Introduction to Quantum Mechanics

Eva-Maria Graefe
Department of Mathematics
Imperial College London

Academic year 2022/23
Autumn Term

Contents

Part I

Hamiltonian and Schrödinger dynamics

Chapter 1

Classical mechanics

Here we summarise the necessary background from classical mechanics, and introduce some physical terminology that will repeatedly occur in the lecture.

1.1 Newton's second law

A classical particle has a *mass* m, and its *state* is described by its *position* $\vec{r} \in \mathbb{R}^3$ and its *velocity* $\vec{v} = \dot{\vec{r}} = \frac{\mathrm{d}\vec{r}}{\mathrm{d}t}$ or, often more conveniently, its *momentum* $\vec{p} = m\vec{v}$. *Newton's second law* connects the change of momentum to the external force $\vec{F}$:

$$\dot{\vec{p}} = m\ddot{\vec{x}} = \vec{F}. \tag{1.1}$$

This is a second order differential equation for the time-dependent position of a classical particle. The solution is uniquely determined by two initial conditions, for example the position and velocity at a given time. The solution $r(\vec{t})$ is called the *trajectory* of the particle.

It is often convenient to study a system's dynamics in the so-called *phase space*, the space of position and momentum coordinates. A *phase-space trajectory*, $(\vec{r}(t), \vec{p}(t))$ is a curve in phase space that position and momentum coordinates trace as time evolves for a particular initial condition. A *phase (space) portrait* is a sketch of a collection of several phase-space trajectories belonging to different initial conditions.

1.2 Hamilton's canonical equations and phase space-dynamics

1.2.1 Hamilton's canonical equations

In many cases of interest the force can be expressed as the negative gradient of a scalar function $V(\vec{r})$, the *potential*, i.e.,

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}). \tag{1.2}$$

Such systems are referred to as *conservative*, as they *conserve* the *energy*.

Remark 1. Remind yourself of the definitions and the physical units of *force*, *energy*, and *action*.

In what follows we shall for simplicity discuss one-dimensional systems, with coordinate $q \in \mathbb{R}$ and momentum $p \in \mathbb{R}$. For a conservative system Newton's second law (1.1) can then be written as

$$\dot{p} = -\frac{\partial V}{\partial q}, \tag{1.3}$$

and by definition we have

$$\dot{q} = \frac{p}{m}. \tag{1.4}$$

Introducing the function

$$H(p,q) = \frac{p^2}{2m} + V(q), \tag{1.5}$$

the dynamical equations (1.3) and (1.4) can be rewritten as

$$\dot{p} = -\frac{\partial H}{\partial q} \quad \text{and} \quad \dot{q} = \frac{\partial H}{\partial p}. \tag{1.6}$$

These are *Hamilton's canonical equations*. The function $H(p,q)$ is called the *Hamiltonian* function, its value is the total energy of a system for a given state (p,q). The term $\frac{p^2}{2m}$ is the *kinetic energy*, $V(q)$ is the *potential energy*. Hamiltonian dynamics conserves the total energy, i.e., $\frac{\mathrm{d}H}{\mathrm{d}t} = 0$.

Remark 2. Hamiltonian mechanics applies more generally than in situations where $p = m\dot{q}$, i.e., q and p do not have to refer to physical coordinate and momentum in all situations (see *canonically conjugate variables* below). It does only apply in situations where the total energy is conserved. Thus, for example systems with friction or other losses *cannot* be described in the standard Hamiltonian formalism.

1.2.2 The potential and characteristics of the dynamics

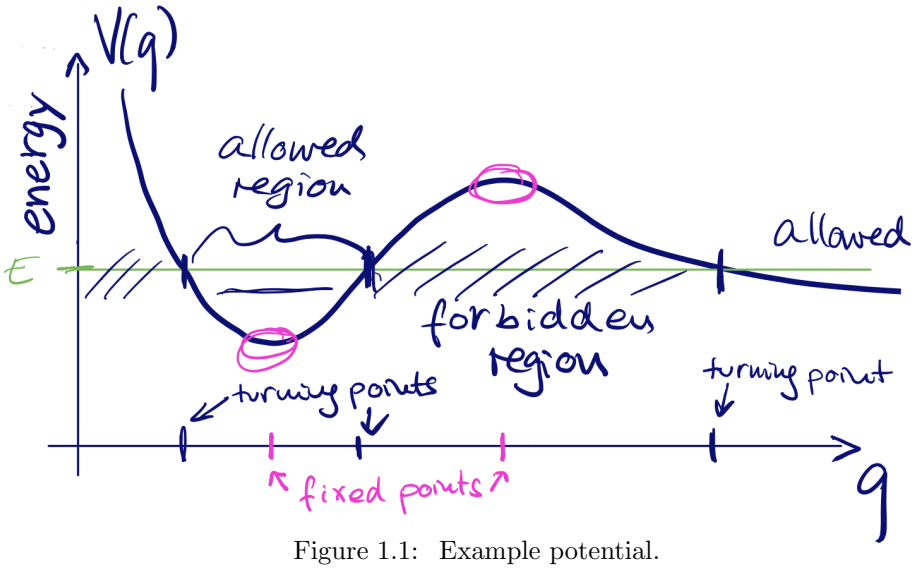
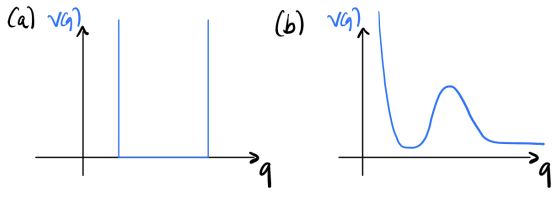


Figure 1.1: Example potential.

For systems whose Hamiltonian is of the form kinetic plus potential energy, as in equation (1.5), one can deduce basic properties of the dynamics from the potential, as sketched in figure 1.1. The extrema of the potential correspond to *fixed points* (or *equilibrium* configurations) of the dynamics. Minima are stable, and maxima are unstable fixed points. For a given energy a particle can only be in a region where its total energy is larger or equal to the potential energy. These regions are called *classically allowed*. They are separated by *turning points* from the *classically forbidden* regions, in which the potential is larger than the total energy available. In a classically allowed region that is bounded by two turning points, a particle oscillates between them. For a region that is only bound by one turning point, the particle comes in from one side, is reflected at the turning point and moves away.

Exercise 1. Sketch the phase space portrait for a particle in one dimension with Hamiltonian of the form $\frac{p^2}{2m} + V(q)$ for the two potentials sketched here on the right.



Example 1 - The free particle. The free particle moves without an external potential, i.e., the Hamiltonian is given by $H = \frac{p^2}{2m}$. Thus, the canonical equations of motion become

$$\dot{p} = 0, \quad \text{and} \quad \dot{q} = \frac{p}{m}, \tag{1.7}$$

and we find $p = const$, and $q(t) = \frac{p}{m}t + q_0$. Sketch the phase space portrait!

Remark 3. $\dot{q} = \frac{p}{m}$ of course always holds for systems whose total energy is of the form $H = \frac{p^2}{2m} + V(q)$.

Remark 4. We could have also used the conservation of energy to deduce that the momentum is conserved here.

Example 2 - The harmonic oscillator. The harmonic oscillator is a particle in a quadratic potential, $V(q) \propto q^2$. It is a very important model system in physics for several reasons. Firstly, it has a simple analytical solution. Secondly, and more importantly, it is a good approximation for many systems that perform small oscillations around an equilibrium configuration. On a mathematical level this approximation can be interpreted as a Taylor expansion of the actual potential up to second order around a minimum, where the first derivative vanishes. The harmonic oscillator is for example a rough approximation for the vibrations of a two-atomic molecule.

Let us here consider the example of a mass m on a spring for small displacements from the equilibrium configuration. At equilibrium the spring counterbalances gravity and there is no net force acting on the mass. If the mass is displaced slightly, the spring responds with a compensation force which is proportional to the displacement (this is known as *Hooke's law*):

$$F = -Dq, \tag{1.8}$$

where D is the spring constant, and q denotes the displacement from equilibrium (i.e. q is positive when the spring is stretched). It is convenient to define a frequency $\omega = \sqrt{\frac{D}{m}}$. The potential is then given by

$$V(q) = \frac{D}{2}q^2 = \frac{m\omega^2}{2}q^2, \tag{1.9}$$

and we have the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2. \tag{1.10}$$

From the conservation of energy we can deduce the phase space portrait as consisting of concentric ellipses.

The equations of motion

$$\dot{p} = -m\omega^2 q, \quad \dot{q} = \frac{p}{m}, \tag{1.11}$$

can be combined into a single second-order differential equation for q,

$$\ddot{q} = -\omega^2 q. \tag{1.12}$$

With the initial conditions $q(t = 0) = q_0$ and $p(t = 0) = p_0$ the solution is given by

$$q(t) = q_0 \cos(\omega t) + \frac{p_0}{m\omega} \sin(\omega t) \tag{1.13}$$

$$p(t) = p_0 \cos(\omega t) - q_0 m\omega \sin(\omega t). \tag{1.14}$$

These equations describe *harmonic* oscillations.

1.3 Time evolution of general dynamical variables - Poisson brackets

The time evolution of any dynamical variable $A(p,q)$ is governed by the equation

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial q}\dot{q} + \frac{\partial A}{\partial p}\dot{p} + \frac{\partial A}{\partial t}. \tag{1.15}$$

Remark 5. For a quantity without explicit time dependence we have $\frac{\partial A}{\partial t} = 0$. All $A(p,q)$ are implicitly time dependent via the time dependence of p and q. Some variables, however, might in addition have an explicit time dependence. Take for example the potential energy in an oscillating potential of the form $V(q,t) = (q - cos(\omega t))^2$.

Inserting Hamilton's canonical equations (1.6) into (1.15) we find

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial A}{\partial t} = \{H, A\} + \frac{\partial A}{\partial t}, \tag{1.16}$$

where we have defined the *Poisson bracket*, $\{\cdot,\cdot\}$ as

$$\{A, B\} = \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} - \frac{\partial A}{\partial q}\frac{\partial B}{\partial p}. \tag{1.17}$$

For N-dimensional systems the Poisson bracket is defined as

$$\{A, B\} = \sum_{j=1}^{N} \frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j} - \frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j}. \tag{1.18}$$

Remark 6. Note that both this definition and the one where p and q are interchanged are used in the literature.

A dynamical variable without explicit time dependence is conserved in time if $\{H, A\} = 0$, such quantities are also referred to as *constants of motion*. For the special case that H does not depend on one of the canonical variables q_j the corresponding p_j is conserved. Then q_j is called a *cyclic variable*.

1.3.1 Properties of Poisson brackets

A bracket $\{\cdot,\cdot\}$ that maps two dynamical variables to another dynamical variable is called a *Poisson bracket* if it fulfils the following requirements:

(i) $\{A, A\} = 0$ for all A

(ii) Linearity in the first element: $\{c_1 A + c_2 B, C\} = c_1\{A, C\} + c_2\{B, C\}$

(iii) Anti-symmetry with respect to interchange of elements: $\{A, B\} = -\{B, A\}$

(iv) Poisson-bracket with constant vanishes: $\{c, A\} = 0$, for c constant

(v) Leibniz rule (product rule): $\{AB, C\} = A\{B, C\} + \{A, C\}B$

(vi) Jacobi identity: $\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0$.

The Poisson bracket between momentum and position coordinates $\{p_j, q_k\} = \delta_{jk}$ is called the *fundamental* Poisson bracket. Every pair $P_j(p_l, q_m)$, $Q_j(p_l, q_m)$ that fulfils $\{P_j, Q_j\} = 1$ is a pair of *canonically conjugate variables* and one can use the Hamiltonian formalism where $H(P_j, Q_j)$ is the total energy of the system, even if this is different from $\frac{P^2}{2m} + V(Q)$. In general of course this also means that $\dot{Q} \neq \frac{P}{m}$.

Exercise 2. Verify that the Poisson bracket defined in (1.17) indeed fulfils all the above properties.

1.4 Historic excursion: The birth of quantum mechanics

In the late 19th century physics was regarded as "complete", in the sense that it was believed that Newton's laws, together with Maxwell's theory of electromagnetism explained *everything*, and all that was left was applying these theories to various situations. However, at the turn of the century some experiments were under way that revealed a fundamental lack of understanding. These included the measurement of atomic spectra, scattering experiments, the photo-electric effect, and the study of the light emitted from hot objects. These experiments and the attempts to build a corresponding theory led to the biggest revolution in modern physics, and the development of quantum mechanics.

Quantum theory is often divided into *old* and *modern* quantum theory. The name *old quantum theory* refers to the theory used in the initial 25 years of quantum theory, for example Bohr's theory of the atom, where electrons are confined to discrete orbits around the atomic nucleus, while there is no actual theoretical reason for this phenomenon given. While *old* quantum theory explained many of the experimental results, and introduced the concept of *energy quantisation*, there was no deeper understanding of the phenomena, and there were a number of unsolved paradoxes. In 1924/25 several breakthroughs were made, which led to the modern formulation of quantum theory (which, I mention such as not to mislead you, still has a lot of unresolved aspects, but we shall not focus on these.).

Werner Heisenberg (who was only 24 then) came up with a strange dynamical equation for observables (measurable quantities) that do not commute:

$$\dot{a} = \frac{2\pi \mathrm{i}}{h}(Wa - aW),$$

where a is an arbitrary observable (for example the position of a particle), and W denotes the energy. It later turned out that the idea of non-commutative objects is of course closely related to matrices, which were not very popular/well-known amongst physicists in the 1920's.

In 1926 Schrödinger came up with his famous wave equation,

$$\mathrm{i}\frac{h}{2\pi}\dot{\psi}(\vec{r},t) = \left(-\frac{h^2}{8\pi^2 m}\nabla^2 + V(\vec{r})\right)\psi(\vec{r},t),$$

where the complex function $\psi(\vec{r},t)$ is meant to somehow encode the position of a quantum particle, and the operator $-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})$ represents the energy. We shall discuss this further in section 2.

The fact that Schrödinger's and Heisenberg's theories are actually two different formulations of the same theory was not immediately clear. Around the same time as Schrödinger formulated his wave-theory, Dirac developed an abstract formulation of quantum theory, which unifies both Heisenberg and Schrödinger picture. All three of them received Nobel prizes for the development of quantum theory (Heisenberg in 1932, Dirac and Schrödinger jointly in 1933).

The development of the theory was very exciting for everyone involved. Legend has it that in the summer of 1925 Heisenberg was thinking hard about quantum mechanics but didn't make much progress. In addition he was suffering from bad hay fever, and decided to escape to Helgoland, an island in the northern sea. There he went for long hikes, read poetry, and also finally found his quantum theory! Later he wrote about this in his autobiography: *It was about 3 o'clock at night when the final result of the calculation lay before me. At first I was deeply shaken. I was so excited that I could not think of sleep. So I left the house and awaited the sunrise on the top of a rock.* He later sent his results to Wolfgang Pauli, with a letter saying: *Everything is still vague and unclear to me, but it seems as if the electrons will no longer move on orbits...*

Chapter 2

Schrödinger dynamics

In this chapter we give a first introductory overview over Schrödinger's wave equation, the meaning and interpretation of the wave function etc. Much of the content will be formalised in chapter 4.

2.1 The Schrödinger equation and the wave function

Parallel to Heisenberg's formulation of quantum mechanics Schrödinger developed a description in terms of the *wave function*. He was inspired by Einstein's and de Broglie's ideas of *wave-particle-duality* and *matter waves*. In modern form this *time-dependent Schrödinger equation* is written as

$$\mathrm{i}\hbar\dot{\psi}(\vec{r},t) = \hat{H}\psi(\vec{r},t), \tag{2.1}$$

where the dot denotes the derivative with respect to time. Here $\hat{H}$ is the *Hamiltonian*, a differential operator acting on the *wave function* ψ, and $\hbar = \frac{h}{2\pi}$, where $h \approx 6.626 \times 10^{-34}$Js is Planck's constant. The Hamiltonian encodes the total energy of the system and is often of the form

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}), \tag{2.2}$$

where ∇^2 denotes the Laplacian, i.e. in three dimensions in Cartesian coordinates, $\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$. The first term encodes the kinetic and the second the potential energy. We will often consider one-dimensional systems for which the Schrödinger equation with a Hamiltonian of the form (2.2) takes the form

$$\mathrm{i}\hbar\dot{\psi}(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x,t). \tag{2.3}$$

Remark 7. In the context of quantum mechanics, we usually denote the spatial coordinate of one-dimensional systems, the argument of the wave function, by x.

2.1.1 The wave function

The wave function $\psi \in \mathbb{C}$ is interpreted as a *probability amplitude*. The square modulus of the wave function is a probability density. Specifically, $|\psi(x,t)|^2\mathrm{d}x$ is the probability of finding a quantum particle (described by the wave function ψ) between x and $x+\mathrm{d}x$ at time t (assuming ψ is *normalised* to one, that is, $\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x = 1$.). The normalisation condition $\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x = 1$ means that the particle is *somewhere* with probability one. The probability to find the particle in the interval $A = [a,b]$ is given by

$$\int_a^b |\psi(x)|^2\mathrm{d}x. \tag{2.4}$$

The wave function encodes everything that one can possibly know about a quantum particle. Knowing $\psi(x,t_0)$ one can deduce $\psi(x,t)$ for all times. That is, the time evolution in quantum mechanics is *deterministic*.

To interpret $|\psi(x,t)|^2$ as a probability distribution we need $\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x$ to be finite. That is, $\psi(x,t)\in L^2$ (we refer to this as *normalisable* or *square integrable*). here we will mostly consider such functions as wave functions, but the theory can be slightly extended to include, for example, functions such as *plane waves*, $\psi\propto \mathrm{e}^{\mathrm{i}kx}$. We shall occasionally use such functions, and fill in a minimum of interpretational background when needed.

Exercise 3. Consider the wave function

$$\psi(x,t)=A\mathrm{e}^{-\lambda|x|}\mathrm{e}^{-\mathrm{i}\omega t},$$

where A,λ,ω are real positive constants.

(a) Find the value of A for which ψ is normalised to one.

(b) What is the probability to find the particle described by $\psi(x,t)$ in the interval $[-\frac{1}{\lambda},\frac{1}{\lambda}]$?

2.1.2 Probability current and conservation of probability

The total probability is conserved in time. That is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x=0. \tag{2.5}$$

We can prove this by using the Schrödinger equation (2.3), and assuming that $\psi\in L^2$: The derivative with respect to time and the integral over the coordinate x commute, that is, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x &= \int_{-\infty}^{\infty}\frac{\mathrm{d}}{\mathrm{d}t}|\psi(x,t)|^2\mathrm{d}x\\ &= \int_{-\infty}^{\infty}\dot\psi^*(x,t)\psi(x,t)+\psi^*(x,t)\dot\psi(x,t)\mathrm{d}x,\end{aligned} \tag{2.6}$$

where the asterisk * denotes complex conjugation. Inserting the Schrödinger equation and its conjugate yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x &= \frac{\mathrm{i}}{\hbar}\int_{-\infty}^{\infty}V(x)\psi^*(x,t)\psi(x,t)-\frac{\hbar^2}{2m}\psi''^*(x,t)\psi(x,t)\mathrm{d}x\\ &\quad -\frac{\mathrm{i}}{\hbar}\int_{-\infty}^{\infty}V(x)\psi^*(x,t)\psi(x,t)-\frac{\hbar^2}{2m}\psi^*(x,t)\psi''(x,t)\mathrm{d}x\\ &= \frac{\mathrm{i}\hbar}{2m}\int_{-\infty}^{\infty}\psi^*(x,t)\psi''(x,t)-\psi''^*(x,t)\psi(x,t)\mathrm{d}x,\end{aligned} \tag{2.7}$$

where we have made use of the fact that the potential $V(x)$ is a real function. Now we recognise the integrand as the first derivative of a function which is proportiional to the so-called *probability current* or *probability flux*, defined as

$$j(x,t):=\frac{\hbar}{2m\mathrm{i}}\left(\psi^*(x,t)\psi'(x,t)-\psi'^*(x,t)\psi(x,t)\right). \tag{2.8}$$

We can integrate to find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}|\psi(x,t)|^2\mathrm{d}x = \frac{\mathrm{i}\hbar}{2m}\left[\psi^*(x,t)\psi'(x,t)-\psi'^*(x,t)\psi(x,t)\right]_{x=-\infty}^{x=\infty}. \tag{2.9}$$

Since $\psi(x,t)\in L^2$, this indeed vanishes, and the overall probability is conserved in time.

Denoting the probability density $|\psi(x,t)|^2$ by $\rho(x,t)$, we can formulate the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0. \tag{2.10}$$

Let us consider an interval $[a,b] \in \mathbb{R}$. The probability to find the particle in this interval at a given time t is $P_{[a,b]}(t) = \int_a^b \rho(x,t)\mathrm{d}t$. The change of this probability in time is given by

$$\frac{\mathrm{d}P_{[a,b]}}{\mathrm{d}t} = \int_a^b \frac{\partial \rho}{\partial t}\mathrm{d}x = -\int_a^b \frac{\partial j(x,t)}{\partial x}\mathrm{d}x = j(a) - j(b). \tag{2.11}$$

That is, the change of probability within the interval equals the difference of the current flowing in on the left side and the current flowing out on the right side. This is in complete analogy to the continuity property of the flow of a fluid or electromagnetic charge.

In three spatial dimensions the probability current is defined as

$$\vec{j}(\vec{r},t) = \frac{\hbar}{2m\mathrm{i}}(\psi^*(\vec{r},t)\nabla\psi(\vec{r},t) - \psi(\vec{r},t)\nabla\psi^*(\vec{r},t)), \tag{2.12}$$

where ∇ denotes the gradient. The continuity equation in three dimensions reads

$$\frac{\partial \rho(x,t)}{\partial t} + \nabla \cdot \vec{j}(\vec{r},t) = 0. \tag{2.13}$$

The probability to find a particle in a certain volume V, $P_V(t) = \int_V \rho(\vec{r},t)\mathrm{d}^3x$, changes in time according to

$$\frac{\mathrm{d}P_V(t)}{\mathrm{d}t} = -\int_V \nabla \cdot \vec{j}(\vec{r},t)\mathrm{d}^3x = -\int_S \vec{j}(\vec{r},t)\cdot \mathrm{d}\vec{s}, \tag{2.14}$$

where we have used Gauss' law to express this as the surface integral over the boundary S of the volume V, and $\mathrm{d}\vec{s}$ denotes the area element. That is, the probability that a particle can be found inside the volume V changes when there is a flux of probability through the boundary of the volume.

2.1.3 Superposition principle

The Schrödinger equation is a linear differential equation. Thus, if $\psi_1(x,t)$ and $\psi_2(x,t)$ are solutions, and $c_1, c_2 \in \mathbb{C}$ are constants, then $c_1\psi_1(x,t) + c_2\psi_2(x,t)$ is also a solutions. (Verify this!) This perhaps trivial sounding fact plays a central role in quantum theory. On a formal level it makes quantum mechanics much "easier" than classical mechanics.

