Quantum Mechanics

Lecture in SS 2022 at the KFU Graz

Axel Maas

Contents

Chapter 1

Introduction

Quantum physics is arguably the most enigmatic, and most challenging, element of modern physics. Of course, this is largely due to the fact that its particular phenomena do manifest in a way not directly perceivable by our senses. But its oblique consequences are ubiquitous, from how the sun burns and how the biology of our own bodies operates up to almost all modern microelectronics and far beyond that. From the point of view of the universe quantum physics is thus rather the norm than any particular exception.

How our usual experience emerges from the quantum case is not entirely trivial, and in fact technically quite involved. This technical complexity is also one of the reasons why it is not fully resolved, and this is probably one of the reasons so many tales and mysteries seem to shroud it. On the other hand, the question whether our mathematical theories of quantum physics are an accurate representation of reality or just a convenient tool to capture reality is fully a question of science philosophy.

Quantum mechanics is the special case of quantum physics when considering the quantum version of classical mechanics. In fact, it is nowadays rather just the case where the effects of special relativity can be ignored, which more or less excludes large parts of electromagnetic effects beyond static systems. Including special relativity is much more complicated and leads to quantum field theory. Still, this is the language in which the foundation of quantum physics in terms of particle physics is nowadays formulated. Thus, to reach this most basic formulation of our current understanding of the laws of nature requires to move to quantum field theory. However, the final, possibly just technical, incorporation of general relativity in quantum physics is not entirely completed, despite substantial progress.

The aim of this lecture is to provide a basic introduction to quantum mechanics. Rather than starting with a full mathematically concise framework[1] first the simplest

[1] In contrast to quantum field theory the mathematical foundation of quantum mechanics is fully

possible system to display quantum behavior will be discussed in chapter 2. In this course the basic rules of quantum mechanics will be explored.

The transition to a full theory will then be made in chapter 3, where the foundation of quantum mechanics will be provided. This will include a discussion of the most pertinent general features of quantum physics.

These newly developed tools will be applied to a multitude of simple systems in chapter 4 and 6. The aim is here to understand how the tools are used, and this will lead to formulating a dynamical principle in chapter 5. It should also serve to build an intuitive idea of quantum mechanics, by experiencing how quantum mechanics operates. Of course, many important consequences of quantum physics will be found, illustrating what are typical phenomena.

This will lead to dealing with the most important prototypical real-world systems in chapter 7, in particular the hydrogen atom. This will also introduce the important differences in angular momentum compared to classical physics. An understanding of the angular momentum will then provide the transition to understanding symmetries in quantum physics in chapter 8. While these were already important in classical physics, they become much more central in quantum physics.

Afterwards, perturbative treatments of quantum physics in chapter 9 will be introduced, which forms one of the mainstays of modern theoretical physics, together with the much more involved numerical simulations. The latter are far too technically involved to include them here, for a lack of time and space. Besides these two primary methods, many other methods have been developed, and claimed their own niches where these two primary methods fail for various reasons. However, any other method is best understood by understanding the language and techniques of perturbation theory, making it a natural first step.

A system of quantum particles acts quite differently than a system of classical particles. A first look of this will be provided in chapter 10. This will be the starting point for many modern applications like solid state physics and quantum optics.

At this point, a understanding of the most important features of quantum physics can be reached. It is thus a good final step to consider that there are actually a multitude of different ways of thinking about quantum physics in chapter 11. In fact, there are many different, but equivalent, ways to set up quantum physics. Each of them emphasizes a different viewpoint. Knowing these different viewpoints is also important to try to grasp the underlying structure, and should also give a first glimpse of what could be behind quantum physics on a more literal interpretation level.

understood in the sense of a well-defined mathematical language.

Due to the long development of more than a hundred years and its continued relevance there is a large number of books on the topics of this lecture, and there are usually several new releases per year. It is thus somewhat futile to give literature recommendations. To give an idea, this lecture has been prepared using (in order of decreasing accessibility)

- Claude et al., "Quantenmechanik" (W de Gruyter)
- Fließbach, "Quantenmechanik" (Spektrum)
- Burkhardt et al., "Foundations of Quantum Physics" (Springer)
- Sakurai, "Modern Quantum Mechanics" (Addison Wesley)

as well as some more advanced texts and original papers. But everyone should find a suitable book for their own style. And certainly, these lecture notes can by no means replace a textbook.

Finally, quantum physics is confusing. Not only at first sight. Do not expect the following to be easy, straightforward, or intuitive. If you have the impression it is, it is more likely than not that you oversimplified something.

Also, quantum mechanics is mostly applied linear algebra, with a slight touch of analysis and differential equations. Especially, the basic formulations are entirely linear algebra. Thus, a good foundation of linear algebra is indispensable for doing quantum mechanics. In particular, quantum mechanics requires firm command of (infinite-dimensional) vector spaces of functions and of the eigenvalue problem. Essentially, technically quantum mechanics boils down to the eigenvalue problem of a particular kind of differential operators. It will be attempted to highlight these requirements, but this cannot replace a familiarity, e. g. with the content of my lecture on linear algebra, or of a reasonably good textbook.

In addition, a few of the advanced concepts of classical mechanics, especially Hamilton's mechanics with the Hamilton function and canonical conjugated variables, will be central. Their quantum analogue is the simplest, but by no means simple, formulation of quantum mechanics. Thus, familiarity with them is indispensable. And, of course, the harmonic oscillator will reappear throughout.

Chapter 2

The two-state system

Quantum mechanics is a vast field, and requires a substantial expansion in how to describe physics in comparison to classical physics, including conceptually. It is therefore very useful to first start out with a comparatively simple example to line out how quantum mechanics actually works, before setting up things on a more formal basis.

While it is indeed possible to start out with a set of, genuinely new, postulates and derive everything, the basic postulates are very different from what one has seen in classical mechanics. Accepting them as a basis for description comes therefore much less natural than, say, Newton's laws. Even the Hamilton-Jacobi formalism, as abstract as it may seem, can be derived step-by-step from Newton's law, and therefore feels natural. This is not the case for quantum physics, and therefore they feel much more strange.

So, it is best to start out relatively simple.

2.1 The Stern-Gerlach experiment

2.1.1 An unexpected result

One of the first experiments to require quantum mechanics to explain is the Stern-Gerlach experiment.

The basic idea of the Stern-Gerlach experiment is comparatively simple. It starts out with a particle source for particles which have a magnetic moment $\vec{\mu}$ of fixed size $|\vec{\mu}|$. The source has no particular orientation. The particles are emitted by the source in a tightly focused beam in z direction.

The particles are then send through a static, magnetic field in x direction for some fixed distance Δz. Classical electrodynamics then dictates that the particles should experience over the distance Δz a force in x-direction $\sim -\mu_x \partial_x B_x$. Thus, depending on the orien-

tation of their magnetic moment they should be more or less deflected by the magnetic field. Thus, if the initial orientation is random, the deflection should be correspondingly distributed. If detecting the particles later along their way, their distribution along the z-axis should be roughly Gaussian distributed, with the maximum along the beam axis.

This is, however, not what is observed.

What happens is rather that the beam is split into two beams of the same intensity, symmetrically around the beam axis. They are well separated, provided they traveled long enough after exiting the magnetic field. This would be only expected if the magnetic momentum would have been polarized before either in positive or negative x-direction. But this did not happen. There is no explanation in classical mechanics and/or electrodynamics which can account for this phenomenon. It is due to quantum mechanics. In fact, there are many other experiments showing correspondingly behaviors unexplainable with classical physics.

The best explanation is that the system has not a continuously distributed magnetic moment, but a discrete one with values only $\mu_x = \pm c\hbar/2$. Herein c is the maximum value of the magnetic moment (essentially a suitable magneton), $\hbar$ is a new constant (with value ($\approx 6.58 \times 10^{-16}$ eVs) with units of an angular momentum or action, and the $\pm$ give the directions for the two different beams. The constant is also called Planck's constant, and its size measures the importance of quantum effects. As will be seen, whenever the characteristic scales are large compared to $\hbar$ classical physics is encountered, and otherwise quantum effects become important[1].

This is then the first surprise: There are two possibilities, rather than a continuum. This is where the quantum comes from: The system is quantized rather than continuous. This is one of the two fundamental important differences between quantum physics and classical physics.

The other difference surfaces when considering sequential Stern-Gerlach experiments.

2.1.2 Sequential Stern-Gerlach experiments

Because of the splitting in two beams, it is, of course, possible to remove one of the two beams. Within classical physics, the following three possibilities can be thought of.

If after the first Stern-Gerlach experiment one of the beam is removed, and the same experiment is repeated, then again only one beam should show up. This is also observed in experiment.

If after the first Stern-Gerlach experiment one beam is removed and afterwards a second Stern-Gerlach experiment is performed, but rotated through $\pi/2$ into the y direction, it

[1] With some exceptions due to numerical effects in both directions.

is expected that again two beams should appear. Of course, the intensity is reduced, as some of the beam is removed, but the distribution on the two final beams is the same. After all, the y and x directions should be independent. This is also observed.

But now remove again one of these beams, and send the remaining beam through a third Stern-Gerlach experiment, again like the first in x direction. Because one of the beams has been eliminated before, this should again only yield a beam in the non-eliminated direction. But this is not observed. Rather, again two beams appear. In classical physics, this should not happen, as indeed the y and x direction should be independent. This is no longer the case in quantum physics - measurements are different. This is the second big difference. It seems that measuring in y direction erases whatever was known in the x direction. It is said that both information cannot be measured simultaneously.

It is probably the latter property of quantum physics what makes it so hard to come to terms with, and not so much the first one. Still both shape how quantum physics works.

2.2 A formulation using linear algebra

It is now necessary to find a more formal description of what is going on. This is a first step towards the quantum version of Newton's law, or more accurately the quantum version of Hamilton's mechanics. It must implement the quantization and that both directions are not independent.

The quantization is a qualitative statement. It is not something which can be smoothly obtained from a continuum. It thus requires something which is conceptually of a similar kind. But real numbers are unsuited for such a thing. So the simplest possibility is to associate it with directions. On the other hand, there should be no independent information on the y direction.

This suggests to use a two-dimensional vector space. Both directions are discrete and independent. Define the two beams after the first Stern-Gerlach experiments as two vectors[2]

$$\begin{aligned} s_+^x &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ s_-^x &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

The measurement and detection of the beams is now done using the unit matrix, $M_x = 1$. Thus, this a combination of the Stern-Gerlach experiment and the screen. This emphasizes

[2]Vectors will be written as small letters without $\vec{}$, and matrices (and later operators) as capital letters.

the very important point that just performing the Stern-Gerlach experiment does not provide a meaningful statement about the system, but a measurement is required. This subtle distinction is very important in quantum physics, and will be returned to in chapter 11.

The blocking of the one beam can then be represented by a matrix

$$B_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

which satisfies $B_x M_x s_+^x = s_+^x$ and $B_x M_x s_-^x = 0$. So this represents the situation after the first block. It could be said that $B_x M_x$ is a linear operator[3] mediating the blocking of one of the beams. Thus, the beam could be beforehand any mixture, e. g. a homogeneous one $b = s_+^x + s_-^x$, and afterwards only one component, s_+^x, would emerge.

The next step is how to represent the second Stern-Gerlach experiment. This requires a description of the y direction. These should again be two independent directions, but should not be correlated with the previous directions. The simplest possibility is

$$\begin{aligned} s_+^y &= \frac{1}{\sqrt{2}}\left(s_+^x + s_-^x\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ s_-^y &= \frac{1}{\sqrt{2}}\left(s_+^x - s_-^x\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

They can be obtained from the beam by the matrices

$$\begin{aligned} M_+^y &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ M_-^y &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned}$$

such that $s_+^y = M_+^y B_x M_x b$ and $s_-^y = M_-^y B_x M_x b$. It is important to note that these operations do not describe the beam directly, but rather the two beams separately.

Blocking out now either of the beams can be obtained using

$$B_y = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

which yield $B_y M_+^y B_x M_x b = s_+^y$ and $B_y M_-^y B_x M_x b = 0$.

[3] A linear operator is any operator, here a matrix, satisfying $M(a+b) = Ma + Mb$, $M(\alpha a) = \alpha M a$, and $M0 = 0$.

Now, finally acting again with a Stern-Gerlach experiment yields $M_x B_y M_+^y B_x M_x b = s_+^y \sim (s_+^x + s_-^x)$, and thus again a homogeneous mixture of both directions, just as observed in the experiment. Hence, representing the beam by a vector and both the experiment and the measurement by matrices did work in giving a description. This will continue onward: The system will be described by vectors, and an action on it by linear operators/matrices. Just the vector space will become more complicated.

This completes the mathematical description of the observation. There are two remarkable differences compared to ordinary mechanical systems. One was that the change in the system was not because of a continuous time development, but by the projection with matrices. The other is that in classical mechanics all measurements are ordinary functions. These commute. Here, measurements are connected to non-commuting matrices[4]. Thus, making the same measurement in different order matters

$$(M_x((B_y M_+^y)(B_x M_x) - (B_x M_x)(B_y M_+^y)))b = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} b = \frac{1}{\sqrt{2}}(s_+^x - s_-^x).$$

In classical mechanics, this should yield zero. Here, it does not. This explains why the order matters, and why the result differs, if the measurements are done in different orders. It is this non-commutativity which lies very much at the heart of what quantum physics is, and how it differs from classical physics.

[4]Two matrices A and B are said to be non-commuting, if $AB - BA \equiv [A, B] \neq 0$.

Chapter 3

Postulates and formulation

In the following, the basic ideas and postulates of quantum physics will be formulated. This will be done iteratively more than once, highlighting different aspects in the process. Especially, this will continued in chapter 11, which will give different possibilities how to formulate the basic postulates, and how they relate.

3.1 The state postulate

To better understand what is actually happening, it is useful to take a step back, and analyze the various components which were involved in the description.

In this process, it is useful to also find a suitable mathematical framework to describe quantum physics.

In the Stern-Gerlach case it became clear that the system is in a certain state at any given instance of time: The state was the composition of the beam. This state was described by a vector. It thus appears useful to introduce that the state of a quantum system is described by vectors in a vector space. Since making statements about the system will certainly require some kind of scalar product, it is postulated that this vector space can be upgraded to a Hilbert space[1]. As it turns out, it will be necessary in general that this Hilbert space is a complex one, i. e. components of vectors and scalars are complex numbers. So it will be treated as such from the outset.

However, there is one notational complications: There are now two vector spaces in the problem. One vector space is the ordinary position space, with position vectors $\vec{x}$. The states are also vectors, but in a different vector space, the state space[2]. To distinguish

[1] A Hilbert space is a vector space with a scalar product and an induced norm derived from this scalar product.

[2] This state space replaces the state space of classical mechanics. Thus it has been given the same

both vector spaces it is useful to have a different notation for vectors from the state space than for vectors from the position space, even though both are mathematically vectors. Therefore a state vector of the state i will be denoted by $|i\rangle$. It will also be given the name ket, rather than vector, to emphasize this difference. But mathematically the kets remain nothing but vectors. In fact, the vectors s_+ and s_- have been such kets, and would now be denoted as $|s_+\rangle$ and $|s_-\rangle$.

With this notation at hand, it is time to formulate the first postulate of quantum physics[3]:

> Everything which can be known about a system is contained in its state. The ray of a state ket is assumed to contain all information about the state. In particular, any measurements performed on a system will be obtained from manipulation of the state ket. This implies that the absolute length of the ket does not contain physical information, only its direction.

3.2 Properties of kets

Because kets are just vectors, they can be added or subtracted. Since this changes the direction this has physical implications. This will lead later on to the concept of superposition. Note however that when adding or subtracting kets, relative lengths will play a role.

Furthermore, even if physical systems are quantized, it can be expected that many physical systems can have a, denumerable or not, infinite number of possible states. Thus, the Hilbert space in question will usually not be a finite one, even though this was true for the Stern-Gerlach experiment.

Since distinguished states are distinguished directions any interaction changing the state of a system needs to be represented by change of a ket. The simplest non-trivial possibility is by acting with linear operators[4] on a ket, i. e. $A|i\rangle$ with A a linear operator. As was shown in linear algebra, such linear operators can always be represented by (infinite-dimensional) matrices acting on vectors. The M and B matrices had been examples of

name.

[3] A ray is the set of all vectors which can be obtained by multiplying a vector with any scalar, i. e. $\{a\vec{v}|\forall a \in \mathbb{C}\}$.

[4] Of course, it seems possible that physics could also proceed via non-linear operators. So far, experiments do not suggest that this would be necessary, although it is not conceptually excluded. During this lecture, only the so far experimentally established (and technically much simpler) case of linear operators will be considered.

such linear operators. This already indicates how the manipulation of the state proceeds by linear operators.

As usual, linear operators can be added and subtracted, with associativity holding. Linear operators can have all the properties one is used to, like hermiticity[5], unitarity[6], and so on. Several of these features will be associated with particular physics in the following.

The first of such features is that a linear operator can have eigenkets[7] $|a\rangle$, i. e. kets $|a\rangle$ satisfying $A|a\rangle = a|a\rangle$, where a is a number, called the eigenvalue. If an operator has more than one eigenket, they are usually denoted by $|a_i\rangle$ where the a_i are the different eigenvalues. Sometimes the index will even be dropped, and then $|a\rangle$ is the set of all eigenkets and a the set of all eigenvalues. Eigenkets to a linear operator are sometimes also called eigenstates. Note that if the linear operator is hermitian, its eigenvectors form a basis of the Hilbert space and its eigenvalues are all real[8].

Since it was assumed that the space is a Hilbert space, it is necessary to define the corresponding scalar product. In analogy to an ordinary complex vector space, this will be created by introducing a uniquely defined dual or adjoint vector $c^*\langle i|$ to any ket $c|i\rangle$, and call it a bra[9]. In an ordinary complex Hilbert space $\mathbb{C}^n$, this would be just the hermitian conjugate $(c\vec{v})^\dagger$ of a vector $c\vec{v}$. However, because of the possibility of non-denumerable-dimensional spaces, it is mathematically not always possible to create the dual vector by just complex transposing it. Thus the different notation and name. This more subtle point will, however, rarely be encountered during this lecture.

The scalar product will then be any map $\langle i|j\rangle \to \mathbb{C}$ such that the map satisfies the conditions of a positive definite scalar product[10]. In the Stern-Gerlach case, this was really just the ordinary scalar product. The norm in the vector space will then be derived as usual from the scalar product. Since the absolute length of a ket has no physical significance, kets will be assumed normalized, $\langle a|a\rangle = 1$, if not stated explicitly differently.

A linear operator acting on the bra will be denoted by $A^\dagger$. However, in general $(A^\dagger)^\dagger \neq A$ in an infinite-dimensional Hilbert space. If $A = A^\dagger$ then the operator is said to be Hermitian, as usual. For the sake of notational simplicity, a linear operator is taken to act to the right on a ket, $A|i\rangle$, and to the left on a bra, $\langle i|A$. Especially, when doing a scalar

[5]Hermiticity implies $A = A^\dagger$. For infinite-dimensional vector spaces there can be subtleties involved, which will not be relevant in this lecture. This requires the more complete concept of self-adjointness.

[6]For an invertible linear operator unitarity implies $A^{-1} = A^\dagger$.

[7]Operators will be denoted by capital letters and eigenvalues as the same small letter.

[8]This is the so-called spectral theorem, and is proven in the linear algebra lecture.

[9]Because the scalar product $\langle a|b\rangle$ contains the bra-kets $\langle\rangle$ this name was chosen.

[10]Essentially $\langle (cu)|(av) + (bw)\rangle = c^*(a\langle u|v\rangle + b\langle u|w\rangle$, $\langle v|w\rangle = \langle w|v\rangle^\dagger$, $\langle v|v\rangle \geq 0$, $\langle v|v\rangle = 0 \leftrightarrow |v\rangle = |0\rangle$, where $|0\rangle$ is the neutral element under vector addition. It induces the norm $\||a\rangle\| = +\sqrt{\langle a|a\rangle}$

product on a modified vector, this can be written as

$$\langle i|(A|j\rangle) = (\langle i|A^\dagger)|j\rangle = \langle i|A|j\rangle,$$

where in the last case it is understood that the linear operator acts to the right.

Since linear operators can be mapped to matrices, it is always possible to define a multiplication AB, which is associative, but not necessarily commutative. Note that just as for matrices, linear operators obey $(AB)^\dagger = B^\dagger A^\dagger$.

In physics terms, AB is a linear operator which represents acting on a state first by B and then by A. If the product is non-commutative, this is in general different from BA, first acting with A and then with B. Thus, even though the state contains all information about the system, the order in which something is done to a system matters! This is what has already been seen in the Stern-Gerlach experiment, and which is one of the central differences from classical physics. In fact it is this non-commutativity which really lies at the qualitative heart of the difference between quantum physics and classical physics.

The order of the manipulations does not matter if the two linear operators commute. It is very convenient to test this by evaluating the so-called commutator

$$[A, B] = AB - BA$$

which encodes this fact compactly without making reference to the states. This will become even more important as will be seen that the commutator is, in fact, intimately related to the classical Poisson brackets, despite being an algebraic rather than a differential operator.

There is also a particular class of linear operators. They are created if two states $|a\rangle$ and $|b\rangle$ are given, and they are denoted as $|b\rangle\langle a|$. They act on a state by determining first the overlap with a and then rescaling the state b by this overlap, i. e.

$$(|b\rangle\langle a|)\,|c\rangle = (\langle a|c\rangle)|b\rangle.$$

Since the state a has, as a vector in a vector space, a direction, this is actually a projection operator, as it will only be non-zero if the state on which it is applied has any component in the direction of a. Furthermore

$$(|b\rangle\langle a|)^\dagger = |a\rangle\langle b|$$

and will therefore exchange the roles of a and b. In the sense of linear algebra, what is formed is actually a tensor product[11] of the states $|a\rangle$ and $|b\rangle$.

[11] In terms of conventional vectors and matrices, this tensor product is the matrix formed from the vectors $\vec{a}$ and $\vec{b}$ as $M_{ij} = b_i a_j$.

So far, what has been done is the construction of states and the transformation of states. However, there is still the need to understand how to determine observables (and to describe the measurement process).

3.3 Hermitian operators

As it will turn out, observables are intimately connected to Hermitian linear operators, i. e. operators satisfying $A = A^\dagger$. It is worthwhile to recall from linear algebra that the spectrum, i. e. set of eigenvalues a_i, of a Hermitian operator is purely real, and that its eigenvectors/kets form a complete basis, where eigenkets to different eigenvalues are orthogonal. Using the Gram-Schmidt procedures, also the possibly non-orthogonal eigenkets to degenerate eigenvalues[12] can be included to finally create an orthonormal basis, i. e. the eigenkets satisfy $\langle a_i|a_j\rangle = \delta_{ij}$ with δ_{ij} the usual Kronecker-δ.

As a consequence, any state $|s\rangle$ can therefore be decomposed into the eigenstates $|a_i\rangle$ of a Hermitian operator A,

$$|s\rangle = \sum_i c_i|a_i\rangle, \tag{3.1}$$

where the expansion coefficients can be obtained from the orthonormality of the basis as $c_i = \langle a_i|s\rangle$. They are called the overlaps of the state s with the states a_i, but they depend, of course, on the choice of basis, i. e. of the operator whose eigenbasis is used. This insight allows to rewrite (3.1) as

$$|s\rangle = \sum_i |a_i\rangle\langle a_i|s\rangle,$$

which corresponds to the spectral decomposition of Hermitian matrices.

However, this also implies

$$1 = \sum |a_i\rangle\langle a_i|.$$

This so-called decomposition of unity is an extremely helpful technical trick, which will

[12] An eigenvalue is degenerate if it appears more than once, i. e. having an algebraic multiplicity of more than one. Because of the spectral theorem the eigenkets form a basis, with eigenkets to different eigenvalues orthogonal. However, if there a multiple eigenvectors to the same eigenket, they do not need to be orthogonal, but need to form a basis of the corresponding subspace with the dimension of the algebraic multiplicity, i. e. having the same geometric multiplicity. As the basis in this subspace is arbitrary, it can always be chosen orthonormal. Note that in this case the eigenkets, in principle, require a second index $|a_i^j\rangle$, enumerating the basis vectors j to the degenerate eigenvalue i, thus satisfying $\langle a_i^j|a_k^l\rangle = \delta_{ik}\delta_{jl}$, where formally for a non-degenerate eigenvalue the second index only takes a single value. However, this second index is usually suppressed, and one uses a single index to denote two different indices, a so-called multiindex.

be used very often in the following[13]. This relation is sometimes also called completeness relation or closure.

To illustrate the use of this decomposition, consider

$$\langle s|s\rangle = \left\langle s \left| \sum_i |a_i\rangle\langle a_i| \right| s \right\rangle = \sum_i |\langle a_i|s\rangle|^2 = \sum_i |c_i|^2, \tag{3.2}$$

which implies that the last sum equals the length of the stateket. If the stateket is normalized, this value is one.

This also allows to create matrix representations of any linear operator X by

$$X = \sum_{ij} |a_i\rangle\langle a_i|X|a_j\rangle\langle a_j|$$

and therefore matrix elements

$$\bar{X}_{ij} = \langle a_i|X|a_j\rangle.$$

of the matrix[14] $\bar{X}$ representing the linear operator X.

3.4 The observables postulate

It is now the time to formulate the next central postulate of quantum mechanics. This is an operational definition for now. In section 11.4 the background of this postulate will be critically appraised, but for all practical purposes it is entirely sufficient. It is:

> Any observable quantity is represented by a Hermitian linear operator. A measurement of this observable will project a state $|s\rangle$ on an eigenstate of this operator and will yield as the measured value the corresponding eigenvalue. This process is not deterministic, and the probability to be projected to any given eigenstate is given by $|\langle a_i|s\rangle|^2$, provided the state and eigenkets are normalized to one, $\langle s|s\rangle = 1$ and $\langle a_i|a_j\rangle = \delta_{ij}$.

The determination of probabilities using (absolute values of) scalar products is also known as Born's rule.

[13]The existence of this decomposition has actually a few more subtleties attached, which will be skipped over here, as they will not be necessary.

[14]Using this it is entirely possible to rewrite the whole of quantum mechanics in terms of ordinary (complex) vectors and matrices, yielding the so-called Heisenberg's matrix mechanics. It is mathematically fully equivalent to the operator language developed here, but has for various reasons not become the language of choice for quantum mechanics, but is occasionally useful.

This encodes the non-deterministic nature of quantum physics. As it is a postulate, this cannot be derived. However, any experimental test of this postulate (and its consequences) so far have not falsified it[15]. This has no parallel to any kind of the deterministic classical physics, and is therefore usually hard to swallow.

Therefore, this postulate will be reexamined more thoroughly to understand what it may signal and what may be behind it in chapter 11. There is also a lot of, not fully solved, problems associated with this postulate, which run very deep into the conceptual way we understand the world around us. In particular, it separates the measurement from the system. I. e. measuring an observable is not described by the formulation of quantum mechanics itself, but as an outside influence. Thus, it gives up at the moment the idea of classical mechanics that there is some 'world-equation' which describes everything, including the measurement process and the measurement apparatus, and person doing the measurement. It should therefore be considered to be an operative definition of measurement of a system isolated from the measurement apparatus.

This problem will be taken up again in chapter 11. However, to do so requires some acquaintance with its consequences and what this operative definition mathematically implies. Therefore it is best to accept it for the moment like a mathematical axiom, and come back to it later when the technical tools are available to dissect it more thoroughly.

Thus a measurement acts as a projection to some (normalized) eigenstate $|a_m\rangle$ of an observable for a given state $|s\rangle$,

$$|s\rangle \to |a_m\rangle,$$

with probability $|\langle a_m|s\rangle|^2$. Since the projection occurs with a probability proportional to the overlap this implies that a state will never be projected into an eigenstate with which it has zero overlap. This makes sense, as if the overlap is zero this just says this state has nothing to do with this eigenstate. Conversely, a state which is already represented by an eigenstate of the observable will remain so, no matter how often the measurement is performed.

This has a very drastic consequences: Any measurement of a state which is not an eigenstate of the observable will change the state. This is meant when it is said that in quantum physics a measurement also changes the state it measures. It is an exact unavoidable mathematical consequence of this postulate. Thus, any non-eigenstates change during a measurement. This may be an arbitrary small change, but a change nonetheless.

This also shows why the observable needs to be Hermitian: Only then form the eigenstates a complete basis, and therefore a state cannot be 'lost' because it has no overlap with any eigenstate of the observable. In addition, only then can the overlap sum (3.2)

[15] As always, verification is conceptually, to the best of our knowledge, impossible.

sum to 1 with positive semi-definite elements each, and therefore can be interpreted as probabilities.

On the other hand, if a state is not an eigenstate, it has components in the direction of multiple eigenstates of the observable. It is said to be in a superposition with respect to this observable. Note that this statement depends on the observable in question. As long as the state is not the zero state there will exists Hermitian operators to which it is an eigenstate, and therefore not be a superposition with respect to them.

Note that from a theoretical perspective it makes sense to say that a (single) state is thrown into an eigenstate with a certain probability, this is not so from an experimental point of view. There, with a single experiment one state just goes into a particular eigenstate, and this for this single measurement with certainty. Only when repeating the experiment with another (identical) starting state[16] it may go to a different eigenstate. And only by repeating the experiment a large enough amount of times it is possible to start measuring probabilities. This is also what makes experiments in quantum physics much more cumbersome than in classical physics. In classical physics, provided the experimental precision is good enough[17], a measurement needs only to be done once, as the outcome is deterministic.

While the actual outcome of a single measurement is not deterministic, it is at least possible to determine what the average outcome of a measurement of the observable is. This will be just the probabilistic expectation value, which is the average of the eigenvalues of the observable weighted by the probability to obtain the eigenvalue when measuring the state s. It is denoted by[18] $\langle A\rangle_s$ and given by

$$\begin{aligned}\langle A\rangle_s &= \sum_i a_i|\langle a_i|s\rangle|^2 = \sum_i a_i\langle s|a_i\rangle\langle a_i|s\rangle = \sum_{ij} a_i\langle s|a_j\rangle\langle a_j|a_i\rangle\langle a_i|s\rangle \\ &= \sum_{ij}\langle s|a_j\rangle\langle a_j|a_i|a_i\rangle\langle a_i|s\rangle = \sum_{ij}\langle s|a_j\rangle\langle a_j|A|a_i\rangle\langle a_i|s\rangle = \sum_{ij} c_i c_j^*\langle a_j|A|a_i\rangle = \langle s|A|s\rangle.\end{aligned}$$

It can therefore be separated into the matrix elements $\langle a_i|A|a_j\rangle$, which do not depend on the state, and in the overlaps c_i, which depend on the state. Alternatively, it is the matrix element[19] $\langle s|A|s\rangle$ of the observable with respect to the state s.

Note that so far the measurement process has been constructed by projecting into a single eigenstate of the observable A. However, for any eigenstate there exists a different

[16] Such an ensemble of identical states is also known as a pure ensemble.

[17] Including effects from chaos, which are just mechanical systems with exponential sensitivity to initial conditions.

[18] Usually, it is clear from the context in which state the expectation value is taken, and then the subscript s will be dropped.

[19] In a $\mathbb{C}^n$, this is the same as $\vec{s}^\dagger A\vec{s}$.

observable A' for which this eigenstate of A is a superposition of its own eigenstates. Thus, a measurement can also be regarded as a measurement throwing a state into this superposition with respect to the observable A'. Thus, whether the final state after a measurement is an eigenstate or a superposition depends on with respect to which observable the result of the measurement is regarded.

3.5 Stern-Gerlach experiment revisited

To get better acquainted with the formalism so far, it is useful to express the Stern-Gerlach experiment of chapter 2 in this language.

The first step is to recognize that there are only two states the system can have, being magnetic moment 'up' or 'down'. Thus, the system has only two states, denoted by the kets $|x+\rangle$ and $|x-\rangle$, where x denotes the direction of the magnetic field. The state space is thus two-dimensional.

Now a Stern-Gerlach filter in y direction splitted the beam in two equally strong beams, the states $|y\pm\rangle$, even if it was before in the $|x+\rangle$ state. Hence, every atom was send into either states with equal probability. Therefore

$$|\langle y + |x+\rangle|^2 = |\langle y - |x+\rangle|^2 = \frac{1}{2},$$

as the states are assumed to be normalized.

But because there are only two states possible, these states must be linear and orthogonal combination of the others,

$$|y\pm\rangle = \frac{1}{\sqrt{2}}|x+\rangle \pm \frac{e^{i\delta}}{\sqrt{2}}|x-\rangle,$$

where δ is an arbitrary phase, which cannot yet be determined. Note that an overall phase is also allowed, even if the state is normalized, and which has been set to one by convention. The corresponding measurement operator is therefore the projection operator

$$S_y = \frac{\hbar}{2}\left(|y+\rangle\langle y+| - |y-\rangle\langle y-|\right) = \frac{\hbar}{2}\left(e^{-i\delta}|x+\rangle\langle x-| + e^{i\delta}|x-\rangle\langle x+|\right).$$

Of course, it does not need to end here. Abstracting from the experimental situation, a corresponding set $|z\pm\rangle$ can also be introduced for the third direction, introducing a new phase α, but otherwise the same.

This gives actually an interesting insight. Because no direction is preferred[20]

$$|\langle z \pm |y\pm\rangle|^2 = \frac{1}{2} = \frac{1}{4}\left|1 \pm e^{i(\delta-\alpha)}\right|^2 = \frac{1}{2}$$

must also hold. But this implies $\delta - \alpha = \pm\pi/2$. This shows that it is still not possible to determine absolute phases, but at least relative phases. This also shows that the components of at least some of the vectors need to be necessarily complex to describe the experiment. This emphasizes the need to use complex vector spaces to describe quantum physics[21].

It is interesting to consider now the three measurement operators S_i, formed from the three directions. The fact that the order of these measurements matter can be seen from calculating their commutator. E. g.

$$\begin{aligned}
[S_y, S_z] &= \frac{\hbar^2}{4}\left(\left(|y+\rangle\langle y+| - |y-\rangle\langle y-|\right)\left(|z+\rangle\langle z+| - |z-\rangle\langle z-|\right)\right. \\
&\quad \left. - \left(|z+\rangle\langle z+| - |z-\rangle\langle z-|\right)\left(|y+\rangle\langle y+| - |y-\rangle\langle y-|\right)\right) \\
&= -\frac{\hbar^2 e^{i(\delta-\alpha)}}{2}\left(|x+\rangle\langle x+| - |x-\rangle\langle x-|\right) = -e^{i(\delta-\alpha)}\hbar S_x,
\end{aligned}$$

where in the second-to-last step all the vectors had been rewritten in the z-basis in a somewhat tedious calculation. There is still an ambiguity with the sign of the differences of the phases. This stems from the fact that it was not defined whether the z and y axis from a right-screw or a left-screw with the x axis. Choosing the usual yields for the difference $\pi/2$, and thus an overall factor of i.

Denoting x, y, z by 1, 2, 3 leads to the generalization

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k. \tag{3.3}$$

This is interesting for two reasons. First, it shows that any change of orientation will affect the measurement. Second, this looks very similar to the expression for ordinary angular momentum in classical mechanics, except for a factor of two and that the Poisson brackets have been replaced by the commutator. This is a pattern, which will turn out to be quite general. It also plays a central role in recovering classical physics from quantum physics.

[20] The actual movement of the beam in the original Stern-Gerlach experiment was a technical device to see the separation. In a gedankenexperiment or using scattering the particles do not need to move, and therefore all three directions are equal.

[21] Of course, any complex vector space can formally be mapped to a real vector space of twice the dimension. Therefore, this could also be phrased by stating that more information is needed than in the classical counterpart - there three real dimensions would be necessary to describe the experiment, but here four real, or two complex ones, are needed. It is also found that with respect to counting, always an even number of real degrees of freedom is necessary, and it is therefore always possible to map everything to a complex vector space of half the dimension, which is also technically superior.

