Quantum physics for teachers

Lecture given by Frank Krauss, summer term 2006,
at Dresden university of technology

Disclaimer:

This script is, to a large extent, influenced by a number of excellent books on quantum mechanics. Specifically, the following books have been used:

- J. J. Sakurai: "Modern Quantum Mechanics"
- R. P. Feynman: "The Feynman Lectures III"

Contents

Chapter 1

Concepts

1.1 The Stern-Gerlach experiment

The Stern-Gerlach experiment, designed by O. Stern in 1921 and carried out by him and his coworker W. Gerlach in 1922 is discussed in nearly every textbook on physics. It is one of the most popular examples for an experiment which cannot be understood by means of classical physics concepts. In some sense, two-state systems like those investigated by the Stern-Gerlach experiment are the least classical and the most quantum-mechanical ones. Therefore it will serve as a first example to introduce some of the fundamental concepts of quantum mechanics.

1.1.1 Physics description

Basic experiment

The experiment itself is quite simple to describe. Silver atoms are heated in an oven with a small hole, through which some of the silver atoms can evade. From them, a beam is produced by a collimating slit. This beam is then subjected to an inhomogeneous magnetic field, which is produced by two irregularly formed poles, one, for instance, may be have the form of a sharp edge. A detector is placed behind the magnetic field and analyzes what happens to the atoms.

In order to gain some feeling about the outcome, let us consider a simplistic model for the silver atoms. Although oversimplified, this model suffices the purpose here. Silver atoms are made up of a nucleus and 47 electrons around

it. In a rough approximation, 46 out of the 47 electrons form a spherically symmetric electron cloud with zero net angular momentum. Ignoring nuclear effects, which are irrelevant here, it can be shown that the total angular momentum of the silver atoms is due to the spin of the 47th electron. The first natural question to ask here is: What is spin? The answer is quite straight-forward although it may not be very vivid to imagine. In contrast to orbital angular momentum, stemming, e.g., from objects circling around themselves or each other, it is an intrinsic angular momentum. There is no pictorial interpretation like "angular momentum of the electron caused by it spinning around itself" or similar. Spin is just an intrinsic property of a potentially point-like particle.
The 47 electrons in the silver atom are bound to the nucleus, which is roughly five orders of magnitude heavier than the electrons around it. However, the spin of the 47th electron acts as a magnetic moment and therefore the atom as a whole possesses a magnetic moment equal to the one of the single electron. In other words: The magnetic moment $\vec{\mu}$ of the silver atom is proportional to the spin $\vec{s}$ of the electron, where the factor of proportionality is, up to an accuracy of 0.2 % equal to $e/m_e c$. Here, e is the electric charge of the electron, m_e its mass and c the speed of light. Thus

$$\vec{\mu} = \frac{e}{m_e c} \cdot \vec{s} \,. \tag{1.1}$$

Therefore, the force exerted by the magnetic field on the silver atoms is

$$\vec{F} = \vec{\nabla}(\vec{\mu} \cdot \vec{B}) \,, \tag{1.2}$$

orienting the magnetic field mainly along the z axis, this yields

$$F_z = \frac{\partial(\vec{\mu} \cdot \vec{B})}{\partial z} = \frac{\mu_z \partial B_z}{\partial z} \,, \tag{1.3}$$

where the derivative of B_z w.r.t. z is taken to be positive, $\partial B_z/\partial z > 0$.
Since the atoms are comparably heavy objects, it may be expected that the notion of a classical trajectory is meaningful - later on this assumption will be checked through the Heisenberg uncertainty principle. Anyways, it is obvious that atoms with $\mu_z > 0$ experience a force upwards, whereas for $\mu_z < 0$ atoms will be deflected downwards. The beam is thus expected to split according to the values of μ_z, and the Stern-Gerlach apparatus therefore measures the z-component of the magnetic moment of the atoms, or, equivalently, of the spin of the 47th electron.

The atoms in the oven have no preferred orientation, they are just thermally produced. Therefore, also their magnetic moment has no orientation and therefore it can be expected that its z-component μ_z assumes all values between in the interval between $-|\mu|$ and $|\mu|$]. This would imply that the beam is continuously spread out by the Stern-Gerlach apparatus. However, the actual experiment had a completely different outcome: The beam of incident silver atoms was split into two distinct beams by the Stern-Gerlach experiment - a phenomenon known as "space quantization" in the old days of quantum mechanics. This implies that not a continuous spectrum of values for μ_z and hence for s_z can experimentally be observed but rather two discrete values only, namely

$$s_z = \pm\frac{\hbar}{2}\,, \tag{1.4}$$

where Planck's constant

$$\hbar = 1.0546 \cdot 10^{-27}\ \mathrm{erg\,s} = 6.5822 \cdot 10^{-16}\ \mathrm{eV\,s} \tag{1.5}$$

has been introduced. The two beams have roughly the same intensity, therefore the values of s_z are roughly evenly distributed.
Of course there is nothing special about the z-axis or the up-direction w.r.t. it. Rather than orienting the magnetic field along the z-axis, the Stern-Gerlach apparatus would split the silver beam into two distinct beams along the x-axis if the magnetic field would point along this axis.

Sequential Stern-Gerlach experiments

Consider now different configurations of sequential Stern-Gerlach experiments:

- First consider two sequential Stern-Gerlach apparatus with magnetic fields oriented along the z-axis. Blocking $s_z = -\hbar/2$-component after the first Stern-Gerlach apparatus and subjecting the remaining $s_z = +\hbar/2$-component to the second one yields only one beam in the detector, namely just the $s_z = +\hbar/2$-component. This is perhaps not too surprising; after all, if the atom spins have been prepared such that they all point upwards (w.r.t. the z-axis, of course) one may expect that they remain so as long as there is no field rotating them between the two Stern-Gerlachs.

- Let a Stern-Gerlach apparatus oriented along the z-axis be followed by a second one, oriented along the x-axis. Blocking again the $s_z = -\hbar/2$-component after the first one, the second Stern-Gerlach experiment just splits the remaining $s_z = +\hbar/2$-beam into two distinct beams of equal intensity, which are readily identified with $s_x = \pm\hbar/2$ beams. How can this be explained? Naively, it may be guessed that the $s_z = +\hbar/2$-beam consists of 50% of $s_x = +\hbar/2$-atoms and 50% of $s_x = -\hbar/2$-atoms. In other words: it could be imagined that the beam coming out of the first Stern-Gerlach apparatus is composed of atoms with $s_z = +\hbar/2$ and $s_x = +\hbar/2$ and of atoms with $s_z = +\hbar/2$ and $s_x = -\hbar/2$. This interpretation, although it would make perfect sense in classical physics and match corresponding expectations, however, runs into severe difficulties, as will be discussed now.

- Let us now add a third Stern-Gerlach apparatus, again in z-direction and let us block the $s_x = -\hbar/2$-beam after the second Stern-Gerlach device, such that only the $s_x = +\hbar/2$-beam enters this third Stern-Gerlach station. It is now observed experimentally that this apparatus again splits the $s_x = +\hbar/2$-beam into two beams, characterized again by $s_z = \pm\hbar/2$. This immediately renders the naive guess before as completely wrong: The beam entering the third Stern-Gerlach station was, by this naive (classically motivated) assumption, composed of atoms with $s_z = +\hbar/2$ and $s_x = +\hbar/2$ and should therefore not develop a $s_z = -\hbar/2$-component.

  Hence, this naive idea of the atoms being characterized in such a way is just wrong. In fact, the phenomenon described here is one of the prime examples for the finding that it is impossible to determine the spin components s_z and s_x simultaneously [1].

[1] Often, spin is visualized as a kind of angular momentum, L, and objects with spin are often thought of behaving analogous to a spinning top. The difference of quantum and classical mechanics here is that indeed, there is no classical spin. Spin always is quantum, and, whereas it is straightforward to simultaneously determine the x and z components of the angular momentum, L_x and L_z of a spinning top in classical mechanics, it is impossible to do so for quantum mechanical angular momenta or spins. It must be understood that this impossibility is not due to incapability of the experimenters but a property of quantum mechanical systems.

Electron spin vs. polarization of light

The property described above looks quite novel, and it is worthwhile to consider better known systems which exhibit a similar behavior. A prime example for this is the polarization of light. Consider a wave of monochromatic light travelling in z-direction, that is linearly polarized in x-direction. Its electric field thus oscillates in x-direction,

$$\vec{E} = E_0 \vec{e}_x \cos(kz - \omega t)\,, \tag{1.6}$$

whereas the electric field of y-polarized light, also travelling in z-direction, would read

$$\vec{E} = E_0 \vec{e}_y \cos(kz - \omega t)\,. \tag{1.7}$$

In both case, $\omega = kc \propto 1/\lambda$. Such light can be produced by passing unpolarized light through a polaroid filter. Depending on the allowed direction, such filters will now be called x- and y-filters.
It is well known that a beam of light cannot pass a y-filter after passing through an x-filter, provided they are 100% efficient. But such an experimental setup can be made more interesting by inserting yet another filter between them, oriented along a new x'-direction, in an angle of $\pi/4$ w.r.t. the original x-direction. This time there is some light passing through the y-direction. How does this work? The light has no y-component after passing the x-filter, but only an x-component. Then, the x'-filter intervenes and selects the corresponding x'-polarization and thus destroys the previous x-polarization. The same happens in the y-filter, which erases the x'-polarization. Note that this situation is in full analogy with the Stern-Gerlach arrangement discussed before, provided the following correspondence is made

$$\begin{aligned} S_z - \pm \text{ atoms} \quad &\leftrightarrow \quad x-,\ y - \text{polarized light} \\ S_x - \pm \text{ atoms} \quad &\leftrightarrow \quad x'-,\ y' - \text{polarized light}\,, \end{aligned} \tag{1.8}$$

where the primed axes are rotated by $\pi/4$ w.r.t. the unprimed ones. In classical electrodynamics the electric fields in the primed directions can be expressed by

$$\begin{aligned} E_0 \vec{e}_{x'} \cos(kz - \omega t) &= E_0 \left[\frac{1}{\sqrt{2}} \vec{e}_x \cos(kz - \omega t) + \frac{1}{\sqrt{2}} \vec{e}_y \cos(kz - \omega t) \right] \\ E_0 \vec{e}_{y'} \cos(kz - \omega t) &= E_0 \left[-\frac{1}{\sqrt{2}} \vec{e}_x \cos(kz - \omega t) + \frac{1}{\sqrt{2}} \vec{e}_y \cos(kz - \omega t) \right]. \end{aligned} \tag{1.9}$$

In other words, the electric fields of the x'- and y'-polarized light can be expressed as linear combination of the unprimed polarizations and vice versa. Hence, after passing the x-filter the light is x-polarized and can thus be described by a linear combination of x'- and y'-components. The x'-filter selects the x'-component, which in turn can be understood as a linear combination of x- and y-components. The y-filter finally projects out the y-component. Without the x'-filter in between there is no way to have primed components being projected out.

1.1.2 Formal description

Applying the correspondence above implies that the spin state of the silver atoms may be represented by some kind of a vector in a new kind of two-dimensional vector space [2]. Just as $\vec{e}_x$ and $\vec{e}_y$ are the base vectors spanning the usual xy-space and are thus used to span the space of polarizations of light, it is reasonable to represent the $S_x\pm$ states by base states, vectors which are called **ket**s. This is the first glimpse of a new notation, the Dirac notation which will be fully developed in the next section. However, the kets are denoted e.g. by $|S_x;+\rangle$ and are constructed as linear combination of the base kets $|S_z;\pm\rangle$. The conjecture is

$$\begin{aligned}
|S_x;+\rangle &= +\frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}|S_z;-\rangle \\
|S_x;-\rangle &= -\frac{1}{\sqrt{2}}|S_z;+\rangle + \frac{1}{\sqrt{2}}|S_z;-\rangle\,. \qquad (1.10)
\end{aligned}$$

This is in analogy to the description of polarized light, cf. Eq. (1.9). Hence the unblocked component of the second Stern-Gerlach apparatus in x-direction can be understood as a linear combination of $|S_z;+\rangle$ and $|S_z;-\rangle$, leaving a component to be projected out for the third z-oriented Stern-Gerlach filter. The next question coming immediately to mind is how to describe the y-spin states. For symmetry reasons the effect of a Stern-Gerlach zyz-sequence should be identical to the zxz-sequence discussed so far. Therefore, also the $|S_y;\pm\rangle$ kets should be written as linear combinations of $S_z;\pm\rangle$, and of course the $S_x\pm$, $S_y\pm$, and the $S_z\pm$ states should be all linearly independent. However, naively it seems as if all options of linear combinations are used up.

[2] This, of course, is a new, abstract vector space, responsible for spins and should not be confused with the usual two-dimensional xy-space.

Again, an analogy with polarized light comes to rescue. This time, consider circularly polarized light, which can be obtained by letting linearly polarized light pass through a quarter-wave plate. Passing a beam of circularly polarized light through an x- or a y-polarimeter the same amount of light passes as in the case of primed polarization passing through an unprimed filter. Yet, it is obvious that circularly polarized light is different from linearly polarized light. But yet, mathematically, how can such light be represented? The answer is the following: circularly polarized light is nothing but a linear combination of x- and y-polarized light where the oscillation of the y-electric field is $\pi/2$ out of phase with the x-component. Hence,

$$\begin{aligned}\vec{E}_r &= E_0\left[\frac{1}{\sqrt{2}}\vec{e}_x\cos\Big(kz-\omega t\Big)+\frac{1}{\sqrt{2}}\vec{e}_y\cos\Big(kz-\omega t+\frac{\pi}{2}\Big)\right]\\ \vec{E}_l &= E_0\left[-\frac{1}{\sqrt{2}}\vec{e}_x\cos\Big(kz-\omega t\Big)+\frac{1}{\sqrt{2}}\vec{e}_y\cos\Big(kz-\omega t+\frac{\pi}{2}\Big)\right].\end{aligned} \tag{1.11}$$

It is more elegant to use a complex notation by introducing a complex vector $\vec{\epsilon}$ such that its real part $\mathcal{R}(\vec{\epsilon})$ is given by

$$\mathcal{R}(\vec{\epsilon}) = \vec{E}\,. \tag{1.12}$$

Then, using $i = \exp(i\pi/2)$,

$$\vec{\epsilon}_r = \left[\frac{1}{\sqrt{2}}\vec{e}_x\,e^{i(kz-\omega t)}+\frac{i}{\sqrt{2}}\vec{e}_y\,e^{i(kz-\omega t)}\right]. \tag{1.13}$$

Extending the analogy of polarizations of light and spins and accepting complex coefficients for the kets yields

$$|S_y;\pm\rangle = \frac{1}{\sqrt{2}}|S_z;+\rangle \pm \frac{i}{\sqrt{2}}|S_z;-\rangle\,. \tag{1.14}$$

This is in fact the only orthogonal combination of base kets for this different spin state.
Thus, in order to describe the electron spin by the two-component formalism of kets, complex prefactors are mandatory. It is remarkable that imaginary coefficients already appear in such a basic example.

1.2 Constructing the Hilbert space

The formal description of the Stern-Gerlach experiment which was initiated in the previous section is further formalized. The vector space being spanned by the kets is constructed. This theory of linear vector spaces used in quantum mechanics of course has been known to mathematicians well before P. A. M. Dirac introduced his ket notation, but Dirac's way of introducing linear vector spaces has many advantages, especially from a physicist's point of view. This theory of a linear vector space is then generalized to the case of a continuous space, i.e. a vector space with a continuous spectrum of base vectors rather than a finite number like the two-state system considered so far. In such a case, for instance for position or momentum, where the number of alternatives is uncountably infinite, the vector space in question is known as **Hilbert space** after D. Hilbert, who studied vector spaces in infinite dimensions. In this section, it will be discussed how the kets span the vector or Hilbert space of physical states. Furthermore, operators acting on these kets will be introduced, scalar products of kets and expectation values of operators will be introduced and finally, the mathematical meaning of measurements will be investigated.

1.2.1 Kets, bras, and operators

Kets and their properties

In quantum mechanics, a physical state, for instance a silver atom with a definite spin orientation is represented by a state vector in a complex vector space. Following Dirac's notation, these vectors are called **ket**s and denoted by $|\alpha\rangle$. The **postulate** is that the state kets carry the complete information about the state - every measurable property of the state is in the ket. Kets can be added, yielding a new one,

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle\,. \tag{1.15}$$

In addition, they can be multiplied with a complex number c from the left or the right, yielding the same results

$$c|\alpha\rangle = |\alpha\rangle c\,. \tag{1.16}$$

In the case, where $c = 0$, the result is the **null ket**. One of the physics postulates of quantum mechanics is that **both** $|\alpha\rangle$ and $c|\alpha\rangle$ represent the same physical state as long as $c \neq 0$. In mathematical language this implies that physicists deal with rays rather than vectors.

Kets, operators and observables

A physical **observable**, such as momentum or spin components or the like can be represented by an **operator**, $\hat{A}$, acting in the vector space in question[3]. In general, operators act on kets from the left,

$$\hat{A} \cdot |\alpha\rangle = \hat{A}|\alpha\rangle = |\beta\rangle \,, \tag{1.17}$$

which results in just another ket. In most cases, the kets $|\alpha\rangle$ and $\hat{A}|\alpha\rangle$ are not proportional. However, there is a class of kets, called **eigenkets** of the operator with the property that

$$\hat{A}|a'\rangle = a'|a'\rangle \;, \quad \hat{A}|a''\rangle = a''|a''\rangle \; etc. \,, \tag{1.18}$$

where the **numbers** a', a'' etc. are eigenvalues of the operator $\hat{A}$.
The physical states connected to the eigenkets are called the **eigenstates** w.r.t. the operator. In the simplest case of a spin-1/2 system, the eigenvalue-eigenket relation reads

$$\hat{S}_z|S_z;\pm\rangle = \pm\frac{\hbar}{2}|S_z;\pm\rangle \,. \tag{1.19}$$

Obviously, the kets $|S_z;\pm\rangle$ are eigenkets of the operator $\hat{S}_z$ and have eigenvalue $\pm\hbar/2$. Similar relations exist also for the operators $\hat{S}_x$ and $\hat{S}_y$; for them, however the eigenkets are $|S_x;\pm\rangle$ and $|S_y;\pm\rangle$, respectively, rather than $|S_z;\pm\rangle$.
Implicitly it has already been stated that the dimensionality of the vector space spanned by the kets is related to the number of alternatives in, e.g. Stern-Gerlach-type experiments. More formally, discussing an N-dimensional vector space spanned by the N eigenkets $|a^{(i=1,\ldots,N)}\rangle$ of $\hat{A}$ it is clear that any ket $|\alpha\rangle$ in this space can be decomposed as linear combination of the eigenkets,

$$|\alpha\rangle = \sum_{i=1}^{N} c_i |a^{(i)}\rangle \tag{1.20}$$

[3]Pictorially, such an operator is nothing but a matrix acting on the vectors of the vector space under consideration.

with complex coefficients c_i. The question of uniqueness of this decomposition will be postponed until the orthogonality of the eigenkets has been proven.

Bras and inner products

So far, only the **ket space** has been dealt with. It is now supplemented with the **bra space**, a vector space which is **dual** to the original one. The postulate is that for every vector in ket space, i.e. for every ket $|\alpha\rangle$, there exists a **bra** $\langle\alpha|$ in **bra space**. This space, in turn, is spanned by some **eigenbras**, which correspond to the **eigenkets**. There is an one-to-one correspondence,

$$\begin{aligned} |\alpha\rangle &\longleftrightarrow \langle\alpha| \\ |a^{(i)}\rangle &\longleftrightarrow \langle a^{(i)}| \\ |\alpha\rangle + |\beta\rangle &\longleftrightarrow \langle\alpha| + \langle\beta| \end{aligned} \tag{1.21}$$

etc.. Roughly speaking, one can imagine bra space to be some kind of mirror of ket space. However, there is one subtlety, namely that the bra dual to $c|\alpha\rangle$ is not $c\langle\alpha|$ but rather

$$c|\alpha\rangle \longleftrightarrow c^*\langle\alpha| \tag{1.22}$$

and therefore

$$c_a|\alpha\rangle + c_b|\beta\rangle \longleftrightarrow c_a^*\langle\alpha| + c_b^*\langle\beta|\,. \tag{1.23}$$

Now an **inner** or **scalar product** of bras and kets is **defined**,

$$\langle\beta|\alpha\rangle = \text{complex number}\,. \tag{1.24}$$

Here, the bra plays the role of a row-vector and the kets can be understood as a column-vector. From this it is clear that the **sequence matters**, having a ket to the left of a bra does not yield a number. Such products are called **dyadic** or **outer products** and will be discussed later. There are two **postulates** for the inner product:

- Behavior under changing the sequence:

$$\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*\,. \tag{1.25}$$

  This implies that the inner product of $\langle\alpha|\alpha\rangle$ is a real number. To prove this property, just let $|\beta\rangle \to |\alpha\rangle$ in Eq. (1.25).

- Positive definiteness of $\langle\alpha|\alpha\rangle$:

$$\langle\alpha|\alpha\rangle \geq 0\,, \tag{1.26}$$

where the equality is for the null ket only. This postulate, from the physicists point of view, is essential for the probabilistic interpretation of quantum mechanics.

The inner product can be used to normalize kets through

$$|\alpha_{\rm norm}\rangle = \frac{1}{\sqrt{\langle\alpha|\alpha\rangle}} \cdot |\alpha\rangle\,, \tag{1.27}$$

where, for obvious reasons, $\sqrt{\langle\alpha|\alpha\rangle}$ is often called the **norm** of the ket $|\alpha\rangle$.

More on operators: Basic properties

There are a number of basic properties related to operators:

- **Identity:**
  Two operators $\hat{X}$ and $\hat{Y}$ are identical, if for any arbitrary ket $|\alpha\rangle$ in the ket space on which they act,

$$\hat{X}|\alpha\rangle = \hat{Y}|\alpha\rangle\,. \tag{1.28}$$

- **Null operators:**
  Operators $\hat{X}$, for which

$$\hat{X}|\alpha\rangle = 0 \tag{1.29}$$

  for any arbitrary ket $|\alpha\rangle$ are called null operators.

- **Addition:**
  Operators can be added; they are **commutative and associative under addition**,

$$\begin{aligned} \hat{X}+\hat{Y} &= \hat{Y}+\hat{X} \\ \hat{X}+(\hat{Y}+\hat{Z}) &= (\hat{X}+\hat{Y})+\hat{Z}\,. \end{aligned} \tag{1.30}$$

- **Acting on bras:**
  Operators always act on bras from the right side, $\langle\alpha|\hat{X}$; in general $\langle\alpha|\hat{X}$ and $\hat{X}|\alpha\rangle$ are not dual to each other. Instead,
  $$\hat{X}|\alpha\rangle \quad\longleftrightarrow\quad \langle\alpha|\hat{X}^\dagger\,, \tag{1.31}$$
  where the **Hermitian adjoint** $\hat{X}^\dagger$ of the operator $\hat{X}$ has been defined. Operators $\hat{X}$ are Hermitian, if
  $$\hat{X} = \hat{X}^\dagger\,. \tag{1.32}$$

- **Multiplication:**
  Multiplication of operators, is associative,
  $$\hat{X}\cdot(\hat{Y}\cdot\hat{Z}) = (\hat{X}\cdot\hat{Y})\cdot\hat{Z} = \hat{X}\cdot\hat{Y}\cdot\hat{Z}\,, \tag{1.33}$$
  but, in general, not commutative,
  $$\hat{X}\cdot\hat{Y} \neq \hat{Y}\cdot\hat{X}\,. \tag{1.34}$$
  Note also that
  $$(\hat{X}\cdot\hat{Y})^\dagger = \hat{Y}^\dagger\cdot\hat{X}^\dagger\,, \tag{1.35}$$
  which, employing the dual relation between bra and ket space, can be seen from
  $$(\hat{X}\hat{Y})|\alpha\rangle = \hat{X}\cdot(\hat{Y}|\alpha\rangle) \quad\longleftrightarrow\quad (\langle\alpha|\hat{Y}^\dagger)\cdot\hat{X}^\dagger = \langle\alpha|(\hat{Y}^\dagger\hat{X}^\dagger)\,. \tag{1.36}$$

- **Outer products:**
  From a bra $\langle\beta|$ and a ket $|\alpha\rangle$, it is also possible to form an outer product, $|\alpha\rangle\langle\beta|$, which can be regarded as an operator, acting on kets from the left and on bras from the right.

- **Associativity in products:**
  Consider
  $$(|\beta\rangle\langle\alpha|)\cdot|\gamma\rangle = |\beta\rangle(\langle\alpha|\gamma\rangle)\,, \tag{1.37}$$
  which translates into the outer product $|\beta\rangle\langle\alpha|$ rotating the ket $|\gamma\rangle$ onto the ket $|\beta\rangle$ times a number, namely $\langle\alpha|\gamma\rangle$. It is quite simple to see that if
  $$\hat{X} = |\beta\rangle\langle\alpha| \tag{1.38}$$

then

$$\hat{X}^\dagger = |\alpha\rangle\langle\beta| \,. \tag{1.39}$$

In addition,

$$\langle\beta| \cdot (\hat{X}|\alpha\rangle) = (\langle\beta|\hat{X}) \cdot |\alpha\rangle = \langle\beta|\hat{X}|\alpha\rangle \,. \tag{1.40}$$

Recalling the dual properties, it is straightforward to see that

$$\langle\beta|\hat{X}|\alpha\rangle = \langle\alpha|\hat{X}^\dagger|\beta\rangle^* \,. \tag{1.41}$$

Eigenkets and base kets

Consider the eigenkets of an Hermitian operator $\hat{A} = \hat{A}^\dagger$. They have the following properties:

- the **eigenvalues of Hermitian operators are real**;
- the **eigenkets of a Hermitian operators corresponding to** *different* **eigenvalues are orthogonal**.

Let's prove this. Recall the eigenvalue equation

$$\hat{A}|a^{(i)}\rangle = a^{(i)}|a^{(i)}\rangle \tag{1.42}$$

and its dual

$$\langle a^{(j)}|\hat{A} = \langle a^{(j)}|a^{(j)*} \,, \tag{1.43}$$

where the $|a^{(i)}\rangle$ are the eigenkets and the $a^{(i)}$ are the corresponding eigenvalues. Multiplying Eq. (1.42) from the left with $\langle a^{(j)}|$ and multiplying Eq. (1.43) from the right with $|a^{(i)}\rangle$ and subtracting both yields

$$\begin{aligned} \langle a^{(j)}|\hat{A}|a^{(i)}\rangle - \langle a^{(j)}|\hat{A}|a^{(i)}\rangle &= \langle a^{(j)}|a^{(i)}|a^{(i)}\rangle - \langle a^{(j)}|a^{(j)*}|a^{(i)}\rangle \\ 0 &= (a^{(i)} - a^{(j)*}) \cdot \langle a^{(j)}|a^{(i)}\rangle \,. \end{aligned} \tag{1.44}$$

There are two cases to consider:

- $i = j$
  Then

$$\langle a^{(j)}|a^{(i)}\rangle = \langle a^{(i)}|a^{(i)}\rangle \neq 0 \tag{1.45}$$

  and, in order to fulfil Eq. (1.44),

$$a^{(i)} = a^{(i)*}\,. \tag{1.46}$$

  Hence, the eigenvalues are real.

- $i \neq j$
  Then, because of $a^{(i)} \neq a^{(j)*}$, the inner product must vanish,

$$\langle a^{(j)}|a^{(i)}\rangle = 0\,. \tag{1.47}$$

  Hence the eigenkets are orthogonal (this is exactly the definition!).

In general, physical observables will be represented by Hermitian operators and thus have real eigenvalues, ultimately leading to real expectation values.

The completeness relation

Using the properties above, it is straightforward to normalize the eigenkets of such an operator such that they form a complete orthonormal base, i.e.

$$\langle a^{(j)}|a^{(i)}\rangle = \delta_{ij}\,. \tag{1.48}$$

Having at hand such a base, any other ket in the ket space under consideration may be expanded as a linear combination of these eigenkets,

$$|\alpha\rangle = \sum_i c_i|a^{(i)}\rangle\,. \tag{1.49}$$

The coefficients $c_i \in \mathbf{C}$ can be obtained using the orthonormality of the base kets: multiplying the equation above from the left with $\langle a^{(j)}|$ yields

$$\langle a^{(j)}|\alpha\rangle = \sum_i c_i\langle a^{(j)}|a^{(i)}\rangle = \sum_i c_i\delta_{ij} = c_j\,. \tag{1.50}$$

In other words,

$$|\alpha\rangle = \sum_i |a^{(i)}\rangle\langle a^{(i)}|\alpha\rangle \tag{1.51}$$

or

$$\mathbf{1} = \sum_i |a^{(i)}\rangle\langle a^{(i)}| \,. \tag{1.52}$$

This last equation is called the **completeness relation** or **closure**. Each term $|a^{(i)}\rangle\langle a^{(i)}|$ in the sum will project out the component parallel to $|a^{(i)}\rangle$ of any arbitrary ket (or similar of any bra). Hence the name **projector** for such a term,

$$\Lambda_i = |a^{(i)}\rangle\langle a^{(i)}| \,, \tag{1.53}$$

where

$$\mathbf{1} = \sum_i \Lambda_i \,. \tag{1.54}$$

To gain an improved feeling for the usefulness of the completeness relation, consider $\langle\alpha|\alpha\rangle$. Inserting the $\mathbf{1}$ in this inner product will not change its value, hence

$$\langle\alpha|\alpha\rangle = \left\langle \alpha \left| \left[\sum_i |a^{(i)}\rangle\langle a^{(i)}| \right] \right| \alpha \right\rangle = \sum_i \left| \langle\alpha|a^{(i)}\rangle \right|^2 \,. \tag{1.55}$$

If the ket $|\alpha\rangle$ is also normalized, this implies that

$$\sum_i |c_i|^2 = \sum_i \left| \langle\alpha|a^{(i)}\rangle \right|^2 = 1 \,. \tag{1.56}$$

Representations

Now that base kets have been specified, it is time to use them for the **representation of operators through square matrices**. To see how this works, consider

$$\hat{X} = \sum_{i,j} |\alpha^{(i)}\rangle\langle\alpha^{(i)}|\hat{X}|\alpha^{(j)}\rangle\langle\alpha^{(j)}| \,. \tag{1.57}$$

If the ket space in question (the one, one which $\hat{X}$ acts) is spanned by N eigenkets $|\alpha^{(i=1,\ldots,N)}\rangle$, then the operator sandwich (the form $\langle\alpha^{(i)}|\hat{X}|\alpha^{(j)}\rangle$)

is described by N^2 numbers. Of course they can be arranged in form of a matrix, such that rows and columns appear as follows:

$$\underset{\text{row}}{\langle\alpha^{(i)}|}\quad \hat{X} \quad \underset{\text{column}}{|\alpha^{(j)}\rangle}$$

indicating that bra's can be understood as row- and kets as column-vectors. Explicitly, the matrix can thus be written as

$$\hat{X} \doteq \begin{pmatrix} \langle\alpha^{(1)}|\hat{X}|\alpha^{(1)}\rangle & \langle\alpha^{(1)}|\hat{X}|\alpha^{(2)}\rangle & \dots \\ \langle\alpha^{(2)}|\hat{X}|\alpha^{(1)}\rangle & \langle\alpha^{(2)}|\hat{X}|\alpha^{(2)}\rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{1.58}$$

where $\doteq$ stands for "is represented by". It is important to stress that this is not the "="-sign, because the operator should be the same, independent of the specific choice of the basis in which it is represented. This is in full analogy with the ket space being the same independent of the choice of base kets spanning it.

The entries in the matrix above, the $\langle\alpha^{(i)}|\hat{X}|\alpha^{(j)}\rangle$, are conveniently called matrix elements, X_{ij} in the "α"-representation. Using Eq. (1.41), it is easy to see that the matrix elements satisfy

$$\langle\alpha^{(i)}|\hat{X}|\alpha^{(j)}\rangle = \langle\alpha^{(j)}|\hat{X}^\dagger|\alpha^{(i)}\rangle^* \tag{1.59}$$

relating the Hermitian adjoint conjugation (indicated by the †) to the transposition of the complex conjugate. Phrased in other words: Employing the matrix representation of an operator $\hat{X}$, Eq. (1.58), it can be seen that the Hermitian conjugate of it is given by the transposed matrix, where the entries are complex conjugated. This implies that, if an operator $\hat{B}$ is Hermitian, i.e. if $\hat{B}^\dagger = \hat{B}$, then

$$\langle\alpha^{(i)}|\hat{B}|\alpha^{(j)}\rangle = \langle\alpha^{(j)}|\hat{B}|\alpha^{(i)}\rangle^* . \tag{1.60}$$

It is straightforward to check that this way of representing the operators is in accordance with operations like multiplication of operators. To see this, check this out

$$\langle\alpha^{(i)}|\hat{X}\hat{Y}|\alpha^{(j)}\rangle = \sum_{k=1,\dots N} \langle\alpha^{(i)}|\hat{X}|\alpha^{(k)}\rangle\langle\alpha^{(k)}|\hat{Y}|\alpha^{(j)}\rangle\,, \tag{1.61}$$

a good example for the power of inserting factors of one.
Of course, matrix representations of operators become particularly simple, if their eigenkets are used as base kets. If the base kets $|a^{(i)}\rangle$ are eigenkets of the operator $\hat{A}$, with eigenvalue $a^{(i)}$, then

$$\begin{aligned}
\hat{A} &= \sum_{i,j=1\dots N} |a^{(i)}\rangle\langle a^{(i)}|\hat{A}|a^{(j)}\rangle\langle a^{(j)}| \\
&= \sum_{i,j=1\dots N} |a^{(i)}\rangle \left(a^{(j)}\langle a^{(i)}|a^{(j)}\rangle\right)\langle a^{(j)}| = \sum_{i,j=1\dots N} |a^{(i)}\rangle \left(a^{(j)}\delta_{ij}\right)\langle a^{(j)}| \\
&= \sum_{i=1\dots N} a^{(i)}|a^{(i)}\rangle\langle a^{(i)}|\,, \qquad (1.62)
\end{aligned}$$

a diagonal matrix with the eigenvalues as entries. Thus, **using eigenkets as base kets for representing operators results in diagonalising them**.