2.1.4 The Hamiltonian and the energy

A quantum system is characterised by its *Hamiltonian* $\hat{H}$, an operator acting on the wave function, whose *expectation value*, defined as

$$\langle \hat{H} \rangle := \frac{\int_{-\infty}^{\infty} \psi^*(x,t)\hat{H}\psi(x,t)\mathrm{d}x}{\int_{-\infty}^{\infty} |\psi(x,t)|^2 \mathrm{d}x}, \quad \text{(and analogously for higher dimensions)} \tag{2.15}$$

gives the *total energy* of the system.

The *eigenenergies*, i.e. the eigenvalues of the Hamiltonian, are real numbers, and the corresponding *eigenstates* (i.e. eigenfunctions) are an *orthonormal* basis for all possible states of a system. The eigenstates $\phi_E(\vec{r})$ are the solutions of the *time-independent Schrödinger equation*

$$\hat{H}\phi_E(\vec{r}) = E\phi_E(\vec{r}). \tag{2.16}$$

Here *orthonormality* means that

$$\int_{\text{all space}} \phi_E^*(\vec{r})\phi_{E'}(\vec{r})\mathrm{d}V = \delta_{EE'}. \tag{2.17}$$

Typically the set of the eigenvalues has a discrete part and a continuum. We shall discuss examples later in the course, specifically in part III.

2.2 Stationary solutions: The time-independent Schrödinger equation

Consider special solutions of the time-dependent Schrödinger equation of the form

$$\psi(x,t) = \phi(x)\chi(t). \tag{2.18}$$

Inserting this ansatz into the Schrödinger equation yields

$$\mathrm{i}\hbar\frac{\dot{\chi}(t)}{\chi(t)} = \frac{\hat{H}\phi(x)}{\phi(x)}. \tag{2.19}$$

Since the left side of this equation is independent of x, so has to be the right side. Similarly, as the right side is independent of t, so is the left side, and in fact, both sides have to be constant. We call the constant E and separate the original equation into two differential equations

$$\mathrm{i}\hbar\dot{\chi}(t) = E\chi(t) \tag{2.20}$$

$$\hat{H}\phi(x) = E\phi(x). \tag{2.21}$$

Remark 8. This method is known as *separation of variables*.

Equation (2.20) is easily solved by direct integration to yield $\chi(t) = \chi(0)\mathrm{e}^{-\mathrm{i}Et/\hbar}$. Equation (2.21) is just the time-independent Schrödinger equation, that is, the eigenvalue equation for $\hat{H}$. The functions $\phi_E(x)$ are the eigenfunctions of $\hat{H}$ to the eigenvalues E. They are also referred to as *eigenstates*. We will solve many examples of this, especially in part III.

The eigenvalue equation can only be solved analytically for a small number of textbook cases. For Hamiltonians of the form $-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$, where the potential has the property $V(x \to \pm\infty) \to \infty$ normalisable solutions to the time-independent Schrödinger equation only exist for special discrete values of E.

Remark 9. These discrete energies are the reason for the discrete atomic spectra, that is, the distinct colours that are emitted by stimulated atoms (think for example of fluorescent lamps).

The time-dependence of separable wave functions is rather trivial:

$$\psi(x,t) = \chi_0\mathrm{e}^{-\mathrm{i}Et/\hbar}\phi_E(x). \tag{2.22}$$

In particular, if the initial wave function at time zero is an eigenfunction of $\hat{H}$ we have $\chi_0 = 1$ and

$$\psi(x,t) = \mathrm{e}^{-\mathrm{i}Et/\hbar}\phi_E(x), \tag{2.23}$$

that is, an initial eigenstate evolves in time only by acquiring a phase, which means for the probability density

$$|\psi(x,t)|^2 = |\phi_E(x)|^2. \tag{2.24}$$

Hence, these wave functions are called *stationary states*.

Linear superpositions of stationary states, however, do have a non-trivial time-dependence. Consider for example an initial state which is a superposition of two eigenstates belonging to two different energies E_1 and E_2,

$$\psi(x,0) = a\phi_{E_1}(x) + b\phi_{E_2}(x).$$

The probability density of the time-dependent state is then given by

$$\begin{aligned}|\psi(x,t)|^2 =& |a|^2|\phi_{E_1}(x)|^2 + |b|^2|\phi_{E_2}(x)|^2 \\ &+ a^*b\,\mathrm{e}^{\mathrm{i}(E_2-E_1)t/\hbar}\phi^*_{E_1}(x)\phi_{E_2}(x) + ab^*\,\mathrm{e}^{\mathrm{i}(E_1-E_2)t/\hbar}\phi_{E_1}(x)\phi^*_{E_2}(x)\end{aligned} \tag{2.25}$$

The last two terms are time-dependent and are often referred to as the *interference terms*. In the special case where $a, b \in \mathbb{R}$, and also the eigenfunctions are real valued, $\phi_{E_j} \in \mathbb{R}$, this reduces to

$$|\psi(x,t)|^2 = a^2|\phi_{E_1}(x)|^2 + b^2|\phi_{E_2}(x)|^2 + 2ab\cos(\tfrac{E_2-E_1}{\hbar}t)\,\phi_{E_1}(x)\phi_{E_2}(x), \tag{2.26}$$

that is, the probability oscillates with a frequency determined by the energy difference of the two states in the superposition.

2.3 The method of stationary states

Consider an initial state that is an arbitrary superposition of eigenstates of the Hamiltonian

$$\psi(x, t = 0) = \sum_n a_n \phi_n(x), \tag{2.27}$$

with

$$\hat{H}\phi_n(x) = E_n \phi_n(x).$$

Due to the linearity of the Schrödinger equation, the time-dependent state is given by

$$\psi(x, t) = \sum_n a_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \phi_n(x). \tag{2.28}$$

Remark 10. That is the superposition principle again!

Given that the eigenstates of a Hamiltonian are a basis for all possible states of a system, and the superposition principle, the time-evolution of an arbitrary initial state can be deduced by representing it in the eigenbasis. This method is often called the *method of stationary states*.

2.3.1 The method of stationary states (for one dimensional systems in coordinate representation)

Given: Hamiltonian $\hat{H}$, initial wave function $\psi(x, t = 0)$
Wanted: Wave function at time t, $\psi(x, t)$

Method:

1. Calculate the eigenvalues E_n and normalised eigenstates $\phi_n(x)$ of the Hamiltonian $\hat{H}$.

2. Expand the initial wave function in the basis of eigenstates $\{\phi_n(x)\}$:

$$\psi(x, 0) = \sum_n a_n \phi_n(x).$$

   The coefficients are given by $a_n = \int_{-\infty}^{\infty} \phi_n^*(x)\psi(x, 0)dx$.

3. The time evolution is then given by

$$\psi(x, t) = \sum_n a_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \phi_n(x).$$

2.4 Example: Particle in a half box

(From David Tannor's Introduction to Quantum Mechanics)

A standard system in quantum mechanics is the so-called *particle in a box*. In one dimension that is a particle moving freely on a certain interval on the real line, here we consider the interval $[0, L]$. The Hamiltonian of the system is given by $\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$, where there is an infinite potential outside this interval, i.e.

$$V(x) = \begin{cases} 0, & 0 \le x \le L \\ \infty, & \text{otherwise.} \end{cases}$$

We shall discuss the solution of the time-independent Schrödinger equation for this system in chapter 6. For now we simply provide the result: The eigenvalues of the Hamiltonian are given by

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}, \tag{2.29}$$

and the corresponding eigenfunctions are

$$\phi_n(x) = \begin{cases} \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right), & 0 \le x \le L \\ 0, \quad \text{otherwise.} \end{cases} \tag{2.30}$$

The first 5 eigenfunctions are schematically depicted at the height of their corresponding energies in figure 2.1.

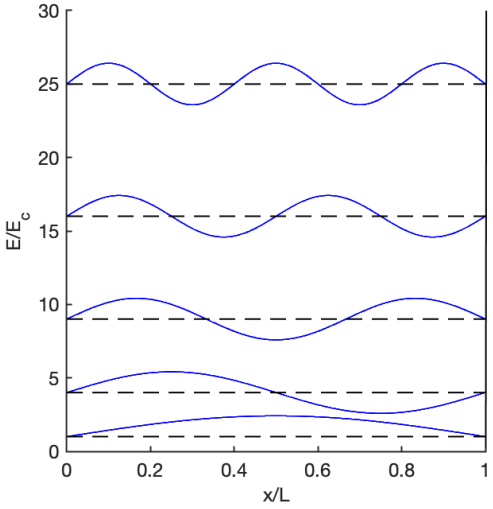


Figure 2.1: Eigenvalues and eigenfunctions of a particle in a box.

Remark 11. This is analogous to the problem of standing waves on a string.

Consider a quantum particle in the lowest eigenstate of a box potential of length L. Assume now that at time $t = 0$ the box is suddenly expanded to the length $2L$. If this is done infinitely fast, the wavefunction is initially not altered. However, the wavefunction is now not an eigenfunction of the new system, and thus it dynamically evolves for times $t > 0$.

Remark 12. Such a fast change of the external potential in an experiment is often called a *quantum quench*, a method often employed to experimentally prepare a system in a state other than an eigenstate.

We can use the method of stationary states to understand the dynamics of the wave function. It turns out that the motion of the wave function while being rather complicated in its details, is periodic. Let us use the method of stationary states to calculate the shortest time τ at which $\psi(x,\tau) = \psi(x,0)$. For this purpose we need to expand the wave function at time zero, which is a normalised eigenfunction of the square well of length L,

$$\psi(x,0) = \phi_0^{(L)}(x), \tag{2.31}$$

in the basis of eigenfunctions $\phi_n^{(2L)}(x)$ of the square well of twice the size, that is, we write

$$\phi_0^{(L)}(x) = \sum_n a_n \phi_n^{(2L)}(x). \tag{2.32}$$

To calculate the actual coefficients a_n we would have to calculate the integrals

$$a_n = \int_{-\infty}^{\infty} \phi_n^{*(2L)}(x)\phi_0^{(L)}(x)\mathrm{d}x = \frac{\sqrt{2}}{L}\int_0^L \sin\left(\frac{n\pi x}{2L}\right)\sin\left(\frac{n\pi x}{L}\right)\mathrm{d}x. \tag{2.33}$$

However, to calculate τ we do not actually need to now all the values of a_n! We know that the wave function at time t has the form

$$\psi(x,t) = \sum_n a_n \mathrm{e}^{-\mathrm{i}E_n^{(2L)}t/\hbar}\phi_n^{(2L)}(x). \tag{2.34}$$

Inserting the energy values for the square well of width $2L$, we have

$$\psi(x,t) = \sum_n a_n \mathrm{e}^{-\mathrm{i}\frac{\pi^2\hbar}{8mL^2}n^2 t}\phi_n^{(2L)}(x). \tag{2.35}$$

Or

$$\psi(x,t) = \sum_n a_n \mathrm{e}^{-\mathrm{i}\omega n^2 t} \phi_n^{(2L)}(x), \tag{2.36}$$

where we have defined $\omega = \frac{\pi^2\hbar}{8mL^2}$. Thus, we have $\psi(x,\tau) = \psi(x,0)$ at times τ for which all $\mathrm{e}^{-\mathrm{i}\omega n^2\tau} = 1$, which is true if all $n^2\omega\tau$ are integer multiples of 2π. Since n are integers here, the shortest such time τ thus fulfils

$$\tau = \frac{2\pi}{\omega} = \frac{16mL^2}{\hbar\pi}. \tag{2.37}$$

Part II

The Principles of quantum mechanics

Chapter 3

Excursion: Mathematical background

After having gained some first impressions of quantum behaviour looking at Schrödinger's wave functions, we shall now move to a more abstract and versatile formulation of quantum mechanics, that will allow us to identify which questions can be answered by quantum mechanics, and which questions the theory does not "allow" us to ask. In particular in chapter 4 we will learn five principles/postulates of quantum mechanics that form the basis of the whole theory and its interpretations. Several mathematical concepts and notations are needed to formulate and understand these principles, and in the present chapter we provide the minimum background. Let us first, however, have a brief look into the first two postulates that we will discuss in more detail in chapter 4 to get an idea which mathematical background is needed and why.

3.1 States and obervables - Hilbert spaces and Hermitian operators

The main ingredients for this theoretical framework of quantum mechanics are the concepts of a state and of observables. A *state* encodes all the information one can have about a quantum system. *Observables* are measurable quantities (such as, for example, the position of a particle or its kinetic energy). In some ways, the same concepts appear in classical mechanics, but we so naturally associate the mathematical objects that describe these properties with the actual physical properties that they are seldom talked about.

Looking back at Hamilton's classical mechanics we see that all measurable quantities, i.e. *observables* are given by real phase-space functions, that is, real functions of the canonical variables. The canonical variables, in turn, are the classical *state*. We implicitly consider this state to be a property of the system (for example a particle) in classical physics, and talk about the *position of an object* or the *momentum of a particle*, in similar terms as when considering its mass, colour or other inherent property. In quantum mechanics we will have to let go of this.

The state of a quantum system, as defined above, encodes all we can in principle measure about a particle, but it makes no claim about measurable quantities being *properties of* a particle. We have already encountered this in the last chapter where we have seen that it makes little sense in Schrödinger's framework of quantum mechanics to talk about a *particle's position*. The only thing we can talk about is the *probability of finding a particle in a certain region at a given time.* Think this over - it is the essence of quantum mechanics!

As Philip Ball argues in his beautiful recent book "Beyond weird", this is neither weird nor strange. It does, however, require a shift away from century-old ways of thinking about the nature of things. This allows us to predict a lot of experimental observations with unprecedented accuracy, and to design technological devices such as transistors and lasers, that would be impossible on the basis of classical physics. A lot can be said (and over the decades a lot has been said, indeed) about the philosophical aspects of quantum theory. Here, however is not the place to do so. Let us instead focus on the mathematical framework and on what we *can* calculate with it.

As we shall see, compared to classical mechanics, the role of position and momentum is shifted from the state to observables. The role of the state is taken over by another mathematical object that does not appear explicitly in classical physics. That brings us to

The first principle of quantum mechanics: The state of a quantum system is described by a non-zero vector in a Hilbert space.

Observables are covered by

The second principle of quantum mechanics: A measurable quantity is described by a Hermitian operator on the Hilbert space. Both the measurable quantities and the operators representing them are often referred to as *observables*.

With this in mind in this chapter we shall review some general concepts from linear algebra, and in particular provide a minimal background on Hilbert spaces and Hermitian operators. We shall also learn about a very powerful notation for vectors and operators on Hilbert spaces that has been developed by Dirac especially for quantum mechanics, aptly called the *Dirac notation*, that we shall make use of in the rest of the course.

3.2 Some linear algebra

Here we summarise a few definitions from linear algebra that are essential in the abstract formulation of quantum mechanics.

Complex vector space: A set of elements (*vectors*) V, forming an abelian group with respect to addition $u+v$, and a scalar multiplication $\mathbb{C}\times V\to V$, satisfying

$$\begin{aligned}(\lambda+\mu)v &= \lambda v+\mu v\\ (\lambda\mu)v &= \lambda(\mu v)\\ \lambda(u+v) &= \lambda u+\lambda v\\ 1v &= v.\end{aligned}$$

for all $\lambda,\mu\in\mathbb{C}$ and $u,v\in V$, is called a vector space.

Remark 13. In general a vector space can be defined over any field $\mathbb{F}$, where the scalars are elements of $\mathbb{F}$. Most commonly encountered in physics are vector spaces over the real or the complex numbers.

Inner product space: An inner product space is a vector space V together with a map $V\times V\to\mathbb{C}$, written as (u,v), such that for all $u,v,w\in V$ and $\lambda,\mu\in\mathbb{C}$:

$$\begin{aligned}(u,v) &= (v,u)^*\\ (u,u) &\geq 0,\quad \text{and}\quad (u,u)=0\Leftrightarrow u=0\\ (u,\lambda v+\mu w) &= \lambda(u,v)+\mu(u,w).\end{aligned}$$

Linear (in)dependence, basis, dimension: A set of vectors $\{u_n\}$ is said to be *linearly independent* if no nontrivial linear combination of them sums to zero. If the set is on the contrary *linearly dependent*, one can express one member of the set as a linear combination of the others: $u_k=\sum_{n\neq k}\alpha_n u_n$.

The maximum number of linearly independent vectors in a space is called the *dimension* of the space.

A maximal set of linearly independent vectors is called a *basis*. That is, every vector in the space can be represented as a linear combination of the basis vectors: $\phi=\sum_n\phi_n u_n$. The ϕ_n are the *coefficients* or *components* of ϕ in the basis $\{u_n\}$. One can represent ϕ as an N-tupel of complex numbers, $(\phi_1,\dots,\phi_N)\in\mathbb{C}^N$, where N is the dimension of the vector space.

Norm: The norm $||v||$ of a vector can be defined via $||v||^2 = (v, v)$. A vector is called *normalised* if $||v|| = 1$.

Orthonormality: Two vectors u, v are called *orthogonal* if $(u, v) = 0$. A set of vectors $\{u_n\}$ is said to be *orthonormal* if each vector is normalised and the vectors are pairwise orthogonal.

An orthonormal maximal set of linearly independent vectors is an *orthonormal basis*.

Coefficients: The coefficients of a vector ϕ in an orthonormal basis $\{u_n\}$ can be found via the orthonormality condition as $\phi_n = (u_n, \phi)$.

Hilbert space: A Hilbert space $\mathcal{H}$ is a complex inner product space which is complete and contains a countable dense subset.

Completeness of a vector space: A vector space $\mathcal{H}$ is complete if the following holds. Whenever $\{u_n \in \mathcal{H}\}$ is a sequence of vectors such that for any positive ϵ there exists an integer N_ϵ with

$$||u_m - u_n|| < \epsilon, \quad \text{for all } m, n > N_\epsilon,$$

then there exists a limit vector $u \in \mathcal{H}$, such that

$$||u_n - u|| \to 0.$$

Dense subset: A subset S of $\mathcal{H}$ is *dense* if for any vector $u \in \mathcal{H}$ and any $\epsilon > 0$ there is a vector $v \in S$ such that

$$||v - u|| < \epsilon.$$

A finite dimensional vector space is automatically complete and contains a dense countable subset, i.e. every finite dimensional vector space is a Hilbert space.

Linear Operator: An operator on a vector space is a map $V \to V$. A *linear operator* A satisfies

$$A(\lambda u + \mu v) = \lambda(Au) + \mu(Av),$$

for all $\lambda, \mu \in \mathbb{C}$, and $u, v \in V$.

A linear operator on a finite dimensional vector space can be represented by a matrix.

Adjoint Operator: The adjoint $A^\dagger$ of a linear operator A is defined by the relation

$$(A^\dagger u, v) = (u, Av), \quad \text{for all } v, u \in V.$$

The adjoint fulfils the following properties:

$$\begin{aligned}(cA)^\dagger &= c^* A^\dagger \\ (A + B)^\dagger &= A^\dagger + B^\dagger \\ (AB)^\dagger &= B^\dagger A^\dagger.\end{aligned}$$

Self-adjointness and Hermiticity: A linear operator defined on a Hilbert space $\mathcal{H}$ is called *Hermitian* or *self-adjoint* if

$$(Au, v) = (u, Av), \quad \text{for all } v, u \in \mathcal{H},$$

that is if $A^\dagger = A$.

For infinite dimensional Hilbert spaces there are subtle differences between self-adjointness and Hermiticity, but we won't go into these. For more details, see, e.g. Ballentine or Cohen Tannoudji.

Eigenvalues and eigenvectors: If an operator acting on a certain vector produces a scalar multiple of that same vector,

$$Au = au,$$

we call the vector u an *eigenvector* and the scalar a an *eigenvalue* of the operator A.

If A is an $N \times N$ matrix the eigenvalue equation only has nontrivial solutions if

$$\det(A - a\mathbb{1}) = 0.$$

This yields the *characteristic polynomial of* A, which is a polynomial of Nth order in a.

3.3 The Dirac notation

3.3.1 The dual space and the Riesz theorem

To any linear vector space V there exists a *dual space* of *linear functionals* on V. A linear functional F assigns a scalar $F(\phi)$ to each vector ϕ, such that

$$F(\alpha\phi + \beta\psi) = \alpha F(\phi) + \beta F(\psi), \tag{3.1}$$

for $\alpha, \beta \in \mathbb{C}$ and $\phi, \psi \in V$.

The set of linear functionals is itself a linear vector space V' with the definition

$$(F_1 + F_2)(\phi) = F_1(\phi) + F_2(\phi). \tag{3.2}$$

Theorem 1 (Riesz representation theorem)**.** *Any Hilbert space $\mathcal{H}$ is (anti-)isomorphic to its dual space $\mathcal{H}'$. That is, there exists a one-to-one correspondence between linear functionals F and vectors f, such that all linear functionals can be represented in the form $F(\phi) = (f, \phi)$ with a fixed vector f and arbitrary vectors ϕ.*

We shall only proof the Riesz theorem for finite dimensions here. Clearly every vector defines a functional via the inner product. The proof for the opposite direction, that every functional can be represented as an inner product with a fixed vector, works by construction:

Proof. Let $\{\phi_n\}$ be an orthonormal basis of $\mathcal{H}$. Let $\psi = \sum_n \psi_n \phi_n$ with $\psi_n \in \mathbb{C}$ be an arbitrary vector in $\mathcal{H}$. Then

$$F(\psi) = \sum_n \psi_n F(\phi_n). \tag{3.3}$$

Defining the vector $f = \sum_n [F(\phi_n)]^* \phi_n$ we see that

$$(f, \psi) = \sum_n F(\phi_n)\psi_n = F(\psi) \tag{3.4}$$

for all $\psi \in \mathcal{H}$. □

3.3.2 The Dirac notation

Vectors in $\mathcal{H}$ are called *ket* vectors, written as $|\phi\rangle$, vectors in the dual space $\mathcal{H}'$ (functionals) are called *bra* vectors, written as $\langle f|$. In this way we have

$$F(\phi) = (f, \phi) = \langle f|\phi\rangle, \tag{3.5}$$

where the last expression is a "bra-ket" – hence the name.

Remark 14. On some level one can think of this *Dirac notation* simply as a notation for scalar products, but it turns out to be useful far beyond this.

There is an *anti-linear correspondence* between bras and kets due to the nature of the scalar product:

$$\alpha^*\langle f_1| + \beta^*\langle f_2| \longleftrightarrow \alpha|f_1\rangle + \beta|f_2\rangle. \tag{3.6}$$

Recalling the definition of the adjoint operator $\hat{A}^\dagger$ of a linear operator $\hat{A}$ we see that the bra associated to the ket vector $\hat{A}|\phi\rangle$ is the vector $\langle\phi|\hat{A}^\dagger$. That is, we can interpret the expression $\hat{A}|\phi\rangle$ as "$|\hat{A}\phi\rangle$" and the expression $\langle\phi|\hat{A}^\dagger$ as "$\langle\hat{A}\phi|$". That is, the notation $\langle\phi|\hat{A}|\psi\rangle$ denotes two equal, but morally different, things:

(i) The inner product of the vector ϕ with the vector resulting from acting with $\hat{A}$ on the vector ψ, i.e. "$\langle\phi|\hat{A}\psi\rangle$";

(ii) The inner product of the vector resulting from acting with $\hat{A}^\dagger$ onto the vector ϕ with the vector ψ, i.e. "$\langle\hat{A}^\dagger\phi|\psi\rangle$".