It is interesting to consider the operator $S^2 = S_x^2 + S_y^2 + S_z^2$. Noting that $[S_i, S_i] = 0$ and

$$\begin{aligned}[AB, C] &= ABC - CAB = ABC - CAB + ACB - ACB = A(BC - CB) + (AC - CA)B \\ &= A[B, C] + [A, C]B\end{aligned}$$

yields immediately

$$[S^2, S_i] = 0$$

for any i. Thus, the operator S^2 can be measured simultaneously with each of the S_i. Since it is not linearly dependent on any of the S_i, it is an additional independent operator. In fact, applying it to any state yields

$$S^2|i\pm\rangle = \frac{3}{4}\hbar^2|i\pm\rangle$$

Thus, it is just a multiple of the identity operator, and therefore seems to be a bit trivial. However, it will turn out that it is related with the total angular momentum of the particle. Physical, this means it is possible to measure at the same time the total value of the spin and one of its components. Measuring another component will change the state, while the total value still remains.

To make this explicit, any maximal set of commuting operators are called compatible operators. The individual components are, however, not compatible.

This also leads to another interesting observation. Since the operators S_i are hermitian, their eigenkets form the whole vector space, and the eigenvalues are $\pm\hbar/2$. However, S^2 is also Hermitian, so its eigenkets also form a basis of the whole vectorspace, but its eigenvalues are degenerate twice $3\hbar^2/4$. This degeneracy is lifted by the presence of any of the compatible operators S_i, as their eigenvalues can distinguish between the degenerate states.

This situation is generic. Given any complete set of linear independent compatible operators, each direction can be uniquely identified by a set of eigenvalues, of which at least one is different for every direction. Thus, here a vector is characterized by the eigenvalues of S^2 (same for all directions) and any of the S_i (different for every direction). The reason is that Hermitian operators always have a full basis as eigenvectors, and Hermitian operators form a group under addition. Thus, by a suitable decomposition of a set of Hermitian operators, it is always possible to find a basis in the operator space to achieve this goal. Whether this is physically useful is an entirely different question.

Note that this does not imply that the measurement of two compatible operators does not distort the state. It just does not alter the result of the measurement. Start with two compatible observables A and B, where there is a degeneracy in the eigenvalues of

A, which is lifted by B. E. g. $A = S^2$ and $B = S_z$. Then a measurement of A will yield some eigenvalue a of A. The resulting state will then be a linear superposition of the eigenkets $|a\rangle$ to the eigenvalue a. Measuring afterwards B will then throw the state into some particular eigenket $|a, b\rangle$ to an eigenvalue b, which is entirely spanned by the eigenkets to the eigenvalue a of A. In particular, a third measurement of A will again yield a.

Note that the compatibility of the operators enforces that if b is measured first for B, then so is a for A afterwards. This does not ensure, however, that this is guaranteed to happen. After all, the outcome of the measurement of B is random. If the result is a different eigenvalue b', with eigenvector not spanned by the kets $|a\rangle$, the subsequent measurement of A will not yield a but some other eigenvalue a'. Thus, compatibility of two operators means only interchangeability if measurements are done on a simultaneous eigenstate $|a, b\rangle$. Otherwise the combination of projection and randomness can still give different results.

The situation with incompatible operators, i. e., those which satisfy $[A, B] \neq 0$, is quite different[22].

First, this implies that if A is measured first, and then B, the outcome of a third measurement of A is again random, at least within the subspace spanned by the eigenkets of B the state is in after a measurement of some eigenvalue b. Thus, a consecutive measurement of two incompatible operator makes the system forget in which state it has been prior. This was the case in the Stern-Gerlach experiment when performing the measurement sequence x-block one-y-x, yielding again two beams, rather than one.

An even more dramatic consequence is encountered if there are three incompatible observables A, B, and C. After a measurement of A, the result is some eigenvalue a and a state $|a\rangle$, which is an eigenstate of A. The probability to find it in the eigenstate b of B is then $|\langle b|a\rangle|^2$. Performing afterwards a measurement of C, the probability to find it in state c is

$$P^{abc} = |\langle c|b\rangle|^2 |\langle b|a\rangle|^2 = \langle b|c\rangle\langle c|b\rangle\langle b|a\rangle\langle a|b\rangle.$$

This corresponds to the fact that only the result b in the measurement of B is accepted for being processed by C. If the measurement of B is performed, but all outcomes are just transferred to C, the probability is

$$P^{aBc} = \sum_b |\langle c|b\rangle|^2 |\langle b|a\rangle|^2 = \sum_b \langle b|c\rangle\langle c|b\rangle\langle b|a\rangle\langle a|b\rangle. \tag{3.4}$$

[22] In some cases, $[A, B]|v\rangle = 0$ for some subset of kets $|v\rangle$. These form a subspace, and in this subspace the operators act as if they were compatible. It will be explicitly noted if such a case is dealt with.

Intuitively, this may seem as if B would be absent. But this is not the case. Consider a setup, where only A and C is present. The probability would then be

$$P^{ac} = |\langle a|c\rangle|^2.$$

To make it comparable to (3.4), it is possible to expand each state into the (complete eigen)basis of B, $|i\rangle = \sum_b \langle b|i\rangle|b\rangle$, yielding

$$P^{ac} = \sum_{b,b'} \langle b|c\rangle\langle c|b'\rangle\langle b|a\rangle\langle a|b'\rangle. \tag{3.5}$$

This is different from (3.4) by the fact that different overlaps of the states a and c with b appear, rather than the same. In fact, it can be shown that if at least two of the three operators are compatible it is possible to expand one of the offending overlaps such that the two expressions again coincide, but otherwise a term involving the commutator remains. Thus, if a measurement is performed or not, even if its result is discarded, matters for incompatible observables in quantum physics! A measurement cannot be undone.

3.6 The uncertainty relation

The probably most famous consequence of the incompatibility of operators is Heisenberg's uncertainty principle for position and momentum. This will appear later. This uncertainty principle is actually just a special case. In fact, any incompatible operators always induce a uncertainty principle. This will be developed here fully. This requires some preparatory definitions for single operators.

Consider first some operator A. In the following, and generally throughout the lecture, it is often necessary to talk about the expectation value of the operator A in some (normalized) state, but which state is either irrelevant or it is uniquely given from the context. In the latter case this is most often the vacuum state. At any rate, in such a situation the state will not be explicitly noted, and the expectation value will be just denoted as $\langle A\rangle$.

The uncertainty principle will be a statement about our ability to measure an operator. This is intrinsically linked to the fact that quantum systems can yield a random result in a single measurement. It is thus necessary to quantify this effect. To do, it is useful to define a fluctuation operator ΔA for any given operator A as

$$\Delta A = A - \langle A\rangle.$$

In particular, $\langle \Delta A\rangle = 0$, by definition, provided the fluctuation operator is evaluated in the same state. If it is evaluated in a different state than the expectation value of the

original operator, this does not need to be the case. But if, the expectation value of any fluctuation operator is zero. Physically, this means that the fluctuation around the expectation value of the original operator average themselves out.

This does not mean there are no fluctuations, only that they do cancel each other on average. They can still have substantial size. To measure this size, it is useful to define the dispersion of the operator A as

$$\langle(\Delta A)^2\rangle = \langle A^2 - 2A\langle A\rangle + \langle A\rangle^2\rangle = \langle A^2\rangle - \langle A\rangle^2,$$

where it has been used that the expectation value of a number is just the number itself, no matter the state.

If the state should be an eigenstate of A, the dispersion vanishes, as $a^2 - a^2 = 0$. If it is not an eigenstate, then this is generally not true. E. g. consider the case $(|a\rangle + |a'\rangle)/\sqrt{2}$, which yields

$$\langle(\Delta A)^2\rangle = \frac{1}{2}(a^2 + a'^2) - \frac{1}{4}(a + a')^2 = \frac{1}{4}\left(a^2 + a'^2 - 2aa'\right).$$

As expected, it vanishes for $a = a'$. E. g., for the S_i, the result in such a superposition state would be $\hbar^2/4$. The appearance of $\hbar$ here is ominous. Apparently, if $\hbar$ would be zero, this effect would vanish. This will play an important role when recovering classical physics in section 5.7.

In some sense, the dispersion measures the effect of superposition. In loose language it could be said that it describes the fuzziness of the state. This is particularly apt for the often encountered situation in practice, that the state is a superposition, but with one state being dominant, in the sense of having a coefficient with norm much larger than all other contributing states. The dispersion is then a measure of how much other states contaminate this dominant state.

Note that since observables are Hermitian operators, and so are their squares, their expectation values are necessarily real. Thus, the dispersion is necessarily positive semi-definite for observables.

After these preliminaries, it is possible to construct the uncertainty principle. Consider two operators A and B describing observables. The uncertainty principle is then the statement

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq \frac{1}{4}\left|\langle[A, B]\rangle\right|^2. \tag{3.6}$$

Before proving this to be correct. It is useful to first understand it. It states that, given some fixed state, the product of the dispersions of the two observables is necessarily larger than the expectation value of their commutator. Thus, for compatible observables the product of the dispersions can be zero. However, for incompatible observables this can

only happen for states with $\langle [A, B] \rangle = 0$, but not in general, implying that in general both dispersions are larger than zero. There are states in which both operators necessarily have a non-vanishing dispersion. As has been seen, the dispersion is zero if the state is an eigenstate. Thus, this implies that for incompatible observables there are states which are not eigenstates of both operators, nor of whatever their commutator is. Alternatively, this implies that the dispersion of the individual operators can only become arbitrarily small if at the same time the dispersion of the incompatible observable is becoming arbitrarily large, and their product is bounded from below by a measure of their incompatibility.

Take as an example the operators S_x and S_y. Their commutator is proportional to S_z. Take as a state an eigenstate of S_z. Then the right-hand-side becomes $(\hbar \cdot \hbar/2)^2/4 = \hbar^4/16$. Thus, the product of the dispersions on the operators on the right-hand side needs to be larger than that. Using that

$$S_x|z+\rangle = \frac{\hbar}{2}\left(|z+\rangle\langle z-| + |z-\rangle\langle z+|\right)|z+\rangle = \frac{\hbar}{2}|z-\rangle$$

and thus

$$S_x^2|z+\rangle = \frac{\hbar^2}{4}\left(|z+\rangle\langle z-| + |z-\rangle\langle z+|\right)|z-\rangle = \frac{\hbar^2}{4}|z+\rangle$$

it follows that

$$\langle (\Delta S_x)^2 \rangle = \frac{\hbar^2}{4} - 0 = \frac{\hbar}{4}.$$

Because of rotational symmetry then also $\langle (\Delta S_y)^2 \rangle = \hbar^2/4$. Thus, in this case even equality holds, as both sides take the value $\hbar^4/16$.

Therefore, the uncertainty principle gives a quantitative measure of how precise it is possible to measure two quantities simultaneously. It is a theoretical optimum, to what extent an experiment can be pushed or how large the effect could be in a naturally existing system.

It is now time to prove that this is correct. The first step is that because the operators considered are Hermitian their dispersion relations are real[23]. Then, the triangle inequality guarantees

$$\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \geq |\langle \Delta A \Delta B \rangle|^2. \tag{3.7}$$

A product of any two operators can be rewritten as

$$AB = \frac{1}{2}[A, B] + \frac{1}{2}\{A, B\}$$

in which

$$\{A, B\} = AB + BA$$

[23]This implies that the uncertainty principle will in general not hold for arbitrary operators.

is the so-called anti-commutator. Note that for non-commuting operators also the anti-commutator is in general non-trivial.

Explicitly inserting the definitions of the dispersion operators yields

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

as numbers drop out in commutators, but not in anti-commutators. The commutator of two Hermitian operators is necessarily anti-hermitian, while the anti-commutator remains hermitian. Thus the expectation values are purely imaginary and purely real, respectively. Therefore taking the absolute square yields

$$\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \geq \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{\Delta A, \Delta B\} \rangle|^2. \tag{3.8}$$

Because the second term is positive or zero, it can be dropped while keeping the inequality, at most making it stronger. This is then (3.6), completing the proof.

3.7 A note on transformations

So far, the operators could be mapped to finite-dimensional matrices. Or build from finite-dimensional matrices.

For the corresponding finite-dimensional (complex) Hilbert spaces changes of basis can be performed using unitary transformations[24]. Given any basis a transformation to the eigenbasis of some observable can therefore always be performed using unitary transformations. Conversely, it is possible to diagonalize observables using unitary transformations. I. e. there exists some unitary transformation operator U build from the eigenvectors of an observable A such that

$$U^\dagger A U = D$$

yields a diagonal matrix D with the eigenvalues of A on the diagonal. Since therefore observables which can be transformed into each other by a unitary transformation do have the same eigenvalues they are said to be unitarily equivalent observables. Such observables have no physical distinction on a conceptual level, and are therefore usually taken to be equivalent.

[24] A unitary matrix satisfies $U^{-1} = U^\dagger$. Any basis transformation which preserves lengths and angles between base vectors can be written using as a unitary matrix U acting on the previous basis $\vec{e}_i$ to create the new basis $\vec{f}_i$ as $\vec{f}_i = U\vec{e}_i$.

This has also an interesting implication for compatible observables. Because a (anti)commutator transforms under a similarity transformation as[25]

$$[U^\dagger AU, U^\dagger BU]_\pm = U^\dagger AUU^\dagger BU \pm U^\dagger BUU^\dagger AU = U^\dagger(AB \pm BA)U = U^\dagger[A,B]_\pm U$$

the compatibility of operators is true in any basis. It is said that compatibility is coordinate-invariant or basis-invariant. In addition, it was noted that compatible operators differ by the fact that there exists degenerate eigenvalues of one operator for which the other operator yields the distinction in form of its eigenvalues. The eigenspace of a degenerate eigenvalue of a Hermitian operator is a subspace of the same dimensionality as the degeneracy. It is therefore always possible to select as basis in this subspace of the eigenstates of the second operator. Thus, the second operators also becomes diagonal in this subspace. This can be repeated if there is a third, fourth,...operator to lift the degeneracies. Thus, all compatible observables can be diagonalized simultaneously. At the same time, after diagonalizing this set, there is no more additional observables, which can still be diagonalized, which is not just unitarily equivalent to any observable of the set. This is a consequence of the spectral theorem.

This maximal set of simultaneously diagonalizable observables, and the search for them, plays a central role in quantum physics. It will be encountered again and again.

[25]The notation $[,]_\pm$ is used to denote both commutator and anti commutators with $[,]_- = [,]$ and $[,]_+ = \{,\}$.

Chapter 4

The free particle

Consider, for example, a quantity like the position or the momentum of a particle. Since also in quantum physics particles should still have both properties, and they should be observable, there must be, according to section 3.4, Hermitian operators representing them. But the observed values must be eigenvalues of these operators, and both quantities are experimentally found to range continuously from $-\infty$ to $+\infty$ for a free particle. Thus, the corresponding operators must have a continuous and infinite set of eigenvalues. How to deal with this is the next necessary step before being able to write down a dynamical law for quantum physics. This is necessary, as the two postulates so far were only descriptive of quantum systems. They are like Newton's first law. What is still missing is the equivalent to Newton's second and third laws. And as a first step it is necessary to formulate the equivalence to Newton's second law, and answer how a particle without external influence behaves. This is the goal of this chapter.

Hence, it is necessary to introduce such infinite-dimensional Hilbert spaces. While the statements in the previous chapters are only proven for finite-dimensional Hilbert spaces in the usual linear algebra lectures, they carry over for almost all relevant cases unchanged to infinite-dimensional Hilbert spaces, and generalized linear operators. The corresponding proofs only involve the definitions of the Hilbert space, and requiring that certain appearing infinite series converge. If need be, the corresponding proofs can be found in the mathematical literature, but will be skipped here for the sake of brevity. There is little to be learned from them for their physics application, except in very special circumstances to which I will come back if need be. But since such cases exist one should always keep this in mind.

4.1 Continuous formulation

So, consider first just a position in one dimension. Then there should exist a position operator X, which has eigenkets $|x\rangle$ with a continuous set of eigenvalues x, i. e.

$$X|x\rangle = x|x\rangle.$$

The first oddity to note is that there then needs to be superpositions of eigenkets,

$$X(a|x\rangle + b|y\rangle) = ax|x\rangle + by|y\rangle.$$

The eigenvalues x and y are just (real) numbers, just like (in general complex) coefficients a and b.

The first non-trivial problem is then orthogonality. Since the eigenvectors are now continuous, the Kronecker-δ must be replaced by its continuous version, the Dirac-δ function,

$$\langle x|y\rangle = \delta(x-y).$$

and correspondingly sums by integrals. Especially, the decomposition of unity now reads

$$1 = \int dx|x\rangle\langle x|.$$

Constructing a state out of eigenstates of the position operator therefore reads

$$|a\rangle = \int dx|x\rangle\langle x|a\rangle. \tag{4.1}$$

The overlaps $\langle x|a\rangle$ are therefore (complex) functions of the eigenvalue x of the eigenket $|x\rangle$, $\langle x|a\rangle = a(x)$. These overlaps are called wave-functions for reasons which will become clear in section 4.3. Likewise, a matrix element of an operator A now reads

$$\langle x|A|y\rangle = a(x,y),$$

where it is important to note that a is now a function of x and y. If either $|x\rangle$ or $|y\rangle$ is an eigenstate of A, this will yield $a(x,y)\delta(x-y)$.

All of this can be generalized straightforwardly to more than one dimension. Note, however, that for the eigenstates of a continuous operator to be a basis requires that the continuous operator is Hermitian. We have required that this is the case for X, $X^\dagger = X$, as we demand that the position is observable. Not all continuous operators need to share this property, and several will be encountered later. Also, as has been seen before, nothing of the above guarantees that momentum and position are compatible operators. As will be seen, they are actually not.

4.2 Position and momentum

Space is three-dimensional. Thus, it is necessary to introduce three position operators X_i for the three directions. These can be combined into an operator-valued vector $\vec{X}$, which has a vector-valued eigenvalue $\vec{x}$ when acting on eigenstates $|\vec{x}\rangle$. It is important to keep in mind that the vector $\vec{x}$ is not an element of the vector space which makes up the states. What happens here is that there are now two vector spaces. One with vectors $|a\rangle$, which describes the states, and one with vectors $\vec{x}$, which describe position eigenvalues, and thus the 'ordinary' space. It is for this reason that two different notations are used to denote the two different vector spaces, the bra-ket notation for the vector space of states, and the ordinary notation for the familiar vector space of positions.

To continue towards a dynamical prescription let us require

$$[X_i, X_j] = 0, \tag{4.2}$$

that is that all three directions are described by compatible operators. This is actually a postulate, which is supported by experiments[1]. By this a position can be described. Note that an eigenstate of $\vec{X}$ is located at an exact position. Thus, measuring an object's position will exactly localize it. Of course, in practice 'exact' is never possible due to the instrumental resolution. This leads to the concept of wave packets to be discussed in more detail in section 5.1. For the moment, though, it is sufficient to work with the idealization of an exact measurement.

Besides position, it will be necessary to describe movement. For this purpose, as in classical mechanics, the momentum is better suited than the speed[2]. However, rather than introduce it directly, it is best to take a short detour by asking the question how a displacement in quantum physics should work. After all, a displacement is the first step towards movement. For simplicity, this will be done again in one spatial dimension first.

Because everything is done in a vector space of states, this question should be phrased as: Is there a (linear) operator T, called the translation or displacement operator, depending on a number dx such that

$$T(dx)|x\rangle = |x + dx\rangle,$$

where again only one dimension is considered? Hence, what $T(dx)$ does is, it transforms an eigenstate $|x\rangle$ of the position operator X with eigenvalue x into the eigenstate $|x + dx\rangle$

[1]This is not necessary. Giving up this feature leads to so-called non-commutative geometry, and actually generically implies the existence of a minimum length. Because of the importance of this for the fundamental laws of nature substantial effort is invested in experimental tests of (4.2).

[2]In fact, the canonically conjugated momentum. But during this lecture this will essentially always be just the conventional one.

with eigenvalue $x + dx$. We assume here that the space is just the flat Euclidean space without boundaries[3]. Before determining it, it is useful to list further desired properties of it.

First of all, this translation should not alter physics in such a flat and infinite position space. Thus, matrix elements should not be affected. Given its action on some arbitrary state

$$T(dx)|a\rangle = T(dx)\int dx|x\rangle\langle x|a\rangle = \int dx|x+dx\rangle\langle x|a\rangle$$

this requires

$$\begin{aligned}\langle a|T^\dagger(dx)T(dx)|b\rangle &= \int dxdy\langle a|x\rangle\langle y|b\rangle\langle x+dx|y+dx\rangle \\ &= \int dxdy\langle a|x\rangle\langle y|b\rangle\delta(x-y) = \int dx\langle a|x\rangle\langle x|b\rangle = \langle a|b\rangle\end{aligned}$$

which implies $T^\dagger T = 1$. Thus, the translation operator is unitary. Note that the only necessary ingredient here was that the operator shifts the eigenstate of a Hermitian operator as its only action. Therefore, nothing special about being the position operator has been used, and the derivation that an operator shifting an eigenstate of a Hermitian operator is unitary is general.

Another feature is that translations form an Abelian group classically. There is no experimental hint that this is not also the case at the quantum level, and therefore it is required that

$$T(dx)T(dy) = T(dx+dy) = T(dy)T(dx). \tag{4.3}$$

As classical translations, this should give a group structure. This can be implemented by requiring $T(0) = 1$ and $T(dx)^{-1} = T(-dx) = T(dx)^\dagger$.

The next requirement is that if dx is really infinitesimal, $T(dx)$ should be continuously differ from 1 just by a linear term in dx,

$$T(dx) = 1 - iGdx + \mathcal{O}(dx^2) \tag{4.4}$$

where the, so defined, new operator G is called the generator[4] of the operator T, i. e. the generator of translations. This satisfies by definition all features of the group structure, except for unitarity. Testing unitary yields

$$T(dx)^\dagger T(dx) = (1 + iG^\dagger dx + \mathcal{O}(dx^2))(1 - iGdx + \mathcal{O}(dx^2)) = 1 - i(G - G^\dagger)dx + \mathcal{O}(dx^2),$$

[3]The quantum physics of systems with boundaries, or even other metric, is involved, and beyond the present scope, but extremely important.

[4]In fact, what is required here, is based on a very deep structure of continuous groups, which is discussed in more detail in the advance mathematics lectures during the master studies.

and thus requires $G = G^\dagger$ to comply with unitarity up to the investigated order. Thus, the generator is necessarily a Hermitian operator, and thus describes an observable! Again, this statement only required the unitarity of T and (4.4), and thus is true for all operators satisfying both requirements. In fact, (4.3) follows now

$$(1 - iGdx + \mathcal{O}(dx^2))(1 - iGdy + \mathcal{O}(dy^2)) = 1 - iG(dx + dy) + \mathcal{O}(dx^2, dxdy, dy^2)$$

However, this does only ensure the Abelian nature up to order dx. In fact, for non-Abelian unitary operators this is where the second equality in (4.3) no longer holds.

From these features now follows a very fundamental relation. First note that

$$[X, T(dx)]|x\rangle = (XT(dx) - T(dx)X)|x\rangle = (x + dx)|x + dx\rangle - x|x + dx\rangle = dx|x + dx\rangle.$$

Since the $|x\rangle$ form a complete basis this can be upgraded to an operator relation

$$[X, T(dx)] = dx. \tag{4.5}$$

Thus, translation and measurement of the position do not commute. This is not surprising: After all, if I measure a position, it should matter where I have put the object I measure.

It is important to note that even if this commutator would have vanished, it would not have made X and T compatible, as T is not hermitian, and thus not an observable! But, G is hermitian. Inserting (4.4) in (4.5) yields

$$[X, 1 - iGdx] = -idx[X, G] = dx,$$

and thus

$$[X, G] = i.$$

Thus, whatever observable G is, it is not compatible with X.

This result can be straightforwardly extended to three dimensions, as then $T = 1 - iG_i dx_i + \mathcal{O}(d\vec{x}^2)$ and thus

$$[X_i, G_j] = i\delta_{ij} \tag{4.6}$$

where now again a Kronecker-δ appears, because it is about the indices, and not the position, and the position dimensions are discrete (and finite).

Now (4.6) is strikingly similar to $\{x_i, p_j\} = \delta_{ij}$, the canonical commutation relation of position and momentum of classical mechanics. In addition, in classical mechanics it was possible to identify the momentum as a generator for a canonical transformation which created translation. This suggests to identify G with $P/\hbar$, where the factor of $\hbar$ is used to correct for the different units only. Then (4.6) reads

$$[X_i, P_j] = i\hbar\delta_{ij}, \tag{4.7}$$

where P is now a Hermitian operator, and thus an observable quantity. It is only provisionally that this is identified with the conventional momentum. However, it will turn out that in the classical limit, discussed in section 5.7, and throughout all known systems, the classical limit of the canonical momentum operator is indeed the canonical momentum. Thus this identification is justified.

It should be noted that this is not a postulate. Rather, a name was given to an operator doing something. It is defined to act in a certain way on a certain eigenbasis of another continuous operator. The relation between both operators is entirely abstract. The names given to it were motivational only, but at no point yet there has been made any connection to an actual experiment and thus physically observable phenomenon. Any arbitrary names could have been given to these two operators. It would not have altered their relations. It will require to use them in the description of a physical system to show that their names are justified. It is only there where a postulate will happen.

On the other hand, this also implies that whenever two operators satisfy (4.7), all the other relations follow.

Before continuing, it is useful to collect further consequences which follow for momentum and position (and any other pairs of operators which fulfill (4.7)), as they will be useful later on.

The first is that (4.7) implies that position and momentum are not compatible observables. In particular, (3.6) implies

$$\langle(\Delta X_i)^2\rangle\langle(\Delta P_i)^2\rangle \geq \frac{\hbar^2}{4}. \tag{4.8}$$

This is the celebrated uncertainty principle of Heisenberg. It states that the position and momentum of a particle (along a certain direction) cannot be measured arbitrarily well simultaneously. It is not possible to know at the same time where a particle exactly is and how much momentum it has. While this is fascinating in itself, and is actually at the root of many properties of quantum systems, it should be noted that this is not a large effect, given that the right-hand side is proportional to $\hbar^2$. Thus, it takes experimental effort to be sensitive to this, but it has been established. It should also be noted that the momentum component and the direction have to be parallel. Orthogonal directions are compatible, as the Kronecker-δ in (4.7) signals.

The second interesting consequence is what this implies for a finite translation. Position is continuous and finite translations can be decomposed into infinitesimal translations. Thus a finite translation can de decomposed in n smaller translations for a total distance $n\Delta x$, and ultimately the limit $n \to \infty$ can be taken while keeping $s = n\Delta x$ fixed. Thus, the corresponding translation is obtained by using the translation operator N times, and

taking the limit. This yields

$$T(s) = \lim_{n\to\infty}\left(1 - \frac{iPs}{n\hbar}\right)^n = \sum_n \frac{1}{n!}\left(-\frac{iPs}{\hbar}\right)^n \equiv e^{-\frac{iPs}{\hbar}}.$$

where the last identification is a definition of what it means to exponentiate an operator - much like the conventional matrix exponential. While this has been done in one dimension, this will work equally well in any other number of dimensions by replacing sP by $\vec{s}\vec{P}$, because the different momentum components are compatible observables, as their commutator vanishes,

$$[P_i, P_j] = 0.$$

This follows as they, by definition, do not affect a position eigenstate in any other direction but their own. Hence,

$$T(\vec{s})|\vec{x}\rangle = |\vec{x} + \vec{s}\rangle$$

even for a finite $\vec{s}$.

The third is the insight that because the momentum operator is hermitian, it has also a complete basis of eigenstates, $|p\rangle$, on which it acts as $P|p\rangle = p|p\rangle$. This can again be generalized to more dimensions. Interestingly, these states are eigenstates of the translation operator

$$T(s)|p\rangle = \sum_i \frac{1}{n!}\left(-\frac{iPs}{\hbar}\right)^n |p\rangle = \sum_n \frac{1}{n!}\left(-\frac{ips}{\hbar}\right)^n |p\rangle = e^{-\frac{ips}{\hbar}}|p\rangle. \tag{4.9}$$

Note that the eigenvalues of T, which is unitary but not hermitian, are complex. They are also of absolute value one, as expected for a unitary operator, but not one, and thus the translation operator is not a special unitary operator. This result reassures the provisional interpretation of P as the momentum: A translation does also classically not affect the momentum of a particle.

The fourth is that everything is actually symmetric between momentum and position. Therefore, there exist a momentum translation operator $\exp(iXq/\hbar)$, which changes the momentum of a momentum eigenstate by q. The difference in the minus sign stems from the fact that the commutator in (4.7) is antisymmetric, rather than symmetric[5]. It is also recognizable that the translation operators have the form of a Fourier transformation. Thus, in dealing with position and momentum Fourier transformations will play a central role.

[5]This sign is rooted in the symplectic phase space of classical mechanics.

4.3 Wave function

Having now position and momentum available, a description of the kinematics of a particle is possible. In this context the matrix element $\langle x|a\rangle$ appearing in (4.1) plays an important role.

This matrix element is the projection of a state $|a\rangle$ onto an eigenstate of the position operator. It therefore makes a statement about the state at the position x, which characterizes therefore how the state at x looks like. Likewise, by switching through all possible values of x, this will give a function of x,

$$\psi(x) = \langle x|a\rangle,$$

which describes the state at all points in space. If not clear from the context, the state is often denoted as an index, ψ_a. In particular, since the states $|x\rangle$ form a complete basis, this function $\psi(x)$ encodes all information about the state. Thus, it is possible to completely encode a state in a function of position[6]. Since states are described in a vector space with the complex numbers as field, these functions are necessarily complex-valued[7]. For reasons to become clear later, this function is called the wave function (of the state a).

Combining this with the axiom of measurement in section 3.4, this implies that

$$|\psi(x)|^2 = |\langle x|a\rangle|^2$$

is a probability density for the state to be localized at position x. Furthermore, if the state $|a\rangle$ describes a particle, this is the probability density to find it at the position x. The quantity $|\langle x|a\rangle|^2 dx$ is likewise the probability to find it in the (infinitesimal) region dx around x and

$$P(y, \Delta y) = \int_{y-\Delta y/2}^{y+\Delta y} dx |\psi(x)|^2$$

is the probability to find it in the area Δy centered at the point y. If Δy is send to $\pm\infty$, the probability should become one, as the particle should be somewhere. If the state is normalized to one, i. e. $\langle a|a\rangle = 1$, this is guaranteed, as

$$1 = \langle a|a\rangle = \int dx \langle a|x\rangle\langle x|a\rangle = \int dx |\psi_a(x)|^2.$$

Note that any function, which, up to a multiplicative constant, satisfies this condition is called square-integrable. Wave functions are thus a special case of square-integrable functions.

[6]Not in all quantum-mechanical problems a notion of position makes sense. See, e. g., the Stern-Gerlach case in section 2.1. Then, of course, another basis is more suitable.

[7]This is the reason why the mathematical theory of functions, so-called function theory, plays an important role in modern theoretical quantum physics.

Of course, all of this can be straightforwardly generalized to more than one dimension. This allows also to rewrite general overlaps using the functions ψ,

$$\langle a|b\rangle = \int dx \langle a|x\rangle\langle x|b\rangle = \int dx \psi_a^*(x)\psi_b(x),$$

where it was used that $\langle a|b\rangle$ is a scalar product on the vector space, and thus $\langle a|b\rangle = \langle b|a\rangle^*$ by the requirements for a scalar product. Hence, the overlap of two states can be rewritten in terms of the spatial overlap of the wave functions corresponding to them. Note that this is not a necessary step - the matrix element $\langle a|b\rangle$ has a meaning without using the x eigenbasis. But it is often a very convenient detour to write overlaps in this way.

This leads also to a different perspective on the wave-function. Consider some other observable B. Any state $|a\rangle$ can be decomposed into eigenstates $|b\rangle$ of this observable,

$$|a\rangle = \sum_b |b\rangle\langle b|a\rangle.$$

Note that if b has a continuous spectrum, the sum will again be an integral. Taking a scalar product with $|x\rangle$ yields

$$\langle x|a\rangle = \sum_b \langle b|a\rangle\langle x|b\rangle = \sum_b a_b \psi_b(x) = \psi_a(x).$$

Thus, the wave functions $\psi_b(x)$ of the eigenstates of B can be used to describe an arbitrary state as well as the eigenstates $|b\rangle$ themselves. This is possible, because both are observables. It is, however, not necessary that they are compatible. Using such basis in terms of wave functions will also be a convenient technical tool later on.

Also the expectation values of an operator, as introduced in section 3.4 can now be determined using wave functions,

$$\langle a|C|b\rangle = \int dxdy \langle a|x\rangle\langle x|C|y\rangle\langle y|b\rangle = \int dxdy \psi_a^*(x)C(x,y)\psi_b(y), \tag{4.10}$$

where the two-argument function $C(x,y) = \langle x|C|y\rangle$ is the expectation value of the quantity in question between position eigenstates. In general, if C really depends independently on x and y, such quantities are called non-local. If

$$C(x,y) = c(x)\delta(x-y) = c(y)\delta(x-y),$$

it is called a local quantity, and (4.10) reduces to

$$\langle a|C|b\rangle = \int dx \psi_a(x)^* c(x)\psi_b(x).$$

As an example, consider some state a and take $C = X$. Then $\langle a|X|a\rangle$ is the expectation value of the position operator in the state a. Since

$$\langle a|X|a\rangle = \int dx \langle a|X|x\rangle\langle x|a\rangle = \int dx x \langle a|x\rangle\langle x|a\rangle = \int dx x |\psi_a(x)|^2$$

the interpretation of this quantity is the following: The integral weights the position with the probability density. Thus, this will be the most likely value of x, i. e. the average value of the position where the particle described by a will be found if enough experiments are made.

Thus, it appears that, in the very truest meaning of the world, the expectation value $\langle X\rangle$ is indeed where the particle is expected to be found. One should, however, be careful with this interpretation, and what it means. Just like the expectation value of a six-sided dice is 3.5, and can therefore never be actually thrown, so can the particle perhaps never be at the position of the expectation value. Consider, e. g., the situation that $|\psi_a(x)|^2 = (\delta(x+a)+\delta(x-a))/2$. Then $\langle a|X|a\rangle = 0$, even though the probability density for actually finding the particle there is zero. It is just that the question was not carefully enough phrased. The question is 'what is the average value of the position after many measurements of the same system'. What one tends to associate with expectation values is 'where will the particle most likely be observed'. And the answer to this questions is the maxima of the probability $|\psi_a(x)|^2$, and not the expectation values of the position operator.

Thus, these function play a central (albeit often technical) role in quantum physics, when it comes to give an idea about how things take place in space.

These functions go by the name of wave functions. Why this is so can be most simply seen by considering them for the translation operator and understand their relation to the momentum operator.

Consider[8]

$$\begin{aligned} T(dx)|a\rangle &= \int dx T(dx)|x\rangle\langle x|a\rangle = \int dx |x+dx\rangle \psi_a(x) \\ &= \int dx |x\rangle \psi_a(x-dx) \approx \int dx |x\rangle \left(\psi_a(x) - dx\frac{\partial \psi_a(x)}{\partial x}\right) + \mathcal{O}(dx^2) \end{aligned} \tag{4.11}$$

In the third step a translation $x \to x - dx$ in the integration variable was made, assuming no contributions due to boundary conditions. In the fourth step a Taylor expansion up to

[8] In principle, also a total derivative could be used. However, generically partial derivatives will appear, and therefore already here they will be used.

linear order was performed. Comparing (4.11) to

$$\begin{aligned} T(dx)|a\rangle &= \left(1 - \frac{iPdx}{\hbar} + \mathcal{O}(dx^2)\right)|a\rangle = \int dx|x\rangle \left(\left\langle x \left| 1 - \frac{iPdx}{\hbar} \right| a \right\rangle\right) + \mathcal{O}(dx^2) \\ &= \int dx|x\rangle \left(\psi_a(x) - dx \left\langle x \left| \frac{iP}{\hbar} \right| a \right\rangle\right) + \mathcal{O}(dx^2) \end{aligned}$$

allows to identify

$$\left\langle x \left| \frac{iP}{\hbar} \right| a \right\rangle = \frac{\partial \psi_a(x)}{\partial x},$$

and thus the expectation value of the momentum operator for any state in terms of the wave function. In particular, this implies

$$\langle x|P|y\rangle = -i\hbar \frac{\partial}{\partial y} \delta(x-y).$$

In the same way the matrix elements of higher powers of the momentum operator can be determined.

