$

Instead of Cartesian coordinates, we will use spherical coordinates

$$\begin{aligned}\left\langle r,\vartheta,\varphi|1 - \frac{i}{\hbar}\delta\varphi L_3|\alpha \right\rangle &= \langle r,\vartheta,\varphi - \delta\varphi|\alpha\rangle \\ &= \langle r,\vartheta,\varphi|\alpha\rangle - \delta\varphi \frac{\partial}{\partial\varphi}\,\langle r,\vartheta,\varphi|\alpha\rangle\ , \end{aligned} \tag{4.80}$$

where we did a Taylor expansion of the wave function

$$\begin{aligned}\langle r,\vartheta,\varphi-\delta\varphi|\alpha\rangle &= \psi_\alpha(r,\vartheta,\varphi-\delta\varphi)\\ &= \psi_\alpha(r,\vartheta,\varphi)-\delta\varphi\frac{\partial}{\partial\varphi}\psi_\alpha(r,\vartheta,\varphi)+\frac{1}{2}(\delta\varphi)^2\frac{\partial^2}{\partial\varphi^2}\psi_\alpha(r,\vartheta,\varphi)+\dots\ .\end{aligned} \tag{4.81}$$

Since $\langle r,\vartheta,\varphi|$ is arbitrary, we can identify

$$\langle \mathbf{x}|L_3|\alpha\rangle = -i\hbar\frac{\partial}{\partial\varphi}\langle\mathbf{x}|\alpha\rangle \tag{4.82}$$

Using $\mathbf{x} = re_r$ and

$$\nabla = e_r\frac{\partial}{\partial r}+e_\vartheta\frac{1}{r}\frac{\partial}{\partial\vartheta}+e_\varphi\frac{1}{r\sin\vartheta}\frac{\partial}{\partial\varphi}\ , \tag{4.83}$$

we obtain

$$\Rightarrow\ \mathbf{x}\times\nabla = e_\varphi\frac{\partial}{\partial\vartheta}-e_\vartheta\frac{1}{\sin\vartheta}\frac{\partial}{\partial\varphi} \tag{4.84}$$

$$\Rightarrow\ \langle\mathbf{x}|L_1|\alpha\rangle = -i\hbar\left(-\sin\varphi\frac{\partial}{\partial\vartheta}-\cos\varphi\cot\vartheta\frac{\partial}{\partial\varphi}\right)\langle\mathbf{x}|\alpha\rangle \tag{4.85}$$

$$\Rightarrow\ \langle\mathbf{x}|L_2|\alpha\rangle = -i\hbar\left(-\cos\varphi\frac{\partial}{\partial\vartheta}-\sin\varphi\cot\vartheta\frac{\partial}{\partial\varphi}\right)\langle\mathbf{x}|\alpha\rangle \tag{4.86}$$

$$\langle\mathbf{x}|L_\pm|\alpha\rangle = -i\hbar e^{\pm i\varphi}\left(\pm i\frac{\partial}{\partial\vartheta}-\cot\vartheta\frac{\partial}{\partial\varphi}\right)\langle\mathbf{x}|\alpha\rangle \tag{4.87}$$

Using $\mathbf{L}^2 = L_3^2+\frac{1}{2}\big(L_+L_-+L_-L_+\big)$, it follows

$$\big\langle x|\mathbf{L}^2|\alpha\big\rangle = -\hbar^2\left[\frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}+\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right)\right]\langle\mathbf{x}|\alpha\rangle \tag{4.88}$$

Apart from $1/r^2$, this is just the angular part of the Laplacian ∇^2 in spherical coordinates. This connection can also be established by reformulating the kinetic-energy operator in spherical coordinates. In the appendix at the end of this chapter, it is explicitly shown that the kinetic energy operator $\frac{\mathbf{p}^2}{2m}$ can be expressed as

$$\frac{1}{2m}\big\langle\mathbf{x}|\mathbf{p}^2|\alpha\big\rangle = -\frac{\hbar^2}{2m}\nabla^2\langle\mathbf{x}|\alpha\rangle = -\frac{\hbar^2}{2m}\left[\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\langle\mathbf{x}|\alpha\rangle-\frac{1}{\hbar^2r^2}\big\langle\mathbf{x}|\mathbf{L}^2|\alpha\big\rangle\right] \tag{4.89}$$

The first term is the radial part of the Laplacian. The last term must be the angular part of the Laplacian, as can explicitly seen in (4.88).

4.5 The Central Potential

With central potential a spherical symmetric potential is meant, i.e., $V \equiv V(|\mathbf{r}|) = V(r)$. Thus, H is rotationally invariant.

$$\begin{aligned}D^\dagger(R)HD(R) &= H\\ \big[H,L_i\big] &= 0 = \big[H,\mathbf{L}^2\big]\end{aligned} \tag{4.90}$$

Therefore, energy eigenkets are also eigenkets of $\mathbf{L}^2$ and L_3. The eigenfunctions of L_3 and $\mathbf{L}^2$ are denoted by

$$\langle \hat{n} | lm \rangle = Y_{lm}(\vartheta, \varphi) \tag{4.91}$$

$Y_{lm}(\vartheta, \varphi)$ are called spherical harmonics. The eigenvalue equations for L_3 and $\mathbf{L}^2$ can be re-written as differential equations.

$$L_3 \, |lm\rangle = m\hbar \, |lm\rangle \tag{4.92}$$

$$\Rightarrow -i\hbar \frac{\partial}{\partial \varphi} Y_{lm}(\vartheta, \varphi) = m\hbar Y_{lm}(\vartheta, \varphi) \tag{4.93}$$

$$\Rightarrow Y_{lm} \propto e^{im\varphi} \tag{4.94}$$

Likewise,

$$\mathbf{L}^2 \, |lm\rangle = l(l+1)\hbar^2 \, |lm\rangle \tag{4.95}$$

defines a differential equation:

$$\left[\frac{1}{\sin\vartheta} \frac{\partial}{\partial \vartheta} \left(\sin\vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{\sin^2\vartheta} \frac{\partial^2}{\partial \varphi^2} + l(l+1) \right] Y_{lm}(\vartheta, \varphi) = 0 \tag{4.96}$$

In general, the spherical harmonics are given by

$$Y_{lm}(\vartheta, \varphi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{2l+1}{4\pi} \frac{(l+m)!}{(l-m)!}} e^{im\varphi} \frac{1}{\sin^m \vartheta} \frac{d^{l-m}}{d\cos^{l-m}\vartheta} \sin^{2l}\vartheta \tag{4.97}$$

They are orthogonal to each other and form a complete set:

$$\int_0^\pi d\vartheta \sin\vartheta \int_0^{2\pi} d\varphi \; Y^*_{lm}(\vartheta, \varphi) Y_{l'm'}(\vartheta, \varphi) = \delta_{ll'} \delta_{mm'} \tag{4.98}$$

Further details about the spherical harmonics are given in the appendix in section 4.7.2. Here we just report the lowest Y_{lm}'s which are

$$Y_{00} = \frac{1}{\sqrt{4\pi}} \qquad Y_{10} = \sqrt{\frac{3}{4\pi}} \cos\vartheta \qquad Y_{11} = -\sqrt{\frac{3}{8\pi}} \sin\vartheta e^{i\varphi} \tag{4.99}$$

For $l = 2$, the functions are slightly more complicated.

$$Y_{20} = \sqrt{\frac{5}{16\pi}} (3\cos^2\vartheta - 1) \qquad Y_{21} = -\sqrt{\frac{15}{8\pi}} \sin\vartheta \cos\vartheta \cdot e^{i\varphi} \qquad Y_{22} = \sqrt{\frac{15}{32\pi}} \sin^2\vartheta \cdot e^{2i\varphi} \tag{4.100}$$

4.5.1 Schrödinger Equation for Central Potential Problems

The Schrödinger equation for a central potential is given by

$$\left[-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{\mathbf{L}^2}{2mr^2} + V(r) \right] \Psi(r, \vartheta, \varphi) = E \Psi(r, \vartheta, \varphi) \tag{4.101}$$

where $V(r)$ has to satisfy $\lim\limits_{r \to 0} r^2 V(r) = 0$.

Since H is spherically symmetric, $\mathbf{L}^2$ commutes with H. Thus there is a common set of eigenfunctions. The only term depending on the angles is $\mathbf{L}^2/(2mr^2)$. In order to solve the Schrödinger equation, we use as an ansatz the *separation of variables*.

$$\Psi(r,\vartheta,\varphi) = R(r)Y_{lm}(\vartheta,\varphi) \tag{4.102}$$

Substituting (4.102) into (4.101) and using $\mathbf{L}^2 Y_{lm} = \hbar^2 l(l+1)Y_{lm}$ leads to a differential equation for the radial part.

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right] R(r) = ER(r) \tag{4.103}$$

(4.103) *almost* looks like a one-dimensional Schrödinger equation except for the kinetic energy term. Substituting $R(r) \equiv \frac{U(r)}{r}$ results in

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) R(r) = \left(\frac{1}{r}\frac{\partial}{\partial r}r\right)^2 \frac{U(r)}{r} = \frac{1}{r}\frac{\partial^2}{\partial r^2}U(r) \tag{4.104}$$

$$\Rightarrow \left[-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right] U(r) = EU(r) \tag{4.105}$$

Thus, we have a one-dimensional problem with the *effective potential*

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} \tag{4.106}$$

There are some constraints on the wave function:

1. **Normalizability:** If we assume that the angular part is already normalized, then

$$\int |\Psi(r,\vartheta,\varphi)|^2 \, d^3r = \int_{\mathbb{R}^+} r^2 \frac{|U(r)|^2}{r^2} dr < \infty$$

$$\Rightarrow \lim_{r\to\infty} |U(r)| \le \frac{a}{r^{\varepsilon+\frac{1}{2}}} \qquad \varepsilon > 0 \tag{4.107}$$

2. $\lim_{r\to 0} U(r) = 0$ because otherwise the Schrödinger equation cannot be satisfied because of $\nabla^2\Psi \propto \nabla^2\frac{U(r)}{r} \propto \frac{1}{r^3}$ for $r \to 0$.

4.5.2 Examples for Spherically Symmetric Potentials

Free Particles

Here, $V(r) \equiv 0$. The general solutions are

$$R(r) = c_1 j_l(\rho) + c_2 \eta_l(\rho) \tag{4.108}$$

where $\rho = kr$ and $k = \frac{1}{\hbar}\sqrt{2mE}$. j_l and η_l are the spherical Bessel and Hankel functions, respectively.

The Coulomb Potential

The potential is

$$V(r) = -\frac{Ze^2}{r} \tag{4.109}$$

This corresponds to a hydrogen-like atom, and for $Z = 1$ it is the hydrogen atom. The solution of the radial Schrödinger equation giving the energy eigenvalues is somewhat involved, but straightforward. Solving this problem was one of the first successes of quantum mechanics since from the eigenvalues the measured hydrogen spectra could be immediately understood.

Here we will not give a derivation of the results, but simply list them. In order to write down the solutions, we first introduce the dimensionless variable

$$\rho \equiv \frac{1}{\hbar}\sqrt{8m_e\,|E|}r \tag{4.110}$$

The energy eigenfunctions are

$$\Psi_{nlm}(r,\vartheta,\varphi) = R_{nl}(r)Y_{lm}(\vartheta,\varphi) \tag{4.111}$$

where

$$R_{nl}(r) = -\sqrt{\left(\frac{2Z}{na_0}\right)^2 \frac{(n-l-1)!}{2n\big((n+l)!\big)^3}}\,e^{-\frac{\rho}{2}}\rho^l L_{n+1}^{2l+1}(\rho) \tag{4.112}$$

The corresponding energy levels are

$$\begin{gathered} E_n = -\frac{mZ^2e^4}{2\hbar^2}\cdot\frac{1}{n^2} = -\frac{Z^2e^2}{2a_0}\cdot\frac{1}{n^2} \quad \text{with} \quad a_0 = \frac{\hbar^2}{m_e e^2} \approx 0.529\ \text{Å} \\ n \geq l+1 \qquad \rho = \frac{2Zr}{na_0}\ , \end{gathered} \tag{4.113}$$

where a_0 is the *Bohr radius*. The $L_p^q(\rho)$ are the *associated Laguerre polynomials*.

$$L_p(\rho) = e^\rho \frac{d^p}{d\rho^p}\left(\rho^p e^{-\rho}\right) \implies L_p^q(\rho) = \frac{d^q}{d\rho^q}L_p(\rho) \tag{4.114}$$

Some radial functions for low n

$$\begin{aligned} R_{10}(r) &= 2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Zr}{a_0}} \\ R_{20}(r) &= 2\left(\frac{Z}{2a_0}\right)^{\frac{3}{2}}\left(1-\frac{Zr}{2a_0}\right)e^{-\frac{Zr}{2a_0}} \\ R_{21}(r) &= \left(\frac{Z}{2a_0}\right)^{\frac{3}{2}}\frac{Zr}{\sqrt{3}a_0}e^{-\frac{Zr}{2a_0}} \end{aligned} \tag{4.115}$$

4.6 Addition of Angular Momentum

4.6.1 Orbital Angular Momentum and Spin $\frac{1}{2}$

For any realistic description of a particle with spin one must take into account both the spin and the spatial degrees of freedom. For a spin $\frac{1}{2}$ system, e.g., the appropriate Hilbert space can be regarded as being a direct product space of the two-dimensional spin space with the infinite-dimensional space of the spatial degrees of freedom. The base kets are then given by

$$|\mathbf{x}, \pm\rangle = |\mathbf{x}\rangle \otimes |\pm\rangle \ , \tag{4.116}$$

where $\otimes$ is the tensor product. Any operator in the space spanned by $\{|\mathbf{x}\rangle\}$ commutes with any operator in the spin space described by $\{|\pm\rangle\}$. The *rotation operator* is still of the form $e^{-\frac{i}{\hbar}\mathbf{J}\cdot\hat{n}\varphi}$, but $\mathbf{J}$ consists of two parts, namely,

$$\mathbf{J} \equiv \mathbf{L} \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{S} \equiv \mathbf{L} + \mathbf{S} \tag{4.117}$$

$$\big[\mathbf{L}, \mathbf{S}\big] = 0 \ \Rightarrow D(R) = D^{\mathrm{orb}}(R) \otimes D^{\mathrm{spin}}(R) \tag{4.118}$$

Since $\mathbf{L}$ and $\mathbf{S}$ commute, $D(R)$ can be written as

$$D(R) = e^{-\frac{i}{\hbar}\mathbf{L}\cdot\hat{n}\varphi} \otimes e^{-\frac{i}{\hbar}\mathbf{S}\cdot\hat{n}\varphi} \tag{4.119}$$

A wave function for a particle with spin $\frac{1}{2}$ has two components:

$$\langle \mathbf{x}, \pm|\alpha\rangle = \Psi_{\pm}(\mathbf{x}) = \begin{pmatrix} \Psi_{+}(\mathbf{x}) \\ \Psi_{-}(\mathbf{x}) \end{pmatrix} \tag{4.120}$$

There are $2(2l+1)$ product states $|l, m, \pm\rangle$ that are eigenstates of $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, and S_z. As we will see later, we can also use a basis made of eigenstates of $\mathbf{J}^2$, $\mathbf{L}^2$, $\mathbf{S}^2$, and J_z. $|j, l, m\rangle$ will denote such an eigenstate.

What is the total angular momentum if we add the orbital and the spin angular momentum? We assume that for a spin $\frac{1}{2}$ particle $j = l \pm \frac{1}{2}$. Does this make sense? In other words, what is the dimension of this space? The number of states is $2(l+\frac{1}{2}) + 1 + 2(l-\frac{1}{2}) + 1 = 2(2l+1)$, so we obtain the right number of states.

As a starting point, we make the ansatz that

$$|l + \frac{1}{2}, l + \frac{1}{2}, l\rangle = |l, l, \uparrow\rangle \ . \tag{4.121}$$

The state of the left-hand side is an eigenstate of $\mathbf{J}^2$, J_z, $\mathbf{L}^2$, and $\mathbf{S}^2$, the state on the right-hand is an eigenstate of $\mathbf{L}^2$, L_z, S_z, and $\mathbf{S}^2$. It is simple to proove that this ansatz makes sense.

Proof

□

Any general state $\left|l+\frac{1}{2},m_j,l\right\rangle$ can be created by successive application of $J_- = L_- + S_-$ (recall that the ladder operators only change the z component of the angular momentum, the total angular momentum remains unchanged!). First, we apply J_- to the state $|l,l,\uparrow\rangle$.

$$J_-\,|l,l,\uparrow\rangle = \sqrt{2l}\hbar\,|l,l-1,\uparrow\rangle + \hbar\,|l,l,\downarrow\rangle \tag{4.125}$$

We normalize the right-hand side.

$$\left|l+\frac{1}{2},l-\frac{1}{2},l\right\rangle = \sqrt{\frac{2l}{2l+1}}\,|l,l-1,\uparrow\rangle + \sqrt{\frac{1}{2l+1}}\,|l,l,\downarrow\rangle \tag{4.126}$$

In general, we get

$$\left|l+\frac{1}{2},m_j,l\right\rangle = \sqrt{\frac{l+m_j+\frac{1}{2}}{2l+1}}\left|l,m_j-\frac{1}{2},\uparrow\right\rangle + \sqrt{\frac{l-m_j+\frac{1}{2}}{2l+1}}\left|l,m_j+\frac{1}{2},\downarrow\right\rangle \tag{4.127}$$

where $-(l+\frac{1}{2}) \le m_j \le l+\frac{1}{2}$.

Analogously, we get

$$\left|l-\frac{1}{2},m_j,l\right\rangle = -\sqrt{\frac{l-m_j+\frac{1}{2}}{2l+1}}\left|l,m_j-\frac{1}{2},\uparrow\right\rangle + \sqrt{\frac{l+m_j+\frac{1}{2}}{2l+1}}\left|l,m_j+\frac{1}{2},\downarrow\right\rangle \tag{4.128}$$

where $-(l-\frac{1}{2}) \le m_j \le l-\frac{1}{2}$. It is simple to check that these states (4.128) are orthogonal to the states (4.127) and that they are indeed eigenstates of $\mathbf{J}^2$ and J_z with eigenvalues $j = l-\frac{1}{2}$ and m_j.

A useful property of the states is the fact that they are eigenstates of the operator $\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\left(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\right)$ which appears in the spin-orbit interaction (see section 5.2.2).

4.6.2 Two Spin $\frac{1}{2}$ Particles

In the case of two spin $\frac{1}{2}$ particles, the total spin operator is given by

$$\mathbf{S} \equiv \mathbf{S}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{S}_2 \equiv \mathbf{S}_1 + \mathbf{S}_2 \tag{4.129}$$

The commutation relations are

$$\left[S_{1x_i}, S_{1x_j}\right] = i\hbar\varepsilon_{ijk}S_{1x_j} \qquad \left[S_{2x_i}, S_{2x_j}\right] = i\hbar\varepsilon_{ijk}S_{2x_j} \tag{4.130}$$

so that we have for the total spin

$$\left[S_x, S_y\right] = i\hbar S_z\ , \tag{4.131}$$

and so forth. The eigenvalues of the spin operators are

$$\begin{aligned} \mathbf{S}^2 = (\mathbf{S}_1+\mathbf{S}_2)^2 &: s(s+1)\hbar^2 \\ S_z = S_{1z}+S_{2z} &: m\hbar \\ S_{1z} &: m_1\hbar \\ S_{2z} &: m_2\hbar \end{aligned} \tag{4.132}$$

Again, there are two different representations of the total spin states.

1. The $\{m_1, m_2\}$ representation based on the eigenkets of S_{1z} and S_{2z}:

$$|\uparrow\uparrow\rangle\,,\ |\uparrow\downarrow\rangle\,,\ |\downarrow\uparrow\rangle\,,\ |\downarrow\downarrow\rangle\ . \tag{4.133}$$

2. The $\{s, m\}$ or triplet-singlet representation based on the eigenkets of $\mathbf{S}^2$ and S_z:

$$\begin{aligned}
|s=1, m=1\rangle &= |\uparrow\uparrow\rangle\ ,\\
|s=1, m=0\rangle &= \left(\frac{1}{\sqrt{2}}\right)(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\ ,\\
|s=1, m=1\rangle &= |\downarrow\downarrow\rangle\ ,\\
|s=0, m=0\rangle &= \left(\frac{1}{\sqrt{2}}\right)(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)\ ,
\end{aligned} \tag{4.134}$$

The $|s=1, m=\pm1, 0\rangle$ states are referred to as the spin triplet, while $|s=0, m=0\rangle$ is called the spin singlet state.

4.6.3 General Case

No we will add two general angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$. J_{1z} and J_{2z} are the z -components of the angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$, respectively. $\mathbf{J} = \mathbf{J}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{J}_2$ is the total angular momentum and J_z its z-component. We can easily show that the components of the total angular momentum satisfy the angular momentum commutation relations:

$$\left[J_i, J_j\right] = i\hbar\varepsilon_{ijk}J_k \tag{4.135}$$

This means that $\mathbf{J}$ is an angular momentum for itself. There are two choices for the chosen base. The first is spanned by the eigenkets $|\alpha\rangle = |j_1, j_2, m_1, m_2\rangle$. Consequently, the eigenvalues of these states are simply given by

$$\mathbf{J}_1^2\,|j_1, j_2, m_1, m_2\rangle = j_1(j_1+1)\hbar^2\,|j_1, j_2, m_1, m_2\rangle \tag{4.136}$$

$$J_{1z}\,|j_1, j_2, m_1, m_2\rangle = m_1\hbar\,|j_1, j_2, m_1, m_2\rangle \tag{4.137}$$

$$\mathbf{J}_2^2\,|j_1, j_2, m_1, m_2\rangle = j_2(j_2+1)\hbar^2\,|j_1, j_2, m_1, m_2\rangle \tag{4.138}$$

$$J_{2z}\,|j_1, j_2, m_1, m_2\rangle = m_2\hbar\,|j_1, j_2, m_1, m_2\rangle \tag{4.139}$$

Equivalently, we could use the base spanned by the eigenvectors of $\mathbf{J}^2$, J_z, $\mathbf{J}_1^2$, and $\mathbf{J}_2^2$. Note that the total angular momentum operator commutes with $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$, i.e. $\left[\mathbf{J}^2, \mathbf{J}_i^2\right] = 0$, which can be directly seen by writing $\mathbf{J}^2$ as

$$\mathbf{J}^2 = \mathbf{J}_1^2 + \mathbf{J}_2^2 + 2J_{1z}J_{2z} + J_{1+}J_{2-} + J_{1-}J_{2+} \tag{4.140}$$

In this representation, the eigenstates are denoted by $|\alpha\rangle = |j, m_j, j_1, j_2\rangle$. Here, the eigenvalue equations are

$$\mathbf{J}^2\,|j, m_j, j_1, j_2\rangle = j(j+1)\hbar^2\,|j, m_j, j_1, j_2\rangle \tag{4.141}$$

$$J_z\,|j, m_j, j_1, j_2\rangle = m_j\hbar\,|j, m_j, j_1, j_2\rangle \tag{4.142}$$

$$\mathbf{J}_1^2\,|j, m_j, j_1, j_2\rangle = j_1(j_1+1)\hbar^2\,|j, m_j, j_1, j_2\rangle \tag{4.143}$$

$$\mathbf{J}_2^2\,|j, m_j, j_1, j_2\rangle = j_2(j_2+1)\hbar^2\,|j, m_j, j_1, j_2\rangle \tag{4.144}$$

$\mathbf{J}^2$ commutes with J_z, but it *does not commute with* J_{1z} or J_{2z}!

Change of base As usual, we can change bases according to

$$|j, m_j, j_1, j_2\rangle = \sum_{m_1} \sum_{m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle \ , \tag{4.145}$$

where we have used that

$$\sum_{m_1} \sum_{m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2| = 1 \tag{4.146}$$

in the ket space of given angular momentum quantum numbers j_1 and j_2. The elements of the transformation matrix $\langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle$ are the so-called *Clebsch-Gordon coefficients*.

Clebsch-Gordon coefficients In fact, many of the Clebsch-Gordon coefficients appearing in (4.145) vanish, which makes it easier to perform the transformation. First of all, they vanish unless $m_j = m_1 + m_2$. The proof is rather simple.

$$\begin{aligned} 0 &= \\ \Rightarrow 0 &= \\ &= \end{aligned} \tag{4.147}$$

Secondly, they also vanish unless $|j_1 - j_2| \leq j \leq |j_1 + j_2|$. This relation seems to be obvious when we visualize $\mathbf{J}$ as the vectorial sum of $\mathbf{J}_1$ and $\mathbf{J}_2$. Its consistency can also be checked when the dimensionality of the space spanned by $\{|j_1, j_2, m_1, m_2\rangle\}$ is compared to that of $\{|j, m_j, j_1, j_2\rangle\}$ which should be the same.

The Clebsch-Gordon coefficients form a unitary matrix. Furthermore, the matrix elements are chosen to be real by convention. A real unitary matrix is orthogonal, which means that

$$\begin{gathered} \sum_{m_1} \sum_{m_2} \langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle \langle j_1, j_2, m_1, m_2 | j', m_j', j_1, j_2\rangle = \delta_{j\,j'} \delta_{m_j\,m_j'} \\ \sum_{j} \sum_{m_j} \langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle \langle j_1, j_2, m_1', m_2' | j, m_j, j_1, j_2\rangle = \delta_{m_1\,m_1'} \delta_{m_2\,m_2'} \end{gathered} \tag{4.148}$$

As an example, we write down the 2×2 matrix formed by the Clebsch-Gordan coefficients for the addition of orbital and spin-angular momentum $\left\langle l, \frac{1}{2}, m_1, m_2 | j, m_j, l, \frac{1}{2}\right\rangle$:

$$\begin{pmatrix} \sqrt{\frac{l+m_j+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l-m_j+\frac{1}{2}}{2l+1}} \\ -\sqrt{\frac{l-m_j+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l+m_j+\frac{1}{2}}{2l+1}} \end{pmatrix} \ , \tag{4.149}$$

which directly follows from (4.127) and (4.128).

Sometimes, *Wigner's 3 j symbol* is used to denote the Clebsch-Gordon coefficients:

$$\langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle = (-1)^{j_1 - j_2 + m_j} \sqrt{2j+1} \begin{pmatrix} j_1 j_2, j \\ m_1 m_2 - m \end{pmatrix} \tag{4.150}$$

Applying $J_\pm \equiv J_{1\pm} + J_{2\pm}$ to the state $|j, m_j, j_1, j_2\rangle$ can be used to define a recursion relation to calculate the Clebsch-Gordon coefficients.

$$J_\pm |j_1, j_2, m_1, m_2\rangle = (J_{1\pm} + J_{2\pm}) \sum_{m_1} \sum_{m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2 | j, m_j, j_1, j_2\rangle \tag{4.151}$$

The addition of angular momenta is of particular importance for the specification of the electronic configurations of atoms. In the absence of external fields, the total angular momentum **J** is a constant of motion because of the spherical symmetry. We will later see that there is a spin-orbit interaction which couples orbital and spin angular momentum. If we neglect this interaction which is in particular for light atoms a good approximation, then the total orbital angular momentum **L** and the total spin angular momentum **S** are individually constants of motion, characterized by the quantum numbers L, S,$\mathrm{M_L}$ and $\mathrm{M_S}$ where we used capital letters to denote the total angular momenta.

Hence we have in total a multiplet of (2L +1)(2S+1) degenerate states. In order to specify a level of definite L, S, and J, spectroscopists use the notation $^{\mathrm{2S+1}}\mathrm{L_J}$. For example, $^3\mathrm{P}_2$ denotes a level with L=1, S=1, J=2, and is read “triplet P two”.