The definition of the adjoint operator is expressed in Dirac notation as

$$\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}^\dagger|\phi\rangle^*. \tag{3.7}$$

3.3.3 Rules for finding the adjoint of an expression

To find the *adjoint* of an expression, one needs to

- complex conjugate scalars $c \leftrightarrow c^*$,
- replace kets with bras and vice versa $|\phi\rangle \leftrightarrow \langle\phi|$,
- replace operators with their adjoints $\hat{A} \leftrightarrow \hat{A}^\dagger$,
- reverse the order of factors.

As an example, consider the expression

$$c|\psi\rangle\langle\phi|A|\psi\rangle\langle\phi|.$$

This expression is an operator. The adjoint operator is given by

$$c^*|\phi\rangle\langle\psi|\hat{A}^\dagger|\phi\rangle\langle\psi|. \tag{3.8}$$

3.3.4 The outer product of two vectors

Besides the *inner product* of a bra $\langle\phi|$ and a ket $|\psi\rangle$, $\langle\phi|\psi\rangle$, which is a scalar, there is also an *outer product*, $|\psi\rangle\langle\phi|$, which is a linear operator.

Exercise 4. What is the adjoint operator of $|\psi\rangle\langle\phi|$?

Remark 15. An operator of the form $P = |\phi\rangle\langle\phi|$ is a *projection* operator, as the mapping $P : |\psi\rangle \mapsto \langle\phi|\psi\rangle|\phi\rangle$ is the projection of $|\psi\rangle$ onto $|\phi\rangle$. We will return to a more general notion of projection operators in the discussion of principle 4 in chapter 4.

3.3.5 Completeness and resolution of identity

An orthonormal set of vectors $\{\phi_n\}$ is complete if an arbitrary $|v\rangle \in V$ can be written as

$$|v\rangle = \sum_n v_n|\phi_n\rangle. \tag{3.9}$$

The coefficients v_n are given by

$$v_n = \langle\phi_n|v\rangle. \tag{3.10}$$

Reinserting this into the expansion (3.9) of $|v\rangle$ gives the identity

$$|v\rangle = \sum_n |\phi_n\rangle\langle\phi_n|v\rangle. \tag{3.11}$$

This, however, is identical to the statement

$$|v\rangle = \left(\sum_n |\phi_n\rangle\langle\phi_n|\right)|v\rangle. \tag{3.12}$$

Since this is valid for all vectors $v \in V$, that means that

$$\sum_n |\phi_n\rangle\langle\phi_n| = \mathbb{1}. \tag{3.13}$$

This property which is fulfilled by any complete orthonormal set, is called the *resolution of identity*.

3.4 More on Hermitian operators

Theorem 2. *All eigenvalues of a Hermitian operator are real.*

Proof. Consider the (non-zero) eigenvector $|u\rangle$ of a Hermitian operator $\hat{A} = \hat{A}^\dagger$, corresponding to the eigenvalue $a \neq 0$:

$$\hat{A}|u\rangle = a|u\rangle.$$

The adjoint equation reads

$$\langle u|\hat{A} = \langle u|a^*.$$

We have

$$\langle u|\hat{A}|u\rangle = \langle u|a|u\rangle = a\langle u|u\rangle.$$

On the other hand we have $\hat{A} = \hat{A}^\dagger$, and thus

$$\langle u|\hat{A}|u\rangle = \langle u|\hat{A}^\dagger|u\rangle = \langle u|a^*|u\rangle = a^*\langle u|u\rangle.$$

It follows that

$$a\langle u|u\rangle = a^*\langle u|u\rangle. \tag{3.14}$$

Since $\langle u|u\rangle \neq 0$ it follows that

$$a = a^* \in \mathbb{R}.$$

□

Theorem 3. *Eigenvectors of a Hermitian operator belonging to different eigenvalues are orthogonal.*

Exercise 5. Prove theorem 3.

Theorem 4. *The expectation values, $\langle v|\hat{A}|v\rangle$, of Hermitian operators are real. On the other hand, if every expectation value of an operator is real, the operator is Hermitian.*

We will prove theorem 4 in the lecture.

Finally, without proof, we state the following important

Theorem 5. *The eigenvectors of a Hermitian operator form a complete orthogonal basis set for the Hilbert space the operator acts upon.*

An operator A is called *anti-Hermitian* if $\hat{A}^\dagger = -\hat{A}$. If $\hat{A}$ is Hermitian then $\mathrm{i}\hat{A}$ is anti-Hermitian. Every operator can be uniquely decomposed into a Hermitian and an anti-Hermitian part, $\hat{A} = \hat{A}_h + \mathrm{i}\hat{A}_a$, with $\hat{A}_h^\dagger = \hat{A}_h$ and $\hat{A}_a^\dagger = \hat{A}_a$.

3.5 Vector/matrix representation

Consider a Hilbert space $\mathcal{H}$ with a countable orthonormal basis $\{|u_n\rangle\}$. The ket vectors $|\psi\rangle \in \mathcal{H}$ can be represented by (possibly infinite) column vectors of their coefficients in this basis:

$$|\psi\rangle = \sum_n a_n |u_n\rangle, \quad |\psi\rangle \simeq \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_j \\ \vdots \end{pmatrix} =: \vec{a}.$$

For the scalar product of two vectors $|\psi\rangle = \sum_n a_n|u_n\rangle$, and $|\phi\rangle = \sum_n b_n|u_n\rangle$ it follows from $\langle\phi| = \sum_n \langle u_n|b_n^*$ that

$$\begin{aligned} \langle\phi|\psi\rangle &= \sum_{m,n} \langle u_m|b_m^* a_n|u_n\rangle = \sum_{m,n} b_m^* a_n \langle u_m|u_n\rangle \\ &= \sum_{m,n} b_m^* a_n \delta_{mn} = \sum_n b_n^* a_n \\ &= (\vec{b}^*)^T \vec{a}\,. \end{aligned}$$

Similarly linear operators on the Hilbert space can be represented by matrices. Consider, e.g., the equation $\hat{M}|\psi\rangle = |\phi\rangle$, that is

$$\sum_n a_n \hat{M}|u_n\rangle = \sum_n b_n |u_n\rangle.$$

Operating on this with $\langle u_m|$ yields

$$\begin{aligned} \sum_n a_n \langle u_m|\hat{M}|u_n\rangle &= \sum_n b_n \langle u_m|u_n\rangle \\ \sum_n a_n \langle u_m|\hat{M}|u_n\rangle &= \sum_n b_n \delta_{mn} \\ \sum_n M_{mn} a_n &= b_m, \end{aligned}$$

which is valid for all m. Thus,

$$\hat{M}|\psi\rangle = |\phi\rangle \;\Leftrightarrow\; M\vec{a} = \vec{b},$$

with $\hat{M} = (M_{mn})$ and $M_{mn} = \langle u_m|\hat{M}|u_n\rangle$. In matrix representation Hermiticity means that $M_{mn} = M_{nm}^*$.

Chapter 4

The Principles of Quantum Mechanics

In this chapter we will learn about the five principles/postulates of quantum mechanics that allow an abstract formulation of the theory. The first two principles that we have already briefly encountered in the previous chapter, concern the description of the *states* of a quantum system and its *observables*, that is, its measurable quantities. The third and fourth principle concern the concept of *measurements* and the final one the *time-evolution* of quantum systems.

4.1 States and observables

A *state* encodes all the information one can have about a quantum system. *Observables* are measurable quantities (such as, for example, the position of a particle or its kinetic energy).

4.1.1 Principle 1 (States):

The state of a quantum system is described by a non-zero vector in a Hilbert space.

Remark 16. The space of square-integrable functions L^2 that we have dealt with in Schrödinger's theory of wave functions is a Hilbert space.

The superposition principle means that the sum of two states is a state, which also follows from the vector space properties. Two vectors describe the same state if they are complex multiples of each other. The equivalence class $\psi \to c\psi$ with $c \in \mathbb{C}$ defines *rays* through the "origin" of the Hilbert space. The space of these rays is called the *projective Hilbert space*.

Remark 17. Some caution is needed when dealing with the concept of the projective Hilbert space; while the states $\mathrm{e}^{\mathrm{i}\phi_1}\psi_1$ and $\mathrm{e}^{\mathrm{i}\phi_2}\psi_2$ are equivalent to the states ψ_1 and ψ_2 respectively, in general $\mathrm{e}^{\mathrm{i}\phi_1}\psi_1 + \mathrm{e}^{\mathrm{i}\phi_2}\psi_2$ describes a different state from $\psi_1 + \psi_2$.

4.1.2 Principle 2 (Observables):

A measurable quantity is described by a Hermitian operator on the Hilbert space. Both the measurable quantities and the operators representing them are often referred to as *observables*.

In particular the operators corresponding to *position* and *momentum*, $\hat{q}$ and $\hat{p}$, fulfil the commutation relation

$$[\hat{q}, \hat{p}] = \mathrm{i}\hbar\mathbb{1}, \tag{4.1}$$

where the *commutator* $[\cdot, \cdot]$ is defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}, \tag{4.2}$$

and $\mathbb{1}$ denotes the identity operator.

4.1.3 Important operators (in "position representation")

Here we consider the Hilbert space of square integrable functions L^2 as our space of states, where states are given by wave functions $\psi(x) \in L^2$, which are interpreted as in Schrödinger's wave mechanics. The *position operator* is then defined by its action on a quantum state $\psi(x) \in L^2$ as

$$\hat{q} : \psi(x) \mapsto x\psi(x). \tag{4.3}$$

With the definition of the position operator, also the *momentum operator* $\hat{p}$ is fixed, because $\hat{p}$ and $\hat{q}$ have to fulfil equation (4.1). The operator

$$\hat{p} : \psi(x) \mapsto -\mathrm{i}\hbar\frac{\partial}{\partial x}\psi(x) \tag{4.4}$$

fulfils this condition with (4.3).

Exercise 6. Verify that the operators (4.4) and (4.3) indeed fulfil the commutation relation (4.1), by applying their commutator onto a test function $\psi(x) \in L^2$.

Remark 18. One could in principle add an arbitrary function of $\hat{q}$ to $\hat{p}$ and still fulfil the commutation relation (4.1). The reason why it is fine to choose this function to be zero is a little subtle, and does not need to concern us here. If you are interested in the details, see e.g. Dirac's textbook.

The *kinetic energy* is described by $E_{kin} = \frac{p^2}{2m}$ in classical mechanics. This is analogous in quantum mechanics, and thus the operator of kinetic energy in the position representation is given by

$$\hat{E}_{kin} = \frac{\hat{p}^2}{2m} : \psi(x) \mapsto -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x). \tag{4.5}$$

4.2 Measurement and dynamics

We have learned in the first two principles that a quantum state is a vector and that measurable quantities are represented by operators. Now we will focus on the actual measurement and how the outcome of a measurement is related to the operator describing an observable and the state describing the system.

4.2.1 Principle 3 ("Born rule"):

The possible outcomes of a measurement of an observable A are given by the eigenvalues of the corresponding operator $\hat{A}$. The probability of obtaining an eigenvalue a_j as a measurement outcome is encoded in the state $|\psi\rangle$ the system is in immediately before the measurement. The probability is given by $p(a_j) = \langle\psi|\hat{P}_j|\psi\rangle$, where $\hat{P}_j$ is the *projection operator* onto the subspace associated to the eigenvalue a_j. For non-degenerate, discrete eigenvalues this reduces to the modulus square $|\psi_j|^2$ of the coefficient of the (normalised) wave function $|\psi\rangle$ in the basis of eigenstates $|\phi_j\rangle$ of $\hat{A}$:

$$|\psi\rangle = \sum_n \psi_n|\phi_n\rangle, \quad \text{with} \quad \hat{A}|\phi_n\rangle = a_n|\phi_n\rangle. \tag{4.6}$$

For continuous spectra we replace the sum with an integral: The probability to obtain a value in the interval I in a measurement is the integral over the modulus squared of the components of the wave function over this interval $P = \int_I |\psi(\alpha)|^2 \mathrm{d}\alpha$, where

$$|\psi\rangle = \int \mathrm{d}\alpha\, \psi(\alpha)|\phi(\alpha)\rangle, \quad \text{with} \quad \hat{A}|\phi(\alpha)\rangle = \alpha|\phi(\alpha)\rangle. \tag{4.7}$$

4.2.2 Projection operator

A linear operator $\hat{P}$ that satisfies $\hat{P}^2 = \hat{P}$ is called a *projection operator*. Examples are operators of the form $\hat{P} = \sum_{n=1}^{N} |\phi_n\rangle\langle\phi_n|$ where $\{|\phi_n\rangle\}$ is an orthonormal set of vectors. We can easily verify that indeed $\hat{P}^2 = \hat{P}$:

$$\begin{aligned}\hat{P}^2 &= \sum_n |\phi_n\rangle\langle\phi_n| \sum_m |\phi_m\rangle\langle\phi_m| = \sum_{n,m} |\phi_n\rangle\langle\phi_n|\phi_m\rangle\langle\phi_m| \\ &= \sum_{n,m} \delta_{mn}|\phi_n\rangle\langle\phi_n| = \sum_n |\phi_n\rangle\langle\phi_n| = \hat{P} \qquad \square\end{aligned} \tag{4.8}$$

Exercise 7. Verify that the eigenvalues λ of a projection operator are either one or zero.

Example 1 - Energies of a particle in a box of length L

The eigenvalues of the Hamiltonian are given by $E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}$. If the state is given by $|\psi\rangle = \sum_n \psi_n|\phi_n\rangle$ with $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$ we will obtain the value E_n in an energy measurement with probability $|\psi_n|^2$. If the system is in the "ground state" (that is, the state with the lowest energy, in this case E_1), for example, we will obtain the value $E_1 = \frac{\pi^2\hbar^2}{2mL^2}$ with probability one.

Example 2 - Position of a particle

The position operator in position representation is defined by its action on L^2 functions as

$$\hat{q} : \psi(x) \mapsto x\psi(x). \tag{4.9}$$

What are the eigenvalues of $\hat{q}$? The eigenvalue equation reads

$$\begin{aligned}\hat{q}\phi_\alpha(x) &= \alpha\phi_\alpha(x) \\ x\phi_\alpha(x) &= \alpha\phi_\alpha(x).\end{aligned} \tag{4.10}$$

There are no functions fulfilling this requirement. Nevertheless, there are distributions that do. The *function* $\phi_\alpha(x)$ would have to be zero everywhere except at $x = \alpha$. *Dirac's δ-distribution* does that. It is defined by the relation

$$\int_{-\infty}^{\infty} f(x)\delta(x) = f(0), \tag{4.11}$$

for functions $f(x)$ that are sufficiently well-behaved (smooth).

Remark 19. The notation in equation (4.11) is an example of the way in which this course differs from pure mathematic courses, as technically we deal with a *distribution* and not a *function*, i.e. the integral in (4.11) is not well defined in the sense of a Riemann or a Lebesgue integral. It is, however, a convenient way to express the essence of the definition of the δ-distribution which is used routinely in the physics literature and can be used without harm in most situations.

The δ-distribution can be expressed as the limit of various well-defined functions, for example as

$$\delta(x) = \lim_{n\to\infty} \frac{\sin(nx)}{\pi x}, \tag{4.12}$$

or as the Fourier transform of a plane wave as

$$\delta(x-\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{ik(x-\alpha)}dk. \tag{4.13}$$

Remark 20. In equation (4.12) we have yet another abuse of notation. Writing $\delta(x) = g(x)$ is a short hand form for the slightly more rigorous statement $\int_{-\infty}^{\infty} f(x)\delta(x)\mathrm{d}x = \int_{-\infty}^{\infty} f(x)g(x)\mathrm{d}x$.

The δ distribution can be viewed as an eigenfunction of $\hat{q}$ in the following sense: We have

$$\int_{-\infty}^{\infty} x\delta(x-\alpha)\mathrm{d}x = \int_{-\infty}^{\infty} \alpha\delta(x-\alpha)\mathrm{d}x, \tag{4.14}$$

In the sense of remark (20) we identify this as the eigenvalue equation (4.10). That is, the position operator has the continuous spectrum $\alpha \in \mathbb{R}$. The corresponding eigenvectors $|\phi_\alpha\rangle$ fulfil the generalised orthonormality condition

$$\langle\phi_\alpha|\phi_\beta\rangle = \delta(\alpha-\beta), \tag{4.15}$$

where the δ-function replaces the *Kronecker delta* $\delta_{\alpha\beta}$ in the familiar orthonormality condition for discrete systems.

The wave function in position representation $\psi(x)$ corresponds to the coefficients of the abstract vector $|\psi\rangle$ in the basis of the eigenfunctions of $\hat{q}$, that is $\psi(x) := \langle x|\psi\rangle$, and

$$|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)|x\rangle\mathrm{d}x. \tag{4.16}$$

Remark 21. In the position representation we calculate directly with $\psi(x)$ and operators acting upon it. When we write $\hat{A}\psi(x)$ this means $\langle x|\hat{A}|\psi\rangle$. In this way we sometimes write $\hat{q}\psi(x) = x\psi(x)$ and $\hat{p}\psi(x) = -\mathrm{i}\hbar\frac{\partial\psi(x)}{\partial x}$.

When measuring a particles position, according to principle 3, the probability of obtaining a value in the interval $I = [a,b]$ is given by $\int_a^b |\psi(x)|^2\mathrm{d}x$, as we have introduced it in section 2. That is, the interpretation of Schrödinger's wave function is a special case of principle 3.

Expectation values

The *mean value* of the outcome of many measurements of an observable A on *identically prepared systems* does not have to be an eigenvalue of the operator $\hat{A}$. It is given by the *expectation value* of $\hat{A}$ in the state $|\psi\rangle$:

$$\langle\hat{A}\rangle := \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{4.17}$$

Let us verify this for the case of discrete eigenvalues: Let us denote the eigenvalues of $\hat{A}$ by a_n and the corresponding eigenvectors by $|\phi_n\rangle$, that is, $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$. The eigenvectors belonging to different eigenvalues are automatically orthogonal, in addition we can choose them normalised. The mean value of many measurements is clearly given by

$$\langle\hat{A}\rangle = \sum_j a_j P(a_j), \tag{4.18}$$

where $P(a_j)$ denotes the probability that a measurement yields the outcome a_j. Let us now verify that this is indeed equal to the expression (4.17), using that according to principle 3 $P(a_j)$ is given by

$$\begin{aligned} P(a_j) &= |\langle\phi_j|\psi\rangle|^2 \\ &= \langle\psi|\phi_j\rangle\langle\phi_j|\psi\rangle, \end{aligned} \tag{4.19}$$

if the system is in the normalised state $|\psi\rangle$. Inserting expression (4.19) for the probability of the measurement outcome a_j into equation (4.18) yields

$$\begin{aligned} \langle\hat{A}\rangle &= \sum_j a_j\langle\psi|\phi_j\rangle\langle\phi_j|\psi\rangle \\ &= \sum_j \langle\psi|\hat{A}|\phi_j\rangle\langle\phi_j|\psi\rangle. \end{aligned} \tag{4.20}$$

Now we recognise the resolution of identity $\sum_j |\phi_j\rangle\langle\phi_j| = \mathbb{1}$ and thus we have

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle, \tag{4.21}$$

as expected.

4.2.3 Principle 4 ("state collapse postulate"):

Consider a system in the quantum state $|\psi\rangle$ on which a measurement of an observable A is performed. If the measurement yields the outcome a_j (an eigenvalue of $\hat{A}$), then the system is in the state $\frac{\hat{P}_j|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_j|\psi\rangle}}$ immediately after the measurement. Here $\hat{P}_j$ is the projection operator onto the subspace belonging to the eigenvalue a_j. For a non-degenerate eigenvalue with eigenvector $|\phi_j\rangle$ that is $\hat{P}_j = |\phi_j\rangle\langle\phi_j|$, which means the state afetr the measurement will be the state $\phi_j\rangle$.

Remark 22. In some ways principle 4 is a logical extension of principle 3: If we were to perform two measurements of the same observable instantly after one another, we would expect to obtain the same outcome in both with probability one.

Remark 23. Having the state collapse postulate in mind one often uses the formulation "to find a system in a state $|\phi\rangle$" in quantum mechanics. Such as for example in "... calculate the probability to find the system in the ground state.".

4.2.4 Principle 5 (dynamics):

The time-evolution of a quantum state $|\psi(t)\rangle$ is governed by the Schrödinger equation

$$\mathrm{i}\hbar|\dot{\psi}\rangle = \hat{H}|\psi\rangle, \tag{4.22}$$

where the operator $\hat{H}$ is the operator corresponding to the total energy of the system.

The solution of the Schrödinger equation can formally be written as

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle, \tag{4.23}$$

where $\hat{U}(t)$ is the *time-evolution operator*. For time independent Hamiltonians it is given by

$$\hat{U}(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\hat{H}t\right). \tag{4.24}$$

Since the Hamiltonian $\hat{H}$ is Hermitian, the time-evolution operator $\hat{U}(t)$ is *unitary*. That is,

$$\hat{U}^{-1} = \hat{U}^\dagger. \tag{4.25}$$

Unitarity means that the inner product is conserved, i.e.

$$\langle\psi_1(t)|\psi_2(t)\rangle = \langle\psi_1(0)|\psi_2(0)\rangle, \tag{4.26}$$

with $|\psi_j(t)\rangle = \hat{U}(t)|\psi_j(0)\rangle$.

Remark 24. Think of "rotations on Hilbert space".

4.2.5 The method of stationary states in Dirac notation (for arbitrary Hilbert spaces)

Given: Hamiltonian $\hat{H}$, initial wave function $|\psi(t=0)\rangle$
Wanted: Wave function at time t, $|\psi(t)\rangle$

Method:

1. Calculate the eigenvalues E_n and normalised eigenstates $|\phi_n\rangle$ of the Hamiltonian $\hat{H}$.

2. Expand the initial wave function in the basis of eigenstates $\{|\phi_n\rangle\}$:

$$|\psi(0)\rangle = \sum_n a_n |\phi_n\rangle.$$

   The coefficients are given by $a_n = \langle\phi_n|\psi(0)\rangle$.

3. The time evolution is then given by

$$|\psi(t)\rangle = \sum_n a_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar} |\phi_n\rangle.$$

Chapter 5

Commutators and uncertainties

The commutator of two observables restricts how accurately they can be measured simultaneously. In this chapter we shall discuss this in some detail. In this context we shall also learn about the famous *Heisenberg uncertainty principle*.