It is indeed tempting to identify the operator P with $-i\hbar\partial_x$. Especially as

$$[X, -i\hbar\partial_x]f(x) = -i\hbar\left(x\partial_x f(x) - \partial_x(xf(x))\right) = i\hbar f(x).$$

If this is true for any admissible function $f(x)$, the identification is justified, as then $-i\hbar\partial_x$ realizes (4.7). It is said that the identification of P with $-i\hbar\partial_x$ is a particular representation of the operator. This representation is also called the position-space representation of the momentum operator. In essentially all cases to be encountered in this lecture, this identification is justified. However, in more general settings this is not always possible.

As the wording above already suggests, there is more than one possible representation of the momentum operator. However, the wording was also not complete. The identification of the operator X by the multiplication operator and the momentum operator as the derivative form a representation of the so-called (operator) algebra, which is the set of all commutators between all involved operators, in this case just these two.

I. e. the general set of commutator relations without reference to some concrete functions or states defines the algebra. A concrete realization of this set of commutators with reference to a vector space is the representation of the operators. Note that the vector space, in this case the square-integrable complex functions, is included in the definition of a representation. Thus, a representation requires not only the algebra, but also the vector space to be fully defined.

As an alternative, span the vector space by eigenfunctions of the momentum operator. Then, because of the symmetry between momentum and position operator, everything

can be repeated, using now the momentum translation operator (4.9). This leads to wave-functions $\phi(p)$, a momentum operator being the multiplication with p, and a position operator $i\hbar\partial_p$, i. e. a derivative with respect to the momentum. The opposite sign comes again from the order in the commutator. This is a different representation, the so-called momentum-space representation.

Actually, it is possible to switch between both representations. Consider the expectation value

$$\langle x|P|p\rangle = p\langle x|p\rangle = -i\hbar\partial_x\langle x|p\rangle. \tag{4.12}$$

This is a first-order differential equation for ϕ_p. Solving it yields

$$\langle x|p\rangle = Ce^{\frac{ipx}{\hbar}}. \tag{4.13}$$

Thus, the wave-function of a momentum eigenstate is a plane wave. This is the origin of the name wave function. Furthermore, because in a normal wave the argument is $2\pi x/\lambda$ with λ the wave-length, this gives rise to the notion of the de-Broglie wave-length defined as[9] $\lambda = 2\pi\hbar/p$. It describes the distance of one full oscillation of the wave-function for a particle of momentum p. This concept is also generalized beyond the plane-wave case.

It should also be noted that the momentum-space wave-function of a position eigenstate is thus exactly the same, up to complex conjugation, since $\langle x|p\rangle = \langle p|x\rangle^*$.

To determine the normalization constant, calculate

$$\delta(p-p') = \langle p'|p\rangle = \int dx\psi_{p'}(x)^*\psi_p(x) = \int dx|C|^2e^{\frac{ix}{\hbar}(p-p')} = 2\pi\hbar|C|^2\delta(p-p'),$$

and thus $C = 1/\sqrt{2\pi\hbar}$, up to an arbitrary phase, which will be set to one by convention. It should still be noted that the state $|p\rangle$ (or $|x\rangle$) is not normalized to one, even if C is finite. Rather they do not have a finite norm, since $\delta(0)$ is infinite. This is a difference to Hermitian operators in finite-dimensional Hilbert spaces, where their eigenstates have a finite norm. Here, they are only finite up to $\delta(0)$, which, in a certain sense, is a controlled infinity.

In fact, an explicit change between position space and momentum space is a Fourier transformation[10]

$$\psi_a(x) = \langle x|a\rangle = \int dp\langle x|p\rangle\langle p|a\rangle = \int \frac{dp}{\sqrt{2\pi\hbar}}e^{\frac{ipx}{\hbar}}\phi_a(p).$$

---

[9] For this reason, $\hbar$ is often called the reduced Planck's constant, and $h = 2\pi\hbar$ is the proper Planck's constant, or Heisenberg's quantum of action. I will not follow this, historically arisen, naming scheme and call $\hbar$ the Heisenberg constant or the quantum of action.

[10] Mathematically, a Fourier transformation is an invertible integral transformation $\omega(y) = \int \frac{dx}{\sqrt{2\pi}}e^{2\pi ixy}\omega(x)$ with inversion $\omega(x) = \int \frac{dy}{\sqrt{2\pi}}e^{-2\pi ixy}\omega(y)$ of continuous functions. A Fourier transformation is a change of basis in a function space, in this case between a basis of momentum eigenstates to one of position eigenstates and back.

As a Fourier transformation is generically nothing but a change of basis in a function vector space, this is not really surprising, but follows here from the imposed structure using Hermitian observables. Each observable has its own eigenbasis, and there must exist base transformations connecting them. Here is now an explicit example. In fact, for any two operators satisfying (4.7) a transformation between their respective eigenbasis will occur by a Fourier transformation.

It is interesting to consider the uncertainty principle (4.8) explicitly for a position eigenstate. As has been noted in section 3.6, the dispersion $\langle(\Delta X)^2\rangle$ vanishes in this case[11]. Since still (4.8) has to hold, this requires that the dispersion of the momentum needs to diverge. Hence, if a particle is located exactly, its momentum is entirely unknown. Conversely, for a momentum eigenstate the momentum can be known exactly, but its position remains unknown. This is arguably the most famous version of (Heisenberg's) uncertainty principle.

[11]In fact, an explicit direct calculation will not be possible, because the expressions are actually ill-defined. This will be remedied in section 5.1.

Chapter 5

The Schrödinger equation

5.1 Wave packets

It is useful to establish first a particle concept before discussing how it moves dynamically. The plane waves do not really do so, as will be seen at the end of this section.

As has been seen in section 4.3, eigenstates of the position operator are not useful to describe a typical experimental situation, where a particle and its momentum are determined both, though of course not precisely and obeying the uncertainty principle (4.8). However, any state must be decomposable into eigenstates of the position operator. Therefore, the wave-function of any state can be written as

$$|a\rangle = \int dx|x\rangle\langle x|a\rangle = \int dx\psi_a(x)|x\rangle,$$

i. e. weighted by the wave-function of the state. Since the wave-function represents the probability density to find a particle at some point x, the situation described above with a reasonable localized particle should be representable by it.

The probably simplest assumption is that a particle is located by a Gaussian probability, with some width d. Hence,

$$\psi_a(x) = \frac{1}{\sqrt{\sqrt{\pi}d}} e^{-\frac{x^2}{2d^2}}$$

will give a particle located at $x = 0$ with a fall-off range of d. The square gives the corresponding probability density. A shift in the position to x_0 could always be obtained by replacing $x \to x - x_0$. States which have such a localized probability to be found at some place are called a wave-packet or a Gaussian wave-packet.

There are now two interesting extensions of this.

The first is to actually determine how this state satisfies the Heisenberg uncertainty principle. This requires to calculate

$$\begin{aligned}\langle a|X|a\rangle &= \int dx x|\psi_a(x)|^2 = 0\\ \langle a|X^2|a\rangle &= \int dx x^2|\psi_a(x)|^2 = \frac{d^2}{2}\end{aligned}$$

and hence the dispersion is

$$\langle(\Delta X)^2\rangle = \langle X^2\rangle - \langle X\rangle^2 = \frac{d^2}{2}$$

and thus, as expected the particle is localized in an area measured by d. This implies that the dispersion in momentum must be at least of order $\hbar/d$.

Calculating it yields

$$\begin{aligned}\langle a|P|a\rangle &= \int dx\langle a|x\rangle\langle x|P|a\rangle = -i\hbar\int dx\psi_a(x)^*\partial_x\psi_a(x) = i\hbar\int dx\frac{x}{d^2}|\psi_a(x)|^2 = 0\\ \langle a|P^2|a\rangle &= \int dx\langle a|P|x\rangle\langle x|P|a\rangle = \hbar^2\int dx\frac{x^2}{d^4}|\psi_a(x)|^2 = \frac{\hbar^2}{2d^2}\\ \langle(\Delta P)^2\rangle &= \langle P^2\rangle - \langle P\rangle^2 = \frac{\hbar^2}{2d^2},\end{aligned}$$

confirming the estimate. Consequently the Heisenberg uncertainty relation (4.8)

$$\langle(\Delta X)^2\rangle\langle(\Delta P)^2\rangle = \frac{\hbar^2}{4}$$

is fulfilled, and actually saturating the inequality. This implies that a Gaussian wave-packet has the minimum possible uncertainty. It is quite important to realize that though the particle has no net momentum ($\langle P\rangle = 0$), its momentum fluctuates ($\langle P^2\rangle > 0$). Thus, the impression arises that the actual particle is 'just' moving around inside its localization region of $d/\sqrt{2}$, all the time changing randomly its momentum such that it, roughly, stays in this region. This impression is enforced by the fact that the particle has a sizable non-zero probability for the positions in its localization range. Finally, its momentum space-wave function is

$$\langle p|a\rangle = \int dx\langle p|x\rangle\langle x|a\rangle = \int\frac{dx}{\sqrt{2\pi\hbar}}e^{i\frac{px}{\hbar}}\frac{1}{\sqrt{\sqrt{\pi}d}}e^{-\frac{x^2}{2d^2}} = \sqrt{\frac{d}{2\pi^{\frac{3}{2}}\hbar}}e^{-\frac{d^2p^2}{2\hbar^2}}.$$

Thus, the probability that the particle has a certain momentum is also Gaussian distributed with a width given by the dispersion. Thus, as the uncertainty relation implies, it has a position and a speed which fluctuates around its central position and zero momentum. Nonetheless, it is crucial to not make this the interpretation of the reality. The

only reality is the one given by the actual measurement, and then there is a position and a momentum attached to the particle. Before, the wave function is just a useful concept to encode these probabilities, but not an actual description of the particle. The information on the state is encoded in the whole wave function, not only in its measured value. This will be investigated in more depth in chapter 11.

The second is how this state reacts to a translation in momentum space, and thus on how to give the particle momentum. Especially, how this alters the wave-function,

$$\langle x|T(p)|a\rangle = \left\langle x \left| e^{i\frac{pX}{\hbar}} \right| a \right\rangle = e^{i\frac{px}{\hbar}}\langle x|a\rangle = \frac{1}{\sqrt{\sqrt{\pi}d}} e^{i\frac{px}{\hbar}-\frac{x^2}{2d^2}} \tag{5.1}$$

Thus, it just multiplies the wave function by a phase. Repeating the previous calculations now yields

$$\begin{aligned}
\langle a|X|a\rangle &= 0 \\
\langle a|X^2|a\rangle &= \frac{d^2}{2} \\
\langle a|P|a\rangle &= \hbar p \\
\langle a|P^2|a\rangle &= \frac{\hbar}{2d^2} + \hbar^2 p^2 \\
\langle(\Delta P)^2\rangle &= \frac{\hbar^2}{2d^2} \\
\langle(\Delta X)^2\rangle\langle(\Delta P)^2\rangle &= \frac{\hbar^2}{4}.
\end{aligned}$$

There are a number of remarkable results. The first is that the uncertainty relation is still fulfilled as an equality. This implies that there is not a single state satisfying equality, but there are many (infinite) possibilities. So a moving Gaussian wave particle still has minimum uncertainty. The second is that the dispersion of the momentum operator did not change, even though the particle has momentum. Thus, the localization in momentum space does not change when the particle is moving. This is not too surprising. After all, without special relativity any speed is as good as any other when it comes to relative statements. Thus, the particle is now just localized at a different point in momentum space.

This should not come as a surprise. After all, a translation in momentum space is nothing but a Galileo transformation.

The one thing which is surprising is that the expectation value of the position $\langle X\rangle$ remains zero. After all, a moving particle is expected to move. The reason is that the translation operator is instantaneous in time. It does not accelerate the particle. It just creates a new state which has this momentum. This is also visible by the fact that neither

the original state nor the state with momentum has any time-dependence. This is because a dynamical principle is still missing, and this still only describes the particle.

Finally, by taking the limit $d \to \infty$ the wave-packet with momentum (5.1) will turn into the plane wave again. This implies that the particle becomes less and less localized in space, as the dispersion diverges, while it becomes better and better localized in momentum space. At the same time, the uncertainty principle is always fulfilled with equality, as it is independent of d. Thus, the plane-wave describes a particle which has the same probability to be everywhere while having a sharp value of the momentum. This is not too surprising, as the plane-wave is the Fourier transform of the $\delta(x)$ function, which represents a sharply localized particle, but which has some arbitrary momentum.

5.2 Time evolution

The first important thing to note is that time is different from space in classical mechanics. The positions defined the trajectory of a particle, and therefore its path, while the time was just a parameter describing the position along the trajectory[1]. This was also manifest by the Galileo group being the invariance group of classical mechanics. It was only after switching to special relativity that this distinction was lifted, and both space and time were treated equally.

Likewise, it is in quantum mechanics also to be expected that without special relativity time will not play the same role as position, and especially will be only a parameter, rather than a dynamical variable. Especially, it will not be promoted to an operator[2]. Rather, time will still be used to describe as a parameter the progress of the evolution of a system.

Of course, at every instance in time, the system will still be described in the same way as a static system, and in particular completely through states. Thus, if the system is described by a state $|a, t_0\rangle$ at some reference time, it will be in some, usually different, state $|a, t\rangle$ at a later time t. However, the vector space of the states will still be the same. Thus, there should exist an operator $U(t, t_0)$, which transforms the two state into each other,

$$U(t, t_0)|a, t_0\rangle = |a, t\rangle.$$

This operator will be called the time-evolution operator.

[1]Of course, a suitable canonical transformation could exchange the time for the position. However, in the end always one quantity played the role of a parameter defining the position along the trajectory. Here, for simplicity, this will always chosen to be time.

[2]In fact, in relativistic quantum physics position becomes demoted to a parameter, and in this way equality between space and time is established, rather than by promoting time to an operator.

If the state a is decomposed into eigenstates of some operator

$$|a, t_0\rangle = \sum_i a_i(t_0)|i\rangle$$

which does not depend explicitly on the time, the eigenstates $|i\rangle$ remain time-independent. Thus, this implies that the $a_i(t)$ become functions of time,

$$a_i(t) = \langle i|a, t\rangle = \langle i|U(t, t_0)|a, t_0\rangle,$$

and thus are the matrix elements of the time-evolution operator between the initial state and the eigenstates of the chosen observable. These overlaps are a measure of the probability to find the state in a given eigenstate of the observable after some time t. At a later time, the total probability to find the system in any eigenstate should not have changed, even if the probability to find it in a particular eigenstate can have changed. This is a demand on the properties of the time-evolution operator, not a consequence.

Thus, the length of the state needs to be invariant under time evolution. But this implies

$$1 = \langle a, t|a, t\rangle = \langle a, t_0|U^\dagger(t, t_0)U(t, t_0)|a, t_0\rangle \stackrel{!}{=} 1 = \langle a, t_0|a, t_0\rangle.$$

This can be only true if the time-evolution operator is unitary, $U^\dagger = U^{-1}$. This argument is essentially the same as for the translation operator in section 4.2. The unitarity of the time-evolution operator therefore ensures the conservation of probability to find the state in some eigenstate of the observable.

Another demand will be that the result should be independent of whether an evolution is performed directly from some time t_0 to t_2, or first from t_0 to t_1 and then from t_1 to t_2, where t_1 is an arbitrary value between t_0 and t_2,

$$U(t_2, t_0) = U(t_2, t_1)U(t_1, t_0). \tag{5.2}$$

This seems reasonable as it implies that there are no special times t_1. Together with the demand that $U(t, t) = 1$, this is sufficient to make time evolution a group. Especially the inverse operator can be considered to evolve the system backwards in time.

This suggests that there exists some Hermitian operator $H/\hbar = H^\dagger/\hbar$, and thus some observable H, which is, as in section 4.2 for space translation, connected with time evolution. The choice of $\hbar$ is convenient, but conventional. To identify it, it is useful to consider an infinitesimal time-translation,

$$U(t_0 + dt, t_0)|a, t_0\rangle = \left(1 - i\frac{H}{\hbar}dt\right)|a, t_0\rangle + \mathcal{O}(dt^2), \tag{5.3}$$

which follows, like for the translation operator, from the group properties.

It is suggestive that in classical mechanics time evolution as a canonical transformation has also an infinitesimal form where the generator is the Hamilton function. However, in classical mechanics it was not possible to derive the Hamilton function. Rather it was necessary to postulate that a system is described by a certain Hamilton function[3]. Here, the same is necessary. For a given system it is necessary to postulate the generator H, which is in analogy called the Hamilton operator. By comparing the results to experiment it can be tested, whether the postulate is correct or not. But there is no way of deriving the fundamental Hamilton operator[4].

Much of the remainder of this lecture will be about postulating some Hamilton operator, and then deriving what the behavior of the corresponding system is.

However, for most of these investigations the current way how time-evolution is expressed is technically inconvenient.

To derive a more convenient formulation[5], note that because of (5.2) and (5.3)

$$U(t+dt,t_0) = U(t+dt,t)U(t,t0) = \left(1 - i\frac{H}{\hbar}dt\right)U(t,t_0).$$

This implies

$$i\hbar\frac{U(t+dt,t_0)-U(t,t_0)}{dt} = i\hbar\partial_t U(t,t_0) = HU(t,t_0). \tag{5.4}$$

This is a first-order differential equation for $U(t,t_0)$ with initial condition $U(t_0,t_0) = 1$. Acting on both sides with $|a,t_0\rangle$ yields

$$i\hbar\partial_t|a,t\rangle = H|a,t\rangle, \tag{5.5}$$

and thus a dynamical equation for the state! This was what has been looked for. This is the celebrated Schrödinger's equation, the equivalent of Hamilton's equation of classical mechanics. One should, of course, keep in mind that H is an operator, and still needs to be postulated for investigations of a particular physical system.

The prior has important practical implications when solving these equations. This is best seen by solving (5.4). Three different cases have to be distinguished.

[3]Thus, in a loose sense, the classical fundamental laws of nature are just a suitably chosen Hamilton function, describing all observed effects in nature, together with Hamilton's principle. Here, the same reasoning applies, but the postulates on states and observables replace Hamilton's principle.

[4]Of course, effective Hamilton operators can be derived. If there is, e. g., a Hamilton operator describing the water molecules, there exists a derived Hamilton operator describing water as a continuous liquid. But the one for the molecules cannot be derived. The current best guess for the fundamental Hamilton operator are the (not yet entirely established) combination of the one of quantum gravity and of the standard model of particle physics.

[5]Note that the $-$ in front of the $iH/\hbar$ term is a convention, and $+$ would be as good. This would yield a sign flip in various exponentials in the remainder of the lecture.

The first case is that the Hamilton operator is time-independent. Then, just as in section 4.2 for the translation operator, the time-evolution operator is just

$$U(t, t_0) = e^{-\frac{iH}{\hbar}(t-t_0)}. \tag{5.6}$$

However, it does not need to be time-independent. In equation (5.3) it was only assumed that some operator H exists at the given instance t_0 to translate the state an infinitesimal amount dt. This may be a different generator for every time t_0. In addition, it is in principle possible that such a time-dependent Hamilton operator does not commute at different times. This is a genuine quantum effect, and there is no analogy in classical physics.

The simpler case is that the Hamilton operator at different times commutes. Then it is still possible to formally integrate (5.4) over time, yielding

$$U(t, t_0) = e^{-\frac{i}{\hbar}\int_{t_0}^{t} H(t')dt'}$$

which for $H(t') = H$ reduces immediately to (5.6).

However, if the Hamilton operator does not commute at different time, this is not possible, and it only remains to stitch it together using (5.3),

$$\begin{aligned} U(t, t_0) &= \left(1 - i\frac{H(t)}{\hbar}dt\right)\left(1 - i\frac{H(t-dt)}{\hbar}dt\right)...\left(1 - i\frac{H(t_0+dt)}{\hbar}dt\right)1 \\ &= 1 + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 ... \int_{t_0}^{t_{n-1}} dt_n H(t_1)...H(t_n), \end{aligned} \tag{5.7}$$

where the second equality is stated, for now, without proof. The second form is called a Dyson series. It will reappear again in section 9.4, where also more will be said on the validity of this expression.

5.3 An explicit example

To see how Schrödinger's equation (5.5) works, it is useful to consider the simplest possible, explicit example, the free particle. This requires to postulate the Hamilton operator H.

Consider first the one-dimensional case. Then the Hamilton function is $H = p^2/2m$. Postulate that p is replaced by the momentum operator P, and then the Hamilton operator reads

$$H = \frac{P^2}{2m}.$$

Inserting this into Schrödinger's equation (5.5) yields

$$i\hbar\partial_t|a, t\rangle = \frac{P^2}{2m}|a, t\rangle.$$

Contracting this with the (time-independent) $\langle x|$ and using (4.12) yields

$$i\hbar\partial_t\langle x|a,t\rangle = -\frac{\hbar^2}{2m}\partial_x^2\langle x|a,t\rangle$$

and thus a (partial) differential equation for the, now time-dependent, wave-function $\Psi_a(t,x)$.

To solve this equation a separation ansatz $\Psi_a(t,x) = \phi(t)\psi(x)$ is useful, since it allows to separate the variable dependence as

$$\frac{i}{\phi(t)}\partial_t\phi(t) = -\frac{\hbar}{2m\psi(x)}\partial_x^2\psi(x).$$

Thus both sides need to equal some constant ω. Solving the left-hand-side yields

$$\phi(t) = a_t e^{-i\omega t}$$

while the right-hand side yields

$$\begin{aligned}\psi(x) &= b_1 e^{ikx} + b_2 e^{-ikx}\\ k &= \sqrt{\frac{2m\omega}{\hbar}}\end{aligned}$$

and thus as total solution

$$\Psi_a(x,t) = a_1 e^{i(\omega t + kx)} + a_2 e^{i(\omega t - kx)} \tag{5.8}$$

with $a_i = a_t b_i$ as two initial conditions characterizing the state, as does ω. Thus, the solutions to the free Schrödinger equation are plane waves. While the space-time dependence is now given by this, it remains to understand the significance of the pre-factors a_i and the quantity ω. One of the pre-factors is fixed by normalization. The other then gives the relative amount of left-propagating and right-propagating waves. The quantity ω is less obvious, but it has actually something to do with the energy of the particle.

To see this, consider the following: The Hamilton function characterizes the energy of a classical system. The Hamilton operator is Hermitian, and therefore observable. It is hence an interesting question what its expectation values are. An explicit calculation for the solution of the Schrödinger equation yields

$$\langle a,t|H|a,t\rangle = \left\langle a,t\left|\frac{P^2}{2m}\right|a,t\right\rangle = -\frac{\hbar^2}{2m}\int dx\Psi_a(t,x)^*\partial_x^2\Psi_a(t,x) = \frac{\hbar^2k^2}{2m}\int dx|\Psi_a(t,x)|^2.$$

To interpret the result it is important to make a short note on the norm of a plane wave. This has been discussed already in section 4.3, and it can be assumed that the integral is

one. Thus, this normalization condition will absorb one of the free constants. Then, the expectation value is $\hbar^2 k^2/(2m) = \hbar\omega$. Therefore ω characterizes the expectation value of the Hamilton operator for the state $\Psi_a(t,x)$.

What does this mean? Well, from a classical plane wave it is known that its kinetic energy is given by $p^2/2m$, where p is its momentum. Since the Hamilton operator is just the momentum operator squared, this suggests to identify $\hbar k$ with the momentum of the free particle and thus $\hbar\omega$ with its energy, just as before the momentum operator has been tentatively identified. This is indeed correct, as will be seen in section 5.7. But to show this requires first a few more formal developments.

Before doing so, there is another interesting observation. The solutions (5.8) are plane waves. Thus, a state described by (5.8) is not localized in any way. This is very far from any notion of a particle imaginable. To recover some resemblance of the notion of a particle, it is useful to combine the results here with those of the wave-packets of section 5.1.

Since (5.8) are the eigenfunctions of an Hermitian operator, any other state can be written as a linear superposition of them. Consider therefore

$$\Psi_l(x,t) = \int dk e^{-\frac{d^2k^2}{2\hbar^2}} e^{i\left(\frac{\hbar k^2}{2m}t + kx\right)} = \frac{\sqrt{2\pi}}{\sqrt{\frac{d^2}{2\hbar^2} - i\frac{\hbar}{m}t}} e^{-\frac{\hbar^2 m x^2}{4(d^2 m - i\hbar^3 t)}}$$

which can be normalized as

$$\Psi_l(x,t) = \sqrt{\frac{d^2}{\hbar\pi^2}} \frac{\sqrt{2\pi}}{\sqrt{\frac{d^2}{2\hbar^2} - i\frac{\hbar}{m}t}} e^{-\frac{\hbar^2 m x^2}{4(d^2 m - i\hbar^3 t)}} \tag{5.9}$$

It describes a probability distribution which is Gaussian around $x = 0$ for all times t. However, the width depends on the time, and increases proportional to time. Thus, physically, it describes a particle at rest at $x = 0$, and it stays also at $x = 0$ on the average. But, as a quantum effect, the probability to find it somewhere else increases with time t. There is a quantum movement involved, and the width of the probability is given by

$$\sigma^2 = \frac{d^4}{\hbar^4} + \frac{\hbar^2}{m^2} t^2. \tag{5.10}$$

It has therefore a minimal width at $t = 0$ which is determined by the width of the Gaussian distribution in momentum space. The increase in width over time is then inversely proportional to the mass, or to be precise the Compton wave-length $c\hbar/m$, where the factor c, the speed of light, is introduced to convert the quantity into a length. This is intuitively clear: The lighter the particle, the more it can move around, and thus spread out its wave function in time.

It is noteworthy that the expectation value of the Hamilton operator is given by

$$\begin{aligned}\langle k|H|k'\rangle &= \frac{d^2}{2m\hbar\pi^2}\int dk'dkdxkk'e^{-\frac{d^2k^2}{2\hbar^2}}e^{-\frac{d^2k'^2}{2\hbar^2}}e^{i\left(\frac{\hbar(k^2-k'^2)}{2m}t+(k-k')x\right)}\\ &= \frac{d^2}{2m\hbar\pi^2}\int dk'dkkk'e^{-\frac{d^2k^2}{2\hbar^2}}e^{-\frac{d^2k'^2}{2\hbar^2}}e^{i\left(\frac{\hbar(k^2-k'^2)}{2m}t\right)}\delta(k-k')\\ &= \frac{d^2}{2m\hbar\pi^2}\int dkk^2e^{-\frac{d^2k^2}{\hbar^2}} = \frac{\hbar^2\sqrt{\pi}}{4dm},\end{aligned}$$

and is therefore just twice the width of the smearing in momentum space. As expected, the energy is time-independent, but non-zero.

5.4 Properties of the Hamilton operator

There are a few very general features pertaining to the Schrödinger equation and the Hamilton operator. The first is that the Hamilton operator, as a Hermitian operator, has real eigenvalues and its eigenvectors form a full basis. As will be seen in section 5.7 the eigenvalues can be identified with energy. Hence, a basis build from the eigenvectors of the Hamilton operator is called the energy basis, and the eigenvalues are denoted by E_n,

$$H|E_n\rangle = E_n|E_n\rangle.$$

But for now these are just names.

If the Hamilton operator is not explicitly time-dependent, $H \neq H(t)$, then the separation ansatz made before works generally. For this note

$$i\hbar\partial_t\sum_n a_n(t)|E_n\rangle = i\hbar\partial_t|a,t\rangle = H|a,t\rangle = H\sum a_n(t)|E_n\rangle = \sum a_n(t)E_n|E_n\rangle$$

Using that the $|E_n\rangle$ are orthogonal and time-independent this yields

$$i\hbar\partial_t a_n(t) = E_n a_n(t),$$

which can be solved by

$$a_n(t) = a_n(0)e^{-i\frac{E_n t}{\hbar}}$$

and thus the energy dependence is entirely contained in these exponential factors. Especially, the solution of Schrödinger's equation is now given by

$$|a,t\rangle = \sum_n a_n(0)e^{-i\frac{E_n t}{\hbar}}|E_n\rangle, \tag{5.11}$$

where the $a_n(0)$ are determined by the initial conditions. Thus, solving Schrödinger's equation has been reduced to the time-independent problem of solving the eigenproblem of the Hamilton operator[6]. The so reduced Schrödinger equation is time-independent and

[6] Although in practice this may not be the most effective way of solving a given problem.

boils down to the eigenvalue problem of the Hamilton operator alone

$$H|E_n\rangle = E_n|E_n\rangle.$$

Therefore, this is called the stationary Schrödinger equation, and the eigenvectors $|E_n\rangle$ the stationary solutions.

In fact, this also helps in determining the time evolution of any state for this case. Because

$$|a,t\rangle = U|a,0\rangle = \sum_n a_n(0)e^{-\frac{iHt}{\hbar}}|E_n\rangle = \sum_n a_n(0)e^{-\frac{iE_nt}{\hbar}}|E_n\rangle$$

the time evolution of any state is given by a superposition of energy eigenstates.

Another interesting aspect arises if there is an operator B which commutes with H, $[B,H] = 0$. As has been seen in section 3.7, there is a maximum set of simultaneously diagonalizable operators. They are distinguished by the fact that for fixed eigenvalue of one operator the other ones have differing eigenvalues. Especially, in the present case this implies that for some or all eigenvalues E_n there are multiple eigenvalues b_n. Assuming that H does not depend explicitly on the time it follows that

$$[B,H]|a,t\rangle = [B,H]U|a,0\rangle = U[B,H]|a,0\rangle = 0$$

where it was used that if $[B,H] = 0$ so follows $[B,f(H)] = 0$, as long as the function $f(H)$ can be written as a Taylor series. This is true, if H does not depend explicitly on the time. Thus, in this case B and H commute for all times.

In particular, this implies

$$U|E_n,b_n,t\rangle = e^{\frac{-iE_nt}{\hbar}}|E_n,b_n,0\rangle,$$

and the time dependence becomes trivial. In particular, the state remains in an eigenstate of both H and B, and therefore measurements of either H or B remain time-independent. Thus, this implies that if H does not depend explicitly on the time, itself and any operator B commuting with H have a special meaning with respect to measurements. This is the quantum version of a constant of motion of the classical physics and the eigenvalue b_n is, alongside the energy E_n, a conserved quantity. However, if the state in question is not a simultaneous eigenstate to B and H, this is no longer the case, and its value will change. Note that this, of course, also applies if there is no second commuting operator. In this case, only H will be a constant of motion in this sense.

In the example of the free particle in section 5.3, the momentum played the same role, albeit this arose from a comparatively simple relation between H and P - it is in general not necessary that B appears in H, only that they commute.

Of course, if there exists one or more operator commuting with H, the dynamical problem is not entirely solved by finding the eigenbasis of H. Rather, it is necessary to also find a full basis of all commuting operators to resolve the degeneracies in the subspaces to fixed eigenvalues E_n. There will be many examples of this in chapter 7.

This has even consequences for operators C which do not commute with H. Consider the expectation value between an energy eigenstate,

$$\langle E_n, t|C|E_n, t\rangle = \langle E_n, 0|U^\dagger C U|E_n, 0\rangle = \langle E_n, 0|e^{\frac{iE_n t}{\hbar}} C e^{-\frac{iE_n t}{\hbar}}|E_n, 0\rangle = \langle E_n, 0|C|E_n, 0\rangle.$$

Thus, even for non-commuting observables, or for that matter any operator, their expectation value in an energy eigenstate is time-independent, except if they themselves would be explicitly time-dependent. Therefore, the situation does not change in time, and that is another reason to call an energy eigenstate a stationary state.

If the state is not an energy eigenstate, the expansion (5.11) shows that

$$\langle a, t|C|a, t\rangle = \sum_{nm} \langle E_m, 0|a_m e^{i\frac{E_m t}{\hbar}} C a_n e^{-i\frac{E_n t}{\hbar}}|E_n, 0\rangle = \sum_{nm} a_n a_m e^{-i\frac{(E_n - E_m)t}{\hbar}} \langle E_m, 0|C|E_n, 0\rangle.$$

Thus, the time-dependence of arbitrary expectation values is given by a superposition of oscillations with frequencies $(E_n - E_m)/\hbar$. Especially, if the eigenspectrum of the Hamilton operator should be discreet, only characteristic frequencies will appear. Though such a situation has not yet been encountered, as for the free particle the spectrum was continuous, this will arise repeatedly in chapter 6.

5.5 Energy-time uncertainty

This leads to an interesting insight when combined with (5.10). The energy is the ratio of the momentum width and the Compton wave-length, up to an irrelevant factor $4c$. Normalizing (5.10) to the Compton-wave length therefore yields

$$(2\sigma)'^2 = E^2 + \left(\frac{\hbar t}{c}\right)^2.$$

Thus, any deviation from the distribution at $t = 0$ becomes only relevant after a time Δt, such that

$$E\Delta t \sim \hbar$$

or, the energy is proportional to the width in momentum space, and thus is actually also a ΔE,

$$\Delta E \Delta t \sim \hbar. \tag{5.12}$$

This looks very much like the uncertainty relation (3.6), but it is not of the same kind. It is a derived quantity, and is not related to any operators, especially as there is no time operator. It is still a very important concept: In quantum physics, also the actual energy and time is not well-defined. Although the particle is at rest, its kinetic energy is non-zero. The size of this uncertain kinetic energy is inverse proportional to the time it needs to spread over a spatial extent of size its Compton wave-length. Reversely, the energy it has cannot be determined more precisely than roughly its Compton wave-length. Hence, in quantum physics, there is always an uncertainty in the kinetic energy of a localized particle. This allows for one of the most spectacular consequences of quantum physics: It is possible to have actually more energy than is present, but this is only possible for a limited amount of time. By localizing the particle in space, it has an uncertainty in momentum, and thus kinetic energy. If this localization is kept for a fixed time, its (kinetic) energy is consequently uncertain by this amount, and can be larger. Thus, there is a probability that the particle does during this time something, which would be forbidden classically by energy conservation, because it actually can have more energy. It is not an independent principle, but follows from (3.6). The most drastic application of this in quantum mechanics will be the tunneling processes to be discussed starting from section 6.4. Beyond quantum mechanics in particle physics this leads to the possibility to investigate particles with a rest mass (much) larger than the available energy. Because again the amount of available energy is not perfect, and can be much larger for a very short amount of time.

In fact, this statement can be generalized. Consider some state a. The probability to find the system at a time t later still in the same state is

$$|\langle a, t|a, 0\rangle|^2 = |C(t)|^2$$

where $C(t)$ is called the correlation amplitude.

If the state a is an eigenstate of the Hamilton operator, this correlation amplitude becomes

$$C(t) = e^{-\frac{iE_a t}{\hbar}},$$

and the state remains an eigenstate. If it is not an eigenstate, it can be expanded in eigenstates $|i\rangle$, yielding

$$C(t) = \sum_i |a_i|^2 e^{-\frac{iE_i t}{\hbar}}$$

with $a_i = \langle i|a\rangle$. These a_i encode the overlap of the state $|a\rangle$ with the eigenstate $|i\rangle$. If the spectrum is continuous, this become an integral

$$C(t) = \int dE \rho(E) |g(E)|^2 e^{-\frac{iEt}{\hbar}},$$

where $\rho(E) \geq 0$ is the density of states in an energy interval dE, and the overlap with the states is encoded in the function $g(E)$, i. e. it plays the role of the overlap coefficients a_i. Note that the state is still assumed to be normalized, implying

$$\int dE \rho(E)|g(E)|^2 = 1$$

and thus any deviation of $C(t)$ from unity is due to the exponential. Especially, $C(0) = 1$. In fact, this implies that the correlation function is the Fourier transform of $\rho(E)|g(E)|^2$.

Assume now that $\rho(E)|g(E)|^2$ is strongly peaked around some value E_0, with width ΔE. On the other hand, for any value very different from E_0, the weight is small, and then the oscillation will reduce it further. Thus, the only appreciable contribution will come from the peak area. This area will be probed if $t \gtrsim \hbar/\Delta E$, or after a time interval $\Delta t = t - 0 \gtrsim \hbar/\Delta E$, as only then the oscillation starts to probe regions where the change in $\rho(E)|g(E)|^2$ starts to become relevant. Thus, to see any effect requires

$$\Delta E \Delta t \gtrsim \hbar$$

and thus the same as (5.12). Hence, this relation tells starting from which time duration a correlation amplitude will change from its initial value. At shorter times, the system behaves as if having no discernible energy structure at all, and all energies contribute in the same way, being the same statement as above in a different disguise: At short enough times, all energies are accessible.