4.7 Appendix

4.7.1 Connection between $\mathbf{L}^2$ and the Laplacian

In section 4.4, it has been shown that $\langle x|\mathbf{L}^2|\alpha\rangle$, apart from a factor $1/r^2$, corresponds to the angular part of the Laplacian ∇^2 in spherical coordinates applied to the wave function. In addition, it was mentioned that this connection can also be established by reformulating the kinetic-energy operator in spherical coordinates. In order to see this, we first note that $\mathbf{L}^2$ can be reexpressed as

$$\begin{aligned}
\mathbf{L}^2 &= \sum_{i,j,l,m,k} \varepsilon_{ijk} x_i p_j \varepsilon_{lmk} x_l p_m = \sum_{i,j,l,m,k} \big(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\big) x_i p_j x_l p_m \\
&= \sum_{i,j,l,m,k} \big[\delta_{il}\delta_{jm} x_i (x_l p_j - i\hbar\delta_{jl}) p_m - \delta_{im}\delta_{jl} x_i p_j (p_m x_l + i\hbar\delta_{lm})\big] \\
&= \mathbf{x}^2\mathbf{p}^2 - i\hbar\mathbf{x}\cdot\mathbf{p} - \sum_{i,j,k,l,m} \delta_{im}\delta_{jl} \big[x_i p_m (x_l p_j - i\hbar\delta_{jl}) + i\hbar\delta_{lm} x_i p_j\big] \\
&= \mathbf{x}^2\cdot\mathbf{p}^2 - i\hbar\mathbf{x}\cdot\mathbf{p} - \big(\mathbf{x}\cdot\mathbf{p}\big)^2 + 3i\hbar\mathbf{x}\cdot\mathbf{p} - i\hbar\mathbf{x}\cdot\mathbf{p} \\
&= \mathbf{x}^2\mathbf{p}^2 - \big(\mathbf{x}\cdot\mathbf{p}\big)^2 + i\hbar\mathbf{x}\cdot\mathbf{p}
\end{aligned} \tag{4.152}$$

Furthermore, we have

$$\langle \mathbf{x}|\mathbf{x}\cdot\mathbf{p}|\alpha\rangle = \mathbf{x}\cdot\big(-i\hbar\nabla\langle\mathbf{x}|\alpha\rangle\big) = -i\hbar r\frac{\partial}{\partial r}\langle\mathbf{x}|\alpha\rangle \ , \tag{4.153}$$

$$\big\langle \mathbf{x}|(\mathbf{x}\cdot\mathbf{p})^2|\alpha\big\rangle = -\hbar^2 r\frac{\partial}{\partial r}\big(r\frac{\partial}{\partial r}\langle\mathbf{x}|\alpha\rangle\big) = -\hbar^2\left(r^2\frac{\partial^2}{\partial r^2}\langle\mathbf{x}|\alpha\rangle + r\frac{\partial}{\partial r}\langle\mathbf{x}|\alpha\rangle\right) \ . \tag{4.154}$$

Consequently

$$\big\langle \mathbf{x}|\mathbf{L}^2|\alpha\big\rangle = r^2\big\langle\mathbf{x}|\mathbf{p}^2|\alpha\big\rangle + \hbar^2\left(r^2\frac{\partial^2}{\partial r^2} + 2r\frac{\partial}{\partial r}\langle\mathbf{x}|\alpha\rangle\right) \tag{4.155}$$

The kinetic energy operator $\frac{\mathbf{p}^2}{2m}$ can now be expressed as

$$\frac{1}{2m}\left\langle \mathbf{x}|\mathbf{p}^2|\alpha\right\rangle = -\frac{\hbar^2}{2m}\nabla^2\left\langle \mathbf{x}|\alpha\right\rangle = -\frac{\hbar^2}{2m}\left[\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\left\langle \mathbf{x}|\alpha\right\rangle - \frac{1}{\hbar^2 r^2}\left\langle \mathbf{x}|\mathbf{L}^2|\alpha\right\rangle\right] \tag{4.156}$$

4.7.2 Properties of the spherical harmonics

The spherical harmonics

$$\langle \hat{n}|lm\rangle = Y_{lm}(\vartheta,\varphi) \tag{4.157}$$

are the eigenfunctions of L_3 and $\mathbf{L}^2$ obeying the eigenvalue equations

$$L_3\left|lm\right\rangle = m\hbar\left|lm\right\rangle \tag{4.158}$$

$$\mathbf{L}^2\left|lm\right\rangle = l(l+1)\hbar^2\left|lm\right\rangle \tag{4.159}$$

They can be expressed as

$$Y_{lm}(\vartheta,\varphi) = (-1)^{\frac{1}{2}(m+|m|)}P_{l|m|}(\cos\vartheta)e^{im\varphi}\cdot\sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}\,, \tag{4.160}$$

where the P_{lm} are the associated Legendre functions

$$P_{lm}(x) = (1-x^2)^{\frac{m}{2}}\frac{d^m}{dx^m}P_l(x) = \frac{(-1)^l}{2^l l!}\sin^m\vartheta\frac{d^{l+m}\sin^{2l}\vartheta}{d\cos^{l+m}\vartheta} \tag{4.161}$$

and the P_l are the Legendre polynomials:

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l = \frac{(-1)^l}{2^l l!}\sin^m\vartheta\frac{d^l\sin^{2l}\vartheta}{d\cos^l\vartheta} \qquad x=\cos\vartheta \tag{4.162}$$

There are two recursion relations in order to determine the next-order Legendre polynomials:

$$\begin{aligned}(l+1)P_{l+1}(x) &= (2l+1)xP_l(x) - lP_{l-1}(x)\\ (1-x^2)\frac{dP_l}{dx} &= l(P_{l-1}-xP_l)\end{aligned} \tag{4.163}$$

The lowest order Legendre polynomials are

$$P_0=1 \qquad P_1=x \qquad P_2=\frac{1}{2}(3x^2-1) \qquad P_3=\frac{1}{2}(5x^3-3x) \qquad \dots \tag{4.164}$$

In general, the spherical harmonics are given by

$$Y_{lm}(\vartheta,\varphi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{2l+1}{4\pi}\frac{(l+m)!}{(l-m)!}}e^{im\varphi}\frac{1}{\sin^m\vartheta}\frac{d^{l-m}}{d\cos^{l-m}\vartheta}\sin^{2l}\vartheta \tag{4.165}$$

They are orthogonal to each other and form a complete set:

$$\int_0^\pi d\vartheta\sin\vartheta\int_0^{2\pi}d\varphi\,Y^*_{lm}(\vartheta,\varphi)Y_{l'm'}(\vartheta,\varphi) = \delta_{ll'}\delta_{mm'} \tag{4.166}$$

Further properties of the spherical harmonics in detail:

Completeness:
$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}^*(\vartheta,\varphi)Y_{lm}(\vartheta',\varphi') = \frac{1}{\sin\vartheta}\delta(\vartheta-\vartheta')\delta(\varphi-\varphi') \tag{4.167}$$

Addition theorem:
$$\sum_{m=-l}^{l} Y_{lm}^*(\vartheta,\varphi)Y_{lm}(\vartheta',\varphi') = \frac{2l+1}{4\pi}P_l(\hat{n}\cdot\hat{n}') \tag{4.168}$$

$$Y_{l,-m}(\vartheta,\varphi) = (-1)^m Y_{lm}(\vartheta,\varphi) \tag{4.169}$$

$$Y_{l,0}(\vartheta,\varphi) = \sqrt{\frac{2l+1}{4\pi}}P_l(\cos\vartheta) \tag{4.170}$$

The lowest Y_{lm}'s are

$$Y_{00} = \frac{1}{\sqrt{4\pi}}, \qquad Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\vartheta, \qquad Y_{11} = -\sqrt{\frac{3}{8\pi}}\sin\vartheta e^{i\varphi}$$
$$Y_{20} = \sqrt{\frac{5}{16\pi}}(3\cos^2\vartheta-1),\ Y_{21} = -\sqrt{\frac{15}{8\pi}}\sin\vartheta\cos\vartheta\cdot e^{i\varphi},\ Y_{22} = \sqrt{\frac{15}{32\pi}}\sin^2\vartheta\cdot e^{2i\varphi} \tag{4.171}$$

5 Approximation Methods

Only very few problems – such as the free particle, the harmonic oscillator and the hydrogen atom – can be solved exactly in quantum mechanics. Hence for most realistic problems we have to rely on approximation methods in order to understand the basic physics, even if we are able to solve the problem numerically.

5.1 Time-Independent Perturbation Theory: Non-Degenerate Case

Consider a Hamiltonian H_0. We assume that we have exactly determined the eigenkets $|n^{(0)}\rangle$ and the eigenvalues $E_n^{(0)}$.

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \tag{5.1}$$

Now we assume that an additional term V is added to the Hamiltonian

$$H = H_0 + V \ . \tag{5.2}$$

The extra term is known as the *perturbation*. We are looking for the solutions of the full Hamiltonian problem

$$H|n\rangle = (H_0 + V)|n\rangle = E_n|n\rangle \tag{5.3}$$

With $\Delta_n \equiv E_n - E_n^{(0)}$ we will denote the difference in the eigenenergies between the perturbed system and the unperturbed system. Usually one writes Eq. (5.3) as

$$(H_0 + \lambda V)|n\rangle = E_n|n\rangle \tag{5.4}$$

As we will see below, we will do a power expansion in terms of the perturbation V. The continuous real parameter λ will help us to keep track of the number of times the perturbation enters. Finally we will then set $\lambda = 1$. The transition from $\lambda = 0 \to \lambda = 1$ can be regarded as “switching on the perturbation”.

Let us first focus on a simple two-state problem with a perturbation $V_{12} = V_{21} = V$, i.e. V is assumed to be real. The unperturbed eigenvalues are given by $E_1^{(0)}$ and $E_2^{(0)}$. In the eigenbase of the unperturbed system, the full Hamiltonian is given by

$$H \equiv H_0 + \lambda \mathbf{V} = \begin{pmatrix} E_1^{(0)} & \lambda V \\ \lambda V & E_2^{(0)} \end{pmatrix} \tag{5.5}$$

The eigenenergies of the perturbed system are easily evaluated:

$$E_{1,2} = \tag{5.6}$$

Now we suppose that $\lambda V \ll \left|E_1^{(0)} - E_2^{(0)}\right|$, i. e. the perturbation is relatively small compared to the energy gap. Using a Taylor expansion to approximate the square root,

$$\sqrt{1+\varepsilon} = \tag{5.7}$$

we obtain

$$\begin{aligned} E_1 &= E_1^{(0)} + \frac{\lambda^2 \left|V\right|^2}{E_1^{(0)} - E_2^{(0)}} + \mathcal{O}(\lambda^4) \\ E_2 &= E_2^{(0)} + \frac{\lambda^2 \left|V\right|^2}{E_2^{(0)} - E_1^{(0)}} + \mathcal{O}(\lambda^4) \end{aligned} \tag{5.8}$$

It is important to note here that there is not always a perturbation series for weak perturbations. However, one can say that perturbation theory usually works whenever the state with finite λ does not differ qualitatively from the state with $\lambda = 0$.

Let us now formally develop the perturbation expansion. In principle, the full problem should be written as

$$H^{(\lambda)}|n\rangle_\lambda = (H_0 + \lambda V)|n\rangle_\lambda = E_n^{(\lambda)}|n\rangle_\lambda \ , \tag{5.9}$$

but we will omit the index λ in the following. Both the eigenstates and the eigenenergies of the full problem are expanded in a power series in λ:

$$|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots \tag{5.10}$$

$$\begin{aligned} E_n &= E_n^{(0)} + \Delta_n \\ &= E_n^{(0)} + \lambda\Delta_n^{(1)} + \lambda^2\Delta_n^{(2)} + \ldots \end{aligned} \tag{5.11}$$

In the following it will be convenient use the normalization $\langle n|n^{(0)}\rangle = 1$ instead of the usual one $\langle n|n\rangle = 1$:

$$\begin{aligned} 0 &= \left\langle n^{(0)}|n\right\rangle - 1 \\ &= \underbrace{\langle n^{(0)}|n^{(0)}\rangle}_{=1} + \lambda\langle n^{(0)}|n^{(1)}\rangle + \lambda^2\langle n^{(0)}|n^{(2)}\rangle + \ldots - 1 \\ &= \lambda\langle n^{(0)}|n^{(1)}\rangle + \lambda^2\langle n^{(0)}|n^{(2)}\rangle + \ldots \end{aligned} \tag{5.12}$$

From that it follows that

$$\langle n^{(0)}|n^{(1)}\rangle = \langle n^{(0)}|n^{(2)}\rangle = \ldots = 0 \ , \tag{5.13}$$

since each term in Eq. (5.12) must vanish individually.

We rewrite the full Hamiltonian problem using the power expansions in λ:

$$\begin{aligned} &\big(H_0 + \lambda V\big)\big(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots\big) \\ &\qquad = (E_n^{(0)} + \lambda\Delta_n^{(1)} + \lambda^2\Delta_n^{(2)} + \ldots)\big(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots\big) \end{aligned} \tag{5.14}$$

Now we match the corresponding coefficients in terms of powers of λ. The first terms are given by

$$\mathcal{O}(\lambda^0) : H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \tag{5.15}$$

$$\mathcal{O}(\lambda^1) : \tag{5.16}$$

$$\vdots$$

$$\mathcal{O}(\lambda^k) : H_0|n^{(k)}\rangle + V|n^{(k-1)}\rangle = E_n^{(0)}|n^{(k)}\rangle + \Delta_n^{(1)}|n^{(k-1)}\rangle + \ldots + \Delta_n^{(k)}|n^{(0)}\rangle \tag{5.17}$$

Multiplying Eq. (5.16) by $\langle n^{(0)}|$ from the left, we obtain

$$\underbrace{\langle n^{(0)}|H_0|n^{(1)}\rangle}_{E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle} + \langle n^{(0)}|V|n^{(0)}\rangle = E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle + \Delta_n^{(1)}\underbrace{\langle n^{(0)}|n^{(0)}\rangle}_{=1} , \tag{5.18}$$

where we have used that $|n^{(0)}\rangle$ is an eigenket of H_0. Thus we have

$$\Rightarrow \langle n^{(0)}|V|n^{(0)}\rangle = \Delta_n^{(1)} \equiv V_{nn} , \tag{5.19}$$

Hence we get the energy shifts to first order in λ and therefore in V:

$$\Delta_n = E_n - E_n^{(0)} = \lambda\langle n^{(0)}|V|n^{(0)}\rangle = \lambda V_{nn} . \tag{5.20}$$

Analogously, by multiplying Eq. (5.17) by $\langle n^{(0)}|$ from the left and using Eq. (5.13) we get the result for all other orders:

$$\Delta_n^{(k)} = \langle n^{(0)}|V|n^{(k-1)}\rangle \tag{5.21}$$

This means that once the change in the state to a certain order is known, the energy change can be found to the next order.

In order to get the perturbed wave function, we expand $|n^{(k)}\rangle$ in terms of eigenstates $|m^{(0)}\rangle$ of H_0:

$$|n^{(k)}\rangle = \sum_m |m^{(0)}\rangle\langle m^{(0)}|n^{(k)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle\langle m^{(0)}|n^{(k)}\rangle \tag{5.22}$$

The term with $m = n$ is not included in the sum because of Eq. (5.13). Multiplying (5.17) from the left with $\langle m^{(0)}|$, we obtain for $m \neq n$

$$\underbrace{\langle m^{(0)}|H_0|n^{(k)}\rangle}_{E_m^{(0)}\langle m^{(0)}|n^{(k)}\rangle} + \langle m^{(0)}|V|n^{(k-1)}\rangle =$$
$$= E_n^{(0)}\langle m^{(0)}|n^{(k)}\rangle + \Delta_n^{(1)}\langle m^{(0)}|n^{(k-1)}\rangle + \ldots + \Delta_n^{(k-1)}\langle m^{(0)}|n^{(1)}\rangle \tag{5.23}$$

Note that in the nondegenerate case that we assumed here, i.e. $E_m^{(0)} \neq E_n^{(0)}$ for $m \neq n$, we have $\langle m^{(0)}|n^{(0)}\rangle = \delta_{mn}$. For $k = 1$, we get

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{1}{E_n^{(0)} - E_m^{(0)}}\langle m^{(0)}|V|n^{(0)}\rangle \tag{5.24}$$

From that we get the first-order approximation for the eigenket $|n\rangle$ of the perturbed system, using (5.10) and (5.22):

$$|n\rangle = |n^{(0)}\rangle + \lambda \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \tag{5.25}$$

The corresponding eigenenergy to first order has already been evaluated in Eq. (5.20):

$$E_n = E_n^{(0)} + \lambda \langle n^{(0)}|V|n^{(0)}\rangle \tag{5.26}$$

to first order.

We get the second-order term of the energy shift by inserting our result for the wave function to first order in the general formula for energy shifts Eq. (5.21):

$$\begin{aligned}\Delta_n^{(2)} &= \langle n^{(0)}|V|n^{(1)}\rangle \\ &= \sum_{m\neq n} \frac{\langle n^{(0)}|V|m^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \langle m^{(0)}|V|n^{(0)}\rangle = \sum_{m\neq n} \frac{\left|\langle n^{(0)}|V|m^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}} = \sum_{m\neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}\end{aligned} \tag{5.27}$$

In order to see the expansion, we write down the whole expression

$$\begin{aligned}E_n &= E_n^{(0)} + \lambda \Delta_n^{(1)} + \lambda^2 \Delta_n^{(2)} + \ldots \\ &= E_n^{(0)} + \lambda V_{nn} + \lambda^2 \sum_{m\neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} + \ldots\end{aligned} \tag{5.28}$$

This expresion should be compared with Eq. (5.8): not surprisingly, the second-order Taylor expansion of the exact eigenvalues of the (2×2) problem coincides with (5.28).

In fact, perturbation theory is usually used just up to second order for the energies. For the sake of completeness, we add the second-order expression for the state kets:

$$\begin{aligned}|n\rangle =& |n^{(0)}\rangle + \lambda \sum_{m\neq n} |m^{(0)}\rangle \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \\ &+ \lambda^2 \left(\sum_{m\neq n} \sum_{l\neq n} |m^{(0)}\rangle \frac{V_{ml} V_{ln}}{\left(E_n^{(0)} - E_m^{(0)}\right)\left(E_n^{(0)} - E_l^{(0)}\right)} - \sum_{m\neq n} |m^{(0)}\rangle \frac{V_{mn} V_{nn}}{\left(E_n^{(0)} - E_m^{(0)}\right)^2} \right) .\end{aligned} \tag{5.29}$$

The second-order perturbation theory expression (5.28) can be found over and over in quantum mechanical derivations and applications in the literature. Let us list some important general properties of this expression:

Remarks For the first-order energy eigenstates, only the eigenstates and eigenenergies of the unperturbed system are needed.

The second order energy shifts tend to ‘repel’ the energy levels E_n and E_m: the lower one decreases by the same amount as the higher one increases.

For the ground state, the second order energy shifts are always negative (which is a direct consequence of the preceding statement).

5.1.1 Harmonic Oscillator

As an elementary example, we consider the harmonic oscillator. To the unperturbed Hamiltonian

$$H_0 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \tag{5.30}$$

we add a harmonic perturbation

$$V(x) = \frac{1}{2}\varepsilon m\omega^2 x^2 \tag{5.31}$$

with $|\varepsilon| \ll 1$. The problem can easily be solved exactly by the substitution $\omega \mapsto \sqrt{1+\varepsilon}\,\omega$. Applying perturbation theory, we obtain

$$\begin{aligned} |0\rangle &= |0^{(0)}\rangle + \sum_{k\neq 0} |k^{(0)}\rangle \frac{V_{k0}}{E_0^{(0)} - E_k^{(0)}} + \dots \\ \Delta_0 &= V_{00} + \sum_{k\neq 0} \frac{|V_{k0}|^2}{E_0^{(0)} - E_k^{(0)}} \end{aligned} \tag{5.32}$$

The relevant matrix elements are

$$V_{00} = \frac{\varepsilon}{2}m\omega^2 \langle 0^{(0)}|x^2|0^{(0)}\rangle = \frac{\varepsilon}{4}\hbar\omega \tag{5.33}$$

$$V_{20} = \frac{\varepsilon}{2}m\omega^2 \langle 2^{(0)}|x^2|0^{(0)}\rangle = \frac{\varepsilon}{2\sqrt{2}}\hbar\omega \tag{5.34}$$

All other matrix elements vanish. The denominators of the non-vanishing terms are $-2\hbar\omega$; thus we obtain

$$|0\rangle = |0^{(0)}\rangle - \frac{\varepsilon}{4\sqrt{2}}|2^{(0)}\rangle + \mathcal{O}(\varepsilon^2) \tag{5.35}$$

$$\Delta_0 = E_0 - E_0^{(0)} = \hbar\omega\big(\frac{\varepsilon}{4} - \frac{\varepsilon^2}{16} + \mathcal{O}(\varepsilon^3)\big) \tag{5.36}$$

Let us compare this expansion with the exact solution. We perform a Taylor expansion of $\sqrt{1+\varepsilon}$ in terms of ε:

$$\begin{aligned} \sqrt{1+\varepsilon} &= \\ \frac{1}{2}\hbar\omega\sqrt{1+\varepsilon} &= \end{aligned} \tag{5.37}$$

This agrees completely with the perturbation result.

5.2 Degenerate Perturbation Theory

Non-degenerate perturbation theory breaks down, if there is degeneracy, i.e. $E_n^{(0)} = E_m^{(0)}$ for $n \neq m$ but $V_{nm} \neq 0$. Then the corresponding term in the sum becomes singular.

If there is a degeneracy, then we are free to choose any suitable base set of unperturbed kets. In fact we will exploit this freedom in such a way that the matrix elements V_{nm} will vanish.

Let $D(E_n^k)$ be the k-dimensional subspace to the eigenvalue $E_n^{(0)}$. Thus, we get

$$\begin{aligned} H_0|l_i^{(0)}\rangle &= E_n^{(0)}|l_i^{(0)}\rangle \\ &= E_D^{(0)}|l_i^{(0)}\rangle \end{aligned} \tag{5.38}$$

for $|l_i^{(0)}\rangle \in D(E_n^k)$. Then choose a set of k degenerate orthonormal states

$$|m_j^{(0)}\rangle = \sum_{i=1}^{k} c_{ji}|l_i\rangle \tag{5.39}$$

such that

$$\langle m_j^{(0)}|V|m_{j'}^{(0)}\rangle = 0 \text{ for } j \neq j' \ . \tag{5.40}$$

This corresponds to the eigenvalue equation in the subspace $D(E_n^k)$:

$$\begin{pmatrix} V_{11} & V_{21} & \dots \\ V_{21} & V_{22} & \\ \vdots & & \ddots \end{pmatrix} \begin{pmatrix} \langle l_1^{(0)}|m_j^{(0)}\rangle \\ \langle l_2^{(0)}|m_j^{(0)}\rangle \\ \vdots \end{pmatrix} = \Delta_j^{(1)} \begin{pmatrix} \langle l_1^{(0)}|m_j^{(0)}\rangle \\ \langle l_2^{(0)}|m_j^{(0)}\rangle \\ \vdots \end{pmatrix} \tag{5.41}$$

The eigenvalues can be found by solving the secular equation

$$\det(V - \Delta^{(1)}\mathbf{1}) \overset{!}{=} 0 \ . \tag{5.42}$$

This means that the first-order energy shifts are simply given by

$$\Delta_j^{(1)} = \langle m_j^{(0)}|V|m_j^{(0)}\rangle \ . \tag{5.43}$$

The eigenket to first order is now given by

$$|m_j\rangle = |m_j^{(0)}\rangle + \lambda \sum_{E_i^{(0)} \neq E_D^{(0)}} \frac{\langle i^{(0)}|V|m_j^{(0)}\rangle}{E_D^{(0)} - E_i^{(0)}} |i^{(0)}\rangle \tag{5.44}$$

Similarly, the second-order energy of the perturbed system is given by

$$E_j = E_j^{(0)} + \lambda\langle m_j^{(0)}|V|m_j^{(0)}\rangle + \lambda^2 \sum_{E_i^{(0)} \neq E_D^{(0)}} \frac{|\langle i^{(0)}|V|m_j^{(0)}\rangle|^2}{E_D^{(0)} - E_i^{(0)}} \tag{5.45}$$

Recipe First, identify the degenerate subspace(s). Then diagonalize the perturbation matrix. The first-order shifts are given by the roots of the secular equation. For higher orders, apply the formulas for non-degenerate perturbation theory, but exclude the degenerate subspace in the summations.

5.2.1 Linear Stark Effect

If we neglect the spin dependence, the bound state energy of the hydrogen atom only depends on the principal quantum number n. Thus for a given value of n, all states

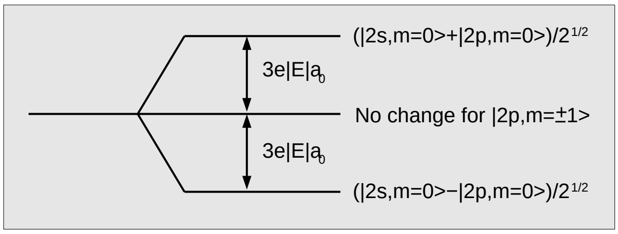


Figure 5.1: Schematic energy-level diagram for the linear Stark effect.

with $0 \leq l < n$ become degenerate. If we apply an uniform electric field in the z-direction, this degeneracy becomes lifted. In order to see this, we have to diagonalize the perturbation which is given by

$$V = -ez\,|\mathbf{E}|\,. \tag{5.46}$$

First, we note that $[V, L_z] \propto [z, L_z] = 0$. Thus, the matrix elements of λV between states with different eigenvalues of L_z vanish. Furthermore, the eigenfunctions of the hydrogen atom are also parity eigenfunctions, i. e. $|\Psi_{nlm}(\mathbf{r})|^2 = |\Psi_{nlm}(-\mathbf{r})|^2$. Hence only states with opposite parity are coupled by the perturbation. Since $z = r\cos\theta$ corresponds to a $l = 1$ function, we have the following selection rule

$$\langle n', l', m'|z|n, l, m\rangle = 0 \qquad \text{unless} \qquad \begin{cases} l' & = l \pm 1 \\ m' & = m \end{cases} . \tag{5.47}$$

Let us take a look at the $n = 2$ energy levels. There is one $l = 0$ (the $2s$ state) and three $l = 1$ states (the $2p$ states). Thus, all entries of the perturbation matrix vanish except for $V_{12} = \langle 2p, m = 0|V|2s\rangle$ and $V_{21} = \langle 2s|V|2p, m = 0\rangle$.

Thus, effectively we have to deal with a 2×2 perturbation matrix. The only two non-vanishing elements of this matrix, which is proportional to the σ_x matrix, are

$$V_{12} = \langle 2s|V|2p, m = 0\rangle = V_{21} = \langle 2p, m = 0|V|2s\rangle = 3a_0\,|\mathbf{E}| \tag{5.48}$$

The energy shifts and the zeroth-order kets that diagonalize V are

$$\begin{aligned} \Delta_\pm^{(1)} &= \pm 3ea_0\,|\mathbf{E}| \\ |\pm\rangle &= \tfrac{1}{\sqrt{2}}\big(|2s, m = 0\rangle \pm |2p, m = 0\rangle\big) \end{aligned} \tag{5.49}$$

The degenerate $2s$ and $2p$ states are now split in three: $|+\rangle$, no change for $|2p, m = \pm 1\rangle$, and $|-\rangle$ (see fig. 5.1).

For non-degenerate states, e. g. the $1s$ state, there is no linear, only a quadratic Stark effect, since the $|1s\rangle$ state has no permanent dipole.

Remark The formalism of degenerate perturbation theory is still useful, if the states are *almost* degenerate.

5.2.2 Spin-Orbit Interaction and Fine Structure

We will now consider Hydrogen-like atoms, i. e. atoms with one valence electron outside a closed shell. The charge of the shell electrons is distributed spherical symmetrical.

Now we have a central potential $V_c = e\Phi(r)$, that is *not* a Coulomb potential anymore. This is due to the interaction between the outer electron and the electrons in the inner shells. As one consequence, the degeneracy of states with same quantum number n but different l is lifted. Higher l states lie higher in energy for a given n since the higher l states are more strongly interacting with the repulsive electron cloud. For lighter atoms, the difference between $|n,l\rangle$ and $|n,l'\rangle$ states is of the order of 0.1 eV where the states with higher l have the higher energy.

Additionally, we have the so-called spin-orbit ($\mathbf{L}\cdot\mathbf{S}$) interaction which is due to relativistic effects and gives rise to the *fine structure*.

Qualitative Derivation

Due to the shell electrons, we have an electric field

$$\mathbf{E} = -\frac{1}{e}\nabla V_c(r) = -\frac{1}{e}\frac{\mathbf{x}}{r}\frac{dV_c}{dr}. \tag{5.50}$$

For a moving charge, there is an effective magnetic field (this is a relativistic effect) which is given by

$$\mathbf{B}_{\text{eff}} = -\frac{1}{c}\mathbf{v}\times\mathbf{E} \tag{5.51}$$

The magnetic moment of the electron is given by

$$\mu = \frac{e}{m_e c}\mathbf{S} \tag{5.52}$$

Now we can express the spin-orbit interaction.

$$-\mu\cdot\mathbf{B}_{\textbf{eff}} = \frac{e\mathbf{S}}{m_e c}\big(\frac{1}{c}\mathbf{v}\times\mathbf{E}\big) \tag{5.53}$$

There is another relativistic effect which can be explained by spin precession called *Thomas precession*. It amounts to $-1/2\cdot$(5.53). In total, we obtain

$$\begin{aligned} V_{LS} &= \frac{e\mathbf{S}}{2m_e c}\big(\tfrac{1}{c}\mathbf{v}\times\mathbf{E}\big) = \frac{e\mathbf{S}}{2m_e c}\left(\frac{\mathbf{p}}{m_e c}\times\Big(-\frac{1}{e}\frac{\mathbf{x}}{r}\frac{dV_c}{dr}\Big)\right) \\ &= \frac{1}{2m_e^2c^2}\frac{1}{r}\frac{dV_c}{dr}\mathbf{L}\cdot\mathbf{S} \end{aligned} \tag{5.54}$$

The unperturbed Hamiltonian is given by

$$H_0 = \frac{\mathbf{p}^2}{2m_e} + V_c(r) \tag{5.55}$$

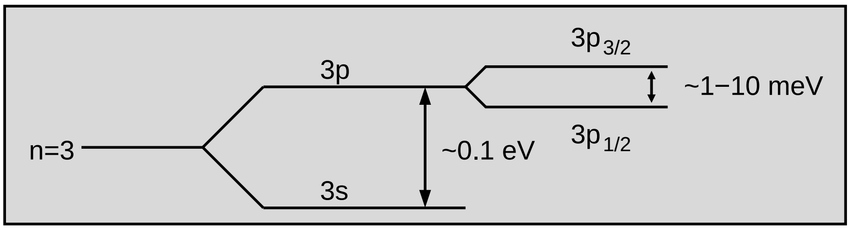


Figure 5.2: Diagram of the fine structure of Sodium

$\mathbf{L}\cdot\mathbf{S}$ commutes with $\mathbf{J}^2$ and J_3, but not with L_3 and S_3. Hence we use the eigenfunctions of $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$, J_3[1].

$$\begin{aligned}\Psi_{nlm} &= R_{nl}(r)Y_{lm}^{j=l\pm 1/2}(\vartheta,\varphi)\\ &= R_{nl}(r)\left(\pm\sqrt{\frac{l\pm m+1/2}{2l+1}}Y_{l,m_j-1/2}(\vartheta,\varphi)\chi_+ + \sqrt{\frac{l\mp m+1/2}{2l+1}}Y_{l,m_j+1/2}(\vartheta,\varphi)\chi_-\right)\end{aligned} \tag{5.56}$$

where we have used (4.127) and (4.128) in order to express the eigenfunctions of the total angular momentum.

Use $\mathbf{L}\cdot\mathbf{S} = \frac{1}{2}\big(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2\big)$ in order to obtain the eigenvalues:

$$= \begin{cases} & j = l + \frac{1}{2} \\ & j = l - \frac{1}{2}\end{cases} \tag{5.57}$$

Now we can calculate the first-order shifts.

$$\Delta_{nlj} = \frac{1}{2m_e^2c^2}\left\langle \frac{1}{r}\frac{dV_c}{dr}\right\rangle \frac{\hbar^2}{2}\begin{cases} l & j = l+\frac{1}{2} \\ -(l+1) & j = l - \frac{1}{2}\end{cases} \tag{5.58}$$

with

$$\left\langle \frac{1}{r}\frac{dV_c}{dr}\right\rangle = \int_0^\infty R_{nl}(r)\frac{1}{r}\frac{dV_c}{dr}R_{nl}(r)r^2\,dr \tag{5.59}$$

Example: Sodium atom (see Fig. 5.2).

5.2.3 van-der-Waals Interaction

We consider the long-range interaction between two hydrogen atoms in their ground state using perturbation theory. As we will see, it is easy to derive the qualitative aspects of the van der Waals interaction using second-order perturbation theory

The unperturbed Hamiltonian is given by the Hamiltonian of two non-interacting hydrogen atoms:

$$H_0 = -\frac{\hbar^2}{2m}\big(\nabla_1^2 + \nabla_2^2\big) - \big(\frac{e^2}{r_1} + \frac{e^2}{r_2}\big) \tag{5.60}$$

---

[1] $\mathbf{J} = \mathbf{S}\otimes\mathbf{1} + \mathbf{1}\otimes\mathbf{L}$

The perturbation term corresponds to the electrostatic interaction between the two hydrogen atoms which are assumed to be aligned along the z-axis. According to Fig. (5.3), this interaction is given by

$$V = \frac{e^2}{r} + \tag{5.61}$$

The lowest-energy solution for H_0 is a combination of the lowest energy solutions for the isolated hydrogen atoms.

$$\Psi_0^{(0)}(\mathbf{r}_1, \mathbf{r}_2) = \Psi_{100}^{(0)}(\mathbf{r}_1)\Psi_{100}^{(0)}(\mathbf{r}_2) \tag{5.62}$$

For large $r \gg a_0$ we expand the perturbation in terms of $r_i/\mathbf{r}$:

$$V = \frac{e^2}{r^3}\big(x_1x_2 + y_1y_2 - 2z_1z_2\big) + \mathcal{O}(r^{-4}) \tag{5.63}$$

The first-order shifts vanish because of symmetry reasons (x_i is odd while $(\Psi_{100}^{(0)}(\mathbf{r}_j))^2$ is even). The second-order perturbation does not vanish:

$$\Delta_0^{(2)} = \frac{e^4}{r^6}\sum_{k\neq 0}\frac{\left|\langle k^{(0)}|x_1x_2 + y_1y_2 - 2z_1z_2|0^{(0)}\rangle\right|^2}{E_0^{(0)} - E_k^{(0)}} \tag{5.64}$$

The result is a long range interaction $\propto r^{-6}$. This the ground-state energy is always lowered by second-order terms, it corresponds to an attraction, the so-called *van der Waals attraction*.