Similar reasoning also applies to kets (and bras). Starting from the relation

$$|\gamma\rangle = \hat{X}|\alpha\rangle\,, \qquad (1.63)$$

multiplying from the left with $\langle\alpha^{(i)}|$ and inserting a $\mathbf{1}$ yields

$$\langle\alpha^{(i)}|\gamma\rangle = \sum_{j=1\dots N} \langle\alpha^{(i)}|\hat{X}|\alpha^{(j)}\rangle\langle\alpha^{(j)}|\alpha\rangle\,. \qquad (1.64)$$

Hence, the two kets $|\gamma\rangle$ and $|\alpha\rangle$ can be represented through

$$|\gamma\rangle \doteq \begin{pmatrix} \langle\alpha^{(1)}|\gamma\rangle \\ \langle\alpha^{(2)}|\gamma\rangle \\ \vdots \end{pmatrix}\,, \quad |\alpha\rangle \doteq \begin{pmatrix} \langle\alpha^{(1)}|\alpha\rangle \\ \langle\alpha^{(2)}|\alpha\rangle \\ \vdots \end{pmatrix}\,. \qquad (1.65)$$

Of course, the same holds true for bras, which are thus represented through

$$\begin{aligned}
\langle\gamma| &\doteq \left(\langle\gamma|\alpha^{(1)}\rangle\,,\ \langle\gamma|\alpha^{(2)}\rangle\,,\ \dots\right) \\
&= \left(\langle\alpha^{(1)}|\gamma\rangle^*\,,\ \langle\alpha^{(2)}|\gamma\rangle^*\,,\ \dots\right)\,. \qquad (1.66)
\end{aligned}$$

Taken together, it can now be understood how the inner product of a bra and a ket can be interpreted as the product of a row and a columns vector,

$$\langle\beta|\gamma\rangle = \sum_{i=1\dots N} \langle\beta|\alpha^{(i)}\rangle\langle\alpha^{(i)}|\gamma\rangle = \left(\langle\alpha^{(1)}|\beta\rangle^*\,,\ \langle\alpha^{(2)}|\beta\rangle^*\,,\ \dots\right)\begin{pmatrix} \langle\alpha^{(1)}|\gamma\rangle \\ \langle\alpha^{(2)}|\gamma\rangle \\ \vdots \end{pmatrix}\,. \qquad (1.67)$$

Outer products can be written as

$$|\gamma\rangle\langle\beta| \doteq \begin{pmatrix} \langle\alpha^{(1)}|\gamma\rangle\langle\alpha^{(1)}|\beta\rangle^* & \langle\alpha^{(1)}|\gamma\rangle\langle\alpha^{(2)}|\beta\rangle^* & \dots \\ \langle\alpha^{(2)}|\gamma\rangle\langle\alpha^{(1)}|\beta\rangle^* & \langle\alpha^{(2)}|\gamma\rangle\langle\alpha^{(2)}|\beta\rangle^* & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} . \tag{1.68}$$

1.2.2 Measurements

The quantum mechanical measurement process

With the formalism of kets at hand, the mathematics of the measurement process can be discussed. Referring to P. A. M. Dirac, "*a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured.*"
How can this be understood? Before a measurement of the observable A, represented by some operator $\hat{A}$ is being made, the system is described by a state vector, the ket $|\alpha\rangle$, which is some linear combination of base kets, $|a^{(i)}\rangle$:

$$|\alpha\rangle = \sum_{i=1,\dots N} c_i|a^{(i)}\rangle . \tag{1.69}$$

These base kets can be conveniently chosen to be the eigenkets of $\hat{A}$. Performing the measurement now implies that the system is "thrown" into one of these eigenstates, say $|a^{(k)}\rangle$, or

$$|\alpha\rangle \stackrel{\text{measurement}}{\longrightarrow} |a^{(k)}\rangle . \tag{1.70}$$

Thus, in most cases, a measurement changes the state of the system, from a linear combination of eigenstates to an eigenstate of the observable that is being measured.
Quantum mechanics **postulates**

1. that it is impossible to know before, into which state a system jumps during the measurement process;

2. that, however, the probability for the system to jump from its state $|\alpha\rangle$ into a state $|a^{(i)}\rangle$ is given by

$$\mathcal{P}_{|\alpha\rangle\to|a^{(i)}\rangle} = \left|\langle a^{(i)}|\alpha\rangle\right|^2 = |c_i|^2 , \tag{1.71}$$

provided both state vectors are normalized, i.e.

$$|\langle\alpha|\alpha\rangle|^2 = \left|\langle a^{(i)}|a^{(i)}\rangle\right|^2 = 1 \tag{1.72}$$

and

$$1 = |\langle\alpha|\alpha\rangle|^2 = \sum_i |c_i|^2 \left|\langle a^{(i)}|a^{(i)}\rangle\right|^2 = \sum_i |c_i|^2 . \tag{1.73}$$

To exemplify this, let's go back to the example of the silver atoms passing through a Stern-Gerlach apparatus. Orienting the Stern-Gerlach apparatus in z-direction implies that the two relevant eigenkets are

$$|S_z;\ +\rangle \equiv |+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |S_z;\ -\rangle \equiv |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{1.74}$$

The Stern-Gerlach apparatus can be represented by the operator $\hat{S}_z$. Employing the same reasoning as before, cf. Eq. (1.58),

$$\hat{S}_z \doteq \sum_{i,j=\pm} \langle i|\hat{S}_z|j\rangle = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.75}$$

Now, each silver atom emitted from the oven and passing through the first Stern-Gerlach filter is in some state $|\alpha\rangle$, a linear combination of the two eigenstates[4]. In principle this state is arbitrary - the silver atoms have been evaporated from the oven through heating - and can thus be represented as a linear combination $|\alpha\rangle$ of the two state kets $|\pm\rangle$,

$$|\alpha\rangle = c_+|+\rangle + c_-|-\rangle . \tag{1.76}$$

Arbitrary, random orientation now implies that $c_+ = c_-$, and normalizing $|\alpha\rangle$ translates into $c_+ = c_- = 1/\sqrt{2}$. Hence the probability for the silver atoms to jump, e.g., into the $+$-state, is given by

$$\mathcal{P}_+ = |\langle+|\alpha\rangle|^2 = |\langle+|\left(c_+|+\rangle + c_-|-\rangle\right)|^2 = |c_+\langle+|+\rangle|^2 = c_+^2 = \frac{1}{2} . \tag{1.77}$$

[4]This, again, is in full analogy to the case of (unpolarized) light, which can be understood as being in a linear combination of two different polarization states. When it goes through a filter, only the corresponding polarization component passes and the other one is being blocked.

The postulates above apply for a single, individual system and, thus, a single, individual measurement process. Since this process involves probabilities, a large number of individual independent measurements on a large ensemble of systems, all described by the same state vector $|\alpha\rangle$ [5] must be performed in order to gain some useful and meaningful information - after all talking about probabilities translates into a statistical process.
It is important to convince oneself here that this makes perfect sense. To this end, consider a few extreme cases:

- Consider a case, where the individual system is already in an eigenstate of the observable, for instance a silver atom in the $|S_z;+\rangle$-state before being analyzed by a Stern-Gerlach apparatus $\hat{S}_z$. Then according to the reasoning above, the probability for being thrown into the same eigenstate ($|S_z;+\rangle$) or for measuring the respective eigenvalue ($\hbar/2$) is unity, just as expected. By repeating the measurement immediately after it has been performed the first time, the same ket will result, of course with unit probability: This is exactly the case encountered for the configuration with two Stern-Gerlach filters pointing into the same direction; there the $|S_z;+\rangle$ beam emerging from the first filter was not split any further by the second one.

- If, on the other hand, the probability for the system jumping into another eigenstate, orthogonal to the original one, is being measured, the result will be 0. This is exemplified by the same example as above, two successive z-oriented Stern-Gerlach filters. There, the probability for silver atoms to jump into the $|S_z;-\rangle$-state in the second filter, after they have been measured to be in the $|S_z;+\rangle$-state by the first filter is 0. From the point of view of quantum mechanics these two states are just mutually exclusive: Systems in one state cannot be in another, orthogonal state at the same time.

- In general, probabilities must be non-negative, and the probabilities of all alternatives must add up to 1. This is given by the definitions above. The definition through an absolute value squared guarantees the probabilities to be non-negative, the demand of normalized states and an orthonormal base of eigenkets ensures that the summed probability equals 1.

[5] Such an ensemble, where all constituents are characterized by the same state ket is known as a **pure ensemble**.

Measurement as expectation values of operators

The **expectation value** of an operator $\hat{A}$ w.r.t. a state $|\alpha\rangle$ is **defined** as

$$\langle A\rangle_\alpha = \langle\alpha|\hat{A}|\alpha\rangle\,. \tag{1.78}$$

As said, this is a definition. It however agrees very well with an intuitive notion of **on-average measured value**, since

$$\langle A\rangle_\alpha = \sum_{i,j}\langle\alpha|a^{(i)}\rangle\langle a^{(i)}|\hat{A}|a^{(j)}\rangle\langle a^{(j)}|\alpha\rangle = \sum_i a^{(i)}\left|\langle a^{(i)}|\alpha\rangle\right|^2\,, \tag{1.79}$$

where, for obvious reasons, the representation based on the eigenkets of $\hat{A}$ has been chosen.
At this point it is very important to stress that **eigenvalues are not expectation values**. For example, the expectation value of the operator $\hat{S}_z$ can, in principle, assume any real value between $\hbar/2$ and $-\hbar/2$ in dependence on the state $|\alpha\rangle$, for instance $0.132\hbar$ whereas the eigenvalues are exactly $\pm\hbar/2$.
To further clarify the concept of a measurement, filters are considered in a more formal language. In the prime example of this chapter, the Stern-Gerlach experiment, only one spin component is allowed to pass individual Stern-Gerlach apparatus whereas the other one is being blocked. In more general terms this is just a device that selects only one of the eigenkets of its corresponding operator and rejects all others. This is exactly what is meant by a **selective measurement** or **filtration**. Such a measurement boils down to applying a projection operator on the state ket of the system under consideration,

$$\Lambda_{a^{(i)}}|\alpha\rangle = |a^{(i)}\rangle\langle a^{(i)}|\alpha\rangle\,. \tag{1.80}$$

Reconstructing kets and operators of the spin-$\frac{1}{2}$ system

In this part, the Stern-Gerlach experiments from the beginning in their various set-ups will be confronted with the formalism developed so far. In order to do so, it is important to first reconstruct the eigenkets $|S_{x,y};\pm\rangle$ and the operators $\hat{S}_{x,y}$ corresponding to the Stern-Gerlach devices in the various directions.
The starting point of this is the case, when the $s_x = +$-beam is subjected to a Stern-Gerlach device in z-direction. Then, the beam splits into two

components with equal intensity. From

$$|\langle +|S_x;\, +\rangle|^2 = |\langle -|S_x;\, +\rangle|^2 = \frac{1}{2} \tag{1.81}$$

it follows that

$$|S_x;+\rangle = \frac{1}{\sqrt{2}}\left[|+\rangle + e^{i\delta_x}|-\rangle\right] , \tag{1.82}$$

where δ_x is real and where an overall phase can be ignored. However, the phase factor between the $|+\rangle$- and the $|-\rangle$-component guarantees generality of the results and is mandatory for the discussion. In a similar fashion, and because of

$$\langle S_x;+|S_x;-\rangle = 0 \tag{1.83}$$

one also finds

$$|S_x;+\rangle = \frac{1}{\sqrt{2}}\left[|+\rangle - e^{i\delta_x}|-\rangle\right] . \tag{1.84}$$

Simple calculation shows that indeed

$$\begin{aligned}\langle S_x;+|S_x;-\rangle &= \frac{1}{\sqrt{4}}\left[\langle +| + e^{-i\delta_x}\langle -|\right]\left[|+\rangle - e^{i\delta_x}|-\rangle\right] \\ &= \frac{1}{2}\left[\langle +|+\rangle - e^{-i\delta_x}e^{i\delta_x}\langle -|-\rangle\right] = \frac{1}{2}[1-1] = 0\,.\end{aligned} \tag{1.85}$$

Having thus found the eigenkets of $\hat{S}_x$, it is straightforward to construct it. In analogy to $\hat{S}_z$ it reads

$$\begin{aligned}\hat{S}_x &= \frac{\hbar}{2}\left(|S_x;+\rangle\langle S_x;+| - |S_x;-\rangle\langle S_x;-|\right) \\ &= \frac{\hbar}{2}\left[e^{-i\delta_x}|+\rangle\langle -| + e^{i\delta_x}|-\rangle\langle +|\right] .\end{aligned} \tag{1.86}$$

Simple inspection shows that this is indeed a Hermitian operator, as is $\hat{S}_x$. This is good news, since Hermitian operators must have real eigenvalues - making them perfect candidates for observables.
Similar reasoning of course holds true for $\hat{S}_y$, leading to

$$|S_y;\pm\rangle = \frac{1}{\sqrt{2}}\left[|+\rangle \pm e^{i\delta_y}|-\rangle\right] . \tag{1.87}$$

and

$$\begin{aligned}\hat{S}_y &= \frac{\hbar}{2}\left(|S_y;+\rangle\langle S_y;+| - |S_y;-\rangle\langle S_y;-|\right)\\ &= \frac{\hbar}{2}\left[e^{-i\delta_y}|+\rangle\langle -| + e^{i\delta_y}|-\rangle\langle +|\right]\,. \end{aligned} \tag{1.88}$$

So far so good. But however nice this looks there is still some problem to solve, namely to find values for the so far arbitrary real phases δ_x and δ_y. For this purpose, the last bit of information can be used, namely the interplay of sequential Stern-Gerlach devices pointing into $x-$ and y-direction. The results of such a configuration finally yields

$$|\langle S_y;\pm|S_x;+\rangle|^2 = |\langle S_y;\pm|S_x;-\rangle|^2 = \frac{1}{2}\,, \tag{1.89}$$

as for the case with sequential z- and x- or z- and y-Stern-Gerlachs. Inserting the respective expression for the state vectors results in

$$\frac{1}{2}\left|1 \pm e^{i(\delta_x-\delta_y)}\right| = \frac{1}{\sqrt{2}} \tag{1.90}$$

or

$$\delta_x - \delta_y = \pm\frac{\pi}{2}\,. \tag{1.91}$$

This immediately implies that not all of the matrix elements of both $\hat{S}_x$ and $\hat{S}_y$ can be real. In fact, choosing $\delta_x = 0$ - a standard choice - implies that $\hat{S}_y$ is purely imaginary.
Having fixed the phases in the standard way ($\delta_x = 0$ and $\delta_y = \pi/2$), yields

$$\begin{aligned}|S_x;\pm\rangle &= \frac{1}{\sqrt{2}}\left[|+\rangle \pm |-\rangle\right]\\ |S_y;\pm\rangle &= \frac{1}{\sqrt{2}}\left[|+\rangle \pm i|-\rangle\right]\end{aligned} \tag{1.92}$$

and

$$\begin{aligned}\hat{S}_x &= \frac{\hbar}{2}\Big[|+\rangle\langle -| + |-\rangle\langle +|\Big]\\ \hat{S}_y &= \frac{i\hbar}{2}\Big[-|+\rangle\langle -| + |-\rangle\langle +|\Big]\end{aligned} \tag{1.93}$$

A simple calculation shows that the **commutator** of two of the operators yields the third one,

$$[\hat{S}_i,\, \hat{S}_j] := \hat{S}_i\hat{S}_j - \hat{S}_j\hat{S}_i = i\epsilon_{ijk}\, \hat{S}_k\,, \tag{1.94}$$

where ϵ_{ijk} is the Levi-Civita symbol with $\epsilon_{123} = 1$. Let's check this for the case $i = 1$ (x) and $j = 2$ (y):

$$\begin{aligned}
[\hat{S}_x,\, \hat{S}_y] &= \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x \\
&= \frac{i\hbar^2}{4}\left\{\Big[|+\rangle\langle-| + |-\rangle\langle+|\Big]\Big[-|+\rangle\langle-| + |-\rangle\langle+|\Big]\right. \\
&\qquad \left. - \Big[-|+\rangle\langle-| + |-\rangle\langle+|\Big]\Big[|+\rangle\langle-| + |-\rangle\langle+|\Big]\right\} \\
&= \frac{i\hbar^2}{4}\left\{\Big[|+\rangle\langle+| - |-\rangle\langle-|\Big] - \Big[-|+\rangle\langle+| + |-\rangle\langle-|\Big]\right\} \\
&= \frac{i\hbar^2}{2}\Big[|+\rangle\langle+| - |-\rangle\langle-|\Big] = i\hbar\hat{S}_z\,,
\end{aligned} \tag{1.95}$$

as claimed. In a similar way, the **anti-commutator** of two of these operators is given by

$$\{\hat{S}_i,\, \hat{S}_j\} := \hat{S}_i\hat{S}_j + \hat{S}_j\hat{S}_i = \frac{\hbar^2}{2}\delta_{ij}\,. \tag{1.96}$$

These commutation relations are the simplest realization of the **angular momentum algebra**, which plays a central role in many quantum mechanical problems.

Related to this algebra are a few more interesting operators, namely, first, the **ladder operators**, defined by

$$\begin{aligned}
\hat{S}_+ &= \hbar\Big[|+\rangle\langle-|\Big] \doteq \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
\hat{S}_- &= \hbar\Big[|-\rangle\langle+|\Big] \doteq \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.
\end{aligned} \tag{1.97}$$

Obviously, these operators are non-Hermitian; more operators of this kind will be encountered during this course. They have the following interpretation: $\hat{S}_+$, acting on a $|-\rangle$-state raises the spin by one unit of $\hbar$ from $-\hbar/2$ to $\hbar/2$ and flips the state to a $|+\rangle$-state; acting on a $|+\rangle$-state it produces the null-ket. In contrast, the $\hat{S}_-$ operator yields the null-ket when acting on a $|-\rangle$ state and produces a $|-\rangle$-ket when acting on a $|+\rangle$-state. So, taking together, these two operators shift the spin by one unit of $\hbar$. They can also be expressed through

$$\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y \,. \tag{1.98}$$

Another important operator is

$$\hat{\vec{S}}^2 = \hat{\vec{S}} \cdot \hat{\vec{S}} = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \,. \tag{1.99}$$

Simple calculation, employing the anti-commutator relation above shows that it is just a multiple of the unit operator, namely

$$\hat{\vec{S}}^2 = \frac{3}{4}\hbar^2 \mathbf{1} \,, \tag{1.100}$$

showing that

$$\left[\hat{\vec{S}}^2,\, \hat{S}_i\right] = 0 \,. \tag{1.101}$$

Compatible observables

Two observables are called **compatible** (or **incompatible**) if the commutator of the corresponding operators $\hat{A}$ and $\hat{B}$ does (or does not) vanishes. For the following discussion the case of **degeneracy** [6] will be ignored.
Let's start with an important theorem: **Commuting hermitian operators have the same eigenkets**, if the eigenvalues are non-degenerate. Mathematically speaking

$$\langle a^{(i)}|\hat{B}|a^{(j)}\rangle \propto \delta_{ij} \quad \text{if } \hat{A}|a^{(i,j)}\rangle \propto |a^{(i,j)}\rangle \quad \text{and } [\hat{A},\, \hat{B}] = 0 \,. \tag{1.102}$$

The proof is extremely simple and straightforward. Using the definition of the commutator yields

$$\begin{aligned}
\langle a^{(i)}|[\hat{A},\, \hat{B}]|a^{(j)}\rangle &= \langle a^{(i)}| \left[\hat{A}\hat{B} - \hat{B}\hat{A}\right] |a^{(j)}\rangle \\
&= (a^{(i)} - a^{(j)})\langle a^{(i)}|\hat{B}|a^{(j)}\rangle \stackrel{!}{=} 0 \,,
\end{aligned} \tag{1.103}$$

[6] **Degeneracy** denotes a situation, where two or more linearly independent eigenkets have the same eigenvalue. In such a case, the theorems above concerning their orthogonality run into trouble and so does the full concept of a system of eigenkets spanning the full vector space. However, in most practical applications the eigenvalues of some other, commuting observables are different and can be used to lift this apparent labelling problem.

where the zero result is enforced by the vanishing commutator. Then, in order to fulfil this equation, either $a^{(i)} = a^{(j)}$ or the matrix element $\langle a^{(i)}|\hat{B}|a^{(j)}\rangle$ must vanish. Hence, this matrix element can be written as

$$\langle a^{(i)}|\hat{B}|a^{(j)}\rangle = \delta_{ij}\langle a^{(i)}|\hat{B}|a^{(i)}\rangle \tag{1.104}$$

indicating that $\hat{B}$ is diagonal when represented in the basis formed by the eigenkets of $\hat{A}$:

$$\hat{B} = \sum_{i=1,\dots N} |a^{(i)}\rangle\langle a^{(i)}|\hat{B}|a^{(i)}\rangle\langle a^{(i)}|\,. \tag{1.105}$$

Multiplying this equation from the right with $|a^{(j)}\rangle$ yields

$$\begin{aligned}\hat{B}|a^{(j)}\rangle &= \sum_{i=1,\dots N} |a^{(i)}\rangle\langle a^{(i)}|\hat{B}|a^{(i)}\rangle\langle a^{(i)}|a^{(j)}\rangle \\ &= \sum_{i=1,\dots N}\Big(|a^{(i)}\rangle\langle a^{(i)}|\hat{B}|a^{(i)}\rangle\delta_{ij}\Big) = \Big(\langle a^{(j)}|\hat{B}|a^{(j)}\rangle\Big)\cdot|a^{(j)}\rangle\,,\end{aligned} \tag{1.106}$$

an eigenvalue equation, where the eigenvalue $b^{(j)} = \langle a^{(j)}|\hat{B}|a^{(j)}\rangle$. Thus, the kets $|a^{(i)}\rangle$ are simultaneous eigenkets of both $\hat{A}$ and $\hat{B}$, and, in order to make this manifest, one may use the notation $|a^{(i)},\, b^{(i)}\rangle$[7].
However, in order to dwell on the concept of compatible observables A and B and understand the physical meaning, let's consider their measurements. Suppose, A is measured first. Then the state vector of the system jumps into the eigenket $|a^{(i)}\rangle$ of the operator $\hat{A}$ corresponding to this observable, and the eigenvalue $a^{(i)}$ is measured. Subsequently, $\hat{B}$ is applied, i.e., B is measured. The system, already being in an eigenstate of $\hat{B}$, won't jump again, and $b^{(i)}$ will be obtained. Measuring A again, will yield $a^{(i)}$ with certainty. This highlights why the term "*compatible*" indeed is appropriate.

Incompatible observables

Let's turn now to incompatible observables. The most important point to keep in mind here is that the corresponding operators **do not have com-**

[7] Although in the proof above, non-degeneracy was assumed, the same statements hold true also for an n-fold degeneracy. In such a case, linear combinations of the original $a^{(i)}\rangle$, diagonalising $\hat{A}$ must be found, that also diagonalise $\hat{B}$. In such a case, the somewhat superfluous notation $|a,\, b\rangle$ introduced above makes sense: Degenerate eigenkets to the same eigenvalue a can be distinguished by the different eigenvalues b. A prime example for this is the case of orbital angular momentum, discussed later.

plete sets of simultaneous eigenkets. The proof of this is through contradiction:
Let's assume that they have simultaneous eigenkets. Then

$$\hat{A}\hat{B}|a^{(i)},\, b^{(i)}\rangle = \hat{A}b^{(i)}|a^{(i)},\, b^{(i)}\rangle = a^{(i)}b^{(i)}|a^{(i)},\, b^{(i)}\rangle \tag{1.107}$$

and

$$\hat{B}\hat{A}|a^{(i)},\, b^{(i)}\rangle = \hat{B}a^{(i)}|a^{(i)},\, b^{(i)}\rangle = a^{(i)}b^{(i)}|a^{(i)},\, b^{(i)}\rangle \tag{1.108}$$

for all i, and hence,

$$\hat{A}\hat{B}|a^{(i)},\, b^{(i)}\rangle = \hat{B}\hat{A}|a^{(i)},\, b^{(i)}\rangle \quad \forall i \tag{1.109}$$

leading to

$$[\hat{A},\, \hat{B}] = 0\,, \tag{1.110}$$

a clear contradiction.
From the construction of operators related to the Stern-Gerlach devices, it is already known that they are incompatible: the respective commutators do not vanish. However, now a more general discussion will be given. Consider a sequence of selective measurements, i.e. filters. A-filters select only a-states, B-filters select only b-states and so on. Normalizing the first beam, passing the A filter to unity, what is the probability of measuring c in the sequence $A \to B \to C$? Because probabilities are multiplicative, the answer obviously is $|\langle c|b\rangle|^2|\langle b|a\rangle|^2$. But what happens if rather than selecting the intermediate $|b\rangle$-state through the B-filter all possible $|b^{(i)}\rangle$-routes are allowed? In terms of actual measurement this boils down to blocking all routes but the $b^{(1)}$-route, recording the probability for obtaining c in the end, then considering $b^{(2)}$ with all other options blocked and so on, or mathematically

$$\sum_i |\langle c|b^{(i)}\rangle|^2|\langle b^{(i)}|a\rangle|^2 = \sum_i \langle c|b^{(i)}\rangle\langle b^{(i)}|a\rangle\langle a|b^{(i)}\rangle\langle b^{(i)}|c\rangle\,. \tag{1.111}$$

How does this compare with no selective B-measurement at all, i.e. having no B-type filter? In such a case the probability is just $|\langle c|a\rangle|^2$. But, inserting identities this can be written as

$$|\langle c|a\rangle|^2 = \left|\sum_i \langle c|b^{(i)}\rangle\langle b^{(i)}|a\rangle\right|^2 = \sum_{i,j} \langle c|b^{(i)}\rangle\langle b^{(i)}|a\rangle\langle a|b^{(j)}\rangle\langle b^{(j)}|c\rangle\,. \tag{1.112}$$

These two expressions, Eq. (1.111) and (1.112) clearly are different! What happened? The important point to note here is that the result after the C-filter depends on whether or not an intermediate b-state has been selected and, thus, fixed. Now, indeed, the beam leaving the A filter in the $|a\rangle$-state can be understood as superposition of $b^{(i)}\rangle$-states,

$$|a\rangle = \sum_i |b^{(i)}\rangle\langle b^{(i)}|a\rangle\,, \tag{1.113}$$

where the sum is over all b-states. But whereas in the case with an intermediate B-measurement the state has been specified, in the case of no B-filtration, the state $|a\rangle$ is **merely imagined to consist of a linear combination of the various b-states**. This difference is, basically, the essence of quantum mechanics.

The uncertainty relation

There is a last issue to be discussed in this first round of presenting basic quantum mechanical concepts, namely the uncertainty relation. To this end, let's introduce the operator

$$\Delta\hat{A}|_a := \hat{A} - \langle\hat{A}\rangle_a\,, \tag{1.114}$$

where the expectation value is taken w.r.t. a state $|a\rangle$. The expectation value of $(\Delta\hat{A})^2$ is known as the **dispersion**, **variance**, or **mean square deviation** of $\hat{A}$. Simple calculation yields

$$\langle(\Delta\hat{A})^2\rangle = \left\langle\left(\hat{A}^2 - 2\hat{A}\Delta\hat{A} + \langle\hat{A}\rangle^2\right)\right\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2\,. \tag{1.115}$$

Clearly, the dispersion vanishes, when the state $|a\rangle$ in question is an eigenstate of $\hat{A}$. For other states, the dispersion is a measure, how well the particular state can be measured: the larger the dispersion the less precise a certain state appears when measured through the operator in question. To exemplify this, consider the $|+\rangle$ state of the Stern-Gerlach experiments. Since it is an eigenstate of $\hat{S}_z$, the corresponding dispersion vanishes:

$$
\begin{aligned}
\langle(\Delta\hat{S}_z|_{|+\rangle})^2\rangle &= \langle\hat{S}_z^2\rangle_{|+\rangle} - \langle\hat{S}_z\rangle^2_{|+\rangle} \\
&= \left\langle + \left| \left[\frac{\hbar}{2}\left(|+\rangle\langle+| - |-\rangle\langle-|\right)\right]^2 \right| + \right\rangle - \left[\left\langle + \left| \frac{\hbar}{2}\left(|+\rangle\langle+| - |-\rangle\langle-|\right) \right| + \right\rangle\right]^2 \\
&= \left\langle + \left| \left[\frac{\hbar^2}{4}\left(|+\rangle\langle+| + |-\rangle\langle-|\right)\right] \right| + \right\rangle - \left[\left\langle + \left| \frac{\hbar}{2}\left(|+\rangle\langle+| - |-\rangle\langle-|\right) \right| + \right\rangle\right]^2 \\
&= \frac{\hbar^2}{4} - \frac{\hbar^2}{4} = 0\,.
\end{aligned}
\tag{1.116}
$$

Hence, the $\hat{S}_z$-apparatus measures this state very well with no probabilistic spread in the distribution. On the other hand, the dispersion of $\hat{S}_x$ does not vanish for the $|+\rangle$ state but is rather

$$
\langle(\Delta\hat{S}_x|_{|+\rangle})^2\rangle = \frac{\hbar^2}{4}\,, \tag{1.117}
$$

which can be easily seen. This is equivalent to the statement that although the mean value (expectation value) of the measurement is zero the average distance of the individually measured value to the expectation value is $\hbar/2$, the square root of the dispersion.
Now, the uncertainty relation will be stated. Let A and B be observables, represented by $\hat{A}$ and $\hat{B}$, respectively. Then, **for any state** $|a\rangle$,

$$
\langle(\Delta\hat{A})^2\rangle_a\langle(\Delta\hat{B})^2\rangle_a \geq \frac{1}{4}\left|\langle[\hat{A},\,\hat{B}]\rangle_a\right|^2\,. \tag{1.118}
$$

To prove this, let's start with three theorems:

1. The Schwarz (triangle) inequality,

$$
\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq |\langle\alpha|\beta\rangle|^2\,, \tag{1.119}
$$

which is like

$$
|\vec{a}^2||\vec{b}^2| \geq |\vec{a}\cdot\vec{b}|^2 \tag{1.120}
$$

in R^3.

To prove this, note that

$$(\langle\alpha| + \lambda^*\langle\beta|)\,(|\alpha\rangle + \lambda|\beta\rangle) \geq 0\,, \tag{1.121}$$

where λ can be any complex number. This inequality must then also hold for the choice

$$\lambda = \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}\,, \tag{1.122}$$

proving the inequality.

2. Expectation values of Hermitian operators are purely real.
   To see this, use Eq. (1.60).

3. Expectation values of purely anti-Hermitian operators, defined through $\hat{C} = -\hat{C}^\dagger$, are purely imaginary.
   To see this, use Eq. (1.60).

Now, these lemmas can be used. First, the Schwarz inequality is used with

$$\begin{aligned} |\alpha\rangle &= \Delta\hat{A}|a\rangle \\ |\beta\rangle &= \Delta\hat{B}|a\rangle\,, \end{aligned} \tag{1.123}$$

resulting in

$$\langle(\Delta\hat{A})^2\rangle_a\langle(\Delta\hat{B})^2\rangle_a \geq \left|\langle\Delta\hat{A}\Delta\hat{B}\rangle_a\right|^2\,. \tag{1.124}$$

Here, it has been used that dispersions are Hermitian by construction, since the operators related to the observables are Hermitian. For the right-hand side of the equation above, the product will be decomposed into a commutator and an anti-commutator,

$$\Delta\hat{A}\Delta\hat{B} = \frac{1}{2}[\Delta\hat{A},\, \Delta\hat{B}] + \frac{1}{2}\{\Delta\hat{A},\, \Delta\hat{B}\}\,. \tag{1.125}$$

The commutator can be cast into

$$\frac{1}{2}[\Delta\hat{A},\, \Delta\hat{B}] = \frac{1}{2}[\hat{A} - \langle\hat{A}\rangle,\, \hat{B} - \langle\hat{B}\rangle] = \frac{1}{2}[\hat{A},\, \hat{B}] \tag{1.126}$$

since the expectation values are just numbers. Simple calculation shows that the commutator clearly is anti-Hermitian,

$$\begin{aligned} [\hat{A},\, \hat{B}]^\dagger &= (\hat{A}\hat{B} - \hat{B}\hat{A})^\dagger = \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger \\ &= \hat{B}\hat{A} - \hat{A}\hat{B} = [\hat{B},\, \hat{A}] = -[\hat{A},\, \hat{B}]\,, \end{aligned} \tag{1.127}$$

where the fact has been used that the operators are Hermitian. In a similar way it can be shown that the anti-commutator is Hermitian.
The expectation value $\langle \Delta\hat{A}\Delta\hat{B}\rangle_a$ can be written as

$$\langle \Delta\hat{A}\Delta\hat{B}\rangle_a = \frac{1}{2}\langle[\Delta\hat{A},\, \Delta\hat{B}]\rangle_a + \frac{1}{2}\langle\{\Delta\hat{A},\, \Delta\hat{B}\}\rangle_a\,, \tag{1.128}$$

where, according to lemmas 2 and 3, the first term is purely imaginary and the second one is purely real. Then

$$\left|\langle \Delta\hat{A}\Delta\hat{B}\rangle_a\right|^2 = \frac{1}{4}\left|\langle[\Delta\hat{A},\, \Delta\hat{B}]\rangle_a\right|^2 + \frac{1}{4}\left|\langle\{\Delta\hat{A},\, \Delta\hat{B}\}\rangle_a\right|^2\,, \tag{1.129}$$

and omitting the second term just yields the inequality

$$\left|\langle \Delta\hat{A}\Delta\hat{B}\rangle_a\right|^2 \geq \frac{1}{4}\left|\langle[\Delta\hat{A},\, \Delta\hat{B}]\rangle_a\right|^2\,. \tag{1.130}$$

This finalises the proof of the uncertainty relation.