5.6 Pictures of time-evolution

After having now the Hamilton operator rather well under control, there is another feature of it which is worth mentioning. For this, remember that the time evolution operator derived from the Hamilton operator is, foremost, a unitary operator.

Any unitary operator U has the property $U^{-1} = U^\dagger$. This implies if states $|a\rangle$ and $\langle b|$ are both acted upon by this operator, their overlap remains unchanged,

$$\langle b|U^\dagger U|a\rangle = \langle b|a\rangle$$

In particular, as seen already before, any overlap remains constant in time if both states evolve in time.

Consider now the expectation value of some operator A. Then

$$(\langle b|U^\dagger)A(U|a\rangle) = \langle b|(U^\dagger A U)|a\rangle. \tag{5.13}$$

is mathematically a tautological identity. But from an interpretation point of view this is rather deep: If a system is acted upon with a unitary transformation, this can be either considered as that the states undergo this transformation or that the operators undergo a similarity transformation using the same unitary transformation.

Consider, for starters, a position eigenket and the translation operator of section 4.2. There, it was usually only considered

$$T(dx)|x\rangle = |x + dx\rangle.$$

and the state was translated, such that

$$X|x + dx\rangle = (x + dx)|x + dx\rangle.$$

Consider now the possibility to act upon the position operator itself,

$$T(dx)^\dagger XT(dx) = \left(1 + \frac{iPdx}{\hbar}\right) X \left(1 - \frac{iPdx}{\hbar}\right) + \mathcal{O}(dx^2) = X + \frac{idx}{\hbar}[P, X] = X + dx + \mathcal{O}(dx^2).$$

which therefore yields the position operator plus a term proportional to the identity operator. Therefore

$$\begin{aligned}
(\langle x|T(dx)^\dagger)X(T(dx)|x\rangle) &= \langle x + dx|X|x + dx\rangle + \mathcal{O}(dx^2) &= x + dx + \mathcal{O}(dx^2) \\
\langle x|(T(dx)^\dagger XT(dx))|x\rangle &= \langle x|X|x\rangle + dx\langle x|1|x\rangle + \mathcal{O}(dx^2) &= x + dx + \mathcal{O}(dx^2)
\end{aligned}$$

and thus the same result, as was to be expected for a mathematical identity. Still, this implements what in classical physics is called active and passive transformations: Either the system (the observable) or the coordinate system (the base kets) are transformed.

As noted, the same holds for any unitary operator, and thus especially also for the time evolution operator. So far, always the states had been transformed. This case is also known as the Schrödinger picture. It has the point of view that the system evolves in time. Considering now the operators to evolve in time is called the Heisenberg picture. The states remain constant in time, but any operator becomes time-dependent

$$A(t) = U(t)^\dagger A(0)U(t), \tag{5.14}$$

where $A(0)$ is the operator as it was in the Schrödinger picture, i. e., all the operators encountered so far. Of course, because of (5.13) any expectation values, and thus any measurements, do not change. Physically, both views are indistinguishable. Thus, they exist only for our convenience.

From (5.14) an interesting consequence arises, provided that both A and H do not have an explicit time-dependence. Then, by acting with a total derivative on this equation arises

$$\begin{aligned}\frac{dA}{dt} &= (\partial_t U^\dagger)A(0)U + U^\dagger A(0)\partial_t U = \frac{1}{i\hbar}\left(U^\dagger A(0)HU - U^\dagger HA(0)U\right) \\ &= \frac{1}{i\hbar}\left(U^\dagger A(0)UU^\dagger HU - U^\dagger HUU^\dagger A(0)U\right) = \frac{1}{i\hbar}[A(t), U^\dagger HU] = \frac{1}{i\hbar}[A(t), H], \end{aligned} \tag{5.15}$$

where in the last step it was used that for H not explicitly time-dependent H and U commute. This equation is the Heisenberg equation of motion.

5.7 Long distances and the emergence of classical physics

Again this looks familiar. In fact, just as with (4.7), this looks deceptively familiar. If not for factors of i and $\hbar$, just replacing the commutator by the Poisson brackets creates the classical canonical commutation relation and Hamilton's equation of motion of mechanics. This observation generalizes in many contexts, and is known as the correspondence principle: Given the classical equations using the Poisson brackets, the quantum version is obtained by replacing the Poisson brackets by commutators and adding an i and $\hbar$. While this is not true in general, it is often so. However, in some cases more elaborate constructions are necessary, and in some cases there is no classical analogue of a quantum system or vice versa. But as a first guess how to find a quantum theory which has a certain classical limit, this is very successful.

However, when following this recipe there is one ambiguity, the operator ordering ambiguity. Classically, x and p, or some other canonically conjugated quantities, commute while the corresponding operators X and P do not. So, in writing the Hamilton operator as a generalization of the Hamilton function requires to make a choice whenever both appear in a product. The first guideline is to make the Hamilton operator Hermitian. XP is not. Quite often this works by replacing xp with $(XP + PX)/2$, which is Hermitian. While again a good guideline, this is not always working.

Beyond that, and especially if there is no classical analogue, it is postulating, or, more profanely put, guessing, the correct Hamilton operator. Experiments are then needed to verify if the guess was correct. In this way, the standard model of particle physics was found, which is the most fundamental Hamilton operator we currently know, and from which all terrestrially observed, non-gravitational physics can be derived, at least in principle, if not in practice for reasons of complexity.

It is now interesting to see how this actually plays out in practice. However, to make it feasible, this will be done with a very humble example, the free particle in one dimension.

Classically, the Hamilton function is $h = p^2/2m$. So, postulate that the quantum version is $H = P^2/2m$. It will be useful to work in the Heisenberg picture. Then

$$\frac{dP}{dt} = \frac{1}{i\hbar}[P, H] = 0, \tag{5.16}$$

and thus the momentum operator is time-independent, very much like its classical analogue. The other relevant object is the position operator X. Its equation of motion reads

$$\frac{dX}{dt} = \frac{1}{i\hbar}[X, H]. \tag{5.17}$$

While this can be done by brute force, a more elegant way is to note that for any polynomial function $F(P)$

$$[X, F(P)] = i\hbar\partial_P F(P)$$

holds. This can be proven by repeatedly using (4.7). As a side-remark,

$$[P, G(X)] = -i\hbar\partial_X G(X) \tag{5.18}$$

holds in the same vain for any polynomial G of X. Thus (5.17) takes the form

$$\frac{dX}{dt} = \frac{P}{m}, \tag{5.19}$$

where P is time-independent, because of (5.16). This equation can be solved immediately and yields

$$X(t) = X(0) + \frac{P}{m}t. \tag{5.20}$$

This is an operator-valued function of time. Especially, the operator $X(t)$ becomes 'P'-like with time. As a consequence, the equal-time commutator

$$[X(0), X(0)] = 0$$

vanishes[7], but not the commutator at different times

$$[X(t), X(0)] = \left[X(0) + \frac{P}{m}t, X(0)\right] = \frac{t}{m}[P, X(0)] = -\frac{i\hbar t}{m}.$$

This implies that the position at different times are not compatible observables. It is thus not possible to locate the particle at two different times exactly. In fact, using (3.8) implies

$$\langle(\Delta X(t))^2\rangle\langle(\Delta X(0))^2\rangle \geq \frac{\hbar^2 t^2}{4m^2} \tag{5.21}$$

[7]As the initial condition can be set at any arbitrary time, this would also be true for any arbitrary time as long as X is evaluated at the same time.

and thus the uncertainty increases quadratically with time. Hence, even though (5.20) is suggestive of an ordinary movement, (5.21) very strongly shows that the location of the particle with respects to its point of origin becomes increasingly less known with time.

However, note that

$$\langle X\rangle(t) = \langle X\rangle(0) + \frac{\langle P\rangle}{m}t$$

the expectation value for the position of the particle actually behaves classically: Since the expectation values of $X(0)$ and P are time-independent, the expectation values are that of a classical particle. It is very important to note here the fact that an expectation value has been taken. Hence, when performing many measurements on identical systems, the average position of the particles will behave like a classical particle. The actual position of a particle in a single experiment will be somewhere, located around the classical trajectory but with an uncertainty increasing with time. Especially, the later the measurement is performed the more likely the particle can be found farther away from the classical trajectory[8]. However, since in (5.21) the uncertainty is multiplied by $\hbar^2$, this effect is quantitatively actually very small. For a macroscopic object of 1 kilogram, it takes about 10^{27} years of undisturbed propagation, before the uncertainty becomes of order one meter.

In fact, this effect is more generally true. Following the quantization procedure, assume that the particle is under the influence of a potential $V(X)$ yielding the Hamilton operator

$$H = \frac{P^2}{2m} + V(X). \tag{5.22}$$

Now, (5.19) still holds, as $V(X)$ commutes with X at equal times. But in this case P is no longer time-independent. Furthermore, the Heisenberg equation of motion (5.15) implies

$$\frac{d^2X}{dt^2} = \frac{1}{i\hbar}\left[\frac{dX}{dt}, H\right] = \frac{1}{i\hbar m}[P, H] = \frac{1}{m}\frac{dP}{dt}.$$

To make progress, use again (5.15) and (5.18)

$$\frac{dP}{dt} = \frac{1}{i\hbar}[P, H] = \frac{1}{i\hbar}[P, V(X)] = -\partial_X V(X)$$

and thus

$$m\frac{d^2X}{dt^2} = -\partial_X V(X).$$

[8]Note that, formally, there is a small, but finite, probability to find the particle arbitrarily far away from its point of origin even after an infinitesimal amount of time. In reality, the effects of special relativity forbids this. However, to incorporate the effect requires to shift a much more powerful formulation of quantum physics, quantum field theory, which is beyond the scope of this lecture.

This is nothing but an operator-valued form of Newton's equation of motion. Of course, this again implies nothing about the actual position of a particle. However,

$$m\frac{d^2}{dt^2}\langle X(t)\rangle = \frac{d}{dt}\langle P(t)\rangle = -\langle \partial_X V(X)\rangle. \tag{5.23}$$

Thus, the expectation values obey Newton's law. This is known as the Ehrenfest relation. If, as above, the deviation of an individual particle position from the average value is small, because $\hbar$ is small and all other parameters are large, it appears that the particle moves as if classical. In this way, classical physics is recovered from quantum physics. This also justifies the identification of P with the momentum operator and of H with the Hamilton operator.

Hence, classical physics can be considered as the limit where $\hbar$ is small, and thus the dispersion is small compared to all other scales. Then the deviation of an individual particle from the average position becomes irrelevant, and all particles appear to behave like the expectation values, and therefore classically. In this sense, classical physics emerges as the long-distance behavior of quantum physics itself. Conversely, as soon as some scale becomes of the same order as $\hbar$, this is no longer true, and the dispersion can no longer be neglected. This also does not explain how to understand things like the measurement process. It only explains why, under certain conditions, particles behave as if classical. This complex will be returned to in chapter 11.

There is another important insight. As long as the potential in (5.22) is not depending on the wave-function itself, the Schrödinger equation is linear in the wave-function, and the Hamilton operator is a linear operator. Thus, any sum of solutions of Schrödinger's equation is also a solution of Schrödinger's equation. This is a realization of the superposition principle. This is a feature which will be kept throughout the lecture. However, it should be recognized that this is not necessarily so, and in situations beyond the quantum-mechanical settings discussed here, this may change. But this requires care to avoid spoiling the consistency, and thus is beyond the scope of this lecture.

5.8 Recognizing a problem: Black-body radiation

While the Stern-Gerlach experiment was used here as the primary example to build up the postulates of quantum physics, it was by far not the only case in which surprising effects were detected. Many others will be discussed throughout this lecture, especially in sections 7.2.3, 7.4, and 7.5. But all of them will first need to discuss problems with more than one spatial dimension, which will be postponed until chapter 7, to avoid at this stage unnecessary technical complications.

There is, however, another example, which should be briefly mentioned, where it became obvious that something was missing, without directly pointing to a quantization. Rather, it signaled the breakdown of a classical description in a way which has by now been established as an important signature for the limitation of a physical theory. And thus an important hint to yet undiscovered physics. This is the appearance of divergences.

Consider a box at a certain temperature. According to classical statistical physics a fixed temperature provides every degree of freedom with an (average) kinetic energy of $k_B T$, where k_B is Boltzmann's constant. The problem arises if electromagnetic radiation is stored in the box, so-called hohlraumstrahlung. Because the energy of an electromagnetic wave is proportional to its frequency and because $\omega = c|\vec{k}|$, where $\vec{k}$ is the mode number, the total energy diverges. This follows, as every mode of the electromagnetic field is a degree of freedom, and the total energy is the sum over all modes, and thus the total energy is

$$E \sim \sum_{\vec{k}} E(k) = \int d\omega \omega^2.$$

Since this was experimentally not the case, this pointed to a serious flaw of the theory describing the combination of statistical physics and electromagnetic radiation.

It is, in fact, quantum physics, which solves the problem. Requiring that the energy is quantized in units of $E = \hbar\omega$, and thus there is a minimal energy, the divergence is absent. Because then there is not an infinite number of degrees of freedom and there is a lowest possible energy, the normal rules of statistical physics yield Planck's black-body radiation formula

$$\frac{dE}{d\omega} = V \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\frac{\hbar\omega}{k_B T}} - 1}. \tag{5.24}$$

Integrating this yields a finite amount of energy stored in the box. It were here two requirements that needed to be met: A finite lowest energy and a countable number of degrees of freedom. The later introduces a cutoff to the spectrum by the box size, as only standing waves can be accommodated.

It is a certain irony that a full description of the problem involves actually more subtleties. Because the full classical theory of electromagnetic waves is Maxwell theory which introduces necessarily special relativity, things become more complicated. Accommodating at the same time special relativity and quantum physics is not possible within quantum mechanics. This is the providence of quantum field theory. Thus, a fully consistent derivation of (5.24) beyond this heuristically argumentation cannot be provided here.

Chapter 6

One-dimensional systems

Before venturing on to discuss higher-dimensional problems, it is useful to first study further quantum phenomena in the one-dimensional case. Thus, an intuition can be slowly build for what kind of new effects quantum physics introduces, before adding the complications due to more dimensions. So far, the treatment covered the free particle in chapter 4.

As discussed in section 5.7 some quantum systems can be obtained by taking the classical Hamilton function and replacing position and momentum by operators. The so obtained Hamilton operator is then postulated to describe the quantum system, with the classical system emerging in the sense of the Ehrenfest relation (5.23). This approach will be used here to discuss various problems known from classical mechanics to see how they change when quantizing them.

6.1 Infinite-well potential

The, arguably, simplest example beyond the free particle is the infinite-well potential

$$V(x) = \begin{cases} 0 \text{ if } 0 \leq x \leq L \\ \infty \text{ otherwise} \end{cases}$$

This potential classically restricts the movement of the particle to inside the range $0 \leq x \leq L$. This requires to decide what happens in the quantum world. Since this potential should describe an impenetrable barrier, it seems to be sensible to require that also in the quantum world nothing is allowed outside. Thus, this implies for the wave-function $|\psi(x)|^2 = 0$ for $x < 0$ or $x > L$: The probability to find the particle outside the box is zero. This can only be satisfied if the wave-function $\psi(x)$ itself vanishes outside the box. These two conditions translate therefore into boundary conditions of the Schrödinger equation (5.5).

They therefore provide the necessary conditions for solving the corresponding differential equation in, e. g., position space.

The Schrödinger equation therefore reads

$$i\hbar\partial_t|E,t\rangle = H|E,t\rangle = \left(\frac{P^2}{2m} + V(X)\right)|E,t\rangle \tag{6.1}$$

where the only quantum number was taken to be the energy. After all, there are no additional operators beyond P and X, and neither commute with the Hamilton operator.

This problem can be most easily solved in position space. Contracting therefore (6.1) with an eigenbasis of the position operator yields the Schrödinger equation in position space

$$i\hbar\partial_t\psi(x,t) = \left(-\frac{\hbar^2}{2m}\partial_x^2 + V(x)\right)\psi(x,t).$$

Because time and space are cleanly separated it is possible to make a separation ansatz

$$\psi(x,t) = e^{-i\frac{Et}{\hbar}}\phi(x) \tag{6.2}$$

which leads to

$$E\phi(x) = \left(-\frac{\hbar^2}{2m}d_x^2 + V(x)\right)\phi(x). \tag{6.3}$$

This is an ordinary second-order differential equation in x. The two necessary boundary conditions are already available. In fact, the solution outside the interval $[0, L]$ is already known, $\phi(x) = 0$. Inside the interval, the equation reads

$$E\phi(x) = -\frac{\hbar^2}{2m}d_x^2\phi(x). \tag{6.4}$$

Deceptively, this reads like the equation for a free particle. But it is not, because of the boundary conditions. Still, as it has the form of a wave-equation, a suitable ansatz, which automatically satisfies the boundary conditions, is

$$\phi(x) = A\sin\left(\frac{n\pi}{L}x\right) \tag{6.5}$$

where $n > 0$ is an integer number. Inserting (6.5) in (6.4) yields

$$E = \frac{n^2\pi^2\hbar^2}{2mL^2} \tag{6.6}$$

as a consistency condition. Thus, the energy must have this values for (6.5) to be a solution. These are the only solutions. The other solution, $\cos(\omega x + \varphi)$, from (6.4) is only satisfying the boundary conditions if the phase shift φ is chosen such that it turns it into (6.5). The result (6.6) has a number of remarkable consequences.

First, in the classical case the energy of the particle in the well could be any value starting from zero. This is no longer the case. Because n in (6.6) is an integer, only certain values are allowed anymore. The allowed energies are quantized. From these kind of effects the name quantum mechanics arises: Where classical physics is continuous, quantum physics may be discrete. This is somewhat at odds with the situation for the free particle, in section 5.3. There, the possible energy values were continuous. The difference between the situations is that the free particle in section 5.3 was not confined to a finite region in space, while here this is the case. This will be the decisive difference: The energies of a particle confined to a finite region of space[1] are quantized, but not necessarily those of a particle which is not. Such states are called bound states, in contrast to the free states. This is the second observation.

The third is that (6.6) already shows why this quantization is not usually seen on macroscopical scales. The relative difference between two neighboring energy levels

$$\frac{\Delta E}{\sum E} = \frac{E(n+1) - E(n)}{E(n+1) + E(n)} = \frac{1+2n}{n^2} \tag{6.7}$$

vanishes like $1/n$ for larger n. The factor of $\hbar^2$ in (6.6) requires n to be very large for even moderately large values of m and L for the energy to be of macroscopical size. Thus, at macroscopical levels of n there is essentially no detectable trace of the quantization left.

The fourth is that n is required to be larger than zero. Thus, the particle cannot be at rest, and there is always a non-zero energy in the system, the so-called vacuum energy. Thus, as could be expected from the uncertainty relation (4.7), a quantum-mechanical particle is never at rest. However, though (4.7) suggests it, this is actually not necessarily true. In fact, there is an arbitrariness involved in defining the total energy. Just as in classical mechanics, the movement of a particle is not altered when replacing H with $H - E_0$. This is also true in quantum mechanics, essentially because any constant drops out of (5.15), and thus for the same reason as in quantum mechanics. Thus, by replacing H by $H - \pi^2\hbar^2/(2mL^2)$ the vacuum energy would again vanish, a so-called renormalization of the Hamilton operator. More generically, only energy differences matter. However, this does not mean that the particle in the lowest level $n = 1$ is at rest.

To see this, first normalize (6.5),

$$\int dx |\psi(x)|^2 = 1 \implies A = \sqrt{\frac{2}{L}}$$

[1] Actually, only exponentially confined, as will be seen in section 6.4.

and then calculate

$$\begin{aligned}
\langle X\rangle_{n=1} &= \int dx x|\psi_{n=1}(x)|^2 = \frac{L}{2} \\
\langle X^2\rangle_{n=1} &= \int dx x^2|\psi_{n=1}(x)|^2 = \frac{L^2(2\pi^2-3)}{6\pi^2} \\
\langle(\Delta X)^2\rangle_{n=1} &= \frac{\pi^2-6}{12\pi^2}L^2 \approx 0.033L^2.
\end{aligned}$$

Thus the particle is centered at the middle of the potential, but lies even in the lowest state not still, but is distributed a little bit around the middle. Thus, it can be found elsewhere. Hence, even though it is possible to shift the value of the energy, the actual behavior of the particle remains the same.

These four features are generic properties of quantum systems with a potential which sustains bound states. They will be encountered repeatedly throughout the rest of the lecture. These features are thus worth to explore a bit more in further examples.

6.2 The harmonic oscillator

And what would for this purpose be better suited than the guinea pig of theoretical physics, the harmonic oscillator? Of course, just as in classical physics, it will not only be relevant for its own sake, but will reappear in many disguises throughout quantum physics. In fact, because it offers so much utility, it will be discussed in two quite different forms. First, it will be discussed very much in the same way as the infinite-well potential of section 6.1. Second, a more abstract approach will be used, which exhibits on the one hand a similar development as in classical mechanics. On the other hand, it is especially this second form which is of great conceptual importance in quantum physics, and actually also its generalization in form of quantum field theory.

The importance of the harmonic oscillator can be realized immediately when one considers the problem of equilibrium. Assume that some particle is affected by a potential $U(x)$. The state of lowest energy is classically the equilibrium position. Of course, lowest energy means that the particle is at the position x_0 with least potential energy, i. e. where the derivative of U with respect to x vanishes.

What happens if the particle is slightly displaced? For small displacements it suffices to regard the potential close by to the minimum, and there the potential can be approximated by its Taylor series

$$U(x) \overset{x\approx x_0}{\approx} U(x_0) + \left.\frac{dU}{dx}\right|_{x=x_0}(x-x_0) + \frac{1}{2}\left.\frac{d^2U}{dx^2}\right|_{x=x_0}(x-x_0)^2 + \mathcal{O}((x-x_0)^3).$$

Now, the first term is just a constant, and can therefore be neglected, as only energy differences can be measured. The second term vanishes, because the potential has a minimum at $x = x_0$. The third term is just an harmonic oscillator potential, if the frequency of the oscillator ω is identified as

$$\omega^2 = \frac{1}{m} \left. \frac{d^2U}{dx^2} \right|_{x=x_0},$$

where m is the mass of the particle. Thus, any[2] small movement around an equilibrium position is described by the harmonic oscillator. In particular, small movements around a minimum maybe affected by quantum effects, even if larger movements are not.

Thus, the interesting object is the Hamilton operator of the harmonic oscillator

$$H = \frac{P^2}{2m} + \frac{m\omega^2}{2} X^2, \tag{6.8}$$

where the minimum has been arbitrarily shifted to zero, for ease of calculations. To fully characterize the quantum harmonic oscillator, it is necessary to determine its energy spectrum and the eigenstates.

6.2.1 Position space

The problem is again separable, and therefore the time-dependence is only a factor of $\exp(iEt/\hbar)$, just as in (6.2). Then, the stationary Schrödinger equation reads

$$H\phi(x) = \left(-\frac{\hbar^2}{2m} d_x^2 + \frac{m\omega^2}{2} x^2 \right) \phi(x) = E\phi(x). \tag{6.9}$$

It is useful to introduce a few shortenings to absorb all dimensionful quantities

$$\begin{aligned} a_H &= \sqrt{\frac{\hbar}{m\omega}} \\ r &= \frac{x}{a_H} \\ \epsilon &= \frac{2E}{\hbar\omega} \end{aligned} \tag{6.10}$$

The quantity a_H has units of length. It is a characteristic length scale of the harmonic oscillator. Then r is the distance in units of this length scale, and ϵ is energy in units of the characteristic energy scale $\hbar\omega$. Since the mass enters the length scale, but not the

[2] Of course, it is assumed that it is a true minimum, and therefore a stable equilibrium, and hence the second derivative is positive and non-zero. There exist also other cases, where either the potential is non-analytic or it has a minimum of higher order. These rather rare cases will not be treated here.

energy scale, it appears reasonable to associate to it the same meaning as the de Broglie or the Compton wave-length, and thus it characterizes the movement of the particle in the oscillator, while ϵ characterizes the oscillator itself. This also requires a Jacobian, as $a_H dr = dx$.

In total, this turns (6.9) into

$$\left(d_r^2 + (\epsilon - r^2)\right)\phi(r) = 0. \tag{6.11}$$

This is an ordinary differential equation, though not a simple one. Since similar cases are encountered frequently both during this lecture as well as in classical and quantum field theory, it is worthwhile to spend some time on this equation, and to solve it systematically.

The first step is to understand the asymptotic behavior of the solution, i. e. what happens at $r \gg \epsilon$ and $r \ll \epsilon$. For $r \gg \epsilon$ the term ϵ can be neglected, and the equation reads

$$(d_r^2 - r^2)\phi(r) = 0.$$

This does not seem to be easier at first sight. But, make an ansatz $\phi(r) \sim \exp(\pm r^2/2)$, which yields

$$d_r^2 e^{\pm\frac{r^2}{2}} = r^2 e^{\pm\frac{r^2}{2}} \pm e^{\pm\frac{r^2}{2}}. \tag{6.12}$$

At large r, the second term is negligible, and therefore the equation is solved. Thus, at larger r ϕ must behave in this way up to subexponential corrections. Since the final wave-function must be normalizable, this implies that the sign in the exponential must be negative.

Likewise, at small r the term r^2 in (6.11) can be neglected. But this is the normal oscillation equation. However, because the frequency is $\sqrt{\epsilon}$, this implies that only the leading behavior in the Taylor series is relevant, as only then $r \ll \epsilon$ is fulfilled. Thus, at small r the solution behaves either linear or constant. But this in itself suggests a polynomial behavior.

Combining both information suggests to make an ansatz

$$\phi(r) = e^{-\frac{r^2}{2}} u(r), \tag{6.13}$$

yielding

$$(d_r^2 - 2rd_r + \epsilon - 1)u(r) = 0. \tag{6.14}$$

Given the information at small r suggests as an ansatz a Taylor series[3] for $u(r)$

$$u(r) = \sum_i a_i r^i. \tag{6.15}$$

[3]There is not yet a guarantee that this will work. It is just a first guess, which ultimately will be vindicated by being a solution. In general, this is not guaranteed, but quite often the case.

Inserting (6.15) into (6.14) yields

$$\sum_{i>1} i(i-1)a_i r^{i-2} - 2\sum_{i>0} i a_i r^i - (1-\epsilon)\sum_i a_i r^i = 0.$$

Nicely, a structure emerges where only two types of powers in r appear. Shifting the summation index to put everything under a single sum yields

$$\sum_i (i+2)(i+1)a_{i+2} r^i - 2\sum_i i a_i r^i - (1-\epsilon)\sum_i a_i r^i = 0.$$

Collecting the prefactors to the same powers of r yields

$$(i+2)(i+1)a_{i+2} + (\epsilon - 2i - 1)a_i = 0,$$

which thus relates the coefficient of the power i with the one of power $i+2$. This can be reformulated as a recursion relation

$$a_{i+2} = \frac{2i+1-\epsilon}{(i+2)(i+1)} a_i, \tag{6.16}$$

and thus the coefficients for larger values of i can be obtained from the ones at smaller values of i. There are now two particularities.

The first is that (6.16) does not connect coefficients for even and odd i, and thus there are two subseries, starting at 1 and r, respectively. Based on the asymptotic analysis above this was expected. Also, because this is a second-order differential equation there should be two independent solutions. They are therefore characterized by the starting values a_0 and a_1. The absolute size of these are not fixed, as the equation (6.14) is homogeneous in $u(r)$. This is expected. After all, as a wave-equation ultimately a normalization to unity to get a probability interpretation should eventually be possible.

The second is that (6.16) implies

$$a_{i+2} \overset{i\to\infty}{\sim} \frac{2a_i}{i}$$

and thus $a_{2i} \sim 1/i!$. Unfortunately, this implies a series like

$$\sum_i \frac{r^{2i}}{i!} \to e^{r^2}$$

which therefore cancels the Gaussian decay factored out in (6.13), and prevents a normalization of the wave function. A similar statement applies to the odd case. Hence, the only possibility to find a physically sensible solution is that the at some point the right-hand

side of (6.16) has to vanish, and therefore forcing all higher coefficients to vanish as well, making the series finite. This is achieved by requiring a termination condition

$$\frac{2E}{\hbar\omega} = \epsilon = 2i + 1$$

for some fixed value of i. Thus, the energy E can only take discrete levels

$$E_i = \hbar\omega\left(i + \frac{1}{2}\right). \tag{6.17}$$

Given the experience in section 6.1, this was to be expected. Also the appearance of a zero-point energy has already there been encountered. So, conceptually everything works in exactly the same way.

The series (6.15) will now terminate at an order becoming higher and higher with the energy level. The resulting polynomials are known as the Hermite polynomials H_i, and can be calculated, e. g. using (6.16). Conveniently normalized, the first few of them read

$$\begin{aligned}
H_0(r) &= 1 \\
H_1(r) &= 2r \\
H_2(r) &= 4r^2 - 2 \\
H_3(r) &= 8r^3 - 12r \\
H_4(r) &= 16r^4 - 48r^2 + 12
\end{aligned}$$

and are, as already shown, either even or odd. The wave-functions can be normalized using known identities for the Hermite polynomials, which will not be detailed here. Note that these polynomials are necessarily orthogonal, since the Hamilton is a Hermitian operator. They form therefore also an alternative base system to the ordinary monomials. They furthermore have a number of nodes equal to n, visible explicitly for $n = 1$ with its single node. They furthermore turn out to be even and odd at large r, i. e. they alternate with n between going to positive values for $r \to \pm\infty$ and to $\pm\infty$ for $r \to \pm\infty$. Since their behavior is therefore either the same for $r \to -r$ or they change their sign, they are said to have even and odd parity, where the lowest one, H_0, starts with even parity.

The total result for the wave-function for the nth level is then

$$\psi_n(x,t) = \frac{1}{\sqrt{a_H}} \frac{\pi^{-\frac{1}{4}}}{\sqrt{2^n n!}} H_n\left(\frac{x}{a_H}\right) e^{-\frac{x^2}{2a_H^2} - i\frac{E_n t}{\hbar}} \tag{6.18}$$

where the characteristic length a_H is given in (6.10). Note that n can be reexpressed by (6.17) in terms of the energy, and therefore not only the time-dependency is driven by the energy-level.

A characteristic feature, which can also be seen from the infinite-well case (6.5), is that the number of nodes, i. e. zero crossings of the wave-function, is given exactly by the energy level. Thus, the ground state has zero nodes, the first state one, and so on. The importance of these nodes is that the probability density of finding the particle there vanishes. This seems quite odd. In a way, this can be seen as a first example of self-interference of the wave describing the particle. This is a feature which will be returned to in section 7.2.3, as it leads to quite intriguing consequences.

However, in comparison to the infinite-well potential of section 6.1, there is also one new feature. Classically, a particle of energy E can only move in the range

$$-\sqrt{\frac{2E}{m\omega^2}} \leq x \leq \sqrt{\frac{2E}{m\omega^2}}. \tag{6.19}$$

But the wave-function (6.18) is non-zero for all values of x for all energies, though because of the reasoning which has lead to (6.12), it will essentially decay Gaussian-like in this area. Still, this means there is a non-zero probability for a particle to be found in a classically forbidden region. On the one hand, this has dramatic consequences, especially for the possibility of classically forbidden decays. This will be studied in more detail later in section 6.3. The other is the direct question of whether the particle can actually be detected in a classical forbidden region. While strictly theoretical the answer appears to be yes, this is in any practical realization a much more subtle problem. If, e. g., the potential is due to a wall, can a detector be placed there? On the other hand, if the potential is realized by a spring, this could mean that quantum-mechanically the spring could just be extended further due to the quantum effects than classically possible.

Alternatively, it is easier to consider the question what the probability is to not find the particle in the classically allowed region. Consider the case $n = 0$ and set a_H to one for ease of calculation. Then

$$\phi_0(x) = \pi^{-\frac{1}{4}} e^{-\frac{x^2}{2}}$$

and the classical points of return are

$$\pm\sqrt{\frac{\hbar}{m\omega}} = \pm\sqrt{a_H} = \pm 1.$$

The probability to find the particle in this range is therefore

$$P_c = \pi^{-\frac{1}{2}} \int_{-1}^{1} dx e^{-x^2} \approx 0.84.$$

Thus, there is a ≈16% chance to find the particle not inside the classical allowed region. In fact, this turns out to be the maximum, and the higher the energy the less likely this

becomes. Note also that the expected position and its uncertainty is

$$\begin{aligned}\langle X\rangle &= 0\\ \langle X^2\rangle &= \frac{1}{2}\\ \langle(\Delta X)^2\rangle &= \frac{1}{2}.\end{aligned}$$

Hence, the expected position, including its uncertainty, is $0\pm 1/2$, and thus well contained inside the classical region $[-1,1]$. Thus, even the probability to find it outside the allowed region is non-zero, it is not expected. Again, these are features which are quite generic for quantum systems.

6.2.2 Operator language

Though the problem of the harmonic oscillator is now fully solved, it is actually neither the most efficient nor most instructive way to solve the problem. Also, it is one which cannot be very well generalized. It is therefore better to investigate an alternative way of solving the problem, now that the answer is known. This will lead to a conceptual important recasting of the problem. This recasting will play a foundational role to developing quantum theory further, especially towards a full quantum field theory in later courses[4].

To proceed, note that the Hamilton operator (6.8) has the structure of almost a square. To complete the square, as usual the mixed term has to be added and subtracted. Since the present case is quantum mechanics, rather the commutator should be added and subtracted. To obtain the right dimension requires an additional pre-factor, which by dimensional analysis has to be ω. In principle, this should be done with an arbitrary constant of proportionality, which will be fixed here with hindsight to be $1/2i$. Thus

$$\begin{aligned}&\frac{P^2}{2m}+\frac{m\omega^2}{2}X^2+\frac{\omega}{2i}([X,P]+[P,X])\\ =\;&\frac{P^2}{2m}+\frac{m\omega^2}{2}X^2+\frac{\omega}{2i}(i\hbar+PX-XP)\\ =\;&\hbar\omega\frac{m\omega}{2\hbar}\left(\frac{P^2}{\omega^2m^2}-\frac{i}{m\omega}(PX-XP)+X^2\right)+\frac{\hbar\omega}{2}\\ =\;&\hbar\omega\frac{m\omega}{2\hbar}\left(X-\frac{iP}{m\omega}\right)\left(X+\frac{iP}{m\omega}\right)+\frac{\hbar\omega}{2}\end{aligned}$$

[4]This is also the historical route to quantum field theory. In modern treatments, an alternative, but fully equivalent approach, the path integral to be discussed in section 11.3, is often found to be superior. But this construction is much more demanding to develop and use in quantum mechanics, and therefore not the primary route of this lecture.

Indeed, up to a constant, this now has the structure of a square. Identifying the hermitian conjugate operators in parentheses as[5]

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(X + \frac{iP}{m\omega}\right) \tag{6.20}$$

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(X - \frac{iP}{m\omega}\right) \tag{6.21}$$

permits to write the harmonic oscillator in a much simpler form as

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(N + \frac{1}{2}\right),$$

where in the second step the Hermitian operator $N = a^\dagger a$ has been defined. This operator commutes with the Hamilton operator, and therefore can be simultaneously diagonalized. Especially, the eigenvalue problem of the Hamilton operator can now be recast into an eigenvalue problem for N and its eigenstates $|n\rangle$ and eigenvalues n as

$$E\,|E\rangle = H\,|E\rangle = \hbar\omega\left(N + \frac{1}{2}\right)|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle \tag{6.22}$$

Though this looks fancy, not much seemed to be gained at this point.

To achieve results, the algebra of the three operators a, $a^\dagger$ and N is the key. First, the commutator of a and $a^\dagger$ can be obtained by their mapping to the momentum and position operators, and then for N by its relation to a and $a^\dagger$ as

$$[a, a^\dagger] = \frac{m\omega}{2\hbar}\left([X,X] + \frac{1}{m^2\omega^2}[P,P] - \frac{i}{m\omega}[X,P] + \frac{i}{m\omega}[P,X]\right) = \frac{1}{2\hbar}\left(0 + 0 + \hbar + \hbar\right) = 1 \tag{6.23}$$

$$[N, a] = [a^\dagger a, a] = a^\dagger[a,a] + [a^\dagger, a]a = -a \tag{6.24}$$

$$\left[N, a^\dagger\right] = [a, N]^\dagger = a^\dagger. \tag{6.25}$$

This is a rather simple algebra. It will now be used to determine the energy spectrum.