5.1 The uncertainty relations

The outcomes of many measurements of an observable A on identically prepared systems have a certain spread. The width of their distribution is characterised by the expectation value of the operator $(\hat{A} - \langle\hat{A}\rangle)^2$, that is, the *variance*

$$\begin{aligned}(\Delta A)^2 &:= \langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle \\ &= \langle\hat{A}^2 - 2\langle\hat{A}\rangle\hat{A} + \langle\hat{A}\rangle^2\rangle \\ &= \langle\hat{A}^2\rangle - 2\langle\hat{A}\rangle^2 + \langle\hat{A}\rangle^2 \\ &= \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2.\end{aligned} \tag{5.1}$$

In quantum mechanics we refer to the *standard deviation*

$$\Delta A := \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}, \tag{5.2}$$

as *uncertainty* of A.

It follows from the principles of quantum mechanics that in general the product of the uncertainties of two observables cannot be arbitrarily small. Instead, it has a lower bound that is related to the commutator of the observables via the following

Theorem 6. *Consider the two Hermitian operators $\hat{A}$ and $\hat{B}$, with $[\hat{A}, \hat{B}] \neq 0$. The product of the uncertainties of the observables A and B fulfils*

$$\Delta A \Delta B \geq \frac{1}{2}|\langle\hat{C}\rangle|, \tag{5.3}$$

where we have defined $\hat{C} := \mathrm{i}[\hat{A}, \hat{B}]$.

Remark 25. The factor of i in the definition of $\hat{C}$ has been introduced such that $\hat{C}$ is Hermitian.

Proof. Let us shift the operators $\hat{A}$ and $\hat{B}$ such that their mean values are zero, that is, let us consider the operators

$$\hat{a} := \hat{A} - \langle\hat{A}\rangle, \quad \text{and} \quad \hat{b} := \hat{B} - \langle\hat{B}\rangle. \tag{5.4}$$

These operators have the same commutator as $\hat{A}$ and $\hat{B}$. That is,

$$[\hat{a}, \hat{b}] = -\mathrm{i}\hat{C}. \tag{5.5}$$

Further we have

$$(\Delta A)^2 = \langle \hat{a}^2 \rangle \quad \text{and} \quad (\Delta B)^2 = \langle \hat{b}^2 \rangle. \tag{5.6}$$

Let us now consider the norm of the state $|\chi\rangle = (\hat{a} - \mathrm{i}\lambda\hat{b})|\psi\rangle$ for $\lambda \in \mathbb{R}$ and an arbitrary normalised state $|\psi\rangle$. We have

$$\begin{aligned} \langle\chi|\chi\rangle &= \langle\psi|(\hat{a} + \mathrm{i}\lambda\hat{b})(\hat{a} - \mathrm{i}\lambda\hat{b})|\psi\rangle \\ &= \langle(\hat{a} + \mathrm{i}\lambda\hat{b})(\hat{a} - \mathrm{i}\lambda\hat{b})\rangle \\ &= \langle\hat{a}^2 + \lambda^2\hat{b}^2 + \mathrm{i}\lambda(\hat{b}\hat{a} - \hat{a}\hat{b})\rangle \\ &= \langle\hat{a}^2\rangle + \lambda^2\langle\hat{b}^2\rangle + \mathrm{i}\lambda\langle\mathrm{i}\hat{C}\rangle \\ &= \langle\hat{a}^2\rangle + \lambda^2\langle\hat{b}^2\rangle - \lambda\langle\hat{C}\rangle \end{aligned} \tag{5.7}$$

On the other hand, we know that $\langle\chi|\chi\rangle \geq 0$. That is, we have

$$\langle\hat{a}^2\rangle + \lambda^2\langle\hat{b}^2\rangle - \lambda\langle\hat{C}\rangle \geq 0, \tag{5.8}$$

for all values of $\lambda \in \mathbb{R}$. Let us now evaluate for which value of λ the norm $\langle\chi|\chi\rangle$ becomes minimal. The location of the minimum is found where

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\chi|\chi\rangle = 2\langle\hat{b}^2\rangle\lambda - \langle\hat{C}\rangle = 0, \tag{5.9}$$

that is

$$\lambda = \frac{\langle\hat{C}\rangle}{2\langle\hat{b}^2\rangle}. \tag{5.10}$$

(Note that this cannot be a maximum, since the function in (5.8) is a quadratic and goes to infinity at $\lambda \to \pm\infty$, and thus has a single minimum at a finite value of λ.) Reinserting this into the condition $\langle\chi|\chi\rangle_{min} \geq 0$ yields

$$\begin{aligned} \langle\chi|\chi\rangle_{min} &= \langle\hat{a}^2\rangle + \frac{\langle\hat{C}\rangle^2}{4\langle\hat{b}^2\rangle^2}\langle\hat{b}^2\rangle - \frac{\langle\hat{C}\rangle^2}{2\langle\hat{b}^2\rangle} \\ &= \langle\hat{a}^2\rangle - \frac{\langle\hat{C}\rangle^2}{4\langle\hat{b}^2\rangle} \geq 0, \end{aligned} \tag{5.11}$$

that is

$$(\Delta A)^2 \geq \frac{\langle\hat{C}\rangle^2}{4(\Delta B)^2}, \tag{5.12}$$

or

$$(\Delta A)(\Delta B) \geq \frac{|\langle\hat{C}\rangle|}{2}. \qquad \square$$

The special case of the uncertainty relation in theorem 6 for the case of position and momentum operator is known as *Heisenberg's uncertainty relation*:

$$\Delta p \Delta x \geq \frac{\hbar}{2}. \tag{5.13}$$

This means that one can never know the exact values of both position and momentum simultaneously.

Remark 26. The uncertainty relation gives a *lower* bound. There are states for which the actual uncertainty product for position and momentum for example is much larger than $\hbar$.

Remark 27. The lower bound of Heisenberg's uncertainty relation is only reached for states of the form

$$\psi(x) \propto \exp\left(-\alpha(x-q)^2 + \frac{\mathrm{i}}{\hbar}p(x-q) + \frac{\mathrm{i}\gamma}{\hbar}\right), \tag{5.14}$$

with $\alpha, q, p \in \mathbb{R}$ and $\gamma \in \mathbb{C}$. These states are called *Gaussian wave packets*. We will discuss their properties in more detail in section 8.2

5.2 Commuting observables

If two observables commute they are in principle simultaneously measurable, according to the uncertainty relation. This is an aspect of the following

Theorem 7. *Commuting operators* $[\hat{A}, \hat{B}] = 0$ *posess a set of common eigenvectors.*

Exercise 8. Prove theorem 7 for the case of non-degenerate eigenvalues. I.e., consider an eigenvector $|\phi\rangle$ of $\hat{A}$ belonging to the eigenvalue $a \in \mathbb{R}$, that is $\hat{A}|\phi\rangle = a|\phi\rangle$, and show that $|\phi\rangle$ is also an eigenvector of $\hat{B}$ if $[\hat{A}, \hat{B}] = 0$.

Remark 28. If the eigenvalue a is degenerate (that is, it is a multiple eigenvalue), we can find a linear superposition of the corresponding eigenvectors that is an eigenvector of $\hat{B}$.

Remember that if we measure an observable A, after the measurement the quantum state of the system will have collapsed into the eigenspace of the operator $\hat{A}$ belonging to the eigenvalue that had been measured. Now if we measure an observable $\hat{B}$ directly afterwards that measurement will project the state of the system onto an eigenspace of $\hat{B}$. That means, only if $\hat{A}$ and $\hat{B}$ share eigenvectors, will the measurements be *compatible*, i.e. the measurement of one of the observables won't *delete* the information about the value of the first observable. Commuting operators have a set of shared eigenvectors, so they can be measured *simultaneously* or without the order of measurement changing the probabilities of the outcomes.

5.2.1 Complete set of commuting observables

Consider a set of Hermitian operators $\hat{A}_{1,2,\ldots}$ on a Hilbert space $\mathcal{H}$. Their simultaneous eigenvectors form a basis of $\mathcal{H}$. They are called a *complete set of commuting observables* if their eigenvalues uniquely define each basis vector (up to a phase).

If an operator $\hat{A}$ has only non-degenerate eigenvalues it is already a complete set of commuting observables on its own. If one or more eigenvalues is/are degenerate one can find a commuting operator to lift, or partially lift the degeneracy, and add further such operators to lift the degeneracies still present in this pair, and so on, until one arrives at a complete set of observables.

Let us consider a simple example for a three dimensional Hilbert space first. The operator

$$\hat{A} = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}, \tag{5.15}$$

with $a_j \in \mathbb{R}$, is clearly Hermitian, and indeed its eigenvectors are the standard vectors e_j, which are a basis of $\mathbb{C}^3$, and could be uniquely labeled here by the eigenvalues of $\hat{A}$, i.e. e_j is the eigenvector of $\hat{A}$ belonging to the eigenvalue a_j. If, however, some of the eigenvalues were degenerate, for example $a_2 = a_1$, or, for clarity, let's consider another operator $\hat{B}$

$$\hat{B} = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & b_2 \end{pmatrix}, \tag{5.16}$$

the situation would be different. We can still find a set of eigenvectors that are an orthogonal basis for $\mathbb{C}^3$, for example still the standard vectors e_j, or alternatively the vector e_3 and two orthogonal linear combinations of e_1 and e_2. However, these basis vectors cannot be uniquely identified by the corresponding eigenvalues of $\hat{B}$, as both e_1 and e_2 belong to the same eigenvalue b_1. If we in addition consider the operator

$$\hat{C} = \begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_2 \end{pmatrix}, \tag{5.17}$$

then the eigenvalues of $\hat{C}$ alone would also not be sufficient to identify the basis vectors, but those of $\hat{B}$ and $\hat{C}$ together would be. In the case here we could label the basis as

$$(b_1, c_1) \leftrightarrow e_1, \quad (b_1, c_2) \leftrightarrow e_2, \quad (b_2, c_2) \leftrightarrow e_3. \tag{5.18}$$

In infinite-dimensional Hilbert spaces it can be rather useful to have an abstract basis labelled according to the eigenvalues of an important observable (for example the energy), and if this has degenerate eigenvalues, it is useful to identify further commuting observables to uniquely label a basis of the Hilbert space. We will come back to this in more concrete physical examples later in the course.

5.2.2 Parity

Consider the Hermitian operator $\hat{S}$ on the Hilbert space of square integrable functions that interchanges x with $-x$, that is, $\hat{S}\psi(x) = \psi(-x)$. This is called the *parity* operator. The eigenfunctions of the parity operator fulfil

$$\hat{S}\phi_s(x) = \phi_s(-x) = s\phi_s(x). \tag{5.19}$$

On the other hand $\hat{S}^2 = \mathbb{1}$, that is,

$$\hat{S}(\hat{S}\phi_s(x)) = s^2\phi_s(x) = \phi_s(x). \tag{5.20}$$

Thus, we have $s^2 = 1$, and therefore the only two eigenvalues of the parity operator are

$$s = \pm 1. \tag{5.21}$$

The corresponding eigenfunctions are all even or odd functions $\phi_\pm(-x) = \pm\phi(x)$. (The parity operator is highly degenerate).

A Hamiltonian of the form $\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$ commutes with $\hat{S}$ if $V(x) = V(-x)$. This can be seen as follows. Consider the action of the commutator on a test function $\psi(x) \in L^2$:

$$\begin{aligned}[\hat{H}, \hat{S}]\psi(x) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(-x) + V(x)\psi(-x) \\ &\quad + \frac{\hbar^2}{2m}\frac{\partial^2}{\partial(-x)^2}\psi(-x) - V(-x)\psi(-x) \\ &= V(x)\psi(-x) - V(-x)\psi(-x),\end{aligned} \tag{5.22}$$

which is zero for arbitrary wave functions $\psi(x)$ if and only if $V(-x) = V(x)$.

According to theorem 7, if the eigenvalues of $\hat{H}$ are non-degenerate, the corresponding eigenfunctions are either odd or even. We shall discuss in the next section that for one-dimensional systems with potentials of the form $V(x \to \pm\infty) \to \infty$ (*binding potentials*) the eigenvalues of the Hamiltonian are always non-degenerate. Thus, for one-dimensional symmetric binding potentials, the eigenfunctions are either even or odd. We have already seen an example when we discussed the particle in a square well in Chapter 2, there the potential was symmetric with respect to $x = \frac{L}{2}$ and the eigenfunctions were even and odd with respect to this point.

Part III

Spectral properties of one-dimensional quantum systems

In this part we will analyse the eigenvalues of the Hamiltonian for a number of important one-dimensional quantum systems. That is, we solve the time-independent Schrödinger equation. Again, we consider systems of the form $\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$. Physically these Hamiltonians describe a particle (in one-dimensional space) of mass m under the influence of an external potential $V(x)$. This could be, for example, an electron in an atom, where the potential is given by the interaction with the nucleus and the other electrons (this is complicated and obviously three-dimensional and we will not discuss details in this chapter). The particle in a box is a good approximation for electrons in so-called *quantum dots*, which are fabricated semiconductor structures in which electrons are confined. Atoms in *optical potentials* (where quantum optical effects are used to confine atoms with laser setups) are another experimental system described by Hamiltonians of the form considered here. Experimentally these systems can often be made *quasi one-dimensional*. Similarly atoms can be trapped magnetically, in approximately harmonic traps described by quadratic potentials. There are countless other examples, even outside quantum physics: The Helmholtz equation for microwaves in flat (quasi two-dimensional) resonators is similar to the Schrödinger equation where the potential is related to the structure of the cavity. Similar analogies exist in optical setups.

Chapter 6

Piecewise constant potentials

6.1 Prelude - Properties of eigenfunctions for general one-dimensional potentials

There are only very few examples of Hamiltonian operators the spectrum (here: set of eigenvalues) of which can be deduced analytically. In general, approximative or numerical methods have to be applied. Yet, a few general remarks can be made about the solutions of time-independent Schrödinger equations of the form

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\phi_E(x) = E\phi_E(x)\,: \tag{6.1}$$

- In general there are two types of solutions of (6.1) that we can interpret in a meaningful way in quantum mechanics: The bound states, for which $\phi(x \to \pm\infty) \to 0$, and another type of states denoted as *continuum states*. We have encountered an example of continuum states in the plane waves, the eigenfunctions for the momentum operator $\mathrm{e}^{\pm \mathrm{i}kx}$. The reason that these states are referred to as continuum states to distinguish them from the bound states, is that the energies of bound states are always discrete, while the other solutions are associated to a continuum of energies.

- For potentials for which $V(x \to \pm\infty) \to \infty$ there are only bound states.

- Eigenfunctions can be chosen real valued.

  Proof. Assume we have a complex solution $\phi(x)$ of the Schrödinger equation. Then also its complex conjugate function $\phi^*(x)$ is a solution. Thus, the linear combination $\phi(x) + \phi^*(x)$ is a real valued solution of the Schrödinger equation. □

- There are no bound states with eigenenergies smaller than the minimum of the potential.

  Proof. We have

  $$\frac{\partial^2\phi_E(x)}{\partial x^2} = \frac{2m}{\hbar^2}(V(x) - E)\phi_E(x). \tag{6.2}$$

  Now if $E < V_{min}(x)$ then the the wave function $\phi_E(x)$ and its second derivative would have the same sign for all x. That is if the wave function would have to curve away from the x-axis everywhere. Such a function however, can not be normalisable. □

- The expectation value of a Hamiltonian is bounded from below by the lowest energy eigenvalue.

  We will prove this in the end of the week lecture.

- The bound state eigenvalues are non-degenerate (that is, there is only one linearly independent solution per eigenvalue).

Proof. Let us first rewrite the Schrödinger equation (6.1) in the form

$$\phi''(x) + k^2(x)\phi(x) = 0, \tag{6.3}$$

where we have introduced the function $k = \sqrt{\frac{2m}{\hbar^2}(E - V(x))}$. Suppose now that $\phi_1(x)$ and $\phi_2(x)$ are two bound state eigenfunctions corresponding to the same energy E. We shall show now that this implies that $\phi_1(x)$ and $\phi_2(x)$ are linearly dependent, i.e. they differ at most by a multiplicative factor.

We have

$$\phi_j''(x)/\phi_j(x) = -k(x)^2, \tag{6.4}$$

for $j = 1, 2$. That is,

$$\phi_1''(x)/\phi_1(x) = \phi_2''(x)/\phi_2(x), \tag{6.5}$$

or

$$\phi_1''(x)\phi_2(x) - \phi_2''(x)\phi_1(x) = 0. \tag{6.6}$$

We realise that the left hand side is equal to the derivative of $\phi_1'(x)\phi_2(x) - \phi_2'(x)\phi_1(x)$, that is we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\phi_1'(x)\phi_2(x) - \phi_2'(x)\phi_1(x)\right) = 0. \tag{6.7}$$

Integrating this yields

$$\phi_1'(x)\phi_2(x) - \phi_2'(x)\phi_1(x) = c, \qquad c \in \mathbb{C}. \tag{6.8}$$

Since we consider bound states we have that $\phi_j(x \to \infty) \to 0$ while the derivative stays bounded, and thus we deduce that $c = 0$. We now integrate the expression

$$\phi_1'(x)/\phi_1(x) = \phi_2'(x)/\phi_2(x), \tag{6.9}$$

to find

$$\ln \phi_1(x) = \ln \phi_2(x) + const. \tag{6.10}$$

that is

$$\phi_1(x) \propto \phi_2(x). \tag{6.11}$$

□

- Combining theorem 7, the fact that for a symmetric potential the Hamiltonian commutes with the parity operator, and the non-degeneracy of the bound state eigenvalues, we conclude that the bound states of one-dimensional systems of the form (6.1) with symmetric potentials are either even or odd.

- The ground state (the state corresponding to the lowest energy eigenvalue) does not have zeros (apart from the asymptotic ones at $x \to \pm\infty$).

- The n-th excited state (the state corresponding to the $(n+1)$th eigenvalue) has n zeros inside the potential.

  For a hand-wavy explanation of this see for example Barton Zwiebach's video lecture on the topic.

Remark 29. Unfortunately most of these statements do not carry over to higher dimensional systems. This concerns in particular the non-degeneracy of eigenvalues and hence the symmetry of eigenfunctions in symmetric potentials.

6.2 Piecewise constant potentials and boundary conditions

For the remainder of this chapter we are specifically looking for the solutions E and $\phi_E(x)$ of the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi_E(x) + V(x)\phi_E(x) = E\phi_E(x), \tag{6.12}$$

where the potential $V(x)$ is piecewise constant. Rewriting this in the form

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi_E(x) = (E - V(x))\phi_E(x), \tag{6.13}$$

emphasises that the second derivative of an eigenfunction is essentially proportional to the value of the potential energy. Thus, at values of x at which the potential makes a finite jump, so does the second derivative of the wave function, which means that the first derivative is continuous, while the wave function itself is continuous and continuously differentiable. If the potential makes an infinite jump the first derivative of the wave function is discontinuous, but the wave function itself is still continuous.

To obtain the eigenfunctions $\phi_E(x)$ we can consider the different regions of constant potential separately and patch the solutions together. Thus, we start out by determining the solutions of the time-independent Schrödinger equation in constant potentials.

6.3 The free particle

The free particle is a particle moving without the influence of an external potential, that is, $V(x) = 0$. However, in quantum (as well as in classical) physics, the zero point of the energy is arbitrary. Thus, the free particle is equivalent to a particle moving under the influence of a constant potential $V(x) = V_0$, with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V_0, \tag{6.14}$$

with $V_0 \in \mathbb{R}$. The eigenvalue equation for the Hamiltonian (6.14), given by

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi_E(x) = (E - V_0)\phi_E(x), \tag{6.15}$$

is formally solved by functions of the form

$$\phi_E(x) = A\mathrm{e}^{\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx}, \tag{6.16}$$

with the *wave number*

$$k = \frac{\sqrt{2m(E - V_0)}}{\hbar}, \tag{6.17}$$

for arbitrary values of $E \in \mathbb{R}$. However, the functions (6.16) diverge at $x \to \pm\infty$ for negative values of $E - V_0$. For the free particle on the whole real axis, we cannot interpret them in any meaningful way, and thus deduce that the eigenvalues of $\hat{H}$ are given by all real numbers $E \geq V_0$.

Even the eigenfunctions corresponding to these energies, however, are not square integrable, and can thus not be interpreted in the usual way as a probability amplitude to find a particle somewhere at a given time. They are however, a basis for all physically more meaningful wave functions in L^2.

Remark 30. We can interpret a wave function of the form

$$\psi_\pm(x) \propto \mathrm{e}^{\pm\mathrm{i}xk}, \tag{6.18}$$

a so-called *plane wave*, as a constant probability beam travelling from minus to plus infinity or plus to minus infinity, respectively. These wave functions are the extreme cases of Heisenberg's uncertainty principle, with accurately defined momentum, and totally unspecified position (describing a "completely *delocalised* particle"). Their probability flux is given by

$$j = \frac{\hbar}{2m\mathrm{i}}\left(\psi^*\psi' - \psi'^*\psi\right) = \pm\frac{\hbar k}{m}. \tag{6.19}$$

They are eigenfunctions of the momentum operator with eigenvalues $p = \hbar k$. Thus, they can be viewed as a stream of particles moving with constant velocity $\frac{p}{m}$.

Remark 31. The plane waves (6.18) fulfil the generalised orthonormality condition

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\mathrm{e}^{\mathrm{i}k'x}\mathrm{d}x = \delta(k-k'). \tag{6.20}$$

Remark 32. If we have a piecewise constant potential then in the separate regions that do not extend to infinity on both sides also the solutions (6.16) with energies $E < V_0$ are meaningful. They are more commonly expressed in the form

$$\phi_E(x) = C\mathrm{e}^{\kappa x} + D\mathrm{e}^{-\kappa x}, \tag{6.21}$$

with

$$\kappa = \frac{\sqrt{2m(V_0 - E)}}{\hbar}. \tag{6.22}$$

For regions that extend to $x_+\infty$ we need $C = 0$ and for regions that extend to $x \to -\infty$ it holds $D = 0$.