5.3 Variational Methods

There is another powerful tool in order to approximately determine eigenenergies, the variational method. In contrast to perturbation methods it does not require the exact solution of an unperturbed Hamiltonian that does not significantly differ from the full Hamiltonian. It is rather based on the fact that the expectation value of H in any state is always greater than or equal to the ground state energy E_0, i.e.

$$\bar{H} \equiv \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_0 \tag{5.65}$$

Proof

Figure 5.3: Two hydrogen atoms

Let us denote the energy eigenkets of H by $|n\rangle$. We can write H in terms of its eigenenergies (i. e. we diagonalize H):

$$H = \sum_n |n\rangle E_n \langle n| \tag{5.66}$$

Thus, we get

$$\begin{aligned}\langle\Psi|H|\Psi\rangle &= \sum_n \langle\Psi|n\rangle E_n \langle n|\Psi\rangle \\ &\geq \sum_n \langle\Psi|n\rangle E_0 \langle n|\Psi\rangle = E_0\langle\Psi|\Psi\rangle\end{aligned} \tag{5.67}$$

Equation (5.65) is a true inequality, if $|\Psi\rangle$ has a component along any state that is not the ground state. □

Equation (5.65) corresponds to the so-called *Rayleigh-Ritz variational principle* which states that $\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle$ will always be an upper bound for the true ground state energy.

In practice, one characterizes a trial ket $|\Psi\rangle$ by one or more parameters $\lambda_1, \lambda_2, \ldots$ and minimizes $\bar{H}$ with respect to the λ_i, i.e.

$$\frac{\partial \bar{H}}{\partial \lambda_1} = 0, \ \frac{\partial \bar{H}}{\partial \lambda_2} = 0, \ \ldots \tag{5.68}$$

We can get closer and closer to the true ground state with more and more suitable test functions Ψ.

Example

Particle in a Box. The potential is

$$V = \begin{cases} 0 & |x| < a \\ \infty & |x| \geq a \end{cases} \tag{5.69}$$

We can solve this analytically and get

$$\begin{aligned}\Psi_0(x) &= \frac{1}{\sqrt{a}} \cos\frac{\pi x}{2a} \\ E_0 &= \frac{\hbar^2\pi^2}{8ma^2}\end{aligned} \tag{5.70}$$

However, let us assume that we did not know the correct solution. Our trial function for the ground state should be symmetric with respect to $x = 0$, furthermore, it should vanish at $x = \pm a$. As our trial function satifying these conditions we choose

$$\Psi_\lambda(x) = |a|^\lambda - |x|^\lambda \tag{5.71}$$

$$\Rightarrow \ \bar{H} = \frac{\langle\Psi_\lambda|H|\Psi_\lambda\rangle}{\langle\Psi_\lambda|\Psi_\lambda\rangle} = \tag{5.72}$$

Minimization with respect to λ yields:

$$\left.\frac{\partial \bar{H}}{\partial \lambda}\right|_{\lambda=\lambda_{\min}} = 0 \;\Rightarrow\; \lambda_{\min} = \frac{1+\sqrt{6}}{2} \approx 1.72 \tag{5.73}$$

$$\bar{H}_{\min} = \frac{5+2\sqrt{6}}{\pi^2} E_0 \approx 1.00298 \cdot E_0 \tag{5.74}$$

Thus we have obtained a surprisingly good estimate of the true ground-state energy.

5.4 Time-Dependent Perturbation Theory

So far we have mainly addressed time-independent Hamiltonians. We will now consider situations with time-dependent potentials, i.e. the Hamiltonian can be expressed as

$$H = H_0 + V(t) \qquad\qquad \frac{\partial H_0}{\partial t} = 0 \tag{5.75}$$

where H_0 does not contain time explicitly. Assume that $V(t)$ is switched on at $t = 0$, $V = \theta(t)V(t)$. Thus the Schrödinger equations for $t < 0$ differs from the one for $t > 0$:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\alpha^{(0)}(t)\rangle &= H_0|\alpha^{(0)}(t)\rangle\ , \qquad t<0\ ,\\ i\hbar\frac{\partial}{\partial t}|\alpha(t)\rangle &= \big(H_0+V(t)\big)|\alpha(t)\rangle\ , \quad t\geq 0\ , \end{aligned} \tag{5.76}$$

subject to the boundary condition $|\alpha^{(0)}(t)\rangle = |\alpha(t)\rangle$ for $t \leq 0$.

We assume that $V(t)$ is a small perturbation, so that the dynamics, i. e. the time-evolution, is mainly determined by H_0. We will take this dependence explicitly:

$$|\alpha(t)\rangle = e^{-i/\hbar H_0 t}|\alpha(t)\rangle_I \tag{5.77}$$

Here the subscript I indicates that the ket is in the *interaction representation* whose time-dependence is solely due to the time-dependent potential $V(t)$, as we see from inserting (5.77) into the Schrödinger equation (5.76):

$$i\hbar\frac{\partial}{\partial t}e^{-iH_0t/\hbar}|\alpha(t)\rangle_I = \big(H_0+V(t)\big)e^{-iH_0t/\hbar}|\alpha(t)\rangle_I$$

$$\Rightarrow$$

$$\Rightarrow i\hbar\frac{\partial}{\partial t}|\alpha(t)\rangle_I = e^{iH_0t/\hbar}V(t)e^{-iH_0t/\hbar}|\alpha(t)\rangle_I \equiv V_I(t)|\alpha(t)\rangle_I \tag{5.78}$$

We will now integrate the last equation,

$$|\alpha(t)\rangle_I = |\alpha^{(0)}(0)\rangle_I + \frac{1}{i\hbar}\int_0^t V_I(t')|\alpha(t')\rangle_I\,dt' \tag{5.79}$$

Now expand $|\alpha(t)\rangle_I$ in "powers" of V_I by iteration. The first-order iteration is given by

$$\begin{aligned} |\alpha(t)\rangle_I &\approx |\alpha^{(0)}(0)\rangle_I + |\alpha^{(1)}(t)\rangle_I \\ &= |\alpha^{(0)}(0)\rangle_I + \frac{1}{i\hbar}\int_0^t dt'\, V_I(t')|\alpha^{(0)}(t=0)\rangle_I \end{aligned} \tag{5.80}$$

For the second-order iteration, we substitute our result back in (5.79) and get

$$|\alpha^{(2)}(t)\rangle_I = \left(\frac{1}{i\hbar}\right)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t')V_I(t'')|\alpha^{(0)}(t=0)\rangle_I \tag{5.81}$$

Similarly, to kth order we get

$$|\alpha^{(k)}(t)\rangle_I = \left(\frac{1}{i\hbar}\right)^k \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} \cdots \int_0^{t_{k-1}} dt_k V_I(t_1)V_I(t_2)\cdots V_I(t_k)|\alpha^{(0)}(t=0)\rangle_I \tag{5.82}$$

Now let us assume that at time $t = 0$ the system is in an eigenstate $|i\rangle$ of H_0 with energy E_i. i.e., $|i\rangle = |i(t=0)\rangle = |i(t=0)\rangle_I$. We will expand the eigenket $|i(t)\rangle$ which is *not* in the interaction picture.

$$\begin{aligned} |i(t)\rangle &= \sum_n |n\rangle\langle n|i(t)\rangle = \sum_n |n\rangle\langle n|e^{-iH_0t/\hbar}|i(t)\rangle_I = \\ &= \sum_n |n\rangle e^{-iE_nt/\hbar}c_n(t) = \sum_n |n\rangle e^{-iE_nt/\hbar}\sum_k c_n^{(k)}(t)\ , \end{aligned} \tag{5.83}$$

where the last sum corresponds to the expansion in the perturbation series. We use (5.80)–(5.82) to determine the coefficients $c_n^{(k)}(t)$ that give the probabilities to find the system in a state $n \neq i$ at later times. We immediately get $c_n^{(0)}(t) = \delta_{ni}$.

$$\begin{aligned} c_n^{(0)}(t) &= \delta_{ni} \\ c_n^{(1)}(t) &= \frac{1}{i\hbar}\int_0^t dt'\langle n|V_I(t')|i\rangle = \frac{1}{i\hbar}\int_0^t dt'\langle n|e^{iH_0t'/\hbar}V(t')e^{-iH_0t'/\hbar}|i\rangle \\ &= \frac{1}{i\hbar}\int_0^t dt' e^{i\omega_{ni}t'}V_{ni}(t') \end{aligned} \tag{5.84}$$

with

$$\omega_{ni} = \frac{(E_n - E_i)}{\hbar}. \tag{5.85}$$

The second-order term is obtained analogously:

$$c_n^{(2)}(t) = \left(\frac{1}{i\hbar}\right)^2 \sum_m \int_0^t dt' \int_0^{t'} dt''\, e^{i\omega_{nm}t'}V_{nm}(t')e^{i\omega_{mi}t''}V_{mi}(t'') \tag{5.86}$$

The transition probability for the system to go from state $|i\rangle$ to $|n\rangle$ for $n \neq i$ is given by

$$P_{i\to n}(t) = \left|c_n^{(1)}(t) + c_n^{(2)}(t) + \ldots\right|^2 \tag{5.87}$$

Example

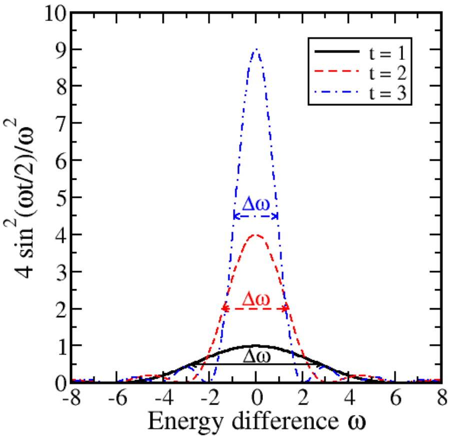


Figure 5.4: Transition probability for various times after a constant perturbation has been switched on as a function of the energy difference assuming that V_{ni} is only weakly dependent on the energy.

Constant Perturbation.

$$V(t) = \begin{cases} 0 & ;\ t < 0 \\ V & ;\ t > 0 \end{cases} \tag{5.88}$$

Assume the initial state at $t = 0$ is $|i\rangle$.

$$\Rightarrow c_n^{(0)} = \delta_{ni}$$

$$c_n^{(1)}(t) = \frac{1}{i\hbar} V_{ni} \int_0^t e^{i\omega_{ni}t'}\,dt' = \frac{V_{ni}}{E_n - E_i}\left(1 - e^{i\omega_{ni}t}\right) \tag{5.89}$$

The transition probability to first-order is

$$P_{i\to n}(t) \approx |c_n^{(1)}(t)|^2 =$$

$$= \frac{4\,|V_{ni}|^2}{|E_n - E_i|^2} \sin^2\left[\frac{E_n - E_i}{2\hbar} t\right] \tag{5.90}$$

For small times, this term grows proportional to t^2.

Assume that there are many states, i.e. that there is a quasi-continuum of energy eigenstates. Define

$$\omega \equiv \frac{E_n - E_i}{\hbar} = \omega_{ni} \tag{5.91}$$

Using this notation, we get $P_\omega(t) \propto 4\sin^2(\omega t/2)/\omega^2$ which is plotted in Fig. (5.4).

For fixed t, the value of $4sin^2(\omega t/2)/\omega^2$ at $\omega = 0$ is t^2 (follows from $sin(\omega t/2) \approx (\omega t/2)$ for small ω. Hence the height of the middle peak as a function of $\omega = (E_n - E_i)/\hbar)$ is proportional to t^2. The first zero of $4sin^2(\omega t/2)/\omega^2$ is at $(\omega t/2) = \pi$, so that the width of the peak is proportional to $1/t$ and the total area of the peak is consequently proportional to t. This means that for large t, assuming that V_{ni} is only weakly dependent on the energy, $|c_n^{(1)}(t)|$ is appreciable only for final states with $\omega \leq (2\pi)/t$:

$$E_n - E_i \leq \frac{2\pi\hbar}{t} \;\Rightarrow\; \Delta t \, \Delta E \sim 2\pi\hbar \, , \tag{5.92}$$

where Δt is the period of time that the perturbation has acted. This is the likely spread in energy as a function of t. For small t, there is a certain amount of energy *non*conservation, while at longer times energy conservation is approximately fulfilled.

It is important to note that switching on a constant perturbation corresponds to an external modification of the system so that energy conservation does not need to be obeyed any more. And indeed, when the constant perturbation is turned on sharply, in first order transitions to any state with $V_{ni} \neq 0$ are possible, but the probability is most significant for transitions that conserve the energy to within $\Delta E = 2\pi\hbar/t$. The total probability for transitions fulfilling this kind of uncertainty principle grows proportional to t.

A quasi-continuum of states is described by a density of states (DOS) $\rho(E)$ where $\rho(E)\, dE$ is the number of states within the intervall $(E, E + dE)$.

The transition probabilty for the states is

$$P(t) = \sum_n |c_n^{(1)}(t)|^2 \, , \tag{5.93}$$

where the sum is taken over final states with $E_n \approx E_i$. This leads to

$$\int dE_n \, \rho(E_n) \left| c_n^{(1)}(t) \right|^2 = 4 \int \sin^2 \left[\frac{E_n - E_i}{2\hbar} t \right] \frac{|V_{ni}|^2}{|E_n - E_i|^2} \rho(E_n) \, dE_n \tag{5.94}$$

Use $\lim\limits_{\alpha \to \infty} \frac{1}{\pi} \frac{\sin^2 \alpha x}{\alpha x^2} = \delta(x)$ to get the asymptotic transition probability:

$$\lim_{t \to \infty} P(t) =$$

$$= \frac{2\pi}{\hbar} t \, |V_{ni}|^2 \, \rho(E_n) \big|_{E_n = E_i} \tag{5.95}$$

Thus the transition probability for large t becomes proportional to t which means that the transition rate $\Gamma \equiv \frac{dP}{dt}$ is constant in time for large t. Γ is given by

$$\Gamma = \frac{dP}{dt} = \frac{2\pi}{\hbar} \, |V_{ni}|^2 \, \rho(E_n) \big|_{E_n = E_i} \tag{5.96}$$

This formula which is of great practical importance is called *Fermi's Golden Rule.* It is often also written as

$$\Gamma = \frac{2\pi}{\hbar} \, |V_{ni}|^2 \, \delta(E_n - E_i) \tag{5.97}$$

where (5.97) has to be integrated with $\int dE_n\,\rho(E_n)$.

The second-order term is

$$
\begin{aligned}
c_n^{(2)}(t) &= \left(\frac{1}{i\hbar}\right)^2 \sum_m V_{nm}V_{mi} \int_0^t dt' e^{i\omega_{nm}t'} \int_0^{t'} dt'' e^{i\omega_{mi}t''} \\
&= \frac{i}{\hbar}\sum_m \frac{V_{nm}V_{mi}}{E_m - E_i} \int_0^t \left(e^{i\omega_{mi}t'} - e^{i\omega_{nm}t'}\right) dt'
\end{aligned}
\tag{5.98}
$$

This is a rapidly oscillating function for $\omega_{mi} \neq \omega_{nm}$. Hence only terms with $E_n \approx E_i$ contribute significantly to the transition probability.

$$
\Gamma = \frac{2\pi}{\hbar}\left|V_{ni} + \sum_m \frac{V_{nm}V_{mi}}{E_i - E_m}\right|^2 \rho(E_n)\big|_{E_n=E_i}
\tag{5.99}
$$

The first-order terms correspond to a *real transitions* that are energy conserving whereas the second-order term involves so-called *virtual transitions* $|i\rangle \to |m\rangle$ and $|m\rangle \to |n\rangle$ that are *not energy conserving*.

We will apply this to harmonic perturbation and stimulated emission and absorption.

Example

Harmonic Perturbation. Consider a harmonic perturbation $V(t) = Ve^{i\omega t} + V^\dagger e^{-i\omega t}$. First-order perturbation theory gives us

$$
\begin{aligned}
c_n^{(1)} &= \frac{1}{i\hbar}\int_0^t \left(V_{ni}e^{i\omega t'} + V_{ni}^\dagger e^{-i\omega t'}\right) e^{i\omega_{ni}t'}\,dt' \\
&=
\end{aligned}
\tag{5.100}
$$

This is similar to the constant perturbation case with the substitution $\omega_{ni} \to \omega_{ni} \pm \omega$. In analogy with the golden rule, we get

$$
\Gamma = \frac{2\pi}{\hbar}\left(|V_{ni}|^2\rho(E_n)\big|_{E_n=E_i-\hbar\omega} + |V_{ni}^\dagger|^2\rho(E_n)\big|_{E_n=E_i+\hbar\omega}\right)
\tag{5.101}
$$

or

$$
\Gamma = \frac{2\pi}{\hbar}\left(|V_{ni}|^2\delta(E_n - E_i + \hbar\omega) + |V_{ni}^\dagger|^2\delta(E_n - E_i - \hbar\omega)\right)
\tag{5.102}
$$

The left term corresponds to stimulated emission, the right one corresponds to absorption (see Fig. 5.5).

Note that $|V_{ni}|^2 = |V_{in}^\dagger|^2$ because $V_{ni}^\dagger = V_{in}^*$ or equivalently $\langle i|V^\dagger|n\rangle = \langle n|V|i\rangle^*$. Thus, the emission rate from some state $|i\rangle$ to some state $|n\rangle$ is given by

$$
\frac{\text{emission rate for } |i\rangle \to |n\rangle}{\text{density of final states for } E_n} = \frac{\text{absorption rate for } |n\rangle \to |i\rangle}{\text{density of final states for } E_i}
\tag{5.103}
$$

This relation is known as the principle of *detailed balance*.

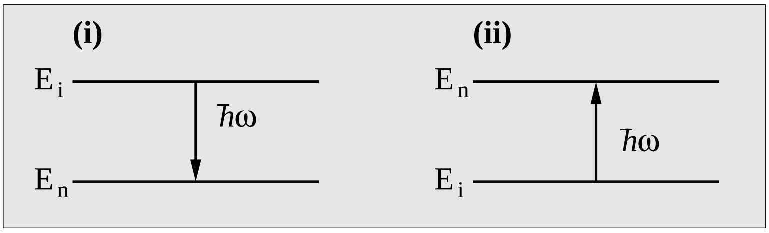


Figure 5.5: Schematic sketch for: (i) stimulated emission, the system gives up $\hbar\omega$; (ii) absorption, the system receives $\hbar\omega$.

6 Symmetry in Quantum Mechanics

We have already dealt with symmetries, e. g. an one-dimensional or spherically symmetrical potential. We have already learned that symmetry properties of the wave function can be used in order to classify the eigenstates of a Hamiltonian. If for example the parity operator commutes with the Hamiltonian, then parity eigenstates are also energy eigenstates.

In this chapter, we will address the relation between the symmetries of a molecule and the classification of the corresponding eigenstates in a more systematic way. This will require to consider some basics of group theory. However, before we will cover a symmetry of nature that is unknown in classical physics, namely the permutation symmetry between identical particles.

There are further important symmetry operations that we wil not consider. For example, there is the lattice translation in a periodic crystal that leads to the existence of so-called *Bloch states* which is important in solid state chemistry.

6.1 Identical Particles

Quantum mechanical identical particles behave completely differently compared to classical physics. In quantum mechanics, identical particles are truly indistinguishable, because we cannot follow the trajectories of particles by position measurements without disturbing the system. Consider the situation shown in Fig. 6.1. Suppose the state of the two particles is written as $|k'\rangle\,|k''\rangle$ which means that particle 1 is in state $|k'\rangle$ and particle 2 in state $|k''\rangle$. However, Fig. 6.1 shows two indistiguishable paths, i. e. $|k'\rangle\,|k''\rangle$ and $|k''\rangle\,|k'\rangle$ describe indistinguishable situations. On the other hand, for $k' \neq k''$, $|k'\rangle\,|k''\rangle$ and $|k''\rangle\,|k'\rangle$ are distinct kets! All kets of the form $c_1\,|k'\rangle\,|k''\rangle + c_2\,|k''\rangle\,|k'\rangle$ lead to an identical set of eigenvalues. This is called *exchange degeneracy*.

We now define the *permutation operator* P_{12}.

$$P_{12}\,|k'\rangle\,|k''\rangle = |k''\rangle\,|k'\rangle \qquad P_{12} = P_{21} \tag{6.1}$$

P_{12} is obviously an involution, i. e. $P_{12}^2 = \mathrm{id}$. Thus, the eigenvalues are ± 1. An

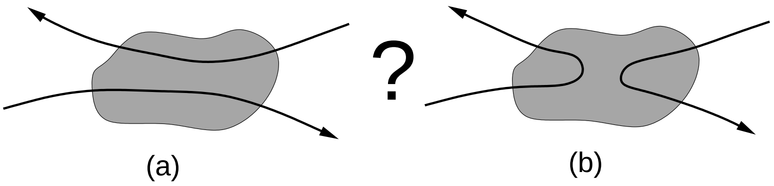


Figure 6.1: Two indistinguishable paths

observable with particle label transforms like

$$P_{12}A_1P_{12}^{-1} = A_2 = P_{12}A_1P_{12} \tag{6.2}$$

Consider a system of two identical particles

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} + V_{\text{pair}}(|\mathbf{x}_1 - \mathbf{x}_2|) + V_{\text{ext}}(\mathbf{x}_1) + V_{\text{ext}}(\mathbf{x}_2) \tag{6.3}$$

$$P_{12}HP_{12}^{-1} = H \iff \big[H, P_{12}\big] = 0 \tag{6.4}$$

Thus, P_{12} is a constant of motion. The two eigenvalues and eigenkets of P_{12} are the *symmetric* one with eigenvalue $+1$

$$|k', k''\rangle_+ = \frac{1}{\sqrt{2}}\big(|k'\rangle\,|k''\rangle + |k''\rangle\,|k'\rangle\big) \tag{6.5}$$

as well as the *antisymmetric* one with eigenvalue -1

$$|k', k''\rangle_- = \frac{1}{\sqrt{2}}\big(|k'\rangle\,|k''\rangle - |k''\rangle\,|k'\rangle\big) \tag{6.6}$$

Now we define the *symmetrizer* and the *antisymmetrizer*.

$$S_{12} \equiv \frac{1}{2}\big(1 + P_{12}\big) \qquad S_{12}\,|k', k''\rangle_+ = |k', k''\rangle_+ \qquad S_{12}\,|k', k''\rangle_- = 0 \tag{6.7}$$

$$A_{12} \equiv \frac{1}{2}\big(1 - P_{12}\big) \qquad A_{12}\,|k', k''\rangle_+ = 0 \qquad A_{12}\,|k', k''\rangle_- = |k', k''\rangle_- \tag{6.8}$$

The proof is left as an exercise.

If we generalize this concept for many identical particles, we get

$$P_{ij}\,|k^1\rangle\,|k^2\rangle\ldots|k^i\rangle\ldots|k^j\rangle\ldots = |k^1\rangle\,|k^2\rangle\ldots|k^j\rangle\ldots|k^i\rangle\ldots \tag{6.9}$$

Again, $P_{ij}^2 = \text{id}$. Note that in general $\big[P_{ij}, P_{kl}\big] \neq 0$.

Example

System of three identical particles. There are $3! = 6$ possible kets that are a combination of $|k'\rangle$, $|k''\rangle$, and $|k'''\rangle$, if they are all different kets.

There is only one *totally symmetrical* and one *totally antisymmetrical* form. These are the states that are symmetric or antisymmetric, respectively, under the interchange of any two arbitrary kets.

$$\begin{aligned} |k'\,k''\,k'''\rangle_\pm \equiv \frac{1}{\sqrt{3!}}\big[\big(&|k'\rangle\,|k''\rangle\,|k'''\rangle + |k''\rangle\,|k'''\rangle\,|k'\rangle + |k'''\rangle\,|k'\rangle\,|k''\rangle\big) \\ &\pm \big(|k''\rangle\,|k'\rangle\,|k'''\rangle + |k'''\rangle\,|k''\rangle\,|k'\rangle + |k'\rangle\,|k'''\rangle\,|k''\rangle\big)\big] \end{aligned} \tag{6.10}$$

These are both simultaneous eigenkets of P_{12}, P_{23}, and P_{13}. It follows that there are four independents states that are neither totally symmetrical nor antisymmetrical. Furthermore, we can introduce an operator P_{123} by

$$P_{123}(|k'\rangle\,|k''\rangle\,|k'''\rangle) = |k''\rangle\,|k'''\rangle\,|k'\rangle \tag{6.11}$$

Note that $P_{123} = P_{12}P_{13}$.

It is also important to realize that it is impossible to create a totally antisymmetric state if two of the three indices are the same.

In nature, N identical particles are either totally symmetrical or totally antisymmetrical under interchange of any pair of kets. In the first case the particles are called *bosons* and the so-called *Bose-Einstein statistics* applies, i. e.

$$P_{ij} \left|N \text{ identical bosons}\right\rangle = + \left|N \text{ identical bosons}\right\rangle \tag{6.12}$$

Fermions, on the other hand, are totally antisymmetrical, i. e. the *Fermi-Dirac statistics* applies:

$$P_{ij} \left|N \text{ identical fermions}\right\rangle = - \left|N \text{ identical fermions}\right\rangle \tag{6.13}$$

Half-integer spin particles are fermions and integer spin particles are bosons (this can only be proved in relativistic quantum mechanics). In particular, electrons are fermions. A direct consequence is the

Theorem Pauli-Exclusion Principle

No two electrons can occupy the same state.

Consider two particles that can occupy only two states. If these particles are fermions, then we get

$$\frac{1}{\sqrt{2}}\left(|k'\rangle\,|k''\rangle - |k''\rangle\,|k'\rangle\right) . \tag{6.14}$$

Alternatively, if we deal with bosons, we have three symmetrical states:

$$|k'\rangle\,|k'\rangle \qquad |k''\rangle\,|k''\rangle \qquad \frac{1}{\sqrt{2}}\left(|k'\rangle\,|k''\rangle + |k''\rangle\,|k'\rangle\right) \tag{6.15}$$

Classical particles satisfying the Maxwell-Boltzmann statistics can be found in four different states.

$$|k'\rangle\,|k''\rangle \qquad |k''\rangle\,|k'\rangle \qquad |k'\rangle\,|k'\rangle \qquad |k''\rangle\,|k''\rangle \tag{6.16}$$

In a sense, fermions “avoid” each other. Two of the three allowed states for bosons are states where both particles occupy the same state, while for “classical” particles, there are only two out of four.

Thus, bosons prefer to be in the same state much more than the other particles. The driving force (there is no “real” force) is a statistical one! The probability for the particle to be in a mutual state is much higher than the one to be in a different state. The most dramatic realization of the symmetric nature of bosons is the *Bose-Einstein condensation*.

6.2 Two-Electron System

Here, we consider two electrons (fermions).

$$\Psi = \sum_{m_{s_1}} \sum_{m_{s_2}} C(m_{s_1}, m_{s_2}) \langle \mathbf{x}_1, m_{s_1}, \mathbf{x}_2, m_{s_2} | \alpha \rangle \tag{6.17}$$

We assume that $\left[\mathbf{S}^2_{\text{total}}, H\right] = 0$, i. e. the total spin operator commutes with the Hamiltonian. Thus, eigenfunctions of $\mathbf{S}^2_{\text{total}}$ are eigenfunctions of the Hamiltonian.

$$\Psi = \Phi(\mathbf{x}_1, \mathbf{x}_2) \chi(m_{s_1}, m_{s_2}) \tag{6.18}$$

with

$$\chi(m_{s_1}, m_{s_2}) = \begin{cases} \chi_{\uparrow\uparrow} & \text{triplet} \\ \frac{1}{\sqrt{2}}\left(\chi_{\uparrow\downarrow} + \chi_{\downarrow\uparrow}\right) & \text{triplet} \\ \chi_{\downarrow\downarrow} & \text{triplet} \\ \frac{1}{\sqrt{2}}\left(\chi_{\uparrow\downarrow} - \chi_{\downarrow\uparrow}\right) & \text{singlet} \end{cases} \tag{6.19}$$

The first three states correspond to the symmetric triplet, the last one to the antisymmetric singlet. Note that $\langle \mathbf{x}_1, m_{s_1}; \mathbf{x}_2, m_{s_2} | P_{12} | \alpha \rangle = \langle \mathbf{x}_2, m_{s_2}; \mathbf{x}_1, m_{s_1} | \alpha \rangle$; hence we get for a fermion

$$\langle \mathbf{x}_1, m_{s_1}; \mathbf{x}_2, m_{s_2} | \alpha \rangle = -\langle \mathbf{x}_2, m_{s_2}; \mathbf{x}_1, m_{s_1} | \alpha \rangle \tag{6.20}$$

Due to the fact that we were able to separate the spin and the position dependence of the wave function, the permutation operator has to decompose into a space and a spin permutation operator:

$$P_{12} = P_{12}^{\text{space}} P_{12}^{\text{spin}} \tag{6.21}$$

In this case, the spin permutation operator is given by

$$\begin{aligned} P_{12}^{\text{spin}} &= \frac{1}{2}\left(1 + \frac{4}{\hbar^2}\mathbf{S}_1 \cdot \mathbf{S}_2\right) \\ |\alpha\rangle &\mapsto P_{12}\,|\alpha\rangle \;\Rightarrow\; \Phi(\mathbf{x}_1, \mathbf{x}_2) \mapsto \Phi(\mathbf{x}_2, \mathbf{x}_1) \qquad \chi(m_{s_1}, m_{s_2}) \mapsto \chi(m_{s_2}, m_{s_1}) \end{aligned} \tag{6.22}$$

The combined wave function has to be antisymmetrical, so if the spatial part is symmetrical, the spin state has to be antisymmetrical and vice versa.