To obtain it, note that

$$\begin{aligned} Na^\dagger\,|n\rangle &= \left([N, a^\dagger] + a^\dagger N\right)|n\rangle = a^\dagger(1+N)\,|n\rangle = (n+1)a^\dagger\,|n\rangle \\ Na\,|n\rangle &= \left([N, a] + aN\right)|n\rangle = a(-1+N)\,|n\rangle = (n-1)a\,|n\rangle \end{aligned}$$

This implies that the states created by a and $a^\dagger$ when acting on an eigenstate of N (and therefore of the Hamilton operator) are states with an eigenvalue of N smaller or larger

[5]Note that, for the sake of compatibility with the prevalent standard notation, the operators a and $a^\dagger$ are not denoted by capital letters.

by one, respectively. Or states with lower and higher energy, respectively. Therefore, these operators are also called annihilation operator and creation operator. They permit to change from one eigenstate to another.

The normalization of these states can be immediately deduced from their norm

$$0 \leq \langle n| a^\dagger a |n\rangle = \langle n| N |n\rangle = n. \tag{6.26}$$

This incidentally shows that n cannot be negative, which will be useful later on. Choosing the normalization to be positive and real yields

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

Similarly it follows that

$$\langle n| aa^\dagger |n\rangle = \langle n| a^\dagger a - a^\dagger a + aa^\dagger |n\rangle = \langle n| N + 1 |n\rangle = n + 1,$$

yielding

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

completing the list.

Now, the harmonic oscillator Hamilton (6.8) is a sum of squares of hermitian operators with positive pre-factors. It is therefore certainly bounded from below. However, it appears at first sight that the annihilation operator can decrease a state arbitrarily often. How does this fit together? Applying the annihilation operator m times to a state n yields

$$a^m |n\rangle = \sqrt{n(n-1)...(n-m+1)} |n-m\rangle .$$

If m is integer, this sequence terminates if n becomes m, since the pre-factor becomes zero. If m is not an integer, the sequence can be arbitrarily extended, eventually yielding negative n, in contradiction to (6.26). Hence, the only admissible values for n are positive integers, or zero. The possibility for zero is consistent as

$$a |0\rangle = \sqrt{0} |-1\rangle = 0,$$

and thus a destroys the zero state and yields the so-called vacuum state $|0\rangle$ as the state of lowest energy.

But now everything is finished: The spectrum of N are the non-negative integers, and therefore because of (6.22) the spectrum of the Hamilton operator is

$$E_n = \hbar\omega \left(n + \frac{1}{2}\right),$$

with n positive or zero, and the eigenstates $|n\rangle$. Hence the energies are evenly spaced, and there is again the zero-point energy. Because the number n counts the energy level, the operator N is also called the counting operator. The evenly spaced energy levels are also called oscillator quanta, and one speaks of the nth level as having n quanta in it.

The new operators also permit a very efficient way of calculating the wave-functions. Start with the ground state wave-function $\psi_0(x) = \langle x|0\rangle$. Since the vacuum is destroyed by the destruction operator a, it follows that

$$0 = \langle x|a|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \left\langle x \left| \left(X + \frac{iP}{m\omega} \right) \right| 0 \right\rangle = \sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{\hbar}{m\omega} \frac{d}{dx} \right) \langle x|0\rangle.$$

This is a first-order linear differential equation for the vacuum wave-function, instead of the second-order Schrödinger equation (6.9). It can therefore be solved by separating variables as

$$0 = \int_y \left(xdx + \frac{x_0^2}{\psi_0(x)} d\psi_0(x) \right) = \int_y \left(xdx + \frac{a_H^2}{\psi_0(x)} \frac{d\psi_0(x)}{dx} dx \right) \overset{y\to x}{=} \frac{1}{2}x^2 + a_H^2 \ln(\psi_0(x)) + C,$$

where C is a constant of integration and

$$a_H = \sqrt{\frac{\hbar}{m\omega}},$$

is the same as (6.10). Solving for the wave function finally yields

$$\psi_0(x) = \exp(-C) \exp\left(-\frac{x^2}{2a_H^2} \right).$$

and normalization finally determines C to be

$$\psi_0(x) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{a_H}} \exp\left(-\frac{x^2}{2a_H^2} \right),$$

the ground state wave-function of the harmonic oscillator. As announced, this wave-function is non-trivial. It is a Gaussian with width a_H, i. e. the width is determined by a combination of the properties of the potential and of the particle.

Generating the other states is also simpler. Since the creation operator needs only to be applied to the ground state to create states of arbitrary high level, the first excited state is given by

$$\langle x|1\rangle = \left\langle x \left| \frac{1}{\sqrt{1}} a^\dagger \right| 0 \right\rangle = \frac{1}{\sqrt{2}a_H} \left(x - a_H^2 \frac{d}{dx} \right) \langle x|0\rangle = \frac{\sqrt{2}}{\pi^{\frac{1}{4}} a_H^{\frac{3}{2}}} x \exp\left(-\frac{x^2}{2a_H^2} \right)$$

and so on for higher levels, yielding already normalized wave functions. The structure implies that this will be a polynomial multiplied with the Gaussian. These polynomials are again the Hermite polynomials, yielding the same result as (6.18).

6.2.3 Chains of oscillators

The new operators therefore make the harmonic oscillator comparatively simple. In fact, in many cases it is possible to map a system onto operators with the same algebra, (6.23-6.25). Since the same algebra implies the same eigensystem, this solves eigenvalue problems very quickly, if the back-transformation is sufficiently simple. They have therefore become a very powerful tool throughout quantum physics. This is similar to the Hamilton-Jacobi theory of classical mechanics.

In particular, if there are M independent oscillators, their Hamilton operator is

$$H = \sum_{i=1}^{M} H_i = \hbar \sum_{i=1}^{M} \omega_i \left(a_i^\dagger a_i + \frac{1}{2} \right).$$

The Hilbert space of this system is then a product space of M one-particle Hilbert spaces, $\mathcal{H} = \otimes_i \mathcal{H}_i$. Especially, a state in this Hilbert space can be denoted by $|\{n_i\}\rangle$, and is therefore uniquely characterized by the components n_i. In this case, the oscillators are all distinguishable. When turning to quantum systems in chapter 10, and later to many-particle systems in the lectures on statistical quantum physics, this will become the prototype of a (non-interacting) system of independent, distinguishable particles.

In this context, the oddity of the vacuum energy has also very far reaching consequences. The spectrum of this chain of oscillators is uniquely given by the n_i, and is thus the sum of all the individual oscillators, say for all $\omega_i = \omega$

$$E = \sum_{i=1}^{M} E_i = M\frac{\hbar\omega}{2} + \hbar\omega \sum_{i=1}^{M} n_i.$$

Hence, irrespective of the actual occupation of the oscillators, there is always a vacuum energy $M\hbar\omega/2$. As before, this is the difference between the vacuum energy of the classical oscillator and the quantum one. If the number of oscillators becomes large, and especially if $M \to \infty$, this vacuum energy grows beyond bounds, no matter what the values of the n_i are. This divergence is the very simplest example of divergences which will be encountered very often when moving to many-particle physics. Here, this can be cured by shifting the energy by the vacuum energy, a process which is called, for historical reasons, renormalization. In the present case, this is not of physical relevance. In particle physics, this will become a hallmark of theories which are only valid at low energies, and one of the strongest hints that our current knowledge of particle physics is fundamentally incomplete. Thus, this is in analogy to section 5.8. In the context of quantum gravity, this is even more involved, as here absolute energy scales will play a role, and renormalization needs to be performed quite differently, and is not yet fully understood.

6.3 Potential step, transmission and reflection

So far, the situation had been either that there was a confinement of the states or no potential at all. A next step will therefore be a potential, which is non-confining, but not everywhere the same. The simplest such example is the potential step

$$V(x) = \begin{cases} 0 \text{ if } x < 0 \\ V_0 \text{ if } x \geq 0 \end{cases}$$

where $V_0 > 0$. Since it is always possible to shift the zero of the energy by adding a constant and to flip for this potential $x \to -x$, this does automatically also cover the case $V_0 < 0$. It is only important that $V_0 \neq 0$.

The Schrödinger equation has again the same form (6.1) and can therefore be immediately reduced to (6.3), the stationary problem without time-dependence. The special point of this problem is $x = 0$ and $E = V_0$. For $E \geq V_0$ the particle can classically move everywhere, while for $0 \leq E \leq V_0$ the particle is classically forbidden to enter the region $x > 0$. However, as the example of the harmonic oscillator in section 6.2.1 has shown, a particle can penetrate this region.

To make an ansatz to start solving the stationary problem, note the following: $V(x)$ has a discontinuous jump at $x = 0$. The stationary Schrödinger equation therefore implies that at least the second derivative of the stationary wave-function $\phi(x)$ needs to have such a jump as well. However, if the first derivative or the wave-function would have such a jump its derivatives become δ-functions or derivatives of δ-functions, which cannot be canceled in the Schrödinger equation. Therefore, the wave-function needs to satisfy

$$\lim_{0<\epsilon\to 0} \phi(x+\epsilon) = \lim_{0<\delta\to 0} \phi(x-\delta) \tag{6.27}$$

$$\lim_{0<\epsilon\to 0} d_x\phi(x+\epsilon) = \lim_{0<\delta\to 0} d_x\phi(x-\delta). \tag{6.28}$$

Thus, it is useful to consider the two cases $x <> 0$ separately.

At $x < 0$ the situation is that of the free particle in chapter 5. The unnormalized wave functions are therefore $\exp(\pm ikx/\hbar)$. Thus, a suitable ansatz will be

$$\phi(x < 0) = \left(Ae^{ikx} + R'e^{-ikx}\right)$$

where the relative factor R' describes the difference of a wave moving towards the potential step, i. e. the first term, and the wave-function moving away from it. Just as in chapter 6 the wave-parameter k has to satisfy

$$E = \frac{\hbar^2 k^2}{2m}.$$

For $x > 0$ there are two options. One is $E > V_0$. Then the particle behaves as a free particle, but with an energy reduced by the constant potential,

$$\begin{aligned} \phi(x>0) &= \left(T' e^{iqx} + B e^{-iqx}\right) \\ E - V_0 &= \frac{\hbar^2 q^2}{2m}. \end{aligned}$$

Of the four free parameters A, B, R', and T' one is determined by the overall normalization. In addition, (6.27-6.28) imply

$$\begin{aligned} A + R' &= B + T' \\ k(A - R') &= q(T' - B) \end{aligned}$$

which yield for R' and T

$$R' = \frac{A(k-q) + 2Bq}{k+q} \tag{6.29}$$

$$T' = B + \frac{2(A-B)k}{k+q}. \tag{6.30}$$

This leaves one unknown coefficient open. This results from the discontinuity of the differential equation, and that effectively two times second-order differential equations had to be solved. This can be used to model different physical situations.

Probably the most interesting case is if there is no incident particle from the right, and thus $B = 0$. This corresponds to the classical situation that a particle is shot above a wall with some kind of speed reduction beyond the wall. Classically, it will just travel on. However, quantum-mechanically (6.29-6.30) yield

$$R = \frac{R'}{A} = \frac{k-q}{k+q} \tag{6.31}$$

$$T = \frac{T'}{A} = \frac{2k}{k+q}. \tag{6.32}$$

where the constants R and T are thus normalized with respect to the total normalization. This implies that a part of the wave, characterized by R, is actually reflected. Thus, quantum-mechanically a probability exist that the particle is reflected, even if it has enough energy to pass the potential step.

The probability density to find a particle on the left-hand side per unit length and normalized to the total normalization factor is then

$$\lim_{a\to-\infty} \frac{1}{a} \int_{-a}^{0} dx |\phi|^2 = \lim_{a\to-\infty} \frac{a + aR^2 + \frac{1}{k}\sin(2ak)}{a} = |A|^2(1+R^2)$$

and on the right

$$\lim_{a\to\infty}\frac{1}{a}\int_0^a dx|\phi|^2 = \lim_{a\to\infty}\frac{aT^2}{a} = |A|^2T^2.$$

Since the particle must be somewhere, this seems to suggest that $|A|$ must be zero. The reason for this is the same as in section 4.3, and is not relevant here. So, while the probability to find a particle to the left per unit length and to the right are not the same, in contrast to the classical case, to continue the interpretation is still not trivial.

To make progress, it is useful to introduce the notion of a probability current. To this end, note that the probability density $\rho(t,x) = |\psi(t,x)|^2$ can change locally with time, i. e. $\partial_t\rho(t,x)$ can be non-zero. On the other hand, the total probability needs to be independent of time,

$$0 = \partial_t 1 = \partial_t \int dx\rho(t,x) = \int dx\partial_t\rho(t,x).$$

This suggests to investigate $\partial_t\rho(t,x)$ closer, yielding

$$\begin{aligned}\partial_t\rho(t,x) &= \psi\partial_t\psi^* + \psi^*\partial_t\psi = \frac{i}{\hbar}\left(\psi H^*\psi^* - \psi^* H\psi\right)\\ &= \frac{i}{\hbar}\left(-\frac{\hbar^2}{2m}\left(\psi\partial_x^2\psi^* - \psi^*\partial_x^2\psi\right) + \psi^*V\psi - \psi^*V\psi\right) = -\frac{\hbar}{2im}\partial_x\left(\psi^*\partial_x\psi - \psi\partial_x\psi^*\right)\\ &= -\partial_x j(x) \qquad (6.33)\end{aligned}$$

where it was assumed that the potential is real. This is a good assumption for most problems, but will be needed to be reconsidered once some sink or source for particles and/or energy appears. This will be noted, if it happens. This defines the probability current

$$j(t,x) = \frac{\hbar}{2im}\left(\psi^*\partial_x\psi - \psi\partial_x\psi^*\right)$$

and establishes the continuity equation

$$\partial_t\rho(t,x) + \partial_x j(t,x) = 0. \qquad (6.34)$$

This can be interpreted as that a local change of probability needs to be accompanied by a current, which transports the probability into the surrounding area. In this sense, probability behaves like a fluid or an (electrical) current. Therefore, this is called (probability) current conservation.

If the situation is stationary, $\psi(t,x) = N\exp(iEt/\hbar)\phi(x)$, the probability current and density are

$$\begin{aligned}j(t,x) &= \frac{\hbar|N|^2}{2im}(\phi(x)^*\partial_x\phi(x) - \phi(x)\partial_x\phi(x)^*)\\ \rho(t,x) &= |N\phi(x)|^2\end{aligned}$$

and thus the probability density is time-independent. Likewise,

$$\partial_x j(x,t) = \frac{\hbar|N|^2}{2im}\left(\partial_x\phi^*\partial_x\phi + \phi^*\partial_x^2\phi - \partial_x\phi(x)\partial_x\phi(x)^* - \phi(x)\partial_x^2\phi(x)^*\right) = 0$$

by virtue of (6.33), as ∂_x^2 can be expressed as $H - V$, which has always been assumed to be Hermitian.

Returning to the original problem, the probability current of a plane wave is

$$j(t,x) \sim \left(e^{-ikx}\partial_x e^{ikx} - e^{ikx}\partial_x e^{-ikx}\right) = 2ik. \tag{6.35}$$

As the factor of proportionality is not finite, it is useful to normalize everything. Thus, there are three relevant currents. One is the incoming current j_0, moving towards the step. Then there is the reflected current j_R and the transmitted current j_T, yielding

$$\begin{aligned} j_0 &= \frac{|A|^2\hbar k}{m} \\ j_R &= \frac{|A|^2R^2\hbar k}{m} \\ j_T &= \frac{|A|^2T^2\hbar q}{m} \end{aligned}$$

which implies

$$\frac{j_R + j_T}{j_0} = R^2 + \frac{q}{k}T^2 = \left(\frac{k-q}{k+q}\right)^2 + \frac{4qk}{(k+q)^2} = \frac{k^2 - 2kq + q^2 + 4qk}{(k+q)^2} = 1.$$

Thus, the incoming probability current equals exactly the sum of reflected and transmitted current. Thus, no probability is lost in the process. The fraction

$$\frac{j_R}{j_T} = \frac{(k-q)^2}{4kq}.$$

also provides an interpretation. If $E \gg V_0$ then $k \approx q$, and the ratio tends to zero. If the particle has a lot of energy, nothing is reflected anymore. However, this is an asymptotic process. Likewise, if $E \approx V_0$ then $q \approx 0$, and the ratio diverges, as j_T tends to zero. In particular for $E = V_0$ q vanishes, and so does j_T. Thus, even though there is a non-zero probability for the particle to be in this case on the other side, the actual transmitted current vanishes. Hence, the particle can no longer cross the step and escape to infinity. In this sense classical physics is again recovered only in the limit of $E \gg V_0$.

It is an interesting question to ask, what happens if $E < V_0$. Classically, the particle cannot penetrate the wall, and will always be reflected. However, already in the case of the harmonic oscillator in section 6.2.1 it was found that this does not remain true

quantum-mechanically. Indeed, the only difference is that the static Schrödinger equation on the right-hand side now reads

$$(E - V_0)\phi(x) = -\frac{\hbar^2}{2m} d_x^2 \phi(x),$$

and thus the sign on the right-hand-side changes. Thus, this is the equation for an exponential decay, and solved by

$$\begin{aligned} \phi(x) &= T e^{-\frac{\kappa x}{\hbar}} \\ V_0 - E &= \frac{\hbar^2 \kappa^2}{2m}. \end{aligned}$$

The second solution is an exponential increasing function, and thus not admissible. Therefore, the situation is very similar to the case of an incident plane-wave, only that the particle can no longer move to infinity. Still, there is a non-zero probability density again in a classically forbidden region, just as with the harmonic oscillator. In this form, the continuity equations (6.29-6.30) have the same form, as do the solutions (6.31-6.32), except that it is necessary to replace everywhere q with $i\kappa$.

What does change is the transmitted current j_T. It now takes the form

$$j_T \sim |T|^2 \left(e^{-\kappa x} \partial_x e^{-\kappa x} - e^{-\kappa x} \partial_x e^{-\kappa x} \right) = 0, \tag{6.36}$$

and therefore no longer particles are transmitted. Thus, it is still true that no particle can move, in a classical sense, to the right, despite the non-zero probability density. Hence, even quantum-mechanically this barrier is 100% efficient in reflecting particles which have too small energy. This can also be seen by the fact that $|R|^2 = 1$, even though $R \neq 1$. Thus, the only thing that happens is that the reflected wave gets a phase shift δ, $R \exp(-ikx/\hbar) = \exp(-i(kx/\hbar - \delta))$,

$$\begin{aligned} e^{i\delta} &= \frac{k - i\kappa}{k + i\kappa} \\ \delta &= \cos^{-1} \frac{k^2 - \kappa^2}{k^2 + \kappa^2} \end{aligned}$$

which goes to 0 when the energy reaches the threshold. Such phase shifts are typical for scattering phenomena, and they play a huge role in experiments.

6.4 δ-potential, bound states, and scattering states

So far, the investigated cases had potentials which either completely forbid a movement to infinity or were (modified) free cases. In classical physics there is a third kind of situations,

potentials which allow both for bound states, i. e. a particle cannot move arbitrarily away from a force center, or unbound states, if the energy is just large enough. The best known example of this type is the gravitational potential, e. g. a planet orbiting a star. While this particular situation will be discussed in chapter 7, the generic situation can already be discussed in a much more simple case. As such bound solutions lead in quantum mechanics to bound (or composite) states, and later particles like the hydrogen atom, nuclei, and much more, this is a conceptually extremely important topic.

The technically simplest such case is the δ-potential,

$$V(x) = V_0\delta(x),$$

where V_0 can be either positive or negative. This can be seen as a highly localized force source. It is, e. g., useful to describe the dominant effect of ions on electrons in certain solids, or also distortions in a crystal structure. Again, this is a static potential, and it will therefore be sufficient to consider the static Schrödinger equation.

However, the δ-function is highly singular, and its integral, the θ-function, is still non-continuous. Therefore, in contrast to the case of the step in section 6.3, only the wave-function itself needs to be continuous. To understand the implications for the derivative, consider the Schrödinger equation

$$E\phi(x) = -\frac{\hbar^2}{2m}d_x^2\phi(x) + V_0\delta(x)\phi(x).$$

This equation can be integrated over the interval $[-\epsilon, \epsilon]$, with ϵ infinitesimal. Because the wave-function needs to be continuous, it is possible to approximate $\phi(x)$ in this interval by a constant, and therefore the left-hand side vanishes. The kinetic term is a total derivative, which yields the primitive upon integration, and the δ-function integrates to $\phi(0)$. This yields

$$0 = -\frac{\hbar^2}{2m}\left((d_x\phi)(\epsilon) - (d_x\phi)(-\epsilon)\right) + V_0\phi(0) \tag{6.37}$$

This implies that $d_x\phi(x)$ is discontinuous at $x = 0$, and the discontinuity, that is the limit of the first term in the limit of $\epsilon \to 0$, needs to satisfy this equation. This will link the derivative of the wave-function and the wave-function at $x = 0$.

Left and right of $x = 0$ the situation is that of the free case. However, as it is not yet clear how the presence of the potential at $x = 0$ will affect the admissible values of E it is best to not yet exclude negative values of E. Furthermore, based on the discussion in section 6.3, it seems useful to make for $E > 0$ the ansatz

$$\begin{aligned}\phi(x) &= N(\theta(-x)\left(e^{ikx} + Re^{-ikx}\right) + \theta(x)Te^{ikx}) \\ E &= \frac{\hbar^2k^2}{2m},\end{aligned}$$

as it reasonable to expect that a potential bump will act on an incident wave as a potential step, as long as the energy is large enough to cross it. Of course, also a wave incident from the right, or a, symmetric or asymmetric, superposition of both cases, could be chosen. But as before, this will only obscure the physical mechanism, while being admissible solutions. The continuity of the wave-function imposes the same condition as in section 6.3, $1 + R = T$.

The situation for the ansatz for $E < 0$ is slightly different than in section 6.3. Now the system is left-right symmetric. Still, away from $x = 0$ there should be no propagation, and thus a suitable normalizable ansatz is

$$\begin{aligned} \phi(x) &= \left(A\theta(-x)e^{\kappa x} + B\theta(x)e^{-\kappa x}\right) \\ E &= -\frac{\hbar^2\kappa^2}{2m}, \end{aligned}$$

The symmetry of the situation suggests that the coefficients A and B must be equal. This is also enforced by the continuity of the wave-function itself, leading to the common factor $A = B$.

The discontinuity condition yields for $E > 0$ and $E < 0$

$$-\frac{i\hbar^2 k}{2m}(T - (1 - R)) + V_0 T = 0 \tag{6.38}$$

$$-\frac{A\hbar^2\kappa}{2m}(-1 - 1) + V_0 A = 0, \tag{6.39}$$

respectively. In the first case, it does not matter on which side $\phi(0)$ is evaluated, as the continuity of the wave-function enforces both cases to be the same. (6.38) then provides a second condition on R and T, and thus it is possible to solve it, yielding

$$\begin{aligned} T &= \frac{1}{1 + i\frac{mV_0}{\hbar^2 k}} \\ R &= \frac{1}{1 - i\frac{\hbar^2 k}{mV_0}}, \end{aligned}$$

and thus both are determined by the ratio $mV_0/(\hbar^2 k)$. The result is a scattering solution, very much as in section 6.3. However, there is one interesting observation. An infinite potential $V_0 > 0$ like this would stop a classical particle in its tracks, as an infinite amount of energy would be necessary to have a positive kinetic energy at zero. Quantum-mechanically, this is not the case, and there is always a non-vanishing transmitted probability current. This is a first example of a tunneling process, which will be discussed in more detail in sections 6.5 and 6.6.

Even more interesting is the case $E < 0$ in (6.39). Here, the prefactor cancels. It will therefore be entirely determined by the normalization. Thus, the condition (6.39) actually reads

$$\frac{\hbar^2\kappa}{m} = -V_0.$$

As κ is positive, this implies that a solution only exists if V_0 is negative. That is not unexpected, as even in quantum-mechanics no solution exists if the energy is everywhere smaller than the potential. It is instructive to rewrite this expression in terms of the energy,

$$E = -\frac{\hbar^2\kappa^2}{2m} = -\frac{mV_0^2}{2\hbar^2} \tag{6.40}$$

This implies that in this case there is only one admissible discrete value of the energy, and not a continuum of positive energy values as in the scattering case. This state is also separated from the continuum by an energy gap. As it is the lowest energy level, it is also the ground-state of the theory. The state is localized, due to the exponential damping: Given any fixed probability, there is a distance at which the particle is, even integrated, only found with a lower probability. Thus the particle is (exponentially) localized at the position of the potential, and thus bound. Therefore, it is indeed a bound state. Moreover, increasing the energy, which is only possible if the energy is added as a quantum at least of size (6.40), it is possible to push the particle out of the potential. Thus, the energy (6.40) is called the binding energy of the bound state (or the particle).

The situation of the δ-potential is prototypical for many problems: There is a negative potential, which admits a number of discrete energy levels, corresponding to localized bound states, which can be set free by providing the binding energy. Thereby they are moved into the continuum, i. e. continuous spectrum, and can move freely. However, the potential still affects their movement, leading to scattering.

6.5 Finite-well potential

It is worthwhile to study the situation once more with a more realistic potential, the so-called finite-well potential. This is essentially the situation of the infinite-well in section 6.1, but now with finite potential walls,

$$V(x) = (\theta(a-x) - \theta(-x-a))V_0.$$

This potential models the situation of an electron trapped in a solid or is an idealization of the Woods-Saxon potential describing nucleons inside a nucleus. Again, the sign of V_0 will determine the behavior, and the situation has a trivial time-dependence.

There are four possibilities. These correspond to the two signs of V_0 and then again how the energy relates to them. The situation with $E > V_0 > 0$ will be not much different from the corresponding situation of the potential step in section 6.3. The same is true for $E > 0 > V_0$. Thus, these both cases will not be treated explicitly, but a solution can be found in the same way as in section 6.3, showing similar behavior. The other two cases show much more interesting behavior.

Consider first $V_0 > E > 0$. Then the particle can move freely to the left and to the right of this potential barrier. But it can penetrate the barrier only exponentially damped. It is again worthwhile to consider the case that the wave is incident from the left. The other situations can again be reconstructed by superposition, if so desired.

A suitable static ansatz, suppressing overall normalization, is

$$\begin{aligned}\phi(x) &= \theta(-x-a)\left(e^{ikx}+Re^{-ikx}\right)+\left(\theta(a-x)-\theta(-x-a)\right)\left(Ae^{qx}+Be^{-qx}\right)+\theta(x-a)Te^{ikx}\\ E &= \frac{\hbar^2k^2}{2m}=\frac{\hbar^2q^2}{2m}+V_0.\end{aligned}$$

It was herein accounted for that one should reasonably expect that if a step reflects a normal wave, it should also be able to reflect an exponentially damped wave. However, this ansatz realized an important insight from section 6.4: It is expected that a particle can cross the classically forbidden region, and travel freely on the other side to infinity. Such a tunneling process is impossible classically, and a genuine quantum phenomenon.

However, for it to take place $T \neq 0$ and $j_T \neq 0$ are required. This requires to solve the continuity equations for the wave-function and its derivatives at the two interfaces, yielding four equations for the four unknowns A, B, R, and T, akin to the situation in section 6.3:

$$\begin{aligned}e^{-ika}+Re^{ika} &= Ae^{-qa}+Be^{qa}\\ Ae^{qa}+Be^{-qa} &= Te^{ika}\\ ik(e^{-ika}-Re^{ika}) &= q(Ae^{-qa}-Be^{qa})\\ q(Ae^{qa}-Be^{-qa}) &= ikTe^{ika}.\end{aligned}$$

This is a linear, inhomogeneous equation, which can be readily solved, though the resulting expressions are lengthy. Here, only T is of interest, which is found to be

$$T=\frac{4ikqe^{2a(q-ik)}}{(q+ik)^2-e^{4aq}(q-ik)^2},$$

which is well-behaved for $a > 0$. In complete analogy to (6.35) the transmitted current is now $\sim k|T|^2$, which is for $a > 0$ always well behaved and non-zero. Thus, indeed a particle

can penetrate through the barrier. It tunnels through it, a classically impossible process. Because

$$|T|^2 = \frac{8k^2q^2}{(k^2+q^2)^2\cosh(4aq) - q^4 - k^4 + 6k^2q^2}$$

this tunneling current decays exponentially with a and $\sqrt{E-V_0}$. This exponential suppression is typical for tunneling phenomena. Since q is an inverse length measured in units of the Compton wave-lengths of the particle, modified by the energy difference, this implies that the penetration depth[6] is roughly of the order of a few Compton wave-lengths. After that, the exponential suppression has reduced the effect to essentially negligible levels. This also shows why the effect is macroscopically usually irrelevant. Still, the effect plays an important role in microelectronics and nanotechnology, in nuclear and subnuclear decays, and some chemical and biological processes.

There are two more side remarks on this topic. One is the time it takes for the particle to tunnel. So far, this has been a static situation. In principle, this can also be done time-dependent, using, e. g., the wave-packets of section 5.1. In quantum mechanics, this turns out to be usually an instantaneous process, because information can be transmitted at arbitrary speeds. A fully relativistic treatment would reveal that it takes a finite time, which can actually be measured. The other is the non-zero probability density inside the barrier. The probability flux is again zero, in the same way as in (6.36). In addition, there is the problem of how to detect a particle, as this would require some kind of experiment to be placed there. But this ultimately requires that the particle can be there, and thus there would be needed some place with non-vanishing probability current, in contradiction to the before statement. Thus, the particle cannot be detected inside this object in the conventional way. This shows that the idea of having a particle moving through the barrier is probably not the right way to discuss tunneling, as this is too much guided by classical intuition, and the wave-function and probability currents are the more suitable concept. This will again be picked up in chapter 11.

The second interesting situation is if $V_0 < E \leq 0$. Based on the experience in section 6.4, it is expected that this could lead again to bound states. Furthermore, based on the results for the infinite well in section 6.1 and the symmetry of the problem an ansatz

$$\begin{aligned}\phi(x) &= (\theta(x+a)-\theta(a-x))\left(A\sin(qx)+B\cos(qx)\right)+C\left(\theta(a-x)e^{kx}+e^{i\phi}\theta(x-a)e^{-kx}\right)\\ E &= -\frac{\hbar^2k^2}{2m} = \frac{\hbar^2q^2}{2m} - V_0\end{aligned} \tag{6.41}$$

[6]Often the penetration depth is formally defined as the distance until the suppression is by one e-fold, i. e. $1/e$.

suggests itself. The boundary conditions then imply

$$
\begin{aligned}
Ce^{-ka} &= B\cos(qa) - A\sin(qa) && (6.42)\\
e^{i\phi}Ce^{-ka} &= B\cos(qa) + A\sin(qa) && (6.43)\\
kCe^{-ka} &= -q(B\sin(qa) + A\cos(qa)) && (6.44)\\
-ke^{i\phi}Ce^{-ka} &= q(A\cos(qa) - B\sin(qa). && (6.45)
\end{aligned}
$$

To have a physical situation the symmetry requires that the probability densities are mirror-symmetric to the origin, allowing only for a phase. Now, since both k and q are unique functions of E, these four equations are non-linear equations for the four unknowns A, B, C, and E. However, it is possible to replace $C' = C\exp(ka)$, to simplify the equations. Equations (6.42-6.43) then imply that only either A can be non-zero or B, and that in these cases ϕ must be either $2\pi n$ or πn, where conventionally n will be set to zero and 1, respectively.

These solutions have then the properties $\phi(x) = \phi(-x)$ and $\phi(x) = -\phi(-x)$, i. e. they are even or odd under mirroring, which is called even and odd parity under a parity transformation $x \to -x$. As required, the probability densities still satisfy $|\phi(x)|^2 = |\phi(-x)|^2$ in accordance with the symmetry of the problem.

Equations (6.44-6.45) are then a linear, homogeneous system for the constants B and C. For a non-trivial solution the determinant of the coefficient matrix has to vanish, which implies

$$k = q\tan(qa).$$

Likewise, for the odd case the condition

$$k = -\frac{q}{\tan(qa)}$$

arises. Unfortunately, neither equations can be solved analytically, and numerical techniques[7] are necessary. It turns out that the solutions are characterized by the quantity

$$l = -\frac{2mV_0a^2}{\hbar^2} > 0$$

which is obtained by replacing q with (6.41). It is found that for $l < (n\pi)^2$ the number of solutions is n. Thus, in contrast to the cases of sections 6.1 and 6.2, with a denumerable infinite number of bound states, and section 6.4, with one bound state, the number of bound states now depends on the details of the potential. Also, in contrast to section 6.4, there is always at least one solution, no matter how shallow the potential is. In fact, the

[7]The simplest 'numerical' technique is to plot both sides and look for crossings.

lowest state has then a energy larger than V_0, and is an even state. Especially, it turns out to have no nodes, and there is just a half-period $2/a$ fitting into the well. The next solution is then odd, and the one afterwards is even, and so on, always increasing the number of half-periods, and thus the number of nodes of the wave-function inside the box, by one. That the number of nodes increases with the level has already been encountered in section 6.1, and will turn out to be a general feature of such problems. This implies also that the alternation of states between even and odd parity, as far as the symmetry allows for it, is typical. This is really connected to the fact that a standing wave's energy increases with the number of oscillation periods.

Note that again there is a non-zero probability density for the particle to be outside the actual well. However, to actually find it the same applies as before to find it inside the potential wall. Furthermore, if V_0 is made large, it is found that for $E \ll V_0$ both the energy eigenvalues and the wave-functions behave more and more like the ones of the infinite-well case in section 6.1. This can be seen by explicitly expanding all equations in E/V_0.

6.6 Double-well potential

The tunneling phenomena can be made even more explicit in case of multiple wells. Consider a potential

$$\begin{aligned}
V(x < -l_-) &= V_0 > 0 \\
V(l_- < x < -l) &= 0 \\
V(-l < x < l) &= V_1 > 0 \\
V(l < x < l_+) &= 0 \\
V(l_+ < x) &= V_0 \\
-l_- < -l < \quad 0 \quad & < l < l_+.
\end{aligned}$$

Thus, there are two wells, separated by a potential barrier, a so-called double-well potential. An explicit solution is quite difficult to find, even in limiting cases. But a qualitative discussion is possible based on the insights of the previous sections.

Consider first the case $V_0 \to \infty$ and $V_1 \to \infty$. In this case, these are two infinite wells of section 6.1. If $l_+ = l_-$, the same wave-lengths are allowed, and the lowest states are just twice the same wave-function as before. But, it is also possible that the wave-function vanishes in either well. Thus, there are now three solutions, two with a vanishing wave-function in either well, and one with the same wave-function in both wells. Thus,

the system appears to become triply degenerate, i. e. there are three solutions for every energy. However, because the solution with a non-zero value in either well is just a linear superposition of the two other solutions, it is not linearly independent. Thus, actually there is only a double degeneracy. The origin of the degeneracy is the parity symmetry of the system.

If $l_+ \neq l_-$, this is no longer the case, as the quantization condition (6.6) can no longer be satisfied simultaneously. Then the wave-functions need to vanish in one well and be non-zero in the other, depending on the energies, and an interleaved spectrum appears, and the degeneracy of the previous case is broken explicitly. This is commensurate with the breaking of parity symmetry. This feature that symmetry and degeneracy are related is a generic feature in quantum systems. Its formal generalization is the Wigner-Eckart theorem, to be (briefly) discussed in section 8.2. It a recurring topic in quantum physics.