6.4 The infinite square well potential ("Particle in a box")

Let us now finally derive the eigenfunctions of a particle in an *infinite square well* (or "box") potential

$$V(x) = \begin{cases} 0, & 0 \le x \le L \\ \infty, & \text{otherwise,} \end{cases} \tag{6.23}$$

that we have already used numerous times. This system is an extreme case of a piecewise constant potential, where the potential is zero in a region $x \in [0, L]$ (the interior of the *box*) and infinite outside this region. Thus, outside the box the wave function has to be identical to zero. Inside the box, that is for $0 \le x \le L$, we have the solutions

$$\phi(x) = A\mathrm{e}^{\mathrm{i}kx} + B\mathrm{e}^{-\mathrm{i}kx}, \tag{6.24}$$

with $k = \frac{\sqrt{2mE}}{\hbar}$. We know, however, that the wave function is continuous, and thus has to fulfil the boundary conditions

$$\phi(0) = 0 = \phi(L). \tag{6.25}$$

This yields the conditions

$$A + B = 0 \tag{6.26}$$

$$A\mathrm{e}^{\mathrm{i}kL} + B\mathrm{e}^{-\mathrm{i}kL} = 0 \tag{6.27}$$

for A and B. Inserting $B = -A$ into (6.27) yields

$$2\mathrm{i}A\sin(kL) = 0. \tag{6.28}$$

The choice $A = 0$ leads to a wave function that vanishes everywhere. The only physical meaningful wave functions exist for $\sin(kL) = 0$, which means

$$\frac{\sqrt{2mE}}{\hbar}L = n\pi, \tag{6.29}$$

for integer values of n, which yields the quantised energies

$$E_n = \frac{\pi^2\hbar^2}{2mL^2}\, n^2. \tag{6.30}$$

The corresponding eigenfunctions are given by

$$\phi_n(x) = 2\mathrm{i}A \sin\left(\frac{\sqrt{2mE_n}}{\hbar}\, x\right) = 2\mathrm{i}A \sin\left(\frac{\pi n x}{L}\right). \tag{6.31}$$

We can now choose A such that the $\phi_n(x)$ are normalised and real valued. The norm of $\phi_n(x)$ in dependence on A is given by

$$\begin{aligned} \int_0^L |\phi_n(x)|^2 &= 4|A|^2 \int_0^L \sin^2\left(\frac{\pi n x}{L}\right) \mathrm{d}x \\ &= 2|A|^2 \int_0^L \left(1 - \cos\left(\frac{2\pi n x}{L}\right)\right) \mathrm{d}x \\ &= 2|A|^2 L. \end{aligned} \tag{6.32}$$

Thus, the choice $A = \frac{-\mathrm{i}}{\sqrt{2L}}$ yields the normalised and real-valued wave functions

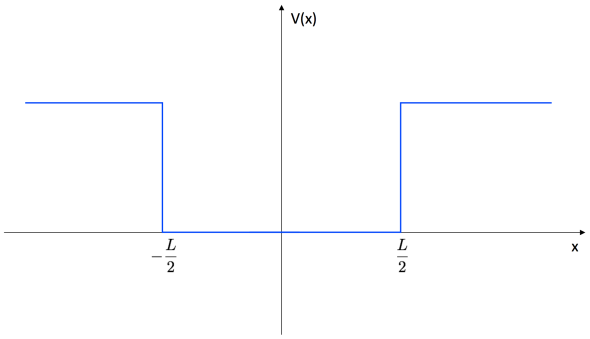
$$\phi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), & 0 \le x \le L \\ 0, & \text{otherwise,} \end{cases} \tag{6.33}$$

as given earlier without derivation in equation (2.30).

6.5 The finite square well potential ("Particle in a finite box")

The *finite square well potential* is similar to the previously considered infinite one, with the difference that the potential outside the central region is non-zero but finite. It is convenient to consider a symmetric version

$$V(x) = \begin{cases} 0, & |x| \le \frac{L}{2} \\ V_0, & \text{otherwise.} \end{cases} \tag{6.34}$$



The solutions of the time-independent Schrödinger equation in the separate regions are of the form $a\mathrm{e}^{\mathrm{i}kx} + b\mathrm{e}^{-\mathrm{i}kx}$ with $k = \sqrt{2m(E - V_j)}/\hbar$, where V_j is the value of the constant potential in region j. For values $E > V_j$, $k \in \mathbb{R}$ this can be equivalently written as $\tilde{a}\cos(kx) + \tilde{b}\sin(kx)$, while for $E < V_j$ we have $\kappa = \sqrt{2m(V_j - E)}/\hbar \in \mathbb{R}$, and the solutions take the form $a\mathrm{e}^{-\kappa x} + b\mathrm{e}^{\kappa x}$. That is we make the ansatz

$$\phi_E(x) = \begin{cases} C_1\mathrm{e}^{\kappa x} + C_2\mathrm{e}^{-\kappa x}, & x \le -L/2 \\ A\cos(kx) + B\sin(kx), & -L/2 < x \le L/2 \\ D_1\mathrm{e}^{\kappa x} + D_2\mathrm{e}^{-\kappa x}, & x > L/2, \end{cases} \tag{6.35}$$

with $k = \sqrt{2mE}/\hbar$ and $\kappa = \sqrt{2m(V_0 - E)}/\hbar$ for the solution of the time-independent Schrödinger equation.

Here we are looking for *bound states*, that is $\phi(x \to \pm\infty) \to 0$. From the general solution (6.35) we deduce that this is only possible for $E < V_0$, and that further we need $C_2 = 0 = D_1$. Thus, the solutions are of the form

$$\phi_E(x) = \begin{cases} C\mathrm{e}^{\kappa x}, & x \le -L/2 \\ A\cos(kx) + B\sin(kx), & -L/2 < x \le L/2 \\ D\mathrm{e}^{-\kappa x}, & x > L/2. \end{cases} \tag{6.36}$$

The wave function $\phi_E(x)$ and its first derivative with respect to x have to be continuous everywhere. This is automatically fulfilled in the separate regions, but it imposes four boundary conditions between the different regions.

Exercise 9. Find the four boundary conditions the eigenfunctions need to fulfil. Use the fact that the symmetry of the potential means that the eigenfunctions are either even or odd eigenfunction, and verify that even eigenfunctions have to fulfil the conditions

$$C\mathrm{e}^{-\kappa L/2} = A\cos(kL/2) \tag{6.37}$$
$$\kappa C\mathrm{e}^{-\kappa L/2} = kA\sin(kL/2), \tag{6.38}$$

and odd eigenfunctions have to fulfil the conditions

$$C\mathrm{e}^{-\kappa L/2} = -B\sin(kL/2) \tag{6.39}$$
$$\kappa C\mathrm{e}^{-\kappa L/2} = kB\cos(kL/2). \tag{6.40}$$

These conditions can only be fulfilled simultaneously for specific values of the energy E, leading to the quantisation conditions

$$\kappa = k\tan(kL/2) \tag{6.41}$$

and

$$\kappa = -k\cot(kL/2), \tag{6.42}$$

for even and odd eigenfunctions, respectively. We can rewrite these in terms of the energy using the identities

$$\begin{aligned} \kappa/k &= \sqrt{\frac{2mV_0}{\hbar^2k^2} - 1} \\ k &= \sqrt{2mE}/\hbar, \end{aligned} \tag{6.43}$$

to find

$$\tan\left(\frac{\sqrt{2m\overline{E}}L}{2\hbar}\right) = \sqrt{\frac{V_0}{E} - 1}\,, \tag{6.44}$$

and

$$-\cot\left(\frac{\sqrt{2m\overline{E}}L}{2\hbar}\right) = \sqrt{\frac{V_0}{E} - 1}\,, \tag{6.45}$$

for even and odd eigenfunctions, respectively.

These conditions have to be solved numerically or graphically to find the quantised energies. Let us introduce the variable $z := \frac{Lk}{2}$ and the constant $c := \sqrt{\frac{mV_0L^2}{2\hbar^2}}$. Then the quantisation conditions become

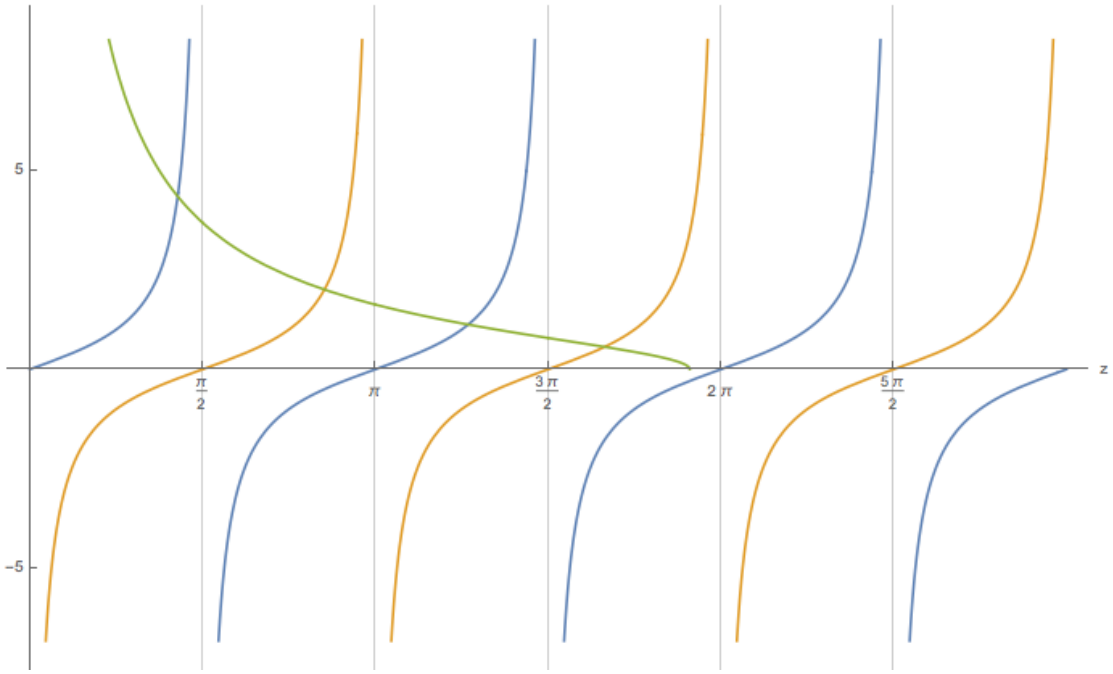
$$\tan(z) = \sqrt{\frac{c^2}{z^2} - 1}, \quad \text{and} \quad -\cot(z) = \sqrt{\frac{c^2}{z^2} - 1}, \tag{6.46}$$

for even and odd eigenfunctions, respectively. Let us now attempt a *graphical* solution:

The blue curves depict $\tan(z)$ (even solutions), the yellow ones $-\cot(z)$ (odd solutions), and the green curve is given by $\sqrt{\frac{c^2}{z^2} - 1}$, (with the arbitrarily chosen value $c = 6$). The values of z where the green curve intersects the $\tan(z)$, and $-\cot(z)$ curves can be converted into the allowed quantised energy values for the even and odd functions respectively.

We make a few observations:

- Even and odd eigenfunctions interchange, where the ground state is even.
- There is always at least one state.



- There is one solution in each interval $[(n-1)\frac{\pi}{2}, n\frac{\pi}{2}]$ as long as $(n-1)\frac{\pi}{2} \leq c$. That is, there is a finite number of bound states

$$N = \left[\frac{2c}{\pi}\right]_< + 1 = \left[\frac{\sqrt{2mV_0}L}{\pi\hbar}\right]_< + 1, \tag{6.47}$$

  where $[X]_<$ denotes the largest integer smaller than X.

- When V_0 is finite, the eigenfunctions do not vanish for $x < -L/2$ and $x > L/2$ (unlike the infinite square well). These are regions that are classically forbidden, yet quantum mechanically, there is a finite probability of finding a particle in a region where $0 < E < V_0 < \infty$. This demonstrates the striking non-classical phenomenon of *barrier penetration*. It is due to the continuity of $\phi(x)$ and $\frac{\mathrm{d}\phi}{\mathrm{d}x}$ at the boundary of a finite potential, a condition derived from the Schrödinger equation (contrast this boundary condition with the infinite square well case). One might think that this effect is purely a mathematical curiosity, however it is very much a real physical phenomenon which we shall explore further at a later point of the course.

- In the limit $V_0 \to \infty$ the quantisation conditions reduce to

$$\tan\left(\frac{L}{2}\frac{\sqrt{2mE}}{\hbar}\right) = \infty,$$

  and

$$-\cot\left(\frac{L}{2}\frac{\sqrt{2mE}}{\hbar}\right) = \infty.$$

  These are equivalent to the conditions $\sin\left(\frac{L}{2}\frac{\sqrt{2mE}}{\hbar}\right) = 0$ and $\cos\left(\frac{L}{2}\frac{\sqrt{2mE}}{\hbar}\right) = 0$, which together are indeed equivalent to

$$\sin\left(L\frac{\sqrt{2mE}}{\hbar}\right) = 0,$$

  that is, the quantisation condition for the infinite square well derived in section 6.4.

Remark 33. In the finite well we only found the (bound) eigenstates for $E < V_0$. The full eigenspace for the Hamiltonian would also have to include eigenstates for $E > V_0$, where the spectrum is continuous and the eigenstates resemble the free particle case (they are not-normalisable), but with different wave numbers inside and outside the well region.

Chapter 7

The harmonic oscillator

Let us now consider the quantum harmonic oscillator, the classical version of which we have studied in 1.2. Similarly to the classical case, the Hamiltonian is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{q}^2. \tag{7.1}$$

We want to solve the eigenvalue equation $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$. First of all we want to deduce the eigenvalues.

Remark 34. One way of solving this problem would be to express the eigenvalue equation in the position representation as $-\frac{\hbar^2}{2m}\frac{\partial^2\phi_E(x)}{\partial x^2} + \frac{1}{2}m\omega^2x^2\phi_E(x) = E\phi_E(x)$, and determine for which values of E this has bound state solutions that fulfil $\phi_E(x \to \pm\infty) \to 0$. Here, however, we use another method that does not require us to use any specific basis/representation at all, and instead only uses algebraic arguments.

7.1 Algebraic solution of the eigenvalue equation

Let us first introduce the dimensionless operators

$$\hat{Q} = \sqrt{\frac{m\omega}{\hbar}}\,\hat{q}, \quad \text{and} \quad \hat{\mathrm{P}} = \sqrt{\frac{1}{\mathrm{m}\omega\hbar}}\,\hat{\mathrm{p}}. \tag{7.2}$$

They fulfil the commutation relation

$$[\hat{Q}, \hat{P}] = \frac{1}{\hbar}[\hat{q}, \hat{p}] = \mathrm{i}\hat{I}. \tag{7.3}$$

In terms of these rescaled operators the Hamiltonian takes the form

$$\hat{H} = \frac{\hbar\omega}{2}\left(\hat{P}^2 + \hat{Q}^2\right). \tag{7.4}$$

Now we introduce the operator

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\hat{Q} + \mathrm{i}\hat{P}\right), \tag{7.5}$$

and its adjoint operator

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\hat{Q} - \mathrm{i}\hat{P}\right). \tag{7.6}$$

These operators fulfil the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = \hat{I}. \tag{7.7}$$

Exercise 10. Verify the commutation relation (7.7) using the commutator(7.3), and rewrite the Hamiltonian in terms of $\hat{a}$ and $\hat{a}^\dagger$.

In terms of $\hat{a}$ and $\hat{a}^\dagger$ the Hamiltonian becomes

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}). \tag{7.8}$$

That is, we need to analyse the spectrum of the Hermitian operator $\hat{N} := \hat{a}^\dagger\hat{a}$.

Theorem 8. *$\hat{N}$ has a discrete spectrum consisting of the non-negative integers.*

To prove this theorem we will need the commutators $[\hat{N}, \hat{a}]$ and $[\hat{N}, \hat{a}^\dagger]$, which follow from $[\hat{a}, \hat{a}^\dagger] = 1$, as

$$\begin{aligned}[\hat{N}, \hat{a}] &= [\hat{a}^\dagger\hat{a}, \hat{a}] \\ &= [\hat{a}^\dagger, \hat{a}]\hat{a} \\ &= -[\hat{a}, \hat{a}^\dagger]\hat{a} \\ &= -\hat{a}\end{aligned} \tag{7.9}$$

and similarly,

$$\begin{aligned}[\hat{N}, \hat{a}^\dagger] &= \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] \\ &= \hat{a}^\dagger.\end{aligned} \tag{7.10}$$

Lemma 1. *The eigenvalues of $\hat{N}$ are non-negative. Further, if ν is an eigenvalue of $\hat{N}$ with eigenvector $|\nu\rangle$ then $\hat{a}|\nu\rangle$ is the zero vector if and only if $\nu = 0$.*

Proof. Consider an eigenstate $|\nu\rangle$ of $\hat{N}$ with eigenvalue ν, that is $\hat{N}|\nu\rangle = \nu|\nu\rangle$. Consider now the norm of the vector $\hat{a}|\nu\rangle$. We have

$$\begin{aligned}||\hat{a}|\nu\rangle||^2 &\geq 0 \\ \langle\nu|\hat{a}^\dagger\hat{a}|\nu\rangle &\geq 0 \\ \langle\nu|\hat{N}|\nu\rangle &\geq 0 \\ \nu\langle\nu|\nu\rangle &\geq 0,\end{aligned} \tag{7.11}$$

from which it follows that $\nu \geq 0$. □

Further, for $\nu = 0$ we have that $||\hat{a}|\nu\rangle||^2 = 0$, which means that $\hat{a}|0\rangle$ is the zero vector, we write $\hat{a}|0\rangle = 0$. (Note that $|0\rangle$ is *not* the zero vector here, it is simply a vector labelled by the index 0.) Also note that $\hat{a}|\nu\rangle$ can be the zero vector for no other value of ν according to equation (7.11).

Lemma 2. *If $|\nu\rangle$ is an eigenvector of $\hat{N}$ corresponding to the eigenvalue ν, and $\hat{a}|\nu\rangle \neq 0$, then $\hat{a}|\nu\rangle$ is also an eigenvector of $\hat{N}$ corresponding to the eigenvalue $\nu - 1$.*

Proof.

$$\begin{aligned}\hat{N}\hat{a}|\nu\rangle &= \left(\hat{N}\hat{a} - \hat{a}\hat{N} + \hat{a}\hat{N}\right)|\nu\rangle \\ &= \left([\hat{N}, \hat{a}] + \hat{a}\hat{N}\right)|\nu\rangle \\ &= \left(-\hat{a} + \hat{a}\hat{N}\right)|\nu\rangle \\ &= \hat{a}(\hat{N} - 1)|\nu\rangle \\ &= (\nu - 1)\hat{a}|\nu\rangle.\end{aligned} \tag{7.12}$$

□

Combining Lemma 1 and 2 we conclude that either the spectrum is empty (i.e. there are no eigenvalues at all) or $\nu = 0$ has to be in the spectrum, and no non-integer value of ν could possibly be in the spectrum (to guarantee that the series ν, $\nu - 1$, $\nu - 2, \ldots$ does not continue to negative values).

Lemma 3. *If $|\nu\rangle$ is an eigenvector of $\hat{N}$ corresponding to the eigenvalue ν, then $\hat{a}^\dagger|\nu\rangle$ is also an eigenvector of $\hat{N}$ corresponding to the eigenvalue $\nu + 1$.*

Exercise 11. Prove Lemma 3.

What we have shown now is that $\nu = 0$ *could be* an eigenvalue of $\hat{N}$. If it is, then starting from $\nu = 0$ via consecutive application of $\hat{a}^\dagger$ we thus deduce that all integers are eigenvalues of $\hat{N}$. We now have to show that there indeed is a vector $|0\rangle$ that is an eigenvector of $\hat{N}$ with eigenvalue 0. Once we have done this we will have deduced that $\hat{N}|n\rangle = n|n\rangle$ with $n = 0, 1, 2, \ldots$.

To show that there indeed is a ground state let us explicitly construct its position representation. We can deduce the functional form of $\phi_0(x)$ from the equation $\hat{a}|0\rangle = 0$ in position representation.

First we remember the definition of $\hat{a}$ in terms of $\hat{q}$ and $\hat{p}$:

$$\hat{a} := \frac{1}{\sqrt{2}}\left(\hat{Q} + \mathrm{i}\hat{P}\right) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\hat{q} + \mathrm{i}\sqrt{\frac{1}{m\omega\hbar}}\hat{p}\right) \tag{7.13}$$

We recall the expressions for $\hat{q}$ and $\hat{p}$ in position representation,

$$\langle x|\hat{q}|\phi\rangle = x\phi(x), \quad \text{and} \quad \langle x|\hat{p}|\phi\rangle = -\mathrm{i}\hbar\frac{\partial}{\partial x}\phi(x), \tag{7.14}$$

to find

$$\langle x|\hat{a}|\phi\rangle = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x + \sqrt{\frac{\hbar}{m\omega}}\frac{\partial}{\partial x}\right)\phi(x). \tag{7.15}$$

Thus, the ground state wave function fulfils the condition

$$\frac{\partial}{\partial x}\phi_0(x) = -\frac{m\omega}{\hbar}x\phi_0(x). \tag{7.16}$$

Thus we find

$$\phi_0(x) = \phi_0(x) = c\,\mathrm{e}^{-\frac{m\omega}{2\hbar}x^2}, \qquad c \in \mathbb{C}. \tag{7.17}$$

We can deduce the normalisation constant c by requiring the wave function to be normalised. From

$$\int_{-\infty}^{\infty}|\phi_0(x)|^2\mathrm{d}x = \int_{-\infty}^{\infty}|c|^2\mathrm{e}^{-\frac{m\omega}{\hbar}x^2}\mathrm{d}x = 1, \tag{7.18}$$

we deduce

$$|c|^2 = \sqrt{\frac{m\omega}{\pi\hbar}}, \tag{7.19}$$

and thus we choose

$$c = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}, \tag{7.20}$$

such that the ground state wave function is normalised, real, and positive. In summary, we have found the position representation of the normalised ground state wave function of the operator $\hat{N}$ (and thus the harmonic oscillator):

$$\phi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}x^2\right). \tag{7.21}$$

This completes our proof of Theorem 8.

Thus, the spectrum of the harmonic oscillator consists of the discrete set of numbers $E_n = \hbar\omega(n + \frac{1}{2})$ where n is a non-negative integer,

$$\hat{H}|n\rangle = \hbar\omega(n + \tfrac{1}{2})|n\rangle. \tag{7.22}$$

Remark 35. The operators $\hat{a}$ and $\hat{a}^\dagger$ are often referred to as *ladder operators*, more specifically $\hat{a}$ is the *lowering* and $\hat{a}^\dagger$ the *raising* operator. Often one also referes to $\hat{a}$ as *annihilation* and $\hat{a}^\dagger$ as *creation* operator. The operator $\hat{N}$ is called the *number* operator.

7.1.1 Normalised eigenvectors

Let us first convince ourselves that the eigenvalues n are non-degenerate, i.e. there is a unique (up to a complex factor) eigenvector $|n\rangle$ for each eigenvalue n. We have seen already that the ground state is a single state. We shall now prove that if the eigenvalue n is non-degenerate so is the eigenvalue $n+1$. Let us prove this by contradiction.