$|\Phi(\mathbf{x}_1, \mathbf{x}_2)|^2 \, d^3x_1 \, d^3x_2$ is the probability for finding electron 1 around $\mathbf{x}_1$ and electron 2 around $\mathbf{x}_2$. Assume that there is no mutual interaction, i. e.

$$H \equiv H(\mathbf{p}_1, \mathbf{p}_2, \mathbf{x}_1, \mathbf{x}_2) = \frac{\mathbf{p}_1^2}{2m_e} + \frac{\mathbf{p}_2^2}{2m_e} + V_{\text{ext}}(\mathbf{x}_1) + V_{\text{ext}}(\mathbf{x}_2) \tag{6.23}$$

In such a case, the eigenfunctions are *separable*:

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}}\left(\varphi_A(\mathbf{x}_1)\varphi_B(\mathbf{x}_2) \pm \varphi_A(\mathbf{x}_2)\varphi_B(\mathbf{x}_1)\right) \tag{6.24}$$

The $+$ sign is for the spin singlet, the $-$ sign for the spin triplet. Thus we get for the probability

$$\begin{aligned} |\Phi(\mathbf{x}_1, \mathbf{x}_2)|^2 \, d^3x_1 \, d^3x_2 = \frac{1}{2}\Big[& |\varphi_A(\mathbf{x}_1)|^2 \, |\varphi_B(\mathbf{x}_2)|^2 + |\varphi_A(\mathbf{x}_2)|^2 \, |\varphi_B(\mathbf{x}_1)|^2 \\ & \pm \underbrace{2\Re\left(\varphi_A(\mathbf{x}_1)\varphi_B(\mathbf{x}_2)\varphi_A^*(\mathbf{x}_2)\varphi_B^*(\mathbf{x}_1)\right)}_{\text{exchange density}} \Big] \, d^3x_1 \, d^3x_2 \end{aligned} \tag{6.25}$$

This implies that in the triplet state $|\Phi(\mathbf{x}_1, \mathbf{x}_2 = \mathbf{x}_1)|^2 = 0$, whereas for a singlet state we get an *enhanced probability* to find both particles at the same point in space: $|\Phi(\mathbf{x}_1, \mathbf{x}_2 = \mathbf{x}_1)|^2 = 2\,|\varphi_A(\mathbf{x}_1)|^2\,|\varphi_B(\mathbf{x}_2)|^2$[1].

6.3 The Helium Atom

The simplest system where the problem of identity plays an important role is the helium atom. Still the two-particle Schrödinger equation with the basic Hamiltonian

$$H = \frac{\mathbf{p}_1^2}{2m_e} + \frac{\mathbf{p}_2^2}{2m_e} - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} \tag{6.26}$$

where $r_1 = |\mathbf{x}_1|$, $r_2 = |\mathbf{x}_2|$, and $r_{12} = |\mathbf{x}_1 - \mathbf{x}_2|$, cannot be solved analytically. Hence this system also serves as an example for the application of approximation methods.

The total spin is a constant of motion, because there is no external magnetic field so that the Hamiltonian commutes with the total spin operator. Hence, the spin state is either a singlet or a triplet. Since there is no analytical solution for this problem, we will first apply perturbation theory with the perturbation ${e^2}/{r_{12}}$. The ground state configuration of the unperturbed system is $1s^2$, i. e. both electrons are in $n = 1$, $l = 0$. The spatial part of the wave function is then symmetric so that only the spin singlet function is allowed.

6.3.1 Ground State

The unperturbed wave function is the product of the single-electron wave functions and the proper spinor:

$$\Psi_{100}(\mathbf{x}_1)\Psi_{100}(\mathbf{x}_2)\chi_{\text{singlet}} = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0}\chi_{\text{singlet}} \tag{6.27}$$

The ground state energy of H_0 is

$$E_0 = -2 \times 4\frac{e^2}{2a_0} = -8\,\mathrm{Ryd} = -108.8\,\mathrm{eV} \tag{6.28}$$

This is considerably less than the experimental value of $E_0^{\exp} = -78.8\,\mathrm{eV}$.

Let's apply first-order perturbation theory:

$$\begin{aligned} \Delta^{(1)}_{1s^2} &= \left\langle \frac{e^2}{r_{12}} \right\rangle_{1s^2} \\ &= \int\!\!\int \frac{Z^6}{\pi^2 a_0^6} e^{-2Z(r_1+r_2)/a_0} \frac{e^2}{r_{12}}\, d^3x_1\, d^3x_2 \end{aligned} \tag{6.29}$$

Let $r_<$ and $r_>$ be the smaller and larger radius of r_1 and r_1 and γ is angle between $\mathbf{x}_1$ and $\mathbf{x}_2$:

$$\begin{aligned} \frac{1}{r_{12}} &= \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\gamma}} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\gamma) \\ &= \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y^*_{lm}(\vartheta_1, \varphi_1) Y_{lm}(\vartheta_2, \varphi_2) \end{aligned} \tag{6.30}$$

---

[1] this does not violate the Pauli exclusion principle, because the electron have opposite spin

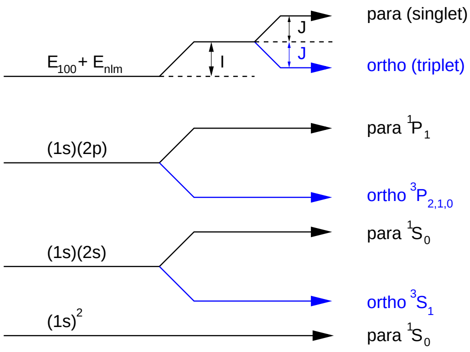


Figure 6.2: Schematic diagram for the energy splitting of the low-lying electronic states of the helium atom. The upper diagram illustrates the splitting of the $(1s)(nl)$ states in general, while the lower panels shows the lowest states explicitly.

After some algebra we get for the first-order shift

$$\begin{aligned} \Delta^{(1)}_{1s^2} &= \frac{5}{2}\frac{e^2}{2a_0} \\ \Rightarrow\ E^{(1)}_0 &= -\big(8-\frac{5}{2}\big)\frac{e^2}{2a_0} \approx -74.8\,\mathrm{eV} \approx 95\% \cdot E^{\mathrm{exp}}_0 \end{aligned} \tag{6.31}$$

Now we apply the variational method. Assume that the nucleus is screened by the charge cloud of the other electron. Thus our trial function is

$$\begin{aligned} \langle \mathbf{x}_1, \mathbf{x}_2 | \tilde{0} \rangle &= \frac{Z^3_{\mathrm{eff}}}{\pi a_0^3} e^{-Z_{\mathrm{eff}}(r_1+r_2)/a_0} \\ \bar{H} &= \left\langle \tilde{0} | \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} | \tilde{0} \right\rangle - \left\langle \tilde{0} | \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} | \tilde{0} \right\rangle + \left\langle \tilde{0} | \frac{e^2}{r_{12}} | \tilde{0} \right\rangle \\ &= \big(2 \cdot \frac{Z^2_{\mathrm{eff}}}{2} - 2ZZ_{\mathrm{eff}} + \frac{5}{8} Z_{\mathrm{eff}}\big)\frac{e^2}{a_0} \end{aligned} \tag{6.32}$$

where Z_{eff} is the optimization parameter. This equation is minimized when $Z_{\mathrm{eff}} = 27/16 = 1.6875$. Hence, $E_0 = -77.5\,\mathrm{eV}$, which is already rather close to the measured value of $E^{\mathrm{exp}}_0 = -78.8\,\mathrm{eV}$

6.3.2 Excited States

Now we will consider the excited states $1s\,nl$. The energy is $E = E_{100} + E_{nlm} + \Delta E$. To first order, we get

$$\Delta E = \left\langle \frac{e^2}{r_{12}} \right\rangle = I \pm J \tag{6.33}$$

Here, we take to + sign for the singlet, the $-$ sign for the triplet. I and J are the direct integral and the exchange integral, respectively, which are given by

$$I = \iint \left|\Psi_{100}(\mathbf{x}_1)\right|^2 \left|\Psi_{nlm}(\mathbf{x}_2)\right|^2 \tfrac{e^2}{r_{12}} d^3x_1\, d^3x_2$$
$$J = \iint \Psi_{100}(\mathbf{x}_1)\Psi_{nlm}(\mathbf{x}_2)\tfrac{e^2}{r_{12}}\Psi^*_{100}(\mathbf{x}_2)\Psi^*_{nlm}(\mathbf{x}_1) d^3x_1\, d^3x_2 \tag{6.34}$$

The resulting energy spectrum is shown schematically in Fig. 6.2. Note that the energies are spin-dependent although there is no spin term in the Hamiltonian. Qualititatively, the spectrum can be understood as follows. In the singlet states which is known as parahelium the spatial part of the wave function is symmetric so that the two electrons tend to be closer to each other. Because of the electrostatic repulsion, the singlet states are therefore higher in energy than the triplet states known as orthohelium where the space funtion is antisymmetric so that the electrons tend to avoid each other.

6.4 Symmetry of molecules: Group theory and representations

In chapter 4 we already learned that there is an intimate relationship between symmetries, the rotational symmetry, and the solution of the corresponding Schrödinger equation, the spherical harmonics. Here we are concerned with molecules. If a molecule is fixed in space, the translational and rotational invariance is broken. There are still symmetry operation that transform the molecule into itself.

In detail, we will mainly consider the following point operations, that leave at least one point in space unchanged:

E	identity
c_n	Rotation through $2\pi/n$ about some axis
σ	Reflection
s_n	Rotation through $2\pi/n$ followed by a reflection in a plane perpendicular to this axis
$i = s_2$	Inversion

If two symmetry operations A and B are applied successively, another symmetry operation C results. One writes: $AB = C$. For a certain number of symmetry operations, the results of the successive application can be listed in a multiplication table, as it is shown in Table 6.1 for the symmetry group C_{2h}. In fact, as it is easy to show, the symmetry operations listed in Table 6.1 obey all the rules necessary to form a *group*, namely the existence of the identity, closure, the existence of inverse elements and associativity (see page 65). Thus these elements form a group of transformations T.

Let us now consider the effect of symmetry operations $A, B, C, \ldots$ on energy eigenfunctions of a Hamiltonian with a discrete spectrum. For the sake of simplicity, we will only take a finite number, say M, of eigenstates into account. Furthermore, let $\psi_1, \psi_2, \ldots, \psi_M$ be a basis which does not necessarily consist of eigenstates. We first apply the symmetry operation A that commutes with the Hamiltonian. This could for example correspond to a rotation. Since A commutes with H, applying A on an *eigenstate* ϕ_1 does not change the energy eigenvalue of the resulting wave function, as can be easily seen from

$$HA\phi_1 \;=\; AH\phi_1 = \varepsilon_1 A\phi_1 \; . \tag{6.35}$$

which means that $A\phi_1$ and ϕ_1 have the same eigenvalue. For the general ket ψ_1 this means that $A\psi_1$ can also be expressed as a sum over the energy eigenstates or, equivalently, in any other suitable basis such as the $\psi_i, i = 1, \dots, M$,

$$A\psi_1 \;=\; \sum_{j=1}^{M} a_{1j}\psi_j \;. \tag{6.36}$$

The same can in fact be done for any of the wavefunctions ψ_i:

$$A\psi_i \;=\; \sum_{j=1}^{M} a_{ij}\psi_j \;, \tag{6.37}$$

which means that the effect of the symmetry operation A on the wave functions defines a matrix $A \rightarrow (a_{ij})$. The same can be done for the other symmetry operations $B, C, \dots$ that form the group of transformations T. It is a simple exercise to show that all the properties of the group formed by $A, B, C, \dots$ is transferred to the matrices $(a_{ij}), (b_{ij}), (c_{ij}), \dots$. This means that the matrices form a representation of the original group $A \rightarrow D(A) = (a_{ij})$ which we call D from the German word for representation, *Darstellung* (for the definition of a representation, see page 65; formally one can say that the matrices of a representation correspond to a group that is homomorphic to the original group). The wave functions $\psi_i, i = 1, \dots, M$, are said to form a basis for the representation D.

If there are degenerate wave functions, any linear combination of these wave functions could be used as a basis for the degenerate subspace. This would of course also influence the representation. However, these representations are closely linked and could be transformed into each other. In fact, two representations D and D' are said to be *equivalent* if a matrix P exists so that

$$D'(T) \;=\; P^{-1}D(T)P \;, \tag{6.38}$$

where T stands for any symmetry operation of the original group, which means that eq. (6.38) should be obeyed for all elements of the group simultaneously.

6.4.1 Irreducible presentations

An important aspect of a representation is its so-called *reducibility* . It is of course desirable to obtain a representation that is rather simple. Preferably this would be a

Table 6.1: Multiplication table of the symmetry operations E, c_2, σ and i forming the symmetry group C_{2h}.

Operation	E	c_2	σ	i
E	E	c_2	σ	i
c_2	c_2	E	i	σ
σ	σ	i	E	c_2
i	i	σ	c_2	E

diagonal form of the matrices with non-zero elements of the matrices only along the diagonal. For a single $(n \times n)$ matrix A this is possible by applying transformations of the type (6.38). However, we have to consider that we have to diagonalize several different matrices with the *same* transformation matrix, and then it is in general not possible to bring all matrices in a diagonal form. Still, it is often possible to bring every matrix $D(T)$ of the representation into a block-diagonal form with smaller matrices $D^{(i)}(T)$ along the diagonal and zeros elsewhere, i.e. into a form that looks like

$$\begin{pmatrix} \begin{pmatrix} & & \\ & D^{(1)} & \\ & & \end{pmatrix} & 0 & 0 & \dots & 0 \\ 0 & \begin{pmatrix} & \\ D^{(2)} & \\ & \end{pmatrix} & 0 & \dots & 0 \\ 0 & 0 & \left(D^{(3)}\right) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \left(D^{(n)}\right) \end{pmatrix} \tag{6.39}$$

If such transformation is possible, the representation D is said to be *reducible* and to contain the representations $D^{(i)}$. This is written as

$$D \;=\; D^{(1)} + D^{(2)} + D^{(3)} + \ldots + D^{(n)} \; . \tag{6.40}$$

How many different irreducible representations of a group exist? One can show that this exactly corresponds to the number of classes of a group which will be discussed below.

The important point here is that the decomposition of the representation into irreducible representation corresponds to finding a basis of the wave functions ϕ_k so that only a subset of the ϕ_k will be transformed into itself upon applying the symmetry operations. The states of this subset will then be energetically degenerate which means that the dimensions of the irreducible presentations directly give the degree of degeneracies of the corresponding states belonging to the particular subset. Likewise, if there are only one-dimensional representations, then there are no energetically degenerate states. The way the ϕ_k transform is furthermore independent of the particular quantum problem, it only depends on the symmetry of the system.

There is a special nomenclature to characterize irreducible representation. One-dimensional representations are denoted by the letters A and B depending on whether the representation is symmetric or antisymmetry with respect to a rotation about the main symmetry axis. Subscripts g and u indicate whether the one-dimensional representations are even (*gerade*) or odd (*ungerade*) with respect to the inversion. The letters E and F stand for two- and three-dimensional representations, respectively. Further subscripts 1 and 2 and one or two apostrophs can be added to indicate further symmetries that will not be discussed here.

6.4.2 Group characters

Before we proceed any further, some basics about group theory shall be addressed. The number of elements in a group is called the *order* of the group and is often denoted by the letter h. For the example of Table 6.1, $h = 4$. One element B of a group G is said to be *conjugate* to the element A if there is an element C in the group so that

$$B \;=\; C^{-1}AC \; , \tag{6.41}$$

For symmetry operations, (6.41) represents nothing else but a transformation. Eq. (6.41) also defines an *equivalence relation* that can be used to separate the group into classes. A *conjugate class* consists of all elemtents of a group that are conjugate to each other. Note that the identity E always forms a class of itself because of $C^{-1}EC = E$ for any C. The group of Table 6.1 is an *abelian* or *commutative* group in which all elements commute, i.e. $AB = BA$. In such a group, each element forms a class of itself.

A central tool in the representation theory is the *character* χ. It is defined as the trace of the representation $D(A)$ of the transformation A:

$$\chi(A) \;=\; \sum_i D_{ii}(A)\;. \tag{6.42}$$

Since the trace remains unchanged by transformations of the type (6.41), the characters of all elements in one class are the same. Furthermore, the characters of equivalent representations are the same, as is directly apparent from (6.38).

The character table of the group C_{2h} is shown in Table 6.2. All representations are one-dimensional which means that there would be no symmetry-related degeneracies. In fact, the dimensionality of a representation is given by $\chi(E)$ since E is in any representation given by the identity matrix. We have used the conventional nomenclature of representations just mentioned in Table 6.2.

An important question is of course, how a reducible representation can be decomposed into its irreducible representations $D^{(1)} + D^{(2)} + D^{(3)} + \ldots + D^{(n)}$. To this end, the characters are very useful. We will sketch here some important relations without giving all the proofs. Please note that usually it is not necessary to perform this decomposition by oneself. For simple groups, the irreducible representations are tabulated in many text books or at least deducible from the tabulated ones, and complicated groups required advanced methods.

First of all, the character of a symmetry operation in a reducible representation is given by the sum of the characters in the irreducible components,

$$\chi(T) \;=\; \sum_i n_i \chi_i(T)\;, \tag{6.43}$$

where n_i is the number of times the irreducible component $D^{(i)}$ occurs in the reduction of D. If a symmetry group, its irreducible representations and a particular basis of a representation are given, then often the n_i can be deduced by inspection or by a trial and error approach.

Equation (6.43) looks similar to the decompostion of a wave function ψ into a linear combination of eigenfunctions ψ_i. In fact, for the characters also an orthogonality

Table 6.2: Character table of the group C_{2h}.

Representation	E	c_2	σ	i
A_g	1	1	1	1
B_g	1	-1	-1	1
A_u	1	1	-1	-1
B_u	1	-1	1	-1

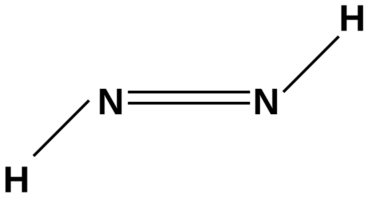


Figure 6.3: Schematic presentation of the N_2H_2 molecule.

relation can be derived which is here given without proof:

$$\frac{1}{h}\sum_{T}\chi_j(T)\chi_i(T) \;=\; \delta_{ij} \;, \tag{6.44}$$

where the sum is to be performed over all symmetry operations T. Equation (6.44) can be used together with eq. (6.43) to determine how often a irreducible representation $D^{(i)}$ is contained in D. Just multiply (6.43) with χ_j and sum over all elements of the group:

$$n_i \;= \frac{1}{h}\sum_{T}\;\chi_i(T)\chi(T) \;== \frac{1}{h}\sum_{Q}\;N_Q\chi_i(T)\chi(T)\;, \tag{6.45}$$

where for the second sum we have used the fact that the characters for all elements of a class are the same, so that we only have to sum over all classes Q with N_Q being the number of elements of a class.

In order to obtain the basis of an irreducible representation from a reducible representation, there is in fact also an useful tool, namely a *projection operator* which can be explicitly constructed using

$$P_i \;= \frac{1}{h}\sum_{T}\;\chi_i(T^{-1})D(T) \tag{6.46}$$

We will illustrate some of these concepts in the next section using the N_2H_2 molecule.

6.4.3 An example: the N_2H_2 molecule

So far we have not considered any molecules. The theory of chemical bonding will only be addressed in the next chapter. Still we will already introduce some of the concepts relevant for the quantum description of molecules in order to illustrate the application of representation theory for the description of molecular properties. In the next chapter we will then present a more eloborate example.

Here we will use the N_2H_2 molecule that is shown in Fig. 6.3 as an example. The symmetry group of the N_2H_2 molecule corresponds in fact to the group C_{2h} whose multiplication table in shown in Tab. 6.1. The symmetry operations besides the identity operation are the rotation by 180° about the molecular center of mass with the axis being perpendicular to the plane of Fig. 6.3, the reflection at the plane of Fig. 6.3, and the inversion at the molecular center of mass.

We will consider a two-dimensional basis that corresponds to both of the NH groups of the molecule. A basis does not necessarily be constructed from functions but could in fact also be based on abstract objects. Applying the identity and the reflection will keep the NH groups unchanged whereas they will be interchange upon the rotation and the inversion. This means that the symmetry operations are represented by the following matrices and characters in this representation, that we may call Γ:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad c_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.47}$$

$$\chi(E) = 2 \qquad \chi(c_2) = 0 \qquad \chi(\sigma) = 2 \qquad \chi(i) = 0 \tag{6.48}$$

In order to find the number of irreducible reprensentations, in such a simple case as in (6.48) often simple combinatorial considerations based on eq. (6.43) are sufficient. One just has to combine two representations listed in Tab. 6.2 in such a way that the sum of the characters of each class are the same in in (6.48). Thus it is easy to see that the Γ can be decomposed into

$$\Gamma \;=\; A_g \;+\; B_u \tag{6.49}$$

We can also check eq. (6.44) yields the same result. We obtain

$$n_{A_g} = \frac{1}{4}\Big(\underbrace{\chi_{A_g}(E)}_{=1}\underbrace{\chi_\Gamma(E)}_{=2} + \underbrace{\chi_{A_g}(c_2)}_{=1}\underbrace{\chi_\Gamma(c_2)}_{=0} + \underbrace{\chi_{A_g}(\sigma)}_{=1}\underbrace{\chi_\Gamma(\sigma)}_{=2} + \underbrace{\chi_{A_g}(i)}_{=1}\underbrace{\chi_\Gamma(i)}_{=0}\Big) \;=\; 1\,, \tag{6.50}$$

$$n_{A_u} = \frac{1}{4}\Big(\underbrace{\chi_{A_u}(E)}_{=1}\underbrace{\chi_\Gamma(E)}_{=2} + \underbrace{\chi_{A_u}(c_2)}_{=1}\underbrace{\chi_\Gamma(c_2)}_{=0} + \underbrace{\chi_{A_u}(\sigma)}_{=-1}\underbrace{\chi_\Gamma(\sigma)}_{=2} + \underbrace{\chi_{A_u}(i)}_{=-1}\underbrace{\chi_\Gamma(i)}_{=0}\Big) \;=\; 0\,, \tag{6.51}$$

Similarly, one obtains $n_{B_g} = 0$ and $n_{B_u} = 1$, i.e. we confirm (6.49). In the next chapter, we will use this concept in order to construct molecular electronic wave functions that correspond to irreducible representations.

7 Theory of chemical bonding

In this chapter, we will provide an introduction into the basic theoretical concepts that are necessary in order to understand the nature of chemical bonding. First the relevant Hamiltonian will be introduced, and the methods in order to determine the corresponding Schrödinger equation will be discussed. In fact, only the simplest molecule, the H_2^+ ion, can still be solved analytically, for all other molecules approximations are required for the solution of the Schrödinger equation which will be presented in this chapter.

7.1 Hamiltonian

In chemistry, the only basic interaction one is concerned with is the electrostatic interaction between the nuclei and electrons. Hence the only terms that enter the Hamiltonian are the kinetic energy of the nuclei and the electrons and the nucleus–nucleus, nucleus-electron and electron-electron electrostatic interaction energy:

$$H \;=\; T_{\rm nucl} \;+\; T_{\rm el} \;+\; V_{\rm nucl-nucl} \;+\; V_{\rm nucl-el} \;+\; V_{\rm el-el} \tag{7.1}$$

Neglecting relativistic effects, the single terms are as follows.
Kinetic energy of the nuclei:

$$T_{\rm nucl} \;=\; \sum_{\rm I=1}^{L} \frac{\vec{P}_{\rm I}^{2}}{2M_{\rm I}}, \tag{7.2}$$

Kinetic energy of the electrons:

$$T_{\rm el} \;=\; \sum_{\rm i=1}^{N} \frac{\vec{p}_{\rm i}^{2}}{2m}, \tag{7.3}$$

Nucleus-nucleus electrostatic energy:

$$V_{\rm nucl-nucl} \;=\; \frac{1}{2} \sum_{\rm I\neq J} \frac{Z_{\rm I}\; Z_{\rm J}\; e^2}{|\vec{R}_{\rm I} - \vec{R}_{\rm J}|}, \tag{7.4}$$

Nucleus-electron electrostatic energy:

$$V_{\rm nucl-el} \;=\; -\sum_{\rm i,I} \frac{Z_{\rm I}\; e^2}{|\vec{r}_{\rm i} - \vec{R}_{\rm I}|}, \tag{7.5}$$

Eelctron-electron electrostatic energy:

$$V_{\rm el-el} \;=\; \frac{1}{2} \sum_{\rm i\neq j} \frac{e^2}{|\vec{r}_{\rm i} - \vec{r}_{\rm j}|}. \tag{7.6}$$

Magnetism can be taken into account either by a full relativistic treatment or by adding magnetic effects explicitly in the Hamiltonian. This Hamiltonian then enters the non-relativistic Schrödinger equation

$$H \; \Phi(\vec{R}, \vec{r}) \;=\; E \; \Phi(\vec{R}, \vec{r}) \tag{7.7}$$

where $\vec{R}$ and $\vec{r}$ denote the *all* nuclear and electronic coordinates, respectively. In principle, we are ready here because by solving the Schrödinger equation and determining the eigenfunctions of the many-body Hamiltonian with taking the proper quantum statistics of the considered particles into account we get the full information about the system. Unfortunately, the solution of the many-body Schrödinger equation in closed form is not possible. In order to make the solution tractable, a hierarchy of approximations is needed.

7.2 Born–Oppenheimer Approximation

The central idea underlying the Born–Oppenheimer or adiabatic approximation is the separation in the time scale of processes involving electrons and atoms. Typical mass ratios between the nuclei and the electron are

$$\begin{aligned} M_{\mathrm{H}}/m_e &\approx 1840 \\ M_{\mathrm{Si}}/m_e &\approx 51520 \\ M_{\mathrm{Cu}}/m_e &\approx 115920 \end{aligned} \tag{7.8}$$

Thus, except for hydrogen and helium, atoms have a mass that is 10^4 to 10^5 times larger than the mass of an electron. Consequently, at the same kinetic energy electrons are 10^2 to 10^3 times faster than the nuclei. Hence one assumes that the electrons follow the motion of the nuclei instantaneously. The electron distribution then determines the potential in which the nuclei moves.

In practice, one splits up the full Hamiltonian and defines the electronic Hamiltonian H_{el} for fixed nuclear coordinates $\vec{R}$ as follows

$$H_{\mathrm{el}}(\{\vec{R}\}) = T_{\mathrm{el}} + V_{\mathrm{nucl-nucl}} + V_{\mathrm{nucl-el}} + V_{\mathrm{el-el}}. \tag{7.9}$$

In (7.9) the nuclear coordinates $\{\vec{R}\}$ do not act as variables but as parameters defining the electronic Hamiltonian. The Schrödinger equation for the electrons for a given fixed configuration of the nuclei is then

$$H_{\mathrm{el}}(\{\vec{R}\})\Psi(\vec{r}; \{\vec{R}\}) = E_{\mathrm{el}}(\{\vec{R}\})\Psi(\vec{r}; \{\vec{R}\}). \tag{7.10}$$

Again, in (7.10) the nuclear coordinates $\{\vec{R}\}$ are not meant to be variables but parameters. In the Born–Oppenheimer or adiabatic approximation the eigenenergy $E_{\mathrm{el}}(\{\vec{R}\})$ of the electronic Schrödinger equation is taken to be the potential for the nuclear motion. $E_{\mathrm{el}}(\{\vec{R}\})$ is therefore called the Born–Oppenheimer energy surface. By finding minima and extrema of the Born-Oppenheimer surface, the geometry and binding energy of molecules can be determined, but also activation barriers for chemical reactions.

In order to describe the dynamics of the nuclear motion within the Born-Oppenheimer approximation, the atomic Schrödinger equation

$$\{T_{\mathrm{nucl}} \;+\; E_{\mathrm{el}}(\vec{R})\} \; \Lambda(\vec{R}) = E_{\mathrm{nucl}}\Lambda(\vec{R}). \tag{7.11}$$

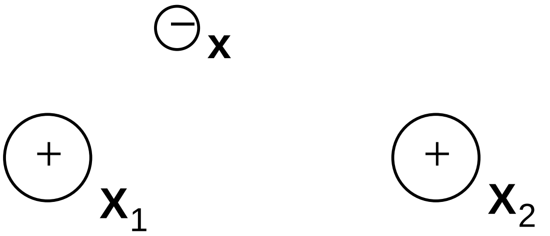


Figure 7.1: The coordinates of the ionized H_2 molecule

has to be solved. Here $E_{\text{el}}(\vec{R})$ can now be regarded as a function of the nuclear coordinates $\vec{R}$. Often the quantum effects in the atomic motion are neglected and the classical equation of motion are solved for the atomic motion:

$$M_I \frac{\partial^2}{\partial t^2} \vec{R}_I = -\frac{\partial}{\partial \vec{R}_I} E_{\text{el}}(\vec{R}). \tag{7.12}$$

The force acting on the atoms can be conveniently evaluated using the Hellmann–Feynman theorem

$$\vec{F}_I = -\frac{\partial}{\partial \vec{R}_I} E_{\text{el}}(\vec{R}) = \langle \Psi(\vec{r}, \{\vec{R}\}) | \frac{\partial}{\partial \vec{R}_I} H_{\text{el}}(\vec{R}) | \Psi(\vec{r}, \{\vec{R}\}) \rangle. \tag{7.13}$$

In principle, in the Born–Oppenheimer approximation electronic transition due to the motion of the nuclei are neglected. Here we do not derive the Born-Oppenheimer approximation in detail (see, e.g., J. Callaway, *Quantum Theory of the Solid State*). The important parameter for the validity of the adiabatic approximation is the mass ratio m_e/M. In fact, the Born–Oppenheimer approximation is very successful in the theoretical description of many properties of solids. Still its true validity is hard to prove because it is very difficult to correctly describe processes that involve electronic transition.

If it takes a finite amount of energy to excite electronic states, i.e., if the adiabatic electronic states are well-separated, then it can be shown that electronically nonadiabatic transitions are rather improbable. This applies to insulator and semiconductor surfaces with a large band gap. In metals, no fundamental band gap exists so that electronic transitions with arbitrarily small excitations energies can occur. Still, the strong coupling of the electronic states in the broad conduction band leads to short lifetimes of excited states and thus to a fast quenching of these states.

7.3 The H_2^+ molecule

In sect. 6.3, we have already seen that a system consisting of two electrons, the helium atom, cannot be solved analytically, but only approximately, for example by a variational method. Here we are concerned with the binding between two atoms mediated through electrons. As a first, relatively easy example we consider the ionized H_2^+ molecule with an electron moving in the attractive potential of two protons which within the Born-Oppenheimer approximation are assumed to be fixed in space. The configuration is illustrated in Fig. 7.1. The Hamiltonian is then given by:

$$H = \frac{\mathbf{p^2}}{2m_e} - \frac{e^2}{|\mathbf{x} - \mathbf{X}_1|} - \frac{e^2}{|\mathbf{x} - \mathbf{X}_2|} + \frac{e^2}{|\mathbf{X}_1 - \mathbf{X}_2|} \tag{7.14}$$

This problem is in fact exactly solvable but requires some coordinate transformation. Here we will use a approximate variational method as in sect. **??** in order to estimate the equilibrium distance of the protons in the H_2^+ molecule. The problem is invariant with respect to a translation of the molecular center of mass and a rotation of the molecular axis. Thus the energy only depends on the distance of the two protons. Since the system has an inversion symmetry with respect to the molecular center of mass which corresponds to the midpoint between the protons, the solutions have to be either symmetric or antisymmetric with respect to this inversion, as group theory tells us (see Sect. 6.4).