If $V_1 < \infty$ tunneling between both wells become possible, as the step in between acts like the potential step in section 6.5. However, because V_0 is still infinite, the situation is similar to the single infinite well of section 6.1 and the harmonic oscillator in section 6.2, and the energy levels will still be quantized. Especially, the continuity equations at the step will yield further conditions on the energy eigenvalues. Because of the generic properties of section 6.5 the lowest eigenstate will have no nodes, and the first a single node. That is, the first solution will be symmetric, and the next solution will be antisymmetric under parity, showing that the parity properties determine the energy levels. Again, there should now be solutions with the particle being mainly located in either of the wells, but any such solution will now have also a non-vanishing probability to find the particle in the other well. Parity breaking, as before, will again lift any degeneracies. Of course, if $E > V_1$ the particles can also move across the step classically, just as in section 6.5.

If $V_0 < \infty$, the situation will be two separate wells, and there will be bound states and scattering states as in section 6.5. More interesting is the situation with still $V_0 \to \infty$ but now, e. g., $l_+ \to \infty$. This corresponds to a hard wall at l_-, which can also quantum mechanically not be penetrated. On the other hand, the particle could move to infinity at large distances. If $E < V_1$, there is a bound state, but now with a tunneling possibility to large positive x, from where it is possible for the particle to escape to infinity. At the same time, there are states describing an incident particle on the step which could tunnel inside. In the static case, these will be solutions with a probability to find the particle inside and outside, at least in the sense of the probability currents of section 6.5. But consider now the case that the wave-function is constructed as a wave-packet in time, i. e. that there is a (discrete) Fourier-transform of the bound-state solution with some weight $f(E)$ with $f(E > V_1) = 0$. It is then possible to construct wave-functions which change in time

from having (essentially) all of their probability distribution inside and later outside. This describes how a particle can escape from a bound state to infinity. This is the prototypical situation for a decay process. Note that because E is uniquely given by the momentum, this is a superposition of momentum eigenstates. Thus, the particle has now a non-zero energy dispersion, and the energy level from which it originates can not be determined uniquely. This will depend on the details of $f(E)$, which will also determine the rate at which a particle can escape to infinity. Therefore, the energy levels of a bound state which can decay do not appear sharp, but have a width, the so-called natural line width.

It also seems to be strange that any such superposition has necessarily even in the infinite past a non-vanishing probability, even though it may be exponentially suppressed, for the particle to be outside. Thus, the bound-state cannot be considered to be stable, and is therefore called a resonance. This distinguishes it from the situation from the potential well of section 6.5, where the bound states could never escape to infinity, and were thus absolutely stable. It also shows that a decay is a probabilistic process, as at any time the particle could be already outside, though the probability increases with time. Usually, the decay time is the time where the probability of the particle to be observed outside has increased to 50% or, depending on conventions, some other values, like $1/e$. A natural observable example of this type are nuclear decays.

Chapter 7

Central-force problems and angular momentum

While the one-dimensional problems of chapter 6 illustrated qualitatively almost all new phenomena encountered in quantum physics, they are usually not describing actual physical situations. These are usually three-dimensional, though under some conditions they may be effectively lower-dimensional. There is also the possibility of lower-(or higher-)dimensional toy theories, or that actually there may be more dimensions in reality. At any rate, a treatment of higher-than-one-dimensional problems is mandatory.

7.1 Generalities in more than one dimension

Simply spoken, this will not alter the Schrödinger equation (5.5)

$$i\hbar\partial_t|a,t\rangle = H|a,t\rangle \tag{7.1}$$

conceptually, as a was standing for any kind of quantum numbers. Where this has an impact is if this is discussed in terms of wave-functions, which now depend on $\vec{x}$ instead of only a single component. Especially, the normalization of position eigenstates now takes the form

$$\langle\vec{x}|\vec{y}\rangle = \delta^d(\vec{x}-\vec{y}) = \Pi_i\delta(x_i - y_i),$$

and likewise for momentum eigenstates.

In position-space the Schrödinger equation therefore reads

$$i\hbar\partial_t\psi(\vec{x},t) = H\psi(\vec{x},t) = \left(-\frac{\hbar^2}{2m}\partial_i^2 + V(\vec{x})\right)\psi(\vec{x},t), \tag{7.2}$$

where in the second step the form for a purely position-dependent potential was used[1].

[1]Note that the Einstein convention is used throughout if nothing is stated otherwise.

In these seemingly harmless generalization a number of important issues have been hidden. The first is that (4.7) is actually valid in this form, and that position and momentum operators in different directions commute with themselves and with each other. This is non-trivial, and e. g. in non-commutative geometries this assumption is relaxed, allowing for different position components to be non-commutative. The second is that time-evolution is isotropic, as in classical mechanics. This implies that (7.2) is a scalar equation with respect to rotations. Also this is a postulate, borrowed from classical mechanics. This is also something which could be given-up, leading to the class of Galileo (or, more appropriately, Lorentz) symmetry-violating theories, or theories which involve an explicit aether. While no experimental indication exist that either postulate is incorrect, they remain postulates.

With (7.2), as it stands, little can be said in general. Especially, if $V(\vec{x})$ depends on the components separately, there are usually no general statements possible. However, in case of the free particle, $V(\vec{x}) = 0$ and for the, in practice paramount important, case of a central potential $V(\vec{x}) = V(|\vec{x}|)$ much can be discussed in general. This will lead to the concept of angular momentum, which plays a much more important conceptual role in quantum physics than in classical physics. This will also elucidate in a much better way what is behind the Stern-Gerlach experiment of chapter 2.

7.2 The free particle in two dimensions

While conceptually not different from three dimensions it is worthwhile to concentrate first on two dimensions, as it is technically much more simple. In the following the case of a free particle in two dimensions will be discussed, leading from (7.2) to

$$i\hbar\partial_t\psi(\vec{x},t) = -\frac{\hbar^2}{2m}\partial_i^2\psi(\vec{x},t).$$

Now, as before, a separation ansatz of type (5.11)

$$\psi(\vec{x},t) = e^{-i\frac{Et}{\hbar}}\phi(\vec{x})$$

leads to the static Schrödinger equation

$$E\phi(\vec{x}) = -\frac{\hbar^2}{2m}\partial_i^2\phi(\vec{x}) \tag{7.3}$$

which will now be solved. It turns out that it is in general quite interesting to discuss the problem in different coordinate systems, as they offer quite different insights. Also here, for technical simplicity it is best to start with two dimensions.

7.2.1 Cartesian basis

Since the directions do not communicate with each other, the first idea is to do a separation ansatz

$$\phi(\vec{x}) = e^{i\frac{\pm p_x x + \pm p_y y}{\hbar}} = \alpha(x)\beta(y). \tag{7.4}$$

Inserting this into (7.3) leads to

$$E\phi(\vec{x}) = \frac{\hbar^2(p_x^2 + p_y^2)}{2m}\phi(\vec{x})$$

and therefore

$$E = \frac{\hbar^2(p_x^2 + p_y^2)}{2m}.$$

Thus, in Cartesian coordinates the wave-function is just a product of one-dimensional plane waves, and the energy is the sum of the energies of both propagation directions.

7.2.2 Circle basis

It is a bit more interesting when using circular coordinates, i. e. radius and angle. Rewriting the Laplacian in (7.3) leads to

$$E\phi(\vec{x}) = -\frac{\hbar^2}{2m}\left(\frac{1}{r}\partial_r r\partial_r + \frac{1}{r^2}\partial_\omega^2\right)\phi(\vec{x}). \tag{7.5}$$

Making again a separation ansatz $\phi(\vec{r}) = \alpha(r)\beta(\omega)$ leads to

$$E\alpha(r)\beta(\omega) = -\frac{\hbar^2}{2m}\left(\frac{\beta(\omega)}{r}\partial_r r\partial_r\alpha(r) + \frac{\alpha(r)}{r^2}\partial_\omega^2\beta(\omega)\right).$$

Separating this in r and ω leads to

$$Er^2 + \frac{\hbar^2 r}{2m\alpha(r)}\partial_r r\partial_r\alpha(r) = -\frac{\hbar^2}{2m\beta(\omega)}\partial_\omega^2\beta(\omega).$$

This is fully separated. Since the right-hand side looks already suspiciously like a free harmonic oscillator, it is useful to set the common constant to $\hbar^2 l^2/2m$. Doing so yields

$$-l^2\beta(\omega) = \partial_\omega^2\beta(\omega)$$

with solutions

$$\beta(\omega) \sim e^{\pm il\omega}. \tag{7.6}$$

These are circular waves, with periods modulated by the factor l. Note that requiring singlevaluedness of the wave function requires l to be an integer. The sign does not matter,

as they are covered by the two independent solutions, so it is sufficient to require that l is a positive or zero integer. This condition will emerge again below for other reasons.

The radial equation takes the form

$$\left(\frac{2mE}{\hbar^2}r^2 - l^2 + r\partial_r r\partial_r\right)\alpha(r) = 0. \tag{7.7}$$

This equation is known as the Bessel equation. Just like with the Hermite polynomials of section 6.2.1 this can be solved by making a series ansatz. This works in essentially the same way, and will therefore not be detailed here, as nothing changes in the technical details. In fact, its three-dimensional equivalent will be considered later, as this case is both more involved and more instructive.

The solutions are

$$\begin{aligned}\alpha(r) &= \alpha_J J_l\left(\sqrt{\epsilon}r\right) + \alpha_Y Y_l\left(l, \sqrt{\epsilon}r\right)\\ \epsilon &= \frac{2mE}{\hbar^2}\end{aligned}$$

where J_l and Y_l are the Bessel functions of the first kind and second kind, respectively. However, only J_l can be finite at the origin, and thus only it is an admissible solution. Furthermore, only if l is a non-negative integer it is finite at $r = 0$, thus restricting the solutions to these cases. This is very consistent with (7.6). For the wave-function to have a unique value at every point in space and time, the so-called requirement of singlevaluedness, l can only be integer, as otherwise (7.6) would not be single-valued.

For this, the remaining Bessel function is given explicitly by

$$J_l(x) = \sum_{i=0}^{\infty}\frac{(-1)^i}{i!(i+l)!}\left(\frac{x}{2}\right)^{2i+l}.$$

This is an oscillating function which decays like $1/\sqrt{x}$ at large x.

Thus, the full solution is

$$\phi_l(\vec{r}) = \left(n_1 e^{il\phi} + n_2 e^{-il\phi}\right) J_l\left(\sqrt{\epsilon}r\right), \tag{7.8}$$

with the n_i normalization constants. This is quite different from the plane waves (7.4), especially as it decays towards large r. Also, instead of $\vec{p}$, it is now parametrized by E and a discrete number l. This seems to be strange at first. However, it turns out that both are equally valid, as they just form a different set of base vectors in the corresponding function space. Thus, it is just a basis transformation which connects both.

What is missing is some interpretation of l. It turns out that this is connected to the angular momentum. Using a canonical quantized angular momentum operator and the

case $n_2 = 0$ for simplicity yields

$$\begin{aligned}\langle XP_y - YP_x\rangle &= -i\hbar\int rd\omega dr\phi_l(\vec{r})^\dagger\left(x\partial_y - y\partial_x\right)\phi_l(\vec{r}) = -i\hbar\int rd\omega dr\phi_l^\dagger(\vec{r})i\hbar\partial_\omega\phi_l(\vec{r})\\ &= \hbar l\int rd\omega dr\phi_l^\dagger(\vec{r})\phi_l(\vec{r}) = \hbar l. \end{aligned}\tag{7.9}$$

In the second step the differential operator has been rewritten in circle coordinates. Thus, the expectation value of the planar angular momentum is $\hbar l$. Actually, it is the single component of the formally defined angular momentum operator. If $n_1 = 0$ instead, the result would have been $-\hbar l$. Since these correspond to forward and backward angular behavior of the wave-function, this is indeed consistent. This will be investigated in more detail in section 7.3.3.

This also explains the difference between (7.4) and (7.8). In (7.8) the states are also eigenstates of the angular momentum operator, as can be seen from the second-to-last step in (7.9). This is not the case for (7.4), as a direct calculation shows. Thus, only (7.8) has a definite angular momentum, and (7.4) should be considered as a superposition of all possible angular momentum states. Conversely, (7.4) are eigenstates of the linear momentum operator, which is not the case for (7.8), as again an explicit calculation shows.

Ultimately, of course, both are equivalent, as they are just different eigenbases of the same Hermitian operator H. Thus, between (7.4) and (7.8) lies only a change of basis. As usual, it will be the concrete problem which determines which is more useful.

7.2.3 Interference, the double slit experiment and wave-particle duality

To have an example where the circular basis (7.8) is actually simpler than the plane-wave one (7.4) consider the famous double-slit experiment. This is also instructive as it will exhibit again how quantum physics differs from classical physics.

The experiment is set up as follows. There is a particle source, a wall at $x = 0$ with only two point-like openings at $y = \pm r$, and a detector plate parallel to the wall some distance $x = s$ after the wall. The wall is taken to be perfectly absorptive, and thus anything can go only through the two openings. Approximate these openings to be point-like. Then the wave-function describing the particle's propagation can only originate from these sources. It will afterwards propagate free, and the detector will yield a signal proportional to $|\phi(s, y)|^2$, i. e. dependent on the probability amplitude.

Since the slits hide all what happens before, the wave-function will emanate in circular waves from the two slits. The total wave-function will then be determined by a superposition of (7.8), with arguments $\vec{x}\pm\vec{r}$, with $\vec{r} = (0, r)^T$. If the source is far away from the wall,

the particle has initially zero angular momentum, and thus $l = 0$, and the wave-functions are otherwise identical. This yields for the detector signal

$$\begin{aligned} |\phi(s,y)|^2 &= |n|^2 \left| J_0\left(\sqrt{\epsilon}\sqrt{s^2+(y-r)^2}\right) + J_0\left(\sqrt{\epsilon}\sqrt{s^2+(y+r)^2}\right)\right|^2 \\ &\overset{y\gg r}{\approx} |n|^2 \left| 2J_0\left(\sqrt{\epsilon}\sqrt{s^2+y^2}\right)\right|^2, \end{aligned} \tag{7.10}$$

where n is the normalization constant. Since J_0 is an oscillatory function, this implies a two-fold effect. First is that even at large distances there is an oscillatory pattern on the screen, showing interference between the two waves, just like ordinary water waves would do. Thus, particles show the wavy phenomena of interference, emphasizing the wave-particle duality and the concept of Compton wave-length of section 5.3. Indeed, if one were to take the limit of a Compton wave-length small compared to s and r, these oscillation would actually become more and more damped, until only a single maximum would remain, just like a classical experiment would be expected to act.

This is, however, not even the most astonishing consequence of quantum mechanics. If the source would be tuned down such that at some point only a single particle is emitted before it reaches the detector, the result would still be the same. Of course, the single particle would hit exactly one spot of the detector. But which spot would still be randomly selected according to the probability distribution (7.10). This is completely at odds with any classical ideas which sees a particle as an nonvolatile, inseparable object. But it is even experimentally possible to detect this phenomenon. Thus, a particle can interfere with itself, while still being detected as a single particle.

While at first looking like a curiosity, this will give rise to a fundamentally different way than the Schrödinger equation to think about, and treat, quantum physics in section 11.3.

7.3 The free particle in three dimensions

Of course, most physical systems have three dimensions. While in many cases quite analogue to the two-dimensional cases the concept of angular momentum becomes different. As it plays a central role both conceptually as well as technically, it deserves a more in-depth treatment than before in two dimensions. Again, most of the relevant conceptual progress can be made for the free particle, which will thus be the first step. The stationary Schrödinger equation in two dimensions (7.3) changes now only in that the Laplacian will be the three-dimensional one.

7.3.1 Cartesian basis

In the Cartesian basis there is little to add to section 7.2.1. A suitable plane-wave ansatz

$$\begin{aligned}\phi(\vec{x}) &= e^{i(\pm p_x x \pm p_y y \pm p_z z)} \\ E &= \frac{\hbar^2(p_x^2+p_y^2+p_z^2)}{2m}\end{aligned}$$

immediately solves the Schrödinger equation. Thus, in Cartesian coordinates the wave-function is again just a product of one-dimensional plane waves, and the energy is the sum of the energies of all three propagation directions. It is not hard to guess that this will also be true for an arbitrary number of dimensions, if desired or needed.

7.3.2 Spherical basis

In classical mechanics central force problems where most readily accessible in spherical coordinates. This will not change in quantum physics, as will be seen in section 7.4. Given the experience from section 7.2.2 immediately suggests that spherical coordinates will be connected to angular momentum. Also this is as would be expected from classical physics. However, it turns out that angular momentum works differently in quantum physics. This will be discussed in section 7.3.3. For now, only the solutions will be derived.

In three-dimensional spherical coordinates the analogue to the two-dimensional case (7.5) reads

$$E\phi(\vec{x}) = -\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\partial_r r^2 \partial_r + \frac{1}{r^2\sin\theta}\left(\partial_\theta \sin\theta\partial_\theta + \frac{1}{\sin\theta}\partial_\omega^2\right)\right)\phi(\vec{x}) \tag{7.11}$$

with the azimuthal angle ω and the polar angle θ.

Again, this can be immediately separated into a radial and angular part by a product ansatz

$$\frac{1}{\alpha(r)}\left(r^2\frac{2mE}{\hbar^2} + \partial_r r^2\partial_r\right)\alpha(r) = \frac{1}{\beta(\theta)\gamma(\omega)\sin\theta}\left(\partial_\theta\sin\theta\partial_\theta + \frac{1}{\sin\theta}\partial_\omega^2\right)\beta(\theta)\gamma(\omega). \tag{7.12}$$

By making the ansatz $\alpha(r) = \eta(r)/\sqrt{r}$ the radial part turns again into the Bessel equation (7.7) for $\eta(r)$, and thus leads ultimately to the same solution.

The angular part is more interesting. Selecting the constant to be κ leads to

$$\kappa\beta(\theta)\gamma(\omega)\sin\theta = \left(\partial_\theta\sin\theta\partial_\theta + \frac{1}{\sin\theta}\partial_\omega^2\right)\beta(\theta)\gamma(\omega).$$

It is once more visible that both angular parts can be separated. But because no explicit dependence on ω appears, the equation for $\gamma(\omega)$ will be the oscillator equation with solutions

$$\gamma(\omega) \sim e^{\pm i l_3 \omega}$$

with some constant l_3, a name chosen with hindsight[2]. Inserting this leads to an ordinary differential equation for $\beta(\theta)$,

$$\left(\sin\theta\partial_\theta \sin\theta\partial_\theta - l_3^2 - \kappa\sin^2\theta\right)\beta(\theta) = 0$$

Because

$$-\sin\theta\partial_\theta = \frac{\partial}{\partial\cos\theta}$$

it is suggestive to make the ansatz $\beta(\theta) = P(\cos\theta)$. Setting furthermore $x = \cos\theta$ the equation takes, after some more algebra, the form

$$(1-x^2)d_x^2 P(x) - 2xd_x P(x) - \left(\kappa + \frac{l_3^2}{1-x^2}\right)P(x) = 0. \tag{7.13}$$

This equation is known as the associated Legendre equation. This equation can actually not be solved by a series ansatz. However, it has still a somewhat regular structure.

Consider therefore for the moment the case $l_3 = 0$, which is known as the Legendre equation itself. This equation can actually be solved by a series ansatz, but a different approach is more useful. Note first that for $\kappa = 0$ $P(x) = 1$ would be a solution. Note further that this is a linear equation, which looks a lot like having factors of $1-x^2$ derived. Make therefore an ansatz

$$P_l(x) = \frac{1}{2^l l!}d_x^l(x^2-1)^l = \sum_k^{\lfloor l/2\rfloor}(-1)^k\frac{(2l-2k)!}{2^l(l-k)!(l-2k)!k!}x^{l-2k}, \tag{7.14}$$

where the second part comes from an evaluation of the binomial structure. The prefactor is here irrelevant, as the differential equation is homogeneous, and thus only there for later convenience. Inserting this ansatz into (7.13) yields

$$(1-x^2)d_x^{l+2}(x^2-1)^l - 2xd_x^{l+1}(x^2-1)^l - \kappa d_x^l(x^2-1)^l = 0$$

Consider first $l = 1$. This yields

$$-2x(2+\kappa) = 0.$$

This can only be fulfilled if $\kappa = -2$. Similarly, if $l = 2$

$$-4(6+\kappa)(3x^2-1) = 0$$

requiring $\kappa = -6$. Thus, it seems that for a fixed κ there can be only one P_l as a solution with a fixed value of l. It also already indicates that the value of κ cannot be arbitrary, and only particular fixed values can solve it. Inserting (7.14) finally yields that $P_l(x)$ is a

[2]In many texts it is called m, which is avoided here to prevent mixing it up with the mass.

solution if $\kappa = -l(l+1)$, with l zero or a positive integer. Again, while formally having other solutions, any other solutions are not well-behaved in a physical sense. Thus, it implies a quantization of κ.

Returning to the full equation (7.13), there is one important insight. The wave-function needs to have at every space-time point a unique solution. Thus, this implies that l_3 needs to be an integer, as otherwise it would not be periodic in ω. With this insight, it is found that the minor modification

$$\begin{aligned} P_l^{l_3 \geq 0}(x) &= (-1)^{l_3}(1-x^2)^{\frac{l_3}{2}} d_x^{l_3} P_l(x) \\ P_l^{l_3 < 0}(x) &= (-1)^{-l_3} \frac{(l+l_3)!}{(l-l_3)!} P_l^{-l_3} \end{aligned}$$

will actually solve (7.13). This can be shown by directly inserting it into (7.13). The product rule then reduces the equation to the $l_3 = 0$ case. There are two remarkable statements following from this. The first is that $P_l^{l_3}$ is no longer a polynomial, as the dependence on x becomes non-analytic. Given that x is originally $\cos\theta$, this implies that for $l_3 = 0$ the functions P_l only depends on $\cos\theta$, while for $l_3 \neq 0$ it also depends on $\sin\theta$, and therefore allows more involved oscillation patterns. The second is that if $|l_3| > l$ the number of appearing derivatives yields $P_l^{|l_3|>l} = 0$. Thus, the maximal absolute value of l_3 is bounded by l. Since there is no bound in l, every value can still be realized, though. More importantly, the behavior in θ and ω are thereby locked, and not every behavior in θ direction can be linked to every behavior in ω direction.

Combining everything together, the angular part are the so-called spherical harmonics

$$Y_{ll_3}(\theta,\omega) = \sqrt{\frac{2l+1}{4\pi}} \sqrt{\frac{(l-l_3)!}{(l+l_3)!}} P_l^{l_3}(\cos\theta) e^{il_3\omega}$$

where the normalization is chosen such that they form an orthonormal basis for functions in θ and ω in spherical coordinates

$$\int_{-1}^{+1} d\cos\theta \int_0^{2\pi} d\omega Y_{ll_3}^*(\theta,\omega) Y_{l'l'_3}(\theta,\omega) = \delta_{l_3 l'_3}\delta_{ll'}.$$

The orthogonality is already guaranteed by the fact that the angular part of the Laplacian is an hermitian operator. Because the eigenfunctions are also a decomposition of unity

$$\sum_{l=0}^{\infty} \sum_{l_3=-l}^{l} Y_{ll_3}^*(\theta,\omega) Y_{ll_3}(\theta',\omega') = \delta(\theta-\theta')\delta(\omega-\omega')$$

must also apply.

7.3.3 Angular momentum

In addition, section 7.2.2 suggests that l_3, and thus probably l, are connected to the angular momentum. The three-dimensional analogue to (7.9) requires the full angular momentum operator. In spherical coordinates it reads

$$\vec{X} \times \vec{P} = \vec{L} = -i\hbar \begin{pmatrix} \cos\omega \cot\theta \partial_\omega + \sin\omega \partial_\theta \\ -\sin\omega \cot\theta \partial_\omega + \cos\omega \partial_\theta \\ \partial_\omega \end{pmatrix}. \tag{7.15}$$

This implies directly

$$\langle L_3 \rangle = \hbar l_3.$$

Thus, l_3 is, up to a factor $\hbar$, the expectation value of the z-component of the angular momentum. It is useful to postpone the determination of L_1 and L_2 for a while.

Consider first the commutator of the L_i. Explicit evaluation, most easily in Cartesian coordinates, yields the angular momentum algebra

$$[L_i, L_j] = i\hbar\epsilon_{ijk} L_k \tag{7.16}$$

which is exactly the same as (3.3), as well as corresponding to the Poisson brackets of the angular momentum in classical mechanics. Thus, in the same way the total angular momentum is given by

$$L^2 = \sum_i L_i^2$$

and it commutes with all three components. Thus L^2 and one of the L_i, usually chosen to be L_3, form a complete set of commuting observables for the angular part. It is thus sufficient to know in addition to L_3 the eigenvalues and expectation value of L^2. This could be obtained by brute force evaluation. However, there is a more elegant way.

First note that (7.15) does not involve the radius. Thus, the radial part is irrelevant to angular momentum. The angular wave-function is uniquely determined by l and l_3, and thus described by a ket $|l, l_3\rangle$. Furthermore, explicit evaluation shows

$$\langle x|L_3|l, l_3\rangle = \hbar l_3 \langle x|l, l_3\rangle,$$

and thus $|l, l_3\rangle$ is an eigenket of L_3 with eigenvalue $\hbar l_3$. This and (7.16) is actually sufficient to determine the eigenvalue of L^2.

Next, define the operators

$$L_\pm = L_1 \pm iL_2.$$

This looks already suspiciously like the harmonic oscillator case in section 6.2.2. First of all note that while the L_i are hermitian operators $L_\pm$ satisfies $L_\pm^\dagger = L_\mp$. Next, note that

$$[L_3, L_\pm] = [L_3, L_1] \pm i[L_3, L_2] = i\hbar\,(L_2 \mp iL_1) = \hbar\,(\pm L_1 + iL_2) = \pm\hbar L_\pm$$

and

$$L^2 = L_\mp L_\pm + L_3^2 \pm \hbar L_3$$

by explicit calculation.

Now

$$L_3 L_\pm|l, l_3\rangle = L_\pm L_3|l, l_3\rangle + [L_3, L_\pm]|l, l_3\rangle = \hbar l_3 L_\pm|l, l_3\rangle \pm \hbar L_\pm|l, l_3\rangle = \hbar(l_3 \pm 1)L_\pm|l, l_3\rangle$$

and thus $L_\pm$ changes the value of l_3 by one. Consider now $l = l_3$. Then

$$L^2|l, l\rangle = (L_- L_+ + L_3^2 + \hbar L_3)|l, l\rangle = \hbar^2(l_3^2 + l_3)|l, l\rangle = \hbar^2 l(l+1)|l, l\rangle$$

Because $[L_\pm, L^2] = 0$ this implies that

$$L^2 L_-^m|l, l\rangle = \hbar^2 l(l+1)L_-^m|l, l\rangle.$$

Thus, the state $|l, l-m\rangle$ has the same eigenvalue for L^2, and thus the states $|l, l_3\rangle$ are $(2l+1)$-times degenerate eigenstates of L^2 to the eigenvalue $\hbar^2 l(l+1)$.

Thus, total angular momentum is also quantized, and most notably $l(l+1) > l^2$. Thus, it is never possible to have all of the angular momentum in only the z-component, and thus by rotational invariance never in a single component only. At the same time, it is not possible to measure more than one component precisely simultaneously. Finally, The number of possible values for the components of the angular momentum are limited, and cannot exceed $2l+1$, with l determined by the total angular momentum. Of course, because they are quantized in steps of $\hbar$, these values are tiny, at macroscopic values the distinction between two quanta of angular momentum becomes negligibly small.

The appearance of any central potential will not change the possibility to make the separation ansatz (7.12). Thus, the behavior of angular momentum is generic for central potentials. In all cases angular momentum will be quantized with quantum numbers l and (conventionally usually chosen) l_3, the latter being the eigenvalue of L_3.

7.4 The Coulomb potential

The arguably most important central potential is

$$V(r) = \frac{\gamma}{r} \tag{7.17}$$

which is shared, e. g., between the gravitational interaction and the Coulomb potential of a point charge. The paradigmatic example of the latter is an atom with a single electron[3], where $\gamma = -\epsilon e^2 Z$, where e is the electric charge of the electron and Ze the electric charge of the nucleus, and ϵ is a suitable constant of proportionality, depending on the system of units employed.

Because the potential is negative, it is expected that bound states with quantized energy levels will appear for $E < 0$ and pure scattering solutions for $E > 0$. The later correspond not to atoms, but electron-nucleus scattering. Though the latter is quite important in nuclear physics and particle physics, it will not be the focus of attention here. Rather, the bound states, which determine the chemical properties of an atom, will be the central issue.

Because of section 7.3 it is already clear that the angular behavior in spherical coordinates can be separated, and that angular momentum becomes quantized. For this it is useful to note that

$$\sum_i \partial_i^2 = \partial_r^2 + \frac{2}{r}\partial_r + \frac{1}{r^2}\left(\frac{1}{\sin\theta}\partial_\theta \sin\theta \partial_\theta + \frac{1}{\sin^2\theta}\partial_\omega^2\right) = \partial_r^2 + \frac{2}{r}\partial_r - \frac{L^2}{\hbar^2 r^2}$$

in analogy to the classical case and using (7.15). This has an important implication. For a central potential, the angular party acts like the angular momentum operator. Thus, after a separation ansatz, L^2 is replaced by $\hbar^2 l(l+1)$. Thus, the radial equation becomes

$$\left(-\frac{\hbar^2}{2m}\left(\partial_r^2 + \frac{2}{r}\partial_r\right) + \frac{\hbar^2 l(l+1)}{2mr^2} + \frac{-\epsilon e^2 Z}{r} - E\right)\alpha_l(r) = 0. \tag{7.18}$$

It is visible how the effective potential of classical mechanics arises again if $\hbar^2 l(l+1) \approx l_c^2$ for $l \gg 1$ and l_c being the classical angular momentum. Of course, a state will now depend on l. However, because the radial equation, which will determine E, does not depend on l_3, the energy will not depend on l_3, and thus every state of the hydrogen atom will be $(2l+1)$-times degenerate.

It is useful to replace $\alpha_l(r)$ by $u_l(r)/r$, as this simplifies the differential operator

$$\left(\partial_r^2 + \frac{2}{r}\partial_r\right)\alpha(r) = \frac{1}{r}\partial_r^2 u_l(r).$$

Furthermore, for $r \to 0$ the equation is dominated by the angular momentum, and reduces to

$$\left(\partial_r^2 - \frac{l(l+1)}{r^2}\right)u_l(r) = 0.$$

[3]Neglecting the finite size of the nucleus as well as that it also reacts to the electric field of the electron, and that this would actually be a two-body problem. Both effects are quite small, though experimentally accessible and have been measured. They can also be taken into account theoretically, except for the size of the nucleus, by using a center-of-mass language, just like in classical physics.

However, this equation is immediately solved by

$$u_l(r) \overset{r\to 0}{=} u_1 r^{l+1} + u_2 r^{-l}. \tag{7.19}$$

This shows already that the angular features also determine the radial behavior. The solution with negative exponent is irregular at the origin. Especially, because already

$$\partial_r^2 \frac{1}{r} = -4\pi\delta(r)$$

it is even for $l = 0$ formally not a solution to the Schrödinger equation.

It is furthermore useful to reexpress the system in suitable units. Note that

$$\begin{aligned} a_B &= \frac{\hbar^2}{\epsilon m e^2} \\ E_B &= \frac{\epsilon e^2}{a_B} \end{aligned}$$

have units of length and energy. They will be characteristic for the Coulomb problem, and are called Bohr length and Bohr energy . It is visible that they depend on the particle, and especially that the relevant length scales decrease with the mass and the energy scales increase with the mass. Thus, heavier particles will create smaller atoms and have larger binding energies. This has been observed by building, e. g., muonic or kaonic atoms[4], i. e. atoms where the electron is replaced by a muon or by a kaon. The former is a 200 times heavier sibling of the electron, while the latter a more involved hadron.

These quantities can be used to define dimensionless quantities

$$\begin{aligned} \rho &= \frac{r}{a_B} \\ \Omega^2 &= -2\frac{E}{E_B} \end{aligned}$$

in which the Schrödinger equation takes the form

$$\left(\partial_\rho^2 - \frac{l(l+1)}{\rho^2} + \frac{2Z}{\rho} - \Omega^2\right) u_l(\rho) = 0. \tag{7.20}$$

For $\rho \to \infty$ the two potential terms become irrelevant. Then only a decay equation is left. Since the exponentially increasing solution is unphysical, the long-distance behavior will be

$$u_l(\rho) \overset{\rho\to\infty}{=} u_3 e^{-\Omega\rho}. \tag{7.21}$$

[4]For such systems the finite extension of the nucleus starts to play a role. In fact, such system have actually been used to study precisely the properties of the nucleus.

Interestingly, the long-distance behavior is not qualitatively determined by the angular momentum, though the value of Ω could, and indeed will, depend on l. This is in contrast to the small-r behavior (7.19), which is qualitatively dominated by l. This is not so surprising, as at large distances the angular terms decays very quickly. In this sense, quantum physics remains local.

Putting (7.19) and (7.21) together yields an ansatz

$$u_l(\rho) = v_l(\rho)\rho^{l+1}e^{-\Omega\rho}.$$

Inserting this into (7.20) yields, after some algebra,

$$\left(\rho\partial_\rho^2 + (2l+2-2\Omega\rho)\partial_\rho + 2Z - 2\Omega(l+1)\right) v_l(\rho) = 0.$$

This has a very similar structure as the harmonic oscillator case in section 6.2.1. It is therefore suggestive to make a series ansatz. In fact, performing another change of variables to $x = 2\Omega\rho$, yielding

$$\left(x\partial_x^2 + (2l+2-x)\partial_x + \left(\frac{Z}{\Omega} - l - 1\right)\right) v_l(x) = 0 \tag{7.22}$$

makes this somewhat easier. This equation is known as Laguerre's equation, and its solution is tabulated. It is, however, instructive to solve it explicitly. The ansatz

$$v_l(x) = \sum_k a_k^l x^k$$

yields

$$\sum_k a_k^l \left(k(k-1)x^{k-1} + k(2l+2)x^{k-1} - kx^k + \left(\frac{Z}{\Omega} - l - 1)x^k\right)\right) = 0.$$

Shifting of the summation indices yields

$$\sum_k \left(a_{k+1}^l \left(k(k+1) + (k+1)(2l+2)\right) + a_k^l \left(\frac{Z}{\Omega} - l - 1 - k\right)\right) x^k = 0$$

and thus requires

$$a_{k+1}^l = \frac{k+l+1-\frac{Z}{\Omega}}{k(k+1)+(k+1)(2l+2)} a_k^l \tag{7.23}$$

for a solution. This recursion relation is comparatively simple.

In particular, at large k $a_{k+1}^l = a_k^l/(k+1)$, and thus a_k^l behaves like $1/k!$. This implies that at large k the series behaves like the exponential series. But then the solution will diverge exponentially for $x \to \infty$, which does no longer allow for a probability interpretation of the wave-function. Provided the Hamilton operator is an adequate description

of the atom, this cannot be allowed. The only alternative is that the series terminates at some k. Since the denominator is positive this can only happen if for some k

$$\frac{Z}{\Omega} - l - 1 = k \tag{7.24}$$

The solution is then a polynomial of order k, the so-called Laguerre polynomials. Especially, (7.23) is solved explicitly by

$$a_k^l = \binom{n + l + \frac{1}{2}}{n - k} \frac{(-1)^k}{k!}$$

as explicit insertion shows.

This has a drastic implication: Z/Ω needs to be an integer for this to be possible. Thus, the energy, which is encoded in Ω, is, as anticipated, quantized. To formulate this, it is quite common to define the maximum value of k as $n-l-1$, i. e. the Laguerre polynomials terminate for n at order $n-l-1$. Reinserting everything, the hydrogen energy levels then become

$$E = -\frac{e^2}{2a_B}\frac{Z}{n^2} \tag{7.25}$$

and n is called the main quantum number, in contrast to the angular quantum numbers l and l_3, as the energy only depends on n. For $Z = 1$ the lowest energy is the famous -13.6 eV. It is an interesting feature that for $n \to \infty$ $E \to 0-$. Thus, there is an infinite number of bound states, even if the energy difference between the lowest state and the highest state is finite.

Moreover, this implies that the maximal value of l is given by $n - 1$. This is not too surprising. After all, rotation implies rotational energy, and given a fixed amount of energy arbitrarily fast rotations should not be possible. Still, his implies that every energy level is

$$d_n = 1 + \sum_{l=1}^{n-1}(2l + 1) \tag{7.26}$$

times degenerate. Thus, the degeneracy level now depends on the energy.