Chapter 2

Position, momentum and angular momentum

2.1 Position and momentum space

2.1.1 Position as a continuous variable

How to deal with continuous variables

So far, only operators with discrete eigenvalues and hence with finite dimensional Hilbert spaces have been considered. This is by no means always the situation: Obviously there are observables and thus operators with a continuous spectrum. Take, for instance, the z-component of momentum, p_z, of a free particle, which in principle can assume any value in $[-\infty, \infty]$, or the x-position of the very same free particle, travelling freely.
Before going to give a brief upshot of how to deal with the emerging, obviously infinite-dimensional vector spaces it is important to note that this subject is far from being trivial when mathematical rigor is asked for. However, many of the results worked out so far still hold true - here, a pragmatic approach will be taken by just stating what can be borrowed without having to pay for it.
Let us start with the analogue of the eigenvalue equation in infinite vector spaces. Similar to the finite-dimensional case, one may write

$$\hat{\Xi}|\xi\rangle = \hat{\xi}|\xi\rangle\,, \tag{2.1}$$

where now ξ is simply a number (real or complex), $\hat{\Xi}$ is the corresponding infinite-dimensional operator acting on a similar infinite-dimensional ket $|\xi\rangle$, here obviously an eigenket. One may choose an orthonormal set of such eigenkets, where now orthonormality is expressed through the **Dirac δ-function** rather than by the Kronecker symbol:

$$\langle\alpha|\alpha'\rangle = \delta_{\alpha\alpha'} \implies \langle\xi|\xi'\rangle = \delta(\xi - \xi')\,. \tag{2.2}$$

Then, the **completeness relation** becomes

$$1 = \sum_\alpha |\alpha\rangle\langle\alpha| \implies 1 = \int \mathrm{d}\xi \xi\rangle\langle\xi|| \tag{2.3}$$

and any state ket $|a\rangle$ can thus be represented as linear combination of the orthonormal base kets by

$$|a\rangle = \sum_\alpha \langle\alpha|a\rangle|\alpha\rangle = \sum_\alpha c_\alpha|\alpha\rangle \implies |a\rangle = \int \mathrm{d}\xi\langle\xi|a\rangle|\xi\rangle = \int \mathrm{d}\xi c(\xi)|\xi\rangle\,, \tag{2.4}$$

If in the equation above, the state ket $|a\rangle$ is normalised the coefficients satisfy

$$\sum_\alpha |c_\alpha|^2 = \sum_\alpha |\langle\alpha|a\rangle|^2 = 1 \implies \int \mathrm{d}\xi|c(\xi)|^2 = \int \mathrm{d}\xi|\langle\xi|a\rangle|^2 = 1\,. \tag{2.5}$$

Orthonormality of eigenkets can then be expressed by

$$\langle\alpha|\hat{A}|\alpha'\rangle = \alpha\delta_{\alpha\alpha'} \implies \langle\xi|\hat{\Xi}|\xi'\rangle = \xi\delta(\xi - \xi')\,. \tag{2.6}$$

Position measurements

In the previous chapter, the measurement process in quantum mechanics has been discussed, where the concept of filtering played a central role. Let us now use the very same concept for an observable with a continuous rather than a discrete spectrum, namely position. To keep things simple, here position in one dimension, say along the x-axis only, will be considered. The generalisation to the three-dimensional position is simple: Just place a vector-arrow over each occurrence of x, $|x\rangle$, or $\hat{x}$. Again the starting point of the discussion is the eigenvalue equation

$$\hat{x}|x\rangle = x|x\rangle \tag{2.7}$$

or

$$\hat{\vec{x}}|\vec{x}\rangle = \vec{x}|\vec{x}\rangle \tag{2.8}$$

in three dimensions. At this point it must be stressed that this equation in fact are three equations, one for each component of the position vector $|\vec{x}\rangle$. Therefore, the eigenvalues $\vec{x}$ of the position operator $\hat{\vec{x}}$ are three numbers. The corresponding Hilbert space, spanned by the kets $|\vec{x}\rangle$ in fact is the product of three infinite-dimensional Hilbert spaces: One for each of the components.

Postulating that the eigenkets $|x\rangle$ form a complete, orthonormal set spanning the (position-) vector space, it is straightforward to expand any state ket $|a\rangle$ in terms of them:

$$|a\rangle = \int\limits_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x|a\rangle\,. \tag{2.9}$$

Let us now discuss how a measurement can be performed. Following the reasoning of the discussion in the previous chapter, this is related to a filtration: During the measurement, the system is forced to “jump” into one of the eigenstates of the operator related to the filtration observable. For systems with a discrete spectrum, this is straightforward, for a continuous system, this is more complicated. In fact, due to the eigenstates being labelled by continuous eigenvalues it seems as if **realistically a measurement can only force the system to jump into an interval of such eigenvalues**. For the specific case at hand, position, this implies that the filtration will force the system to jump into a state related to any eigenvalue in some interval $[x - \Delta x/2,\, x + \Delta x/2]$. Hence

$$|a\rangle = \int\limits_{-\infty}^{\infty} \mathrm{d}x'|x'\rangle\langle x'|a\rangle \stackrel{\text{measurement}}{\longrightarrow} \int\limits_{x-\Delta x/2}^{x+\Delta x/2} \mathrm{d}x'|x'\rangle\langle x'|a\rangle\,. \tag{2.10}$$

Assuming that the coefficient $\langle x'|a\rangle$ does not change in the interval, the probability to measure the system in this interval is $|\langle x'|a\rangle|^2 Deltax$. Replacing Δx with $\mathrm{d}x'$, i.e. going to the limit of small intervals, the probability to measure the system at any position is given by

$$\mathcal{P}_{\text{somewhere}} = \int\limits_{-\infty}^{\infty} \mathrm{d}x'|\langle x'|a\rangle|^2\,. \tag{2.11}$$

If the state ket $|a\rangle$ has been properly normalised this implies that this probability is unity,

$$\langle a|a\rangle = 1 \Longrightarrow \mathcal{P}_{\text{somewhere}} = \int\limits_{-\infty}^{\infty} \mathrm{d}x'|\langle x'|a\rangle|^2 = \int\limits_{-\infty}^{\infty} \mathrm{d}x'\langle a|x'\rangle\langle x'|a\rangle = 1\,. \tag{2.12}$$

Here, one may already identify $\langle x'|a\rangle$ with the wave function in position space related to the system. This will be discussed in the next section.
As indicated above, the generalisation to three-dimensional positions is straightforward. In this case, the expansion of a state ket $|a\rangle$ reads

$$|a\rangle = \int\limits_V \mathrm{d}^3x|\vec{x}\rangle\langle\vec{x}|a\rangle\,, \tag{2.13}$$

where V denotes the volume of space. In order to make the three-dimensional nature manifest, let us replace

$$|\vec{x}\rangle = |x,\, y,\, z\rangle = |x_1,\, x_2,\, x_3\rangle\,. \tag{2.14}$$

Hence, $\vec{x}$ is a simultaneous eigenvalue of the three operators $\hat{x}$, $\hat{y}$, and $\hat{z}$, which can conveniently be written as one, $\hat{\vec{x}}$. In order that this construction of a simultaneous eigenvalue works at all, the three operators must be compatible, i.e. they must commute,

$$[\hat{x},\, \hat{y}] = 0 \quad \text{or, more generally} \quad [\hat{x}_i,\, \hat{x}_j] = 0\,. \tag{2.15}$$

Infinitesimal translations

Let us now discuss an extremely important concept in quantum mechanics and, in fact, in theoretical physics, namely infinitesimal translations. To see, what is meant by it, consider a system, for instance a particle, well-localised around $\vec{x}$. Let us consider a transformation which moves this particle into another well-localised state around $\vec{x} + \mathrm{d}\vec{x}$, where all other properties, like, e.g., spin, remain unchanged. Such an operation is dubbed a translation, if $\mathrm{d}\vec{x}$ is infinitely 'small, it is obviously an infinitesimal translation[1]. The

[1]At this point it is noteworthy that in some textbooks translation is introduced in an alternative fashion, namely as an operation acting on the observer's system, i.e. shifting the origin of the coordinate system. It should be plain clear that these two views should not be mixed up, although they finally yield the same results.

corresponding operator $\hat{\mathcal{T}}(\mathrm{d}\vec{x})$ thus connects the two state vectors $|\vec{x}\rangle$ and $|\vec{x}+\mathrm{d}\vec{x}\rangle$,

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x})|\vec{x}\rangle = |\vec{x}+\mathrm{d}\vec{x}\rangle \tag{2.16}$$

and for obvious reasons

$$\hat{\mathcal{T}}(-\mathrm{d}\vec{x})|\vec{x}+\mathrm{d}\vec{x}\rangle = |\vec{x}\rangle\,. \tag{2.17}$$

In both cases, an arbitrary phase which in principle could emerge, has been set to 0.
Expanding an arbitrary state ket $|a\rangle$ in terms of position eigenkets, the action of the translation operator on such kets becomes more transparent:

$$\begin{aligned} |a\rangle \Longrightarrow \hat{\mathcal{T}}(\mathrm{d}\vec{x})|a\rangle &= \hat{\mathcal{T}}(\mathrm{d}\vec{x})\int \mathrm{d}^3x|\vec{x}\rangle\langle\vec{x}|a\rangle \\ &= \int \mathrm{d}^3x|\vec{x}+\mathrm{d}\vec{x}\rangle\langle\vec{x}|a\rangle = \int \mathrm{d}^3x|\vec{x}\rangle\langle\vec{x}-\mathrm{d}\vec{x}|a\rangle \end{aligned} \tag{2.18}$$

where the last equation holds true, since the shift $\mathrm{d}\vec{x}$ can be thought of some arbitrary, but constant shift and d^3x is just a continuous integration variable.

Properties of the translation operator

Let us now list the properties of the translation operator $\hat{\mathcal{T}}$:

- **Limit:**
  It is obvious to assume that the limit of vanishing translation vector is identical to no operation at all,

$$\lim_{\mathrm{d}\vec{x}\to 0}\hat{\mathcal{T}}(\mathrm{d}\vec{x}) = 1\,. \tag{2.19}$$

- **Unitarity:**
  This is another statement of **probability conservation**. Demanding that the ket $|\alpha\rangle$ is normalised to unity, it is clearly reasonable to demand that also its translated partner $\hat{\mathcal{T}}(\mathrm{d}\vec{x})|\alpha\rangle$ is normalised to unity[2]. Hence

$$\langle\alpha|\alpha\rangle = \langle\alpha|\hat{\mathcal{T}}^\dagger(\mathrm{d}\vec{x})\hat{\mathcal{T}}(\mathrm{d}\vec{x})|\alpha\rangle = 1 \quad\Longrightarrow\quad \hat{\mathcal{T}}^\dagger(\mathrm{d}\vec{x})\hat{\mathcal{T}}(\mathrm{d}\vec{x}) = 1\,. \tag{2.20}$$

[2] To see the reason for this, just consider the limiting case, $\lim\limits_{\mathrm{d}\vec{x}\to 0}\hat{\mathcal{T}}(\mathrm{d}\vec{x})|\alpha\rangle$.

- **Additivity:**
  Consider two successive infinitesimal translations (infinitesimal is important!), not necessary into the same direction. The overall result should just be identical to one translation, into the direction of both vectors added:

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x}_1)\hat{\mathcal{T}}(\mathrm{d}\vec{x}_2) = \hat{\mathcal{T}}(\mathrm{d}\vec{x}_1 + \mathrm{d}\vec{x}_2)\,. \tag{2.21}$$

- **Inversion:**
  It is also reasonable to assume that a translation into the opposite direction is just the inverse of the original translation, i.e.

$$\hat{\mathcal{T}}(-\mathrm{d}\vec{x}) = \hat{\mathcal{T}}^{-1}(\mathrm{d}\vec{x})\,. \tag{2.22}$$

It can be demonstrated that all these properties can be satisfied, when the operator related to an infinitesimal translation can be written as

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x}) = 1 - i\hat{\vec{K}} \cdot \mathrm{d}\vec{x}\,, \tag{2.23}$$

where the components $\hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$ of $\hat{\vec{K}}$ are Hermitian operators, which do not need to be specified more precisely here. In fact, this will be dealt with in the next section. Then the first property, the zero limit is obvious. Unitarity can be checked through

$$\begin{aligned}\hat{\mathcal{T}}^\dagger(\mathrm{d}\vec{x})\hat{\mathcal{T}}(\mathrm{d}\vec{x}) &= \left(1 + i\hat{\vec{K}}^\dagger \cdot \mathrm{d}\vec{x}\right)\left(1 - i\hat{\vec{K}} \cdot \mathrm{d}\vec{x}\right) \\ &= 1 - i(\hat{\vec{K}} - \hat{\vec{K}}^\dagger) \cdot \mathrm{d}\vec{x} + \mathcal{O}[(\mathrm{d}\vec{x})^2] \approx 1\,,\end{aligned} \tag{2.24}$$

where terms of infinitesimal second order are safely neglected and where the definition of unitary operators, cf. Eq. (1.32) has been used. In order to check for the additivity, consider

$$\begin{aligned}\hat{\mathcal{T}}(\mathrm{d}\vec{x}_1)\hat{\mathcal{T}}(\mathrm{d}\vec{x}_2) &= \left(1 - i\hat{\vec{K}} \cdot \mathrm{d}\vec{x}_1\right)\left(1 - i\hat{\vec{K}} \cdot \mathrm{d}\vec{x}_2\right) \\ &= 1 - i\hat{\vec{K}} \cdot (\mathrm{d}\vec{x}_1 + \mathrm{d}\vec{x}_2) + \mathcal{O}[(\mathrm{d}\vec{x})^2] \approx \hat{\mathcal{T}}(\mathrm{d}\vec{x}_1 + \mathrm{d}\vec{x}_2)\,,\end{aligned} \tag{2.25}$$

where again terms of higher order have been ignored. The inversion is also obvious.

Accepting that Eq. (2.23) indeed is the correct form of the translation operator, fundamental relations between the operators $\hat{\vec{K}}$ and $\hat{\vec{x}}$, the position operator, can be found. First, note that applying the position operator on a translated position ket yields

$$\hat{\vec{x}}\hat{\mathcal{T}}(\mathrm{d}\vec{x}')|\vec{x}'\rangle = \hat{\vec{x}}|\vec{x}' + \mathrm{d}\vec{x}'\rangle = (\vec{x}' + \mathrm{d}\vec{x}')|\vec{x}' + \mathrm{d}\vec{x}'\rangle\,, \tag{2.26}$$

since $\vec{x}' + \mathrm{d}\vec{x}'$ is the eigenvalue of the position operator corresponding to the shifted position. On the other hand,

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x}')\hat{\vec{x}}|\vec{x}'\rangle = \vec{x}'\hat{\mathcal{T}}(\mathrm{d}\vec{x}')|\vec{x}'\rangle = \vec{x}'|\vec{x} + \mathrm{d}\vec{x}\rangle\,. \tag{2.27}$$

Subtracting these two equations results in

$$\begin{aligned}\left[\hat{\vec{x}}\,,\hat{\mathcal{T}}(\mathrm{d}\vec{x}')\right]\vec{x}'\rangle &:= \left[\hat{\vec{x}}\hat{\mathcal{T}}(\mathrm{d}\vec{x}') - \hat{\mathcal{T}}(\mathrm{d}\vec{x}')\hat{\vec{x}}\right]|\vec{x}'\rangle \\ &= \left[(\vec{x}' + \mathrm{d}\vec{x}') - \vec{x}'\right]|\vec{x}' + \mathrm{d}\vec{x}'\rangle = \mathrm{d}\vec{x}'|\vec{x}' + \mathrm{d}\vec{x}'\rangle \approx \mathrm{d}\vec{x}'|\vec{x}'\rangle\,,\end{aligned} \tag{2.28}$$

where terms of higher order in $\mathrm{d}\vec{x}'$ have been ignored. $\vec{x}'$ can be any position, and thus $|\vec{x}'\rangle$ can be any position eigenket in the complete set of position eigenkets spanning the Hilbert space. Therefore, the relation between the two operators must be an operator identity, i.e. a valid relation irrespective of the ket to the right. Hence

$$\left[\hat{\vec{x}}\,,\hat{\mathcal{T}}(\mathrm{d}\vec{x}')\right] = \mathrm{d}\vec{x}'\cdot\mathbf{1} \tag{2.29}$$

or

$$-i\hat{\vec{x}}\left(\hat{\vec{K}}\cdot\mathrm{d}\vec{x}'\right) + i\left(\hat{\vec{K}}\cdot\mathrm{d}\vec{x}'\right)\hat{\vec{x}} = \mathrm{d}\vec{x}'\cdot\mathbf{1}\,. \tag{2.30}$$

In both cases $\mathbf{1}$ is the identity operator (in the ket space spanned by the $|\vec{x}'\rangle$). Let us choose now $\mathrm{d}\vec{x}'$ in the direction $\vec{e}_j$ and form the let us form the scalar product with $\vec{e}_i$, where the $\vec{e}_{i,j}$ are unit vectors spanning the coordinate system, i.e. $\vec{e}_x$, $\vec{e}_y$, or $\vec{e}_z$. Then, with

$$\mathrm{d}\vec{x}' = \mathrm{d}x\vec{e}_i \tag{2.31}$$

and by factoring out and throwing away the infinitesimal number $\mathrm{d}x$, the above equation reads

$$\vec{e}_j\cdot\left[-i\hat{\vec{x}}\left(\hat{\vec{K}}\cdot\vec{e}_i\right) + i\left(\hat{\vec{K}}\cdot\vec{e}_i\right)\hat{\vec{x}}\right] = \vec{e}_j\cdot\vec{e}_i\cdot\mathbf{1}\,. \tag{2.32}$$

Straightforward calculation yields

$$\vec{e}_j \cdot \left[-i\hat{\vec{x}}\hat{K}_i + i\hat{K}_i\hat{\vec{x}}\right] = \left[-i\hat{x}_j\hat{K}_i + i\hat{K}_i\hat{x}_j\right] = -i\left[\hat{x}_j,\, \hat{K}_i\right] = \delta_{ij}\cdot\mathbf{1} \quad (2.33)$$

or

$$\left[\hat{x}_j,\, \hat{K}_i\right] = i\delta_{ij}\,. \tag{2.34}$$

2.1.2 Momentum

Eq. (2.34) is the fundamental commutation relation between the position operators $\hat{x}_i = \hat{x}$, $\hat{y}$, and $\hat{z}$ and the operators of translation $\hat{K}_i = \hat{K}_x$, $\hat{K}_y$, and $\hat{K}_z$. So far these operators have been defined in terms of infinitesimal translations along the respective axis. But what is the physical significance of them? In the following, the fact that in classical mechanics momentum is the generator of infinitesimal translations will be borrowed and its consequences will be exploited.

Momentum as generator of translation

In classical mechanics, an infinitesimal translation

$$\vec{x}_{\text{new}} \equiv \vec{X} = \vec{x} + \mathrm{d}\vec{x}\;, \quad \vec{p}_{\text{new}} \equiv \vec{P} = \vec{p} \tag{2.35}$$

can be regarded as a canonical transformation. This transformation is obtained from the generating function

$$F(\vec{x},\, \vec{P}) = \vec{x}\cdot\vec{P} + \vec{p}\cdot\mathrm{d}\vec{x}\,, \tag{2.36}$$

where $\vec{p}$ and $\vec{P}$ refer to the corresponding momenta.
The similarity of this equation with the infinitesimal translation operator in Eq. (2.23) becomes more than striking, when recalling that the term $\vec{x}\cdot\vec{P}$ is the generating function for the identity transformation

$$\vec{X} = \vec{x}\;, \quad \vec{P} = \vec{p}\,. \tag{2.37}$$

This immediately leads to the speculation that the operator $\hat{\vec{K}}$ is somehow related to the momentum operator in quantum mechanics.
Now, the question naturally arises, whether the operator $\hat{\vec{K}}$ can be identified directly with the momentum operator? Unfortunately, on dimensional

grounds this is impossible: Because $\hat{\vec{K}} \cdot d\vec{x}$ must have the same physical dimension as the identity operator (no dimension at all), $\hat{\vec{K}}$ **must have the dimension of inverse length**. Nevertheless, it is perfectly legitimate to set

$$\hat{\vec{K}} = \frac{1}{\text{constant}} \hat{\vec{p}}, \tag{2.38}$$

where the constant is universal and has the dimensions of action, and where $\hat{\vec{p}}$ is the momentum operator[3]. This universal constant turns out to be the same as $\hbar$ appearing in de Broglie's relation from 1924,

$$\frac{2\pi}{\lambda} = \frac{p}{\hbar}, \tag{2.39}$$

which connects the wavelength λ of a particle with its momentum p.

In other words, the $\hat{\vec{K}}$ operator is the quantum mechanical operator that corresponds to the wavenumber - 2π times the reciprocal wavelength, usually denoted by k. With this identification, the operator of infinitesimal translation reads

$$\hat{\mathcal{T}}(d\vec{x}) = 1 - \frac{i}{\hbar}\hat{\vec{p}} \cdot d\vec{x}\,. \tag{2.40}$$

The commutation relation Eq. (2.34) thus reads

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}\,. \tag{2.41}$$

This commutation relation implies that, for example, $\hat{x}$ and $\hat{p}_x$ are incompatible observables which do not have a set of simultaneous eigenkets. The general formalism developed so far then immediately implies that

$$\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \geq \hbar^2/4\,, \tag{2.42}$$

the **position-momentum uncertainty relation** of W. Heisenberg.

[3]Note that, from the fundamental postulates of quantum mechanics there is no way to determine the actual numerical value of the constant. Instead, it is needed, because classical mechanics with its system of units and measures for the macroscopic world was developed well before quantum mechanics. If, instead, quantum mechanics describing the microscopic world would have been developed before classical mechanics, physicists for sure would have set this constant to unity. This, in fact is the reason for a textbook convention, setting $\hbar = 1$. This has some analogy with electrostatics. The interaction strength of two particles with charge e and distance r is proportional to e^2/r, where in unrationalised Gaussian units the constant of proportionality is 1, whereas in the rationalised *mks*-system, used, e.g., by engineers, this constant factor is $1/4\pi\epsilon_0$.

Finite translations

Up to now, only infinitesimal translations have been considered. But how can, equipped with the formalism developed so far, finite translations be realised? In the following it will be shown that this is possible by successive infinitesimal translations. To start with, consider a finite translation $\Delta x'$ along the x-axis:

$$\hat{\mathcal{T}}(\Delta x'\vec{e}_x)|\vec{x}'\rangle = |\vec{x}' + \Delta x'\vec{e}_x\rangle\,. \tag{2.43}$$

Let us write this as a compound of N infinitesimal translations, each in the direction of the x-axis with a shift of $\Delta x'/N$, and let's take the limit $N \to \infty$. Then,

$$\hat{\mathcal{T}}(\Delta x'\vec{e}_x) = \lim_{N\to\infty}\left(1 - \frac{i\hat{p}_x\Delta x'}{N\hbar}\right)^N = \exp\left(-\frac{i\hat{p}_x\Delta x'}{\hbar}\right)\,. \tag{2.44}$$

Here, the exponential is understood as an operator-valued function of the operator, which is defined through the Taylor-series of the exponential

$$\exp(\hat{X}) = 1 + \hat{X} + \frac{1}{2}\hat{X}^2 + \dots\,. \tag{2.45}$$

A fundamental property of translation is that translations in different directions commute; this is implied in the additivity property above. For two finite translations in the x and y-direction this boils down to

$$\hat{\mathcal{T}}(\Delta x'\vec{e}_x)\hat{\mathcal{T}}(\Delta y'\vec{e}_y) = \hat{\mathcal{T}}(\Delta y'\vec{e}_y)\hat{\mathcal{T}}(\Delta x'\vec{e}_x) = \hat{\mathcal{T}}(\Delta x'\vec{e}_x + \Delta y'\vec{e}_y)\,. \tag{2.46}$$

To see this is not as trivial as one may expect[4]. However, treating $\Delta x'$ and $\Delta y'$ up to second order, shows that if the commutator of the finite translations must vanish,

$$\begin{aligned}
&\left[\hat{\mathcal{T}}(\Delta y'\vec{e}_y)\,,\hat{\mathcal{T}}(\Delta x'\vec{e}_x)\right] \\
&= \left[\left(1 - \frac{i\hat{p}_y\Delta y'}{\hbar} - \frac{\hat{p}_y^2(\Delta y')^2}{\hbar} + \dots\right)\,,\left(1 - \frac{i\hat{p}_x\Delta x'}{\hbar} - \frac{\hat{p}_x^2(\Delta x')^2}{\hbar} + \dots\right)\right] \\
&= -\frac{(\Delta x')(\Delta y')}{\hbar^2}\,[\hat{p}_y,\,\hat{p}_x]
\end{aligned} \tag{2.47}$$

[4]For **rotations around different axes**, for instance, which do **not commute**, such a relation does not exist.

the commutator of the momenta must vanish as well. Hence, in general

$$[\hat{p}_i,\, \hat{p}_j] = 0\,. \tag{2.48}$$

This is a **direct consequence of the requirement that finite translations do commute**. Whenever the generators of transformations (the basic operators spanning the finite transformations through exponentiation) commute, the respective group is called **Abelian**. Hence, the group of translations in three spatial dimensions is **Abelian**.
The commutator equation above implies that the three momentum operators $\hat{p}_i$ are mutually compatible. Therefore simultaneous sets of eigenkets can be constructed, namely

$$|\vec{p}'\rangle = |p'_x,\, p'_y,\, p'_z\rangle \tag{2.49}$$

such that

$$\hat{\vec{p}}|\vec{p}'\rangle = \vec{p}'|\vec{p}'\rangle \tag{2.50}$$

or

$$\hat{p}_x|\vec{p}'\rangle = p'_x|\vec{p}'\rangle\ ,\quad \hat{p}_y|\vec{p}'\rangle = p'_y|\vec{p}'\rangle\ ,\quad \hat{p}_z|\vec{p}'\rangle = p'_z|\vec{p}'\rangle\,. \tag{2.51}$$

It is instructive to work out the effect of an infinitesimal translation on such a momentum eigenket.

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x}')|\vec{p}'\rangle = \left(1 - \frac{i\hat{\vec{p}}\cdot\mathrm{d}\vec{x}'}{\hbar}\right)|\vec{p}'\rangle = \left(1 - \frac{i\vec{p}'\cdot\mathrm{d}\vec{x}'}{\hbar}\right)|\vec{p}'\rangle\,. \tag{2.52}$$

The momentum eigenket thus remains unchanged, however, it suffers from a slight change of phase given by the scalar product $\vec{p}'\cdot\mathrm{d}\vec{x}'$. Hence, unlike the position eigenkets $|\vec{x}'\rangle$, the momentum eigenkets are eigenkets of the operator of infinitesimal translation. This could have been anticipated, since

$$\left[\hat{\vec{p}},\, \hat{\mathcal{T}}(\mathrm{d}\vec{x}')\right] = 0\,, \tag{2.53}$$

which can be seen from the definition of $\hat{\mathcal{T}}$. The eigenvalues of this operator are complex - again, this could have been anticipated because, although $\hat{\mathcal{T}}$ **is unitary, it is not Hermitian**.

Canonical commutation relations

The set of commutation relations

$$\begin{aligned} [\hat{x}_i,\, \hat{p}_j] &= i\hbar\delta_{ij} \\ [\hat{x}_i,\, \hat{x}_j] &= [\hat{p}_i,\, \hat{p}_j] = 0 \end{aligned} \tag{2.54}$$

is dubbed the **canonical commutation relations** or **fundamental commutation relations**. It was W. Heisenberg, who, in 1925, realised that the combination rule for different atomic transition lines known at that time could be best understood, if arrays of numbers obeying certain multiplication rules were associated with these frequencies. Immediately after that, M. Born and P. Jordan realised that these multiplication rules are those of matrix algebra, and this triggered a theory based on matrix analogues of the commutation relations above, known as **matrix mechanics**. In the same year, P. A. M. Dirac observed that various quantum mechanical relations, including the ones above, can be guessed from classical mechanics. In order to do so, the Poisson brackets of classical mechanics would just have to be replaced by the commutators of quantum mechanics,

$$[\ ,\ ]_{\text{classical}} \longrightarrow \frac{-i}{\hbar}\,[\ ,\ ]_{\text{quantum}}\ . \tag{2.55}$$

Recalling that in classical mechanics the Poisson bracket of a pair of functions $A(q,\, p)$ and $B(q,\, p)$ is defined as

$$[A(q,\, p),\, B(q,\, p)] = \frac{\partial A}{\partial q}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} \tag{2.56}$$

the canonical commutator relations follow immediately.
Dirac's rule is quite plausible, because both the Poisson bracket and the commutator enjoys similar algebraic identities, namely

$$\begin{aligned} [A,\, A] &= 0 \\ [A,\, B] &= -[B,\, A] \\ [A,\, c] &= 0 \\ [A+B,\, C] &= [A,\, C]+[B,\, C] \\ [A,\, BC] &= [A,\, B]\,C + B\,[A,\, C] \\ 0 &= [A,\, [B,\, C]] + [B,\, [C,\, A]] + [C,\, [A,\, B]]\ , \end{aligned} \tag{2.57}$$

where A, B, and C are either functions or operators and c is just a number. The last identity is called the Jacobi identity. However, although the properties of Poisson brackets and commutators are identical, there are still important differences: First of all, due to its definition through differentiation, the Poisson bracket has a different dimension than the commutator. Second, the Poisson bracket of two real functions is also real, but the commutator of two Hermitian operators is anti-Hermitian. To take care of these differences, the factor $-i/\hbar$ needs to be inserted.

2.2 Wave functions

2.2.1 Position and momentum space

To simplify the discussion here, let us consider the one-dimensional case only; the generalisation to three spatial dimensions is straightforward.

Position space wave functions

The base kets in position space are labelled by the continuous position x', and satisfy the eigenvalue equation

$$\hat{x}|x'\rangle = x'|x'\rangle \,. \tag{2.58}$$

They form a complete orthonormal system of base kets,

$$\langle x''|x'\rangle = \delta(x'' - x') \tag{2.59}$$

and can be used to decompose any ket $|\alpha\rangle$ according to

$$|\alpha\rangle = \int \mathrm{d}x'|x'\rangle\langle x'|\alpha\rangle \,. \tag{2.60}$$

The expansion coefficients $\langle x'|\alpha\rangle$ can be interpreted such that

$$|\langle x'|\alpha\rangle|^2 \,\mathrm{d}x' \tag{2.61}$$

yields the probability to find the particle in a narrow interval $\mathrm{d}x'$ around x', if the state kets $|\alpha$ are normalised to unity. This inner product $\langle x'|\alpha\rangle$ is usually called the **wave function** of the state $|\alpha$ in position space[5]:

$$\psi_\alpha(x') = \langle x'|\alpha\rangle \,. \tag{2.62}$$

In full analogy to the cases considered so far, the inner product of two states can be rewritten as convolution of wave functions,

$$\langle\beta|\alpha\rangle = \int \mathrm{d}x' \langle\beta|x'\rangle\langle x'\alpha\rangle = \int \mathrm{d}x' \psi^*_\beta(x')\psi_\alpha(x') \,, \tag{2.63}$$

and statements about vector spaces spanned by base kets translate into statements about function spaces spanned by base functions. However, $\langle\beta|\alpha\rangle$ characterises the **overlap of two wave functions**[6]. The more general interpretation of $\langle\beta|\alpha\rangle$, independent of the representation, is that of a transition amplitude between the two states.
Let us interpret now the expansion of $|\alpha\rangle$

$$|\alpha\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha\rangle \tag{2.64}$$

in terms of the complete set of eigenkets a' of an operator $\hat{A}$ in the language of wave functions. To this end, the above equation is multiplied from the left with $\langle x'|$, leading to

$$\psi_\alpha(x') = \langle x'|\alpha\rangle = \sum_{a'} \langle x'|a'\rangle\langle a'|\alpha\rangle = \sum_{a'} u_{a'}(x') c_{a'} \,, \tag{2.65}$$

where some base wave functions $u_{a'}(x')$ are combined as a linear combination with coefficients $c_{a'}$. These base wave functions clearly are eigenfunctions of the operator $\hat{A}$ with eigenvalue a',

$$u_{a'}(x') = \langle x'|a'\rangle \,. \tag{2.66}$$

[5]In elementary wave mechanics and some textbooks on quantum mechanics starting with wave functions, the probabilistic interpretation of the expansion coefficients $c_\alpha(a) = \langle a|\alpha\rangle$ and of the wave functions $\psi_\alpha(x)$ are often presented as separate postulates. It is one of the advantages of Dirac's formalism that they are treated on identical footing as two versions - finite and continuous - of the same postulate. This clearly shows the unity of quantum mechanical formulation for systems with discrete and continuous states.

[6]It is important to stress here that the scalar product $\langle\beta|\alpha\rangle$ is not defined as the overlap integral, in fact the identification is a consequence of the completeness postulate of the position kets $|x'\rangle$.