Assume that the eigenvalue n is non-degenerate, but there are two linearly independent eigenstates $|\phi_{n+1}\rangle$ and $|\chi_{n+1}\rangle$ belonging to the eigenvalue $n+1$. Applying the lowering operator to these two states we can produce the states $|\phi_n\rangle = \hat{a}|\phi_{n+1}\rangle$ and $|\chi_n\rangle = \hat{a}|\chi_{n+1}\rangle$. Now, since the eigenvalue n is non-degenerate these two vectors have to be multiples of each other, i.e. there exists a $\lambda \in \mathbb{C}$ such that

$$|\phi_n\rangle = \lambda|\chi_n\rangle. \tag{7.23}$$

Now acting on this relation with the raising operator $\hat{a}^\dagger$ yields:

$$\hat{a}^\dagger|\phi_n\rangle = \hat{a}^\dagger\hat{a}|\phi_{n+1}\rangle = \hat{N}|\phi_{n+1}\rangle = \lambda\hat{a}^\dagger|\chi_n\rangle = \lambda\hat{a}^\dagger\hat{a}|\chi_{n+1}\rangle = \lambda\hat{N}|\chi_{n+1}\rangle, \tag{7.24}$$

that is $\hat{N}|\phi_{n+1}\rangle = \lambda\hat{N}|\chi_{n+1}\rangle$. But since $|\phi_{n+1}\rangle$ and $|\chi_{n+1}\rangle$ are eigenstates of $\hat{N}$ with eigenvalue $(n+1)$, we have

$$|\phi_{n+1}\rangle = \lambda|\chi_{n+1}\rangle, \tag{7.25}$$

which means the two assumed to be linearly independent eigenvectors are linearly dependent after all, and the eigenvalue $n+1$ is also non-degenerate. Thus, starting from $n=0$ we have shown that the eigenvalues are non-degenerate. □

We know that $\hat{a}^\dagger|n\rangle = c_n|n+1\rangle$. Let us now deduce the constants c_n assuming that the eigenvectors are normalised. We consider the norm of $|n+1\rangle$:

$$\langle n+1|n+1\rangle = \frac{1}{|c_n|^2}\langle n|\hat{a}\hat{a}^\dagger|n\rangle = \frac{1}{|c_n|^2}\langle n|\hat{N}+1|n\rangle = \frac{n+1}{|c_n|^2}\langle n|n\rangle. \tag{7.26}$$

Thus, if $|n\rangle$ is normalised, so is $|n+1\rangle$ for the choice $c_n = \sqrt{n+1}$. That is,

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \tag{7.27}$$

from which it follows that

$$|n\rangle = \frac{1}{\sqrt{n!}}\hat{a}^{\dagger n}|0\rangle. \tag{7.28}$$

Remark 36. The corresponding equation to (7.27) for $\hat{a}$ is

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle. \tag{7.29}$$

7.1.2 Matrix representations in the harmonic oscillator basis

1. **Matrix representation of $\hat{a}^\dagger$:** From equation (7.27) it follows for the matrix elements of $\hat{a}^\dagger$ in the harmonic oscillator basis:

$$\langle m|\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,\delta_{m,n+1}. \tag{7.30}$$

2. **Matrix representation of $\hat{a}$:** We can deduce the matrix elements for $\hat{a}$ either directly from equation (7.29) or from the fact that $\hat{a}$ is the adjoint of $\hat{a}^\dagger$, and thus their matrix elements are related by

$$\langle m|\hat{a}|n\rangle = \overline{\langle n|\hat{a}^\dagger|m\rangle}. \tag{7.31}$$

From the matrix elements of $\hat{a}^\dagger$ we thus find the matrix elements of $\hat{a}$ as

$$\langle m|\hat{a}|n\rangle = \sqrt{n}\,\delta_{m,n-1} \tag{7.32}$$

3. **Matrix representation of $\hat{H}$:** The Hamiltonian matrix is diagonal in the eigenbasis, with diagonal elements

$$\langle m|\hat{H}|n\rangle = \hbar\omega(n+\frac{1}{2})\,\delta_{m,n} \tag{7.33}$$

Exercise 12. Matrix representation of $\hat{x}$ and $\hat{p}$.

(a) What are the matrix elements of the operators $\hat{q}$ and $\hat{p}$ in the basis of harmonic oscillator eigenstates $\{|n\rangle\}$?

(b) Verify that these matrices fulfil the correct commutation relation.

7.2 The eigenfunctions in position representation

While we have an abstract expression for the chain of eigenvectors starting from the ground state in equation (7.28), we would now like to find the position representations of these eigenvectors, $\phi_n(x) := \langle x|n\rangle$. Using expression (7.28) we can deduce all $\phi_n(x)$ from $\phi_0(x)$. To find the position representations of the higher excited states, we need to apply $\hat{a}^\dagger$ onto the ground state:

$$\langle x|n\rangle = \phi_n(x) = \frac{1}{\sqrt{n!}}\langle x|(\hat{a}^\dagger)^n|0\rangle. \tag{7.34}$$

That is,

$$\begin{aligned}\phi_n(x) &= \frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{\partial}{\partial x}\right)\right)^n \phi_0(x)\\ &= \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\frac{1}{\sqrt{2^n}}\left(\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{\partial}{\partial x}\right)^n \exp\left(-\frac{m\omega}{2\hbar}x^2\right).\end{aligned} \tag{7.35}$$

Explicitly the $\phi_n(x)$ can be expressed in terms of *Hermite polynomials* $\mathrm{H}_n(x)$ as

$$\phi_n(q) = \sqrt{\frac{\sqrt{m\omega}}{2^n n!\sqrt{\pi\hbar}}}\,\mathrm{H}_n(q)\,\mathrm{e}^{-q^2/2}, \tag{7.36}$$

with $q = \sqrt{\frac{m\omega}{\hbar}}x$. The Hermite polynomials are defined as

$$\mathrm{H}_n(q) = \mathrm{e}^{q^2/2}\left(q - \frac{\mathrm{d}}{\mathrm{d}q}\right)^n \mathrm{e}^{-q^2/2}. \tag{7.37}$$

They are polynomials of degree n with alternating parity (even, odd, even, ...). They can alternatively be defined by the recursion relation

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x),$$

with $H_0(x) = 1$, and $H_1(x) = 2x$. They further fulfil the relation

$$H_n'(x) = 2nH_{n-1}(x).$$

7.3 Harmonic oscillator basis - A simple numerical method for general potentials

The matrix representation of position and momentum operator in the harmonic oscillator basis forms the basis of a simple numerical method to calculate eigenvalues of one-dimensional systems with Hamiltonians that are functions of position and momentum operators. Here the matrices are

terminated at finite size, leading to an approximation. For more details on this method see H. J. Korsch and M. Glück, *Computing quantum eigenvalues made easy*, Eur. J. Phys. **23**, 413 (2002).

The following matlab program gives out the 8 lowest energy states of the harmonic oscillator Hamiltonian $\hat{H} = \hat{p}^2/2 + \hat{x}^2/2$:

```
clear all
N=100; %matrix size

%matrices for postition and momentum operators
n=1:N-1;
k=sqrt(n);

Q=sqrt(0.5)*(diag(k,1)+diag(k,-1));
P=i*sqrt(0.5)*(diag(k,-1)-diag(k,1)):

%Hamiltonian
H=0.5*P^2+0.5*Q^2;

%eigenvalues
EigSort=sort(eig(H));
EigSort(1:8)
```

This programme can be modified to calculate the spectrum of many other Hamiltonians with different potential functions. In class I will show you the example of a quartic potential $V(x) = \lambda x^4$.

Part IV

Quantum Dynamics and Angular momentum

Chapter 8

Quantum dynamics

In this chapter we shall turn away from the spectral problems, and focus our attention on the dynamical features of quantum theory. The most interesting features of observable things are often those that change in time, after all. We shall discuss dynamics of expectation values, which yields a version of the famous *Heisenberg equations* of motion, which can be viewed as an alternative to Schrödinger's dynamical equation in quantum mechanics. Further, we shall investigate the dynamical behaviour of a free Gaussian wave packet (i.e. for vanishing potential). We then turn to the discussion of quantum dynamics from a Lie algebraic perspective.

8.1 Dynamics of expectation values

From the time-dependent Schrödinger equation, $\mathrm{i}\hbar|\dot{\psi}\rangle = \hat{H}|\psi\rangle$, we can deduce a dynamical equation for observable expectation values $\langle\hat{A}\rangle$: We have by definition

$$\langle\hat{A}\rangle = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{8.1}$$

Thus, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{A}\rangle = \frac{\langle\dot{\psi}|\hat{A}|\psi\rangle + \langle\psi|\dot{\hat{A}}|\psi\rangle + \langle\psi|\hat{A}|\dot{\psi}\rangle}{\langle\psi|\psi\rangle} - \frac{\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle^2}\langle\psi|\hat{A}|\psi\rangle. \tag{8.2}$$

From the Schrödinger equation we know

$$|\dot{\psi}\rangle = -\frac{\mathrm{i}}{\hbar}\hat{H}|\psi\rangle, \quad \text{and} \quad \langle\dot{\psi}| = \frac{\mathrm{i}}{\hbar}\langle\psi|\hat{H}, \tag{8.3}$$

thus, the overall norm is conserved, $\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\psi\rangle = 0$ and we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{A}\rangle &= \frac{\frac{\mathrm{i}}{\hbar}\langle\psi|\hat{H}\hat{A}|\psi\rangle + \langle\psi|\dot{\hat{A}}|\psi\rangle - \frac{\mathrm{i}}{\hbar}\langle\psi|\hat{A}\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \\ &= \frac{\mathrm{i}}{\hbar}\langle\hat{H}\hat{A} - \hat{A}\hat{H}\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle.\end{aligned} \tag{8.4}$$

Or, using the commutator,

$$\frac{\mathrm{d}\langle\hat{A}\rangle}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}\langle[\hat{H},\hat{A}]\rangle + \langle\frac{\partial\hat{A}}{\partial t}\rangle. \tag{8.5}$$

This is one formulation of the famous *Heisenberg equation* of motion for observables, which can be used as the foundational dynamical equation of quantum mechanics, as an alternative of Schrödinger's equation. Comparing the Heisenberg equation to the classical equation of motion (1.16) for a phase space variable $A(p,q)$ we observe a striking similarity if we associate expectation values with classical variables and interpret the commutator, or more precisely $\mathrm{i}/\hbar$ times the commutator, as a *quantum Poisson bracket*.

The Ehrenfest theorem

The special case of the equations of motion for the expectation values of position and momentum for a system with Hamiltonian of the type

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(\hat{q}) \tag{8.6}$$

forms the basis of what is referred to as the *Ehrenfest theorem*, which we shall consider in what follows. From equation (8.5) and the form of the Hamiltonian (8.6) we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{q}\rangle &= \frac{\mathrm{i}}{\hbar}\langle[\frac{\hat{p}^2}{2m}, \hat{q}]\rangle \\ &= \frac{\mathrm{i}}{2\hbar m}\langle[\hat{p}^2, \hat{q}]\rangle \\ &= \frac{\mathrm{i}}{2\hbar m}\langle\hat{p}[\hat{p}, \hat{q}] + [\hat{p}, \hat{q}]\hat{p}\rangle. \end{aligned} \tag{8.7}$$

Recalling the commutation relation between position and momentum operator, $[\hat{q}, \hat{p}] = \mathrm{i}\hbar$, this reduces to

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{q}\rangle = \frac{\langle\hat{p}\rangle}{m}. \tag{8.8}$$

For the momentum operator we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{p}\rangle = \frac{\mathrm{i}}{\hbar}\langle[\hat{V}(\hat{q}), \hat{p}]\rangle. \tag{8.9}$$

To calculate the commutator between the momentum operator and the potential, it is convenient to use the position representation, in which we have

$$[\hat{V}(\hat{q}), \hat{p}] = -\mathrm{i}\hbar[V(x), \frac{\partial}{\partial x}], \tag{8.10}$$

applied to a test function $\psi(x)$,

$$\begin{aligned} -\mathrm{i}\hbar\,[V(x), \frac{\partial}{\partial x}]\,\psi(x) &= -\mathrm{i}\hbar V(x)\psi'(x) + \mathrm{i}\hbar\frac{\partial}{\partial x}(V(x)\psi(x)) \\ &= -\mathrm{i}\hbar V(x)\psi'(x) + \mathrm{i}\hbar V'(x)\psi(x) + \mathrm{i}\hbar V(x)\psi'(x) \\ &= \mathrm{i}\hbar V'(x)\psi(x) \end{aligned} \tag{8.11}$$

Translating this back to the abstract operator notation this is

$$[\hat{V}(\hat{q}), \hat{p}] = \mathrm{i}\hbar\frac{\mathrm{d}V(\hat{q})}{\mathrm{d}\hat{q}}, \tag{8.12}$$

and thus we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{p}\rangle = -\langle\frac{\partial\hat{V}}{\partial\hat{q}}\rangle. \tag{8.13}$$

In summary we have obtained the set of equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{q}\rangle = \frac{\langle\hat{p}\rangle}{m}, \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{p}\rangle = -\langle\frac{\partial\hat{V}}{\partial\hat{q}}\rangle, \tag{8.14}$$

which are strikingly similar in form to the classical equations of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}q = \frac{p}{m}, \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}p = -\frac{\partial V}{\partial q}, \tag{8.15}$$

a fact that is often referred to as the *Ehrenfest theorem*.

Notwithstanding the formal similarity, there are of course important differences between quantum and classical dynamics. On the level of the Ehrenfest theorem, this is due to the fact that in general

$$\langle \frac{\partial \hat{V}}{\partial \hat{q}} \rangle \neq \frac{\partial V(\langle \hat{q} \rangle)}{\partial \langle \hat{q} \rangle}. \tag{8.16}$$

Consider for example the case $\hat{V}(\hat{q}) = \hat{q}^4$. Here we have

$$\frac{\partial \hat{V}(\langle \hat{q} \rangle}{\partial \langle \hat{q} \rangle} = 4\langle \hat{q} \rangle^3, \tag{8.17}$$

while

$$\langle \frac{\partial \hat{V}}{\partial \hat{q}} \rangle = 4\langle \hat{q}^3 \rangle. \tag{8.18}$$

The two expressions in equation (8.16) are, however, similar for potentials that are polynomials of $\hat{q}$ of order two or lower. For more general potentials equations (8.14) in fact not even form a closed set of equations (as the left side of the time derivative of $\langle \hat{p} \rangle$ is not only a function of $\langle \hat{q} \rangle$ and $\langle \hat{p} \rangle$).

Literary Excursion

In the following I provide a translation of a short excerpt from the novel *Dirac* by Dietmar Dath (Surkamp 2006) to provide some further insight into the development of quantum mechanics, albeit in the form of a literary supplement.

Just like a melody, which the whistling lips are faster to grasp than the conscious mind, Heisenberg's mathematics returns to the focus of Dirac's attention. Non-commutativity, this times that is different from that times this. A pale-blue bird thrusts itself onto the strong wind just above Dirac's head, struggling, staggering, spinning, re-balancing, it positions itself tangentially to the air stream, and then rises again, very fast and very high. Dirac is viewing the bird in a coordinate plane, seeing it rotating, and fleeing upwards.

Position and velocity, where the bird is thrown towards, and how it is flapping its wings: Sky and puddle, wind and water, bird and – fish? "Poisson", Dirac says.

That's it! Poisson brackets. Siméon Denis Poisson, seventeen hundred eighty one to eighteen hundred forty. Non-commutativity... Where has he read about this? What was it? There seems to be a connection. Or is he just grabbing for thin air? Heisenberg's commutator: $uv - vu$, and the relation between the two quantities in this formalism for phase space transforms. ...This is only the faintest memory: Old mathematics, which might come to rescue Heisenberg's new physics. Dirac is leaping over the road-side ditch without a thought on what he is doing, and he is now walking right across the muddy meadow. Every step is making slurping sounds, Dirac is in a hurry: He has to go home, has to look it up, find out what's going on there... Poisson, non-commutativity – everything.

Dirt splashing on his trousers, his joints cracking, he is hurrying on. It can't be true. The similarity of these two forms must be a coincidence, just like a rhyme in a poem or a harmony of two voices. It would be too weird for such a coincidence to actually mean something. Already he has a feeling that this thought, which had only just occurred to him, might be implausible, but his excitement is growing at the same time: What if there was something to it? After a few hundred metres of cross-country walking he is now reaching a footpath and finding himself panting for air, staggering, dizzy: He had been running. He has left the rural area behind and is now in the town. He is pushed forward into his thoughts by a tailwind, which is seizing him just as it had seized the bird before him.

Finally he reaches his door, hectically searching his coat pocket for his key. As he tries to unlock the door too fast, the key gets stuck. He is swearing under his breath. He has to get in there, look it up, clear definitions, order! He is rattling at the lock, which finally releases the key. Dirac immediately reinserts it and this time the door opens. Without even taking off his coat Dirac is rushing through his room, searching his books, his lecture notes, kneeling on the floor, opening all

the drawers.

The storm that had been raging outside earlier is revived again in this room; sheets are flying, eyes are searching, but nothing: canonical transformations, linear mappings, differential equations, textbooks, mathematical tables. All of this is much too fundamental. Nothing on Poisson brackets, no hint. But he did see it. Did read it: In a book, not in a lecture note.

Finally he is sitting in the midst of the chaos he has produced, panting. He takes off his coat and it makes him chuckle: Sunday night. All libraries closed. He runs his fingers through his hair, as if what he is looking for might be stuck in there.

Dreadful, euphoric night: Dirac's heart is pounding in his throat, half asleep he is imagining himself as a little bird in the wind, throwing himself from left to right. If it really was possible to express things this way, if one could formulate a type of Hamiltonian dynamics using non-commutative Poisson-brackets... If that was the key...

Dawn is breaking. Dirac gets up. His joints are painful. He is extraordinarily weary, as if he had just finished some sort of hard work. After a quick superficial wash he gets dressed and leaves the house, walking straight towards the faculty library. He is the first one there. The librarian allows him in, shaking his head. Dirac is rushing past long rows of bookshelves, he knows his way around here.

He takes a left turn at the big cross-bar window in front of one of the columns. There, at breast height, is Whittaker's "Analytical Dynamics". Dirac quickly finds what he was looking for. The letters are flickering, and he is finding it difficult to read. Yet, his intuition had not failed him. He now has his own key to the latest physics...

8.2 Dynamics of wave packets

8.2.1 Gaussian wave packets

We consider wave functions of the form

$$\psi(x) = N \exp\left(-\alpha(x-q)^2 + \frac{\mathrm{i}}{\hbar}p(x-q) + \frac{\mathrm{i}}{\hbar}\gamma\right), \tag{8.19}$$

where $q, p \in \mathbb{R}$ and $\alpha, \gamma \in \mathbb{C}$, with $\mathrm{Re}(\alpha) > 0$ are parameters, and N is a normalisation constant.

Demanding that the wave function (8.19) is normalised, that is $\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1$, and choosing N real and positive, we find

$$N = \left(\frac{2\mathrm{Re}(\alpha)}{\pi}\right)^{1/4} \mathrm{e}^{\mathrm{Im}(\gamma)/\hbar}. \tag{8.20}$$

The expectation values of $\hat{q}$ and $\hat{p}$ are given by

$$\langle \hat{q} \rangle = q, \qquad \langle \hat{p} \rangle = p. \tag{8.21}$$

For the uncertainty Δq of the position we find

$$\Delta q = \sqrt{\frac{1}{4\mathrm{Re}(\alpha)}}, \tag{8.22}$$

and for the momentum uncertainty we obtain

$$\Delta p = \frac{\hbar|\alpha|}{\sqrt{\mathrm{Re}(\alpha)}}. \tag{8.23}$$

Thus we have

$$\Delta q \Delta p = \frac{\hbar}{2}\frac{|\alpha|}{\mathrm{Re}(\alpha)} \geq \frac{\hbar}{2}. \tag{8.24}$$

That is, the lower bound of Heisenberg's uncertainty relation is reached for $\alpha \in \mathbb{R}$.

The so-called Gaussian wave packets (8.19) are of great importance for a number of reasons. Among them are:

- The Gaussian is the only form of a wave function that in principle allows to reach the lower bound of Heisenberg's uncertainty relation for $\hat{q}$ and $\hat{p}$.
- A Gaussian (8.19) can be interpreted as a localised particle with a rather well defined momentum, as they have a single maximum at the expectation value of $\hat{q}$ and similarly their momentum representation has a single maximum at the expectation value of $\hat{p}$.
- They are form-invariant under time-evolution generated by Hamiltonians of degree two or lower in $\hat{q}$ and $\hat{p}$.
- Gaussian integrals lead to simple analytic expressions for expectation values, uncertainties etc.

8.2.2 Time evolution of the free Gaussian wave packet

Let us now study the time evolution of a free particle according to quantum mechanics. That is, we aim to solve the time-dependent Schrödinger equation in position representation

$$\mathrm{i}\hbar\dot{\psi}(x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x), \tag{8.25}$$

for a particular given initial condition. Here we consider the initial condition to be a Gaussian wave packet, as this is the closest to the notion of a particle with a well-defined position and momentum. That is, we solve equation (8.25) for the initial condition

$$\psi(x) = N\exp\left(-\alpha_0(x-q_0)^2 + \frac{\mathrm{i}}{\hbar}p_0(x-q_0) + \frac{\mathrm{i}}{\hbar}\gamma_0\right). \tag{8.26}$$

We shall show that the time-dependent wave function indeed stays Gaussian, that is, it is of the form (8.19) with time dependent parameters α_t, q_t, p_t, γ_t. This can be seen directly by inserting this ansatz into the time-dependent Schrödinger equation and find the time-dependence of the parameters. (Remember that the Schrödinger equation is a first order differential equation with respect to time, and thus has a unique solution for a given initial state at a specific time.)

To insert the ansatz

$$\psi(x,t) = \exp\left(-\alpha_t(x-q_t)^2 + \frac{\mathrm{i}}{\hbar}p_t(x-q_t) + \frac{\mathrm{i}}{\hbar}\gamma_t\right), \tag{8.27}$$

into the time-dependent Schrödinger equation we first calculate

$$\frac{\partial^2}{\partial x^2}\psi(x,t) = \left[\left(-2\alpha_t(x-q_t) + \frac{\mathrm{i}}{\hbar}p_t\right)^2 - 2\alpha_t\right]\psi(x,t) \tag{8.28}$$

and

$$\frac{\partial}{\partial t}\psi(x,t) = \left[-\dot{\alpha}_t(x-q_t)^2 + 2\alpha_t(x-q_t)\dot{q}_t + \frac{\mathrm{i}}{\hbar}\dot{p}_t(x-q_t) - \frac{\mathrm{i}}{\hbar}p_t\dot{q} + \frac{\mathrm{i}}{\hbar}\dot{\gamma}_t\right]\psi(x,t). \tag{8.29}$$

This yields

$$\begin{aligned}\left[-\mathrm{i}\hbar\dot{\alpha}_t(x-q_t)^2\right. &+2\mathrm{i}\hbar\alpha_t(x-q_t)\dot{q}_t - \dot{p}_t(x-q_t) + p_t\dot{q} - \left.\dot{\gamma}_t\right]\psi(x,t)\\ &= \left[-\frac{2\hbar^2}{m}\alpha_t^2(x-q_t)^2 + \frac{2\mathrm{i}\hbar}{m}\alpha_t p_t(x-q_t) + \frac{1}{2m}p_t^2 + \frac{\hbar^2}{m}\alpha_t\right]\psi(x,t).\end{aligned} \tag{8.30}$$

This equation has to hold for all values of x, and in particular for all orders of $(x-q_t)$ separately. That is, we can deduce three separate equations from (8.30): The quadratic term (i.e. the coefficient of $(x-q_t)^2$) yields

$$-\mathrm{i}\hbar\dot{\alpha}_t = -\frac{2\hbar^2}{m}\alpha_t^2. \tag{8.31}$$

From the linear term we find

$$2\mathrm{i}\hbar\alpha_t\dot{q}_t - \dot{p}_t = \frac{2\mathrm{i}\hbar}{m}\alpha_t p_t, \tag{8.32}$$

and the constant term yields

$$p_t\dot{q}_t - \dot{\gamma}_t = \frac{1}{2m}p_t^2 + \frac{\hbar^2}{m}\alpha_t. \tag{8.33}$$

Equation (8.31) directly gives a dynamical equation for the complex parameter α_t that does not depend on any of the other parameters:

$$\dot{\alpha}_t = -\frac{2\mathrm{i}\hbar}{m}\alpha^2. \tag{8.34}$$

Equation (8.32) encodes dynamical equations for the two real parameters q_t and p_t. To obtain separate equations for the two parameters we divide the equation into its real and imaginary parts. The imaginary part yields

$$2\hbar\mathrm{Re}(\alpha_t)\dot{q}_t = \frac{2\hbar}{m}\mathrm{Re}(\alpha_t)p_t, \tag{8.35}$$

that is

$$\dot{q}_t = \frac{p_t}{m}, \tag{8.36}$$

just as in the corresponding classical problem.