As the initial guess we take the symmetric and antisymmetric superposition of hydrogen $1s$ functions:

$$\psi_\pm \ = \ C_\pm \left[\psi_1(\mathbf{x}) \pm \psi_2(\mathbf{x})\right] , \tag{7.15}$$

where the two $1s$ functions located at the nuclei $\mathbf{X}_1$ and $\mathbf{X}_2$ are (see sec. 4.5)

$$\langle \mathbf{x} | 1s_{1,2} \rangle = \psi_{1,2}(\mathbf{x}) \ = \ \frac{1}{\sqrt{\pi a_0^3}} \exp\left(-|\mathbf{x} - \mathbf{X}_{1,2}|/a_0\right) \tag{7.16}$$

The normalization constants $C_\pm$ follow from the overlap of the two wavefunctions ψ_1 and ψ_2:

$$1 = \int d^3x |\psi_\pm(\mathbf{x})|^2 = C_\pm^2 \, 2(1 \pm S(R)) , \tag{7.17}$$

where the overlap integral $S(R)$ as a function of the distance $R = |\mathbf{X}_1 - \mathbf{X}_2|$ is given by

$$S(R) = \int d^3x \ \psi_1^*(\mathbf{x})\psi_2(\mathbf{x}) = \left(1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2}\right) \exp(-R/a_0) . \tag{7.18}$$

The proof of this relation is left as an exercise.

The symmetric and the antisymmetric combination of the atomic orbitals are illustrated in Fig. 7.2. The symmetric combination ψ_+ is larger in the region between the two nuclei than the antisymmetric one ψ_-. Since there is an attraction between the positively charged nuclei and the electrons, we expect that ψ_+ will be energetically more favorable than ψ_-.

In order to check this expectation, we have to evaluate the expectation value of the Hamiltonian with respect to these states 7.15 which can be expressed as

$$\begin{aligned} \langle H \rangle_\pm &= (2 \pm 2S)^{-1} (\langle 1s_1|H|1s_1\rangle + \langle 1s_2|H|1s_2\rangle \pm 2 \langle 1s_1|H|1s_2\rangle) \\ &= (1 \pm S)^{-1} (\langle 1s_1|H|1s_1\rangle \pm \langle 1s_1|H|1s_2\rangle). \end{aligned} \tag{7.19}$$

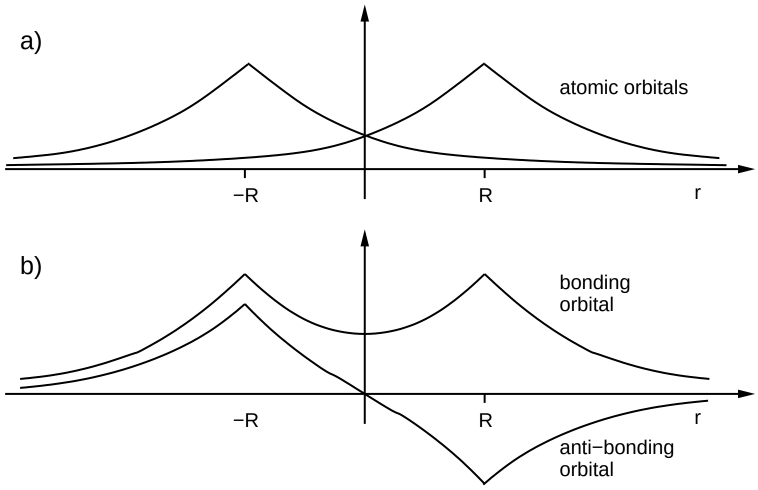


Figure 7.2: Illustration of the difference between the symmetric and the anti-symmetric combination of atomic orbitals.

The Hamiltonian matrix elements can be evaluated as follows:

$$\begin{aligned}\langle 1s_1|H|1s_1\rangle &= \int d^3x\ \psi_1^*(\mathbf{x})H\psi_1(\mathbf{x}) \\ &= E_0 - \int d^3x\ \psi_1^2(\mathbf{x})\frac{e^2}{|\mathbf{x}-\mathbf{X}_2|} + \frac{e^2}{R} \\ &= E_0 - \frac{e^2}{R}\left(1 - e^{-2R/a_0}\left(\frac{R}{a_0}+1\right)\right) + \frac{e^2}{R} \\ &= E_0 + \frac{e^2}{R}\left(\frac{R}{a_0}+1\right)e^{-2R/a_0} \qquad (7.20)\end{aligned}$$

where E_0 is the ground state energy of -1 Rydberg $= -13.6$ eV of the hydrogen atom. Furthermore,

$$\begin{aligned}(\langle 1s_1|H|1s_2\rangle &= \int d^3x\ \psi_1^*(\mathbf{x})H\psi_2(\mathbf{x}) \\ &= \left(E_0 + \frac{e^2}{R}\right)S(R) - \int d^3x\ \psi_1(\mathbf{x})\psi_2(\mathbf{x})\frac{e^2}{|\mathbf{x}-\mathbf{X}_1|} \\ &= \left(E_0 + \frac{e^2}{R}\right)S(R) - \frac{e^2}{a_0}\left(\frac{R}{a_0}+1\right)e^{-R/a_0} \qquad (7.21)\end{aligned}$$

where the last integral in the second line

$$A(R) = \int d^3x\ \psi_1(\mathbf{x})\psi_2(\mathbf{x})\frac{e^2}{|\mathbf{x}-\mathbf{X}_2|} = \frac{e^2}{a_0}\left(\frac{R}{a_0}+1\right)e^{-R/a_0} \qquad (7.22)$$

is called the exchange integral. In total, the energy of the hydrogen ion as a function

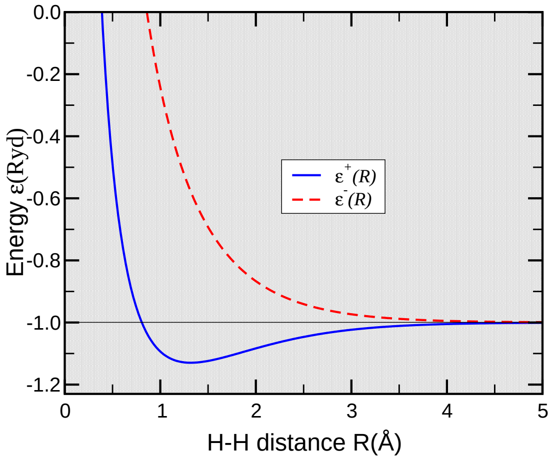


Figure 7.3: The energy of the H_2^+ molecule as a function of the interatomic H-H distance R in the symmetric bonding state ε_+ and the anti-symmetric antibonding state ε_-.

of the interatomic distance R is given by

$$\begin{aligned}\varepsilon_\pm(R) \equiv \langle H\rangle_\pm = (1\pm S)^{-1}\Bigg[E_0 + \frac{e^2}{R}\left(\frac{R}{a_0}+1\right)e^{-2R/a_0} \\ \pm\left(\left(E_0+\frac{e^2}{R}\right)S(R) - \frac{e^2}{a_0}\left(\frac{R}{a_0}+1\right)e^{-R/a_0}\right)\Bigg] \qquad (7.23)\end{aligned}$$

With $E_0 = -\frac{e^2}{2a_0}$, this can be rewritten as

$$\begin{aligned}\varepsilon_\pm(R) \equiv \langle H\rangle_\pm =&(1\pm S)^{-1}E_0\left[1-2\left(1+\frac{a_0}{R}\right)e^{-2R/a_0}\right. \\ &\left.\pm\left(\left(1-2\frac{a_0}{R}\right)S(R)+2\left(\frac{R}{a_0}+1\right)e^{-R/a_0}\right)\right] \qquad (7.24)\end{aligned}$$

The potential curve for the symmetrical superposition ε_+ that is plotted in Fig. 7.3 has a minimum as a function of R while ε_- is purely repulsive. This is due to the fact that the symmetric wave function is bonding while the antisymmetric wave function is antibonding because ψ_+ is larger than ψ_- in the region between the nuclei where ψ_- changes sign which is illustrated in Fig. 7.2

The variational results for the binding energy -1.76 eV and the nuclear distance 1.32 Å are in fact rather inaccurate. The exact results are -2.79 eV and 1.06 Å, respectively. This shows that the variational approach used here is not really appropriate.

7.4 The H_2 molecule

Now we address the simplest neutral system in chemistry where a bonding between two atoms occurs: the hydrogen molecule H_2. The Hamiltonian is

$$\begin{aligned} H =& \frac{\mathbf{p^2}_1}{2m_e} + \frac{\mathbf{p^2}_2}{2m_e} - \frac{e^2}{|\mathbf{x}_1 - \mathbf{X}_1|} - \frac{e^2}{|\mathbf{x}_1 - \mathbf{X}_2|} \\ & - \frac{e^2}{|\mathbf{x}_2 - \mathbf{X}_1|} - \frac{e^2}{|\mathbf{x}_2 - \mathbf{X}_2|} + \frac{e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} + \frac{e^2}{|\mathbf{X}_1 - \mathbf{X}_2|} \end{aligned} \tag{7.25}$$

Although the Hamiltonian is still relatively easy, such a system can not been solved exactly, as has already be mentioned in the context of the helium atom (see sect. 6.3). This is due to the electron-electron interaction which is the second-last term on the right-hand side of eq. 7.25. Qualitatively, we can already say that this term leads to an extra repulsion. This additional repulsion will lower the binding energy, so that one will get in comparison to the H_2^+ ion

$$|E_{\mathrm{H}_2}| \; < \; 2|E_{\mathrm{H}_2^+}| \,. \tag{7.26}$$

We will use the H_2 molecule in order to address the basics in the quantum mechanical description of chemical bonding. There are basically two methods two treat the chemical bonding of the H_2 molecule, the methods of molecular orbitals and the Heitler-London method. We will present both of them.

However, first we note that the total electronic wave function of the H_2 molecule describes two electrons which means that we have to take into account the quantum statistics of the electrons. The Hamiltonian (7.25) does not depend on the spin state of the electrons. Therefore, the two-electron wave function solving this Hamiltonian can be separated in a spatial and a spin part, as already shown in chapter 6.

$$\Psi = \psi(\mathbf{x}_1, \mathbf{x}_2)\, |m_{s_1} m_{s_2}\rangle \tag{7.27}$$

The spin part consists of a linear combination of the four spin states

$$|\uparrow\uparrow\rangle\,,\; |\uparrow\downarrow\rangle\,,\; |\downarrow\uparrow\rangle\,,\; |\downarrow\downarrow\rangle \;\;. \tag{7.28}$$

The $\{s, m\}$ or triplet-singlet representation based on the eigenkets of the total spin operator $\vec{S}^2 = (\vec{S}_1 + \vec{S}_2)^2$ and its component S_z along the z-axis:

$$\begin{aligned} |s = 1, m = 1\rangle &= |\uparrow\uparrow\rangle \;\;, \\ |s = 1, m = 0\rangle &= \left(\frac{1}{\sqrt{2}}\right) (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \;\;, \\ |s = 1, m = -1\rangle &= |\downarrow\downarrow\rangle \;\;, \\ |s = 0, m = 0\rangle &= \left(\frac{1}{\sqrt{2}}\right) (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \;\;, \end{aligned} \tag{7.29}$$

The $|s = 1, m = \pm 1, 0\rangle$ states are referred to as the spin triplet, while $|s = 0, m = 0\rangle$ is called the spin singlet state.

The Pauli exclusion principle requires that the total wave function Ψ is antisymmetric with respect to the interchange of both particles. The spin singlet state is antisymmetric; hence it has to be coupled with a symmetric spatial wave function. The triplet state, on the other hand, is symmetric with respect to the interchange of the two particles; therefore the spatial part has to be antisymmetric. Hence the eigenstates can be characterized according to their spin state, and triplet and singlet states will in general have different eigenenergies since their corresponding spatial parts have not the same symmetry. It is important to recall that this splitting is a consequence of a spin-independent Schrödinger equation.

7.4.1 Molecular orbitals

We first construct the singlet state in the method of molecular orbitals which corresponds to the product of the symmetric wavefunction ψ_+ (7.15) of the H_2^+ molecule,

$$\begin{aligned}\Psi_s(1,2) = &\quad [\psi_1(\mathbf{x_1}) + \psi_2(\mathbf{x_1})]\,[\psi_1(\mathbf{x_2}) + \psi_2(\mathbf{x_2})]\,\chi_{\text{singlet}}\,\frac{1}{2(1+S(R))} && (7.30)\\ = &\quad \psi_+(\mathbf{x_1})\psi_+(\mathbf{x_2})\,\chi_{\text{singlet}}, && (7.31)\end{aligned}$$

where χ_{singlet} describes the spin singlet part of the wave function. The expression in the square bracket will be denoted as a molecular orbital that is constructed as a *linear combination of atomic orbitals*. Therefore this approach is also called the LCAO method.

A triplet state can be constructed by using the antisymmetric combination of the H_2^+ states

$$\Psi_t(1,2) \;=\; [\psi_+(\mathbf{x_1})\psi_-(\mathbf{x_2}) \;-\; \psi_-(\mathbf{x_1})\psi_+(\mathbf{x_2})]\,\chi_{\text{trip}}\,\frac{1}{\sqrt{2}}\,. \qquad (7.32)$$

This state includes anti-bonding functions and does therefore not correspond to the ground state.

The approach (7.32) is in fact not too well-suited for describing the hydrogen molecule. For short distances R of the two protons the electronic wave function should rather be like the helium wave function (6.27) instead of being a product of $1s$ hydrogen wave functions. Furthermore, if we consider the singlet wavefunction (7.32) with all terms explicitly multiplied,

$$\begin{aligned}\Psi_s(1,2) = &\big[(\psi_1(\mathbf{x_1})\psi_1(\mathbf{x_2}) \;+\; \psi_2(\mathbf{x_1})\psi_2(\mathbf{x_2}))\\ &\;+\; (\psi_1(\mathbf{x_1})\psi_2(\mathbf{x_2})) \;+\; \psi_1(\mathbf{x_2})\psi_2(\mathbf{x_1}))\big]\chi_{\text{singlet}}\,\frac{1}{2(1+S(R))}\,, \qquad (7.33)\end{aligned}$$

we realize that in the two terms of the first bracket both electrons are concentrated at the same atom describing a situation that corresponds to a proton plus a hydrogen anion H^-. For large distances of the two nuclei this term should be absent, however, since two hydrogen atoms $H + H$ are energetically more favorable that $p{+}H^-$. Still the wave function (7.32) gives a good upper bound for the total energy of the hydrogen molecule because at the molecular distance of the two atoms the approach is better than for very short and very large distances.

7.4.2 Heitler-London method

In the Heitler-London method, the singlet and triplet wave fucntions are written as symmetric and anti-symmetric combinations of atomic states,

$$\Psi_s(1,2) = \frac{1}{\sqrt{2(1+S^2)}}\,[\psi_1(\mathbf{x_1})\psi_2(\mathbf{x_2}) \;+\; \psi_2(\mathbf{x_1})\psi_1(\mathbf{x_2})]\,\chi_{\text{singlet}} \tag{7.34}$$

$$\Psi_t(1,2) = \frac{1}{\sqrt{2(1-S^2)}}\,[\psi_1(\mathbf{x_1})\psi_2(\mathbf{x_2}) \;-\; \psi_2(\mathbf{x_1})\psi_1(\mathbf{x_2})]\,\chi_{\text{triplet}}\,. \tag{7.35}$$

For large distances, this ansatz describes separated hydrogen atoms, but for very small distances the same problems as for the molecular orbital method (7.32) occurs. The quantitative difference in the total energies for the equilibrium molecular distance between the orbitals (7.32) and (7.34) is in fact small.

The expectation values

$$\varepsilon_{s,t} \equiv \langle H\rangle_{s,t} \tag{7.36}$$

of the Hamiltonian (7.25) in the singlet and triplet state can be separated into different parts. First of all there are the one particle terms corresponding to isolated hydrogen atoms

$$\begin{aligned}\sum_{i=1}^{2}\int d^3x_i\,\psi_i^*(\mathbf{x}_i)\left(-\frac{\hbar^2}{2m}\nabla_i^2 + \frac{e^2}{|\mathbf{x}_i-\mathbf{X}_i|}\right)\psi_i(\mathbf{x}_i) &= \sum_{i=1}^{2}\int d^3x_i\,\psi_i^*(\mathbf{x}_i)E_0\psi_i(\mathbf{x}_i)\\ &= 2E_0\end{aligned} \tag{7.37}$$

Furthermore, there is the so-called Coulomb term which corresponds to the classical Coulomb energy of the charge distribution

$$\begin{aligned}Q =& \int d^3x_1\int d^3x_2\,|\psi_1(\mathbf{x}_1)|^2|\psi_2(\mathbf{x}_2)|^2\\ &\times\left[\frac{e^2}{|\mathbf{x}_1-\mathbf{x}_2|}-\frac{e^2}{|\mathbf{x}_1-\mathbf{X}_2|}-\frac{e^2}{|\mathbf{x}_2-\mathbf{X}_1|}+\frac{e^2}{|\mathbf{X}_1-\mathbf{X}_2|}\right]\\ =& -2\int d^3x_1\,|\psi_1(\mathbf{x}_1)|^2\frac{e^2}{|\mathbf{x}_1-\mathbf{X}_2|} \;+\; \frac{e^2}{R}\\ &-\int d^3x_1\int d^3x_2\,|\psi_1(\mathbf{x}_1)|^2\frac{e^2}{|\mathbf{x}_1-\mathbf{x}_2|}|\psi_2(\mathbf{x}_2)|^2\,.\end{aligned} \tag{7.38}$$

Here the terms are two times the Coulomb interaction of the electron centered at the first atom with the second nucleus which equals the interaction of the electron centered at the second atom with the first nucleus, and then the Coulomb repulsion between the two nuclei and the two electrons.

Finally, because the electrons are indistinguishable, there is so-called exchange term

$$\begin{aligned}A =& \int d^3x_1\int d^3x_2\,\psi_1^*(\mathbf{x_1})\psi_2^*(\mathbf{x_2})\psi_1(\mathbf{x_2})\psi_2(\mathbf{x_1})\\ &\times\left[\frac{e^2}{|\mathbf{x}_1-\mathbf{x}_2|}-\frac{e^2}{|\mathbf{x}_1-\mathbf{X}_2|}-\frac{e^2}{|\mathbf{x}_2-\mathbf{X}_1|}+\frac{e^2}{|\mathbf{X}_1-\mathbf{X}_2|}\right]\\ =& -2S\int d^3x_1\,\psi_1^*(\mathbf{x}_1)\psi_2(\mathbf{x}_1)\frac{e^2}{|\mathbf{x}_1-\mathbf{X}_2|} \;+\; S^2\frac{e^2}{R}\\ &+\int d^3x_1\int d^3x_2\,\psi_1^*(\mathbf{x}_1)\psi_2(\mathbf{x}_1)\frac{e^2}{|\mathbf{x}_1-\mathbf{x}_2|}\psi_2^*(\mathbf{x}_2)\psi_1(\mathbf{x}_2)\,.\end{aligned} \tag{7.39}$$

This term is a consequence of the quantum nature of the electrons and their Coulomb interaction. In total, the energies of the singlet and triplet state are given by

$$\varepsilon_s = 2E_0 \; + \; \frac{Q + A}{1 + S^2} \tag{7.40}$$

$$\varepsilon_t = 2E_0 \; + \; \frac{Q - A}{1 - S^2} \tag{7.41}$$

Both the Coulomb energy Q as well as the exchange energy A are functions of the distance of the nuclei. The Coulomb energy is positive, but due to the compensation of the different terms relatively small. The exchange energy is except for small distances negative and overcompensates the Coulomb energies, and therefore the singlet state is bonding. *This shows that it is in fact the exchange energy that is mainly responsible for the bonding in the* H_2 *molecule.*

Using the Heitler-London method, a molecular distance of $R = 0.80$ Å results which should be compared to the experimental value of $R = 0.74$ Å. Furthermore, the Heitler-London method also leads to an underestimation of the H_2 binding energy which is given by the difference of the energy at the equilibrium distance and the energy for infinite separation ($= 2E_0$ here). The Heitler-London value for the H_2 binding energy is 3.14 eV which is much smaller than the experimentally derived value of 4.75 eV.

7.4.3 Covalent-ionic resonance

According to the ansatz (7.34) for the wave function in the Heitler-London method, the two electrons are always located at opposite nuclear positions. This correspond to a covalent type of bonding, i.e, we can rename the spatial part of the wave function as

$$\Psi_{\rm cov}(1,2) = \; \psi_1(\mathbf{x_1})\psi_2(\mathbf{x_2}) \; + \; \psi_2(\mathbf{x_1})\psi_1(\mathbf{x_2}) \; . \tag{7.42}$$

However, it might well be that there is a certain probability that both electrons are located at the same atom. This would correspond to the situation of a negatively charged hydrogen atom plus a proton. Therefore this situation is called *ionic* which is described by

$$\Psi_{\rm ion}(1,2) = \; \psi_1(\mathbf{x_1})\psi_1(\mathbf{x_2}) \; + \; \psi_2(\mathbf{x_1})\psi_2(\mathbf{x_2}) \; , \tag{7.43}$$

where we have taken into account that both hydrogen nuclei have to be treated equivalently. The true situation in the hydrogen molecule is neither fully covalent nor fully ionic. Hence a more appropriate description of the H_2 electronic wave function will be a superposition of the covalent and the ionic function,

$$\Psi(1,2) = N(\Psi_{\rm cov}(1,2) + c\Psi_{\rm ion}(1,2)) \; , \tag{7.44}$$

where N is the normalisation constant and c is a parameter that should be determined in such a way that the expection value of the energy becomes minimal. In fact, all the three approaches for the construction of the H_2 wave function can be regarded as special cases of a more general ansatz in which we add to any of the atomic wave function a contribution of the wave function that is localized at the other atom, i.e., we make the substitution

$$\psi_1 \; \rightarrow \psi_1 + b\,\psi_2 \; , \qquad \qquad \psi_2 \; \rightarrow \psi_2 + b\,\psi_1 \; . \tag{7.45}$$

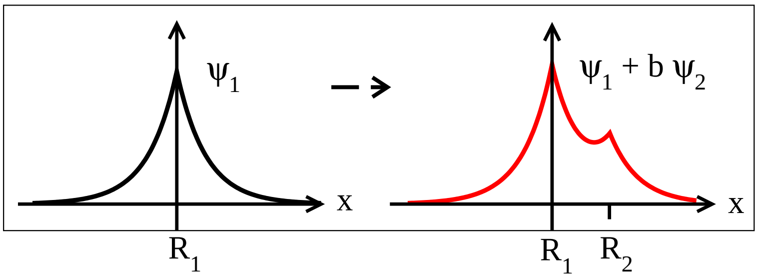


Figure 7.4: Illustration of the substitution of the atomic wave function by a wave function with an additional contribution from the second atom.

This substitution is illustrated in Fig. 7.4. Thus our new ansatz for the singlet wave function can be written as

$$\begin{aligned}\Psi_s(1,2) &= N\,[(\psi_1(\mathbf{x_1}) + b\,\psi_2(\mathbf{x_1}))\,(\psi_2(\mathbf{x_2}) + b\,\psi_1(\mathbf{x_2})) \\ &\quad + (\psi_1(\mathbf{x_2}) + b\,\psi_2(\mathbf{x_2}))\,(\psi_2(\mathbf{x_1}) + b\,\psi_1(\mathbf{x_1}))]\,\chi_{\text{singlet}} \\ &= N\,[(1+b^2)(\psi_1(\mathbf{x_1})\psi_2(\mathbf{x_2}) + \psi_1(\mathbf{x_2})\psi_2(\mathbf{x_1})) \\ &\quad + 2b(\psi_1(\mathbf{x_1})\psi_1(\mathbf{x_2}) + \psi_2(\mathbf{x_1})\psi_2(\mathbf{x_2}))]\,\chi_{\text{singlet}}\,. \qquad (7.46)\end{aligned}$$

Comparing eq. (7.46) with the previous expressions for the wave function, we see that molecular orbital expression (7.32) results from eq. (7.46) by setting $b = 1$, the Heitler-London method (7.34) is recovered for $b = 0$, and the mixed covalent-ionic description of eq. (7.44) can be obtained for

$$\frac{2b}{1+b^2} = c\,. \qquad (7.47)$$

7.5 Hybridization

So far we have assumed that in the bonding only s functions are involved. All bonding orbitals were constructed from a superposition of such s functions. However, there are situations where by mixing s and p functions the overlap between adjacent atomic orbitals can be increased and where the resulting energy gain due to the greater overlap in the bonding region between the two nuclei overcompensates the energetic price that is to pay in order to partially occupy the energetically higher lying p orbital. Such a scenario is illustrated in Fig. 7.5.

The Li_2 molecule is an example where such orbitals do occur. The outermost electron is a $2s$ electrons, but it takes relatively little energy to excite this electron to a $2p$ state. As trial wave function for the bonding, we assume an orbital of the form

$$|\psi\rangle = |2s\rangle + \lambda\,|2p\rangle\ . \qquad (7.48)$$

Using a variational approach, one finds that $\lambda \approx 0.3$ leads to a minimum in the Li_2 binding energy which means that the probability of the electron being in the $2p$ state is $0.3^2 \approx 10\%$. This process of forming orbitals from a combination of states with different angular momenta is called *hybridization*.

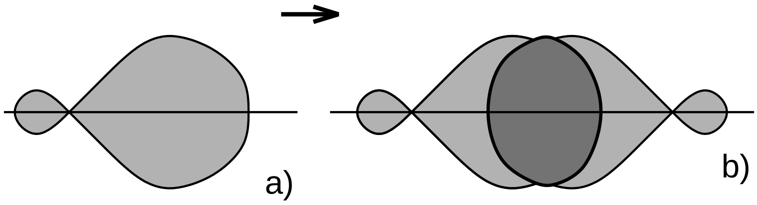


Figure 7.5: A hybridized state formed from a superposition of a s and a p state shown in (a) can lead to a greater overlap in a molecular bond (b).

A particular important atom forming the basis of organic chemistry is the carbon atom that has the electronic ground state $(1s)^2(2s)^2(2p)^2$. In this state, carbon is divalent. However, the configuration $(1s)^2(2s)(2p)^3$ is very close in energy to the ground state. In this state, carbon has a valency of four, and it is is this state that carbon forms most of its compounds. For example, in methane, CH_4, the $2s$ and $2p$ states of carbon a completely hybridized, forming a so-called sp^3 hybridization.

$$\begin{aligned}
|1\rangle &= \frac{1}{2}(|2s\rangle + |2p_x\rangle + |2p_y\rangle + |2p_z\rangle) \\
|2\rangle &= \frac{1}{2}(|2s\rangle + |2p_x\rangle - |2p_y\rangle - |2p_z\rangle) \\
|3\rangle &= \frac{1}{2}(|2s\rangle - |2p_x\rangle + |2p_y\rangle - |2p_z\rangle) \\
|4\rangle &= \frac{1}{2}(|2s\rangle - |2p_x\rangle - |2p_y\rangle + |2p_z\rangle) \qquad (7.49)
\end{aligned}$$

Each of these orbitals binds one hydrogen atom. The methane moclecule has a tetragonal structure with a bond angle of exactly $109.6°$ following from symmetry. The ethylene molecule, C_2H_4, on the other hand, has a planar structure. The orbitals of the carbon atom are hybridized as folllows:

$$\begin{aligned}
|1\rangle &= |2p_z\rangle \\
|2\rangle &= \sqrt{\frac{1}{3}}\,|2s\rangle + \sqrt{\frac{2}{3}}\,|2p_x\rangle \\
|3\rangle &= \sqrt{\frac{1}{3}}\,|2s\rangle - \sqrt{\frac{1}{6}}\,|2p_x\rangle + \sqrt{\frac{1}{2}}\,|2p_y\rangle \\
|4\rangle &= \sqrt{\frac{1}{3}}\,|2s\rangle - \sqrt{\frac{1}{6}}\,|2p_x\rangle - \sqrt{\frac{1}{2}}\,|2p_y\rangle \qquad (7.50)
\end{aligned}$$

The first state is oriented along the $\pm z$ direction while the other three orbitals form angles of $120°$ in the xy plane. The electrons in the orbitals $|3\rangle$ and $|4\rangle$ bind to hydrogens, while those in $|1\rangle$ and $|2\rangle$ form bonds with the other carbon atom (see Fig. 7.6) whose orbitals $|n'\rangle$ are mirror images of those of the first carbon atom. The orbitals $|2\rangle$ and $|2'\rangle$ of the two carbon atoms correspond to a singlet state such as (7.34) having a large overlap. Since this singlet state is now occupied, the second carbon electron could populate a triplett state with the other carbon atom. However, this antibonding state is energetically rather unfavorable. It is more favorable if the second electron is in a orbital that is perpendicular to the C-C axis and the xy plane. Therefore this

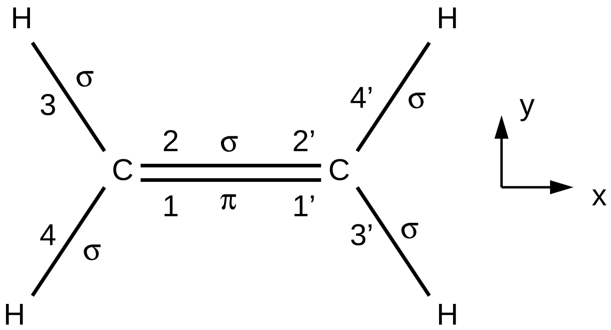


Figure 7.6: The structure of the ethylene molecule, C_2H_4, including the σ and the π bonds.

electron is in a $|2p_z\rangle$ state and forms a spin singlet with the other $|2p_z\rangle$ electron of the second carbon atom. While there is not as much overlap as for the $|2\rangle$ states, this configuration still contributes to the bonding.

Note that the angular momentum of the bond formed by the orbitals $|2\rangle$ and $|2'\rangle$ along the C-C axis, the z axis, is zero. Such a bond is called a σ bond (corresponding to a s state). In contrast, the angular momentum of the $|2p_z\rangle$ along the z axis is 1 so that the electrons in the bond formed by the two $|2p_z\rangle$ states have a angular momentum of magnitude 1. This bond is called a π bond (as in p state). π bonds are in general weaker than σ bonds because of the smaller overlap. To a first approximation, the bond energies of σ and π bonds can be added in the total energy because of their localisation.

7.6 Molecular orbitals and symmetry properties

In the last chapter, in section 6.4, we have introduced some basic concepts of group and representation theory and have illustrated them using a very simple example. Here we will now apply these concepts in order to construct one-electron functions for a more complex molecule, water (H_2O).