This can be used to write

$$\langle\beta|\hat{A}|\alpha\rangle = \int \mathrm{d}x'\mathrm{d}x''\langle\beta|x'\rangle\langle x'|\hat{A}|x''\rangle\langle x''|\alpha\rangle$$
$$= \int \mathrm{d}x'\mathrm{d}x''\psi^*_\beta(x')\langle x'|\hat{A}|x''\rangle\psi_\alpha(x'') = \int \mathrm{d}x'\mathrm{d}x''\psi^*_\beta(x')A(x',\, x'')\psi_\alpha(x'')\,, \tag{2.67}$$

where, apart from the wave functions, the function $A(x,\, y)$ in position space has been introduced. The natural question is now, how to construct this function. It is clear that an enormous simplification takes place, if $\hat{A}$ can be written as a function of position operators, because they can just be replaced by the corresponding eigenvalues. In particular consider e.g.

$$\hat{A} = \hat{x}^2\,. \tag{2.68}$$

Such an operator would, for example, appear in the Hamiltonian of a simple harmonic oscillator. For this, the matrix element is given by

$$\langle x'|\hat{A}|x''\rangle = \langle x'|\hat{x}^2|x''\rangle = \langle x'|(\hat{x}^2|x''\rangle) = \langle x'|x''\rangle(x'')^2 = {x'}^2\delta(x'-x'')\,. \tag{2.69}$$

This occurrence of the δ-function actually reduces the double integration over x' and x'' to a single integration over x':

$$\langle\beta|\hat{A}|\alpha\rangle = \int \mathrm{d}x'\mathrm{d}x''\langle\beta|x'\rangle\langle x'|\hat{x}^2|x''\rangle\langle x''|\alpha\rangle$$
$$= \int \mathrm{d}x'\mathrm{d}x''\psi^*_\beta(x'){x'}^2\delta(x'-x'')\psi_\alpha(x'') = \int \mathrm{d}x'\psi^*_\beta(x'){x'}^2\psi_\alpha(x')\,. \tag{2.70}$$

Momentum operator in position space

Let us now investigate, what the momentum operator may look like in position space. The starting point of this investigation is the definition of momentum as the generator of infinitesimal translations,

$$\left(1-\frac{i\hat{p}\mathrm{d}x'}{\hbar}\right)|\alpha\rangle = \int \mathrm{d}x'\hat{\mathcal{T}}(\mathrm{d}x')|x'\rangle\langle x'|\alpha\rangle$$
$$= \int \mathrm{d}x'|x'+\mathrm{d}x'\rangle\langle x'|\alpha\rangle = \int \mathrm{d}x'|x'\rangle\langle x'-\mathrm{d}x'|\alpha\rangle$$
$$= \int \mathrm{d}x'|x'\rangle\left(\langle x'|\alpha\rangle - \mathrm{d}x'\frac{\partial}{\partial x'}\langle x'|\alpha\rangle\right)\,. \tag{2.71}$$

Comparing both side yields

$$\hat{p}|\alpha\rangle = -i\hbar \int \mathrm{d}x'|x'\rangle \frac{\partial}{\partial x'}\langle x'|\alpha\rangle \tag{2.72}$$

or

$$\langle x'|\hat{p}|\alpha\rangle = -i\hbar \frac{\partial}{\partial x'}\langle x'|\alpha\rangle\,, \tag{2.73}$$

where the orthonormality of position eigenkets has been used. Hence, the momentum operator in the x-representation looks like

$$\langle x'|\hat{p}|x''\rangle = -i\hbar \frac{\partial}{\partial x'}\delta(x'-x'')\,. \tag{2.74}$$

From the same reasoning the following, important identity can be deduced

$$\langle\beta|\hat{p}|\alpha\rangle = -i\hbar \int \mathrm{d}x'\langle\beta|x'\rangle \frac{\partial}{\partial x'}\langle x'|\alpha\rangle = -i\hbar \int \mathrm{d}x'\psi^*_\beta(x')\frac{\partial}{\partial x'}\psi_\alpha(x')\,. \tag{2.75}$$

In the formalism developed so far, this is not a postulate, it has been derived and is a direct consequence of the properties of translation and, hence, of momentum. In the same fashion, one finally arrives at

$$\begin{aligned}
\langle x'|\hat{p}^n|\alpha\rangle &= (-i\hbar)^n \frac{\partial^n}{\partial^n x'}\langle x'|\alpha\rangle \\
\langle\beta|\hat{p}^n|\alpha\rangle &= (-i\hbar)^n \int \mathrm{d}x'\psi^*_\beta(x')\frac{\partial^n}{\partial^n x'}\psi_\alpha(x')\,.
\end{aligned} \tag{2.76}$$

Momentum space wave functions

So far, only the position basis has been employed. But there is some complete symmetry between x and p-space, apart from occasional factors of i or -1. Let us therefore now dwell on this and construct the momentum representation and its connection to position space.
As in position space, the base kets in momentum space enjoy eigenvalue and completeness relations,

$$\begin{aligned}
\hat{p}|p'\rangle &= p'|p'\rangle \\
\langle p'|p''\rangle &= \delta(p'-p'')\,.
\end{aligned} \tag{2.77}$$

Since also the momentum kets span the Hilbert space, again arbitrary state kets can be decomposed as

$$|\alpha\rangle = \int \mathrm{d}p'|p'\rangle\langle p'|\alpha\rangle = \int \mathrm{d}p'|p'\rangle\phi_\alpha(p')\,, \tag{2.78}$$

where a **wave function $\phi_\alpha(p')$ for the state $|\alpha\rangle$ in momentum space has been introduced**. Again it has the same properties and interpretation of the wave function in position space - up to the replacement of position by momentum. In particular,

$$|\langle p'|\alpha\rangle|^2\,\mathrm{d}p' = |\phi_\alpha(p')|^2\,\mathrm{d}p' \tag{2.79}$$

denotes the probability of a particle in state $|\alpha\rangle$ to be found in a momentum interval $\mathrm{d}p'$ around p'.
Let us now establish the connection of both representations. Recalling that in the case of discrete spectra and a finite number of base kets such the connection was just a base transformation with all information encoded in the scalar product of the base vectors, it can be anticipated that the important information about the connection of x- and p-space is contained in $\langle x'|p'\rangle$. This scalar product is a function of both x' and p' and usually called the transformation function related to the base transformation. To derive this function, replace first $|\alpha\rangle$ with $|p'\rangle$ in Eq. (2.73):

$$\langle x'|\hat{p}|p'\rangle = p'\langle x'|p'\rangle = -i\hbar\frac{\partial}{\partial x'}\langle x'|p'\rangle\,. \tag{2.80}$$

This is nothing but a differential equation for $\langle x'|p'\rangle$, namely

$$\left(p' + i\hbar\frac{\partial}{\partial x'}\right)\langle x'|p'\rangle = 0 \tag{2.81}$$

with the solution

$$\langle x'|p'\rangle = N\exp\left(\frac{ip'x'}{\hbar}\right)\,, \tag{2.82}$$

where N is a normalisation constant that will be determined next.
But before doing so, let us dwell on the interpretation of the finding above: Just forget for a moment that the transformation function $\langle x'|p'\rangle$ in fact is a function of both x' and p'. Let us just pretend for a moment that p' is

fixed - the transformation function can then be viewed as the probability amplitude to find a particle at x' if its momentum is p'. In other words, the transformation function, interpreted from this angle is just the wave function for p' in position space or, vice versa, the complex conjugate for x' in momentum space. These two wave functions are often called the **momentum eigenfunction in position space** or vice versa. This implies that the momentum eigenfunction in position space is just a plane wave. It is fun to note that with the formalism developed so far plane wave functions have been constructed, even without writing down anything like the Schrödinger equation.

Anyways, let us fix the normalisation now. To do so, consider

$$\begin{aligned}
\langle x'|x''\rangle &= \delta(x'-x'') \\
&= \int \mathrm{d}p' \langle x'|p'\rangle\langle p'|x''\rangle = |N|^2 \int \mathrm{d}p' \exp\left(\frac{ip'(x'-x'')}{\hbar}\right) \\
&= 2\pi\hbar\,|N|^2\,\delta(x'-x'')
\end{aligned} \tag{2.83}$$

and thus, choosing N to be purely real and positive

$$N = \frac{1}{\sqrt{2\pi\hbar}}\,. \tag{2.84}$$

Hence,

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ip'x'}{\hbar}\right)\,. \tag{2.85}$$

This can be used to connect wave functions in momentum and position space. By starting with

$$\begin{aligned}
\langle x'|\alpha\rangle &= \int \mathrm{d}p' \langle x'|p'\rangle\langle p'|\alpha\rangle \\
\langle p'|\alpha\rangle &= \int \mathrm{d}x' \langle p'|x'\rangle\langle x'|\alpha\rangle
\end{aligned} \tag{2.86}$$

immediately results in

$$\begin{aligned}
\psi_\alpha(x') &= \frac{1}{\sqrt{2\pi\hbar}} \int \mathrm{d}p' \exp\left(\frac{ip'x'}{\hbar}\right) \phi_\alpha(p')\rangle \\
\phi_\alpha(p') &= \frac{1}{\sqrt{2\pi\hbar}} \int \mathrm{d}x' \exp\left(-\frac{ip'x'}{\hbar}\right) \psi_\alpha(x')\rangle\,.
\end{aligned} \tag{2.87}$$

This pair of equations states that the **two wave functions in momentum and position space are connected by a Fourier transformation**.

2.2.2 Gaussian wave packets

Wave packets of minimum uncertainty

Consider now a physical example to illustrate the findings above, namely an entity known as **Gaussian wave packet**. Its x-space wave function is given by

$$\langle x'|\alpha\rangle = \left(\frac{1}{\pi^{1/4}\sqrt{d}}\right)\exp\left(ikx' - \frac{x'^2}{2d^2}\right)\,. \tag{2.88}$$

This is a plane wave with wave number k, modulated by a Gaussian profile centered on the origin. The probability for observing the particle vanishes very fast for $|x'| > d$: The probability density

$$\mathrm{d}\mathcal{P}(x') = |\langle x'|\alpha\rangle|^2 \tag{2.89}$$

has a Gaussian shape with width d. Computing the expectation values for x, x^2, p, and p^2 is straightforward and yields

$$\langle x\rangle = \int\limits_{-\infty}^{\infty} \mathrm{d}x'\langle\alpha|x'\rangle x'\langle x'|\alpha\rangle = \int\limits_{-\infty}^{\infty} \mathrm{d}x' x'|\langle x'|\alpha\rangle|^2 = 0 \tag{2.90}$$

for symmetry reasons. Similarly,

$$\begin{aligned}\langle x^2\rangle &= \int\limits_{-\infty}^{\infty} \mathrm{d}x'\langle\alpha|x'\rangle x'^2\langle x'|\alpha\rangle = \int\limits_{-\infty}^{\infty} \mathrm{d}x' x'^2|\langle x'|\alpha\rangle|^2 \\ &= \frac{1}{\sqrt{\pi}d}\int\limits_{-\infty}^{\infty} \mathrm{d}x' x' \exp\left(-\frac{x'^2}{d}\right) = \frac{d^2}{2}\,.\end{aligned} \tag{2.91}$$

Hence the dispersion of x is

$$\langle(\Delta x)^2\rangle = \langle x^2\rangle - \langle x\rangle^2 = \frac{d^2}{2}\,. \tag{2.92}$$

The expectation values of p and p^2 are given by

$$\begin{aligned}\langle p\rangle &= \hbar k \\ \langle p^2\rangle &= \frac{\hbar^2}{2d^2} + \hbar^2 k^2\end{aligned} \tag{2.93}$$

such that

$$\langle(\Delta p)^2\rangle = \langle p^2\rangle - \langle p\rangle^2 = \frac{\hbar^2}{2d^2}\,. \tag{2.94}$$

Taken together, this yields that for Gaussian wave packets Heisenberg's uncertainty product is given by

$$\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle = \frac{\hbar^2}{4}\,, \tag{2.95}$$

the minimal uncertainty. This is a statement independent of d. Therefore Gaussian wave packets are sometimes called **minimum uncertainty wave packets**.

Gaussian wave packets in momentum space

Going to momentum space results in an integration over x', namely

$$\begin{aligned}
\langle p'|\alpha\rangle &= \int\limits_{-\infty}^{\infty} \mathrm{d}x' \langle p'|x'\rangle\langle x'|\alpha\rangle \\
&= \left(\frac{1}{\sqrt{2\pi\hbar}}\right)\left(\frac{1}{\pi^{1/4}\sqrt{d}}\right)\int\limits_{-\infty}^{\infty} \mathrm{d}x' \exp\left(-\frac{ip'x'}{\hbar} + ikx' - \frac{x'^2}{2d^2}\right) \\
&= \left(\frac{1}{\sqrt{2\pi\hbar}}\right)\left(\frac{1}{\pi^{1/4}\sqrt{d}}\right) \\
&\qquad \int\limits_{-\infty}^{\infty} \mathrm{d}x' \exp\left[-\left(\frac{x'}{\sqrt{2}d} + \frac{ip'd - i\hbar kd}{\sqrt{2}\hbar}\right)^2 - \left(\frac{ip'd - i\hbar kd}{\sqrt{2}\hbar}\right)^2\right] \\
&= \left(\frac{1}{\sqrt{2\pi\hbar}}\right)\left(\frac{1}{\pi^{1/4}\sqrt{d}}\right)\left(\sqrt{2\pi}d\right)\exp\left[-\left(\frac{ip'd - i\hbar kd}{\sqrt{2}\hbar}\right)^2\right] \\
&= \sqrt{\frac{\sqrt{d}}{\hbar\sqrt{\pi}}}\exp\left[-\frac{(p'-\hbar k)^2 d^2}{2\hbar^2}\right]\,.
\end{aligned} \tag{2.96}$$

This momentum-space function yields an alternative method for finding the expectation values of p and p^2. Thus, the probability for finding the particle with momentum p' is centered around $\hbar k$, just as the probability for finding the particle at x' centers around 0. The widths of the two probability distributions are inversely proportional to each other. Therefore, the wider the spread in p the less the spread in x and vice versa. To illustrate this, lets go to extreme cases, namely $d \to \infty$ and $d \to 0$. In the former case, position is extremely sharp (a δ function) around 0, but the momentum has infinite width. In the latter case, momentum is extremely sharply determined, but position is essentially unknown.

2.3 Angular momentum

2.3.1 Angular momentum as generator of rotation

Finite rotations

It is known from classical physics and also easy to convince oneself that **finite rotations around the same axis commute, whereas finite rotation around different axes do not**. In a first round, let us work out, how this failure can be embedded in a more abstract framework. To this end, consider matrices R that mediating rotations in space (i.e. orthogonal or $SO(3)$ transformations),

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \tag{2.97}$$

It is well-known that these matrices R have real entries and that

$$RR^T = R^T R = 1 . \tag{2.98}$$

This property ensures that

$$\vec{x'}^2 = \vec{x}^2 . \tag{2.99}$$

To be definite, let us consider a rotation around the z-axis by some angle ϕ. At this point, some convention needs to be fixed, namely the difference between rotations affecting the physical system itself or the coordinate system beneath it. The difference between both is a change in the sign of ϕ. Here, the convention is chosen such that the angle ϕ is positive, if the rotation is counterclockwise around the positive z-axis. Then

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{2.100}$$

For rotations around the x- or y-axis, the corresponding matrices look similar, namely

$$\begin{aligned} R_x(\phi) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \\ R_y(\phi) &= \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix} . \end{aligned} \tag{2.101}$$

It is now simple to check that there is a difference in performing $R_x(\phi_x)R_y(\phi_y)$ or $R_y(\phi_y)R_x(\phi_x)$. More formally

$$[R_i(\phi_i),\, R_j(\phi_j)] \neq 0\,. \tag{2.102}$$

Infinitesimal rotations

In the spirit of the previous treatment of the momentum operator, let us focus now in particular on infinitesimal rotations around the three axis, such that the angle is ε. Then

$$\begin{aligned} R_x(\varepsilon) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\frac{\varepsilon^2}{2} & -\varepsilon \\ 0 & \varepsilon & 1-\frac{\varepsilon^2}{2} \end{pmatrix} \\ R_y(\varepsilon) &= \begin{pmatrix} 1-\frac{\varepsilon^2}{2} & 0 & \varepsilon \\ 0 & 1 & 0 \\ -\varepsilon & 0 & 1-\frac{\varepsilon^2}{2} \end{pmatrix} \\ R_z(\varepsilon) &= \begin{pmatrix} 1-\frac{\varepsilon^2}{2} & -\varepsilon & 0 \\ \varepsilon & 1-\frac{\varepsilon^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} . \end{aligned} \tag{2.103}$$

Let us now compare the effect of sequence when applying an infinitesimal rotation first around the x- and then around the y-axis or vice versa. Simple calculation shows that

$$R_x(\varepsilon)R_y(\varepsilon) = \begin{pmatrix} 1-\frac{\varepsilon^2}{2} & 0 & \varepsilon \\ \varepsilon^2 & 1-\frac{\varepsilon^2}{2} & -\varepsilon \\ -\varepsilon & \varepsilon & 1-\varepsilon^2 \end{pmatrix} \tag{2.104}$$

and

$$R_y(\varepsilon)R_x(\varepsilon) = \begin{pmatrix} 1-\frac{\varepsilon^2}{2} & \varepsilon^2 & \varepsilon \\ 0 & 1-\frac{\varepsilon^2}{2} & -\varepsilon \\ -\varepsilon & \varepsilon & 1-\varepsilon^2 \end{pmatrix}. \tag{2.105}$$

Simple comparison reveals that

$$[R_x(\varepsilon),\, R_y(\varepsilon)] = \begin{pmatrix} 0 & -\varepsilon^2 & 0 \\ \varepsilon^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_z(\varepsilon^2) - 1\,. \tag{2.106}$$

This can be generalised to

$$[R_i(\varepsilon),\, R_j(\varepsilon)] = \epsilon_{ijk} R_k(\varepsilon^2) - 1\,. \tag{2.107}$$

Introducing angular momentum in quantum mechanics

From classical mechanics, it is known that angular momentum acts as generator for infinitesimal rotations in the same way as momentum generates infinitesimal translations. In quantum mechanics, then the momentum operator generates infinitesimal translations of state kets,

$$\hat{\mathcal{T}}(\mathrm{d}\vec{x}) = 1 - \frac{i}{\hbar}\hat{\vec{p}}\cdot \mathrm{d}\vec{x}\,. \tag{2.108}$$

cf. Eq. (2.40). Therefore, it seems logical **to define the operator of angular momentum** around the axis k, $\hat{J}_k$, in such a way that the operator for an infinitesimal rotation around the kth axis by the infinitesimal angle $\mathrm{d}\phi$ is given by the Hermitian operator[7]

$$\mathcal{D}(\vec{e}_k,\, \mathrm{d}\phi) = 1 - \frac{i}{\hbar}\hat{J}_k \cdot \mathrm{d}\phi\,. \tag{2.110}$$

[7] It is important here to stress that a definition of angular momentum through relations like

$$\vec{J} = \vec{x}\times\vec{p} \tag{2.109}$$

For arbitrary orientation of the axis, $\vec{e}_n$, therefore rotations are given by

$$\mathcal{D}(\vec{e}_n,\, \mathrm{d}\phi) = 1 - \frac{i}{\hbar}\left(\vec{e}_n\; \vec{\hat{J}}_k\right) \cdot \mathrm{d}\phi\,. \tag{2.111}$$

2.3.2 Properties of the angular momentum operator

Finite rotations

As in the case of finite translations, finite rotations can be obtained by infinitely many infinitesimal rotations,

$$\mathcal{D}_z(\phi) = \lim_{N\to\infty}\left[1 - \frac{i\hat{J}_z}{\hbar}\frac{\phi}{N}\right]^N = \exp\left[-\frac{i\hat{J}_z\phi}{\hbar}\right]. \tag{2.112}$$

Now, each rotation around an axis can be obtained by such a form. The trick again is that **the explicit form of the operator $\hat{J}_k$ depends on the state ket it is acting on**. What is important however, is that irrespective of the representation in question, the rotations enjoy group properties (existence of a 1-element, existence of an inverse element for each element, associativity and closure). These group properties then fix a good part of the representation-independent fundamental properties of the $\hat{J}_k$. In this respect it makes sense to demand that the **abstract** rotations $\mathcal{D}$ in arbitrary spaces and representations have the same fundamental properties as the well-known rotations represented by $SO(3)$ matrices.

Commutator relations and consequences

One of these properties is their behaviour under commutation. Borrowing what has been found already for infinitesimal rotations in three dimensions, namely

$$[R_i(\varepsilon),\, R_j(\varepsilon)] = \epsilon_{ijk}R_k(\varepsilon^2) - 1\,. \tag{2.113}$$

known from classical mechanics is not good - defined in such a way there is no handle on how to treat spin and angular momentum in a consistent manner; after all spin has nothing to do with position or momentum, since it is an internal degree of freedom. This definition, however, is of course perfectly legitimate for the definition of **orbital angular momentum**, see below.

Translating this into the language of rotations in quantum mechanics, this implies e.g. for the x and y rotations

$$[\mathcal{D}_x(\varepsilon),\, \mathcal{D}_y(\varepsilon)] = \mathcal{D}_z(\varepsilon^2) - 1 \tag{2.114}$$

or

$$\begin{aligned}
&\left[\exp\left(-\frac{i\hat{J}_x\varepsilon}{\hbar}\right),\, \exp\left(-\frac{i\hat{J}_y\varepsilon}{\hbar}\right)\right] \\
= &\quad \left(1 - \frac{i\hat{J}_x\varepsilon}{\hbar} - \frac{\hat{J}_x^2\varepsilon^2}{\hbar^2} + \ldots\right)\left(1 - \frac{i\hat{J}_y\varepsilon}{\hbar} - \frac{\hat{J}_y^2\varepsilon^2}{\hbar^2} + \ldots\right) \\
&\quad - \left(1 - \frac{i\hat{J}_y\varepsilon}{\hbar} - \frac{\hat{J}_y^2\varepsilon^2}{\hbar^2} + \ldots\right)\left(1 - \frac{i\hat{J}_x\varepsilon}{\hbar} - \frac{\hat{J}_x^2\varepsilon^2}{\hbar^2} + \ldots\right) \\
= &\quad -\left(\frac{\hat{J}_x\hat{J}_y}{\hbar^2} - \frac{\hat{J}_y\hat{J}_x}{\hbar^2}\right)\varepsilon^2 + \mathcal{O}(\varepsilon^3) \equiv 1 - \frac{i\hat{J}_z}{\hbar}\varepsilon^2 + \mathcal{O}(\varepsilon^3) - 1\,.
\end{aligned} \tag{2.115}$$

Comparing terms of order ε^2 on both sides, this implies that

$$\left[\hat{J}_x,\, \hat{J}_y\right] = i\hbar\hat{J}_z\,. \tag{2.116}$$

The same kind of argument for the other configurations results in the general commutator relations for the operators of angular momentum,

$$\left[\hat{J}_i,\, \hat{J}_j\right] = i\hbar\epsilon_{ijk}\hat{J}_k\,. \tag{2.117}$$

The first consequence that can be derived from the commutation relations is the existence of a new operator $\hat{\vec{J}}^2$ given by

$$\hat{\vec{J}}^2 = \hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z\,. \tag{2.118}$$

It is simple to show that this new operator commutes with any of the $\hat{J}_k$,

$$\left[\hat{\vec{J}}^2,\, \hat{J}_k\right] = 0\,. \tag{2.119}$$

Let's check this for $\hat{J}_z$.

$$
\begin{aligned}
&\left[\hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z,\, \hat{J}_z\right] \\
&= \hat{J}_x\left[\hat{J}_x,\, \hat{J}_z\right] + \left[\hat{J}_x,\, \hat{J}_z\right]\hat{J}_x + \hat{J}_y\left[\hat{J}_y,\, \hat{J}_z\right] + \left[\hat{J}_y,\, \hat{J}_z\right]\hat{J}_y \\
&= \hat{J}_x(-i\hbar\hat{J}_y) + (-i\hbar\hat{J}_y)\hat{J}_x + \hat{J}_y(i\hbar\hat{J}_x) + (i\hbar\hat{J}_x)\hat{J}_y = 0\,.
\end{aligned}
\tag{2.120}
$$

Because each of the $\hat{J}_k$ commutes with $\hat{\vec{J}}^2$, but since they do not commute with each other, it is possible to simultaneously diagonalise J^2 together with only one of the $\hat{J}_k$. Usually, for this $\hat{J}_z$ is selected.

Eigenvalues of angular momentum

Since there are two operators that can simultaneously be diagonalised, the eigenstates of angular momentum are characterised by two numbers: their eigenvalues. Denoting these numbers for the time being by a and b (details to come)

$$
\begin{aligned}
\hat{\vec{J}}^2|a,\, b\rangle &= a|a,\, b\rangle \\
\hat{J}^z|a,\, b\rangle &= b|a,\, b\rangle\,.
\end{aligned}
\tag{2.121}
$$

Now it is time to determine these values. To this end, it turns out to be convenient to work with two non-Hermitian operators, dubbed **ladder operators**

$$
\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y\,. \tag{2.122}
$$

It is easy to check that they satisfy

$$
\begin{aligned}
\left[\hat{J}_+,\, \hat{J}_-\right] &= 2\hbar\hat{J}_z \\
\left[\hat{J}_z,\, \hat{J}_\pm\right] &= \pm\hbar\hat{J}_\pm \\
\left[\hat{\vec{J}}^2,\, \hat{J}_\pm\right] &= 0\,.
\end{aligned}
\tag{2.123}
$$

What is their physical meaning? to answer this consider their action on the eigenkets.

$$
\hat{J}^z\hat{J}_\pm|a,\, b\rangle = \left[\hat{J}_z,\, \hat{J}_\pm\right]|a,\, b\rangle + \hat{J}_\pm\hat{J}^z|a,\, b\rangle = (b\pm\hbar)\hat{J}_\pm|a,\, b\rangle\,, \tag{2.124}
$$

so they shift the eigenvalue of $\hat{J}_z$ by one unit of $\hbar$. This explains the term "ladder operator". In contrast, they do not change the eigenvalue of $\hat{\vec{J}}^2$.

Let us now explicitly determine the simultaneous eigenkets of $\hat{\vec{J}}^2$ and $\hat{J}_z$. Starting at some of these kets and applying one of the ladder operators n times will result in another simultaneous eigenket, this time to an eigenvalue of $\hat{J}_z$ increased or decreased by $n\hbar$. In principle one could assume that this procedure can be repeated *ad infinitum.* In practise, however, it turns out that there exists an upper limit to b, namely $a \geq b^2$. To prove this, note that

$$\begin{aligned} \hat{\vec{J}}^2 - \hat{J}_z^2 &= \hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y \\ &= \frac{1}{2}\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right) = \frac{1}{2}\left(\hat{J}_+\hat{J}_+^\dagger + \hat{J}_+^\dagger\hat{J}_+\right) . \end{aligned} \tag{2.125}$$

Now, both $\hat{J}_+\hat{J}_+^\dagger$ and $\hat{J}_+^\dagger\hat{J}_+$ must have non-negative eigenvalues, because

$$\hat{J}_+^\dagger|a,\, b\rangle \leftrightarrow \langle a,\, b|J_+ \,. \tag{2.126}$$

and hence the norm of the ket $\hat{J}_+^\dagger|a,\, b\rangle$ is given by

$$|\hat{J}_+^\dagger|a,\, b\rangle| = \sqrt{\langle a,\, b|J_+\hat{J}_+^\dagger|a,\, b\rangle} \geq 0\,. \tag{2.127}$$

Therefore

$$\langle a,\, b|\hat{\vec{J}}^2 - \hat{J}_z^2|a,\, b\rangle = \langle a,\, b|\hat{\vec{J}}^2|a,\, b\rangle - \langle a,\, b|\hat{J}_z^2|a,\, b\rangle = |a|^2 - |b|^2 > 0 \tag{2.128}$$

and thus there must be a b_{max} in dependence on the value of a. Then, applying $\hat{J}_+$ on a state with maximal b must yield 0 - since b is maximal it cannot be further increased,

$$\hat{J}_+|a,\, b_{\text{max}}\rangle = 0\,. \tag{2.129}$$

Then also

$$\hat{J}_-\hat{J}_+|a,\, b_{\text{max}}\rangle = (\hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|a,\, b_{\text{max}}\rangle = 0\,. \tag{2.130}$$

Since the eigenket by definition is not the null ket, this implies that

$$a^2 - b_{\text{max}}^2 - \hbar b_{\text{max}} = 0 \quad\Longrightarrow\quad a = b_{\text{max}}(b_{\text{max}} + h)\,. \tag{2.131}$$

In a similar fashion, there should also be a minimal b, $b_{\rm min}$, such that

$$\hat{J}_-|a,\, b_{\rm min}\rangle = 0 \tag{2.132}$$

and therefore

$$\hat{J}_+\hat{J}_-|a,\, b_{\rm min}\rangle = (\hat{\vec{J}}^2 - \hat{J}_z^2 + \hbar\hat{J}_z)|a,\, b_{\rm min}\rangle = 0\,. \tag{2.133}$$

implying that

$$a^2 - b_{\rm min}^2 - \hbar b_{\rm min} = 0 \quad\Longrightarrow\quad a = b_{\rm min}(b_{\rm min} - h)\,. \tag{2.134}$$

For symmetry reasons, $b_{\rm min} = -b_{\rm max}$ and thus

$$-b_{\rm max} \le b \le b_{\rm max}\,. \tag{2.135}$$

Now, for each given a, there must be a finite number of steps between $b_{\rm min}$ and $b_{\rm max}$, say n steps, therefore

$$b_{\rm max} = b_{\rm min} + n\hbar = -b_{\rm max} + n\hbar\,. \tag{2.136}$$

As a result

$$b_{\rm max} = \frac{n\hbar}{2}\,. \tag{2.137}$$

Conventionally, the eigenvalue j is defined by

$$j = \frac{b_{\rm max}}{\hbar} = \frac{n}{2}\,. \tag{2.138}$$

Then the maximal eigenvalue of $\hat{J}_z$ is $j\hbar$, where j is either integer or half-integer. In addition then

$$a = \hbar^2 j(j+1)\,. \tag{2.139}$$

Defining in addition $b = m\hbar$, where $m \in [-j,\, j]$ and

$$m = -j, \text{ or } \ -j+1,\ \ldots\ , \text{ or } \ j \tag{2.140}$$

allows to label the simultaneous eigenkets of $\hat{\vec{J}}^2$ and $\hat{J}_z$ by the **total angular momentum quantum number j and by the magnetic quantum number** m, such that

$$\begin{aligned} \hat{\vec{J}}^2|j,\, m\rangle &= j(j+1)\hbar^2|j,\, m\rangle \\ \hat{J}_z|j,\, m\rangle &= m\hbar|j,\, m\rangle\,. \end{aligned} \tag{2.141}$$

At that point it is important to realise that **the quantisation of angular momentum is a consequence of the commutation relations for the operators of angular momentum only, which has been inferred using merely the symmetry properties of rotations**.

Matrix elements of angular momentum operators

Now it is time to give the matrix elements of the angular momentum operators. Assuming that the basis $|j,\, m\rangle$ is orthonormal, the basic relations are

$$\begin{aligned} \langle j', m'|\hat{\vec{J}}^2|j,m\rangle &= j(j+1)\hbar^2\delta_{jj'}\delta_{mm'} \\ \langle j', m'|\hat{J}_z|j,m\rangle &= m\hbar\delta_{jj'}\delta_{mm'}\,. \end{aligned} \tag{2.142}$$

For the ladder operators, consider first that

$$\begin{aligned} &\langle j,m|\hat{J}_+^\dagger\hat{J}_+|j,m\rangle = \langle j,m|\hat{J}_-\hat{J}_+|j,m\rangle \\ = &\ \langle j,m|\left(\hat{\vec{J}}^2 - \hat{J}_z^2 - \hbar\hat{J}_z\right)|j,m\rangle = \hbar^2[j(j+1) - m^2 - m]\,. \end{aligned} \tag{2.143}$$

But also,

$$\hat{J}_+|j,m\rangle = c^+_{jm}|j,m+1\rangle\,, \tag{2.144}$$

where the constant c^+_{jm} is given by

$$|c^+_{jm}|^2 = \hbar^2[j(j+1) - m^2 - m] = \hbar^2(j-m)(j+m+1)\,. \tag{2.145}$$

Using only real, non-negative numbers for the various c^+ then implies

$$\begin{aligned} \hat{J}_+|j,m\rangle &= \hbar\sqrt{(j-m)(j+m+1)}|j,m+1\rangle \\ \hat{J}_+|j,m\rangle &= \hbar\sqrt{(j+m)(j-m+1)}|j,m-1\rangle \end{aligned} \tag{2.146}$$

and, finally,

$$\langle j', m'|\hat{J}_\pm|j,m\rangle = \hbar\sqrt{(j\pm m)(j\mp m+1)}\delta_{jj'}\delta_{m,m'\pm 1}\,. \tag{2.147}$$

2.3.3 Orbital angular momentum

Definition and commutator relations

In contrast to the definition above, defining angular momentum by being the generator of infinitesimal rotations, one may also introduce orbital angular momentum through its classical counterpart. It must be stressed here, however, that this approach is incapable of describing spin and is therefore valid only for spin-less particles or the orbital component of angular momentum. Defining the operator of orbital angular momentum by

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}} \tag{2.148}$$

or

$$\hat{L}_k = \epsilon_{ijk} \hat{x}_i \hat{p}_i \tag{2.149}$$

it is straightforward to check that the same commutation relations hold true for the $\hat{L}_k$ and for the $\hat{J}_k$, namely

$$\begin{aligned}
\left[\hat{L}_i,\, \hat{L}_j\right] &= \hbar\epsilon_{ijk}\hat{L}_k \\
\left[\hat{L}_+,\, \hat{L}_-\right] &= 2\hbar\hat{L}_z \\
\left[\hat{L}_z,\, \hat{L}_\pm\right] &= \pm\hbar\hat{L}_\pm \\
\left[\hat{\vec{L}}^2,\, \hat{L}_\pm\right] &= 0\,,
\end{aligned} \tag{2.150}$$

where the ladder operators are given by

$$\hat{L}_\pm = L_x \pm i\hat{L}_y \tag{2.151}$$

and where

$$\hat{\vec{L}}^2 = \hat{L}_x\hat{L}_x + \hat{L}_y\hat{L}_y + \hat{L}_z\hat{L}_z\,. \tag{2.152}$$

Since the last three commutation relations above are a consequence of the first, it suffices to check this first relation.

$$\begin{aligned}
&\left[\hat{L}_x,\, \hat{L}_y\right] = [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y,\, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z,\, \hat{z}\hat{p}_x] - [\hat{z}\hat{p}_y,\, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z,\, \hat{x}\hat{p}_z] + [\hat{z}\hat{p}_y,\, \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z,\, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y,\, \hat{x}\hat{p}_z] = (\hat{y}\hat{p}_x - \hat{p}_y\hat{x})\,[\hat{p}_z,\, \hat{z}] = i\hbar\hat{L}_z\,.
\end{aligned} \tag{2.153}$$

Generating rotations

In order to see whether the orbital angular momentum operator can be interpreted as the infinitesimal rotation operator, let us check how the infinitesimal rotation

$$1 - i\frac{\delta\phi}{\hbar}\hat{L}_z = 1 - i\frac{\delta\phi}{\hbar}(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) \tag{2.154}$$

acts on kets in position space. Using the fact that the momentum operators transmit infinitesimal translations, one finds

$$
\begin{aligned}
\left[1 - i\frac{\delta\phi}{\hbar}\hat{L}_z\right]|\vec{x}'\rangle &= \left[1 - i\frac{\delta\phi}{\hbar}(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x)\right]|x',\, y',\, z'\rangle \\
&= \left[1 - i\frac{\hat{p}_y}{\hbar}(\delta\phi x') + i\frac{\hat{p}_x}{\hbar}(\delta\phi y')\right]|x',\, y',\, z'\rangle \\
&= |x' - \delta\phi y',\, y' + \delta\phi x',\, z'\rangle
\end{aligned}
\tag{2.155}
$$

verifying our assumption. In fact, the position of the state ket has been altered, in full analogy what could be expected from an infinitesimal rotation around the z-axis.