The real part of (8.32) is given by

$$-2\hbar\mathrm{Im}(\alpha_t)\dot{q}_t - \dot{p}_t = -\frac{2\hbar}{m}\mathrm{Im}(\alpha_t)p_t \tag{8.37}$$

Inserting $\dot{q}_t = \frac{p_t}{m}$ yields

$$-\frac{2\hbar}{m}\mathrm{Im}(\alpha_t)p_t - \dot{p}_t = -\frac{2\hbar}{m}\mathrm{Im}(\alpha_t)p_t, \tag{8.38}$$

that is,

$$\dot{p}_t = 0, \tag{8.39}$$

which is also equivalent to the corresponding classical dynamics. (Remember Ehrenfest's theorem?!)

Inserting (8.36) into the equation from the constant term (8.33) yields a dynamical equation for γ_t:

$$\dot{\gamma}_t = -\frac{\hbar^2}{m}\alpha_t + \frac{p_t^2}{2m}, \tag{8.40}$$

which is coupled to the dynamics of α_t and p_t. In summary we can deduce the time-dependent state from the solutions of the system of coupled differential equations of first order

$$\dot{\alpha}_t = -\frac{2\mathrm{i}\hbar}{m}\alpha^2, \quad \dot{q}_t = \frac{p_t}{m}, \quad \dot{p}_t = 0, \quad \text{and} \quad \dot{\gamma}_t = -\frac{\hbar^2}{m}\alpha_t + \frac{p_t^2}{2m}. \tag{8.41}$$

The equations for q_t and p_t are trivially solved. We have

$$p_t = p_0, \quad \text{and} \quad q_t = q_0 + \frac{p_0}{m}t. \tag{8.42}$$

The equation for α_t can be integrated to yield

$$\alpha_t = \frac{\alpha_0}{1 + \frac{2\mathrm{i}\hbar}{m}\alpha_0\, t}. \tag{8.43}$$

That then yields an explicitly time-dependent dynamical equation for γ_t:

$$\dot{\gamma}_t = -\frac{\frac{\hbar^2\alpha_0}{m}}{1 + \frac{2\mathrm{i}\hbar\alpha_0}{m}\, t} + \frac{p_0^2}{2m}, \tag{8.44}$$

which is straight-forwardly integrated to yield

$$\gamma_t = \frac{p_0^2}{2m}t + \frac{\mathrm{i}\hbar}{2}\ln\left(1 + \frac{2\mathrm{i}\hbar\alpha_0}{m}t\right) + \gamma_0. \tag{8.45}$$

Thus we found the time-dependent wave function of the free particle, assuming the initial state was a Gaussian wave packet. The expectation values of position and momentum move exactly as in the classical case here. In addition, the uncertainty product increases in time. Consider for example an initial minimum uncertainty wave packet, i.e., $\alpha_0 \in \mathbb{R}$. We then find

$$\Delta q = \frac{1}{2}\sqrt{\frac{m^2 + 4\hbar^2\alpha_0^2 t^2}{m^2\alpha_0}}, \tag{8.46}$$

while Δp stays constant, $\Delta p = \hbar\sqrt{\alpha_0}$, and thus we have an increase of the uncertainty product in time as

$$\Delta q \Delta p = \frac{\hbar}{2}\sqrt{1 + \frac{4\hbar^2\alpha_0^2}{m^2}\,t^2}. \tag{8.47}$$

This spread of a free quantum wave packet is known as *wave packet dispersion.*

Remark 37. The dispersion of a wave packet is in some sense a *mild* quantum effect, as it is entirely due to the initial uncertainty. A classical propagation of an ensemble of initial conditions reflecting the initial uncertainty of the quantum state would result in the identical dynamics for expectation values and uncertainties.

8.3 Lie groups and Lie algebras in quantum dynamics

The time evolution in quantum mechanics is closely related to the concept of (Lie) groups. In this section, I will give a brief summary of the most important concepts in this context.

We begin with a brief reminder of the concept of a

Group. A *group* is a set $\mathcal{G}$ of elements, together with a map $\circ : \mathcal{G}\times\mathcal{G} \to \mathcal{G}$, which fulfil the following three properties:

- The map is associative, i.e. $(a \circ b) \circ c = a \circ (b \circ c)$, for all $a, b, c \in \mathcal{G}$,
- The group contains an identity element e with $e \circ a = a \circ e = a$ for all $a \in \mathcal{G}$,
- Every element has an inverse element in the group, often denoted by a^{-1}, such that $a \circ a^{-1} = a^{-1} \circ a = e$.

A *representation* of a group can be thought of as a realisation of group elements as linear operators on a linear vectorspace $\mathcal{V}$, that is, a map (*a group homomorphism*) $g \to \hat{D}(g)$, from $\mathcal{G}$ to $\mathcal{V}$, which fulfils

$$\hat{D}(g_1 \circ g_2) = \hat{D}(g_1)\hat{D}(g_2). \tag{8.48}$$

Thus we have $\hat{D}(g^{-1}) = (\hat{D}(g))^{-1}$, and $\hat{D}(e) = \hat{I}$, where $\hat{I}$ denotes the identity on $\mathcal{V}$. A representation is *faithful* if the map $g \to \hat{D}(g)$ is invertible (an isomorphism).

A *continuous group* is a group with elements that continuously depend on parameters, i.e. the group elements are functions of a set of parameters. The set of unitary operations on a Hilbert space together with the consecutive application of unitary operations as the group map, form a continuous group.

A special type of continuous groups are *Lie groups.* Whole lectures and textbooks are devoted to the topic of Lie groups. We will only devote a couple of paragraphs, sacrificing rigour for an intuitive idea. Broadly speaking, an n-dimensional *Lie group* is a continuous group, the elements g of which can be represented as C^n functions of n real parameters, with $g(0) = e$. Thus, Lie groups are both groups and analytical manifolds.

If $\hat{D}(g)$ is a representation of a Lie group element, we write for short $\hat{D}(x) = \hat{D}(g(x))$, and $\hat{D}(0) = \hat{I}$. Since a Lie group element is a differentiable function of the parameters, for small values of x we can expand an element into a power series with the first two terms

$$\hat{D}(x) = \hat{I} + \mathrm{i}\sum_{j=1}^{n} x_j \hat{T}_j + \ldots, \tag{8.49}$$

with

$$\hat{T}_j = -\mathrm{i}\left.\frac{\partial \hat{D}}{\partial x_j}\right|_{x=0}. \tag{8.50}$$

Remark 38. The factor of i in (8.49) and (8.50) is common in physics, but is not usually used in the mathematics literature.

Remark 39. The higher order terms in (8.49) are in general complicated.

The operators $\hat{T}_j$ in (8.50) are called the *generators* of the group. Let us for simplicity consider a one-dimensional group with representation $\hat{D}(x)$ with $x \in \mathbb{R}$. When the generator $\hat{T}$ is known the group structure allows to represent group elements not only for small but for arbitrary x, since $\hat{D}(2\Delta x) = \hat{D}(\Delta x)\hat{D}(\Delta x) = (\hat{D}(\Delta x))^2$ and $\hat{D}(m\Delta x) = (\hat{D}(\Delta x))^m$, and thus, if we define $x = m\Delta x$ we have

$$\hat{D}(x) = (\hat{D}(\Delta x))^m \approx (1 + \mathrm{i}\Delta x\hat{T})^m = (1 + \mathrm{i}\frac{x}{m}\hat{T})^m \xrightarrow[m\to\infty]{} \mathrm{e}^{\mathrm{i}x\hat{T}}. \tag{8.51}$$

Remark 40. Here the similarity to the relation between our time-evolution operator and our Hamiltoniann for time-independent systems should become apparent.

The n generators of an n dimensional Lie group form a *Lie algebra*. A *Lie algebra* is a vector space G with a map $[\cdot,\cdot] : G \times G \to G$, a so-called *Lie bracket*, that maps each ordered pair of algebra elements to a third algebra element, and that is anti-symmetric, bilinear, and fulfils the Jacobi identity $[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0$ (just like our Poisson brackets and commutators).

A simple example of a finite Lie algebra we encounter in quantum mechanics consists of the position and the momentum operators and the identity, since we have

$$[\hat{q},\hat{p}] = \mathrm{i}\hbar\hat{I}. \tag{8.52}$$

This algebra corresponds to the group of translations in phase space.

Another example is the algebra associated to the harmonic oscillator, spanned by the identity, $\hat{a}$, $\hat{a}^\dagger$, and $\hat{N}$, with

$$[\hat{a},\hat{a}^\dagger] = \hat{I}, \quad [\hat{N},\hat{a}] = -\hat{a}, \quad [\hat{N},\hat{a}^\dagger] = \hat{a}^\dagger. \tag{8.53}$$

The corresponding group consists of translations and rotations in phase space.

An important consequence for quantum mechanics is that if a Hamiltonian is an element of a Lie algebra, that is, a linear combination of the Lie algebra elements, $\hat{K}_j$,

$$\hat{H} = \sum_j c_j \hat{K}_j, \tag{8.54}$$

then the corresponding time evolution operator can be expressed as

$$\hat{U} = \mathrm{e}^{\mathrm{i}\sum_j \alpha_j \hat{K}_j}, \tag{8.55}$$

in what is called the *Magnus* form, or alternatively as

$$\hat{U} = \prod_j \mathrm{e}^{\mathrm{i}\beta_j \hat{K}_j}, \tag{8.56}$$

in the so-called *Wei Norman* form. Importantly this still holds if the coefficients c_j in the Hamiltonian are time dependent.

When considering the time evolution of expectation values (in the Heisenberg equation), we encounter expressions of the form $\hat{U}^\dagger \hat{A} \hat{U}$. Explicitly evaluating these can be challenging unless the Hamiltonian and the observable $\hat{A}$ are elements of a nice Lie group. In general we have

Theorem 9 (Hadamard lemma)**.**

$$\mathrm{e}^{s\hat{X}}\hat{Y}\mathrm{e}^{-s\hat{X}} = \hat{Y} + s[\hat{X},\hat{Y}] + \frac{s^2}{2!}[\hat{X},[\hat{X},\hat{Y}]] + \frac{s^3}{3!}[\hat{X},[\hat{X},[\hat{X},\hat{Y}]]] + ..., \tag{8.57}$$

where $\hat{X}$ and $\hat{Y}$ are linear operators and $s \in \mathbb{C}$ is a complex constant.

Proof. We consider $f(s) := \mathrm{e}^{s\hat{X}}\hat{Y}\mathrm{e}^{-s\hat{X}}$ as an operator valued function of the complex variable s, and Taylor expand around $s = 0$:

$$\hat{f}(s) = \sum_n \hat{f}^{(n)}(0)\frac{s^n}{n!}. \tag{8.58}$$

We have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{s\hat{X}}\hat{Y}\mathrm{e}^{-s\hat{X}} &= \mathrm{e}^{s\hat{X}}\hat{X}\hat{Y}\mathrm{e}^{-s\hat{X}} - \mathrm{e}^{s\hat{X}}\hat{Y}\hat{X}\mathrm{e}^{-s\hat{X}} \\ &= \mathrm{e}^{s\hat{X}}[\hat{X},\hat{Y}]\mathrm{e}^{-s\hat{X}}, \end{aligned} \tag{8.59}$$

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}s^2}\mathrm{e}^{s\hat{X}}\hat{Y}\mathrm{e}^{-s\hat{X}} &= \mathrm{e}^{s\hat{X}}\hat{X}[\hat{X},\hat{Y}]\mathrm{e}^{-s\hat{X}} - \mathrm{e}^{s\hat{X}}[\hat{X},\hat{Y}]\hat{X}\mathrm{e}^{-s\hat{X}} \\ &= \mathrm{e}^{s\hat{X}}[\hat{X},[\hat{X},\hat{Y}]]\mathrm{e}^{-s\hat{X}}, \end{aligned} \tag{8.60}$$

and so forth. Inserting this into the Taylor expansion yields equation (8.57). □

Exercise 13. Use the Hadamard lemma and the fundamental commutator to show that

$$\mathrm{e}^{\mathrm{i}a\hat{p}_1/\hbar}\hat{q}_1\mathrm{e}^{-\mathrm{i}a\hat{p}_1/\hbar} = \hat{q}_1 + a. \tag{8.61}$$

Chapter 9

Angular momentum

9.1 Definition of angular momentum - classical and quantum

In this chapter we move beyond one-dimensional systems and consider angular momentum. In classical mechanics the angular momentum of a particle moving in three spatial dimensions with spatial coordinate $\vec{q} \in \mathbb{R}^3$ and momentum $\vec{p} \in \mathbb{R}^3$ is given by

$$\vec{L} = \vec{q} \times \vec{p}, \tag{9.1}$$

or in components

$$L_1 = q_2 p_3 - q_3 p_2, \quad L_2 = q_3 p_1 - q_1 p_3, \quad \text{and} \quad L_3 = q_1 p_2 - q_2 p_1. \tag{9.2}$$

In quantum mechanics, we can as well consider a particle in three dimensional space $\mathbb{R}^3$, with position and momentum operators $\hat{q} = (\hat{q}_1, \hat{q}_2, \hat{q}_3)^T$, and $\hat{p} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)^T$, whose components fulfil the canonical commutation relations

$$[\hat{q}_j, \hat{p}_k] = \mathrm{i}\hbar\delta_{jk}, \qquad [\hat{q}_j, \hat{q}_k] = 0 = [\hat{p}_j, \hat{p}_k]. \tag{9.3}$$

In analogy to classical mechanics we can then define an operator associated to the *orbital angular momentum* as

$$\hat{L} = \hat{q} \times \hat{p}, \tag{9.4}$$

that is, an operator with the three components

$$\begin{aligned} \hat{L}_1 &= \hat{q}_2\hat{p}_3 - \hat{q}_3\hat{p}_2 \\ \hat{L}_2 &= \hat{q}_3\hat{p}_1 - \hat{q}_1\hat{p}_3 \\ \hat{L}_3 &= \hat{q}_1\hat{p}_2 - \hat{q}_2\hat{p}_1. \end{aligned} \tag{9.5}$$

The components of the orbital angular momentum fulfil the commutation relations

$$[\hat{L}_1, \hat{L}_2] = \mathrm{i}\hbar L_3, \quad [\hat{L}_2, \hat{L}_3] = \mathrm{i}\hbar L_1, \quad [\hat{L}_3, \hat{L}_1] = \mathrm{i}\hbar L_2, \tag{9.6}$$

or in more compact form

$$[\hat{L}_j, \hat{L}_k] = \mathrm{i}\hbar \sum_l \epsilon_{jkl} \hat{L}_l, \tag{9.7}$$

where

$$\epsilon_{jkl} = \begin{cases} 1, & j,k,l \text{ cyclic} \\ -1, & j,k,l \text{ anticyclic} \\ 0, & \text{otherwise,} \end{cases} \tag{9.8}$$

is the *Levi-Civita symbol*.

In quantum mechanics, we make the following definition:

Quantum angular momentum An operator with three Hermitian operator components $\hat{J}_{1,2,3}$ fulfilling the commutation relations

$$[\hat{J}_j, \hat{J}_k] = \mathrm{i}\hbar \sum_l \epsilon_{jkl} \hat{J}_l, \tag{9.9}$$

is an *angular momentum operator*.

The total angular momentum operator is then defined as

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2. \tag{9.10}$$

It can be easily verified that $[\hat{J}^2, \hat{J}_j] = 0$.

Remark 41. In the following we shall use the letter J to denote generic quantum angular momentum, while we reserve the letter L to denote orbital angular momentum, which can be defined as (9.5) in some given coordinate system.

9.2 Angular momentum and rotations

The above definition of angular momentum relates it to the group $SU(2)$, which, is a *double cover* of the group $SO(3)$, which can be represented as rotations in $\mathbb{R}^3$. Intuitively it makes perfect sense, that angular momentum is thus related to rotations. Technically speaking, the angular momentum operators are the generators of infinitesimal rotations. Here we shall investigate an intuitive connection between angular momentum and rotations, by showing that the time-evolution where the angular momentum operators act as Hamiltonians, are rotations of the coordinate system.

9.2.1 Reminder: The group $SU(2)$

The commutation relations (9.9) are in fact the defining equations of the Lie algebra $su(2)$, which is a generating algebra of the Lie group $SU(2)$. You have probably encountered the group $SU(2)$ before, as the group of complex unitary 2×2 matrices with determinant one, i.e. the set of matrices

$$SU(2) = \{S \in \mathbb{C}^{2\times 2} | S^\dagger = S^{-1} \text{ and } \det S = 1\}, \tag{9.11}$$

which is a Lie group under matrix multiplication.

In the mathematics literature, the associate Lie algebra is commonly defined as the set of matrices generating the group elements via the exponential map, which in the case of $SU(2)$ consists of complex anti-Hermitian 2×2 matrices with trace zero. In the physics literature we tend to include a factor of i into the exponent in the exponential map resulting in the algebra of complex Hermitian 2×2 matrices with trace 1. A basis for these is given by the matrices

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.12}$$

These matrices are also known as the *Pauli matrices*. They fulfil the commutation relations

$$[\hat{\sigma}_j, \hat{\sigma}_k] = 2\mathrm{i} \sum_l \epsilon_{jkl} \hat{\sigma}_l. \tag{9.13}$$

We observe that up to a factor of $\frac{\hbar}{2}$ this is the defining equation of quantum angular momentum. That is the martices $\hat{s}_j = \frac{\hbar}{2}\hat{\sigma}_j$ are a type of quantum angular momentum operators for the Hilbert space $\mathbb{C}^2$. The group $SU(2)$ of special unitary transformations of vectors in $\mathbb{C}^2$ is closely related to the group $SO(3)$ of rotations in R^3, as we shall remind ourselves of in what follows.

$SU(2)$ is commonly associated with complex 2×2 matrices. However, it has higher dimensional representations which will be related to quantum mechanical angular momentum, as we shall see in what follows.

9.2.2 Reminder: Rotations in $\mathbb{R}^3$

The rotations in $\mathbb{R}^3$ are a non-commutative group. A rotation in $\mathbb{R}^3$ can be described by three parameters: the direction of the axis of rotation (two parameters) and the angle of rotation, and can be expressed by an orthogonal (real unitary) 3×3 matrix. There are three fundamental rotations that can be used as a basis for all other rotations, the rotations around the x, y- and z-axes, whose matrices are given by:

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}, \quad R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix},$$

and

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Applying for example $R_x(\phi)$ to $\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ yields:

$$R_x(\phi) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y\cos\phi - z\sin\phi \\ y\sin\phi + z\cos\phi \end{pmatrix},$$

i.e. a rotation around the x-axis by the angle ϕ.

9.2.3 Action of the $\mathrm{e}^{\mathrm{i}\phi\hat{L}_j/\hbar}$ on the position operators

Consider a quantum system in $\mathbb{R}^3$ described by the wave function $|\psi\rangle$. We shall now show that the time evolution with a Hamiltonian $\hat{H} = \hat{L}_j$ (orbital angular momentum) corresponds to a rotation of the coordinate system. The time-evolved wave function is given by $|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}\hat{L}_j t/\hbar}|\psi(0)\rangle$. Thus, the expectation value of an observable $\hat{A}$ at time t is given by

$$\langle\psi(t)|\hat{A}|\psi(t)\rangle = \langle\psi(0)|\mathrm{e}^{\mathrm{i}\hat{L}_j t/\hbar}\hat{A}\mathrm{e}^{-\mathrm{i}\hat{L}_j t/\hbar}|\psi(0)\rangle. \tag{9.14}$$

Alternatively, we can interpret this as the expectation value in the wave function at time zero of the operator $\mathrm{e}^{\mathrm{i}\hat{L}_j t/\hbar}\hat{A}\mathrm{e}^{-\mathrm{i}\hat{L}_j t/\hbar}$. Let us consider the expectation values of the position operators $\hat{q}_{1,2,3}$. Denote the expectation value of $\hat{q}$ at time zero by the vector $\vec{Q} = (Q_1, Q_2, Q_3)^T \in \mathbb{R}^3$.

To calculate $\mathrm{e}^{\mathrm{i}\hat{L}_j t/\hbar}\hat{q}_k\mathrm{e}^{-\mathrm{i}\hat{L}_j t/\hbar}$ we can use the Hadamard lemma 8.57. From the definition of the components of the orbital angular momentum (9.5) we find the commutators

$$[\hat{L}_j, \hat{q}_k] = i\hbar\epsilon_{jkl}\hat{q}_l. \tag{9.15}$$

Using these and the Hadamard lemma we can calculate the $\mathrm{e}^{i\phi\hat{L}_j/\hbar}\hat{q}_k\mathrm{e}^{-i\phi\hat{L}_j/\hbar}$. Let us consider the examples $\mathrm{e}^{i\phi\hat{L}_1/\hbar}\hat{q}_j\mathrm{e}^{-i\phi\hat{L}_1/\hbar}$. Since $[\hat{L}_j, \hat{q}_j] = 0$, we trivially have

$$\mathrm{e}^{i\phi\hat{L}_1/\hbar}\hat{q}_1\mathrm{e}^{-i\phi\hat{L}_1/\hbar} = \hat{q}_1. \tag{9.16}$$

To evaluate $\mathrm{e}^{i\phi\hat{L}_1/\hbar}\hat{q}_2\mathrm{e}^{-i\phi\hat{L}_1/\hbar}$ we need to calculate the nested commutators $[\hat{L}_1, [\hat{L}_1, [..., [\hat{L}_1, \hat{q}_2]...]]]$. This can be done iteratively, starting from $[\hat{L}_1, \hat{q}_2] = i\hbar\hat{q}_3$:

$$\begin{aligned} [\hat{L}_1, [\hat{L}_1, \hat{q}_2]] &= i\hbar[\hat{L}_1, \hat{q}_3] = \hbar^2\hat{q}_2 \\ [\hat{L}_1, [\hat{L}_1, [\hat{L}_1, \hat{q}_2]]] &= \hbar^2[\hat{L}_1, \hat{q}_2] = i\hbar^3\hat{q}_3 \\ [\hat{L}_1, [\hat{L}_1, [\hat{L}_1, [\hat{L}_1, \hat{q}_2]]]] &= i\hbar^3[\hat{L}_1, \hat{q}_3] = \hbar^4\hat{q}_2 \\ &\vdots \end{aligned}$$

Remembering the Taylor expansions of sin and cos:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + ...$$
$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + ...$$

we realise that

$$e^{i\phi\hat{L}_1/\hbar}\hat{q}_2 e^{-i\phi\hat{L}_1/\hbar} = \hat{q}_2\cos(\phi) - \hat{q}_3\sin(\phi).$$

Similarly we find

$$e^{i\phi\hat{L}_1/\hbar}\hat{q}_3 e^{-i\phi\hat{L}_1/\hbar} = \hat{q}_3\cos(\phi) + \hat{q}_2\sin(\phi),$$

and so on. That is, we have

$$\vec{Q}(t) := \langle\psi(0)|e^{i\hat{L}_j t/\hbar}\hat{q}e^{-i\hat{L}_j t/\hbar}|\psi(0)\rangle = R_j(t)\vec{Q}, \tag{9.17}$$

where $R_j(t)$ is the rotation matrix around the axis j by an angle t, and $\vec{Q}$ is the vector of expectation values of the position operators at time zero.