In the previous section, we constructed the molecular orbitals for ethylene more or less *ad hoc*. For the water molecule, we will do this more schematically. Again, we will start with atomic orbitals, but we will try to determine the cofficients in the LCAO approach based on group theoretical consideration that take into account the symmetry of the molecule. As a first step, we have to determine the symmetry group of the water molecule.

In Fig. 7.7, the water molecule and its symmetry operations are illustrated which form the symmetry group C_{2v}. Both the xz and the yz plane are symmetry planes at which the reflections σ_v' and σ_v, respectively, can be performed. In addition, the z-axis corresponds to a twofold rotation axis.

Both the multiplication and the character table of the symmetry group C_{2v} are displayed in Teb. 7.1. This is an Abelian group with four elements. Consequently, there are four one-dimensional representations.

In order to derive a set of molecular orbitals, we have first to determine which atomic orbitals are involved. As before, for the hydrogen atom we just consider the $1s$

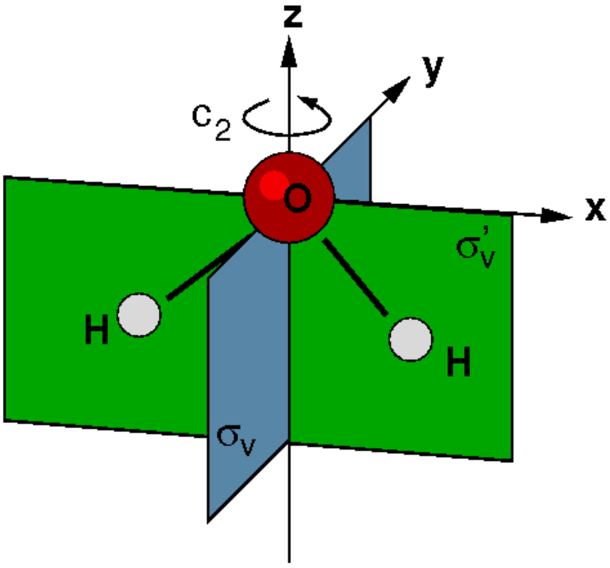


Figure 7.7: The H_2O molecules and its symmetry operations.

orbitals $\left|1s_1^{\mathrm{H}}\right\rangle \left|1s_2^{\mathrm{H}}\right\rangle$. As far as the oxygen atom is concerned, the $1s$ states form a filled shell that does practically not contribute to any bonding. Hence we just consider the $2s$ and $2p$ orbitals $\left|2s^{\mathrm{O}}\right\rangle$, $\left|2p_x^{\mathrm{O}}\right\rangle$, $\left|2p_y^{\mathrm{O}}\right\rangle$ and $\left|2p_z^{\mathrm{O}}\right\rangle$. In principle, we should now set up the matrix representation of the symmetry operation using the six atomic orbitals just mentioned. However, as it turns out, the hydrogen and the oxygen orbitals only weakly couple. Therefore, in a first approximation, we can regard the hydrogen-derived and the oxygen derived orbitals separately, so that the reducible representation matrices already break up in a (2×2) block related to the hydrogen orbitals that we denote by D^{H} and in a (4×4) block related to the oxygen orbitals that we denote by D^{O}.

In the following, we will omit the superscripts denoting the elements. If we for example consider the effect of the rotation c_2 on the hydrogen orbitals, we find that

$$c_2 \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \tag{7.51}$$

since the rotation by 180° exchanges the two $1s$ orbitals.

In total, we obtain the following representation matrices and the corresponding

Table 7.1: Multiplication and character table of the symmetry group C_{2v}.

C_{2v}	E	c_2	σ_v	σ_v'		E	c_2	σ_v	σ_v'
E	E	c_2	σ_v	σ_v'	A_1	1	1	1	1
c_2	c_2	E	σ_v'	σ_v	A_2	1	1	-1	-1
σ_v	σ_v	σ_v'	E	c_2	B_1	1	-1	1	-1
σ_v'	σ_v'	σ_v	c_2	E	B_2	1	-1	-1	1

characters

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad c_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_v = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_v' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.52}$$

$$\chi(E) = 2 \qquad \chi(c_2) = 0 \qquad \chi(\sigma_v) = 0 \qquad \chi(\sigma_v') = 2 \tag{7.53}$$

As in section 6.4.3, we can use simple combinatorial considerations in order to find the irreducible reprensentation included in the reprensentation D^{H} that is displayed in (7.52). Using the characters as illustrated in Tab. 7.2, we find that

$$D^{\mathrm{H}} = A_1 + B_2 \; . \tag{7.54}$$

In order to find out the wave functions that correspond to these irreducible representations, we will use the projection operators introduced in Eq. (6.46)

$$P_i = \frac{1}{h} \sum_T \chi_i(T^{-1}) D(T) \; .$$

Applying this formula, we find that

$$\begin{aligned} P_{A_1} &= \frac{1}{4} \left\{ 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \; . \end{aligned} \tag{7.55}$$

Equivalently, we obtain

$$\begin{aligned} P_{B_2} &= \frac{1}{4} \left\{ 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \; . \end{aligned} \tag{7.56}$$

If we apply the projection operators to the basis of the atomic $1s$ functions, we obtain

$$P_{A_1} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} s_1 + s_2 \\ s_1 + s_2 \end{pmatrix} \tag{7.57}$$

$$P_{B_2} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} s_1 - s_2 \\ -s_1 + s_2 \end{pmatrix} \tag{7.58}$$

Table 7.2: Characters of the representation D^{H} and of the corresponding irreducible representations that are contained in D^{H}.

Representation	E	c_2	σ_v	σ_v'
D^{H}	2	0	0	2
A_1	1	1	1	1
B_2	1	-1	-1	1
$A_1 + B_2$	2	0	0	2

Thus we obtain the following basis functions of the irreducible representations A_1 and B$_2$:

$$
\begin{aligned}
A_1 : |1\rangle &= \frac{1}{2}(|1s_1\rangle + |1s_2\rangle) && (7.59)\\
B_2 : |2\rangle &= \frac{1}{2}(|1s_1\rangle - |1s_2\rangle) && (7.60)\\
& && (7.61)
\end{aligned}
$$

This is nothing else but the symmetric and antisymmetric combination of the hydrogen $1s$ orbitals, similar to the case of the H_2^+ molecule (Eq. (7.15)), however, here we have derived it using the technology of group and representation theory.

It is also interesting to consider what happens if we apply Eq. (6.46) for a irreducible representation that is not contained in D^{H}. For example, for the irreducible representation A_2 we obtain

$$
\begin{aligned}
P_{A_2} &= \frac{1}{4}\left\{1\cdot\begin{pmatrix}1&0\\0&1\end{pmatrix} + 1\cdot\begin{pmatrix}0&1\\1&0\end{pmatrix} - 1\cdot\begin{pmatrix}0&1\\1&0\end{pmatrix} - 1\cdot\begin{pmatrix}1&0\\0&1\end{pmatrix}\right\}\\
&= \frac{1}{4}\begin{pmatrix}0&0\\0&0\end{pmatrix}\,. \qquad (7.62)
\end{aligned}
$$

This is an confirmation that A_2 is not included in D^{H}.

Let us now consider the slightly more complicated case of the four oxygen-derived orbitals. In fact, it turns out that the representation matrices D^{O} in this case are rather simple since all oxygen orbitals are centered at the origin which remains fixed under all symmetry operations of the group C_{2v}. All representation matrices are diagonal leading to the following relations:

$$
\begin{aligned}
&E\begin{pmatrix}2s\\2p_x\\2p_y\\2p_z\end{pmatrix} = \begin{pmatrix}2s\\2p_x\\2p_y\\2p_z\end{pmatrix}, \qquad
c_2\begin{pmatrix}2s\\2p_x\\2p_y\\2p_z\end{pmatrix} = \begin{pmatrix}2s\\-2p_x\\-2p_y\\2p_z\end{pmatrix},\\
&\sigma_v\begin{pmatrix}2s\\2p_x\\2p_y\\2p_z\end{pmatrix} = \begin{pmatrix}2s\\-2p_x\\2p_y\\2p_z\end{pmatrix}, \qquad
\sigma_v'\begin{pmatrix}2s\\2p_x\\2p_y\\2p_z\end{pmatrix} = \begin{pmatrix}2s\\2p_x\\-2p_y\\2p_z\end{pmatrix}, \qquad (7.63)\\
&\chi(E) = 4 \qquad \chi(c_2) = 0 \qquad \chi(\sigma_v) = 2 \qquad \chi(\sigma_v') = 2 \qquad (7.64)
\end{aligned}
$$

The irreducible representations that are contained in D^{O} can again be obtained by combinatorial considerations or by applying Eq. (6.45). This will be left as an exercise for the reader. It turns that

$$
D^{\mathrm{O}} = 2A_1 + B_1 + B_2\,. \qquad (7.65)
$$

The wave functions that belong to the different irreducible representations are listed in table 7.3. Thus it becomes evident which functions belong to the same irreducible representations. Why is this helpful?

The wave functions listed in table 7.3 do not necessarily correspond to molecular orbitals of the water molecule. Group theory does not tell us anything about the energy levels,h owever, in order to construct specific molecular orbitals, we only have to

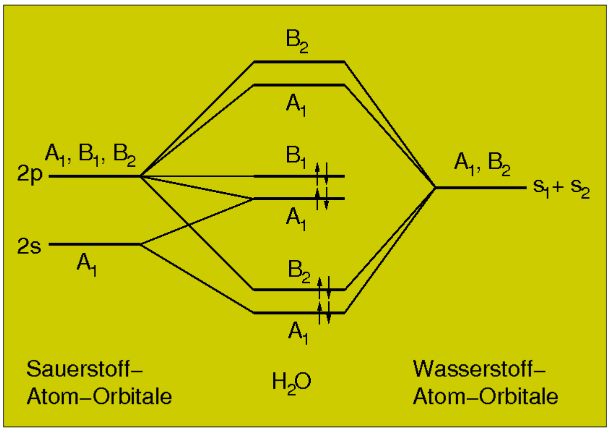


Figure 7.8: Energy scheme of H_2O.

consider wave functions that belong to the same irreducible representation. The easiest case is the state belonging to the B_1 representation where the molecular orbitals just consists of the oxygen p_y orbitals. Since this orbital is oriented perpendicular to the plane spanned by the hydrogen and oxygen atoms, it does not contribute to the bonding, i.e., it corresponds to a non-bonding orbital of an intermediate energy.

In the case of the B_2 representation, two wave functions have to be considered, the oxygen p_x state and the antisymmetric hydrogen-derived state $|2\rangle$. The coupling between these two states leads to a bonding and an anti-bonding state. For A_1, there are three wave functions that lead to three different energy levels. The qualitative energy scheme of the water molecule illustrating the energetic arrangement of the state is displayed in Fig. 7.8.

How are the energy levels shown in Fig. 7.8 related to hybridized orbitals shown in the previous section? It is important to realize that the hybridized orbitals are not necessarily eigenfunctions of the Hamiltonian. If we for example consider the non-bonding A_1 and B_2 orbitals of the water molecule, it is immediately clear that they are not equivalent since they belong to different irreducible representation; furthermore, their energies are different. However, from these two orbitals two hybridized orbitals can be constructed that are equivalent and that correspind to two so-called lone pairs, i.e., these are orbitals that do not contribute to the intra-molecular bonding. However, these molecules play a very important role in the interaction with other molecules, i.e., in the inter-molecular bonding. They are in fact responsible for the hydrogen bonding between water molecules and lead to the well-known hexagonal structure of ice.

Table 7.3: Basis functions of the irreducible representations of the water molecule.

Representation	O orbitals	H orbitals
A_1	$\lvert 2s^{\rm O}\rangle$, $\lvert 2p_z^{\rm O}\rangle$	$\lvert 1\rangle = \frac{1}{2}(\lvert 1s_1\rangle + \lvert 1s_2\rangle)$
A_2		
B_1	$\lvert 2p_y^{\rm O}\rangle$	
B_2	$\lvert 2p_x^{\rm O}\rangle$	$\lvert 2\rangle = \frac{1}{2}(\lvert 1s_1\rangle - \lvert 1s_2\rangle)$

7.7 Energy levels of diatomic molecules: vibrations and rotations

In the preceding sections we have focused on the electronic structure of molecules and the resulting potential energy surface. Using this potential energy surface, the energy levels of the motion of the nuclei can be determined using the atomic Schrödinger equation (7.11). Here we illustrate how the vibrational and rotational states of a diatomic molecule can be determined. In principle, a diatomic molecule has six degrees of freedom. Thus the Schrödinger equation is given by

$$\left[-\frac{\hbar^2}{2M_1}\nabla_1^2 - \frac{\hbar^2}{2M_2}\nabla_2^2 + V(r)\right]\Psi(\mathbf{X_1},\mathbf{X_2}) \;=\; E_{\text{tot}}\,\Psi(\mathbf{X_1},\mathbf{X_2})\;. \qquad (7.66)$$

However, the potential V only depends on the distance $r = |\mathbf{X_1} - \mathbf{X_2}|$ of the two atoms and not on their orientation. Therefore, by introducing center-of-mass coordinates $\mathbf{X} = (M_1\mathbf{X}_1 + M_2\mathbf{X}_2)/(M_1 + M_2)$ and relative coordinates $\mathbf{x} = \mathbf{X}_1 - \mathbf{X}_2$, the Schrödinger equation can be rewritten after some rearrangement as

$$\left[-\frac{\hbar^2}{2M}\nabla_{\mathbf{X}}^2 - \frac{\hbar^2}{2m}\nabla_{\mathbf{x}}^2 + V(r)\right]\Psi(\mathbf{X},\mathbf{x}) \;=\; E_{\text{tot}}\,\Psi(\mathbf{X},\mathbf{x})\;. \qquad (7.67)$$

Here, $m = M_1M_2/(M_1 + M_2)$ is the *reduced mass*. This form shows that the center-of-mass motion and the relative motion are independent from each other since the potential only acts on the relative coordinate. Using a product ansatz

$$\Psi(\mathbf{X},\mathbf{x}) \;=\; \Phi(\mathbf{X})\,\psi(\mathbf{x})\;, \qquad (7.68)$$

the center-of-mass motion can be separated from the relative motion of the two atoms resulting in a effective one-particle equations

$$\left[-\frac{\hbar^2}{2M}\nabla_{\mathbf{X}}^2\right]\Phi(\mathbf{X}) = E_{\mathbf{X}}\,\Phi(\mathbf{X})\;, \qquad (7.69)$$

$$\left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{x}}^2 + V(r)\right]\psi(\mathbf{x}) = E\,\psi(\mathbf{x})\;. \qquad (7.70)$$

The center of mass moves as a free particle according to eq. (7.69). Again, it is important to note that the effective potential energy does only depend on the distance $r = |\mathbf{x}|$ of the two atoms and not on their orientation. Thus we have a central potential problem analogous to the one already treated in section **??** (see eq. (4.101). Introducing the angular momentum operator $\mathbf{L}$, the Schrödinger equation (7.70) for the relative motion can be rewritten as

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + \frac{\mathbf{L}^2}{2mr^2} + V(r)\right]\psi(r,\vartheta,\varphi) = E\psi(r,\vartheta,\varphi)\;. \qquad (7.71)$$

Using the ansatz $\psi(\mathbf{x}) = R_{nl}(r)Y_{lm}(\vartheta,\varphi)$ we obtain

$$\left[-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right) + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r)\right]R_{nl}(r) = ER_{nl}(r)\;. \qquad (7.72)$$

The potential terms can be merged into an effective potential

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2mr^2}\;. \qquad (7.73)$$

As in section **??**, we perform the substitution

$$R_{nl}(r) \ = \ \frac{U_{nl}}{r} \ , \tag{7.74}$$

resulting in a simplified one-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + V_{\text{eff}}(r)\right] u_{nl}(r) \ = \ E\, u_{nl}(r) \tag{7.75}$$

For small angular momentum quantum numbers l the effective potential $V_{\text{eff}}(r)$ has a minimum at the position r_l which in general depends on the rotational quantum number. Now we expand $V_{\text{eff}}(r)$ close to r_l in a Taylor series:

$$\begin{aligned} V_{\text{eff}}(r) &= V_{\text{eff}}(r_l) + \frac{1}{2}m\omega_l^2(r-r_l)^2 + \dots \\ &= \varepsilon(r_l) + \frac{\hbar^2 l(l+1)}{2mr_l^2} + \frac{1}{2}m\omega_l^2(r-r_l)^2 + \dots \ , \end{aligned} \tag{7.76}$$

where we introduced the notations $\varepsilon(r_l) = V(r_l)$ and

$$m\omega_l^2 \ = \ \left.\frac{d^2V_{\text{eff}}(r)}{dr^2}\right|_{r=r_l} \ . \tag{7.77}$$

For small displacements, the Taylor expansion (7.76) can be terminated after the quadratic term. With the definition $x \equiv r - r_l$ eq. (7.75) then becomes the Schrödinger equation of a one-dimensional harmonic oscillator:

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \varepsilon(r_l) + \frac{\hbar^2 l(l+1)}{2mr_l^2} + \frac{m\omega_l^2}{2}x^2\right] u(x) = Eu(x) \ . \tag{7.78}$$

This means that the eigenvalues are given by

$$E \ = \ \varepsilon(r_l) + \frac{\hbar^2 l(l+1)}{2I_l} + \hbar\omega_l\left(v + \frac{1}{2}\right) \ , \tag{7.79}$$

with the effective moment of inertia $I_l = mr_l^2$ and the vibrational quantum number v. These energy eigenvalues consist of the effective electronic energy, the rotational energy and the vibrational energy. The vibrations correspond to the fastest motion; therefore they have the highest energies with $\hbar\omega_l \approx 0.5$ eV. Because of this high vibrational frequency, at room temperature they are *frozen* which means that they are not populated and therefore do not contribute to the specific heat of the hydrogen molecule. The rotations, on the other hand, have excitations energies that are much lower, $E_{\text{rot}} \geq 0.05$ eV.

In reality, the potential curve of a diatomic molecule as a function of the interatomic distance is not harmonic. For smaller distance than the equilibrium distance the Pauli repulsion between the electronic clouds leads to an enhanced repulsion while at large distances the potential curves becomes flat once the dissociation has occured. It is common practice to describe the potential curve of molecules in terms of a so-called Morse potential

$$\begin{aligned} V(r) \ &= D_e\left[\left(1 - e^{-\alpha(r-r_{eq})}\right)^2 \ - \ 1\right] \\ &= D_e\left[e^{-2\alpha(r-r_{eq})} \ - \ 2e^{-\alpha(r-r_{eq})}\right] \ , \end{aligned} \tag{7.80}$$

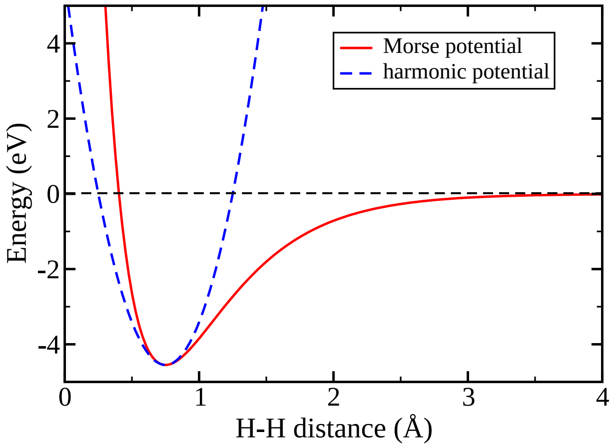


Figure 7.9: Morse potential curve of the H_2 molecule. The dashed line corresponds to the potential curve of a harmonic oscillator with the same curvature at the minimum.

which is plotted in Fig. 7.9 with parameters appropriate for the H_2 molecule. The inverse length α determines the range of the potential. It is related to the frequency and the reduced mass of a harmonic oscillator with the same curvature at the equilibrium position r_{eq} via

$$\alpha = \sqrt{\frac{m}{2D_e}}\,\omega \,. \tag{7.81}$$

In Fig. 7.9, the potential of the corresponding harmonic potential with the same curvature at the equilibrium position is also plotted. Close to the minimum there is a good agreement between the harmonic and the Morse potential, but for larger distances from the minimum the harmonic approximation breaks down. The energy levels of the Morse potential are given by

$$\begin{aligned} E_v &= D_e + \hbar\omega(v+\frac{1}{2}) - \chi_e\hbar\omega(v+\frac{1}{2})^2 \\ &= D_e + \hbar\omega(v+\frac{1}{2})\,[1-\chi_e(v+\frac{1}{2})]\,, \end{aligned} \tag{7.82}$$

where χ_e is the so-called anharmonicity constant

$$\chi_e = \frac{\hbar\omega}{4D_e}\,, \tag{7.83}$$

which is typically of the order of 0.01. The frequencies of the vibrations in the Morse potential are thus given by

$$\omega_v = \omega\,[1-\chi_e(v+\frac{1}{2})] \tag{7.84}$$

which means that the vibrational level spacing is linearly decreasing with the vibrational quantum number v. The expression (7.84) including the anharmonicity gives a more realistic description of the vibrational levels of molecules. Furthermore, the observed rotational levels also do not follow the simple relation given in eq. (7.79). This is caused by the vibrational-rotational coupling since vibrations and rotations are not really independent. First of all one notes that in the anharmonic potential the mean H-H distance does not only depend on the rotational state but also on the vibrational state: For larger vibrational quantum numbers v the mean distance R increases leading to a smaller moment of inertia. Furthermore, the molecule is not really a rotator with a fixed bond length. The centrifugal forces acting on the rotating molecules lead to a further elongation that depends on the vibrational and rotational state and results in a additional $\mathbf{L}^4$ term. Thus the total energy should be written as

$$E \;=\; \varepsilon(r_{l,v}) + \frac{\hbar^2 l(l+1)}{2I_{l,v}} + D_{l,v} l^2 (l+1)^2 + \hbar\omega_l \left(n + \frac{1}{2}\right) - \chi_e \hbar\omega_l (v + \frac{1}{2})^2 \; . \quad (7.85)$$

8 The many-electron problem and quantum chemistry

So far we have discussed the chemical bonding within an one-electron picture, i.e. we have just considered electronic orbitals and populated them with electrons. Total energies had been estimated so far only for two-electron systems such as the hydrogen molecule. We will now introduce electronic structure methods that allow to calculate the energies of many-electron systems. These are actually still only approximate methods, but quantum chemists have developed a whole machinery of different methods that can give accurate and reliable results. We note in passing that we will find out that the oribtals that we have used in the previous chapter are in fact only an useful auxiliary construction, but the basic quantity is the many-body wave function.

In this chapter, we will first discuss the basics about quantum chemistry methods based on Hartree-Fock theory, but we will also give a brief sketch of density functional theory which has become quite popular in recent years.

8.1 Hartree Theory

We start the sections about electronic structure methods with the so-called Hartree and Hartree-Fock methods. This does not only follow the historical development, but it also allows to introduce important concepts such as self-consistency or electron exchange and correlation. In this whole chapter we are concerned with ways of solving the time-independent electronic Schrödinger equation

$$H_{\rm el}\Psi(\vec{r}) = E_{\rm el}\Psi(\vec{r}) \ . \tag{8.1}$$

Here we have omitted the parametric dependence on the nuclear coordinates (c.f. (7.10)) for the sake of convenience. As already stated, except for the simplest cases there is no way to solve (8.1) in a close analytical form. Hence we have to come up with some feasible numerical scheme to solve (8.1). Mathematically, it corresponds to a second order partial differential equation. There are methods to directly integrate partial differential equations. However, if N is the number of electrons in the system, then we have to deal with a partial differential equation in $3N$ unknowns with N commonly larger than 100. This is completely intractable to solve. The way out is to expand the electronic wave function in some suitable, but necessarily finite basis set whose matrix elements derived from (8.1) can be conveniently determined. This will then convert the partial differential equation into a set of algebraic equations that are much easier to handle. Of course, we have to be aware that by using a finite basis set we will only find an approximative solution to the true many-body wave function. However, by increasing the size of the basis set we have a means to check whether our results are converged with respect to the basis set. Hence this corresponds to a controlled approximation because the accuracy of the calculations can be improved in a systematic way.

Furthermore, for the moment we are mainly interested in the electronic ground-state energy E_0. There is an important quantum mechanical principle – the *Rayleigh–Ritz variational principle* – that provides a route to find approximative solutions for the ground state energy. It states that the expectation value of the Hamiltonian in any state $|\Psi\rangle$ is always larger than or equal to the ground state energy E_0, i.e.

$$E_0 \leq \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \; . \tag{8.2}$$

Hence we can just pick some suitable guess for $|\Psi\rangle$. Then we know that $\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle$ will always be an upper bound for the true ground state energy. By improving our guesses for $|\Psi\rangle$, preferentially in a systematic way, we will come closer to the true ground state energy.

Before we proceed, we note that the potential term $V_{\text{nucl}-\text{el}}$ (7.5) acts as an effective external one-particle potential for the electrons. Hence we define the external potential for the electrons as

$$v_{ext}(\vec{r}) = -\sum_{\text{I}} \frac{Z_{\text{I}}e^2}{|\vec{r}-\vec{R}_{\text{I}}|} \; . \tag{8.3}$$

Now let us assume that the number of electrons in our system is N and that we have already determined the N lowest eigenfunctions $|\psi_i\rangle$ of the one-particle Schrödinger equation

$$\left\{-\frac{\hbar^2}{2m}\nabla^2 + v_{ext}(\vec{r})\right\}\psi_i(\vec{r}) = \varepsilon_i^o\psi_i(\vec{r}) \; . \tag{8.4}$$

Here we have completely neglected the electron-electron interaction. Still, we might simply consider the product wave function

$$\Psi_{\text{H}}(\vec{r}_1,\ldots,\vec{r}_{\text{N}}) = \psi_1(\vec{r}_1)\cdot\ldots\cdot\psi_{\text{N}}(\vec{r}_{\text{N}}) \; , \tag{8.5}$$

in which every one-particle state is only occupied once, as a first crude guess for the true many-particle wave function. Note that here we have not explicitly included the spin. For the sake of simplicity we consider two wave functions with the same spatial part, but orthogonal spin parts to be different.

We can now determine the expection value of the electronic Hamiltonian (7.9) using the wave function (8.5). Thus we obtain

$$\begin{aligned}\langle\Psi_{\text{H}}|H|\Psi_{\text{H}}\rangle &= \sum_{i=1}^{N}\int d^3r\psi_i^*(\vec{r})\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{ext}(\vec{r})\right)\psi_i(\vec{r}) \\ &+\frac{1}{2}\sum_{i,j=1}^{N}\int d^3rd^3r'\frac{e^2}{|\vec{r}-\vec{r}'|}|\psi_i(\vec{r})|^2|\psi_j(\vec{r}')|^2 + V_{\text{nucl}-\text{nucl}} \; .\end{aligned} \tag{8.6}$$

Now we would like to minimize the expectation value (8.6) with respect to more suitable single-particle functions $\psi_i(\vec{r})$ under the constraint that the wave functions are normalized. This is a typical variational problem with the constraint taken into account via Lagrange multipliers. If we consider the wave functions $\psi_i(\vec{r})$ and $\psi_i^*(\vec{r})$ as independent, we can minimize (8.6) with respect to the ψ_i^* under the constraint of normalization via

$$\frac{\delta}{\delta\psi_i^*}\left[\langle\Psi_{\text{H}}|H|\Psi_{\text{H}}\rangle - \sum_{j=1}^{N}\{\varepsilon_i(1-\langle\psi_j|\psi_j\rangle)\}\right] = 0 \; . \tag{8.7}$$

The ε_i act as the so-called Lagrange multipliers ensuring the normalisation of the eigen functions. The variational problem (8.7) can be reformulated using a function $Q(\{\psi_i^*\},\{\psi_i^*\})$ which is defined as

$$Q(\{\psi_i^*\},\{\psi_i\}) \;=\; \langle\Psi_\mathrm{H}|H|\Psi_\mathrm{H}\rangle \;-\; \sum_{j=1}^{N}\{\varepsilon_j(1-\langle\psi_j|\psi_j\rangle)\}\;. \tag{8.8}$$

Eq. (8.7) thus corresponds to the relation

$$\delta Q(\{\psi_i^*\},\{\psi_i\}) \;=\; Q(\{\ldots,\psi_k^*+\delta\psi_k^*,\ldots\},\{\psi_i\}) \;-\; Q(\{\psi_i^*\},\{\psi_i\}) \;=\; 0 \tag{8.9}$$

for a arbitrary variation $\delta\psi_k^*$, $k=1,\ldots,N$. Using (8.6) and (8.8), we obtain

$$\begin{aligned}&\int d^3r\;\delta\psi_k^*(\vec{r})\left(-\frac{\hbar^2}{2m}\nabla^2+v_{ext}(\vec{r})\right)\psi_k(\vec{r})\\ &+2\cdot\frac{1}{2}\sum_{j=1}^{N}\int d^3rd^3r'\frac{e^2}{|\vec{r}-\vec{r}'|}\delta\psi_k^*(\vec{r})\psi_k(\vec{r})|\psi_j(\vec{r}')|^2\\ &\qquad\qquad-\varepsilon_k\int d^3r\;\delta\psi_k^*(\vec{r})\psi_k(\vec{r})=0\;.\end{aligned} \tag{8.10}$$

Now the variation $\delta\psi_k^*$ is arbitrary, i.e., (8.10) is valid for any variation. The variational scheme thus leads to the so-called *Hartree equations*:

$$\left\{-\frac{\hbar^2}{2m}\nabla^2+v_{ext}(\vec{r})+\sum_{j=1}^{N}\int d^3r'\frac{e^2}{|\vec{r}-\vec{r}'|}|\psi_j(\vec{r}')|^2\right\}\psi_k(\vec{r}) \;=\; \varepsilon_k\psi_k(\vec{r})\;. \tag{8.11}$$

The Hartree equations correspond to a mean-field approximation. Equation (8.11) shows that an effective one-particle Schrödinger equation is solved for an electron embedded in the electrostatic field of *all electrons including the particular electron itself*. This causes the so-called *self interaction* which is erroneously contained in the Hartree equations.