Matrix elements

This can be made even more explicit by changing to spherical coordinates for the state kets under consideration

$$
\langle x',\, y',\, z'|\alpha\rangle \longrightarrow \langle r',\, \theta',\, \phi'|\alpha\rangle\,. \tag{2.156}
$$

Here the polar angle is taken w.r.t. the z-axis and the azimuthal angle describes the angle around the z-axis. Then,

$$
\left\langle x',\, y',\, z'\left|\left[1 - i\frac{\delta\phi}{\hbar}\hat{L}_z\right]\right|\alpha\right\rangle = \langle x' - \delta\phi y',\, y' + \delta\phi x',\, z'|\alpha\rangle \tag{2.157}
$$

becomes

$$
\begin{aligned}
\left\langle r',\, \theta',\, \phi'\left|\left[1 - i\frac{\delta\phi}{\hbar}\hat{L}_z\right]\right|\alpha\right\rangle &= \langle r',\, \theta',\, \phi' - \delta\phi|\alpha\rangle \\
&= \langle r',\, \theta',\, \phi'|\alpha\rangle - \delta\phi\frac{\partial}{\partial\phi'}\langle r',\, \theta',\, \phi'|\alpha\rangle\,.
\end{aligned}
\tag{2.158}
$$

This allows to finally identify

$$
\langle\vec{x}'|\hat{L}_z|\alpha\rangle = -i\hbar\frac{\partial}{\partial\phi}\langle\vec{x}'|\alpha\rangle\,. \tag{2.159}
$$

Repeating the same steps for the operators $\hat{L}_x$ and $\hat{L}_y$ yields

$$\begin{aligned}\langle\vec{x}'|\hat{L}_x|\alpha\rangle &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta}-\cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)\langle\vec{x}'|\alpha\rangle \\ \langle\vec{x}'|\hat{L}_y|\alpha\rangle &= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta}-\cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\langle\vec{x}'|\alpha\rangle\,. \end{aligned}\tag{2.160}$$

Employing this for the ladder operators, one has

$$\langle\vec{x}'|\hat{L}_\pm|\alpha\rangle = -i\hbar e^{\pm i\phi}\left(\pm i\frac{\partial}{\partial\theta}-\cot\theta\frac{\partial}{\partial\phi}\right)\langle\vec{x}'|\alpha\rangle\,. \tag{2.161}$$

Connection of $\hat{\vec{L}}^2$ with angular part of ∇^2

Then, using

$$\hat{\vec{L}}^2 = \hat{L}_z^2+\frac{1}{2}\left(\hat{L}_+\hat{L}_-+\hat{L}_-\hat{L}_+\right) \tag{2.162}$$

results in

$$\langle\vec{x}'|\hat{\vec{L}}^2|\alpha\rangle = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\phi}\right]\langle\vec{x}'|\alpha\rangle\,. \tag{2.163}$$

But, up to a factor $1/r^2$, this is just the spherical part of the Laplace operator ∇^2 expressed in spherical coordinates.
This relation can also be seen by an alternative derivation. To this end let's prove and use the following (operator) identity

$$\hat{\vec{L}}^2 = \hat{\vec{x}}^2\hat{\vec{p}}^2-(\hat{\vec{x}}\cdot\hat{\vec{p}})^2+i\hbar\hat{\vec{x}}\cdot\hat{\vec{p}}\,. \tag{2.164}$$

The proof is a bit tedious but straightforward when using components. Then

$$\begin{aligned}\hat{\vec{L}}^2 &= \sum_{ijlmk}\varepsilon_{ijk}\hat{x}_i\hat{p}_j\varepsilon_{lmk}\hat{x}_l\hat{p}_m = \sum_{ijlm}\left[\sum_k\varepsilon_{ijk}\varepsilon_{lmk}\right]\hat{x}_i\hat{p}_j\hat{x}_l\hat{p}_m \\ &= \sum_{ijlm}\left[\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}\right]\hat{x}_i\hat{p}_j\hat{x}_l\hat{p}_m = \sum_{ij}\left[\hat{x}_i\hat{p}_j\hat{x}_i\hat{p}_j-\hat{x}_i\hat{p}_j\hat{x}_j\hat{p}_i\right] \\ &= \sum_{ij}\left[\hat{x}_i\left(\hat{x}_i\hat{p}_j-\delta_{ij}\hbar\right)\hat{p}_j-\hat{x}_i\hat{p}_j\left(\hat{p}_i\hat{x}_j+\delta_{ij}\hbar\right)\right] \\ &= \sum_{ij}\left[\hat{x}_i^2\hat{p}_j^2-\hat{x}_i\hat{p}_i\left(\hat{x}_j\hat{p}_j-\hbar\delta_{jj}\right)\right] = \hat{\vec{x}}^2\hat{\vec{p}}^2-(\hat{\vec{x}}\cdot\hat{\vec{p}})^2+i\hbar\hat{\vec{x}}\cdot\hat{\vec{p}}\,. \end{aligned}\tag{2.165}$$

Considering now this expression sandwiched between $\langle\vec{x}'|$ and $|\alpha\rangle$ is simple after identifying

$$\begin{aligned}\langle\vec{x}'|\hat{\vec{x}}\cdot\hat{\vec{p}}|\alpha\rangle &= \langle\vec{x}'|\sum_i \hat{x}_i\cdot\hat{p}_i|\alpha\rangle = \sum_i x_i'\langle\vec{x}'|\hat{p}_i|\alpha\rangle \\ &= \sum_i x_i'\left(-i\hbar\frac{\partial}{\partial x_i'}\right)\langle\vec{x}'|\alpha\rangle = -i\hbar\vec{x}_i'\cdot\vec{\nabla}'\langle\vec{x}'|\alpha\rangle = -i\hbar r\frac{\partial}{\partial r}\langle\vec{x}'|\alpha\rangle\end{aligned} \tag{2.166}$$

and, in a similar way,

$$\begin{aligned}&\langle\vec{x}'|(\hat{\vec{x}}\cdot\hat{\vec{p}})^2|\alpha\rangle \\ &= -\hbar^2 r\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\langle\vec{x}'|\alpha\rangle\right) = -\hbar^2\left(r^2\frac{\partial}{\partial r^2}\langle\vec{x}'|\alpha\rangle + r\frac{\partial}{\partial r}\langle\vec{x}'|\alpha\rangle\right).\end{aligned} \tag{2.167}$$

Taken together, this implies that

$$\langle\vec{x}'|\hat{\vec{L}}^2|\alpha\rangle = r^2\langle\vec{x}'|\hat{\vec{p}}^2|\alpha\rangle + \hbar^2\left(r^2\frac{\partial}{\partial r^2}\langle\vec{x}'|\alpha\rangle + 2r\frac{\partial}{\partial r}\langle\vec{x}'|\alpha\rangle\right) \tag{2.168}$$

and thus the kinetic energy (the term with the squared momentum) of the state $|\alpha\rangle$ in position space can be written as

$$\begin{aligned}\frac{1}{2m}\langle\vec{x}'|\hat{\vec{p}}^2|\alpha\rangle &= -\frac{\hbar^2}{2m}\nabla'^2\langle\vec{x}'|\alpha\rangle \\ &= -\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial r^2}\langle\vec{x}'|\alpha\rangle + \frac{2}{r}\frac{\partial}{\partial r}\langle\vec{x}'|\alpha\rangle - \frac{1}{\hbar^2 r^2}\langle\vec{x}'|\hat{\vec{L}}^2|\alpha\rangle\right).\end{aligned} \tag{2.169}$$

Comparing this with the well-known expression for the kinetic energy in spherical coordinates yields the result above.

Chapter 3

Dynamical situations

So far, the dynamical evolution of quantum mechanical systems with time has not been discussed at all. This blind spot of the lecture will be cured in this chapter, which is devoted exclusively to the dynamics development of state kets and/or the operators corresponding to the observables. The first important point in this respect is to realise that **in Quantum Mechanics, time is just a parameter - it is by no means an operator!** Therefore, it is absolute non-sensical to talk about a time-operator in the same way as talking about the position operator[1].

3.1 Schrödinger equation

3.1.1 Time evolution

The time-evolution operator

The first question to be answered is, how a state ket does change with time. Suppose a system has been prepared to be in state $|\alpha_0\rangle$ at some time t_0. In general, at later times, the system will be in a different state $|\alpha\rangle$. In order

[1]In the historical development of quantum mechanics, however, E. Schrödinger and L. de Broglie both were guided by some kind of covariant analogy of position and momentum on one hand and time and energy on the other hand. But when looking at quantum mechanics as it appears now as a mature textbook subject, there is no trace left of this idea: Position and time are treated on different footing. A more symmetric treatment leads to relativistic quantum mechanics, to quantum field theory.

to exhibit this let us write

$$|\alpha\rangle = |\alpha(t)\rangle \quad \text{with} \quad |\alpha_0\rangle = |\alpha(t_0)\rangle \,. \tag{3.1}$$

Assuming time to be a continuous parameter, it is natural to expect

$$\lim_{t\to 0} |\alpha(t)\rangle = |\alpha(t_0)\rangle = |\alpha_0\rangle \,. \tag{3.2}$$

Our basic task in the following is to study the time evolution of a state ket, i.e.

$$|\alpha_0\rangle = |\alpha(t_0)\rangle \quad \overset{\text{timeevolution}}{\longrightarrow} \quad |\alpha(t)\rangle \,. \tag{3.3}$$

Similar to the case of translation the starting point is to define a **time-evolution operator** $\hat{\mathcal{U}}(t,\, t_0)$ with

$$|\alpha(t)\rangle = \hat{\mathcal{U}}(t,\, t_0)|\alpha(t_0)\rangle \,. \tag{3.4}$$

What are the properties of this operator?

- **Unitarity:**
  First of all, it is natural to expect unitarity to hold true, i.e. to demand that

  $$\hat{\mathcal{U}}^\dagger(t,\, t_0)\hat{\mathcal{U}}(t,\, t_0) = 1 \,. \tag{3.5}$$

  In order to see this demanding that the state kets at times t are normalised to one, if they are normalised at preparation time t_0 implies

  $$\begin{aligned} 1 &= \langle\alpha(t)|\alpha(t)\rangle = \langle\alpha(t_0)|\hat{\mathcal{U}}^\dagger(t,\, t_0)\hat{\mathcal{U}}(t,\, t_0)|\alpha(t_0)\rangle \\ &\equiv \langle\alpha(t_0)|\alpha(t_0)\rangle = 1 \end{aligned} \tag{3.6}$$

  and thus yields unitarity of the time-evolution operator.

  This requirement of **unitarity is quite synonymous to demanding the conservation of probability**: Assume the state $|\alpha_0\rangle$ to be a linear combination of orthogonal eigenvectors of some observable $\hat{A}$,

  $$|\alpha_0\rangle = |\alpha(t_0)\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha(t_0)\rangle = \sum_{a'} |a'\rangle c_{a'}(t_0) \tag{3.7}$$

where the time dependence of the coefficients is explicitly displayed. Of course, they are normalised to one, if both the base vectors $|a'\rangle$ and the state $|\alpha_0$ is normalised,

$$1 = \sum_{a'} |c_{a'}(t_0)|^2 \,. \tag{3.8}$$

At some later time t, the decomposition of the state vector looks like

$$|\alpha(t)\rangle = \sum_{a'} |a'\rangle\langle a'|\alpha(t)\rangle = \sum_{a'} |a'\rangle c_{a'}(t) \,, \tag{3.9}$$

where

$$1 = \sum_{a'} |c_{a'}(t)|^2 \,. \tag{3.10}$$

In general now, the moduli of the individual coefficients may differ,

$$|c_{a'}(t_0)| \neq |c_{a'}(t)| \,, \tag{3.11}$$

since the system may have a completely different decomposition w.r.t. the individual states. However, the sum of the squares remains identical to one.

To exemplify this, consider a spin-1/2 system, subjected to a uniform magnetic field in z-direction. To be specific, let's start with the system to be in the $S_x = +1/2$ state, i.e. the spin being oriented along the positive x-axis. As time goes on, the spin will experience some precession in the x-y-plane (see below), and thus the probability for observing the system in the $S_x = +1/2$-state will not remain unity. Instead, for $t > t_0$ there will be a finite probability to also observe $S_x = -1/2$. However, the sum of probabilities to observe the system in either $S_x = +1/2$ or $S_x = -1/2$ is always unity: The system must have either of both with certainty.

- **Composition:**
  A further natural property of the time-evolution operator is to demand that it is possible to compose a time-evolution from, say, t_0 to t_2 by two such operators between t_0 and t_1 and then from t_1 to t_2:

$$\hat{\mathcal{U}}(t_2,\, t_0) = \hat{\mathcal{U}}(t_2,\, t_1)\hat{\mathcal{U}}(t_1,\, t_0) \quad (t_2 \geq t_1 \geq t_0) \,. \tag{3.12}$$

It is important here to realise that the sequence of operators acts on a state ket to the right, meaning that the state ket is first evolved from t_0 to t_1 before a second evolution brings him to t_2.

- **Infinitesimal time-evolution:**
  Similar to the case of the translation operator, it makes sense to consider infinitesimal time-evolution. In this case,

$$\hat{\mathcal{U}}(t_0 + \mathrm{d}t,\, t_0)|\alpha(t_0)\rangle = |\alpha(t_0 + dt)\rangle \tag{3.13}$$

  and, for obvious reasons,

$$\lim_{\mathrm{d}t\to 0} \hat{\mathcal{U}}(t_0 + \mathrm{d}t,\, t_0) = 1\,. \tag{3.14}$$

  As in the case of translation, the difference between the identity 1 and $\hat{\mathcal{U}}(t_0 + \mathrm{d}t,\, t_0)$ is expected to be of first order in $\mathrm{d}t$.

Constructing the infinitesimal time-evolution operator

Similar to the case of translations, the infinitesimal time-evolution operator will be constructed first. In this case,

$$\hat{\mathcal{U}}(t_0 + \mathrm{d}t,\, t_0) = 1 - i\hat{\Omega}\mathrm{d}t\,, \tag{3.15}$$

where, in full analogy to the translation case, the operator $\hat{\Omega}$ is Hermitian,

$$\hat{\Omega}^\dagger = \hat{\Omega}\,, \tag{3.16}$$

i.e., it has real eigenvalues. It is simple to prove that this operator has the correct limit and also satisfies the other requirements above, which have been formulated for finite time evolution[2]. Obviously, $\hat{\Omega}$ (or its eigenvalues) should have reciprocal dimension of time, or dimension of frequency. Recalling that angular frequency ω of, say a photon, corresponds to an energy of

$$E = \hbar\omega\,, \tag{3.17}$$

the **Planck-Einstein relation**, it is natural to expect the operator to be related to energies. Adding in that in classical mechanics, the **Hamilton function** (or **Hamiltonian**) is the generator of time-evolution, the choice

$$\hbar\hat{\Omega} = \hat{H} \tag{3.18}$$

[2] The corresponding proofs are in full analogy to the translation case, just consider factors of $\mathrm{d}t$ up to first order.

seems logical. To sum up, the infinitesimal time-evolution operator can be written as

$$\hat{\mathcal{U}}(t_0 + \mathrm{d}t,\, t_0) = 1 - \frac{i}{\hbar}\hat{H}\mathrm{d}t\,, \tag{3.19}$$

where the Hamiltonian (**Hamilton operator**) is assumed to be Hermitian. This is quite a sensible choice, since measurable energies of any system are expected to be real numbers. One may also wonder, whether the $\hbar$ showing up here is the same as in the translation case. This question can be answered by comparing equations of motion in quantum mechanics, to be derived later, and in classical theory. It turns out that unless both $\hbar$'s are identical, it is impossible to derive something like

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = \frac{\vec{p}}{m} \tag{3.20}$$

as the classical limit of quantum mechanics.

3.1.2 A first glance at the Schrödinger equation

Schrödinger equation for the time-evolution operator

The reasoning so far allows to derive the fundamental differential equation fixing the time-evolution operator. Exploiting the composition property of $\hat{\mathcal{U}}$ allows to write

$$\hat{\mathcal{U}}(t + \mathrm{d}t,\, t_0) = \hat{\mathcal{U}}(t + \mathrm{d}t,\, t)\hat{\mathcal{U}}(t,\, t_0) = \left(1 - \frac{i}{\hbar}\hat{H}\mathrm{d}t\right)\hat{\mathcal{U}}(t,\, t_0). \tag{3.21}$$

Therefore,

$$\hat{\mathcal{U}}(t + \mathrm{d}t,\, t_0) - \hat{\mathcal{U}}(t,\, t_0) = -\frac{i}{\hbar}\hat{H}\mathrm{d}t\hat{\mathcal{U}}(t,\, t_0)\,, \tag{3.22}$$

or, in differential form,

$$\frac{i}{\hbar}\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{U}}(t,\, t_0) = \hat{H}\hat{\mathcal{U}}(t,\, t_0)\,. \tag{3.23}$$

This is the **Schrödinger equation for the time-evolution operator**.

From there it is only a short step to the **Schrödinger equation for a state ket**: Just multiply with $|\alpha(t_0)\rangle$ from the right to obtain

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{U}}(t,\,t_0)|\alpha(t_0)\rangle = \hat{H}\hat{\mathcal{U}}(t,\,t_0)|\alpha(t_0)\rangle \tag{3.24}$$

or

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\alpha(t)\rangle = \hat{H}|\alpha(t)\rangle\,. \tag{3.25}$$

In principle, this is the equation to solve. In practice, however, it is as good to know how the time-evolution operator acts on the initial state ket $|\alpha(t_0)\rangle$. Then, rather than bothering with the Schrödinger equation for state kets, the time evolution operator can directly be applied to obtain the state ket at any time $t > t_0$. However, in order to derive formal solutions for the Schrödinger equation for the time-evolution operator, there are three cases to consider:

1. The Hamilton operator is explicitly independent of time:
   This means that changing the parameter t does not imply any change in $\hat{H}$. Then, the differential equation - the Schrödinger equation - can easily be solved and the time-evolution operator reads

   $$\hat{\mathcal{U}}(t,\,t_0) = \exp\left[-\frac{i\hat{H}(t-t_0)}{\hbar}\right]\,. \tag{3.26}$$

   That this in fact is a solution can be proven by expanding the exponential in a series as

   $$\exp\left[-\frac{i\hat{H}(t-t_0)}{\hbar}\right] = 1 - \frac{i\hat{H}(t-t_0)}{\hbar} + \frac{1}{2}\frac{i^2\hat{H}^2(t-t_0)^2}{\hbar^2} + \ldots\,. \tag{3.27}$$

   The partial derivative w.r.t. t therefore reads

   $$\begin{aligned}\frac{\partial}{\partial t}\exp\left[-\frac{i\hat{H}(t-t_0)}{\hbar}\right] &= -\frac{i\hat{H}}{\hbar} + 2\frac{1}{2}\frac{i^2\hat{H}^2(t-t_0)}{\hbar^2} + \ldots\\ &= -\frac{i\hat{H}}{\hbar}\left[1 - \frac{i\hat{H}(t-t_0)}{\hbar} + \ldots\right]\end{aligned} \tag{3.28}$$

   and hence the expression above satisfies the differential equation.

An alternative way to construct the time evolution operator is to understand it as a compound of infinitely many infinitesimal time-evolution steps, in a fashion quite similar to the way the finite translation operator has been constructed before.

2. The Hamilton operator depends on time, but the Hamilton operators at different times commute:
   The formal solution to this case reads

$$\hat{\mathcal{U}}(t,\, t_0) = \exp\left[-\frac{i}{\hbar}\int\limits_{t_0}^{t} \mathrm{d}t'\hat{H}(t')\right] . \tag{3.29}$$

   The proof is similar to the one of case 1 - just a bit more tedious.

3. Hamilton operators at different times do not commute:
   The formal solution to this case is known as **Dyson series** and reads

$$\hat{\mathcal{U}}(t,\, t_0) = 1 + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int\limits_{t_0}^{t}\mathrm{d}t_1 \int\limits_{t_0}^{t_1}\mathrm{d}t_2 \ldots \int\limits_{t_0}^{t_{n-1}}\mathrm{d}t_n \hat{H}(t_1)\hat{H}(t_2)\ldots\hat{H}(t_n)\, . \tag{3.30}$$

In most cases - and for sure in all cases considered during this course - only case 1 is of practical interest. Therefore throughout this course, $\hat{H}$ is assumed to be time-independent.

Energy eigenkets

Considering the case of time-independent Hamiltonians only, the task now is to find a viable solution for the time-evolution of an arbitrary state ket $|\alpha\rangle$. This problem can be simplified by using a suitable base to expand $|\alpha\rangle$ in. A particularly useful set of base kets are the eigenkets of an operator $\hat{A}$, which commutes with the Hamiltonian,

$$\left[\hat{A},\, \hat{H}\right] = 0\, . \tag{3.31}$$

In this case, the eigenkets $|a'\rangle$ of $\hat{A}$ are simultaneous eigenkets also of $\hat{H}$ - therefore they are often called **energy eigenkets**, and their eigenvalues thus are the **energy eigenvalues**,

$$\hat{H}|a'\rangle = E_{a'}|a'\rangle\,. \tag{3.32}$$

The time evolution operator has a particularly simple form in this base:

$$\begin{aligned}
\exp\left(-\frac{i\hat{H}t}{\hbar}\right) &= \sum_{a',a''}|a''\rangle\left\langle a''\middle|\exp\left(-\frac{i\hat{H}t}{\hbar}\right)\middle|a'\right\rangle\langle a'| \\
&= \sum_{a'}|a'\rangle\exp\left(-\frac{iE_{a'}t}{\hbar}\right)\langle a'|\,,
\end{aligned} \tag{3.33}$$

where the fact has been used that in analytical functions of an operator this operator can be replaced by corresponding eigenvalues, if eigenkets are acting on it.
However, written in this form, the solution for the time-evolution of any initial state ket $|\alpha(t)\rangle$ is quite simple, if its expansion in terms of energy eigenkets is known. In this case (and setting the initial time $t_0 = 0$), the expansion reads

$$|\alpha(t_0 = 0)\rangle = \sum_{a'}|a'\rangle\langle a'|\alpha(t_0 = 0)\rangle = \sum_{a'}c_{a'}(t_0 = 0)|a'\rangle \tag{3.34}$$

and the time-evolution of the initial state is thus given by

$$|\alpha(t)\rangle = \exp\left(-\frac{i\hat{H}t}{\hbar}\right)|\alpha(0)\rangle = \sum_{a'}|a'\rangle\exp\left(-\frac{iE_{a'}t}{\hbar}\right)\langle a'|\alpha(0)\rangle\,, \tag{3.35}$$

leading to a time-evolution of the expansion coefficients namely

$$c_{a'}(t_0) \longrightarrow c_{a'}(t) = \exp\left(-\frac{iE_{a'}t}{\hbar}\right)c_{a'}(0)\,. \tag{3.36}$$

This is maybe a bit surprising: **During time-evolution only the phases of the (energy-eigenket-) expansion coefficients of an initial state ket change**. Note however that the **oscillation frequencies of the different energy eigenmodes may vary**.

An instructive special case is where the initial state ket $|\alpha\rangle$ already is an energy eigenket $|a\rangle$. In this case, the expansion coefficients are given by

$$c_{a'}(0) = \delta_{aa'} \tag{3.37}$$

and therefore, at a later time

$$|\alpha(t)\rangle = \exp\left(-\frac{iE_a t}{\hbar}\right)|a\rangle\,, \tag{3.38}$$

so, if the system initially is in an simultaneous eigenstate of $\hat{A}$ and $\hat{H}$, it remains so at all times. The only thing that happens is a phase modulation, characterised by $E_a/\hbar$. In this sense, **observables, which are compatible with the Hamiltonian are constants of motion**. This will be encountered again, when discussing the Heisenberg picture of quantum dynamics.
However, to some extent, quantum dynamics thus boils down to finding all observables that are compatible with the Hamiltonian and evaluating their eigenvalues and finding their eigenkets. Once this is done, the initial state merely has to be expanded in terms of these eigenkets, and then the time-evolution operator can happily be applied. This step then only leads to a phase modulation of the individual modes.

Time-evolution of expectation values

In a next step let us study the time-evolution of an expectation value of an observable $\hat{B}$. To ease this task, let's assume that at $t = 0$ the initial state is one of the eigenstates $|a\rangle$ of $\hat{A}$, where $\hat{A}$ is compatible with the Hamiltonian. Then, at later times,

$$|a(t)\rangle = \hat{\mathcal{U}}(t,\, 0)|a(0)\rangle\,. \tag{3.39}$$

The expectation value of $\hat{B}$ in such a state is given by

$$\begin{aligned}
\langle\hat{B}\rangle_{a(t)} &= \langle a(0)|\hat{\mathcal{U}}^\dagger(t,\, 0)\hat{B}\hat{\mathcal{U}}(t,\, 0)|a(0)\rangle \\
&= \left\langle a(0)\left|\exp\left(\frac{iE_a t}{\hbar}\right)\hat{B}\exp\left(-\frac{iE_a t}{\hbar}\right)\right|a(0)\right\rangle \\
&= \langle a(0)|\hat{B}|a(0)\rangle = \langle\hat{B}\rangle_{a(0)}\,.
\end{aligned} \tag{3.40}$$

Hence, the expectation value of any observable in an energy eigenstate is independent of time. For this reason, energy eigenstates are also known as **stationary states**.

Of course, the situation is more interesting when the expectation value with respect to a superposition of energy eigenstates, which does not need to be a stationary state. With

$$|\alpha(0)\rangle = \sum_{a'} c_{a'}(0)|a'\rangle \tag{3.41}$$

the expectation value of $\hat{B}$ reads

$$\begin{aligned}
\langle \hat{B} \rangle_{\alpha(t)} &= \langle \alpha(0)|\hat{\mathcal{U}}^{\dagger}(t,\,0)\hat{B}\,\hat{\mathcal{U}}(t,\,0)|\alpha(0)\rangle \\
&= \left[\sum_{a'} \langle a'|c^*_{a'}(0)\exp\left(\frac{iE_{a'}t}{\hbar}\right)\right]\hat{B}\left[\sum_{a''} c_{a''}(0)\exp\left(\frac{-iE_{a''}t}{\hbar}\right)|a''\rangle\right] \\
&= \sum_{a',a''} c^*_{a'}(0)c_{a''}(0)\exp\left(\frac{-i(E_{a''}-E_{a'})t}{\hbar}\right)\langle a'|\hat{B}|a''\rangle\,. \qquad (3.42)
\end{aligned}$$

Simple example: Spin precession

Let us, for illustrative purposes, consider an extremely simple example system, namely the Hamiltonian for a spin-1/2 system subject to an external magnetic field $\vec{B}$. If the magnetic moment for the system is $e\hbar/2mc$ (e is the charge, m the mass of the system, c is the speed of light), the Hamiltonian reads

$$\hat{H} = -\frac{e}{mc}\hat{\vec{S}}\cdot\vec{B}\,. \tag{3.43}$$

Let's assume that the magnetic field is a static and uniform field in z-direction, $\vec{B} = B\vec{e}_z$. Then

$$\hat{H} = -\frac{eB}{mc}\hat{S}_z\,. \tag{3.44}$$

From this it is simple to see that the Hamiltonian and $\hat{S}_z$ just differ by a constant factor and that they thus commute. Therefore, the S_z-eigenstates are also energy eigenstates with corresponding energy eigenvalues

$$E_{\pm} = \mp\frac{e\hbar B}{2mc} \quad \text{for} \quad S_z = |\pm\rangle\,. \tag{3.45}$$

Defining a frequency ω such that it is just the difference of both energy eigenvalues,

$$\omega = -\frac{e\hbar B}{mc} \tag{3.46}$$

the Hamiltonian can be rewritten as

$$\hat{H} = \hbar\omega \hat{S}_z\,. \tag{3.47}$$

All information on time-evolution is now encoded in the time-evolution operator

$$\hat{\mathcal{U}}(t,\, t_0) = \exp\left(-\frac{i\hat{H}(t-t_0)}{\hbar}\right) = \exp\left(-\frac{i\hbar\omega\hat{S}_z(t-t_0)}{\hbar}\right)\,. \tag{3.48}$$

In order to apply this to an initial state, it must be expanded in terms of the S_z-eigenkets $|\pm\rangle$, which, as said before, are now also energy eigenstates. If, at $t = 0$ the system is in a state characterised by

$$|\alpha(0)\rangle = c_+(0)|+\rangle + c_-(0)|-\rangle\,, \tag{3.49}$$

it will be in the state

$$|\alpha(t)\rangle = \exp\left(-\frac{i\omega t}{2}\right) c_+(0)|+\rangle + \exp\left(\frac{i\omega t}{2}\right) c_-(0)|-\rangle \tag{3.50}$$

at a later time t. Here, the eigenvalue equation

$$\hat{H}|\pm\rangle = \pm\frac{\hbar\omega}{2}|\pm\rangle \tag{3.51}$$

has been used.
Let us now consider two cases:

1. $\alpha(0)\rangle = |+\rangle$:
   In this case, $c_+(0) = 1$ and $c_-(0) = 0$. At later times, c_- will still vanish, and

   $$c_+(t) = \exp\left(-\frac{i\omega t}{2\hbar}\right)\,. \tag{3.52}$$

   Hence the system will remain in the $|+\rangle$-state.

2. $\alpha(0)\rangle = |S_x = +\rangle$:
   In this case,

   $$c_+(0) = c_-(0) = \frac{1}{\sqrt{2}}\,. \tag{3.53}$$

It is straightforward to find the probabilities for the system to be in the $|S_x = \pm\rangle$ state at later times t:

$$
\begin{aligned}
&|\langle\alpha(t) = S_x;\, \pm|\alpha(0)\rangle|^2 = |\langle S_x;\, \pm|\alpha(t)\rangle|^2 \\
&= \left|\left[\frac{1}{\sqrt{2}}\left(\langle +| \pm \langle -|\right)\right]\right| \left\{\frac{1}{\sqrt{2}}\left[\exp\left(-\frac{i\omega t}{2}\right)|+\rangle + \exp\left(\frac{i\omega t}{2}\right)|-\rangle\right]\right\}\Bigg|^2 \\
&= \left|\frac{1}{2}\left[\exp\left(-\frac{i\omega t}{2}\right) \pm \exp\left(\frac{i\omega t}{2}\right)\right]\right|^2 \\
&= \begin{cases} \cos^2\frac{\omega t}{2} & \text{for } S_x+ \\ \sin^2\frac{\omega t}{2} & \text{for } S_x- \;. \end{cases}
\end{aligned}
\tag{3.54}
$$

What does this imply? Let's answer this with a simple example, $|\alpha(0)\rangle = |S_x+\rangle$. In this case, although initially the system is in the $s_x = +1/2$-state, the magnetic field causes it to rotate around the z-axis. Thus, there is a finite probability to find the system in the $s_x = -1/2$-state at later times. However, the sum of both probabilities always equals unity. This is no surprise, since at every time the system must have a spin w.r.t. the x-axis. However, considering the expectation value of $\hat{S}_x$ it is straightforward to calculate it to be

$$
\langle\hat{S}_x\rangle_t = \left(\frac{\hbar}{2}\right)\cos^2\frac{\omega t}{2} + \left(\frac{-\hbar}{2}\right)\sin^2\frac{\omega t}{2} = \left(\frac{\hbar}{2}\right)\cos(\omega t)\,. \tag{3.55}
$$

In other words, when measured the spin in x-direction will oscillate with time with an angular frequency corresponding to the energy difference of the two energy eigenstates. A similar calculation shows that

$$
\langle\hat{S}_y\rangle_t = \left(\frac{\hbar}{2}\right)\sin(\omega t) \tag{3.56}
$$

and that

$$
\langle\hat{S}_z\rangle_t = 0 \tag{3.57}
$$

implying that the spin of the system experiences some precession around the z-axis.