9.2.4 Action of the $e^{i\phi\hat{J}_j/\hbar}$ on the angular momentum operators

Similarly to the considerations above we can use the Hadamard lemma to calculate $e^{i\phi\hat{J}_j/\hbar}\hat{J}_k e^{-i\phi\hat{J}_j/\hbar}$: Trivially we have

$$e^{i\phi\hat{J}_j/\hbar}\hat{J}_j e^{-i\phi\hat{J}_j/\hbar} = \hat{J}_j. \tag{9.18}$$

Let us now consider the example $e^{i\phi\hat{J}_1/\hbar}\hat{J}_2 e^{-i\phi\hat{J}_1/\hbar}$. Again, we need to calculate the nested commutators $[\hat{J}_1,[\hat{J}_1,[...,[\hat{J}_1,\hat{J}_2]...]]]$. Similarly to above, this can be done iteratively, starting from $[\hat{J}_1,\hat{J}_2] = i\hbar\hat{J}_3$:

$$[\hat{J}_1,[\hat{J}_1,\hat{J}_2]] = i\hbar[\hat{J}_1,\hat{J}_3] = -\hbar^2\hat{J}_2$$
$$[\hat{J}_1,[\hat{J}_1,[\hat{J}_1,\hat{J}_2]]] = -\hbar^2[\hat{J}_1,\hat{J}_2] = -i\hbar^3\hat{J}_3$$
$$[\hat{J}_1,[\hat{J}_1,[\hat{J}_1,[\hat{J}_1,\hat{J}_2]]]] = -i\hbar^3[\hat{J}_1,\hat{J}_3] = \hbar^4\hat{J}_2$$
$$\vdots$$

And thus,

$$e^{i\phi\hat{J}_1/\hbar}\hat{J}_2 e^{-i\phi\hat{J}_1/\hbar} = \hat{J}_2\cos(\phi) - \hat{J}_3\sin(\phi).$$

Similarly we find

$$e^{i\phi\hat{J}_1/\hbar}\hat{J}_3 e^{-i\phi\hat{J}_1/\hbar} = \hat{J}_3\cos(\phi) + \hat{J}_2\sin(\phi),$$

and so on. In summary we have

$$e^{i\hat{J}_j t/\hbar}\hat{J}e^{-i\hat{J}_j t/\hbar} = R_j(t)\hat{J}, \tag{9.19}$$

where $R_j(t)$ is the rotation matrix around the axis j by an angle t.

Exercise 14. Prove that $\hat{J}_2$ and $\hat{J}_3$ have the same eigenvalues.

9.3 Spectral properties of angular momentum

We now continue to determine the eigenvalues of quantum angular momentum. Similar to the case of the harmonic oscillator, this can be done by purely algebraic methods. We start out with some considerations about orbital angular momentum, where we shall deduce that the possible eigenvalues of the components of orbital angular momentum are integer multiples of $\hbar$. We shall then move on to find that the commutation relations defining angular momentum would also allow for half-integer values. Half-integer angular momentum does indeed occur in nature, though it cannot be related to orbital angular momentum of particles, as the latter would have to be integer valued. Half-integer valued angular momentum appears as an internal degree of freedom, in the *spin* of quantum particles. We shall briefly discuss this quantum spin in section 9.5.

9.3.1 Orbital angular momentum

Let us consider the possible eigenvalues of $\hat{L}_3 = \hat{q}_1\hat{p}_2 - \hat{q}_2\hat{p}_1$. One can show that these will be given by integer multiples of $\hbar$. This is done by introducing the new set of operators $\hat{x}$ and $\hat{y}$, and $\hat{p}_x$ and $\hat{p}_y$, defined as

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{q}_1 + \hat{p}_2), \qquad \hat{p}_x = \frac{1}{\sqrt{2}}(\hat{p}_1 - \hat{q}_2)$$
$$\hat{y} = \frac{1}{\sqrt{2}}(\hat{q}_1 - \hat{p}_2), \qquad \hat{p}_y = \frac{1}{\sqrt{2}}(\hat{p}_1 + \hat{q}_2),$$

which also fulfil the fundamental commutation relation of position and momenta, and rewriting $\hat{L}_3$ in terms of these operators as

$$\hat{L}_3 = \frac{1}{2}\left(\hat{p}_x^2 + \hat{x}^2\right) - \frac{1}{2}\left(\hat{p}_y^2 + \hat{y}^2\right). \tag{9.20}$$

In this way, we can interpret $\hat{L}_3$ as the difference of two harmonic oscillators $\hat{H}_x$ and $\hat{H}_y$, with masses $m = 1$ and frequencies $\omega = 1$. Remembering that the eigenvalues of the harmonic oscillator $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{q}^2$ are given by $\hbar\omega(n + \frac{1}{2})$, where n are the non-negative integers, we thus deduce that the eigenvalues of $\hat{L}_3$ are given by

$$\lambda_n = \hbar(n_x + \frac{1}{2}) - \hbar(n_y + \frac{1}{2}) = \hbar(n_x - n_y), \tag{9.21}$$

where n_j are non-negative integers. Thus, the eigenvalues of $\hat{L}_3$ are integer multiples of $\hbar$, if $\hat{L}_3$ is associated to orbital angular momentum. The same argument can be used to deduce the eigenvalues of $\hat{L}_1$ (and $\hat{L}_2$), which are also integer multiples of $\hbar$.

9.3.2 Algebraic derivation of the spectral properties

Since the three components of angular momentum do not commute, we cannot measure them simultaneously. The operator $\hat{J}^2$ commutes with all of the components, and thus, it is meaningful to talk of the total angular momentum and the component in one direction. We recall Theorem 7 from Section 5.2, stating that two commuting observables have simultaneous eigenvectors, and consider the simultaneous eigenvectors of $\hat{J}^2$ and one selected component of the angular momentum. In what follows we arbitrarily choose $\hat{J}_3$:

$$\hat{J}^2|\beta, m\rangle = \hbar^2\beta|\beta, m\rangle, \quad \text{and} \quad \hat{J}_3|\beta, m\rangle = \hbar m|\beta, m\rangle, \tag{9.22}$$

and ask the question: What are the possible values of β and m?

Remark 42. Due to the symmetry of the problem, the eigenvalues of $\hat{J}_1$ and $\hat{J}_2$ are the same as those of $\hat{J}_3$.

We will prove the following

Theorem 10. *The possible values of β are given by $j(j+1)$ with $2j \in \mathbb{N}$, and for each value of β m can take the values $-j, -j+1, \ldots, j-1, j$.*

We first prove

Lemma 4. *For a given β we have $m^2 \leq \beta$, that is, m is bounded from above and below.*

Proof. Assume that the states $|\beta, m\rangle$ are normalised to one. Consider the sum of the norms of the states $\hat{J}_j|\beta, m\rangle$, which are positive by definition. That is, we have

$$\sum_j \langle\beta, m|\hat{J}_j|\beta, m\rangle = \langle\beta, m|\hat{J}^2|\beta, m\rangle \geq 0. \tag{9.23}$$

On the other hand we have

$$\langle \beta, m|\hat{J}^2|\beta, m\rangle = \hbar^2\beta, \quad \text{and} \quad \langle \beta, m|\hat{J}_3^2|\beta, m\rangle = \hbar^2 m^2. \tag{9.24}$$

Now we don't know what the $\langle \beta, m|\hat{J}_{1,2}^2|\beta, m\rangle$ are exactly, but we do know that

$$\langle \beta, m|\hat{J}_{1,2}^2|\beta, m\rangle \geq 0, \tag{9.25}$$

and thus

$$\begin{aligned} \sum_j \langle \beta, m|\hat{J}_j|\beta, m\rangle &= \langle \beta, m|\hat{J}^2|\beta, m\rangle \\ \hbar^2\beta &\geq \hbar^2 m^2. \end{aligned} \tag{9.26}$$

□

The remainder of the proof of Theorem 10 proceeds very similarly to the case of the harmonic oscillator treated in section 7.1. We first introduce the set of adjoint operators

$$\hat{J}_\pm = \hat{J}_1 \pm \mathrm{i}\hat{J}_2. \tag{9.27}$$

We will need the commutators

$$[\hat{J}_3, \hat{J}_\pm] = \pm\hbar J_\pm \quad \text{and} \quad [\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_3, \tag{9.28}$$

and the identities

$$\hat{J}_-\hat{J}_+ = \hat{J}^2 - \hat{J}_3^2 - \hbar\hat{J}_3, \quad \text{and} \quad \hat{J}_+\hat{J}_- = \hat{J}^2 - \hat{J}_3^2 + \hbar\hat{J}_3. \tag{9.29}$$

Exercise 15. Verify the commutation relations (9.28) and (9.29) using the defining commutator relations (9.9).

The operators $\hat{J}_\pm$ now play the same role as the $\hat{a}^{(\dagger)}$ for the harmonic oscillator. We have the following two lemmas:

Lemma 5. *If $|\beta, m\rangle$ is an eigenvector of $\hat{J}_3$ with eigenvalue $\hbar m$, then $\hat{J}_+|\beta, m\rangle$ is either the zero vector or also an eigenvector of $\hat{J}_3$ with eigenvalue $\hbar(m+1)$.*

Lemma 6. *If $|\beta, m\rangle$ is an eigenvector of $\hat{J}_3$ with eigenvalue $\hbar m$, then $\hat{J}_-|\beta, m\rangle$ is either the zero vector or also an eigenvector of $\hat{J}_3$ with eigenvalue $\hbar(m-1)$.*

Exercise 16. Proof Lemma 5 and Lemma 6.

Since $m^2 \leq \beta$ there has to be a maximal value $m_{\max} = j$, for which

$$\hat{J}_+|\beta, j\rangle = 0, \tag{9.30}$$

and a minimal value $m_{\min} = k$, for which

$$\hat{J}_-|\beta, k\rangle = 0, \tag{9.31}$$

and m runs in integer values from k to j. Applying $\hat{J}_-$ to equation (9.30), and using that $\hat{J}_-\hat{J}_+ = \hat{J}^2 - \hat{J}_3^2 - \hbar\hat{J}_3$ yields

$$\begin{aligned} \hat{J}^2 - \hat{J}_3^2 - \hbar\hat{J}_3|\beta, j\rangle &= 0 \\ &= \left(\hbar^2\beta - \hbar^2 j^2 - \hbar^2 j\right)|\beta, j\rangle, \end{aligned} \tag{9.32}$$

where $|\beta, j\rangle$ is assumed not to be the zero vector. That is,

$$\beta = j^2 + j = j(j+1). \tag{9.33}$$

On the other hand, applying $\hat{J}_+$ to equation (9.31), and using that $\hat{J}_+\hat{J}_- = \hat{J}^2 - \hat{J}_3^2 + \hbar\hat{J}_3$ yields

$$\begin{aligned}\hat{J}^2 - \hat{J}_3^2 + \hbar\hat{J}_3|\beta, k\rangle &= 0 \\ &= \left(\hbar^2\beta - \hbar^2k^2 + \hbar^2k\right)|\beta, k\rangle,\end{aligned} \tag{9.34}$$

where $|\beta, k\rangle$ is not the zero vector either, and thus

$$\beta = k^2 - k = k(k-1). \tag{9.35}$$

Combining equations (9.33) and (9.35) yields

$$k(k-1) = j(j+1), \tag{9.36}$$

which has the two possible solutions

$$k = j + 1 \quad \text{or} \quad k = -j, \tag{9.37}$$

we know, however, that k is the minimal value of m and j the maximal, that is $k \leq j$, which means we have

$$k = -j. \tag{9.38}$$

Thus we have deduced that m runs in integer steps from $-j$ to j, from which it follows that $2j$ is an integer, and the value of β is related to j by $\beta = j(j+1)$, which completes the proof of Theorem 10.

Remark 43. The eigenstates are usually labelled by the numbers j and m, rather than β and m.

Remark 44. For any given value of j there are $2j + 1$ values of m.

9.4 Matrix representation in the "standard basis"

The so-called *standard basis* of angular momentum is given by the set of states $|j, m\rangle$ fulfilling

$$\begin{aligned}\hat{J}^2|j, m\rangle &= \hbar^2 j(j+1)|j, m\rangle, \\ \hat{J}_3|j, m\rangle &= \hbar m|j, m\rangle,\end{aligned} \tag{9.39}$$

with $2j \in \mathbb{N}$ and m running in integer steps from $-j$ to j. We shall now deduce the representation of the angular momentum operators in the standard basis. By definition the matrices of $\hat{J}^2$ and $\hat{J}_3$ in the standard basis are diagonal. We have

$$\langle j', m'|\hat{J}^2|j, m\rangle = \hbar^2 j(j+1)\delta_{j'j}\delta_{m'm}, \tag{9.40}$$

and

$$\langle j', m'|\hat{J}_3|j, m\rangle = \hbar m\delta_{j'j}\delta_{m'm}, \tag{9.41}$$

We can deduce the matrix elements of $\hat{J}_1$ and $\hat{J}_2$ from those of $\hat{J}_\pm$. We know that $\hat{J}_+|j, m\rangle \propto |j, m+1\rangle$. We can find the proportionality constant from the normalisation condition. We have

$$\begin{aligned}\langle j, m+1|j, m+1\rangle &= \frac{1}{c^2}||\hat{J}_+|j, m\rangle||^2 = \frac{1}{c^2}\langle j, m|\hat{J}_-\hat{J}_+|j, m\rangle \\ &= \frac{1}{c^2}\langle j, m|\hat{J}^2 - \hat{J}_3^2 - \hbar\hat{J}_3|j, m\rangle \\ &= \frac{1}{c^2}\left(\hbar^2 j(j+1) - \hbar^2m^2 - \hbar^2m\right)\langle j, m|j, m\rangle \\ &= \frac{1}{c^2}\left(\hbar^2 j(j+1) - \hbar^2m(m+1)\right)\langle j, m|j, m\rangle.\end{aligned} \tag{9.42}$$

Assuming that $\langle j,m|j,m\rangle = 1$, $|j,m+1\rangle$ is normalised if $c = \hbar\sqrt{j(j+1)-m(m+1)}$, that is,

$$\hat{J}_+|j,m\rangle = \hbar\sqrt{j(j+1)-m(m+1)}|j,m+1\rangle. \tag{9.43}$$

Similarly we find

$$\hat{J}_-|j,m\rangle = \hbar\sqrt{j(j+1)-m(m-1)}|j,m-1\rangle. \tag{9.44}$$

That is we have for the matrix elements of $\hat{J}_+$ and $\hat{J}_-$:

$$\begin{aligned}\langle j',m'|\hat{J}_+|j,m\rangle &= \hbar\sqrt{j(j+1)-m(m+1)}\delta_{j'j}\delta_{m',m+1},\\ \langle j',m'|\hat{J}_+|j,m\rangle &= \hbar\sqrt{j(j+1)-m(m-1)}\delta_{j'j}\delta_{m',m-1}.\end{aligned} \tag{9.45}$$

The matrix elements of $\hat{J}_1$ and $\hat{J}_2$ then follow from

$$\hat{J}_1 = \frac{1}{2}\left(\hat{J}_+ + \hat{J}_-\right), \quad \text{and} \quad \hat{J}_2 = \frac{1}{2\mathrm{i}}\left(\hat{J}_+ - \hat{J}_-\right). \tag{9.46}$$

9.4.1 The 2×2 representation - Pauli matrices again

The angular momentum matrices for the total angular momentum $j = \frac{1}{2}$ are given by

$$\hat{s}_z = \hbar\begin{pmatrix}\frac{1}{2} & 0\\ 0 & -\frac{1}{2}\end{pmatrix}, \quad \hat{s}_x = \hbar\begin{pmatrix}0 & \frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix}, \quad \hat{s}_y = \mathrm{i}\hbar\begin{pmatrix}0 & -\frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix} \tag{9.47}$$

The matrices $\hat{s}_j$ are related to the Pauli matrices $\hat{\sigma}_j$ (9.12) via $\hat{s}_j = \frac{\hbar}{2}\hat{\sigma}_j$.

9.5 Quantum Spin

There is a quantum observable whose three components fulfil the angular momentum commutation relations, which does not describe orbital angular momentum. This observable is the so-called *spin* of a particle. It corresponds to an internal degree of freedom. The values of the spin of a particle in a given direction can be integer or half integer multiples of $\hbar$. Although it does not relate to an actual rotation or "spinning" motion of the particle, it can be transformed into orbital angular momentum. The spin is a fundamental quantity of a quantum particle, similar to for example, its mass. That is, the total value (i.e. the value of $\hat{S}^2$) is a conserved quantity.

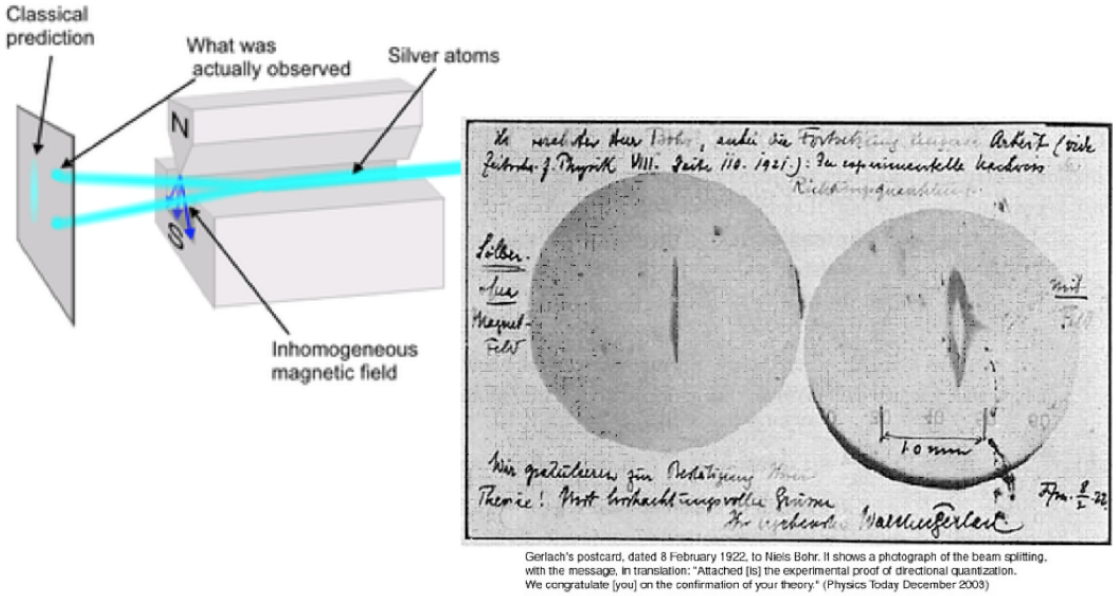
Remark 45. There are two types of particles in nature, those with integer spin, the so-called *bosons*, and those with half-integer spin, the so-called *fermions*, which behave very differently from each other. The details of this difference could easily lead us into many more lectures and thus we will not comment further at this point.

Remark 46. The spin degree of freedom of a spin $\frac{1}{2}$ particle is described by the 2×2 representation of angular momentum, i.e. the Pauli matrices. Together with the identity matrix the Pauli matrices form a basis for all 2×2 matrices. That is, every 2×2 matrix can be written as a superposition of Pauli-matrices, and thus every quantum system with two states can be interpreted as a spin $\frac{1}{2}$ system.

9.5.1 A short history

In 1922 a now famous experiment was performed by Stern and Gerlach, which was meant to demonstrate the quantisation of angular momentum. We now know that this was the first experimental verification of the quantum spin, even though at the time of the experiment that was not clear, and even the notion of quantum spin did not exist at the time. In the experiment, the setup of which is schematically depicted in figure 9.1 on the left, silver atoms were heated in an oven and allowed to escape and fly through an inhomogeneous magnetic field. The interaction between a charged particle with a given angular momentum and an inhomogeneous magnetic field leads to a

force that is given by the so-called *magnetic moment* of the particle multiplied with the gradient of the magnetic field. The magnetic moment is proportional to the charge and the internal angular momentum of the particle. Thus, one expects that after crossing the magnetic field the silver atoms will be displaced from the straight line on which they entered the magnetic field by a distance proportional to their angular momentum. Stern and Gerlach were lucky to have chosen silver for their experiment, as the electrons in a silver atom have total orbital angular momentum of zero, and the total angular momentum of the atom is entirely due to the spin of the *outermost* electron in the atom, which has spin $\frac{1}{2}$ (that is, $j = 1/2$) and thus the beam split into two. Stern and Gerlach were aware they had measured the quantisation of angular momentum, and sent a postcard to Bohr (see figure 9.1), congratulating him on the experimental verification of his quantum theory. They were, however, not aware that they had in fact measured the internal spin degree of the electron, a concept which didn't even exist at the time. Most certainly they had not done our calculation in Section 9.3.1, or otherwise they would have been rather surprised to find the initial beam of silver atoms had split into an even number of beams. It took until 1927 for the Stern-Gerlach experiment to be connected to the electron spin.



Gerlach's postcard, dated 8 February 1922, to Niels Bohr. It shows a photograph of the beam splitting, with the message, in translation: "Attached [is] the experimental proof of directional quantization. We congratulate [you] on the confirmation of your theory." (Physics Today December 2003)

Figure 9.1: Stern-Gerlach experiment, setup, and postcard with the results from Gerlach to Bohr.

Uhlenbeck and Goudsmith proposed the idea of a *spin* of the electron around 1925 in their calculations of atomic spectra. At approximately the same time Pauli used the idea of a double-valued observable and introduced the *Pauli-matrices*, but said this had "nothing to do with Uhlenbeck's and Goudsmith's spin"... Not much later Dirac introduced a relativistic version of quantum theory, compatible with special relativity. In this theory, the spin comes out automatically.

Today the quantum spin is at the basis of many technologies, and gains ever more importance. Spin dynamics is the basic principle behind medical techniques such as MRI, there is an ever increasing interest in so-called *spin-tronics*, and the spin is the key ingredient (*qubit*) in quantum information.