Using the electron density

$$n(\vec{r})=\sum_{i=1}^{N}|\psi_i(\vec{r})|^2\;, \tag{8.12}$$

the *Hartree potential* v_H can be defined:

$$v_\mathrm{H}(\vec{r})=\int d^3r'n(\vec{r}')\frac{e^2}{|\vec{r}-\vec{r}'|}\;. \tag{8.13}$$

It corresponds to the electrostatic potential of all electrons. With this definition the Hartree equations can be written in a more compact form as

$$\left\{-\frac{\hbar^2}{2m}\nabla^2+v_{ext}(\vec{r})+v_\mathrm{H}(\vec{r})\right\}\psi_i(\vec{r})=\varepsilon_i\psi_i(\vec{r})\;. \tag{8.14}$$

The Hartree equations have the form of one-particle Schrödinger equations. However, the solutions $\psi_i(\vec{r})$ of the Hartree equations entering the effective one-particle

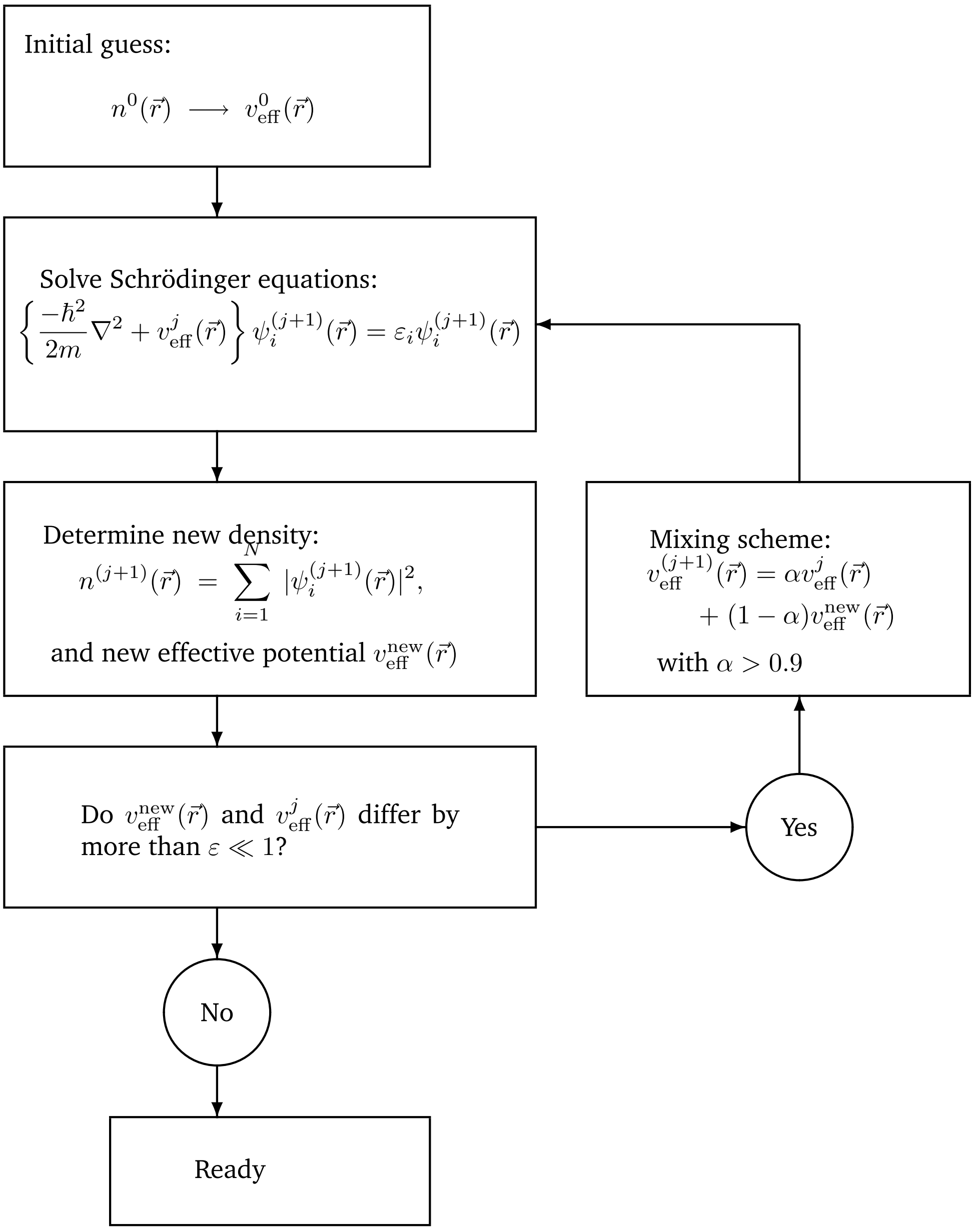


Figure 8.1: Flow-chart diagram of a self-consistent scheme to solve the Hartree equations

Hamiltonan are in fact used to obtain the solutions. Consequently, the Hartree equations can only be solved in an iterative fashion: One starts with some initial guess for the wave functions which enter the effective one-particle Hamiltonian. The Hartree equations are then solved and a new set of solutions is determined. This cycle is repeated so often until the iterations no longer modify the solutions, i.e. until self-consistency is reached. Approximative methods such as the Hartree approximation that include a self-consistency cycle are also known as *self-consistent field* (SCF) approximations.

Such a self-consistent scheme is illustrated in a flow-chart diagram in Fig. 8.1 where we have combined the external and the Hartree potential to an effective potential $v_{\mathrm{eff}}(\vec{r}) = v_{\mathrm{ext}}(\vec{r}) + v_{\mathrm{H}}(\vec{r})$. Furthermore, we have included a mixing scheme between the new effective potential and the one of the previous step for the construction of the effective potential entering the next iteration cycle. Usually a mixing scheme speeds up the convergence of the iteration scheme significantly; sometimes convergence can even not be reached without a mixing scheme. Note that the general self-consistency cycle depicted in Fig. 8.1 is not only appropriate for the solution of the Hartree scheme but for any method that requires a self-consistent solution of one-particle equations.

The expectation value of the total energy in the Hartree approximation E_{H} can be written as

$$\begin{aligned} \langle \Psi_{\mathrm{H}} | H | \Psi_{\mathrm{H}} \rangle &= \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \int d^3r d^3r' \frac{e^2 n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} + V_{\mathrm{nucl-nucl}} \\ &= \sum_{i=1}^{N} \varepsilon_i - V_{\mathrm{H}} + V_{\mathrm{nucl-nucl}} = E_{\mathrm{H}} \end{aligned} \tag{8.15}$$

The integral in (8.15) is the so-called *Hartree energy* V_{H}. It corresponds to the classical (or mean-field) electrostatic energy of the electronic charge distribution. It is contained twice in the Hartree eigenvalue; in order to correct for this double-counting it has to be subtracted in (8.15). In fact, the total energy in (8.15) would only be a sum over single-particle energies if the particles were non-interacting (except for the term $V_{\mathrm{nucl-nucl}}$, which in this context for fixed nuclei just acts as an energy renormalization constant). If we evaluate the total energy for interacting particles by self-consistently solving a set of effective single-particle equations, the total energy is not just a sum over single-particle energies, but there will always be correction terms reflecting the interaction between the particles.

8.2 Hartree–Fock Theory

Since electrons are fermions, they have to obey the Pauli principle: Any state (n, σ) where σ denotes the spin can only be occupied by one electron. In other words, any state characterized by the quantum number n can be occupied by two electrons, a spin up (↑) and a spin down (↓) electron. Furthermore, electrons have to satisfy the Fermi-Dirac statistics which requires that the many-body electron wave function is totally antisymmetrical. This means that the many-body wave function has to be antisymmetrical with respect to the exchange of any two particles.

The Hartree ansatz obeys the Pauli principle only to some extent by populating each electronic state once. However, it does not take into account the anti-symmetry of the wave function. This can be achieved by forming antisymmetrical linear combinations of the products

$$\begin{aligned} &\psi_1(\vec{r}_1\sigma_1) \cdot \psi_2(\vec{r}_2\sigma_2) \cdot \psi_3(\vec{r}_3\sigma_3) \cdot \ldots \cdot \psi_N(\vec{r}_N\sigma_N) \;, \\ &\psi_1(\vec{r}_2\sigma_2) \cdot \psi_2(\vec{r}_1\sigma_1) \cdot \psi_3(\vec{r}_3\sigma_3) \cdot \ldots \cdot \psi_N(\vec{r}_N\sigma_N) \;, \qquad \text{etc.} \;, \end{aligned} \tag{8.16}$$

where we now explicitly include the spin. Note that the spin-up (↑) and a spin down (↓) states are often denoted as α-spin and β-spin, respectively, by quantum chemists.

The easiest way to get a totally antisymmetrical wave function is by using the so-called *Slater determinant* , which is constructed by taken the determinant of the products of the single-particle wave function

$$\Psi_{\rm HF}(\vec{r}_1\sigma_1,\ldots,\vec{r}_{\rm N}\sigma_{\rm N}) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\vec{r}_1\sigma_1) & \psi_1(\vec{r}_2\sigma_2) & \ldots & \psi_1(\vec{r}_{\rm N}\sigma_{\rm N}) \\ \psi_2(\vec{r}_1\sigma_1) & \psi_2(\vec{r}_2\sigma_2) & \ldots & \psi_2(\vec{r}_{\rm N}\sigma_{\rm N}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{\rm N}(\vec{r}_1\sigma_1) & \psi_{\rm N}(\vec{r}_2\sigma_2) & \ldots & \psi_{\rm N}(\vec{r}_{\rm N}\sigma_{\rm N}) \end{vmatrix} . \tag{8.17}$$

Now we follow the same procedure as for the Hartree ansatz: first we write down the expectation value of the total energy:

$$\begin{aligned} \langle\Psi_{\rm HF}|H|\Psi_{\rm HF}\rangle \;&=\; \sum_{i=1}^{N}\int d^3r\psi_i^*(\vec{r})\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{ext}(\vec{r})\right)\psi_i(\vec{r}) \\ &+\frac{1}{2}\sum_{i,j=1}^{N}\int d^3rd^3r'\frac{e^2}{|\vec{r}-\vec{r}'|}|\psi_i(\vec{r})|^2|\psi_j(\vec{r}')|^2 + V_{\rm nucl-nucl} \\ &-\frac{1}{2}\sum_{i,j=1}^{N}\int d^3rd^3r'\frac{e^2}{|\vec{r}-\vec{r}'|}\delta_{\sigma_{\rm i}\sigma_{\rm j}}\psi_i^*(\vec{r})\psi_i(\vec{r}')\psi_j^*(\vec{r}')\psi_j(\vec{r})\;. \end{aligned} \tag{8.18}$$

There is now an additional negative term for electrons with the same spin. This extra term is called the exchange energy $E_{\rm x}$. Note that the total-energy expression (**??**) is *self-interaction free* because the diagonal terms of $E_{\rm x}$ with $i = j$ exactly cancel the corresponding terms in the Hartree energy $E_{\rm H}$.

Again, we minimize the expression (**??**) with respect to the ψ_i^* under the constraint of normalization. This yields the *Hartree–Fock equations*:

$$\begin{aligned} &\left\{-\frac{\hbar^2}{2m}\nabla^2 \;+\; v_{ext}(\vec{r}) \;+\; v_{\rm H}(\vec{r})\right\}\;\psi_i(\vec{r}) \\ &-\sum_{j=1}^{N}\int d^3r'\;\frac{e^2}{|\vec{r}-\vec{r}'|}\;\psi_j^*(\vec{r}')\psi_i(\vec{r}')\psi_j(\vec{r})\delta_{\sigma_{\rm i}\sigma_{\rm j}} \;=\; \varepsilon_i\;\psi_i(\vec{r}). \end{aligned} \tag{8.19}$$

The additional term, called the *exchange term*, introduces quite some complexity to the equations. It is of the form $\int V(\vec{r},\vec{r}')\psi(\vec{r}')d^3r'$, i.e., it is an integral operator. In more compact form, the expectation value of the total energy in the Hartree-Fock approximation is given by

$$\langle\Psi_{\rm HF}|H|\Psi_{\rm HF}\rangle \;=\; E_{\rm HF} \;=\; \sum_{i=1}^{N}\varepsilon_i \;-\; V_{\rm H} \;-\; E_{\rm x} \;+\; V_{\rm nucl-nucl}\;. \tag{8.20}$$

Analogously to (8.15), the Hartree energy $V_{\rm H}$ and the exchange energy $E_{\rm x}$ have to be substracted since they enter the Hartree–Fock eigenvalues twice.

Introducing the notations

$$h_{p,q} \;=\; \int d^3r\;\psi_p^*(\vec{r})\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{ext}(\vec{r})\right)\psi_q(\vec{r}) \tag{8.21}$$

$$V_{lm,pq} \;=\; \int d^3rd^3r'\;\psi_l^*(\vec{r})\psi_m^*(\vec{r}')\;\frac{e^2}{|\vec{r}-\vec{r}'|}\;\psi_p(\vec{r})\psi_q(\vec{r}')\;, \tag{8.22}$$

the total energy in Hartree-Fock theory can be re-expressed as

$$\begin{aligned} E_{\mathrm{HF}} &= \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \sum_{i,j=1}^{N} \left(V_{ij,ij} - V_{ij,ji}\delta_{\sigma_{\mathrm{i}}\sigma_{\mathrm{j}}}\right) + V_{\mathrm{nucl-nucl}} \\ &= \sum_{i=1}^{N} h_{i,i} + \frac{1}{2} \sum_{i,j=1}^{N} \left(V_{ij,ij} - V_{ij,ji}\delta_{\sigma_{\mathrm{i}}\sigma_{\mathrm{j}}}\right) + V_{\mathrm{nucl-nucl}} \,. \end{aligned} \tag{8.23}$$

Note that sometimes also the notations

$$(p,q) = h_{p,q}, \quad (pq,rs) = V_{pq,rs} \tag{8.24}$$

are used.

The equations (**??**), (8.19) and (8.20) give a general expression of the Hartree-Fock energy. Solving them is usually called the unrestricted Hartree-Fock (UHF) method. However, in cases where there is an even number of states and all spatial states are occupied twice by a spin-up and a spin-down electron, it is sufficient just to focus on the doubly-occupied spatial orbitals. Before doing so, we first introduce a shorter way to write down a Slater determinant:

$$\Psi_{\mathrm{HF}} = |\psi_1\psi_2 \ldots \psi_M\bar{\psi}_{M+1}\bar{\psi}_{M+2} \ldots \bar{\psi}_N| \,, \tag{8.25}$$

where the index i labels the spatial part of the wave function, ψ_i is a α or spin-up state, whereas the bar ($\bar{\psi}_j$) denotes the β or spin-down state, M is the number of occupied spin-up states, and $(N-M)$ the number of ocuupied spin-down states.

If we now have an even number N of electrons and the lowest $n/2$ electron levels are occupied twice with a pair of electrons of opposite spin, the corresponding Slater determinant is given by

$$\Psi_{\mathrm{RHF}} = |\psi_1\psi_2 \ldots \psi_{N/2}\bar{\psi}_1\bar{\psi}_2 \ldots \bar{\psi}_{N/2}| \,, \tag{8.26}$$

This means that in the expression (8.23) for the total energy every one-particle energy $h_{i,i}$ appears twice, while in the double sum $V_{ij,ij}$ appears four times. However, $V_{ij,ij}$ has to be counted only twice since the spin quantum numbers of the states have to be the same. Hence one obtains in the so-called restricted Hartree-Fock (RHF) method

$$E_{\mathrm{RHF}} = 2\sum_{i=1}^{N/2} h_{i,i} + \sum_{i,j=1}^{N/2} \left(2V_{ij,ij} - V_{ij,ji}\right) + V_{\mathrm{nucl-nucl}} \,. \tag{8.27}$$

The advantage of the restricted Hartree-Fock method is that in principle only half as many electronic states have to be explicitly considered compared to the unrestricted Hartree-Fock method which leads to a significant reduction of the computational effort.

Note that even if there is an even number of electrons in an atom or molecule, not all states are necessarily occupied twice. Prominent examples are the oxygen atom and the oxygen molecule O_2 which are both in a spin-triplet state with the two uppermost degenerate orbitals each only occupied once.

In the most general case of atoms or molecules with open shells, i.e., if the Hartree-Fock wave function is given by the Slater determinant (8.25), we have to take into account that the exchange term only couples electrons with the same spin. Hence the total energy in the unrestricted Hartree-Fock method is given by

$$E_{\mathrm{UHF}} = \sum_{i=1}^{N} h_{i,i} + \frac{1}{2}\sum_{i,j=1}^{N} V_{ij,ij} - \frac{1}{2}\left(\underbrace{\sum_{i,j=1}^{M} V_{ij,ji}}_{\text{spin up}} + \underbrace{\sum_{i,j=M+1}^{N} V_{ij,ji}}_{\text{spin down}} \right) + V_{\mathrm{nucl-nucl}} \,. \tag{8.28}$$

The problem with the general Slater determinant (8.25) is that it is not neccessarily an eigenstate of the total spin operator $\mathbf{S}^2$. This can be most easily seen regarding the generalisation of eq. (**??**) for the total spin of N spins:

$$\begin{aligned} \mathbf{S}^2 &= \left(\sum_{i=1}^{N} \mathbf{S}_i\right)^2 \\ &= \sum_{i=1}^{N} \mathbf{S}_i^2 + \sum_{i<j}^{N} \left(2S_{iz}S_{jz} + S_{i+}S_{j-} + S_{i-}S_{j+}\right) \end{aligned} \tag{8.29}$$

The ladder operators in the second sum switch the spin of the i-th and j-th electrons, if they are different. If, for example, in Ψ only the one-electron state ψ_k is contained, but not $\bar{\psi}_k$, then $\mathbf{S}^2\Psi$ is not proportional to Ψ, i.e., Ψ is no eigenstate of $\mathbf{S}^2$. On the other hand, since the total electronic Hamiltonian (7.9) does not explicitly depend on the spin coordinates, it commutes with $\mathbf{S}^2$ and $\mathbf{S}_z$; thus the eigenfunctions of the electronic Hamiltonian are also eigenfunctions of both spin operators. This means that in general open-shell systems cannot be represented by one single Slater determinant since it is contaminated by different spin states. This *spin contamination* can be measured by determining the mean expectation value per electron of the spin $\mathbf{S}^2$ of an UHF determinant. If the mean value is exactly or close to the one expected for an eigenstate, there is no or very little spin contamination. If, one the other hand, the UHF wave function is contaminated by another spin state by more than 3%, it should be rejected since it does not give a proper description of the open-shell system.

In order to treat open-shell systems within the Hartree-Fock method, one can only consider wave functions that are eigenfunctions of $\mathbf{S}^2$. This method is then called restricted open-shell Hartree-Fock (ROHF). If we consider for example a triplet state where all states are doubly occupied except for the two highest m and n, the triplet wave function with $S_z = -1$ is given by

$$ {}^{3}_{-1}\Psi^{n}_{m} = |\psi_1\bar{\psi}_1 \ldots \psi_M\bar{\psi}_M\bar{\psi}_m\bar{\psi}_n| \,, \tag{8.30}$$

where we have used the notation ${}^{(2S+1)}_{S_z}\Psi$. By applying the ladder operator S_+, we obtain, disregarding any normalization, the ${}^{3}_{0}\Psi$ state:

$$ {}^{3}_{0}\Psi^{n}_{m} = \left(|\psi_1\bar{\psi}_1 \ldots \psi_M\bar{\psi}_M\psi_m\bar{\psi}_n| - |\psi_1\bar{\psi}_1 \ldots \psi_M\bar{\psi}_M\bar{\psi}_m\psi_n|\right) \;.\,, \tag{8.31}$$

This shows that a $S_z = 0$ state of a triplet can not be represented by a single Slater determinant. For the sake of completeness, we also give the ${}^{3}_{1}\Psi$ state:

$$ {}^{3}_{1}\Psi^{n}_{m} = |\psi_1\bar{\psi}_1 \ldots \psi_M\bar{\psi}_M\psi_m\psi_n| \,, \tag{8.32}$$

The expectation value of the total energy of this triplett state which is independent of the value of S_z is then given by

$$\begin{aligned} E_{\mathrm{ROHF}} &= \underbrace{2\sum_{i=1}^{M} h_{i,i} + \sum_{i,j=1}^{M} (2V_{ij,ij} - V_{ij,ji})}_{\text{closed shell}} + \underbrace{h_{m,m} + h_{n,n} + V_{mn,mn} - V_{mn,nm}}_{\text{open shell}} \\ &\quad \underbrace{+ \sum_{i=1}^{M} (2V_{im,im} - V_{im,mi}) + \sum_{i=1}^{M} (2V_{in,in} - V_{in,ni})}_{\text{interaction open } - \text{ closed shell}} \end{aligned} \qquad (8.33)$$

In general, Hartree-Fock theory is a good starting point for the theoretical description of molecules and clusters. The exact total energy of a molecule is often reproduced to up to 99%. The Hartree ansatz does not take into account the antisymmetry of the electronic many-body wave function; consequently it does not yield a proper solution of the many-body Schrödinger equation. The Hartree–Fock method, on the other hand, incorporates the antisymmetry requirement and leads to a reduction of the total energy. Hence it should be a more appropriate solution for the true ground-state already on the basis of the Rayleigh-Ritz variational principle. However, in the Hartree–Fock ansatz, electrons of opposite spin are still not correlated. If these electrons are also avoiding each other, the energy can be further reduced. This additional effect is called *electron correlation*. The *electron correlation energy* is defined as the difference between the exact energy of the system and the Hartree–Fock energy. The distinction between electron correlation and exchange is somehow artificial because the Hartree–Fock exchange is in principle also an electron correlation effect.

As mentioned above, Hartree-Fock theory gives usually a good estimate for the total energy of a molecule. Unfortunately, the missing part, namely the electron correlation energy, is rather important for a reliable description of chemical bond formation. The correlation energy per electron can easily be more than 1 eV. Chemists often demand an accuracy of 1 kcal/mol $\approx$ 0.04 eV (“chemical accuracy”) for energy differences in order to consider the calculations useful. If the uncertainty of the total energies is already above several eV, then only by a fortuitous cancellation of errors chemical accuracy can be achieved. Quantum chemists have therefore developed a whole machinery of so-called post-Hartree-Fock methods that treat the electron correlation at various levels of sophistication. Some of the most important concepts will be introduced in the next section.

8.3 Post–Hartree–Fock Methods

The post-Hartree-Fock methods methods can be devided into to two categories, the so-called *single-reference* and the *multiple-reference* methods. Both methods take into acount the fact that the electrons do not only experience the *mean field* of all other electrons, but that they are constantly interacting which each other. In real space, this corresponds to the consideration of collisions of the electrons which can be described as the excitation of virtual excited orbitals. The most probable and therefore most important processes are the collision of two electrons which is represented by double or pair excitations.

In the single-reference methods, one starts with the Slater determinant that is the solution of the Hartree–Fock equations. This means that the electron correlation between electrons with opposite spin is neglected. One way to introduce correlation effects is by considering virtually excited states. These can be generated by replacing occupied orbitals in the Slater determinant by unoccupied ones, i.e., by states that do not correspond to the N lowest Hartree–Fock eigenvalues. By replacing one, two, three, four or more states, single, double, triple, quadruple or higher excitations can be created. In the Møller–Plesset theory, these excitations are treated perturbatively by regarding the Hartree–Fock Hamiltonian as the unperturbed Hamiltonian and the difference to the true many-body Hamiltonian as the perturbation.

To derive the Møller-Plesset theory, we first rewrite the Hartree–Fock equations (8.19) as

$$h_i^{\mathrm{HF}} \; \psi_i(\vec{r}) = \varepsilon_i \; \psi_i(\vec{r}) \; , \tag{8.34}$$

where the effective one-electron operator h_i^{HF} acts on the i-th electron. If we neglect the nuclear-nuclear interaction $V_{\mathrm{nucl-nucl}}$ for a moment, the Hartree–Fock Hamiltonian H_{HF} can be defined as

$$H_{\mathrm{HF}} = \sum_i h_i^{\mathrm{HF}} - V_{\mathrm{H}} - E_{\mathrm{x}} \; . \tag{8.35}$$

Here V_{H} and E_{x} are just treated as constants determined with the true Hartree–Fock ground state wave-function. Note that H_{HF} does not correspond to the correct many-body electronic Hamiltonian H_{el} (7.9) since the correlation effects between electron with opposite spins are not included. The difference between H_{el} and H_{HF} is now treated as the perturbation H':

$$H' = H_{\mathrm{el}} - H_{\mathrm{HF}} \; . \tag{8.36}$$

The expression for the ground-state energy in second-order perturbation theory becomes

$$E^{(2)} = E_{\mathrm{HF}} + \langle \Psi_0 | H' | \Psi_0 \rangle + \sum_{l \neq 0} \frac{|\langle \Psi_l | H' | \Psi_0 \rangle|^2}{E_0 - E_l} \; . \tag{8.37}$$

The sum over states l other than the ground state corresponds to Slater determinants with single, double, triple, etc. excitations. In fact it can be easily shown that the Hartree–Fock theory is correct to first order, i.e., the first-order correction $\langle \Psi_0 | H' | \Psi_0 \rangle$ vanishes. The second-order expression is given by

$$E^{(2)} = E_{\mathrm{HF}} - \sum_{l<m}^{occ} \sum_{p<q}^{unocc} \frac{|V_{lm,qp} - V_{lm,pq}|^2}{\varepsilon_l + \varepsilon_m - \varepsilon_p - \varepsilon_q} \; , \tag{8.38}$$

where the sum is performed over all occupied and unoccupied (virtual) orbitals. Second-order Møller–Plesset theory is usually denoted by MP2. If higher-order corrections are included, the methods are named MP3, MP4 and so forth.

MP2 is a very popular method due to its conceptual simplicity. However, due to the perturbative treatment of electron correlation its applicability is still limited.

In order to go beyond pertubation theory, we note that a general Slater determinant for any given set of wave functions $\chi_{K_1}, \chi_{k_2}, \ldots \chi_{k_N}$ can be written as

$$\Psi_{k_1, k_2 \ldots, k_N} = |\chi_{k_1} \chi_{k_2} \cdots \chi_{k_N}| \; , \tag{8.39}$$

Here we assume that the index k_i also includes the spin quantum number and that the quantum numbers k_i are ordered in some suitable form

$$k_1 < k_2 < \ldots < k_N \tag{8.40}$$

Now, if the χ_{k_i} form a complete set of one-particle functions, then it can be shown that the Slater determinants (8.39) form a complete set for any antisymmetric wave function Ψ of N electrons. This means that any N-electron wave function can be expressed as

$$\Psi \;=\; \sum_{k_1<k_2<\ldots<k_N} C_{k_1 k_2 \ldots k_N} \Psi_{k_1,k_2,\ldots,k_N} \tag{8.41}$$

Using this expansion in the Schrödinger equation $H\Psi = E\Psi$, one obtains after some calculations a set of linear equations for the coefficients $C_{k_1 k_2 \ldots k_N}$:

$$\begin{aligned} &\sum_{i=1}^{N} \sum_{k'_i} h_{k_i,k'_i} \delta_{\sigma_{\mathrm{k_i}} \sigma_{\mathrm{k'_i}}} C_{k_1 \ldots k'_i \ldots k_N} \\ &\quad + \sum_{i,j=1}^{N} \sum_{k'_i k''_j} V_{k_i k_j, k'_i k''_j} \delta_{\sigma_{\mathrm{k_i}} \sigma_{\mathrm{k'_i}}} \delta_{\sigma_{\mathrm{k_j}} \sigma_{\mathrm{k''_j}}} C_{k_1 \ldots k'_i \ldots k''_j \ldots k_N} \;=\; E C_{k_1 \ldots k_N} \;. \end{aligned} \tag{8.42}$$

Every wave function $\Psi_{k_1,\ldots,k_2}$ is regarded as a configuration. Therefore this method is called *configuration interaction* (CI). Eq. (8.42) corresponds to a set of linearly coupled equations which can be solved numerically using standard routines. Usually this method is coupled with the LCAO approach, i.e., every orbital is expanded in a set of atomic wave functions

$$\chi_{k_i} \;=\; \sum_l c_l^{(k_i)} \phi_l \tag{8.43}$$

Hence both the coefficients $C_{k_1 k_2 \ldots k_N}$ and $c_l^{(k_i)}$ have to be determined. The matrix elements h_{k_i,k'_i} and $V_{k_i k_j, k'_i k''_j}$ are then expressed as a sum over the corresponding integrals of the atomic orbitals which are then stored and which enter the linear equations (8.42). The numerical effort does of course sensitively depend on the number of atomic orbitals and the number of configurations $\Psi_{k_1,\ldots,k_2}$ included in the expansion.

If only single (S) and double (D) excitations are included in the sum (8.41), one refers to the method as CISD:

$$\begin{aligned} \Psi_{\mathrm{CISD}} = \Psi_{\mathrm{HF}} &+ \sum_{i,k'_i} C^{\mathrm{S}}_{k_1 \ldots k'_i \ldots k_N} \Psi^{\mathrm{S}}_{k_1 \ldots k'_i \ldots k_N} \\ &+ \sum_{ij,k'_i k''_j} C^{\mathrm{D}}_{k_1 \ldots k'_i \ldots k''_j \ldots k_N} \Psi^{\mathrm{D}}_{k_1 \ldots k'_i \ldots k''_j \ldots k_N} \;, \end{aligned} \tag{8.44}$$

where Ψ^{S} and Ψ^{D} correspond to Slater determinants with one or two occupied orbitals replaced by unoccupied ones, respectively. The CISD method represents an improvement with respect to the pure Hartree-Fock method, however, this approach does not obey one important desirable property of electronic structure methods, namely the so-called *size consistency* or *size extensivity*. In principle, with size extensitivity the linear scaling of the energy with the number of electrons is meant. For infinitely separated systems, this comes down to additive energies of the separated components.

This property is not only important for large systems, but even for small molecules. The CISD method does not fulfil size extensitivity because the product of two fragment CISD wave functions contains triple and quadruple excitations and is therefore no CISD function. One elegant way to recover size extensitivity is to exponentiate the single and double excitations operator:

$$\Psi_{\mathrm{CCSD}} = \exp(T_1 + T_2)\ \Psi_{\mathrm{HF}}\ . \tag{8.45}$$

This approach is called coupled cluster (CC) theory, the limitation to single and double excitations is denoted by CCSD. If also triple excitations are included, the method is called CCSDT. However, the computational effort of this method has an eighth-power dependence on the size of the system and is therefore rather impractical. The scaling is more favorable if the triple excitations are incorporated perturbatively still, this CCSD(T) method is very accurate.

The single-reference methods can be very reliable in the vicinity of equilibrium configurations, but they are often no longer adequate to describe a bond-breaking process. One Slater determinant plus excited states derived from this determinant are not sufficient because the dissociation products should be described by a linear combination of two Slater determinants taking into account the proper spin state. Any many-particle electronic wave function can in principle be represented by a sum over Slater determinants. Hence by considering more and more configurations in the calculations, the accuracy can be systematically improved. If all possible determinants are included in an electronic structure calculation, the method is called full configuration interaction (FCI). Because of the large computational effort required, FCI calculations are limited to rather small systems. In particular, the treatment of larger clusters necessary to model surface science problems is not possible with FCI. Hence approximate multi-reference methods are needed.

In the multiconfigurational self-consistent field (MCSCF) approach, a relatively small number of configurations is selected and both the orbitals and the configuration interactions coefficients are determined variationally. The selection of the configurations included in a MCSCF calculation cannot be done automatically; usually chemical insight is needed which, however, might introduce a certain bias in the calculations. This can be avoided by the complete active space (CAS) approach. In a CASSCF calculations a set of active orbitals is identified and all excitations within this active space are included. This method is again computationally very costly. After all, if the active space is increased to include all electrons and orbitals, we end up with a FCI calculation.