3.2 Schrödinger vs. Heisenberg

So far, the time-evolution of a system has been understood as an effect affecting the state ket describing the system, in contrast operators have been taken

as constant in time. This approach to the time development of a system is known as the **Schrödinger picture**. There is however another formulation of dynamical situations in the framework of quantum mechanics, known as the **Heisenberg picture**. In this formulation, the state kets are constant in time and the operators change.
The connection of both pictures becomes transparent when considering the quantum-mechanical measurement process: There, expectation values of operators are central, i.e. something like

$$\langle \hat{O} \rangle_\alpha = \langle \alpha | \hat{O} | \alpha \rangle \, . \tag{3.58}$$

Assuming time-dependent state kets, the expectation value will evolve like

$$\begin{aligned} \langle \hat{O} \rangle_{\alpha(t)} = \langle \alpha(t) | \hat{O} | \alpha(t) \rangle &= \Big(\langle \alpha(t_0) \hat{\mathcal{U}}^\dagger(t,\, t_0) \Big) \, |\hat{O}| \, \Big(\hat{\mathcal{U}}(t,\, t_0) \alpha(t_0) \rangle \Big) \\ &= \langle \alpha(t_0) | \hat{\mathcal{U}}^\dagger(t,\, t_0) \hat{O} \hat{\mathcal{U}}(t,\, t_0) | \alpha(t_0) \rangle \, . \end{aligned} \tag{3.59}$$

It is quite easy now to understand the basic idea when going from the Schrödinger to the Heisenberg picture: Just drop the dependence of the state kets $|\alpha\rangle$ on the initial time t_0 and assume that the operator $\hat{O}$ becomes time-dependent,

$$\hat{O}(t) = \hat{\mathcal{U}}^\dagger(t,\, t_0) \hat{O} \hat{\mathcal{U}}(t,\, t_0) = \hat{\mathcal{U}}^\dagger(t,\, t_0) \hat{O}(t_0) \hat{\mathcal{U}}(t,\, t_0) \tag{3.60}$$

This implies that the constant operator $\hat{O}$ in the Schrödinger picture is used as the operator $\hat{O}(t_0)$ at the initial time t_0 in the Heisenberg picture.

3.2.1 Unitary operators

Base transformations

This re-interpretation of state-kets and operators as being time dependent or not is related to the question of on which object the time-evolution operator $\hat{\mathcal{U}}(t,\, t_0)$ acts. However, before continuing, it is worthwhile to consider unitary transformations induced by unitary operators such as, e.g., the time-evolution operator $\hat{\mathcal{U}}(t,\, t_0)$ or the translation operator $\hat{T}(\mathrm{d}\vec{x})$.
In order to see the effect of such an unitary operator and to understand the formal significance of them, it is important to realise that applying them is equivalent to a **change of basis** (or, to spell it out even better, a **change of representation**). In order to see this, consider an orthonormal set of

eigenkets $\{|a'\rangle\}$ of an operator $\hat{A}$. Obviously, in the base spanned by these eigenkets, $\hat{A}$ is diagonal. In order to change from this $\hat{A}$-diagonal base to a base spanned by the set of orthonormal eigenkets $\{|b'\rangle\}$ of an operator $\hat{B}$, an unitary operator $\hat{U}$ can be employed.
To phrase it more formally: Given two complete sets of orthonormal base kets, $\{|a'_i\rangle\}$ and $\{|b'_i\rangle\}$, there exists a unitary operator $\hat{U}$ such that

$$|b'_i\rangle = U|a'_i\rangle \quad \forall\, i\,. \tag{3.61}$$

This can be proven by explicit construction. First of all, it must be realised that $\hat{U}$ can be written as

$$\hat{U} = \sum_i |b'_i\rangle\langle a'_i|\,. \tag{3.62}$$

Orthonormality of the $|a'_i\rangle$ guarantees that applying $\hat{U}$ on any of them yields the corresponding $|b'_i\rangle$. Using that

$$\hat{U}^\dagger = \sum_i |a'_i\rangle\langle b'_i| \tag{3.63}$$

immediately implies that

$$\hat{U}^\dagger\hat{U} = \sum_{i,j} |a'_i\rangle\langle b'_i|b'_j\rangle\langle a'_j| = \sum_{i,j} |a'_i\rangle\langle a'_j|\delta_{ij} = \sum_i |a'_i\rangle\langle a'_i = 1\,. \tag{3.64}$$

Hence, $\hat{U}$ is unitary, i.e.

$$\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = 1\,. \tag{3.65}$$

Matrix elements of the transformation operator

The matrix elements of this operator $\hat{U}$ in the old $\hat{A}$-diagonal basis read

$$\langle\hat{U}\rangle_{ij} = \langle a'_i|\hat{U}|a'_j\rangle = \langle a'_i|b'_j\rangle\,. \tag{3.66}$$

Hence, the matrix elements of the transformation operator are just the inner products of old base bras and new base kets[3].

[3] This can be exemplified by the rotation matrix in three dimensions, which rotates the old set of unit base vectors $\vec{e}_x$, $\vec{e}_y$, and $\vec{e}_z$ into a new set of such vectors, $\vec{e}_{x'}$, $\vec{e}_{y'}$, and $\vec{e}_{z'}$. It is simple to check that the corresponding transformation matrix has the form

$$\hat{R} = \begin{pmatrix} \vec{e}_x\cdot\vec{e}_{x'} & \vec{e}_x\cdot\vec{e}_{y'} & \vec{e}_x\cdot\vec{e}_{z'} \\ \vec{e}_y\cdot\vec{e}_{x'} & \vec{e}_y\cdot\vec{e}_{y'} & \vec{e}_y\cdot\vec{e}_{z'} \\ \vec{e}_z\cdot\vec{e}_{x'} & \vec{e}_z\cdot\vec{e}_{y'} & \vec{e}_z\cdot\vec{e}_{z'} \end{pmatrix}\,. \tag{3.67}$$

How can this be used to represent arbitrary kets $|\alpha\rangle$ in the new rather than in the old base? The answer is quite straightforward. In the old base, $|\alpha\rangle$ is represented by

$$|\alpha\rangle = \sum_i |a'_i\rangle\langle a'_i|\alpha\rangle\,, \tag{3.68}$$

pictorially this amounts to a vector with the ith entry given by the inner product (scalar product) of the ith base vector and $|\alpha\rangle$, i.e. $\langle a'_i|\alpha\rangle$. Correspondingly, in the new base, the representation is given by

$$|\alpha\rangle = \sum_i |b'_i\rangle\langle b'_i|\alpha\rangle\,. \tag{3.69}$$

The expansion coefficients, i.e. the entries in the vector representation of the ket, can be related through

$$\langle b'_i|\alpha\rangle = \sum_j \langle b'_i|a'_j\rangle\langle a'_j\alpha\rangle = \sum_j \langle a'_i|\hat{U}^\dagger|a'_j\rangle\langle a'_j\alpha\rangle\,. \tag{3.70}$$

In matrix notation, this is identical to the statement that the column vector representing $|\alpha\rangle$ in the new base can be obtained from the one in the old base by multiplying form the left with $\hat{U}^\dagger$:

$$(\text{new}) = \hat{U}^\dagger(\text{old})\,. \tag{3.71}$$

It is important to note, however, that **under a change of base it is only the representation that changes and not the state ket itself**.

Similarity transformations

In a similar fashion, matrix elements of operators $\hat{O}$ in the old and new base can be related:

$$\begin{aligned}\langle b'_i|\hat{O}|b'_j\rangle &= \sum_{l,k}\langle b'_i|a'_l\rangle\langle a'_l|\hat{O}|a'_k\rangle\langle a'_k|b'_j\rangle \\ &= \sum_{l,k}\langle a'_i|\hat{U}^\dagger|a'_l\rangle\langle a'_l|\hat{O}|a'_k\rangle\langle a'_k|\hat{U}|a'_j\rangle\,,\end{aligned} \tag{3.72}$$

or, in matrix form

$$\hat{O}' = \hat{U}^\dagger\hat{O}\hat{U}\,, \tag{3.73}$$

where the primed and unprimed operator refers to the different representation. This transformation is known as **similarity transformation** in linear algebra.
It is simple to show that under similarity transformations the trace of the operator, defined as the sum of its diagonal elements,

$$\mathrm{Tr}(\hat{O}) = \sum_i \langle a'_i|\hat{O}|a'_i\rangle \,, \tag{3.74}$$

remains constant,

$$\begin{aligned}\mathrm{Tr}(\hat{O}) &= \sum_i \langle a'_i|\hat{O}|a'_i\rangle = \sum_{i,j,k} \langle a'_i|b'_j\rangle\langle b'_j|\hat{O}|b'_k\rangle\langle b'_k|a'_i\rangle \\ &= \sum_{j,k} \langle b'_k|b'_j\rangle\langle b'_j|\hat{O}|b'_k\rangle = \sum_{j,k} \delta_{kj}\langle b'_j|\hat{O}|b'_k\rangle = \sum_j \langle b'_j|\hat{O}|b'_j\rangle \,.\end{aligned} \tag{3.75}$$

In a similar fashion, the following identities can be asserted:

$$\begin{aligned}\mathrm{Tr}(\hat{X}\hat{Y}) &= \mathrm{Tr}(\hat{Y}\hat{X}) \\ \mathrm{Tr}(\hat{X}\hat{Y}\hat{Z}) &= \mathrm{Tr}(\hat{Y}\hat{Z}\hat{X}) = \mathrm{Tr}(\hat{Z}\hat{X}\hat{Y}) \\ \mathrm{Tr}(\hat{U}^\dagger\hat{X}\hat{U}) &= \mathrm{Tr}(\hat{X}) \\ \mathrm{Tr}(|a'_i\rangle\langle a'_j|) &= \delta_{ij} \,.\end{aligned} \tag{3.76}$$

3.2.2 Heisenberg equation of motion

Interpretation of unitary transformations

Having formalised how a change of basis can be achieved through application of unitary operators, the focus will now shift again on time-evolution as a result of applying the unitary time-evolution operator. Keeping in mind that in general unitary transformations change the kets but not their scalar products,

$$\langle\beta|\alpha\rangle \to \langle\beta'|\alpha'\rangle = \Big(\langle\beta\hat{U}^\dagger\Big) \,|\, \Big(\hat{U}\alpha\rangle\Big) = \langle\beta|\hat{U}^\dagger\hat{U}|\alpha\rangle = \langle\beta|\alpha\rangle \,, \tag{3.77}$$

it can be inferred how $\langle\beta|\hat{O}|\alpha\rangle$ must change, namely

$$\langle\beta|\hat{O}|\alpha\rangle \to \langle\beta'|\hat{O}|\alpha'\rangle = \Big(\langle\beta\hat{U}^\dagger\Big) \,|\hat{O}|\, \Big(\hat{U}\alpha\rangle\Big) = \langle\beta|\hat{U}^\dagger\hat{O}\hat{U}|\alpha\rangle \,. \tag{3.78}$$

This mathematical identity suggests two approaches to unitary transformation like the ones considered so far (time-evolution and translation):

- state kets change, $|\alpha\rangle \to \hat{U}|\alpha\rangle$, operators remain unchanged, or
- operators change, $\hat{O} \to \hat{U}^\dagger\hat{O}\hat{U}$, state kets remain unchanged.

It is instructive now to take a step back and remember that in classical physics, there are no state kets, although concepts like time-evolution, translation etc. are important there as well. This is possible there, because these operations are applied to observable quantities such as position, momentum and the like. In quantum mechanics, these quantities are represented by Hermitian operators. This justifies the suspicion that a closer connection with classical mechanics emerges when choosing approach 2 above.
This can be illustrated by a simple example, namely by the infinitesimal translations $\hat{\mathcal{T}}(\mathrm{d}\vec{x}')$. So far, they have been treated according to approach 1, i.e. they have been understood as acting on the state kets rather than on the position operator:

$$\begin{aligned} |\alpha\rangle &\to |\alpha'\rangle = \left(1 - \frac{i\hat{\vec{p}}\cdot\mathrm{d}\vec{x}'}{\hbar}\right)|\alpha\rangle \\ \hat{\vec{x}} &\to \hat{\vec{x}}' = \hat{\vec{x}}\,. \end{aligned} \tag{3.79}$$

In contrast, in approach 2,

$$\begin{aligned} |\alpha\rangle &\to |\alpha'\rangle = |\alpha\rangle \\ \hat{\vec{x}} &\to \hat{\vec{x}}' = \left(1 + \frac{i\hat{\vec{p}}\cdot\mathrm{d}\vec{x}'}{\hbar}\right)\hat{\vec{x}}\left(1 - \frac{i\hat{\vec{p}}\cdot\mathrm{d}\vec{x}'}{\hbar}\right) \\ &= \hat{\vec{x}} + \frac{i}{\hbar}\left[\hat{\vec{p}}\cdot\mathrm{d}\vec{x},\,\hat{\vec{x}}\right] = \hat{\vec{x}} + \mathrm{d}\vec{x}\,. \end{aligned} \tag{3.80}$$

It is then simple to show that the expectation value in both approaches lead to the same result, namely

$$\langle\hat{\vec{x}}\rangle \to \langle\hat{\vec{x}}'\rangle = \langle\hat{\vec{x}}\rangle + \langle\mathrm{d}\hat{\vec{x}}\rangle\,. \tag{3.81}$$

Heisenberg equation of motion

In order to derive the Heisenberg equation of motion, let us introduce the following notation: Quantities in the Schrödinger picture are indicated by a superscript $^{(S)}$; the state kets in this picture depend on time and are therefore denoted by $|\alpha^{(S)}(t)\rangle$, whereas operators in this picture do not depend on time, and are therefore denoted by $\hat{A}^{(S)}$. In contrast, in the Heisenberg picture the state kets do not depend on time and are thus denoted by $|\alpha^{(H)}\rangle$, whereas the operators do depend on time and are denoted by $\hat{A}^{(H)}(t)$. To connect both pictures assume that at some time $t_0 = 0$ the quantities in both pictures coincided,

$$
\begin{aligned}
|\alpha^{(H)}\rangle &= |\alpha^{(S)}(t=0)\rangle \\
\hat{A}^{(H)}(t=0) &= \hat{A}^{(S)}\,.
\end{aligned}
\tag{3.82}
$$

The time evolution of relevant quantities is now given by

$$
|\alpha^{(S)}(t)\rangle = \hat{\mathcal{U}}(t,\,0)|\alpha^{(S)}(0)\rangle \tag{3.83}
$$

and

$$
\hat{A}^{(H)}(t) = \hat{\mathcal{U}}^\dagger(t,\,0)\hat{A}^{(H)}(0)\hat{\mathcal{U}}(t,\,0)\,, \tag{3.84}
$$

respectively, a dramatic difference! However, the expectation values in both pictures coincide:

$$
\begin{aligned}
\langle\alpha^{(S)}(t)|\hat{A}^{(S)}|\alpha^{(S)}(t)\rangle &= \langle\alpha^{(S)}(t=0)|\hat{\mathcal{U}}^\dagger(t,\,0)\hat{A}^{(S)}\hat{\mathcal{U}}(t,\,0)|\alpha^{(S)}(t=0)\rangle \\
&= \langle\alpha^{(H)}|\hat{\mathcal{U}}^\dagger(t,\,0)\hat{A}^{(H)}(t=0)\hat{\mathcal{U}}(t,\,0)|\alpha^{(H)}\rangle \\
&= \langle\alpha^{(H)}|\hat{A}^{(H)}(t)|\alpha^{(H)}\rangle\,.
\end{aligned}
\tag{3.85}
$$

Using the above relation between operators in both pictures,

$$
\hat{A}^{(H)}(t) = \hat{\mathcal{U}}^\dagger(t,\,0)\hat{A}^{(H)}(t=0)\hat{\mathcal{U}}(t,\,0) = \hat{\mathcal{U}}^\dagger(t,\,0)\hat{A}^{(S)}\hat{\mathcal{U}}(t,\,0) \tag{3.86}
$$

and differentiating w.r.t. time yields

$$
\begin{aligned}
\frac{\mathrm{d}\hat{A}^{(H)}}{\mathrm{d}t} &= \frac{\mathrm{d}\hat{\mathcal{U}}^\dagger}{\mathrm{d}t}\hat{A}^{(S)}\hat{\mathcal{U}} + \hat{\mathcal{U}}^\dagger\hat{A}^{(S)}\frac{\mathrm{d}\hat{\mathcal{U}}}{\mathrm{d}t} \\
&= \frac{i}{\hbar}\hat{\mathcal{U}}^\dagger\hat{H}\hat{\mathcal{U}}\hat{\mathcal{U}}^\dagger\hat{A}^{(S)}\hat{\mathcal{U}} - \frac{i}{\hbar}\hat{\mathcal{U}}^\dagger\hat{A}^{(S)}\hat{\mathcal{U}}\hat{\mathcal{U}}^\dagger\hat{H}\hat{\mathcal{U}} \\
&= \frac{i}{\hbar}\hat{\mathcal{U}}^\dagger\hat{H}\hat{\mathcal{U}}\hat{A}^{(H)} - \frac{i}{\hbar}\hat{A}^{(H)}\hat{\mathcal{U}}^\dagger\hat{H}\hat{\mathcal{U}} = \frac{i}{\hbar}\left[\hat{\mathcal{U}}^\dagger\hat{H}\hat{\mathcal{U}},\,\hat{A}^{(H)}\right]\,.
\end{aligned}
\tag{3.87}
$$

Here the explicit independence on time of the operator in the Schrödinger picture has been used and the Schrödinger equation for the time evolution operator, Eq. (3.23), and its complex conjugate

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{U}} &= -\frac{i}{\hbar}\hat{H}\hat{\mathcal{U}} \\
\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{U}}^\dagger &= \frac{i}{\hbar}\hat{H}\hat{\mathcal{U}}^\dagger
\end{aligned}
\tag{3.88}
$$

have been employed.
Since the Hamiltonian originally was introduced in the Schrödinger picture, it is tempting to write

$$\hat{H}^{(H)} = \hat{\mathcal{U}}^\dagger \hat{H}^{(S)} \hat{\mathcal{U}} \,. \tag{3.89}$$

In most applications, however, the Hamiltonian is explicitly independent of time and the time evolution operator is given by Eq. (3.26) and the two operators thus commute. As a result of this,

$$\hat{H}^{(H)} = \hat{\mathcal{U}}^\dagger \hat{H}^{(S)} \hat{\mathcal{U}} = \hat{\mathcal{U}}^\dagger \hat{\mathcal{U}} \hat{H}^{(S)} = \hat{H}^{(S)} := \hat{H} \,. \tag{3.90}$$

Therefore it seems correct to write

$$\frac{\mathrm{d}\hat{A}^{(H)}}{\mathrm{d}t} = -\frac{i}{\hbar}\left[\hat{A}^{(H)},\, \hat{H}\right] \,. \tag{3.91}$$

This form of the equation is known as the **Heisenberg equation of motion**. Comparing this with the classical equation of motion, expressed through the Poisson bracket,

$$\frac{\mathrm{d}A(p,\,q)}{\mathrm{d}t} = \left[A(p,\,q),\, H(p,\,q)\right] \,, \tag{3.92}$$

it can be recognized that Dirac's quantisation rule

$$[\,,\,]_{\text{classical}} \longrightarrow -\frac{i}{\hbar}[\,,\,]_{,\text{quantum}} \tag{3.93}$$

leads to the correct behavior of the operators in the Heisenberg picture. It is in this sense that the Heisenberg picture is closer to classical physics than the Schrödinger picture. It should be stressed however that **the Heisenberg equation of motion is valid independent of whether the quantum mechanical operator in question has a classical counterpart or not**.

Simple example: Free particle

Independent of whether the Schrödinger or the Heisenberg picture is employed, it is paramount to be able to construct a Hamiltonian to work with. For systems with classical analogues, it seems logical to assume that the Hamiltonian has the same form as the classical Hamilton function, apart from replacing positions and momenta with the corresponding operators.

With this assumption the correct classical equations of motion in the classical limit can be reproduced. As a rule, whenever there are some ambiguities due to non-commuting operators, these are resolved by the requirement that the Hamilton operator is Hermitian. This implies that the classical product xp is replaced by

$$xp \longrightarrow \frac{1}{2}(\hat{x}\hat{p} + \hat{p}\hat{x})\,. \tag{3.94}$$

If there is no classical analogue the only way towards a correct Hamiltonian is guessing: Just try any structure until the results are in agreement with observation.

Let us first try to apply the Heisenberg equation of motion to a free particle of mass m. The Hamiltonian reads

$$\hat{H} = \frac{1}{2m}\,\hat{\vec{p}}^{\,2} = \frac{1}{2m}\,\left(\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2\right)\,. \tag{3.95}$$

Working in the Heisenberg picture implies that its the operators which depend on time, leading to

$$\frac{\mathrm{d}\hat{p}_i}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{p}_i,\,\hat{H}] = 0\,. \tag{3.96}$$

Thus, for a free particle, the momentum is a constant of motion, as in classical mechanics. The Heisenberg equation of motion for the position operator reads

$$\frac{\mathrm{d}\hat{x}_i}{\mathrm{d}t} = -\frac{i}{\hbar}[\hat{x}_i,\,\hat{H}] = -\frac{i}{\hbar}\frac{1}{2m}[\hat{x}_i,\,\hat{p}_i^2] = -\frac{i}{\hbar}\frac{1}{2m}2i\hbar\hat{p}_i = \frac{\hat{p}_i}{m}\,. \tag{3.97}$$

This implies that

$$\hat{x}_i(t) = \hat{x}_i(0) + t\frac{\hat{p}_i(0)}{m}\,, \tag{3.98}$$

again as in classical mechanics. It is important to note in this context that in the Heisenberg picture, position operators at different times do not commute with each other, although they commute at equal times. This can be seen already with this simple example: Although

$$[\hat{x}_i(0),\,\hat{x}_j(0)] = 0 \tag{3.99}$$

$$[\hat{x}_i(t),\, \hat{x}_i(0)] = \frac{t}{m}[\hat{x}_i(0),\, \hat{p}_i(0)] = -\frac{i\hbar t}{m}\,. \tag{3.100}$$

Applying the uncertainty relation to these operators thus yields

$$\langle(\Delta\hat{x}_i)^2\rangle_t\langle(\Delta\hat{x}_i)^2\rangle_{t=0} \geq \frac{\hbar^2 t^2}{4m^2}\,. \tag{3.101}$$

3.3 The harmonic oscillator

The harmonic oscillator is one of the most important problems in quantum mechanics. Often it is used to illustrate all basic concepts and methods in a first, non-trivial example. In addition, it has application in nearly all branches of modern physics, from spectroscopy, over solid state physics, nuclear structure, quantum optics, quantum statistics to quantum field theory. From a historical perspective it was Planck's idea of associating discrete units of energy with radiation oscillators that ultimately lead to quantum physics.

3.3.1 Diagonalising the Hamilton operator

Creation and annihilation operators

Let us start the discussion following Dirac's elegant idea. The basic Hamiltonian of the simple harmonic oscillator in one dimension reads

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2\hat{x}^2}{2}\,. \tag{3.102}$$

Here ω is the angular frequency of the classical oscillator. It is related to the mass and the spring constant k of Hooke's law through $\omega = \sqrt{k/m}$. Of course, the operators $\hat{x}$ and $\hat{p}$ are Hermitian: They correspond to observable quantities. It is convenient to introduce two non-Hermitian operators,

$$\begin{aligned}
\hat{a} &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right) \\
\hat{a}^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - \frac{i\hat{p}}{m\omega}\right)\,,
\end{aligned} \tag{3.103}$$

which are known as **annihilation and creation operator**, respectively. The reasoning behind these names will become apparent soon. Using the canonical commutation relation of position and momentum, cf. Eq. (2.54),

$$[\hat{x},\, \hat{p}] = i\hbar \tag{3.104}$$

implies

$$[\hat{a},\, \hat{a}^\dagger] = \frac{m\omega}{2\hbar}\left[\left(\hat{x} + \frac{i\hat{p}}{m\omega}\right),\, \left(\hat{x} - \frac{i\hat{p}}{m\omega}\right)\right] = -\frac{i}{2\hbar}\left([\hat{x},\, \hat{p}] - [\hat{p},\, \hat{x}]\right) = 1\,. \tag{3.105}$$

Introducing also the obviously Hermitian **number operator**

$$\hat{N} = \hat{a}^\dagger \hat{a}\,, \tag{3.106}$$

it is straightforward to show that

$$\hat{N} = \left(\frac{m\omega}{2\hbar}\right)\left(\hat{x}^2 + \frac{\hat{p}^2}{m^2\omega^2} + \frac{i[\hat{x},\, \hat{p}]}{m\omega}\right) = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2}\,. \tag{3.107}$$

Thus, the Hamiltonian can be expressed by the number operator as

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right)\,, \tag{3.108}$$

and it is simple to see that both operators commute. Therefore, $\hat{N}$ and $\hat{H}$ can simultaneously be diagonalised, and energy eigenvalues and eigenkets can be labelled by the eigenvalues n of $\hat{N}$,

$$\hat{N}|n\rangle = n|n\rangle\,. \tag{3.109}$$

In the next section it will be shown that the n must be nonnegative integers. Here it suffices to state that the energy eigenvalues read

$$\hat{H}|n\rangle = E_n|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle\,. \tag{3.110}$$

Therefore, in contrast ot the classical case, where the energy given by the maximal amplitude can take any continuous value, in quantum mechanics the energy of the oscillator must take any of the discrete values

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)\,. \tag{3.111}$$

More on the number operator

Let us now turn again to the number operator and its eigenvalues. First, consider the commutators of the number with the creation and annihilation operator:

$$\begin{aligned}\left[\hat{N},\,\hat{a}\right] &= \left[\hat{a}^\dagger\hat{a},\,\hat{a}\right] = \hat{a}^\dagger\left[\hat{a},\,\hat{a}\right] + \left[\hat{a}^\dagger,\,\hat{a}\right]\hat{a} = -\hat{a} \\ \left[\hat{N},\,\hat{a}^\dagger\right] &= \left[\hat{a}^\dagger\hat{a},\,\hat{a}^\dagger\right] = \hat{a}^\dagger\left[\hat{a},\,\hat{a}^\dagger\right] + \left[\hat{a}^\dagger,\,\hat{a}^\dagger\right]\hat{a}^\dagger = \hat{a}^\dagger\,. \end{aligned} \tag{3.112}$$

As a result,

$$\begin{aligned}\hat{N}\hat{a}^\dagger|n\rangle &= \left(\left[\hat{N},\,\hat{a}^\dagger\right] + \hat{a}^\dagger\hat{N}\right)|n\rangle = (1+n)\hat{a}^\dagger|n\rangle \\ \hat{N}\hat{a}|n\rangle &= \left(\left[\hat{N},\,\hat{a}\right] + \hat{a}\hat{N}\right)|n\rangle = (-1+n)\hat{a}|n\rangle\,. \end{aligned} \tag{3.113}$$

These relations imply that, if $|n\rangle$ is an eigenket of $\hat{N}$ with eigenvalue n, then $\hat{a}^\dagger|n\rangle$ and $\hat{a}|n\rangle$ are also eigenkets of $\hat{N}$, but with eigenvalue $n\pm1$, respectively. It is because of this increase and decrease of the eigenvalue by one unit of $\hbar\omega$ that the term creation or annihilation operator is appropriate.
The equation above implies that $\hat{a}|n\rangle$ and $|n-1\rangle$ are the same up to a multiplicative constant. Hence, let us write

$$\hat{a}|n\rangle = c|n-1\rangle\,, \tag{3.114}$$

where c is a constant complex number. Requiring that both states $|n\rangle$ and $|n-1\rangle$ are normalised, this implies that

$$\begin{aligned}\langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{N}|n\rangle &= |c|^2\langle n-1||n-1\rangle \\ n &= |c|^2\,. \end{aligned} \tag{3.115}$$

Conventionally c is taken to be real and positive and therefore

$$\begin{aligned}\hat{a}|n\rangle &= \sqrt{n}|n-1\rangle \\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle\,. \end{aligned} \tag{3.116}$$

Suppose now that the annihilation operator is applied over and over to a state $|n\rangle$.

$$\begin{aligned}\hat{a}|n\rangle &= \sqrt{n}|n-1\rangle \\ \hat{a}^2|n\rangle &= \sqrt{n(n-1)}|n-2\rangle \\ \hat{a}^3|n\rangle &= \sqrt{n(n-1)(n-2)}|n-3\rangle \\ &\vdots \end{aligned} \tag{3.117}$$

At some point, this series is bound to terminate, if n initially was a positive integer. This is just because then one of the factors $(n-m)$ under the square-root will equal zero and then the overall result will be zero. It could be argued now that rather than starting with an integer positive n one could start with a positive non-integer. But this would lead to eigenkets with negative value of n. But there is also the positivity requirement for the norm of any state, i.e.

$$0 \leq \langle n|\hat{a}^\dagger \hat{a}|n\rangle = n \,, \tag{3.118}$$

implying that n cannot be negative! This forces us to assume that the sequence above **must** terminate for $n = 0$ and therefore the allowed values of n are the non-negative integers.

Energy eigenstates

Since the smallest possible number of n is zero, the ground state, i.e. the state of least energy, of the simple harmonic oscillator is given by

$$E_0 = \frac{1}{2}\hbar\omega \,. \tag{3.119}$$

Successively applying the creation operator on the ground state yields a series of excited states,

$$\begin{aligned}
|1\rangle &= \hat{a}^\dagger|0\rangle \\
|2\rangle &= \frac{\hat{a}^\dagger}{\sqrt{2}}|1\rangle = \frac{1}{\sqrt{2!}}(\hat{a}^\dagger)^2|0\rangle \\
|3\rangle &= \frac{\hat{a}^\dagger}{\sqrt{3}}|2\rangle = \frac{1}{\sqrt{3!}}(\hat{a}^\dagger)^3|0\rangle \\
&\vdots \\
|n\rangle &= \frac{\hat{a}^\dagger}{\sqrt{n}}|n-1\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle \,.
\end{aligned} \tag{3.120}$$

In this way, simultaneous eigenkets of $\hat{N}$ and $\hat{H}$ can be constructed iteratively from the ground state.

To proceed in constructing wave functions for these states it is mandatory to have the representation of the various operators in the base spanned by the energy eigenkets. For the creation and annihilation operators this is straightforward:

$$\langle n'|\hat{a}|n\rangle = \sqrt{n}\delta_{n',n-1}\ , \quad \langle n'|\hat{a}^\dagger|n\rangle = \sqrt{n+1}\delta_{n',n+1}\,, \tag{3.121}$$

and it shouldn't come as a surprise that they have entries only in one of the secondary diagonals. Inverting the definition of these two operators yields

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\,(\hat{a}+\hat{a}^\dagger) \quad \text{and} \quad \hat{p} = -i\sqrt{\frac{m\hbar\omega}{2}}\,(\hat{a}-\hat{a}^\dagger)\,. \tag{3.122}$$

Therefore, their matrix elements in the energy eigenket base reads

$$\begin{aligned} \langle n'|\hat{x}|n\rangle &= \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right) \\ \langle n'|\hat{p}|n\rangle &= -i\sqrt{\frac{m\hbar\omega}{2}}\left(\sqrt{n}\delta_{n',n-1} - \sqrt{n+1}\delta_{n',n+1}\right)\,. \end{aligned} \tag{3.123}$$

Notice that, unsurprisingly, neither of the operators is diagonal in the $\hat{N}$-diagonal representation; after all neither of them commutes with $\hat{N}$.
Let us now construct wave functions in position space. Starting with the ground state, defined by,

$$\hat{a}|0\rangle = 0\,, \tag{3.124}$$

or, in the x-representation,

$$\langle x'|\hat{a}|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\left\langle x'\left|\left(\hat{x}+\frac{i}{m\omega}\,\hat{p}\right)\right|0\right\rangle = 0\,, \tag{3.125}$$

and recalling Eq. (2.75), yields a differential equation for the ground state wave function:

$$\left(x' + \frac{\hbar^2}{m^2\omega^2}\frac{\mathrm{d}}{\mathrm{d}x'}\right)\langle x'|0\rangle = 0\,. \tag{3.126}$$

Introducing a characteristic length scale of the oscillator, x_0 through

$$x_0^2 = \frac{\hbar^2}{m^2\omega^2} \tag{3.127}$$

it can be shown that the normalised solution of the differential equation above reads

$$\langle x'|0\rangle = \left(\frac{1}{\pi^{1/4}\sqrt{x_0}}\right)\exp\left[-\frac{1}{2}\left(\frac{x'}{x_0}\right)^2\right]. \tag{3.128}$$

The wave functions of the excited states can be obtained through

$$\begin{aligned}
\langle x'|1\rangle &= \langle x'|\hat{a}^\dagger|0\rangle = \left(\frac{1}{\sqrt{2}\,x_0}\right)\left(x' - x_0^2\frac{\mathrm{d}}{\mathrm{d}x'}\right)\langle x'|0\rangle \\
\langle x'|2\rangle &= \frac{1}{\sqrt{2!}}\langle x'|(\hat{a}^\dagger)^2|0\rangle = \frac{1}{\sqrt{2!}}\left(\frac{1}{\sqrt{2}\,x_0}\right)^2\left(x' - x_0^2\frac{\mathrm{d}}{\mathrm{d}x'}\right)^2\langle x'|0\rangle \\
&\ \vdots \\
\langle x'|n\rangle &= \frac{1}{\sqrt{n!}}\langle x'|(\hat{a}^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}\,x_0}\right)^n\left(x' - x_0^2\frac{\mathrm{d}}{\mathrm{d}x'}\right)^n\langle x'|0\rangle
\end{aligned} \tag{3.129}$$

leading to

$$\langle x'|n\rangle = \left(\frac{1}{\pi^{1/4}\sqrt{2^n n!}}\right)\left(\frac{1}{x_0^{n+1/2}}\right)\left(x' - x_0^2\frac{\mathrm{d}}{\mathrm{d}x'}\right)^n\exp\left[-\frac{1}{2}\left(\frac{x'}{x_0}\right)^2\right]. \tag{3.130}$$

3.3.2 Dynamical properties

Expectation values

It is instructive to check the ground-state expectation values for $\hat{x}^2$ and $\hat{p}^2$. From

$$\hat{x}^2 = \frac{\hbar}{2m\omega}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger\right) \tag{3.131}$$

only the last term survives being sandwiched between the ground state bra and ket. Thus,

$$\langle\hat{x}^2\rangle_0 = \frac{\hbar}{2m\omega} = \frac{x_0^2}{2}. \tag{3.132}$$

Likewise

$$\langle \hat{p}^2 \rangle_0 = \frac{m\hbar\omega}{2} \,. \tag{3.133}$$

Therefore, the expectation values of kinetic and potential energy are identical,

$$\begin{aligned}
\left\langle \frac{\hat{p}^2}{2m} \right\rangle_0 &= \frac{\hbar\omega}{4} = \frac{1}{2}\,\langle \hat{H} \rangle \\
\left\langle \frac{\hat{m}\omega^2 x^2}{2} \right\rangle_0 &= \frac{\hbar\omega}{4} = \frac{1}{2}\,\langle \hat{H} \rangle \,,
\end{aligned} \tag{3.134}$$

a quantum mechanical realisation of the virial theorem of classical mechanics. In addition,

$$\langle \hat{x} \rangle_0 = \langle \hat{p} \rangle_0 = 0 \,, \tag{3.135}$$

which also holds true for the excited states.
Anyways, taking together these findings implies that

$$\begin{aligned}
\langle (\Delta\hat{x})^2 \rangle_0 &= \langle \hat{x}^2 \rangle_0 = \frac{\hbar}{2m\omega} \\
\langle (\Delta\hat{p})^2 \rangle_0 &= \langle \hat{p}^2 \rangle_0 = \frac{m\hbar\omega}{2}
\end{aligned} \tag{3.136}$$

and thus

$$\langle (\Delta\hat{x})^2 \rangle_0 \langle (\Delta\hat{p})^2 \rangle_0 = \frac{\hbar^2}{4} \,, \tag{3.137}$$

implying that the ground state is a state of minimal uncertainty. This comes as no surprise, after all its wave function has Gaussian shape. For the excited states, the uncertainty products are given by

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

Time development

Before leaving the harmonic oscillator for good, let us dwell on the time evolution of the state kets or observables such as position or momentum.