Putting everything together, the Laguerre polynomials, now of order n, take the form

$$L_{n-l-1}^{2l+1}(x) = \sum_{k=0}^{n-l-1} \binom{n + l}{n - l - 1 - k} \frac{(-1)^k}{k!} x^k,$$

where the labels are chosen to coincide with standard notation The full eigenfunctions of the single-electron atom, after proper normalization, become

$$\Psi_{nll_3} = \frac{2}{n^2}\sqrt{\frac{Z^3}{a_B^3}}\sqrt{\frac{(n-1-1)!}{(n+l)!}}\left(\frac{2Zr}{na_B}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2Zr}{na_B}\right) e^{-\frac{Zr}{na_B}} Y_{ll_3}(\theta, \omega).$$

As noted in the beginning, the characteristic size is a_B/Z.

It is conventional to give names to different l values. Especially, $l = 0$ is called s, $l = 1$ p, $l = 2$ d, $l = 3$ f, and therefore, ignoring l_3, the energy levels are called $1s$, $2s$, $2p$, $3s$, $3p$, $3d$, and so on this. This is the so-called spectroscopic notation.

There are a few additional remarks to be made.

The first is that it thus seems that the lowest energy is $1s$, and thus all electrons of an atom should reside in this lowest level. However, electrons belong to a particular class of particles, so-called fermions. These particles have the special feature that it is not allowed that any two fermions have the same quantum numbers, the so-called Pauli principle. As these are n, l, and l_3, only a single one of them seems to be allowed in every level. In fact, electrons carry another quantity, spin, which does not enter at this level, and which has two possible values for electrons. Thus, in real atoms only two electrons can be in any given level, giving rise to the shell model of atoms. This immediately also explains the magic shell numbers of particular chemical stability. At these values, calculable by $2d_n$ to be 2, 8, and so on, it is necessary that the next electron has a higher value of n than the previous one, and thus a higher energy. Thus, any activity before this requires to lift an electron from a much lower energy level. Thus, more energy is required. Below these magic shells there is a space available, and electrons can shift around at the same n. Therefore, at the magic shells the jump in energy makes chemistry more complicated.

Why fermions have these properties is at the level of quantum mechanics an empirical fact. At the level of relativistic quantum physics it can be shown that besides fermions, another kind of particles exists, so-called bosons, which are not subject to the Pauli principle, and these two options exhaust all possibilities for a causal, relativistic quantum theory with a structure compatible with all experiments so far. This will be discussed in more detail in section 8.4, but many further details about this go substantially beyond the scope of this lecture, and are the providence of quantum field theory. Thus, the existence of fermions and the Pauli principle should be taken here as an empirical finding which needs to be included into the definition of the Hamilton operator.

The second issue is the degeneracy in angular momentum, which is actually also typical for central potential problems. It is not observed in nature, and the atomic energy levels show different values for different values of l, and even for l_3. The reason for these are manifold. However, they are related to the fact that (7.17) is an idealization. On the one hand relativistic effects will split levels with different values of the sum of the angular momentum and the spin of the electrons, the latter acting like an additional angular momentum, but of size $\hbar/2$. This will be discussed in more detail in section 8.4. The other comes from the fact, because of what has been said in section 5.3, that it is possible

to have additional quantum fluctuations. This, so-called, Lamb-shift will yield that all possible combination of l and spin get different corrections, and thus different energy levels. Finally, also the nucleus can have spin, including contributions of angular momentum of the nucleon in the nucleus. Interaction effects between it and the electron can finally lead to differences depending on l_3, the so-called hyperfine splitting. The latter can also be achieved by external magnetic and electric fields, the so-called Zeeman effect and Stark effect. Thus, the degeneracy comes from treating an idealization of the interactions to (7.17). Neglected are also effects due to the finite extent of the nucleus, especially if the nucleus is not spherical symmetric.

The third issue is that the results for (7.17) even quantum mechanically are only valid if there is only a single electron. If there are multiple electrons their electromagnetic repulsion can also alter the energy levels substantially, already for such a case as helium. Treating this effect is analytically also not possible, similar to the three-body problem of classical mechanics. However, using the perturbation theory of chapter 9, or other methods, allow to take this into account approximately.

Fourth, this immediately explains the atomic series, like the Lyman series or Balmer series. The difference in energies between two levels scale like $1/n^2 - 1/k^2$, and thus create the observed regularities. As the energy differences between such levels are, more or less, in the optical regime, at least for most values of Z and n, these differences are observable optically. This explains the classical observations of discrete lines in the light emitted by atoms, with such regularly spaced differences in the frequencies.

7.5 The Franck-Hertz experiment

Another confirmation of the atomic energy levels of section 7.4 is the Franck-Hertz experiment.

In this experiment an electron source creates a beam of electrons. This beam is then accelerated with a voltage. After passing a certain distance, the voltage is reversed. In the vacuum, depending on the size of the two voltages, the electrons will either reach or not reach a final detector.

Then, the path of the beam will be filled with an atomic gas, in the original experiment quicksilver. Classically, the electrons will loose energy by collisions with the gas atoms continuously. Thus, when increasing the acceleration voltage or the stopping voltage, the current will be continuously increased or decreased, depending on the density of the gas, the relative size of the voltages, and the thermal spread of the beam. The underlying physical process is a scattering of the atoms as a whole, like with balls.

In quantum mechanics the internal structure of the atoms, the shells of section 7.4, allow for another process. The electrons can loose energy by moving a bound electron of an atom to a different energy level. However, this mechanism will decrease the energy of the beam electrons not continuously, but by the energy difference between the original and final energy level of the bound electron.

Thus, the beam current will no longer change only continuously. Of course, there are still continuous changes when the collision is taking place not with a particular bound electron, but with the atom as a whole as a collective effect[5]. But as soon as the electron energy increases beyond the threshold of the difference of energies between two shells a discontinuous drop in the detected beam current occurs, as now a new loss mechanism becomes available. This is indeed observed.

Though a quantitative description requires much more details, this effect experimentally shows, like the line spectrum, that energy in atoms is quantized.

7.6 The harmonic oscillator

It is also possible to construct a central potential version of the harmonic oscillator of section 6.2. This is done by having three independent oscillators in the three directions, classically described by the potential

$$V(\vec{r}) = \sum_i \frac{m\omega_i^2}{2} x_i^2.$$

If the three frequencies are the same, $\omega_i = \omega$, this is called the isotropic oscillator. Only then it is a central potential.

This potential is an interesting quantity for two reasons. The first is that it is, in its isotropic formulation, equally well accessible in both Cartesian and spherical coordinates. It thus allows to conceptually highlight how the two formulations are connected. The second is that it is useful in many physical systems. E. g. it is a reasonable first approximation to nuclear interactions and describes very well harmonic traps in cold atom physics.

7.6.1 Cartesian formulation

In the Cartesian formulation it is rather straightforward to deal even with an anisotropic oscillator. Because the Hamilton operator in position space

$$H = \sum_i \left(-\frac{\hbar^2}{2m} \partial_i^2 + \frac{m\omega_i^2}{2} x_i^2 \right) \tag{7.27}$$

[5]Which is not yet treated.

does not mix the different directions, it can immediately be identified with a set of three independent oscillators of section 6.2.3. Thus, its eigenfunctions decompose as

$$\phi(\vec{r}) = \phi^h_{n_x}(x)\phi^h_{n_y}(y)\phi^h_{n_z}(z)$$

where ϕ^h_n are the eigenfunctions of the one-dimensional harmonic oscillator of section 6.2.1. The total energy is

$$E = \sum_i \hbar\omega_i \left(n_i + \frac{1}{2}\right). \tag{7.28}$$

Likewise, it could be solved in the operator language of section 6.2.2 by rewriting the Hamilton operator as

$$H = \sum_i \hbar\omega_i \left(a_i^\dagger a_i + \frac{1}{2}\right)$$

where the creation and annihilation operator $a_i^\dagger$ and a_i, respectively, are constructed and act only in one of the three directions. Thus, by all means, the system has independent directions.

7.6.2 Angular momentum formulation

Consider now the isotropic case and spherical coordinates. First of all, the angular part is not changed compared to the hydrogen atom or the free particle. Thus, the angular eigenfunctions will again be the spherical harmonics, see section 7.3.2, and the remaining radial Schrödinger equation can be read off from (7.18) to be

$$\left(-\frac{\hbar^2}{2m}\left(\partial_r^2 + \frac{2}{r}\partial_r\right) + \frac{\hbar^2 l(l+1)}{2mr^2} + \frac{m\omega^2}{2}r^2 - E\right)\alpha_l(r) = 0.$$

Defining

$$\begin{aligned} a_h &= \sqrt{\frac{\hbar}{m\omega}} \\ E_h &= \frac{E}{\hbar\omega} \\ \rho &= \frac{r}{a_h} \end{aligned} \tag{7.29}$$

the Schrödinger equation takes the form

$$\left(\partial_\rho^2 - \frac{l(l+1)}{\rho^2} - \rho^2 + 2E_h\right)u_l(\rho) = 0, \tag{7.30}$$

where u_l is again α_l/r. The situation is now very analogous to the Coulomb potential of section 7.4, except that a few numerical factors are different.

At short distances again only the angular momentum part contributes, and it does not matter that the potential is different than the Coulomb part. This, in itself, is an interesting observation. The qualitative behavior of two fundamentally different systems - after all the harmonic oscillator has only bound states while the hydrogen atoms has also scattering states - looks still identical at short distances. This is a first example of universality, the property of systems to show the same qualitative behavior in certain limits. Of course, quantitatively, e. g. the energy levels, are still different. After all, the values of E are already known and need to coincide with (7.28), and are therefore different from (7.25). But since the $2E_h$ term is also irrelevant at $\rho \to 0$ this does not change the result qualitatively.

Such features explain why some very different systems show qualitatively similar behavior in certain limits. This explains a lot of experimental observations. At the same time, this is technically useful, as this allows to reuse previous insights, and in some cases allow for the first foothold in very complex systems, if they reduce in certain limits to much more simpler (known) systems.

This said, the opposite limit $\rho \to \infty$, is fundamentally different, as here the involved potentials are very different - the free one in the hydrogen case and the harmonic oscillator one in the present case. This explicitly demonstrates that universality is not a feature which works outside the limits[6]. Fortunately, it is still a known system. Due to the $1/r$ rescaling of the eigenfunction this is just the one-dimensional harmonic oscillator potential of section 6.2.1. Thus, a suitable ansatz, taking the asymptotic behavior into account, is

$$u_l(\rho) = v_l(\rho)\rho^{l+1}e^{-\frac{\rho^2}{2}}.$$

Inserting this into (7.30) yields

$$\left(\rho\partial_\rho^2 + \left(l + \frac{3}{2} - \rho\right)\partial_\rho + \frac{E_h - l - \frac{3}{2}}{2}\right) v_l(\rho) = 0. \tag{7.31}$$

Comparison with (7.22) immediately shows that both equations are, up to some differing constants, identical. Thus, also (7.31) is solved by the Laguerre polynomials.

In particular, identifying constant by constant yields that the main quantum number n needs to be defined as

$$n = \frac{E_h - l - \frac{3}{2}}{2}$$

and thus implies again a quantization of the energy. Moreover, the energy (7.29) then reads

$$E_h = 2n + l + \frac{3}{2}$$

[6]Of course, and in fact, there can be now a third system in which the same behavior is observed, like the free system in case of the Coulomb potential.

and thus receives the same zero-point energy as (6.15), and has the same level spacing. Also, in this case the value of l influences the energy, in contrast to the hydrogen atom (7.25), and for every n every l is admissible. It is noteworthy that any solution with an odd number of energy levels has to have non-zero angular momentum, even if, like in the case of $n_x = n_y = n_z = 1$, the system is isotropic in the three Cartesian directions.

The full, properly normalized, solution is

$$\begin{aligned}
\phi_{nll_3} &= \sqrt{\frac{2n!}{(n+l+\frac{1}{2})!}}\frac{r^l}{a_h^{l+\frac{3}{2}}}L_n^{l+\frac{1}{2}}\left(\frac{r}{a_h}\right)e^{-\frac{r^2}{2a_h^2}}Y_{ll_3}(\theta,\omega) \\
\left(k+\frac{1}{2}\right)! &= \sqrt{\pi}\left(k+\frac{1}{2}\right)\left(k-\frac{1}{2}\right)...\left(\frac{1}{2}\right).
\end{aligned}$$

Interestingly, the counting is now different. There is a solution also for $n = l = 0$, which has just the vacuum energy, and is only a radial Gaussian. However, this agrees with the lowest level of the Cartesian solution, and is thus consistent.

Thus, very different looking solutions are obtained in Cartesian and spherical coordinates which, however, still describe the same physics, as it must be.

7.6.3 Coherent states

No matter the solutions of section 6.2 or of the current section, there is one surprising feature: The wave-functions have nodes, and here they can have nodes even for a whole spherical shell. At nodes, the probability density to find the particle vanishes. This is very at odds with what is expected for a classical particle in an harmonic oscillator. Moreover, The Gaussian always yields that the particle will be most probably found at the center of the oscillator. But this is precisely the point were classically it is least!

So, what does this mean? Well, this is very similar to the difference between the plane waves and the wave packets of the free particle in section 5.3. Plane waves were the eigenfunctions of the free Hamilton operator, but they were not suitable to describe particles. Rather, wave packets were necessary to describe a state which had the properties expected from a classical particle. Here, the results show that this is a generic feature of quantum systems. Classical physics emerges not only as a long-distance phenomena of expectation values of the Ehrenfest relations in section 5.7. It also necessary to use a superposition of states which exhibits suitable particle properties.

Such states are known as coherent states. The wave-packet (5.9) is an example of such a coherent state. It is useful to see how such states are constructed for the harmonic oscillator. This is the simplest such example beyond the free particle, but the same

guidelines reemerge in other cases. For this purpose, it is best to go back to the one-dimensional harmonic oscillator once more.

Thus, the general question is, if there are states, which are necessarily superpositions of eigenstates of the Hamilton operator, which exhibit the behavior of a classical particle in some sense, especially when it comes to the position of the particles. This is also important from an experimental point of view, as an initial condition is usually prepared with particles, and which has therefore to be described accordingly.

First, note that the ground state of the harmonic oscillator is already a Gaussian wave packet. It therefore already exhibits the desired features. The question is, whether this can be used. In the harmonic oscillator case there are three different operators available. The Hamilton operator, and the creation and annihilation operators. Since the Hamilton operator did not help, maybe the others would. In this context, it is interesting to note that

$$a|0\rangle = 0|0\rangle, \tag{7.32}$$

and thus the ground state can be actually considered to be an eigenstate to the annihilation operator to eigenvalue zero. The annihilation operator is not hermitian, and therefore its eigenstates are not a complete basis. But this should not stop one. After all, already in the free particle case the coherent states, the Gaussian wave packets, did not form a full basis. This is also intuitively clear, as there is no reason why classical physics should dictate the structure of the state space.

Generalize therefore (7.32) to

$$a|z\rangle = z|z\rangle = \frac{\zeta}{\sqrt{2}a_H}|z\rangle \tag{7.33}$$

where $z = \zeta/(a_H\sqrt{2})$, which for a non-hermitian operator is complex in general, labels as eigenvalues the eigenstates. The decomposition in a_H (defined as in (6.10)) and ζ is given for convenience in the following.

In coordinate space (7.33) reads

$$\left(\frac{x}{a_H} + a_H\partial_x\right)\psi(x) = \frac{\zeta}{a_H}\psi(x).$$

This first-order ordinary differential equation is the one of a shifted Gaussian, and thus immediately solved by

$$\psi(x) = Ne^{-\frac{(x-\zeta)^2}{2a_H^2}}, \tag{7.34}$$

with N a suitable normalization constant. However, it is important to keep in mind that ζ is still potentially complex, even though it has units of length. Taking this into account,

N is found to be

$$N = \sqrt{\frac{1}{\sqrt{\pi} a_H} e^{-\frac{(\Im\zeta)^2}{2a_H^2}}}$$

and thus is itself Gaussian. Physically, $\psi(x)$ are at $t = 0$ just wave-packets, or ground-state wave functions of the harmonic oscillator, centered at $\Re\zeta$, at fixed time. Thus, by construction, they behave at fixed time like a particle described by the wave packet of section 5.3. Note that ζ is not restricted, and thus there is a continuously infinite number of eigenstates, all corresponding to a product of a Gaussian wave-packet times a x-dependent phase.

This leaves to understand their behavior in the context of the harmonic oscillator. This is best done by expanding the coherent state into eigenstates of the harmonic oscillator. Start with the static case. Because the latter form a complete basis, this must be possible. Thus, if

$$|z\rangle = \sum_{n=0} b_n |n\rangle \tag{7.35}$$

then (7.33) leads to

$$\frac{\zeta}{\sqrt{2} a_H} |z\rangle = a|z\rangle = \sum_{n=0} b_n a|n\rangle = \sum_{n=1} b_n \sqrt{n} |n-1\rangle.$$

This requires

$$b_{n+1} = \frac{\zeta}{a_H \sqrt{2(n+1)}} b_n.$$

This recursion relation can be explicitly solved, taking the normalization condition into account, by

$$b_n = \frac{1}{\sqrt{n!}} \left(\frac{\zeta}{\sqrt{2} a_H} \right)^n e^{-\frac{|\zeta|^2}{4a_H^2}}. \tag{7.36}$$

As all b_n are different from zero for non-zero ζ this implies that a coherent state is a superposition of all eigenstates, though with factorial decreasing contributions for higher n once $\sqrt{n!} \gtrsim |\zeta|$.

Since (7.35) rewrites the coherent state as eigenstates of the harmonic oscillator it is now possible to use the time evolution operator and the general statements of section 5.4, and especially (5.11), to find the time dependence of a coherent state as

$$|z, t\rangle = e^{-\frac{|\zeta|^2}{4a_H^2}} e^{-\frac{i\omega t}{2}} \sum_n \frac{1}{\sqrt{n!}} \left(\frac{\zeta e^{-i\omega t}}{\sqrt{2} a_H} \right)^n |n\rangle$$

where it was used that $\exp(in\omega t) = (\exp(i\omega t))^n$. In comparison to (7.35) and (7.36) this amounts to an overall phase $\exp(-i\omega t/2)$ and replacing $\zeta \to \zeta e^{-i\omega t}$.

This can now be used to explicitly determine the properties of the coherent state, and especially if its properties behave classically. Because $P \sim a - a^\dagger$ and $X \sim a + a^\dagger$ according to (6.20-6.21) it can be used that the coherent state is an eigenstate of the annihilation operator. Thus

$$\langle P \rangle = -\frac{i\hbar}{\sqrt{2}a_H}\langle z|a - a^\dagger|z\rangle = -\frac{i\hbar}{\sqrt{2}a_H}\left(\langle z|\left(a|z\rangle\right) - \left(\langle z|a\right)|z\rangle\right) = -\frac{i\hbar}{2a_H^2}(\zeta - \zeta^*) = \frac{\hbar}{a_H^2}\Im\zeta. \tag{7.37}$$

Likewise, it follows that

$$\langle X \rangle = \Re\zeta. \tag{7.38}$$

Thus, the real and imaginary part of ζ characterize the expectation values for position and momentum.

As a side remark, it follows in a very similar way that

$$\begin{aligned}
\langle X^2 \rangle &= (\Re\zeta)^2 + a_H^2 \\
\langle (\Delta X)^2 \rangle &= \frac{a_H^2}{2} \\
\langle P^2 \rangle &= \frac{\hbar^2}{a_H^4}(\Im\zeta)^2 + \frac{\hbar^2}{2a_H^2} \\
\langle (\Delta P)^2 \rangle &= \frac{\hbar^2}{2a_H^2}
\end{aligned}$$

and thus

$$\langle (\Delta X)^2 \rangle \langle (\Delta P)^2 \rangle = \frac{\hbar^2}{4}.$$

Hence the coherent states actually fulfill the uncertainty relation as minimally as possible. Thus, the particle is located in phase space as stringently as quantum-mechanically possible.

Still, it needs to be seen if the coherent state has a 'classical' probability density. Because (7.37) and (7.38) are the expectation values also at time zero, ζ can be written in terms of the values of these expectation values at time zero, which can be cast as initial conditions x_0 and p_0,

$$\zeta = x_0 + \frac{ia_H^2 p_0}{\hbar}.$$

Finally, the full wave-functions can now be combined from this, the identification $\zeta \to \zeta e^{i\omega t}$, the form (7.34), and a multiplication with a phase $\exp(i\omega t/2)$. This requires some algebra. Setting for simplicity $p_0 = 0$, i. e. the particle is released from rest at some displacement x_0, the wave-function becomes

$$\psi(x,t) = \frac{1}{\sqrt{\sqrt{\pi}a_H}} e^{-\frac{(x - x_0\cos(\omega t))^2}{2a_H^2} - i\frac{x_0(x - x_0\cos(\omega t))\sin(\omega t)}{a_H^2}}$$

which leads to a probability density

$$|\psi(x,t)|^2 = \frac{1}{\sqrt{\pi}a_H} e^{-\frac{(x-x_0\cos(\omega t))^2}{a_H^2}}.$$

Thus, the particle is indeed Gaussian located along the classical path. As this has the minimum uncertainty, this is the state closest to a classical behavior, which can be constructed quantum-mechanically. The price to be paid was that it is a superposition of all energy states, though higher states contribute factorial less with increasing energy.

Chapter 8

Symmetries

Already in classical mechanics symmetries played a central role. There, they appeared as cyclic coordinates, simplifying calculations considerable. They were also associated with conserved quantities, integrals of motion. In fact, for continuous symmetries Noether's theorem guaranteed the existence of a conserved quantity.

Symmetries play an even larger role in quantum physics. As will be seen degeneracies are connected to symmetries. Symmetries will also dictate the vanishing of various overlaps, leading to forbidden processes. They also allow to separate kinematic features from genuine dynamical features, thus reducing problems considerably.

8.1 Symmetries and internal quantum numbers

Symmetries are foremost transformations which do not change the physics. Translations of section 4.2 and rotations are examples of such symmetries. Thus, under the application of a symmetry quantities like overlaps may not change. The value of symmetries is that this requirement will imply the existence of conserved quantities an the vanishing of certain overlaps.

However, this also means that a symmetry is necessarily described by a unitary operator S, $S^\dagger = S^{-1}$. If a symmetry is a feature of a system then it cannot change with time[1], and thus

$$[H, S] = 0. \tag{8.1}$$

The symmetry is a continous symmetry, if it can be parametrized by a continuous parameter ϵ, $S(\epsilon)$. Then, in complete analogy to section 4.2, this implies the existence of an

[1]Of course, it may be interesting to consider systems where the symmetry is time-dependent. That is possible, but will not be considered here. The symmetry is then, strictly speaking, not a part of the system, but only of the system during some time.

Hermitian operator $G = G^\dagger$

$$S = 1 - \frac{i\epsilon}{\hbar} G + \mathcal{O}(\epsilon^2) \tag{8.2}$$

if ϵ is small[2]. Inserting (8.2) into (8.1) implies

$$[H, G] = 0$$

and thus the observable G is independent of time, $G(t) = G(0)$. Moreover, this implies

$$\langle t|G|t\rangle = \left\langle 0 \left| e^{\frac{iHt}{\hbar}} G e^{-\frac{iHt}{\hbar}} \right| 0 \right\rangle = \langle 0|G|0\rangle.$$

and thus expectation values of G are conserved under time evolution, and thus are conserved quantities. Alternatively, consider some eigenstate of the operator G. Then

$$GU|g\rangle = UG|g\rangle = gU|g\rangle,$$

and thus any eigenstate of G is conserved under time-evolution with the time-evolution operator U.

This transfers Noether's theorem to quantum mechanics: For every continuous symmetry there is a conserved observable. This is why symmetries are so important in quantum physics. E. g. if the Hamilton operator is invariant under translations, momentum is conserved. If it is invariant under rotation, angular momentum is conserved.

8.2 The Wigner-Eckart theorem and degeneracies

These insights lead to an extremely important theorem: The Wigner-Eckart theorem.

Consider again a continuous symmetry $S(\epsilon)$ of a theory. Because

$$H(S|E\rangle) = SH|E\rangle = ES|E\rangle \tag{8.3}$$

the state $|E\rangle$ and $S|E\rangle$ have the same energy. If $|E\rangle \neq S|E\rangle$, perhaps up to a constant factor, i. e. the transformed state and the original state are not identical, there must exist a degeneracy in energy. An example of this is the central potential problem, and especially the hydrogen atom, of section 7, where the rotational symmetry and angular momentum, lead to the degeneracy in angular momentum quantum numbers.

[2]This dependency on a single parameter and the fact that $S(0) = 1$ can be generalized using the tools of group theory. This does not lead, at the current level, to fundamentally different insights, and is hence a unnecessary complication. However, at a deeper level, this provides a vast richness of new phenomena.

Assume now that the symmetry is continuous. Then there is the observable generator G with eigenvalues g. Thus, any eigenstate of the Hamilton operator for a fixed energy E can be written as

$$|E\rangle = \sum_i^{n(E)} |E, g_i(E)\rangle,$$

where the multiplicity $n(E)$ and the actual appearing values of eigenvalues $g_i(E)$ will depend on the energy in question. This was seen for the hydrogen atom, where the multiplicity was given by the main quantum number in (7.26), and the $g_i(E)$ where the corresponding l and l_3 values.

Now, because H and S commute and because of (8.3) it follows necessarily that

$$S|E, g_j\rangle = \sum_i^{n(E)} s_{ji}(E, \epsilon)|E, g_i(E)\rangle$$

because the energy has not been changed. Because the base kets are still orthonormal, the coefficients can be calculated

$$\langle E', g'_k|S|E, g_j\rangle = s_{jk}(E, \epsilon)\delta_{EE'}\delta_{gg'}$$

and are thus the matrix elements of the symmetry operator. Note, however, all matrix elements between different energy levels vanish.

Now let O be a linear operator which commutes with $S(\epsilon)$ for all ϵ, $[O, S(\epsilon)] = 0$. In this case, also all matrix elements of the commutator vanishes. Inserting a one yields

$$0 = \langle E, g_k| \, [O, S] \, |E', g'_j\rangle = s_{jk}\delta_{EE'}\delta_{gg'}\langle E, g_k|O|E', g'_j\rangle - s_{kj}\delta_{EE'}\delta_{gg'}\langle E, g_k|O|E', g'_j\rangle.$$

Now there are two possibilities, either E and/or g do not coincide with E' and g'. Or $E = E'$ and $g = g'$. But then this reduces to

$$s_{jk}\langle E, g_k|O|E, g_j\rangle = s_{kj}\langle E, g_k|O|E, g_j\rangle.$$

But since this must be true for any values of j and k, this can only be possible if $\langle E, g_k|O|E, g_j\rangle \sim \delta_{jk}$, and thus a unit matrix. Thus

$$\langle E, g_k|O|E, g_j\rangle = o(E, g)\delta_{jk}$$

and thus the expectation value of an invariant operator O cannot depend on the values of j and k, and thus is the same for any quantum number degenerate because of a symmetry. Operators, which are symmetry invariant do not change under a symmetry transformation. Also, this implies their eigenvalues needs to be degenerate with respect to the size of the subspace at fixed E and g.

Although this may seem trivial, this Wigner-Eckart theorem has very far-reaching consequences. E. g. it can be used to deduce the degeneracies of the hydrogen atom without solving the Schrödinger equation, just from the properties of rotational invariance. Also, it predicts the existence of the proton and the neutron based on the symmetries of the standard model of particles physics. These, and many more, statements show that the Wigner-Eckart theorem is, together with Noether's theorem, a defining property of physics. Much more of it will be heard in the lectures to come.

8.3 Discrete symmetries

Though continuous symmetries are therefore of central importance to quantum physics, also discrete symmetries, i. e. symmetries without a continuous parameter ϵ, play a fundamental role. In fact, they will be eventually related to fundamental concepts like causality or the possible spins of particles and the Pauli principle.

There are very many discrete symmetries, but most of them are special to particular systems. However, there are a few which appear in (almost) all, and they also act as role models to more specific cases. Therefore, these will be discussed here in turn.

8.3.1 Parity

The first of these is parity, which is defined to be a transformation which mirrors positions at the origin, classically just $\vec{x} \to -\vec{x}$, while time is not changed. In quantum physics, there must exist an operator Π which does this, and which acts on kets

$$|\alpha\rangle \to \Pi|\alpha\rangle.$$

As it should implement the classical idea of a space inversion, this requires that it should act on expectation values of X as

$$\langle\alpha|\Pi^\dagger X \Pi|\alpha\rangle = -\langle\alpha|X|\alpha\rangle.$$

Taking the same road as in the Heisenberg picture this implies

$$\Pi^\dagger X \Pi = -X \tag{8.4}$$

and thus parity indeed changes the sign of the X operator. Because a parity transformation should be a symmetry, $\Pi^\dagger = \Pi^{-1}$, and thus

$$\{X, \Pi\} = 0$$

and thus the position operator anticommutes with the parity transformation.

Selecting now specifically $|\alpha\rangle = |x\rangle$ (8.4) implies that

$$\Pi|x\rangle = e^{i\delta}|-x\rangle.$$

As absolute phases are irrelevant in quantum physics δ can be chosen at will, and is conventionally chosen to be $\delta = 0$. It follows then

$$\Pi^2|x\rangle = \Pi|-x\rangle = |x\rangle$$

and thus $\Pi^2 = 1$. This implies that

$$\Pi^\dagger = \Pi^\dagger\Pi^{-1}\Pi = \Pi^\dagger\Pi^\dagger\Pi = (\Pi\Pi)^\dagger\Pi = 1^\dagger\Pi = \Pi$$

and thus Π is also hermitian. This also implies that Π has only two eigenvalues, ± 1, as its eigenvalues are now necessarily real, while those of a unitary operator can be in general a complex phase. Hence, if a system is invariant under parity, there can be a twofold degeneracy, labeled by the observable eigenvalues of the parity operator. This also justifies the discussion on the double-well potential in section 6.6. However, because parity is not a continuous operator, Noether's theorem does not apply, and this must not necessarily happen[3].

This can be understood in more detail. Consider some wave function $\psi(x) = \langle x|\alpha\rangle$. Because

$$\langle -x|\alpha\rangle = \langle x|\Pi|\alpha\rangle = \psi(-x)$$

this implies that the wave-function of $\Pi|\alpha\rangle$ is the same wave function as that of $|\alpha\rangle$ evaluated at $-x$. Now assume that $|\alpha\rangle$ is an eigenket of Π. Then

$$\Pi|\alpha\rangle = \pm|\alpha\rangle$$

and thus

$$\psi(-x) = \langle x|\Pi|\alpha\rangle = \pm\langle x|\alpha\rangle = \pm\psi(x).$$

Thus, such a wave-function of an eigenket of parity changes sign under parity, but not its argument. Depending on whether the sign is $+1$ or -1 the wave function is said to be even or odd under parity. In the same way it can be shown that eigenkets of angular momentum behave as

$$\Pi|lm\rangle = (-1)^l|lm\rangle \tag{8.5}$$

[3]Examples are the single well of section 6.1 and the harmonic oscillator in section 6.2. A parity transformation does not lead to a linearly independent solution of the Schrödinger equation, because in both cases only one solution survived the physicality cut of normalizability. Thus, no double degeneracy appears because the non-trivial partner solution is unphysical.

by explicit calculation.

Consider now the cases where these are eigenstates of the Hamilton operator. Then $[\Pi, H] = 0$. If $H|n\rangle = E_n|n\rangle$ is a non-degenerate eigenket of H, it is also an eigenket of Π. Because of the non-degeneracy,

$$|n\rangle = \Pi\Pi|n\rangle = \Pi(\Pi|n\rangle), \tag{8.6}$$

and because $H\Pi|n\rangle = E_n\Pi|n\rangle$ $\Pi|n\rangle$ and $|n\rangle$ need to be the same ket. It is thus an eigenket of Π, and thus is either even or odd under parity. Note that while degenerate eigenstates can still be eigenstates of the parity operator, this is no longer guaranteed. E. g., because of (8.5) for $l > 0$ $a|ll\rangle + b|l - l\rangle$ will be an eigenstate of the Hamilton operator, but is not an eigenstate of the parity operator, while $|lm\rangle$ is. In fact, e. g. for the hydrogen atom $a|31l_3\rangle + b|32l_3'\rangle$ does not even have a definite behavior under parity for general values of a and b, and the question whether it is even or odd does not make sense. However, this seems to suggest that it is at least possible to construct an eigenbasis out of parity eigenstates. Again, this is sometimes possible, but again not guaranteed to be possible.

Just like (8.4) also other operators will change under a parity transformation, just as in classical physics also other quantities are affected. This has to be calculated for each other operator (or observable) for every case. However, it is instructive to see how this works in some examples.

Consider first the momentum operator. As it is defined in terms of the translation operator, which in turn is defined by the momentum eigenstates, it is possible to combine the results of section 4.2 with (8.4). This implies

$$T(-dx) = \Pi^\dagger T(dx)\Pi = \Pi^\dagger \left(1 - \frac{iPdx}{\hbar}\right)\Pi = \left(1 + \frac{iPdx}{\hbar}\right).$$

As everything else are just numbers this implies

$$\{\Pi, P\} = 0,$$

and thus the same as for the position operator. Of course, as all of this works in every direction in the same way, all of this also applies to three dimensions[4], for $\vec{X}$ and $\vec{P}$.

From this, it can be deduced for angular momentum, because of its decomposition into X and P as in section 7.3.3, that

$$[\Pi, \vec{L}] = 0$$

[4]Likewise, it would be possible to define a parity transformation acting only in a single direction. However, in practice there is little use for this.

and thus both operations commute. This implies that any angular momentum eigenstate will have a parity associated with it. It also implies that rotations and parity transformations commute. Classically, this does not surprise, as this creates the rotation group O(3) from the proper rotations SO(3) together with the parity group Z_2.

The parity of some system can have far-reaching consequences. Take a situation with two parity eigenstates $|\alpha\rangle$ and $\beta\rangle$. Let their parity eigenvalues be π_α and π_β. Then

$$\langle\beta|X|\alpha\rangle = \langle\beta|\Pi^\dagger\Pi X\Pi^\dagger\Pi|\alpha\rangle = -\pi_\alpha\pi_\beta\langle\beta|X|\alpha\rangle.$$

But this can only be true if either $\pi_\alpha\pi_\beta = -1$ or if $\langle\beta|X|\alpha\rangle = 0$. Thus, just because of parity properties, it can be deduced whether a matrix element is zero or non-zero. This is known as a selection rule: A selection rule makes a statement on whether a matrix element can be non-zero based on its symmetries. As will be seen in section 9.4 changes in a system can be described by matrix elements. Therefore such selection rules restrict possible physical processes.

This also leads to statements reminiscent of the Wigner-Eckart theorem of section 8.2. Consider a parity-invariant Hamilton operator. Then for its nondegenerate eigenstates follows

$$\langle n|X|n\rangle = \langle n|\Pi^\dagger\Pi X\Pi^\dagger\Pi|n\rangle = -\langle n|X|n\rangle.$$

Especially, as eX, with e the electric charge, is the observable describing an electric dipole moment, this implies than non-degenerate eigenstates of parity-invariant systems cannot carry an electric dipole moment. This is a very far-reaching statement for solids or elementary particles.

8.3.2 Time-reversal

The second archetypal discrete symmetry is time-reversal. The name is actually a historical name[5]. It could also be called movement reversal, because what is actually also happening is that $\vec{p} \to -\vec{p}$ without $\vec{x} \to -\vec{x}$, and its differs in this respect with parity. If $\vec{p} = d\vec{x}/dt$, this can be implemented by $(d)t \to -(d)t$, and this is how the name originates. Factually, it poses the question if the system is stopped at $t = 0$ and reversed, if it then just backtracks as it would move forward[6]. Classically, the question of time-reversal invariance can be

[5]Things become different once special or general relativity enters the game. This will not be covered here.

[6]This is what is usually not expected, as just reversing a movie shows. However, what in the macroscopical world is usually perceived as the absence of such time-reversal symmetry has actually to do with the initial conditions and evolution of a statistical system of many particles. This leads to an apparent deviation from time-reversal symmetry, though, in actuality, the initial conditions are at fault, which

better formulated as follows: If $\vec{x}(t)$ is a solution to the equation of motion, is then $\vec{x}(-t)$ also? If a potential exists, which is not explicitly time-dependent, the answer will always be yes.