For a proper treatment of the electronic correlation, not only the appropriate method has to been chosen, but also the basis set has to be sufficiently large enough. Quantum chemical methods usually describe the electrons by a localized basis set derived from atomic orbitals. The preferred basis functions are Gaussian functions because they allow the analytical evaluation of the matrix elements necessary to perform an electronic structure calculations. In fact, the most popular electronic structure code used in quantum chemistry, the GAUSSIAN program, is named after the type of basis functions employed in the program.

Quantum chemists have developed a particular nomenclature to describe the quality of a basis set. It is beyond the scope of this book to give a full introduction into the terminology. I will only give a short overview. The simplest choice of just one atomic orbital per valence state is called “minimal basis set” or “single zeta”. If two or more orbitals are included, the basis set is called “double zeta”(DZ), “triple zeta” (TZ), “quadruple zeta” (QZ) and so on. Often polarization functions are added which correspond to one or more sets of d functions on first row atoms. Then a “P” is added to the acronym of the basis set resulting in, e.g., DZ2P. These polarization functions describe small displacements of the atomic orbitals from the nuclear centers. For rather delocalized states such as anionic or Rydberg excited states, further diffuse functions are added.

In quantum chemical methods the accuracy of the treatment of electron correlation can be improved in a systematic way by either choosing a more refined method or by increasing the basis set. This is a very attractive property of these wave function based methods. Unfortunately, the accuracy is paid by an immense computational effort. For example, the CCSD(T) method exhibits an N^7 scaling in the computational effort where N is the number inlcuded elecrons. Also the convergence with the size of the basis set is rather slow. If one gives cardinal numbers X for the basis sets according to DZ:$X = 2$, TZ:$X = 3$, QZ:$X = 4, \ldots$, then the error of the calculations ΔE_{err} shows a scaling with the basis set size that goes like

$$\Delta E_{\mathrm{err}} \propto X^{-4} , \tag{8.46}$$

which means that any new reliable digit in the calculations requires 10 000 times more CPU time. Hence accurate calculations in quantum chemistry are usually limited to a rather small number of atoms, typically about 10–20.

In recent years one method has become more and more popular that is not based on a representation of the many-body wave function, but on the electron density which is therefore computationally much more efficient: density functional theory.

8.4 Density Functional Theory

The huge computational effort of quantum chemistry methods is related to the fact that they are based on the many-body wave function which depends on $3N$ coordinates, where N is the number of electrons. Let us consider, e.g., the oxygen atom with 8 electrons. If we use 10 entries for each coordinate, then we need 10^{24} entries for the whole wave function. Such an amount of data is almost impossible to store. It would be much more convenient if not the many-body wave function, but the electron density could be used as the basic quantity in electronic structure calculations.

Such a connection between the electronic density and the many-body Hamiltonian is in fact provided by the Hohenberg–Kohn theorem that density functional theory (DFT) is based upon. This theorem states that the ground-state density $n(\vec{r})$ of a system of interacting electrons in an external potential uniquely determines this potential. The proof for this theorem which is rather simple will be presented here in order to demonstrate that also theories that are based on simple ideas can lead to a Nobel prize (Walter Kohn, Nobel prize for chemistry 1998). However, the formulation of its rigorous mathematical foundations was only completed several years after the first presentation of the Hohenberg–Kohn theorem.

Let us assume for the sake of simplicity that the system of interacting electrons has a nondegenerate ground state (the extension to degenerate cases is straightforward). Let the wave function Ψ_1 be the nondegenerate ground state of the Hamiltonian H_1 with external potential $v_1(\vec{r})$ and the corresponding ground-state density $n(\vec{r})$. The ground-state energy E_1 is then given by

$$\begin{aligned} E_1 &= \langle \Psi_1 \mid H_1 \mid \Psi_1 \rangle \\ &= \int v_1(\vec{r})n(\vec{r})\, d^3\vec{r} \;+\; \langle \Psi_1 \mid (T+U) \mid \Psi_1 \rangle \,. \end{aligned} \tag{8.47}$$

Here T and U are the operators of the kinetic and the interaction energy. Now let us assume that there is a second potential $v_2(\vec{r})$ which differs from $v_1(\vec{r})$ by not just a constant, i.e. $v_2(\vec{r}) \neq v_1(\vec{r}) + \text{const.}$, but leads to the same electron density $n(\vec{r})$. The corresponding ground state energy is

$$\begin{aligned} E_2 &= \langle \Psi_2 \mid H_2 \mid \Psi_2 \rangle \\ &= \int v_2(\vec{r})n(\vec{r})\, d^3\vec{r} \;+\; \langle \Psi_2 \mid (T+U) \mid \Psi_2 \rangle \,. \end{aligned} \tag{8.48}$$

Now we can apply the Rayleigh-Ritz variational principle. Since the ground state Ψ_1 is assumed to be nondegenerate, we obtain the true inequality

$$\begin{aligned} E_1 &< \langle \Psi_2 \mid H_1 \mid \Psi_2 \rangle \\ &= \int v_1(\vec{r})n(\vec{r})\, d^3\vec{r} \;+\; \langle \Psi_2 \mid (T+U) \mid \Psi_2 \rangle \\ &= E_2 \;+\; \int (v_1(\vec{r}) \;-\; v_2(\vec{r}))\, n(\vec{r})\, d^3\vec{r} \,. \end{aligned} \tag{8.49}$$

Equivalently, we can use the Rayleigh–Ritz variational principle for H_2. We have not explicitly assumed that Ψ_2 is nondegenerate, hence we obtain

$$\begin{aligned} E_2 &\leq \langle \Psi_1 \mid H_2 \mid \Psi_1 \rangle \\ &= E_1 \;+\; \int (v_2(\vec{r}) \;-\; v_1(\vec{r}))\, n(\vec{r})\, d^3\vec{r} \,. \end{aligned} \tag{8.50}$$

If we add (8.49) and (8.50), we end up with the contradiction

$$E_1 + E_2 < E_1 + E_2 \,. \tag{8.51}$$

Hence the initial assumption that two different external potential can lead to the same electron density is wrong. This concludes the proof of the Hohenberg–Kohn theorem.

Since the density $n(\vec{r})$ is uniquely related to the external potential and the number N of electrons via $N = \int n(\vec{r}) d^3\vec{r}$, it determines the full Hamiltonian. Thus in principle it determines all quantities that can be derived from the Hamiltonian such as, e.g., the electronic excitation spectrum. However, unfortunately this has no practical consequences since the dependence is only implicit.

In the derivation of the Hartree and the Hartree–Fock methods we have used the Rayleigh-Ritz variational principle. This demonstrated the importance of variational principles. In fact, there is also a variational principle for the energy functional, namely that the exact ground state density and energy can be determined by the minimisation of the energy functional $E[n]$:

$$E_{\text{tot}} = \min_{n(\vec{r})} E[n] = \min_{n(\vec{r})} (T[n] + V_{\text{ext}}[n] + V_{\text{H}}[n] + E_{\text{xc}}[n]) \,. \tag{8.52}$$

$V_{\text{ext}}[n]$ and $V_H[n]$ are the functionals of the external potential and of the classical electrostatic interaction energy that corresponds to the Hartree energy, while $T[n]$

is the kinetic energy functional for non-interacting electrons, i.e. the kinetic energy functional of a non-interacting reference system that is exposed to the same external potential as the true interacting system. All quantum mechanical many-body effects are contained in the so-called exchange-correlation functional $E^{\mathrm{xc}}[n]$. Yet, this non-local functional is not known; probably it is even impossible to determine it exactly in a closed form. However, it has the important property that it is a well-defined universal functional of the electron density, i.e., it does not depend on any specific system or element. Instead of using the many-body quantum wave function which depends on $3N$ coordinates now only a function of three coordinates has to be varied. In practice, however, no direct variation of the density is performed. One of the reasons is that the kinetic energy functional $T[n]$ is not well-known either.

The density is rather expressed as a sum over single-particle states

$$n(\vec{r}) = \sum_{i=1}^{N} |\psi_i(\vec{r})|^2 \, . \tag{8.53}$$

Now we make use of the variational principle for the energy functional and minimize $E[n]$ with respect to the single particle states under the constraint of normalization. This procedure is entirely equivalent to the derivation of the Hartree and the Hartree–Fock equations (8.11) and (8.19), respectively. Thus we obtain the so-called Kohn–Sham equations

$$\left\{ -\frac{\hbar^2}{2m}\nabla^2 + v_{ext}(\vec{r}) + v_{\mathrm{H}}(\vec{r}) \; + \; v_{\mathrm{xc}}(\vec{r}) \right\} \psi_i(\vec{r}) = \varepsilon_i \; \psi_i(\vec{r}) \, . \tag{8.54}$$

Thus the effective one-electron potential acting on the electrons is given in the Kohn–Sham formalism by

$$v_{\mathrm{eff}}(\vec{r}) = v_{\mathrm{ext}}(\vec{r}) + v_{\mathrm{H}}(\vec{r}) + v_{\mathrm{xc}}(\vec{r}) \, . \tag{8.55}$$

The exchange-correlation potential $v_{\mathrm{xc}}(\vec{r})$ is the functional derivative of the exchange-correlation functional $E_{\mathrm{xc}}[n]$

$$v_{\mathrm{xc}}(\vec{r}) = \frac{\delta E_{\mathrm{xc}}[n]}{\delta n} \, . \tag{8.56}$$

The ground state energy can now be expressed as

$$E = \sum_{i=1}^{N} \varepsilon_i + E_{\mathrm{xc}}[n] - \int v_{\mathrm{xc}}(\vec{r}) n(\vec{r}) \; d^3\vec{r} - V_{\mathrm{H}} + V_{\mathrm{nucl-nucl}} \, . \tag{8.57}$$

Here we have added the term $V_{\mathrm{nucl-nucl}}$ in order to have the correct total energy of the electronic Hamiltonian (7.9). In solid-state applications, the sum over the single-particle energies in (8.57) is often called the band-structure energy. However, it is important to keep in mind that the "single-particle energies" ε_i enter the formalism just as Lagrange multipliers ensuring the normalisation of the wave functions. The Kohn–Sham states correspond to quasiparticles with no specific physical meaning except for the highest occupied state. Still it is almost always taken for granted that the Kohn–Sham eigenenergies can be interpreted, apart from a rigid shift, as the correct electronic one-particle energies. This is justified by the success since the Kohn–Sham eigenenergy spectrum indeed very often gives meaningful physical results, as will be shown in the next chapters.

Note that if the exchange-correlation terms E_{xc} and v_{xc} are neglected in (8.54)–(8.57), we recover the Hartree formulation of the electronic many-body problem.

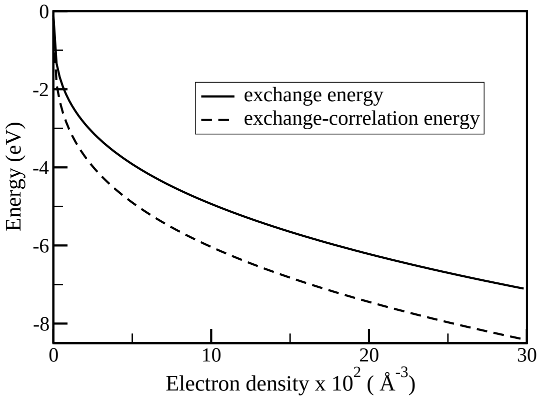


Figure 8.2: Exchange and exchange-correlation energy per particle in the homogeneous electron gas. For example, typical valence electron densities of metals are in the range of 5–10×10^2 Å^{-3}.

Hence the Kohn–Sham theory may be regarded as a formal extention of the Hartree theory. In contrast to the total energy expression in the Hartree and the Hartree–Fock approximation, the ground-state energy (8.57) is in principle exact. The reliability of any practical implementation of density functional theory depends crucially on the accuracy of the expression for the exchange-correlation functional.

The exchange-correlation functional $E_{\mathrm{xc}}[n]$ can be written as

$$E_{\mathrm{xc}}[n] = \int d^3\vec{r}\, n(\vec{r})\, \varepsilon_{\mathrm{xc}}[n](\vec{r}) \,, \tag{8.58}$$

where $\varepsilon_{\mathrm{xc}}[n](\vec{r})$ is the exchange-correlation energy per particle at the point $\vec{r}$, but depends on the whole electron density distribution $n(\vec{r})$. In order to discuss the properties of $E_{\mathrm{xc}}[n]$, it is helpful to introduce the exchange-correlation hole distribution

$$n_{\mathrm{xc}}(\vec{r}, \vec{r}') = g(\vec{r}, \vec{r}') - n(\vec{r}') \,, \tag{8.59}$$

where $g(\vec{r}, \vec{r}')$ is the conditional density to find an electron at $\vec{r}'$ if there is already an electron at $\vec{r}$. Every electron creates a hole corresponding to exactly one electron out of the average density $n(\vec{r})$. This is expressed through the sum rule

$$\int d^3\vec{r}'\; n_{\mathrm{xc}}(\vec{r}, \vec{r}') = -1 \,. \tag{8.60}$$

Furthermore, the exchange-correlation hole vanishes for large distances:

$$n_{\mathrm{xc}}(\vec{r}, \vec{r}') \underset{|\vec{r}-\vec{r}''| \to \infty}{\longrightarrow} 0 \,, \tag{8.61}$$

and there is an asymptotical result for the integral

$$\int d^3\vec{r}'\; \frac{n_{\mathrm{xc}}(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} \underset{|\vec{r}| \to \infty}{\longrightarrow} -\frac{1}{|\vec{r}|} \,. \tag{8.62}$$

Since the exchange-correlation functional $E_{\mathrm{xc}}[n]$ is not known in general, the exchange-correlation energy $\varepsilon_{\mathrm{xc}}[n](\vec{r})$ cannot be exactly derived either. What is known is the

Table 8.1: O_2 binding energy obtained by DFT calculations using LDA and different GGA exchange-correlation functionals

functional	LDA	PW91	PBE	RPBE	Exp.
O_2 binding energy (eV)	7.30	6.06	5.99	5.59	5.23

exchange-correlation energy for the homogeneous electron gas, i.e. for a system with a constant electron density. This energy is plotted in Fig. 8.2. In the so-called Local Density Approximation (LDA), the exchange-correlation energy for the homogeneous electron gas is also used for non-homogeneous situations,

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d^3\vec{r}\; n(\vec{r})\; \varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(n(\vec{r}))\;, \tag{8.63}$$

As (8.63) shows, at any point in space the local exchange-correlation energy $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}(n(\vec{r}))$ of the homogeneous electron gas is used for the corresponding density, ignoring the non-locality of the true exchange-correlation energy $\varepsilon_{\mathrm{xc}}[n]$.

In a wide range of bulk and surface problems the LDA has been surprisingly successful. This is still not fully understood but probably due to a cancellation of opposing errors in the exchange and the correlation expression in the LDA. Furthermore, the LDA satisfies the sum rule (8.60) which is apparingly also very important. For chemical reactions in the gas phase and at surfaces, however, the LDA results are not sufficiently accurate. Usually LDA shows *over-binding*, i.e. binding and cohesive energies turn out to be too large compared to experiment. This overbinding also leads to lattice constants and bond lengths that are smaller than the experimental values. These shortcominigs of LDA were the reason why many theoretical chemists were rather reluctant to use DFT for a long time. There had been attempts to formulate a Taylor expansion of the exchange-correlation energy $\varepsilon_{\mathrm{xc}}[n]$, but these first attempts had not been successful because by a straightforward gradient expansion (8.60) is violated. Only with the advent of exchange-correlation functionals in the Generalized Gradient Approximation (GGA) this situation has changed. In the GGA the gradient of the density is also included in the exchange-correlation energy,

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n] = \int d^3\vec{r}\; n(\vec{r})\; \varepsilon_{\mathrm{xc}}^{\mathrm{GGA}}(n(\vec{r}), |\nabla n(\vec{r})|)\;, \tag{8.64}$$

but the dependence on the gradient is modified in such a way as to satisfy the sum rule (8.60). In addition, general scaling properties and the asymptotic behavior of effective potentials are taken into account in the construction of the GGA. DFT calculations in the GGA achieve chemical accuracy (error ≤ 0.1 eV) for many chemical reactions. This improvement in the accuracy of DFT calculations finally opened the way for Walter Kohn to be honored with the Nobel prize in chemistry in 1998 for the development of DFT which is somewhat paradox because DFT was accepted in the physics community much earlier than in the chemistry community.

Still there are important exceptions where the GGA also does not yield sufficient accuracy. In Table 8.1 DFT results for the O_2 binding energy obtained using LDA and different GGA exchange-correlation functionals are compared to the experimental value. The LDA result shows the typical large overbinding. The GGA functional by Perdew and Wang (PW91) and by Perdew, Burke and Ernzerhof (PBE) have been constructed to give similar results. The revised PBE functional (RPBE) follows the same construction scheme as the PBE functional, just a different interpolation that is not specified by the construction scheme is used. This leads to a difference of almost half an eV for the O_2 binding energy. This is a rather unsatisfactorily result because this means that there is an intrinsic uncertainty of up to half an eV for energy differences obtained within the generalized gradient approximation. And still the theoretical O_2 binding energies are much larger than measured in experiment.

The binding energy of O_2 is not the only case where DFT calculations are rather inaccurate. A list of the failures of DFT with present-day functionals includes: (i) van der Waals forces are not properly described, (ii) negative ions are usually not bound, i.e. electron affinities are too small, (iii) the Kohn–Sham potential falls off exponentially for large distances instead of $\propto 1/r$, (iv) band gaps are underestimated in both LDA and GGA by approximately 50%, (v) cohesive energies are overestimated in LDA and underestimated in GGA, (vi) strongly correlated solids such as NiO and FeO are predicted as metals and not as antiferromagnetic insulators.

The problem in the development of a more accurate exchange-correlation functional is the reliable representation of the non-locality of this functional. One could say that all present formulations of the exchange-correlation functional in principle still represent an uncontrolled approximation. There is no systematic way of improving the functionals since there is no expansion in some controllable parameter.

Still the development of more accurate exchange-correlation function is a very active research field. One route is to include to some extent "exact exchange" in addition to a standard functional. Another *ansatz* is the development of so-called *meta-GGA*'s that include higher-order powers of the gradient or the local kinetic energy density. Very accurate results for small molecules can be obtained by methods based on *orbital functionals* such as the optimized potential method (OPM) or the optimized effective potential (OEP) method. In this approach, the exchange-correlational functional does not explicitly depend on the density but on the individual orbitals. Thus the self-interaction can be avoided. It is still true that all improved functionals mentioned above require a significant increase in the computational effort. Therefore they have not been used yet in standard applications of DFT calculations for surface science problems.

In any practical implementation of DFT the computational effort of course increases significantly with the number of electrons that have to be taken into account. However, most chemical and solid-state properties are determined almost entirely by the valence electrons while the influence of the core electrons on these properties is negligible. Indeed there is a way to replace the effect of the core electrons by an effective potential so that they do not have to be taken into account explicitly, namely by constructing so called *pseudopotentials*. Since this significantly reduces the number of electrons that have to be taken into account, the use of pseudopotentials leads to an enormous saving of computer time.

9 Scattering Theory

The interaction potential between two particles cannot be measured directly, only its consequences in scattering. Therefore it is very important to understand the outcome of scattering processes. By comparing measured and calculated scattering distributions, the interaction potentials between the scattering particles might be derived. In this short chapter we will only cover the basic concepts of scattering theory.

9.1 Wave Packets

We now consider the scattering of particles by a fixed "short range" potential. This means the particle is free except for the time of interaction. As a consequence, Coulomb scattering is excluded which requires a special treatment.

Assume that an incident particle with average momentum $\langle \mathbf{p} \rangle = \hbar \mathbf{k}$ at $t = t_0$ is described by a wave packet

$$\Psi(\mathbf{r}, t) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} a_{\mathbf{k}} \tag{9.1}$$

We expand $\Psi(\mathbf{r}, t)$ in exact eigenstates of the potential problem.

$$\Psi(\mathbf{r}, t) = \int \frac{d^3k}{(2\pi)^3} \Psi_{\mathbf{k}}(\mathbf{r}) A_{\mathbf{k}} \tag{9.2}$$

with

$$\left(\frac{\hbar^2}{2m} \nabla^2 + E_{\mathrm{k}} \right) \Psi_{\mathbf{k}}(\mathbf{r}) = V(\mathbf{r}) \Psi_{\mathbf{k}}(\mathbf{r}) \tag{9.3}$$

with $E_k = \frac{\hbar^2 \mathbf{k}^2}{2m} > 0$. At later times $t > t_0$, we have

$$\Psi(\mathbf{r}, t) = \int \frac{d^3k}{(2\pi)^3} \Psi_k(\mathbf{r}) A_{\mathbf{k}} e^{-iE_k(t-t_0)/\hbar} \tag{9.4}$$

We construct the energy eigenstates $\Psi_{\mathbf{k}}$ by using the *Green's functions approach*,

$$\left(\frac{\hbar^2}{2m} \nabla^2 + E_{\mathrm{k}} \right) \mathcal{G}(\mathbf{r}, k) = \delta(\mathbf{r}) \tag{9.5}$$

Insertion into the Schrödinger equation leads to

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \Phi_0(\mathbf{r}) + \int \mathcal{G}(\mathbf{r} - \mathbf{r}', k) V(\mathbf{r}') \, \Psi_{\mathbf{k}}(\mathbf{r}') \, d^3r' \tag{9.6}$$

where $\Phi_0(\mathbf{r})$ is a solution of the homogenous Schrödinger equation, i. e. for a free particle without any potential acting on it.

Proof

Operate with $\frac{\hbar^2}{2m} \nabla^2 + E_{\mathrm{k}}$ on both sides of equation (9.6).

□

We interpret this result as follows: the total wave $\Psi_{\mathbf{k}}(\mathbf{r})$ consists of the incident wave $\Phi_0(\mathbf{r})$ plus the scattered wave. The second term $V(\mathbf{r})\Psi_{\mathbf{k}}(\mathbf{r})$ acts as a source for the scattered wave. The propagator $\mathcal{G}(\mathbf{r}-\mathbf{r}', \mathbf{k})$ gives us the amplitude of the scattered wave at $\mathbf{r}$ due to a unit source at $\mathbf{r}'$. It is given explicitly by solving the equation

$$\left[\frac{\hbar^2}{2m}\nabla^2 + E_k\right]\int \frac{d^3p}{(2\pi)^3}\frac{e^{i\mathbf{p}\cdot\mathbf{r}}}{E_k - \frac{\hbar^2}{2m}\mathbf{p}^2} = \delta(\mathbf{r}) \tag{9.7}$$

$$\mathcal{G}(\mathbf{r},\mathbf{k}) = \int \frac{d^3p}{(2\pi)^3}\frac{e^{i\mathbf{p}\cdot\mathbf{r}}}{E_k - \frac{\hbar^2}{2m}\mathbf{p}^2} = -\frac{m}{2\pi^2 ir\hbar}\int_{-\infty}^{\infty}\frac{pe^{ipr}}{p^2-k^2}\,dp \tag{9.8}$$

The integral is evaluated by the theory of complex functions.

$$\mathcal{G}(\mathbf{r},\mathbf{k}) = -\frac{m}{2\hbar^2\pi^2 ir}\cdot 2\pi i\cdot\frac{ke^{ikr}}{2k} = -\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r} \tag{9.9}$$

Let $\mathbf{r}$ be far from the potential.

$$k\left|\mathbf{r}-\mathbf{r}'\right| \approx kr - k\mathbf{r}'\cdot\hat{r} = kr - \mathbf{k}'\cdot\mathbf{r}'\ , \tag{9.10}$$

where $\mathbf{k}' = k\hat{r}$ is the wave vector in the far field. The asymptotic solutions are the of the form

$$\Psi_k(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{ikr}}{r}f_{\mathbf{k}}(\Omega_r) \tag{9.11}$$

where Ω_r specifies the direction of r. $f_{\mathbf{k}}(\Omega_r)$ is called the *scattering amplitude*

$$f_{\mathbf{k}}(\Omega_r) = -\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{k}'\cdot\mathbf{r}'}V(\mathbf{r}')\Psi_k(\mathbf{r}')\,d^3r' \tag{9.12}$$

Insert (9.11) in (9.4) and use $a_{\mathbf{k}} = A_{\mathbf{k}}$:

$$\Psi(\mathbf{r},t) = \Psi_0(\mathbf{r},t) + \int\frac{d^3k}{(2\pi)^3}a_{\mathbf{k}}\frac{e^{i\left(kr-E_k(t-t_0)/\hbar\right)}}{r}f_{\mathbf{k}}(\Omega_r) \tag{9.13}$$

where

$$\Psi_0(\mathbf{r}) = \int\frac{d^3k}{(2\pi)^3}a_{\mathbf{k}}e^{i(\mathbf{k}\cdot\mathbf{r}-E_k(t-t_0)/\hbar)} \tag{9.14}$$

is the wave packet at time t if there were no potential present.

Now let us take a closer look at the second term of (9.13). Assume that $a_{\mathbf{k}}$ is strongly peaked about $\mathbf{k}_0$ and that $f_{\mathbf{k}}(\Omega_r)$ slowly varies around $\mathbf{k}_0$. Then we can replace $e^{ikr}f_{\mathbf{k}}(\Omega_r)$ by $e^{i\mathbf{k}\cdot\hat{\mathbf{k}}_0 r}f_{\mathbf{k}_0}(\Omega_r)$ and obtain

$$\Psi(\mathbf{r},t) = \Psi_0(\mathbf{r},t) + \frac{f_{\mathbf{k}_0}(\Omega_r)}{r}\Psi_0(\hat{\mathbf{k}}_0 r,t)\ . \tag{9.15}$$

This equation says that the total wave function after scattering is given by the wave packet which one would have in the absence of any scattering plus a scattered part. The scattered term has a pictorial interpretation. $\Psi_0(\hat{\mathbf{k}}_0 r,t)$ corresponds to the value of the wave function at point $\mathbf{r}$ if all the potential did was to bend the trajectory of the particle from the forward direction towards $\mathbf{r}$. The probability amplitude that this bending has occured is just given by the factor $f_{\mathbf{k}_0}(\Omega_r)/r$.

9.2 Cross Sections

Experimental results are usually expressed in terms of cross sections. The differential cross section is defined as

$$\frac{d\sigma}{d\Omega} \equiv \frac{\text{\# of particles scattered into unit solid angle per unit of time}}{\text{\# of incident particles crossing a unit area per unit of time}} \tag{9.16}$$

With respect to the wave function (9.11), we note that the number of incident particles crossing a unit area per unit of time is just proportional to the probability flux due to the first term of Eq. (9.11). Equivalently, the number of particles scattered into an unit solid angle per unit of time is given by the probability flux according to the second term of Eq. (9.11). Thus we have

$$\frac{d\sigma}{d\Omega} = \frac{r^2\,|\mathbf{j}_{\text{scatt}}|}{|\mathbf{j}_{\text{incid}}|} = |f_{\mathbf{k}_0}(\Omega_r)|^2 \ . \tag{9.17}$$

The total cross section given by $d\sigma/d\Omega$ integrated over all angles is the total probability of a particle being scattered with respect to the total probability that crossed a unit area in front of the target. Thus

$$\sigma = \int |f_{\mathbf{k}_0}(\Omega_r)|^2 \ d\Omega \tag{9.18}$$

Note that cross sections have the unit of an area.

9.3 Partial Waves

Consider a spherical symmetric potential. Then it is convenient to expand the plane waves in angular momentum eigenfunctions. Assuming that $\mathbf{k}$ is parallel to the $\hat{z}$-axis, the expansion is given by

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{l=0}^{\infty} i^l(2l+1)P_l(\cos\vartheta)j_l(kr) \\ &= \frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)P_l(\cos\vartheta)\big(h_l(kr)+h_l^*(kr)\big) \end{aligned} \tag{9.19}$$

where j_l is a Bessel function and h_l are the Hankel functions. Only $m = 0$ terms enter because of $\mathbf{k} \parallel \hat{z}$:

$$P_l(\cos\vartheta) = \sqrt{\frac{4\pi}{2l+1}}Y_{l0}(\vartheta,\varphi) \tag{9.20}$$

Consequently, $f_{\mathbf{k}}(\Omega_r)$ depends only on ϑ. Asymptotically, the energy eigenstates of the potential problem can be written as

$$\begin{aligned} \Psi_k(\mathbf{r}) &= \frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)P_l(\cos\vartheta)\big(h_l^*(kr)+e^{i2\delta_l}h_l(kr)\big) \\ &= e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{1}{2}\sum_{l=0}^{\infty} i^l(2l+1)P_l(\cos\vartheta)(e^{i2\delta_l}-1)h_l(kr) \end{aligned} \tag{9.21}$$

where we have used the representation (9.19) for $e^{i\mathbf{k}\cdot\mathbf{r}}$. The factors δ_l appearing in the exponential term are called phase shifts. Thus, the scattering process is completely described in terms of the phase shifts. $e^{i2\delta_l}$ is called the partial wave scattering amplitude.

Total cross sections can be expressed as

$$\sigma = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l = \sum_{l=0}^{\infty}\sigma_l \tag{9.22}$$

One can show that if $0 < \delta_l < \pi/2$, the potential is attractive, i. e. $V < 0$. Here, the frequency of the wave functions increases. If $\delta_l < 0$, the potential is predominantly repulsive.

9.4 Born Approximation

We assume that the potential has only a small effect on all partial waves. Approximate $\Psi_{\mathbf{k}}(\mathbf{r})$ by $e^{i\mathbf{k}\cdot\mathbf{r}}$ in (9.12). Then we have

$$f_{\mathbf{k}}(\Omega_r) = -\frac{m}{2\pi\hbar^2}\int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}'}V(\mathbf{r}')\,d^3r' = -\frac{m}{2\pi\hbar^2}V(\mathbf{k}'-\mathbf{k})\;, \tag{9.23}$$

where $V(\mathbf{k})$ corresponds to the Fourier transform of $V(\mathbf{r})$. This approximation is known as the *Born approximation*. As an example, we will apply this approximation to the so-called *Yukawa potential.*:

$$\begin{aligned} V(\mathbf{r}) = a\,\frac{e^{-\kappa r}}{r} \quad &\Longrightarrow \quad V(\mathbf{q}) = \frac{4\pi a}{q^2+\kappa^2} \\ \Rightarrow \frac{d\sigma}{d\Omega} &= \frac{a^2}{4E_k\sin^2\frac{\vartheta}{2} + \hbar^2\frac{\kappa^2}{2m}} \end{aligned} \tag{9.24}$$

For $a = e^2$ and $\kappa \to 0$ this results reduces to the classical Rutherford cross section for Coulomb scattering. However, this is just accidental since the Born approximation is not valid for the long-range Coulomb potential.