Everything that has been done so far is for some fixed time, say $t = 0$. At that time either the operators in he Heisenberg picture have been fixed or the state kets in the Schrödinger picture. However, in this section, the Heisenberg picture will be chosen, i.e. the state kets (labelled by the number n) are regarded as stationary. In contrast, the operators $\hat{x}$, $\hat{p}$, $\hat{a}$, or $\hat{a}^\dagger$ are all time-dependent, although this dependence has not been denoted so far. Then, the Heisenberg equations of motion governing their time-evolution read

$$\begin{aligned} \frac{\mathrm{d}\hat{p}}{\mathrm{d}t} &= \frac{i}{\hbar}[\hat{H},\,\hat{p}] = -m\omega^2\hat{x} \\ \frac{\mathrm{d}\hat{x}}{\mathrm{d}t} &= \frac{i}{\hbar}[\hat{H},\,\hat{x}] = \frac{\hat{p}}{m}\,, \end{aligned} \tag{3.139}$$

cf. Eq. (3.91). This pair of coupled differential equations is equivalent to two uncoupled differrential equations for the creation and annihilation operator, namely

$$\frac{\mathrm{d}\hat{a}}{\mathrm{d}t} = \sqrt{\frac{m\omega}{2\hbar}}\left(\frac{\hat{p}}{m} - i\omega\hat{x}\right) = -i\omega\hat{a} \tag{3.140}$$

and, analogously,

$$\frac{\mathrm{d}\hat{a}^\dagger}{\mathrm{d}t} = i\omega\hat{a}^\dagger\,. \tag{3.141}$$

The solutions for these two equations can be obtained straightforwardly and read

$$\hat{a}(t) = e^{-i\omega t}\hat{a}(0) \quad \text{and} \quad \hat{a}^\dagger(t) = e^{i\omega t}\hat{a}^\dagger(0)\,. \tag{3.142}$$

It should not come as an surprise that this shows that both the number operator $\hat{N}$ and the Hamiltonian $\hat{H}$ are thus independent of time - this can be checked by simply multiplying the two operators and cancelling the time-dependence residing in the exponentials. The connection of the two operators with position and momentum can be used to obtain the time evolution of the latter. Just identify

$$\begin{aligned} \hat{a}(t) &= \hat{x}(t) + \frac{i}{m\omega}\hat{p}(t) = e^{-i\omega t}\hat{x}(0) + \frac{ie^{-i\omega t}}{m\omega}\hat{p}(0) \\ \hat{a}^\dagger(t) &= \hat{x}(t) - \frac{i}{m\omega}\hat{p}(t) = e^{i\omega t}\hat{x}(0) - \frac{ie^{i\omega t}}{m\omega}\hat{p}(0)\,. \end{aligned} \tag{3.143}$$

Combining these two equations finally yields

$$\begin{aligned}\hat{x}(t) &= \hat{x}(0)\cos(\omega t) + \frac{1}{m\omega}\hat{p}(0)\sin(\omega t)\\ \hat{p}(t) &= -m\omega\hat{x}(0)\sin(\omega t) + \hat{p}(0)\cos(\omega t)\,.\end{aligned}\tag{3.144}$$

These two equations look the same as their classical counterparts, and, obviously, both position and momentum oscillate just as their classical analogues. The same results can also be obtained in an laternative fashion. Rather than solving the Heisenberg E.o.M., which is sysnonymous to sloving the coupled differential equations above, one may use the time-evolution operator and directly evaluate

$$\hat{x}(t) = \exp\left(\frac{i\hat{H}t}{\hbar}\right)\,\hat{x}(0)\,\exp\left(-\frac{i\hat{H}t}{\hbar}\right)\,.\tag{3.145}$$

In order to do so an extremely important identity, known as the **Baker-Hausdorff lemma**, will be used. It reads

$$\begin{aligned}\exp\left(i\lambda\hat{G}\right)\,\hat{A}\,\exp\left(-i\lambda\hat{G}\right) &= \hat{A} + i\lambda[\hat{G},\,\hat{A}] + \left(\frac{i^2\lambda^2}{2!}\right)[\hat{G},\,[\hat{G},\,\hat{A}]] + \dots\\ &\quad \dots + \left(\frac{i^n\lambda^n}{n!}\right)[\hat{G},\,[\hat{G},\,\dots[\hat{G},\,\hat{A}]\dots]] + \dots\,,\end{aligned}\tag{3.146}$$

where $\hat{G}$ is Hermitian and λ is a real parameter. To apply this formula, one thus has to calculate the commutators of position or momentum with the Hamiltonian,

$$\begin{aligned}\left[\hat{H},\,\hat{x}(0)\right] &= -\frac{i\hbar}{m}\,\hat{p}(0)\\ \left[\hat{H},\,\hat{p}(0)\right] &= i\hbar m\omega^2\,\hat{x}(0)\,.\end{aligned}\tag{3.147}$$

Inserting this yields

$$\begin{aligned}
&\exp\left(\frac{i\hat{H}t}{\hbar}\right)\,\hat{x}(0)\,\exp\left(-\frac{i\hat{H}t}{\hbar}\right)\\
&= \hat{x}(0)+\left(\frac{it}{\hbar}\right)\left[\hat{H},\,\hat{x}(0)\right]+\left(\frac{i^2t^2}{2!\hbar^2}\right)\left[\hat{H},\,\left[\hat{H},\,\hat{x}(0)\right]\right]+\ldots\\
&= \hat{x}(0)-\left(\frac{i\hbar}{m}\cdot\frac{it}{\hbar}\right)\hat{p}(0)-\left(i\hbar m\omega^2\cdot\frac{i\hbar}{m}\cdot\frac{i^2t^2}{2!\hbar^2}\right)\hat{x}(0)+\ldots\\
&= \hat{x}(0)+\frac{t}{m}\hat{p}(0)-\frac{t^2\omega^2}{2!}\hat{x}(0)-\frac{t^3\omega^2}{m}\hat{p}(0)+\ldots\\
&= \hat{x}(0)\cos(\omega t)+\frac{\hat{p}(0)}{m\omega}\sin(\omega t)\,. \qquad (3.148)
\end{aligned}$$

This, of course, is in full agreement with the result obtained from the previous derivation.
From the findings so far, one may be tempted to conclude that with the operators also the expectation values $\langle\hat{x}\rangle$ and $\langle\hat{p}\rangle$ oscillate with angular frequency ω. This inference, however, is wrong. To see this, consider any energy eigenstate $|n\rangle$. The expectation value $\langle n|\hat{x}(t)|n\rangle$ vanishes, because $\hat{x}(0)$ can be written as linear combination of the creation and annihilation operators and thus change n by one to yield states $|n\pm 1\rangle$, which are orthogonal with $|n\rangle$. To observ e oscillations like the classical oscillator, **superpositions of energy eigenstates**, such as

$$|\alpha\rangle = a_0|0\rangle + a_1|1\rangle \qquad (3.149)$$

must be considered. The expectation value of $\hat{x}(t)$ taken w.r.t. $|\alpha\rangle$ does indeed oscillate. Then the natural question arises: How can such a superposition of energy eigenstates be constructed that most closely imitates the oscillatory behaviour of the classical harmonic oscillator? Re-expressed in the language of wave-functions this is synonymous to constructing a wave-packet that bounces back and forth without spreading in shape. Such a state - no spread - is called a **coherent state**. It turns out that such a coherent state has the following properties:

- When expressed as a superposition of energy eigenstates,

$$|\lambda\rangle = \sum_{n=0}^{\infty} f(n)|n\rangle\,, \qquad (3.150)$$

the coefficients $f(n)$ enjoy the property that their squares $|f(n)|^2$ are distributed according to a Poisson-type equation around some mean value $\bar{n}$,

$$|f(n)|^2 = \left(\frac{\bar{n}^n}{n!}\right) \exp(-\bar{n}) . \tag{3.151}$$

- The state $|\lambda\rangle$ can be obtained by translating the ground state $|0\rangle$ about some finite distance.

- It satisfies the minimum uncertainty product at all times.

3.4 Wave mechanics

3.4.1 Schrödinger's wave equation, revisited

Time-dependent wave equation

Let us turn back to the Schrödinger picture and examine, once again, the time evolution of state kets $|\alpha(t)\rangle$. In x-representation, they are nothing but the wave functions used as starting point in many textbooks,

$$\psi(\vec{x}', t) = \langle\vec{x}'|\alpha(t)\rangle . \tag{3.152}$$

If the Hamilton operator is given by

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m} + \hat{V}(\vec{x}) , \tag{3.153}$$

where the potential V, being a function of the Hermitian operator $\hat{\vec{x}}$ itself is a Hermitian operator, dependent on the position $\vec{x}$, as indicated by

$$V(\hat{\vec{x}}) \leftrightarrow \hat{V}(\vec{x}) . \tag{3.154}$$

It is also local (i.e. it depends on **one** position $\vec{x}$ only),

$$\langle\vec{x}''|\hat{V}(\vec{x})|\vec{x}'\rangle = V(\vec{x}')\delta^3(\vec{x}' - \vec{x}'') . \tag{3.155}$$

In order to guarantee Hermiticity, in the equation above, $V(\vec{x}')$ is a real function of $\vec{x}'$. In order to derive Schrödinger's time-dependent wave function, it suffices to write down the Schrödinger equation for the state ket $|\alpha(t)\rangle$, cf. Eq. (3.24),

$$i\hbar\frac{\partial}{\partial t}\langle\vec{x}'|\alpha(t)\rangle = \langle\vec{x}'|\hat{H}|\alpha(t)\rangle\,. \tag{3.156}$$

Since the position eigenkets in the Schrödinger picture do not change with time, it is thus save to multiply from the right with them. Using

$$\left\langle \vec{x}' \left|\hat{\vec{p}}^2\right| \alpha(t)\right\rangle = -\hbar^2\vec{\nabla}'^2\langle\vec{x}'|\alpha(t)\rangle\,, \tag{3.157}$$

cf. Eq. (2.76), and the x-representation of the potential,

$$\langle\vec{x}'|\hat{V}(\vec{x}) = V(\vec{x}')\langle\vec{x}'|\,, \tag{3.158}$$

where $V(\vec{x}')$ is no longer an operator (Eq. (3.155), yields

$$i\hbar\frac{\partial}{\partial t}\langle\vec{x}'|\alpha(t)\rangle = -\frac{\hbar^2}{2m}\vec{\nabla}'^2\langle\vec{x}'|\alpha(t)\rangle + V(\vec{x}')\langle\vec{x}'|\alpha(t)\rangle \tag{3.159}$$

or

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{x}',\,t) = -\frac{\hbar^2}{2m}\vec{\nabla}'^2\psi(\vec{x}',\,t) + V(\vec{x}')\psi(\vec{x}',\,t)\,, \tag{3.160}$$

the celebrated **time-dependent Schrödinger equation**. Quantum mechanics based on this equation usually is called wave mechanics, and in most cases this is how quantum mechanics is presented. In the ket-formalism presented here, this equation, however, is just the Schrödinger equation for a state ket written explicitly in the position representation.

Time-independent wave equation

This equation, Eq. (3.160), can now be specialised for wave functions corresponding to a stationary state. Such a state is usually connected to energy eigenfunctions. The time-dependence of such an energy eigenstate $|a'(t)\rangle$ can be written as

$$\langle\vec{x}'|a'(t)\rangle = \langle\vec{x}'|a'\rangle\exp\left(-\frac{iE_{a'}t}{\hbar}\right)\,, \tag{3.161}$$

where it is understood that the system initially is prepared in some simultaneous eigenstate of $\hat{H}$ and $\hat{A}$ with eigenvalues $E_{a'}$ and a', respectively. Substituting this into Eq. (3.160) immediately results in

$$E_{a'}\langle\vec{x}'|a'\rangle = -\frac{\hbar^2}{2m}\vec{\nabla}'^2\langle\vec{x}'|a'\rangle + V(\vec{x}')\langle\vec{x}'|a'\rangle\,. \tag{3.162}$$

At that point it is important to note that the labelling of energy eigenstates through some reference operator $\hat{A}$ is not really needed - any operator commuting with the Hamiltonian, including the Hamiltonian itself, will do the job. Therefore, the labels a' may be just ommitted. Denoting the energy eigenfunctions $\langle\vec{x}'|a'\rangle$ by $u_E(\vec{x}')$ then allows to replace the equation above by

$$Eu_E(\vec{x}') = -\frac{\hbar^2}{2m}\vec{\nabla}'^2 u_E(\vec{x}') + V(\vec{x}')u_E(\vec{x}')\,. \tag{3.163}$$

This is the original form of the **time-independent Schrödinger equation**. It has been announced by E. Schrödinger in the first of his four monumental papers laying the foundations of Quantum Mechanics in 1926. In the same paper he immediately applied this equation to the hydrogen atom, which will be dealt with in the next chapter.

Solutions for the time-independent wave equation

To solve Eq. (3.163), boundary conditions must be imposed. Suppose the solution for a bound state problem is wanted. This implies that

$$E < \lim_{|\vec{x}'|\to\infty} V(\vec{x}')\,, \tag{3.164}$$

translating into

$$u_E(\vec{x}') \to 0 \quad \text{for} \quad |\vec{x}'| \to \infty\,. \tag{3.165}$$

From the theory of partial differential equations it is well known that such a constraint applied to equations like the time-independent wave equation, Eq. (3.163) allows **non-trivial solutions only for a discrete set of values for** E and thus leads to a **quantisation of energy levels**. Once, the partial differential equation is written out, i.e. once the potential is fixed, finding them is a straightforward exercise that can be left to a computer. The problem as such is of the same type as that of finding characteristic frequencies of vibrating strings or membranes, it is a boundary-value problem. Let us, at that point, go back a step and investigate the nature of the constraint imposed on u_E. Using that

$$u_E(\vec{x}') = \langle \vec{x}' | a' \rangle \,, \tag{3.166}$$

Eq. (3.165) can be understood as a condition of having properly normalised states $|a'\rangle$:

$$\begin{aligned} 1 &= \langle a' | a' \rangle = \int \mathrm{d}^3 x' \langle a' | \vec{x}' \rangle \langle \vec{x}' | a' \rangle \\ &= \int \mathrm{d}^3 x' |u_E(\vec{x}')|^2 = \int \mathrm{d}x' \mathrm{d}\rho \mathrm{d}\phi |x' u_E(x',\, \rho,\, \phi)|^2 \,, \end{aligned} \tag{3.167}$$

where in the last step a transformation on polar coordinates has been perfomed. This is a condition, which in general can only be satisfied, if the function $u_E(\vec{x}')$ vanishes faster than $1/x'$ for infinite x'.

Chapter 4

Selected time-independent problems

In this chapter a few selected problems will be discussed, which can be solved by means of the time-independent Schrödinger equation.

4.1 Box potentials

4.1.1 Rigid wall potential in one dimension

Anatomy of the differential equation

As a first example for applications of the time-independent Schrödinger equation in one dimension,

$$\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}^2x}u_E(x) = (V(x) - E)u_E(x)\,, \tag{4.1}$$

cf. Eq. (3.163), consider a particle in a potential of the form

$$V(x) = \begin{cases} 0 & \text{for } x \in [-a,\, a] \\ V & \text{else}\,, \end{cases} \tag{4.2}$$

where for the beginning the case $V = \infty$ is investigated. In this case the infinitely high potential confines the particle to the region between $-a$ and a, it cannot penetrate the potential walls. According to the probabilistic interpretation of the wave function, its value

$$\psi(x,\, t) = e^{-iE/\hbar t}u_E(x) = 0 \quad \text{for } x \notin [-a,\, a] \tag{4.3}$$

translates into the particle being in the region. Therefore, suitable boundary conditions for the wave function inside the region $[-a,\, a]$ are

$$u_E(-a) = u_E(a) = 0\,. \tag{4.4}$$

First solution of the differential equation

A suitable *ansatz* for its solution is

$$u_E(x) = B^+ \exp(ikx) + B^- \exp(-ikx)\,, \tag{4.5}$$

where

$$k = \sqrt{\frac{2mE}{\hbar}}\,. \tag{4.6}$$

Plugging this into the Schrödinger equation, translates into

$$\begin{aligned} &-\frac{\hbar^2}{2m}\,(ik)^2\,[B^+ \exp(ikx) + B^- \exp(-ikx)] \\ &= \;\; E[B^+ \exp(ikx) + B^- \exp(-ikx)] \end{aligned} \tag{4.7}$$

in the region $x \in [-a,\, a]$, proving the validity of the ansatz. Adding in boundary conditions

$$u_E(\pm a) = B^+ \exp(\pm ika) + B^-(\mp ika) = 0 \tag{4.8}$$

implies there are two sets of solutions, namely

- Sine-type solutions:

$$\begin{aligned} &B^+ = -B^- \equiv A^- \quad \text{and} \quad k = \frac{2n\pi}{2a} \\ &\Longrightarrow \quad u_E(x) = A^- \left[\exp\left(\frac{2ni\pi x}{2a}\right) - \exp\left(-\frac{2ni\pi x}{2a}\right)\right] \end{aligned} \tag{4.9}$$

- Cosine-type solutions:

$$\begin{aligned} &B^+ = B^- \equiv A^+ \quad \text{and} \quad k = \frac{(2n-1)\pi}{2a} \\ &\Longrightarrow \quad u_E(x) = A^+ \left[\exp\left(\frac{(2n-1)i\pi x}{2a}\right) + \exp\left(-\frac{(2n-1)i\pi x}{2a}\right)\right] \end{aligned} \tag{4.10}$$

In both cases, $n = 1, 2, 3, \ldots$.
These two sets of solutions, +- and −-types, correspond to sinoidal and cosinoidal solutions and, ultimately, to different parities. Here, parity is the property of an object under spatial reflection, i.e. under the transformation $\vec{x} \to -\vec{x}$, or, in one dimension, $x \to -x$. More precisely, parity is defined through

$$\mathcal{O} \stackrel{\vec{x}\to-\vec{x}}{\longrightarrow} P\mathcal{O}\,. \tag{4.11}$$

Then clearly, the ±-parts have parity $P = \pm$. Thus the ±-label also labels the parity of the respective eigenmode.
The corresponding energy eigenmodes have kinetic energies only (the potential vanishes), namely

$$\begin{aligned} E_n^- &= \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\frac{(2n)^2\pi^2}{(2a)^2} \\ E_n^+ &= \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\frac{(2n-1)^2\pi^2}{(2a)^2}\,. \end{aligned} \tag{4.12}$$

Hence, each individual mode can also be written as

$$\begin{aligned} u^-_{E_n}(x) &= 2A_n^- \sin\frac{2n\pi x}{2a} \\ u^+_{E_n}(x) &= 2A_n^+ \cos\frac{(2n-1)\pi x}{2a}\,. \end{aligned} \tag{4.13}$$

To fix the amplitudes, it must be taken into account that the energy eigenmodes are also normalised, implying

$$1 \equiv \int\limits_{-a}^{a} \mathrm{d}x |u^{\pm}_{E_n}(x)|^2 = 4a|A\pm_n|^2 \tag{4.14}$$

leading to

$$A_n = \frac{1}{2\sqrt{a}} \tag{4.15}$$

Taken everything together yields

$$u_n^{\pm}(x) = \frac{1}{\sqrt{a}} \begin{cases} \cos\dfrac{(2n-1)\pi x}{2a} \\ \sin\dfrac{2n\pi x}{2a}\,. \end{cases} \tag{4.16}$$

for each allowed value $n \in \mathbf{N}$. This can easily be generalised to the case of more than one dimension, i.e. a three-dimensional box with dimensions a, b and c.

4.1.2 Square-well potential in one dimension

Anatomy of the differential equation

Let us consider a case now, where the potential is not infinite but finite. For simplicity, the region of vanishing potential now is given by

$$V(x) = \begin{cases} V_0 < 0 & \text{for } x \in [-a,\, a] \\ 0 & \text{else,} \end{cases} \tag{4.17}$$

Then bound states are those, whose total energy E is smaller than 0 - that is the potential energy is negative and overwhelms the kinetic energy. The other case, E larger than 0, will be considered later. For the bound state, the time-independent Schrödinger equation reads

$$\begin{aligned} \frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}^2 x}u_E(x) &= (V(x) - E)u_E(x) \\ &= \begin{cases} -(V_0 + E)u_E(x) & \text{for } x \in [-a,\, a] \\ -Eu_E(x) & \text{else}\,. \end{cases} \end{aligned} \tag{4.18}$$

Because the potential is not continuous, three regions must be distinguished, namely I, II, and III. The total solution then must be composed of solutions in each of the regions, such that

$$u_E(x) = \begin{cases} u_E^{(I)}(x) & \text{for} \quad x < -a \\ u_E^{(II)}(x) & \text{for} \quad -a < x < a \\ u_E^{(III)}(x) & \text{for} \quad x > a\,. \end{cases} \tag{4.19}$$

In addition, there are two conditions to be imposed on $u_E(x)$, namely

1. Normalisation:
   Since the function $u_E(x)$ corresponds to a state vector, it must be possible to normalise it. Only this guarantees the probabilistic interpretation of the solution. Hence,

$$\int\limits_{-\infty}^{\infty} \mathrm{d}x |u_E(x)|^2 < \infty\,. \tag{4.20}$$

2. Continuity:
   Both $u_E(x)$ and its first derivative w.r.t. x must be continuous at the borders of the regions. This is due to the structure of the Hamilton operator, exhibiting a double differentiation w.r.t. x (the ∇^2).

It should be noted here that both requirements were trivially fulfilled in the case above.

Constructing the individual solutions

Introducing

$$\kappa^2 = -\frac{2mE}{\hbar^2} \quad \text{and} \quad k^2 = \frac{2m(E-V_0)}{\hbar^2} \tag{4.21}$$

allows to impose the following differential equations in the three regions:

$$\begin{aligned} {u''}_E^{(I)} - \kappa^2 u_E^{(I)} &= 0 \\ {u''}_E^{(II)} + k^2 u_E^{(II)} &= 0 \\ {u''}_E^{(III)} - \kappa^2 u_E^{(III)} &= 0\,. \end{aligned} \tag{4.22}$$

Then, *ansätze* for the solutions in the three regions are readily formulated,

$$\begin{aligned} u_E^{(I)}(x) &= A_1 e^{\kappa x} + B_1 e^{-\kappa x} \\ u_E^{(II)}(x) &= A_2 e^{ikx} + B_2 e^{-ikx} \\ u_E^{(III)}(x) &= A_3 e^{\kappa x} + B_3 e^{-\kappa x}\,. \end{aligned} \tag{4.23}$$

The first condition, normalisation of the function $u_E(x)$ immediately implies that

$$B_1 = A_3 = 0\,, \tag{4.24}$$

because the corresponding exponentials diverge for $x \to \mp\infty$. The requirement of continuous transitions at the borders for the functions and their first derivatives (the Hamiltonian includes a second derivative, ∇^2),

$$\begin{aligned} u_E^{(I)}(-a) &= u_E^{(II)}(-a)\,, & {u'}_E^{(I)}(-a) &= {u'}_E^{(II)}(-a)\,, \\ u_E^{(II)}(a) &= u_E^{(III)}(a)\,,\text{and} & {u'}_E^{(II)}(a) &= {u'}_E^{(III)}(a) \end{aligned} \tag{4.25}$$

result in the following linear homogeneous system of equations

$$\begin{aligned}
A_1 e^{-\kappa a} - A_2 e^{-ika} - B_2 e^{ika} &= 0 \\
\kappa A_1 e^{-\kappa a} - ikA_2 e^{-ika} + ikB_2 e^{ika} &= 0 \\
A_2 e^{ika} + B_2 e^{-ika} - B_3 e^{-\kappa a} &= 0 \\
ikA_2 e^{ika} - ikB_2 e^{-ika} + \kappa B_3 e^{-\kappa a} &= 0\,. \qquad (4.26)
\end{aligned}$$

In order then to have non-trivial solutions for the remaining A_i and B_i, the determinant of coefficients must vanish, hence

$$\begin{aligned}
&\det \begin{vmatrix} e^{-\kappa a} & -e^{-ika} & -e^{ika} & 0 \\ \kappa e^{-\kappa a} & -ike^{-ika} & ike^{ika} & 0 \\ 0 & e^{ika} & e^{-ika} & -e^{-\kappa a} \\ 0 & ike^{ika} & -ike^{-ika} & \kappa e^{-\kappa a} \end{vmatrix} \\
&= \; e^{-2\kappa a}\left[-(\kappa + ik)^2 e^{-2ika} + (\kappa - ik)^2 e^{2ika}\right] = 0 \qquad (4.27)
\end{aligned}$$

or

$$\begin{aligned}
&(\kappa^2 - k^2)\left(e^{2ika} - e^{-2ika}\right) - 2ik\kappa\left(e^{2ika} + e^{-2ika}\right) \\
&= \; \left(e^{2ika} - e^{-2ika}\right)\left[(\kappa^2 - k^2) + 2k\kappa\cot(2ka)\right] = 0\,. \qquad (4.28)
\end{aligned}$$

This last equation is quadratic in k and κ and has the solutions

$$\kappa = \begin{cases} +k\tan(ka) \\ -k\cot(ka)\,. \end{cases} \qquad (4.29)$$

Introducing two new variables, $\xi = ka$ and $\eta = \kappa a$ allows to rewrite this as

$$\eta = \begin{cases} +\xi\tan\xi \\ -\xi\cot\xi\,. \end{cases} \qquad (4.30)$$

Constructing the combined solution

There is, however, yet another relation between κ and k, encoded in Eq. (4.21). Eliminating E from it yields

$$\kappa^2 + k^2 = -\frac{2mV_0}{\hbar^2}\,. \qquad (4.31)$$

These equations show that there is only a limited set of values of k (and hence κ and, ultimately E), allowing for a solution of the two equations above. In fact there is a case, where only one energy eigenvalue exists for either of the two sets of solutions. to construct it, let's go back to the ξ and η and rephrase the equation above as a condition in the ξ-η plane. There, the radius of the characteristic circle gets multiplied by a, and thus

$$\xi^2+\eta^2 = -\frac{2mV_0a^2}{\hbar^2}\,. \tag{4.32}$$

Comparing this with the solutions of Eq. (4.30) shows that there is one valid solution only, if the radius in smaller than $\pi/2$ resulting in

$$-V_0a^2 \le \frac{\pi^2\hbar^2}{8m}\,. \tag{4.33}$$

Calculating the coefficients A_i and B_i from the two solutions is now straight-forward. Just replace κ with either $k\tan(ka)$ or with $-k\cot(ka)$ in the set of equations. Then, with the first solution for κ, the set of equations Eq. (4.26) boils down to

$$B_2 = A_2\,,\quad B_3 = A_1\,,\ \text{ and }\ 2A_2\cos(ka) = A_1e^{-\kappa a} \tag{4.34}$$

leading to a solution with positive parity,

$$u_E(x) = u_E(-x) = \begin{cases} A_1e^{\kappa x} & \text{for}\quad x<-a \\ A_1e^{-\kappa a}\frac{\cos(kx)}{\cos(ka)} & \text{for}\quad -a<x<a \\ A_1e^{-\kappa x} & \text{for}\quad x>a\,. \end{cases} \tag{4.35}$$

For the second solution for κ, the set of equations Eq. (4.26) results in

$$B_2 = -A_2\,,\quad B_3 = -A_1\,,\ \text{ and }\ 2iA_2\sin(ka) = -A_1e^{-\kappa a} \tag{4.36}$$

leading to a solution with negative parity,

$$u_E(x) = -u_E(-x) = \begin{cases} A_1e^{\kappa x} & \text{for}\quad x<-a \\ -A_1e^{-\kappa a}\frac{\sin(kx)}{\sin(ka)} & \text{for}\quad -a<x<a \\ -A_1e^{-\kappa x} & \text{for}\quad x>a\,. \end{cases} \tag{4.37}$$

In order to fix the so far unknown amplitude A_1, the normalisation condition is invoked,

$$1 = \int\limits_{-\infty}^{\infty}\mathrm{d}x|u_E(x)|^2 = \int\limits_{-\infty}^{-a}\mathrm{d}x|u_E^{(I)}(x)|^2 + \int\limits_{-a}^{a}\mathrm{d}x|u_E^{(II)}(x)|^2 + \int\limits_{a}^{\infty}\mathrm{d}x|u_E^{(III)}(x)|^2\,. \tag{4.38}$$

In the first case this leads to

$$
\begin{aligned}
1 &= |A_1|^2 \left[\frac{e^{-2\kappa a}}{2\kappa} + \frac{e^{-2\kappa a}}{\cos^2(ka)} \left(a + \frac{\sin(2ka)}{2k} \right) + \frac{e^{-2\kappa a}}{2\kappa} \right] \\
&= |A_1^2| e^{-2\kappa a} \left[\frac{1}{\kappa} + a[1 + \tan^2(ka)] + \frac{\tan(ka)}{k} \right] \\
&= |A_1^2| e^{-2\kappa a} \left[\frac{1}{\kappa} + a + \frac{k\kappa^2 + \kappa}{k^2} \right] , \qquad (4.39)
\end{aligned}
$$

where the identity $\kappa = k \tan(ka)$ has been used. Hence

$$
|A_1|^2 = \frac{e^{2\kappa a}}{a \left(1 + \frac{1}{a\kappa} + \frac{a\kappa^2 + \kappa}{ak^2} \right)} . \qquad (4.40)
$$

In the second case this leads to

$$
\begin{aligned}
1 &= |A_1|^2 \left[\frac{e^{-2\kappa a}}{2\kappa} + \frac{e^{-2\kappa a}}{\sin^2(ka)} \left(a - \frac{\sin(2ka)}{2k} \right) + \frac{e^{-2\kappa a}}{2\kappa} \right] \\
&= |A_1|^2 e^{-2\kappa a} \left[\frac{1}{\kappa} + a[1 + \cot^2(ka)] - \frac{\cot(ka)}{k} \right] \\
&= |A_1|^2 e^{-2\kappa a} \left[\frac{1}{\kappa} + a + \frac{a\kappa^2 + \kappa}{k^2} \right] , \qquad (4.41)
\end{aligned}
$$

where this time the identity $\kappa = -k \cot(ka)$ has been used. Hence, again

$$
|A_1|^2 = \frac{e^{2\kappa a}}{a \left(1 + \frac{1}{a\kappa} + \frac{a\kappa^2 + \kappa}{ak^2} \right)} . \qquad (4.42)
$$

Taken everything together, in both case, the amplitude is given by

$$
|A_1|^2 \quad = \quad \frac{e^{2\kappa a}}{a \left(1 + \frac{1}{a\kappa} + \frac{a\kappa^2 + \kappa}{ak^2} \right)} . \qquad (4.43)
$$

Tunneling, first exposure

Since the energy eigenfunctions do not vanish in the regions I and III, there is a finite probability to find the particle in these regions, where classically it is forbidden (remember, $E < 0!$). This effect is known as tunneling. It should be noted that the probability density (the probability per unit length) for finding the particle vanishes quickly (exponentially) with the distance from the border, namely with

$$\begin{aligned}
\mathcal{P}^{(I)} = \mathcal{P}^{(III)} &= |u_E^{(I,III)}|^2 = |A_1|^2\, e^{-2\kappa|x|} \\
&= |A_1|^2 \exp\left(-\sqrt{\frac{-8mE}{\hbar^2}}\,|x|\right)\,.
\end{aligned} \tag{4.44}$$

4.1.3 Scattering at a potential barrier

Anatomy of the differential equation

Let us now consider a different type of problem, where a beam of particles coming from $x = -\infty$ hits a potential at $x = 0$, given by

$$V(x) = V_0\vartheta(x)\,, \tag{4.45}$$

where

$$\vartheta(x) = \begin{cases} 1 & \text{for} \quad x \geq 0 \\ 0 & \text{else} \end{cases} \tag{4.46}$$

is the Heavyside function. Again, the interesting case is, where the kinetic energy of the particles E is smaller than the potential, $E < V_0$.

Constructing the individual solutions

Again there are disjoint regions with different equations, namely

$$\begin{aligned}
u''^{(I)}_E(x) + \kappa^2 u_E^{(I)} &= 0 \quad \text{for } x < 0 \\
u''^{(II)}_E(x) - k^2 u_E^{(II)} &= 0 \quad \text{for } x > 0\,,
\end{aligned} \tag{4.47}$$

where

$$\kappa^2 = \frac{2mE}{\hbar^2} > 0 \quad \text{and} \quad k^2 = \frac{2m(V_0 - E)}{\hbar^2} \tag{4.48}$$

have been introduced. Then *ansätze* for the solutions in the two regions are again of exponential form, namely

$$\begin{aligned}
u_E^{(I)}(x) &= A_1 e^{i\kappa x} + B_1 e^{-i\kappa x} \\
u_E^{(II)}(x) &= A_2 e^{kx} + B_2 e^{-kx}\,.
\end{aligned} \tag{4.49}$$

They have to be compared with the physical situation.
First of all, the energy eigenfunction in region I must describe a particle wave approaching from $x = -\infty$, i.e. the total wave function should be of the form

$$\Psi(x,\, t) = Ae^{i(\kappa x - \omega t)} := u_E^{(I)} e^{-i\omega t}\,. \tag{4.50}$$

Hence, the contribution $\propto B_1$ describes the **reflected part** of the incoming wave. In addition for $x \to \infty$, only a finite (i.e. vanishing) amount of particles should appear, implying that $A_2 = 0$. This is a physics requirement, which has nothing to do with normalisation of the wave function. Since the particles in region I follow **a plane wave, the total wave function cannot be normalised to one**. However, the usual continuity conditions still hold true,

$$u^{(I)}(0) = u^{(II)}(0) \quad \text{and} \quad u'^{(I)}(0) = u'^{(II)}(0) \tag{4.51}$$

and thus result in

$$\begin{aligned} A_1 + B_1 &= B_2 \\ i\kappa(A_1 - B_1) &= -kB_2 \end{aligned} \tag{4.52}$$

and finally in

$$\begin{aligned} B_1 &= -\frac{k + i\kappa}{k - i\kappa}\, A_1 \\ B_2 &= -\frac{2i\kappa}{k - i\kappa}\, A_1\,. \end{aligned} \tag{4.53}$$

The reflection coefficient

Defining the reflection coefficient $\mathcal{R}$ through the ratio of the intensities of the outgoing and the incoming wave in region I, one finds

$$\mathcal{R} = \frac{|B_1|^2}{|A_1|^2} = \frac{|k + i\kappa|^2}{|k - i\kappa|^2} = 1\,. \tag{4.54}$$

Hence, **in the case considered here,** $0 < E < V_0$**, all particles are reflected at the potential barrier**. This is in full agreement with classical theory. There is, however, one subtlety, namely the fact that there is a finite probability to find particles in region II. The corresponding **probability density** is given by the square of the wave function in that region,

$$\begin{aligned}\mathcal{P}^{(II)} &= |u_E^{(II)}(x)|^2 = |B_2|^2 e^{-2kx} \\ &= \frac{4\kappa^2}{k^2+\kappa^2}|A_1|^2 e^{-2kx} \\ &= \frac{4E}{V_0}\,|A_1|^2 \exp\left(-\frac{2\sqrt{2m(V_0-E)}\,x}{\hbar}\right) .\end{aligned} \tag{4.55}$$

The case $E > V_0$

The results for this case can be constructed from the results obtained so far with the replacement $k \to ik'$ and

$$k'^2 = \frac{2m(E-V_0)}{\hbar} > 0\,. \tag{4.56}$$

4.1.4 Tunneling: Scattering at a potential wall

Anatomy of the differential equation and individual solutions

As a last problem consider the case where the potential barrier is of finite size, say a, such that

$$V(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ V_0 > 0 & \text{for} \quad 0 < x < a \\ 0 & \text{for} \quad x > a\,. \end{cases} \tag{4.57}$$

Setting again

$$\kappa^2 = \frac{2mE}{\hbar^2} > 0 \quad \text{and} \quad k^2 = \frac{2m(V_0-E)}{\hbar^2} \tag{4.58}$$

the differential equations in the three regions are given by

$$\begin{aligned}u''^{(I)}_E(x) + \kappa^2 u_E^{(I)} &= 0 \quad \text{for } x < 0 \\ u''^{(II)}_E(x) - k^2 u_E^{(II)} &= 0 \quad \text{for } 0 < x < a\,, \\ u''^{(III)}_E(x) + \kappa^2 u_E^{(III)} &= 0 \quad \text{for } x > a\,.\end{aligned} \tag{4.59}$$

As before, the following, exponential *ansätze*

$$\begin{aligned} u_E^{(I)}(x) &= A_1 e^{i\kappa x} + B_1 e^{-i\kappa x} \\ u_E^{(II)}(x) &= A_2 e^{kx} + B_2 e^{-kx} \\ u_E^{(III)}(x) &= A_3 e^{i\kappa x} + B_3 e^{-i\kappa x} \end{aligned} \tag{4.60}$$

have to be compared with the physical situation. Since there is no incoming wave from $x = \infty$, $B_3 = 0$. Then there remain four coefficients (B_1, A_2, B_2 and A_3) to be determined in dependence on A_1. Again this can be done by employing the continuity conditions at the borders.

Transmission vs. reflection

After some straightforward but tedious algebra, it turns out that the coefficient A_3 does not vanish, therefore it makes sense to define a transmission coefficient $\mathcal{T}$ in addition to the reflection coefficient. Both are then given by defining ratios of transmitted over incoming and of reflected over incoming wave intensity

$$\begin{aligned} \mathcal{R} &= \frac{|B_1|^2}{|A_1|^2} \\ \mathcal{T} &= \frac{|A_3|^2}{|A_1|^2}. \end{aligned} \tag{4.61}$$

The sum of both must add to one - after all the individual particles can be either transmitted through the potential in region I by tunnelling or they are reflected. The tunneling probability is given by the transmission coefficient and reads

$$\mathcal{T} = \left[\cosh^2(ka) + \frac{1}{4}\left(\frac{\kappa}{k} - \frac{k}{\kappa}\right)^2 \sinh^2(ka)\right]^{-1}. \tag{4.62}$$

For a high potential wall, $V_0 \gg E$, $ka \gg 1$ and $k \gg \kappa$ and therefore

$$\mathcal{T} \to \frac{16E}{V_0} e^{-2ka} = \frac{16E}{V_0} \exp\left(-\frac{2a}{\hbar}\sqrt{2m(V_0 - E)}\right). \tag{4.63}$$

If, on the other hand, E approaches V, $k \to 0$ and $\mathcal{T} \to 1$.

4.1.5 Summary of the findings

To summarise the findings of the previous examples:

1. Similar to classical physics, bound states are different from free particle states. For free particle plane-wave states a normalisation of the position space wave function is impossible; bound states in contrast populate only a finite region and their states are normalisable. This **normalisation condition immediately implies a discretisation of the allowed energy states - the quantisation of bound state energies is thus a result of normalisation**. In fact there may be a finite or countable infinite number of energy eigenstates of the bound particle - depending on the exact form of the potential.

2. Energy eigenstates can also be characterised by their symmetry properties - in the examples here, the behaviour of the wave function under spatial reflections, parity, has been investigated. There is a growing number of knots (points where the wave function equals 0) with rising energy of the eigenstate in question.

3. There is a finite probability for a particle to penetrate into the classically forbidden region. This tunneling effect is due to the fact that the wave function must be continuous over the borders, because the Hamiltonian usually involves derivatives.

4. Heisenbergs uncertainty relation implies that the lowest energy eigenstate is always above the minimum of the potential.

4.2 The hydrogen atom

This section is devoted to the treatment of the hydrogen atom with the methods developed so far. There are two tricky, connected points, which will not be discussed in detail here, namely the fact that angular momentum is also quantised and the fact that the electron has a spin interacting with its angular momentum.

4.2.1 Schrödinger equation

Anatomy of the differential equation and individual solutions

This time, the time-independent Schrödinger equation in three dimensions must be investigated

$$\frac{\hbar^2}{2m}\vec{\nabla}^2 u_E(\vec{r}) = (V(\vec{r}) - E)u_E(\vec{r})\,, \tag{4.64}$$

where the potential takes the spherically symmetric form

$$V(\vec{r}) = -\frac{e^2}{|\vec{r}|}\,. \tag{4.65}$$

In choosing this form of the Schrödinger equation, the fact has been used that the proton (or, more generally) the nucleus is much heavier than the electron - in fact by a factor of roughly 2000. Therefore, when replacing the original two-body problem (electron+proton) by the equivalent effective one-body problem with a reduced mass, it turns out that effectively the reduced mass is more or less identical to the electron. Therefore, to first order it is safe to assume the proton to be fixed at the origin with the electron surrounding it. Since the potential depends on the distance from the origin only, it is reasonable to move on to spherical coordinates,

$$\vec{r} = (r\sin\theta\cos\phi,\, r\sin\theta\sin\phi,\, r\cos\theta)^T\,. \tag{4.66}$$

In these coordinates, the Laplace operator ∇^2 acting on a function $f(r,\,\theta,\,\phi)$ assumes the form

$$\vec{\nabla}^2 f(r,\,\theta,\,\phi) = \frac{1}{r}\frac{\partial^2(rf)}{\partial r^2} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}\right]\,. \tag{4.67}$$

This translates into the time-independent Schrödinger equation to read

$$\begin{aligned}
&-\frac{2m}{\hbar^2}\left(E + \frac{e^2}{r}\right)u_E(\vec{r}) \\
&= \frac{1}{r}\frac{\partial^2[ru_E(\vec{r})]}{\partial r^2} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u_E(\vec{r})}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 u_E(\vec{r})}{\partial\phi^2}\right]\,.
\end{aligned} \tag{4.68}$$

Spherically symmetric solutions

Let us first try to find a set of simple functions satisfying the equation above, Eq. (4.68). Although in general the energy eigenfunction will depend on the distance and both angles, it is worthwhile to investigate whether there is a set of **special solutions**, which do not depend on the angles. These solutions are thus characterised by having total angular momentum equals to zero. The corresponding states are usually called "s"-wave states, where "s" stands for spherically symmetric.
In this case, Eq. (4.68) reduces to

$$-\frac{2m}{\hbar^2}\left(E+\frac{e^2}{r}\right)u_E(r)=\frac{1}{r}\frac{\mathrm{d}^2[ru_E(r)]}{\mathrm{d}r^2}\,. \tag{4.69}$$

Before actually solving it, let us rewrite this equation in a nicer fashion and introduce the Bohr radius r_B and the Rydberg energy E_R

$$r_B=\frac{h^2}{me}\approx 0.528\,\text{Angstrom}\quad\text{and}\quad E_R=\frac{me^4}{2\hbar^2}\approx 13.6\,\text{eV}\,. \tag{4.70}$$

Rescaling r and E by these quantities allows to rewrite Eq. (4.69) as

$$\frac{\mathrm{d}^2[\rho u_E(\rho)]}{\mathrm{d}\rho^2}=-\left(\varepsilon+\frac{2}{\rho}\right)\rho u_E(\rho)\,, \tag{4.71}$$

where the unit-free quantities

$$\rho=r/r_B\quad\text{and}\quad\varepsilon=E/E_R\,. \tag{4.72}$$

To solve the remaining differential equation is not trivial. However, employing the *ansatz*

$$\rho u_E(\rho)=e^{-\alpha\rho}\,g(\rho)\,, \tag{4.73}$$

where α is an arbitrary number and g is a new function, yields

$$\frac{\mathrm{d}^2g(\rho)}{\mathrm{d}\rho^2}-2\alpha\frac{\mathrm{d}g(\rho)}{\mathrm{d}\rho}+\left(\varepsilon+\alpha^2+\frac{2}{\rho}\right)g(\rho)=0 \tag{4.74}$$

for Eq. (4.71). Choosing $\alpha^2=-\varepsilon$ then yields

$$\frac{\mathrm{d}^2g(\rho)}{\mathrm{d}\rho^2}-2\alpha\frac{\mathrm{d}g(\rho)}{\mathrm{d}\rho}+\frac{2}{\rho}g(\rho)=0\,. \tag{4.75}$$

At first, this doesn't look like a great achievement, since it seems that the complicated differential equation for one function just has been replace by another, complicated differential equation for another function. However, the second one can be solved by making the *ansatz*

$$g(\rho) = \sum_{k=1}^{\infty} a_k \rho^k \,, \tag{4.76}$$

where the α_k are constant coefficients of the Taylor expansion. Then Eq. (4.74) becomes

$$\begin{aligned}
\sum_{k=1}^{\infty} k(k-1)a_k\rho^{k-2} - \sum_{k=1}^{\infty} 2\alpha k a_k \rho^{k-1} + \sum_{k=1}^{\infty} 2a_k\rho^{k-1} &= \\
\sum_{k=1}^{\infty}(k+1)k a_{k+1}\rho^{k-1} - \sum_{k=1}^{\infty} 2\alpha k a_k \rho^{k-1} + \sum_{k=1}^{\infty} 2a_k\rho^{k-1} &= \\
\sum_{k=1}^{\infty}\left[(k+1)ka_{k+1} - 2\alpha k a_k + 2a_k\right]\rho^{k-1} &= 0\,,
\end{aligned} \tag{4.77}$$

where in the transition from the first to the second line the fact has been employed that in the first sum the first term vanishes. Since the sum goes up to infinity this allows to shift the k's by one in this contribution.
The last line then implies that the coefficients a_k must satisfy the relation

$$(k+1)ka_{k+1} - 2\alpha k a_k + 2\alpha a_k = 0 \tag{4.78}$$

or

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1)} a_k \tag{4.79}$$

for all values of k. This means that, choosing some initial value for, say, a_1, $g(\rho)$ and hence the full solution $u_E(\rho)$ can be written as a potential series, dependent on this initial value and the energy $\alpha^2 = -\varepsilon$.

Boundary conditions: Interpreting the s-wave solutions

So, there is a solution now. Approximating that for large $k \gg 1$

$$a_{k+1} \approx \frac{2\alpha}{k} a_k \quad \Longrightarrow \quad a_{k+1} \approx \frac{(2\alpha)^k}{k!} a_1 \tag{4.80}$$

this allows to approximate the solution by

$$\begin{aligned}\rho u_E(\rho) &= e^{-\alpha\rho}\, g(\rho) \approx e^{-\alpha\rho} a_1 \sum_{k=1}^{\infty} \frac{(2\alpha)^k}{k!} \rho^k \\ &= a_1 e^{-\alpha\rho} e^{2\alpha\rho} = a_1 e^{\alpha\rho} \stackrel{\rho\to\infty}{\longrightarrow} \infty\,. \end{aligned} \tag{4.81}$$

This is bad news! It implies that the wave functions of bound states diverges for large distances from its center - the electron is most probable infinitely away from the proton. Of course, this is wrong. The only question is, whether this game of rising coefficients a_k can be altered such that the solutions look friendly. The answer is buried in Eq. (4.79),

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1)} a_k. \tag{4.82}$$

If for some $k = n \in N$, $\alpha = 1/n$ the recursion terminates (the numerator becomes zero). Then the function $g(\rho)$ is just a polynomial of finite order n and thus **smaller** than the exponential for $\rho \to \infty$. Thus, the bound states in the s-wave are characterised by

$$\alpha = \frac{1}{n} \quad \text{with} \quad n \in N \tag{4.83}$$

and therefore by

$$-\varepsilon = -\frac{E}{E_R} = \frac{1}{n^2} = 1,\, \frac{1}{4},\, \frac{1}{9},\, \ldots . \tag{4.84}$$

In fact, this finding was one of the first big successes of quantum mechanics, since it allowed for the first time to describe the gross features in the spectrum of the hydrogen from first principles (and not by some artificial “fix”).

Energy eigenfunctions of the s-wave solutions

Having obtained the solutions of the s-wave energy eigenfunctions it makes sense to investigate their properties. For each n, the energy is

$$E_n = -\frac{1}{n^2} E_R = -\frac{me^4}{2n^2\hbar^2}\,. \tag{4.85}$$

The energy eigenfunction looks like

$$u_E(\rho) = \frac{e^{-\rho/n}}{\rho} \sum_{k=1}^{n} a_k \rho^k \,, \tag{4.86}$$

where the recursion can be used to fix a_1 by demanding the wave function to be normalised to 1. Rather than doing this, let us just set $a_1 = 1$ and check the shape of the energy eigenfunctions in units of the Bohr radius. The ground state ($1s$) and the two first excited states ($2s$ and $3s$) are given by $n = 1, 2, 3$, respectively, where

$$\begin{aligned} u_{E_1}(\rho) &= e^{-\rho} \\ u_{E_1}(\rho) &= e^{-\rho/2}\left(1 - \frac{\rho}{2}\right) \\ u_{E_1}(\rho) &= e^{-\rho/3}\left(1 - \frac{2\rho}{3} + \frac{2\rho^2}{27}\right) \,. \end{aligned} \tag{4.87}$$

The pattern is clear: All energy eigenstates are exponentially damped for large distances ρ; the electron thus is pretty much localised around the proton. Higher excitations have a number of knots, i.e. places where the wave function (and thus the probability to find the electron) are zero. In fact, the ith excitation of the ground state has i knots.

4.2.2 Angular dependence: Spherical harmonics

Definition of spherical harmonics

In order to study the full problem, i.e. include also those states which are not spherically symmetric, let us go back to Eq. (4.68),

$$\begin{aligned} &-\frac{2m}{\hbar^2}\left(E + \frac{e^2}{r}\right) u_E(\vec{r}) \\ &= \frac{1}{r}\frac{\partial^2[r u_E(\vec{r})]}{\partial r^2} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u_E(\vec{r})}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 u_E(\vec{r})}{\partial\phi^2}\right] \,. \end{aligned} \tag{4.88}$$

Apparently, the term

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u_E(\vec{r})}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 u_E(\vec{r})}{\partial\phi^2}\right] \tag{4.89}$$

contains only derivatives w.r.t. angles. It seems natural to use an *ansatz*, which separates the wave function in spherical coordinates such that there is a radial and an angular part. Characterising the eigenstates by energy eigenvalues n, and by the quantum numbers of orbital angular momentum l and m, the wave function in position space (spherical coordinates) can then be written as

$$u_{nlm}(\vec{x}') = \langle \vec{x}' | n,\, l,\, m \rangle = R_{nl}(r) Y_l^m(\theta,\, \phi)\,. \tag{4.90}$$

The functions $Y_l^m(\theta,\, \phi)$ are known as spherical harmonics and they encode all the angular dependence of the wave function; even more, for spherically symmetric problems (the Hamiltonian commuting with both $\hat{\vec{L}}^2$ and $\hat{L}_z$), they build a basis of functions in which a solution may be formulated. Hence, their importance in physics is hard to overestimate. These functions can then be defined by the matrix elements

$$\langle \vec{e}_r | l,\, m \rangle = Y_l^m(\theta,\, \phi)\,, \tag{4.91}$$

where $\vec{e}_r$ is a unit vector, labelling the direction eigenkets. In this way, the **spherical harmonic $Y_l^m(\theta,\, \phi)$ is the amplitude to find a state characterised by l and m to point in the direction of $\vec{e}_r$, specified by θ nd ϕ.**

Determination of spherical harmonics

In order to actually determine them, let's start with the relations involving the operators of angular momentum. The eigenvalue equation

$$\hat{L}_z | l,\, m \rangle = \hbar m | l,\, m \rangle \tag{4.92}$$

immediately yields

$$\langle \vec{e}_r | \hat{L}_z | l,\, m \rangle = -i\hbar \frac{\partial}{\partial \phi} \langle \vec{e}_r | l,\, m \rangle = \hbar m \langle \vec{e}_r || l,\, m \rangle \tag{4.93}$$

or

$$-i\hbar \frac{\partial}{\partial \phi} Y_l^m(\theta,\, \phi) = \hbar m Y_l^m(\theta,\, \phi)\,. \tag{4.94}$$

This implies that the ϕ-dependence of the spherical harmonics is determined by $\exp(im\phi)$.

The behaviour with θ is a bit more complicated. To start with, remember the eigenvalue equation

$$\hat{\vec{L}}^2|l,\, m\rangle = \hbar^2 l(l+1)|l,\, m\rangle \tag{4.95}$$

and the fact that

$$\langle\vec{x}'|\hat{\vec{L}}^2|\alpha\rangle = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\phi}\right]\langle\vec{x}'|\alpha\rangle\,, \tag{4.96}$$

cf. Eq. (2.163). Taken together and replacing $\vec{x}'$ with $\vec{e}_r$, thus

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\phi}+l(l+1)\right]Y_l^m(\theta,\,\phi) = 0\,. \tag{4.97}$$

To obtain the $Y_l^m(\theta,\,\phi)$, consider the case $l=m$ first. In this case the ladder operator applied on the state yields 0,

$$\hat{L}_+|l,\, l\rangle = 0\,. \tag{4.98}$$

Remembering the matrix element for $\hat{L}_+$, Eq. (2.161),

$$\langle\vec{x}'|\hat{L}_+|\alpha\rangle = -i\hbar e^{i\phi}\left(i\frac{\partial}{\partial\theta}-\cot\theta\frac{\partial}{\partial\phi}\right)\langle\vec{x}'|\alpha\rangle\,, \tag{4.99}$$

then translates into (again $\vec{x}'\to\vec{e}_r$)

$$-i\hbar e^{i\phi}\left(i\frac{\partial}{\partial\theta}-\cot\theta\frac{\partial}{\partial\phi}\right)Y_l^l(\theta,\,\phi) = 0\,. \tag{4.100}$$

Straightforward calculation asserts that

$$Y_l^l(\theta,\,\phi) = c_l e^{il\phi}\sin^l\theta \tag{4.101}$$

solves this differential equation, and the only task left is to determine the normalisation constants c_l. This can be done, using that the states are all normalised,

$$\begin{aligned}\langle l',\, m'|l,\, m\rangle &= \int\limits_0^\pi \mathrm{d}\theta\sin\theta\int\limits_0^{2\pi}\mathrm{d}\phi\,\langle l',\, m'|\vec{e}_r\rangle\langle\vec{e}_r|l,\, m\rangle\\ &= \int\limits_0^\pi \mathrm{d}\theta\sin\theta\int\limits_0^{2\pi}\mathrm{d}\phi\, Y_{l'}^{m'*}(\theta,\,\phi)Y_l^m(\theta,\,\phi) = \delta_{ll'}\delta_{mm'}\,.\end{aligned} \tag{4.102}$$

Then

$$\int_0^\pi \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi\, |c_l|^2 \sin^{2l}\theta = 2\pi|c_l|^2 \int_{-1}^{1} \mathrm{d}\cos\theta (1-\cos^2\theta)^l = 1 \quad (4.103)$$

leading to

$$|c_l|^2 = \frac{1}{2\pi}\left[\int_{-1}^{1} \mathrm{d}x(1-x^2)^l\right]^{-1} \quad (4.104)$$

and thus

$$c_l = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)\,(2l)!}{4\pi}}\,. \quad (4.105)$$

where an arbitrary phase $(-1)^l$ has been added for convenience. In particular, for the first few values of l,

$$\begin{aligned} c_0 &= \sqrt{\frac{1}{4\pi}} \\ c_1 &= -\frac{1}{2}\sqrt{\frac{6}{4\pi}} \\ c_2 &= \frac{1}{8}\sqrt{\frac{120}{4\pi}}\,. \end{aligned} \quad (4.106)$$

It is worth noting that the normalisation is the same for all values of m for the Y_l^m with identical l - after all, the m-dependence resides completely in a phase that is irrelevant for the normalisation. Having thus fixed Y_l^m for the largest value of m, namely $m = l$, the ladder operator $\hat{L}_-$ can repeatedly be applied in order to determine all other Y_l^m with m down to $m = -l$. To see this, consider

$$\begin{aligned} \langle \vec{e}_r | l,\, m-1\rangle &= \frac{1}{\hbar\sqrt{(l+m)(l-m+1)}}\langle \vec{e}_r|\hat{L}_-|l,\, m\rangle \\ &= \frac{1}{\hbar\sqrt{(l+m)(l-m+1)}}e^{-i\phi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right)\langle \vec{e}_r|l,\, m\rangle\,. \end{aligned} \quad (4.107)$$

l	m	ang. dependence	Name of orbital	Number of states	Parity
0	0	1	s	1	+
1	+1 0 -1	$-\frac{1}{\sqrt{2}}\sin\theta\, e^{i\phi}$ $\cos\theta$ $\frac{1}{\sqrt{2}}\sin\theta\, e^{-i\phi}$	p	3	-1
2	+2 +1 0 -1 -2	$\frac{\sqrt{6}}{4}\sin^2\theta\, e^{2i\phi}$ $\frac{\sqrt{6}}{2}\sin\theta\cos\theta\, e^{i\phi}$ $\frac{1}{2}(3\cos^2\theta - 1)$ $-\frac{\sqrt{6}}{2}\sin\theta\cos\theta\, e^{-i\phi}$ $\frac{\sqrt{6}}{4}\sin^2\theta\, e^{-2i\phi}$	d	5	+

Table 4.1: The spherical harmonics for the first few (integer) values of orbital angular momentum l. In addition to the angular part of the Y_l^m also the spectroscopic name of the corresponding orbital, the number of states for each l (2l+1), and the parity of the orbital $(-1)^l$ are displayed.

This can then be used to recursively construct all Y_l^m starting from Y_l^l. The result of this is

$$Y_l^m(\theta,\,\phi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l+m)!}{(l-m)!}}\,e^{im\phi}\,\frac{1}{\sin^m\theta}\frac{\mathrm{d}^{l-m}}{\mathrm{d}(\cos\theta)^{l-m}}\sin^{2l}\theta\,. \tag{4.108}$$

In Table 4.1, the first few of them are listed.
The spherical harmonics satisfy

$$Y_l^{-m}(\theta,\,\phi) = (-1)^m[Y_l^m(\theta,\,\phi)]^* \tag{4.109}$$

and for $m = 0$ they are connected with the Legendre polynomials P_l such that

$$Y_l^0(\theta,\,\phi) = \sqrt{\frac{2l+1}{4\pi}}\,P_l(\cos\theta)\,. \tag{4.110}$$

Obviously, they do not depend on ϕ. The first few of the Legendre polynomials are

$$\begin{aligned}
P_0(\cos\theta) &= 1 \\
P_1(\cos\theta) &= \cos\theta \\
P_2(\cos\theta) &= \frac{1}{2}(3\cos^2\theta - 1) \\
P_3(\cos\theta) &= \frac{1}{2}(5\cos^3\theta - 3\cos\theta)\,.
\end{aligned} \tag{4.111}$$

Orbitals

Having dwelt on the angular part, it is time to use it in order to determine the full form of the orbitals of the spin-less, non-relativistic hydrogen atom. To do so, let us go back one more to Eq. (4.68),

$$\begin{aligned}
&-\frac{2m}{\hbar^2}\left(E+\frac{e^2}{r}\right)u_E(\vec{r}) \\
&= \frac{1}{r}\frac{\partial^2[ru_E(\vec{r})]}{\partial r^2} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u_E(\vec{r})}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 u_E(\vec{r})}{\partial\phi^2}\right]\,.
\end{aligned} \tag{4.112}$$

Plugging in the *ansatz* of Eq. (4.90), where the radial and angular part have been seprated,

$$u_{nlm}(\vec{x}') = \langle\vec{x}'|n,\,l,\,m\rangle = R_{nl}(r)Y_l^m(\theta,\,\phi)\,. \tag{4.113}$$

yields

$$\begin{aligned}
-\frac{2m}{\hbar^2}\left(E+\frac{e^2}{r}\right)R_{nl}(r)Y_l^m(\theta,\,\phi) &= \frac{1}{r}\frac{\partial^2[rR_{nl}(r)]}{\partial r^2}Y_l^m(\theta,\,\phi) \\
+\frac{R_{nl}(r)}{r^2}&\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y_l^m(\theta,\,\phi)}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y_l^m(\theta,\,\phi)}{\partial\phi^2}\right]\,.
\end{aligned} \tag{4.114}$$

Now, the square bracket in the second line is just the l.h.s. of

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\phi}\right]\,Y_l^m(\theta,\,\phi) = -l(l+1)\,Y_l^m(\theta,\,\phi)\,, \tag{4.115}$$

cf. Eq. (4.97), and thus

$$
\begin{aligned}
&-\frac{2m}{\hbar^2}\left(E+\frac{e^2}{r}\right)R_{nl}(r)Y_l^m(\theta,\,\phi)\\
&= \frac{1}{r}\frac{\partial^2[rR_{nl}(r)]}{\partial r^2}Y_l^m(\theta,\,\phi)-\frac{R_{nl}(r)}{r^2}l(l+1)Y_l^m(\theta,\,\phi)\,.
\end{aligned}
\tag{4.116}
$$

Shuffling terms then results in a differential equation for the radial part (the angular part is a constant as far as the differential equation above is concerned), namely

$$
-\frac{2m}{\hbar^2}\left(E+\frac{e^2}{r}-\frac{l(l+1)\hbar^2}{2mr^2}\right)R_{nl}(r)=\frac{1}{r}\frac{\partial^2[rR_{nl}(r)]}{\partial r^2}\,.
\tag{4.117}
$$

In contrast to the differential equation that describes the spherically symmetric solutions, cf. Eq. (4.69), a new term has appeared. At first it seems to be the result of some mathematical magic only, but it has some physics in it. To see how this works, consider the classical problem of a mass point in a spherically symmetric potential $V(r)$. The total energy of the particle amounts to

$$
E_{\text{class.}}=\frac{m}{2}\vec{v}^2+V(r)=\frac{m}{2}\left(\vec{v}_r^2+r^2\omega^2\right)+V(r)\,,
\tag{4.118}
$$

where $\vec{v}$ is the velocity, v_r its radial component and ω the angular frequency. But now, since the potential is spherically symmetric the angular momentum $L=mr^2\omega$ is a constant of motion. This allows to rewrite the energy as

$$
E_{\text{class.}}=\frac{m}{2}\vec{v}_r^2+V(r)\frac{L^2}{2mr^2}\,,
\tag{4.119}
$$

and the L-dependent term looks like a term being added to the potential energy rather than a part of the kinetic energy, since it does not depend on the velocity. But this is nearly the extra term above, the only difference is that $l(l+1)\hbar^2$ appears instead of $L^2\hbar^2$.

Going back to construct a solution for Eq. (4.117), it is clear that the same *ansatz* for the radial part as in the case of the spherically symmetric solutions will work again, namely the expansion in a series times an exponential,

$$
\rho R_{nl}(\rho)=e^{-\alpha\rho}\sum_{k=1}^{\infty}a_k\rho^k\,.
\tag{4.120}
$$

The only difference to the spherically symmetric case is the occurrence of the $l(l+1)$-term, which can be written as

$$
\begin{aligned}
-\frac{l(l+1)}{\rho^2}[\rho R_{nl}(\rho)] &= -l(l+1)e^{-\alpha\rho}\sum_{k=1}^{\infty} a_k \rho^{k-2} \\
&= -l(l+1)e^{-\alpha\rho}\left(\frac{a_1}{\rho}+\sum_{k=1}^{\infty} a_{k+1}\rho^{k-1}\right) . \quad (4.121)
\end{aligned}
$$

Sticking everything together results in the following equation for the a_k:

$$
\sum_{k=1}^{\infty}\left\{[k(k+1)-l(l+1)]\, a_{k+1} - 2(\alpha k - 1)a_k\right\} - \frac{l(l+1)a_1}{\rho} = 0 . \quad (4.122)
$$

There is only one term $\propto 1/\rho$, which therefore must vanish. Hence, either $l = 0$ and then the spherically symmetric solutions from above are described, or $a_1 = 0$, which must hold true for all solution with an angular dependence. All other terms vanish, if

$$
a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1) - l(l+1)}\, a_k . \quad (4.123)
$$

The only difference now is the appearance of the $l(l+1)$-term in the denominator. But again, this series must terminate at some point in order to have a physically meaningful interpretation: So, again, the series must terminate for $k = n$, if $\alpha n = 1$. Hence, again $\alpha n = 1$ implies the same condition as before for α, namely

$$
\alpha = 1/n , \quad (4.124)
$$

but this time there is a new condition, namely that $k \neq l$ in order to keep a non-vanishing denominator. This implies that since $a_1 = 0$ that all $a_k = 0$ for $k < l + 1$. Hence the k's start at $k = l + 1$ and end at $k = n$. The overall result is that for every value of l, there are various solutions, but for all $n > l$. Each of these solutions has the energy

$$
E_{n,l} = -\frac{me^4}{2\hbar^2}\,\frac{1}{n^2} . \quad (4.125)
$$

The wave function for these states is given by

$$\psi_{n,\,l,\,m}(\vec{r}) = Y_l^m(\theta,\ \phi)\, R_{n,\,l}(\rho)\,, \tag{4.126}$$

where

$$[\rho R_{n,\,l}(\rho)] = e^{-\alpha\rho} \sum_{k=1}^{\infty} a_k \rho^k\,, \tag{4.127}$$

with coefficients given by the corresponding recursion above.