However, the reason is that the classical equations of motion are, even when written in the Hamilton formalism, second order in time, and thus $(-dt)^2 = dt^2$. The Schrödinger equation is, however, of first order in time, and this may mean all the difference in the world. Especially, for

$$i\hbar\partial_t\psi(x,t) = \left(-\frac{\hbar^2}{2m}\partial_x^2 + V(x)\right)\psi(x,t)$$

$\psi(x,-t)$ will not be a solution, as the left-hand-side will change sign, but not the right-hand side. However, because

$$-i\hbar\partial_t\psi(x,t)^* = \left(-\frac{\hbar^2}{2m}\partial_x^2 + V(x)\right)\psi(x,t)^* \tag{8.7}$$

is also true, $\psi(x,-t)^*$ is a solution to the Schrödinger equation. Now, this is odd, as complex conjugation appears. On the other hand, this is a feature, which does not exist in classical mechanics, and may point the way to how to get the first-order nature of the Schrödinger equation into line with time-reversal symmetry. But this needs somehow to fit into the picture of what transformations are.

At this point a realization is necessary. Unitary operators have so far been natural as they do maintain all measurements because they do not change matrix elements. However, this may actually be too restrictive, as not really matrix elements play a role, but either expectation values of hermitian observables, which are real, or probability amplitudes, which are absolute squares. In all cases the information about reality is guaranteed. Thus, instead of requiring

$$\langle\alpha'|\beta'\rangle = \langle\alpha|\beta\rangle$$

it may be sufficient to alternatively also accept

$$\langle\alpha'|\beta'\rangle = \langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*,$$

i. e. it to allow for complex conjugation of matrix elements under a transformation. Of course, this cannot be achieved by a unitary transformation. Operators, which lead to such a behavior are said to be antiunitary. It can be shown that this exhausts the possibilities in quantum mechanics: Any transformation can be represented either by a unitary operator

violate the equilibrium condition, and the fact that nature is not in equilibrium. From this has to be distinguished that the best microscopical theory, the standard model of particle physics, does indeed violate this. However, the effect is extremely tiny, and requires sophisticated experiments to detect.

or an antiunitary operator. Such operators Θ can no longer be linear, but actually act antilinear,

$$\Theta\left(a|\alpha\rangle + b|\beta\rangle\right) = a^*\Theta|\alpha\rangle + b^*\Theta|\beta\rangle,$$

i. e. they act as complex conjugation on numbers.

It can now be shown, which is stated without proof, than any such operator Θ can be written as a product $\Theta = TC$, where T is a conventional unitary operator, and C acts as complex conjugation operator. The latter acts on an arbitrary ket, which can be expanded in some basis $|a\rangle$, as

$$C|\alpha\rangle = \sum \langle a|\alpha\rangle^* |a\rangle, \tag{8.8}$$

i. e., it does not alter the base kets, but complex conjugates all expansion coefficients. This, essentially, defines the action of C. Note that it is important here that $|a\rangle$ are base kets. If they are expanded in a different base $|b\rangle$, this implies

$$C|a\rangle = \sum \langle b|a\rangle^* |b\rangle.$$

Thus, while the base kets are unchanged, their expansion in different bases change under the action of C.

With this at hand, it is now possible to introduce the time-reversal operator Θ. Again, it should be emphasized that rather the defining property should be

$$\Theta|p\rangle = e^{i\delta}|-p\rangle, \tag{8.9}$$

i. e. is motion reversal, up to some phase. Considering an arbitrary state ket $|\alpha, t\rangle$, a system symmetric under motion-reversal needs to satisfy

$$|\alpha, \delta t\rangle = \Theta|\alpha, -\delta t\rangle$$

where it was chosen arbitrarily to work at $t = 0$ and δt infinitesimal. Since

$$|\alpha, \delta t\rangle = \left(1 - \frac{iH}{\hbar}\delta t\right)|\alpha, 0\rangle = \Theta\left(1 + \frac{iH}{\hbar}\delta t\right)|\alpha, 0\rangle \tag{8.10}$$

and $\Theta|\alpha, 0\rangle = |\alpha, 0\rangle$, this implies that

$$-iH\Theta = \Theta iH. \tag{8.11}$$

If Θ would be unitary then i would drop out, and Θ and H would anticommute, just like in the case of parity. But this cannot be, as then

$$H\Theta|E\rangle = -\Theta H|E\rangle = -E\Theta|E\rangle. \tag{8.12}$$

But this would mean that there is a negative-energy eigenstate for every positive energy eigenstate. As, e. g., the free-particle is certainly symmetric under motion reversal, this leads immediately to a contradiction. In fact, it would imply that not only the eigenvalues of P change sign, but also of P^2, which should be a positive operator. Thus, Θ cannot be unitary, but needs to be antiunitary, and thus the sign-change in (8.11) comes from the complex conjugation of i. Then, $[H, \Theta] = 0$, as expected from a symmetry of the theory, and there is no sign change in (8.12), restoring consistency.

Thus, Θ needs to be antiunitary. Of course, motion-reversal needs to be reversible again, returning the original value. Thus, an operator Θ^{-1} needs to exist. However, as an antiunitary operator this operator will in general not equal $\Theta^\dagger$. In fact, $\Theta^\dagger$ is not even well-defined, as no notion of complex conjugation of complex conjugation itself exists, and therefore will be avoided. However, it is still sensible to identify

$$\langle\alpha|A|\beta\rangle = \langle\alpha|\Theta^{-1}\Theta A\Theta^{-1}\Theta|\beta\rangle = \langle\Theta\alpha|\Theta A^\dagger\Theta^{-1}|\Theta\beta\rangle$$

which follows by inserting ones and using (8.8) repeatedly. The notion $|\Theta\alpha\rangle = \Theta|\alpha\rangle$ is used here to avoid a feeling 'as if' the time-reversal operator has been pulled to the bra by using hermitian conjugation, which was not the case. The formula is constructed by using (8.8), and acting to the right. This is quite subtle.

If the operator A is hermitian, $A^\dagger = A$, it is useful to define A to be even or odd under time-reversal if

$$\Theta A\Theta^{-1} = \pm A,$$

similar to the case of parity. If an operator has such a property, its matrix elements behave as

$$\langle\beta|A|\beta\rangle = \pm\langle\Theta\beta|A|\Theta\beta\rangle^*.$$

In particular, because of (8.9), it is necessary that P is odd under time-reversal, since only then

$$P\Theta|p\rangle = -\Theta P\Theta^{-1}\Theta|p\rangle = Pe^{i\delta}|-p\rangle = -pe^{i\delta}|-p\rangle = -p\Theta|p\rangle$$

follows, as is required by the definition of what time-reversal actually is. On the contrary, since $|x\rangle$ should not be modified, it can be shown in the same way that X needs to be even. Similarly, the properties of other operators can be deduced. E. g., orbital angular momentum is also odd, while, e. g., $X + \alpha P$ would not have a definite property.

Because the position operator is not affected by Θ and because of (8.8), it follows that

$$\Theta|\alpha\rangle = \int dx|x\rangle\langle x|\alpha\rangle^*$$

and thus the wave function of $\Theta|\alpha\rangle$ is $\psi(x)^*$, as anticipated from (8.7). Also, from this it can be shown by explicit computation, just as with parity,

$$\Theta|l,m\rangle = (-1)^m|l,-m\rangle.$$

Similarly as with parity, these and other features can be used to construct selection rules, or make powerful statements about degeneracies if $[\Theta, H] = 0$. However, because Θ is antiunitary this does not imply any conservation laws. After all, Heisenberg's equation, which was used to derive Noether's theorem, is based on the assumption that the symmetry operator is unitary, but Θ is antiunitary.

8.4 Spin

Before continuing on, it is worthwhile to note a further symmetry, which is spin. While this topic will be discussed in much more detail in the advanced quantum mechanics lectures, it is is for many reasons a feature to be noted already now.

Spin is a generalization of the orbital angular momentum of section 7.3.3. Thus, its hermitian operators S_i, which form a vector $\vec{S}$, obey the algebra

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k,$$

in analogy to (7.16). However, the generalized spin is not related to X and P, and cannot be reexpressed by it. It describes an additional quantum number, s, and accordingly a third component s_3. These are internal quantum numbers, in the sense that they are not build from X and P. They are intrinsic features of the particles described by a state. This is fundamentally different from the previous cases.

Therefore, the arguments of section 7.3.3 cannot be repeated to obtain the allowed eigenvalues for $\vec{S}^2$ and S_3, $\hbar^2 s(s+1)$ and $\hbar s_3$, respectively. At the level of quantum physics, in fact, the values cannot be restricted, rather they have to be measured. It is found that $s = n/2$, where n is a positive integer or zero. Still, because it follows from the algebra, s_3 will then vary between $-s$ and s in steps of $\hbar$. In the context of quantum field theory, i. e. when introducing special relativity, it can be derived that this is the case[7]. Thus, the range of values is different than for orbital angular momentum, which could have only integer multiplies of $\hbar$.

[7]There is an exception: In one space and one time dimension, s can be continuous. But since all observed particles exist in a world with three space dimensions, this does not apply to known particles. However, this behavior can be studied in toy theories, or appears in effectively one-dimensional systems as an emergent feature.

It is now again an observation in quantum mechanics, which can be derived in quantum field theory, that only particles having a half-integer spin are affected by the Pauli principle. As has been discussed in section 7.4, this has far-reaching consequences for the properties of atoms, and thus matter. In fact, that matter is stable and does not collapse can be traced back to this feature. However, in quantum mechanics, this has to be taken as another postulate and observation, and it will be treated as such.

It is not simple to make sense out of spin, as it is a genuine quantum phenomenon, and there is no corresponding quantity in classical physics. All angular momentum in classical physics is entirely orbital angular momentum. Most of the understanding of spin will therefore be relegated to the advanced quantum mechanics course.

One idea, however, which shows why spin is different, is the following. In classical mechanics angular momentum is connected to rotations. This is also true in quantum physics. As angular momentum is a Hermitian operator, it is possible to construct from it a unitary rotation operator in its infinitesimal version as

$$D(\vec{\alpha}) = 1 - i\frac{\vec{\alpha}\vec{L}}{\hbar}$$

where for now it will not be specified whether $\vec{L}$ is an orbital angular momentum or a spin. Just like for translation, it is necessary to construct a scalar to create a unitary operator, which invokes a parameter vector α. In contrast to, e. g., translations, finite rotations are not so simply created by exponentiation, as the L_i do not commute, in contrast to the P_i.

Though this can be done, consider for a moment the special case $\vec{\alpha} = \alpha\vec{e}_z$. Then only one operator is involved, and exponentiation is possible,

$$D(\alpha) = e^{-i\frac{\alpha L_3}{\hbar}}.$$

Acting on some angular momentum eigenstate yields

$$D(\alpha)|l, l_3\rangle = e^{-i\alpha l_3}|l, l_3\rangle.$$

First of all, if $\vec{L}$ is orbital angular momentum l_3 is an integer. Thus, if $\alpha = 2\pi$ this will just leave the state unchanged, to be expected for a rotation. This is as it ought to be. A full rotation should not change physics. To confirm that this is actually a rotation around the z-axis, note

$$\langle x|D(\alpha)|l, l_3\rangle = e^{-i\alpha l_3}\langle x|l, l_3\rangle = \sqrt{\frac{2l+1}{4\pi}}\sqrt{\frac{(l-l_3)!}{(l+l_3)!}}P_l^{l_3}(\cos\theta)e^{il_3(\omega+\alpha)}$$

and it thus indeed just shifts the angle ω by α. And thus $\alpha = 2\pi$ is a full rotation.

But this is no longer true for a half-integer spin. Now, without proof, the ket transforms to

$$D(\alpha)|s, s_3\rangle = e^{-i\alpha s_3}|s, s_3\rangle$$

and thus does change under a full rotation, albeit only by a pure phase, if s_3 is half-integer. This is not possible classically, and a genuine quantum effect. Only a rotation by 4π will return the original state. Of course, as it is just a phase, this will in general not be relevant, as absolute phases are not relevant. But as a relative phase between a rotated and an unrotated system this will create interference effects, which are typical for spin-half particles. They will not occur for integer spin particles.

Because of these distinct features particles with half-integer or integer spin received special names, and are called fermions and bosons, respectively[8]. The reason for their existence is deeply ingrained in the structure of (Minkowski) space-time, and can be naturally understood from there. But for here this is an observation.

[8]The special case of continuous spin in one space dimension and one time dimension discussed in footnote 7 is called anyon.

Chapter 9

Approximation methods

After having now an idea of the foundation of quantum physics, as well as its prevalent phenomena, it is time to dive again more into practicalities.

One of the very annoying features of most contemporary physics is that it is technically very complicated. Especially, the luxuries of chapters 6 and 7 that problems can actually be completely and exactly solved is an almost not happening luxury. To deal with this there are two common solutions.

One is the use of computer simulations. In principle, they are possible for any system, and their limiting factor is only the available computing resources. They limit the final accuracy of the results, as simulations can achieve with a finite amount of resources always only a finite accuracy, because a computer is digital, and thus discretization is necessary. Also, a computer is finite, and thus only finite system sizes are possible.

However, in practice there are some problems which require exponential amounts of resources, and are thus not feasible. On the other hand, even if the resources required are only polynomial, the achievable accuracy with real resources can still be too limited for the purpose of comparing to experiment.

Thus, while computer simulations are, in principle, able to solve all problems, they are not able to do so for all interesting cases in practice. Moreover, any simulation only produces numbers, not functions. The latter can at best be reconstructed by fitting from the numbers, which incurs another loss of accuracy. Thus, the insight which can be gained by simulations can be limited.

Performing computer simulations for quantum-mechanical systems needs suitable simulation algorithms, which require deep insights into numerics and also reasonably good knowledge of computers. Together, this is beyond the scope of this lecture, and therefore relegated to specialized other lectures.

An alternative to such simulations are analytical approximation methods. Many of

these exist, each with their own advantages and disadvantages, not only among themselves, but also in comparison to simulations. Thus, there is not 'the' approximation method. Hence, a reasonable toolbox of analytical approximation methods is needed for modern physics. To this end, some of the most commonly useful methods for quantum mechanics will be introduced here.

9.1 Time-independent perturbation theory

One of the time-honoured methods already in classical physics to solve problems is that of perturbation theory. The basic idea of perturbation theory is that a complicated system, which cannot be solved exactly, can be described by a small deviation from an exactly solvable system. Then, whatever parametrizes the 'small' can be used to set up a perturbative expansion, i. e. a kind of series representation, which allows to calculate the deviations systematically. This approach is called perturbation theory.

As the problems in quantum mechanics can in most cases be reduced, using the methods of section 5.4, to a stationary problem exactly, the main question is of course how to treat stationary problems perturbatively. This will be done here. Treating effects due to time evolution perturbatively is more involved, and will be discussed in section 9.4.

9.1.1 The non-degenerate case

9.1.1.1 Initial considerations

The appearance of degeneracies actually complicates perturbation theory substantially. Therefore, it is best to start first with a situation without degeneracies, to avoid obstructing a view on the underlying mechanisms.

The starting point is now the following. Given is the Hamilton operator of the system H. It is now necessary that it can be split as

$$H = H_0 + \lambda V.$$

Herein is H_0 an Hamilton operator of which the exact solution is known. The operator λV, with λ a real number, is the remainder, which makes the problem hard.

Now, an essential point is the choice of H_0. Of course, in principle any choice of H_0 is possible, by merely defining

$$\lambda V = H - H_0.$$

A possible choice is, e. g., always the free particle. However, to make perturbation theory efficient the choice of H_0 should yield something which is 'close' to the correct solution.

However, this cannot be guaranteed a-priori, as for this the solution should be known. Sometimes experiment can provide guidance in that observations seem to be close to some known H_0. More often, this is not the case. One possible, but not always successful, guidance is then the parameter λ. While any operator O can be arbitrarily split as $O = \lambda V$, the situation is most interesting if λ is a number much smaller than 1, while everything else in V is of order one[1]. As will be seen, perturbation theory, just as in classical physics, leads to expressions involving increasing powers of λ, which therefore become small prefactors if λ is less than one.

It is also often quite useful to think of λ as an additional parameter of H in the sense, that it can be tuned from zero up to its actual value. By this a continuous deformation of the Hamilton operator from the known one H_0 to the full one H is possible. This assumes that the limit $\lambda \to 0$ is continuous, and one should be wary that this is not always the case, though in quantum mechanics this is rarely encountered. This possibility also allows to treat Hamilton operators where no constant of the desired properties appears. Then a λ is introduced and the actual solution is recovered by taking the limit $\lambda \to 1$. Of course, this may spoil convergence, as 1 is of the same size as 1^n, and this possibility has to be treated with care.

By assumption, the properties of H_0 are known. Especially, the eigenproblem

$$H_0|n^{(0)}\rangle = E^{(0)}_{n^{(0)}}|n^{(0)}\rangle$$

is fully solved. Here, because a non-degenerate situation is assumed, the quantum numbers $n^{(0)}$ are in a one-to-one correspondence with the eigenvalues $E^{(0)}_{n^{(0)}}$. Separating both explicitly already now will be useful in section 9.1.2, when degeneracies will be admitted.

Now, the aim will be to solve

$$H|n\rangle = (H_0 + \lambda V)|n\rangle = E|n\rangle. \tag{9.1}$$

As noted the aim will not be to find an exact solution at once, but rather try to express the unknown quantities E and $|n\rangle$ in terms of the known quantities $E^{(0)}_{n^{(0)}}$ and $|n^{(0)}\rangle$.

9.1.1.2 Lessons from exact solutions

To understand how this can actually be possible, it is useful to first investigate a system where actually the eigenproblem of H can be exactly solved, such that an approach can be

[1]The size of an operator is a tricky problem. In principle, the size can be measured by the size of the eigenvalues, which, however, are not always known. Thus, if nothing is known about the operators making up V, any guess is as good as any other. In practice, quite often V is some combination of operators whose individual properties are known to some extent, giving some estimate.

developed which builds up the solution. While any of the cases of chapters 6 and 7 would be suitable, it is better to look at an even simpler problem first. Consider the Hamilton operator

$$H = H_0 + \lambda V = \begin{pmatrix} E_1^{(0)} & 0 \\ 0 & E_2^{(0)} \end{pmatrix} + \lambda \begin{pmatrix} 0 & V \\ V & 0 \end{pmatrix}$$

with V a constant and $E_1^{(0)} < E_2^{(0)}$. H_0 is the free case, which is already diagonal. This problem can be exactly solved, yielding as the full energies

$$E_{1,2} = \frac{E_1^{(0)} + E_2^{(0)}}{2} \pm \sqrt{\left(\frac{E_1^{(0)} - E_2^{(0)}}{2}\right)^2 + \lambda^2 |V|^2}. \tag{9.2}$$

To have a perturbative series in λ, this requires to recast (9.2) into a series in λ. Expanding the square root yields

$$E_{1,2} = E_{1,2}^{(0)} + \sum_{n=1}^{\infty} \binom{2n}{n} \frac{(-1)^{n+1}}{4^n(2n-1)} \left(\frac{\lambda^2|V|^2}{\pm E_1^{(0)} \mp E_2^{(0)}}\right)^n \tag{9.3}$$

$$\approx E_{1,2}^{(0)} \pm \frac{\lambda^2|V|^2}{\pm E_1^{(0)} \mp E_2^{(0)}} + \mathcal{O}\left(\left(\frac{\lambda^2|V|^2}{\pm E_1^{(0)} \mp E_2^{(0)}}\right)^2\right). \tag{9.4}$$

The last line gives the leading order perturbative correction to the H_0 solutions[2]. Beyond this leading-order corrections of power 4, 6, etc. in λ would be called next-to-leading order, next-to-next-to-leading order etc..

A few remarks have to be added. The first is that the expansion (9.3) for (9.2) is only meaningful if the square-root can be described by its Taylor series, i. e. if the argument is within its radius of convergence. The radius of convergence for the square root is given by

$$\frac{2\lambda^2|V|^2}{|E_1^{(0)} - E_2^{(0)}|} < 1.$$

That means that a perturbative calculation is not always possible, or at the very least not every H_0 admits for every λV a useful perturbative expansion. The corrections must, as noted above, be 'small' in a sense, which needs to be made precise in the sense of a radius of convergence. Likewise, the same quantity will also control if the approximation (9.4) is quantitatively good, or whether higher orders need to be included as well.

Another interesting observation is that in (9.4) the first corrections pushes the lower energy lower and the next energy higher. This will be found to be also quite generic.

[2]This leading corrections to the case of H_0 is also called tree-level, as perturbation theory can be cast into graph theory. This leading contribution then corresponds to a graph without loops, and thus a tree, while any higher order usually involves loops, and which are thus called loop graphs.

Especially, as only the absolute square of λV enters. This is also important in a different aspect. Consider the case that H_0 is a free particle, and V is now describing a square well, like in section 6.5. But the ground state of the free particle has energy zero, and thus generically a bound state would be created perturbatively even if the square well potential is repulsive. This is another generic feature obstructing perturbation theory: At threshold were bound states are created it becomes generically unreliable.

Even though these two observations may suggest that perturbation theory can easily be misleading in many practically cases this is not so. Thus, perturbation theory is a powerful tool, and a mainstay of modern quantum physics. These statement should be just a reminder to not apply perturbation theory blindly, and always make basic sanity checks, both mathematically (e. g. radius of convergence) and physically (e. g. appearance of bound states for repulsive potentials), before accepting the results.

9.1.1.3 Formulation

Now, after having obtained the desired result (9.3) starting from the known result (9.2), the next step is to construct (9.3) without knowing the exact solution. The aim will be to somehow construct the full solution constructively out of the known properties of H_0 as a power series in λ. To this end, define first

$$\Delta_n = E_n - E_n^{(0)} \tag{9.5}$$

the difference between the true energy eigenvalue and the energy eigenvalues of H_0. Here the $^{(0)}$ on the n can be dropped, as because of the non-degeneracy every energy level of H_0 and H need to be in an ordered one-to-one correspondence. It will still be kept in the states, though, to distinguish eigenstates of H and H_0, respectively. Then (9.1) can be recast as

$$(E_n^{(0)} - H_0)|n\rangle = (\lambda V - \Delta_n)|n\rangle, \tag{9.6}$$

and thus a known operator acts on the left-hand side. The operator $E_n^{(0)} - H_0$ is an hermitian operator. Unfortunately, hermiticity does not guarantee the existence of an inverse. In fact, in general this operator is not invertible, as it vanishes when acting on at least one eigenket of H_0. Thus, it is somehow necessary to avoid this.

Fortunately, one piece of information is known. The right-hand side of (9.6) is orthogonal to $\langle n^{(0)}|$, because the left-hand side of (9.6) is,

$$\langle n^{(0)}|(\lambda V - \Delta_n)|n\rangle = \langle n^{(0)}|(E_n^{(0)} - H_0)|n\rangle = \langle n^{(0)}|(E_n^{(0)} - E_n^{(0)})|n\rangle = 0. \tag{9.7}$$

Because the eigenvectors of H_0 form an orthonormal basis, the right-hand side will not be orthogonal in the complement of $|n^{(0)}\rangle$. It is here where the non-degeneracy plays a crucial role.

Define thus a projection operator to this complement,

$$\Phi_n = 1 - |n^{(0)}\rangle\langle n^{(0)}| = \sum_{k^{(0)}\neq n^{(0)}} |k^{(0)}\rangle\langle k^{(0)}|. \tag{9.8}$$

The inverse[3] of $E_n^{(0)} - H_0$ exists on this orthogonal complement. It can be constructed as

$$(E_n^{(0)} - H_0)^{-1}\Phi_n = \frac{1}{E_n^{(0)} - H_0}\Phi_n = \sum_{k^{(0)}\neq n^{(0)}} \frac{1}{E_n^{(0)} - E_k^{(0)}}|k^{(0)}\rangle\langle k^{(0)}|. \tag{9.9}$$

Note that writing the operator as an actual inverse is a common notation, but it is not a fraction. It is also not uncommon to write the combined operator as

$$\frac{\Phi_n}{E_n^{(0)} - H_0}$$

with the understanding that it acts to the right. This will not be done here. Since the inverse does not change the subspace into which Φ_n projects, another identity is also

$$\frac{1}{E_n^{(0)} - H_0}\Phi_n = \Phi_n \frac{1}{E_n^{(0)} - H_0}\Phi_n$$

However, having Φ_n to the left makes only sense if the operator is applied to a bra, and one should be wary there. That (9.9) is correct can be seen immediately as

$$\begin{aligned}
(E_n^{(0)} - H_0)(E_n^{(0)} - H_0)^{-1}\Phi_n &= \sum_{k^{(0)}\neq n^{(0)}} \frac{E_n^{(0)} - H_0}{E_n^{(0)} - E_k^{(0)}}|k^{(0)}\rangle\langle k^{(0)}| \\
&= \sum_{k^{(0)}\neq n^{(0)}} \frac{E_n^{(0)} - E_k^{(0)}}{E_n^{(0)} - E_k^{(0)}}|k^{(0)}\rangle\langle k^{(0)}| = \sum_{k^{(0)}\neq n^{(0)}} |k^{(0)}\rangle\langle k^{(0)}| \\
&= \Phi_n.
\end{aligned}$$

However, as noted above, this is only true because of Φ_n, and thus the result is Φ_n, rather than 1.

Because of (9.7) it follows that

$$(\lambda V - \Delta_n)|n\rangle = \Phi_n(\lambda V - \Delta_n)|n\rangle \tag{9.10}$$

as the difference in any matrix element with any state $\langle\alpha|$ is at most

$$\begin{aligned}
\langle\alpha|(\lambda V - \Delta_n)|n\rangle &= \langle\alpha|1(\lambda V - \Delta_n)|n\rangle = \langle\alpha|\left(\Phi_n + |n^{(0)}\rangle\langle n^{(0)}|\right)(\lambda V - \Delta_n)|n\rangle \\
&= \langle\alpha|\Phi_n(\lambda V - \Delta_n)|n\rangle + \langle\alpha|n^{(0)}\rangle\langle n^{(0)}|(\lambda V - \Delta_n)|n\rangle \\
&= \langle\alpha|\Phi_n(\lambda V - \Delta_n)|n\rangle + 0
\end{aligned}$$

[3]Sometimes also called resolvent.

and thus vanishing. As this is true for any state, this implies the identity (9.10).

Inserting now on the left-hand side of (9.6) a judicious 1 yields

$$(E_n^{(0)} - H_0)1|n\rangle = (E_n^{(0)} - H_0)(\Phi_n + |n^{(0)}\rangle\langle n^{(0)}|)|n\rangle = (E_n^{(0)} - H_0)\Phi_n|n\rangle + |n^{(0)}\rangle\langle n^{(0)}|n\rangle.$$

Pulling everything together yields

$$|n\rangle = |n^{(0)}\rangle\langle n^{(0)}|n\rangle + \frac{1}{E_n^{(0)} - H_0}\Phi_n(\lambda V - \Delta_n)|n\rangle. \tag{9.11}$$

That this actually solves (9.6) can be seen by explicit calculation,

$$\begin{aligned} & (E_0^{(0)} - H_0)\left(|n^{(0)}\rangle\langle n^{(0)}|n\rangle + \frac{1}{E_n^{(0)} - H_0}\Phi_n(\lambda V - \Delta_n)|n\rangle\right) \\ = \; & 0 + \Phi_n(\Lambda V - \Delta_n)|n\rangle = (\lambda V - \Delta_n)|n\rangle, \end{aligned}$$

thus showing that (9.11) indeed satisfies (9.6). As (9.6) is linear in $|n\rangle$, and so is (9.11), this actually does not say anything about the normalization of $|n\rangle$. The usually choice is to normalize $\langle n|n\rangle = 1$ afterwards. However, it will be technically convenient to rather normalize n such that

$$\langle n^{(0)}|n\rangle = 1, \tag{9.12}$$

removing the prefactor of the first term in (9.11). If needed, the result can still be normalized differently after it has been determined.

While (9.11) is formally correct, it is not yet very helpful. After all, this is now still an eigenproblem, though for a seemingly (and usually actually) much more complicated operator than the original Hamilton operator. To make progress it is necessary to somehow get rid of $|n\rangle$ on the right-hand side. For this, note that (9.7) implies together with (9.12)

$$\Delta_n = \lambda\langle n^{(0)}|V|n\rangle, \tag{9.13}$$

and thus Δ_n is actually a matrix element of V.

So far, however, everything are still exact statements. Now, assume that both $|n\rangle$ and Δ_n can be expanded in a power series in λ

$$|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + ... \tag{9.14}$$

$$\Delta_n = \lambda\Delta_n^{(1)} + \lambda^2\Delta_n^{(2)} + ... \tag{9.15}$$

That the expansion for Δ_n starts at one power higher than of $|n\rangle$ results from the relation (9.5), as this requires $\Delta_n = 0$ for $\lambda = 0$, and thus the constant term has to vanish.

Inserting (9.14-9.15) in (9.13), and using that (9.13) needs to be fulfilled at every order of λ individually, yields

$$\Delta_n^{(i)} = \langle n^{(0)}|V|n^{(i-1)}\rangle, \tag{9.16}$$

and thus the $\Delta_n^{(i)}$ can be determined from the solution of the $|n^{(j)}\rangle$ of one order less of the expansion in λ. Thus, finding $|n^{(i)}\rangle$ is the important task.

This can be done by inserting (9.14-9.15) into (9.11). Comparing the expression at different orders in λ is then an exercise in book keeping. First of all, consistently, to order λ^0

$$|n^{(0)}\rangle = |n^{(0)}\rangle$$

reemerges. This also shows how the normalization condition (9.12) was convenient, as otherwise also the expansion of the prefactor in (9.11) in the first term would have been necessary.

At order λ

$$|n^{(1)}\rangle = \frac{1}{E_n^{(0)} - H_0}\Phi_n V|n^{(0)}\rangle \tag{9.17}$$

$$\Delta_n^{(1)} = \langle n^{(0)}|V|n^{(0)}\rangle \tag{9.18}$$

emerges. This has now the desired structure. On the right hand-side only known quantities enter. However, this is not necessarily a simple calculation. As, according to (9.5), Φ_n is still an infinite sum. And because $|n^{(0)}\rangle$ is usually not an eigenstate of V, it may transform it into some rather complicated state. But this is a practical problem. The conceptual is completed. Especially, the energy shift to this order, $\Delta_n^{(1)}$ is just the expectation value of the potential V in the eigenstates of H_0.

Using (9.16) in (9.17) then determines already the second-order change of the energy levels,

$$\Delta_n^{(2)} = \left\langle n^{(0)} \left| V\frac{1}{E_n^{(0)} - H_0}\Phi_n V \right| n^{(0)} \right\rangle .$$

Thus, also the second-order energy shift is an expectation value in the eigenstates of H_0, though now for a more complicated operator.

This leads to the order λ^2 change in the states. This is generic, and actually not surprising. After all, H_0 is an hermitian operator, and thus all information must be encodeable by its eigenkets and operators. To find a generic structure it is useful to explicitly investigate what the appearance of (9.8) in (9.16), (9.17), and higher orders imply. Because (9.8) inserts either a one or projectors on eigenkets of H_0, it can actually be quite well decomposed. For this purpose, define the matrix elements

$$V_{nk} = \langle n^{(0)}|V|k^{(0)}\rangle. \tag{9.19}$$

Then

$$\begin{aligned}\Delta_n^{(1)} &= V_{nn}\\ \Delta_n^{(2)} &= \sum_{n\neq k}\left\langle n^{(0)}\left|V\frac{1}{E_n^{(0)}-H_0}\right|k^{(0)}\right\rangle\left\langle k^{(0)}\left|V\right|n^{(0)}\right\rangle = \sum_{n\neq k}\frac{|V_{nk}|^2}{E_n^{(0)}-E_k^{(0)}}\end{aligned} \tag{9.20}$$

Thus, the perturbative series is a series determined by the matrix element of the perturbation weighted by the energy difference of the kets in which the matrix elements are taken. Higher orders will emerge likewise. It is also characteristic that in each order the powers of the matrix elements V_{kn} increase. In the same manner

$$|n^{(1)}\rangle = \sum_{k\neq n}\frac{V_{kn}}{E_n^{(0)}-E_k^{(0)}}|k^{(0)}\rangle.$$

Thus, the state kets acquire contributions from all other eigenkets for which there is a non-vanishing overlap V_{nk}. Thus, indeed, they become a superposition with respect to H_0, justifying the caution exercised in inverting $E^{(0)}-H_0$. Note that because only λ is small but the V_{nk} do not need to be, there is no a-priori statement about whether only states close in n will contribute, even if the denominator becomes quickly small if n and k become widely separated.

At the next order, something interesting happens,

$$|n^{(2)}\rangle = \sum_{k\neq n,l\neq n}\frac{V_{kl}V_{ln}}{\left(E_n^{(0)}-E_k^{(0)}\right)\left(E_n^{(0)}-E_l^{(0)}\right)}|k^{(0)}\rangle - \sum_{k\neq n}\frac{V_{nn}V_{kn}}{\left(E_n^{(0)}-E_k^{(0)}\right)^2}|k^{(0)}\rangle. \tag{9.21}$$

Besides a direct quadratic term there is now also a term which has an element V_{nn} in it, but which is subtracted. Thus, the perturbative expansion yields a structure, which is graphically interpretable. The first term describes a movement from the kth to the lth level, and then on to the nth level. The second term is a movement directly from kth to the lth level, but in addition a so-called vacuum bubble of a movement of n to itself. Both combinations are then summed over all possible level movements. This graphical view is a first example of a so-called Feynman graph. Such graph-theoretical recastings are generic for such series techniques, and ubiquitous in quantum physics.

Higher orders are constructed in a very similar fashion, though they become increasingly annoying exercises in bookkeeping. The graph-theoretical approach is here much better suited to construct them, though the technology is straightforward.

To complete the perturbative determination requires still to normalize $|n\rangle$. This requires the multiplication of it by a suitable constant, a so-called wave-function renormalization constant $+\sqrt{Z_n}$,

$$|n\rangle \to Z_n^{\frac{1}{2}}|n\rangle$$

where the square-root is taken, as the norm should be one. This constant can also be determined in perturbation theory, as

$$|Z_n|\langle n|n\rangle = 1. \tag{9.22}$$

Because

$$|Z_n|^2 = \left|\left\langle n^{(0)}\left|Z_n^{\frac{1}{2}}\right|n\right\rangle\right|^2$$

the wave-function renormalization constant gives the probability that the perturbed state is the eigenstate of H_0.

Inserting into (9.22) (9.14) and solving for Z_n yields a perturbative series also for Z_n,

$$Z_n^{-1} = 1 + \lambda^2 \sum_{k\neq n} \frac{|V_{kn}|^2}{\left(E_n^{(0)} - E_k^{(0)}\right)^2} + \mathcal{O}(\lambda^3).$$

Because the perturbation is small, the expression can be inverted and expanded in a Taylor series to yield

$$Z_n = 1 - \lambda^2 \sum_{k\neq n} \frac{|V_{kn}|^2}{\left(E_n^{(0)} - E_k^{(0)}\right)^2} + \mathcal{O}(\lambda^3). \tag{9.23}$$

Hence the second term describes the amount of probability lost to other than the unperturbed state. Because this is a sum over positive-definite numbers, this implies that Z_n is one or smaller, as it must be to be a probability, which it is by construction. While it looks accidental here that

$$Z_n = \left.\frac{\partial E_n}{\partial E_n^{(0)}}\right|_{V_{nk}}$$

this is generally true, and stated here without proof.

9.1.1.4 Examples

This completes the construction of non-degenerate perturbation theory. It is time for some examples.

As before, it is best to first start with a case where the exact solution can also be determined, to see how it becomes build up from perturbation theory. To this end, consider the harmonic oscillator of section 6.2. Add to it a perturbation

$$V = \lambda \frac{1}{2} m\omega^2 X^2,$$

i. e. the perturbation is here by another oscillator potential of strength λ. The exact solution is given by

$$\omega \to \sqrt{1+\lambda}\omega,$$

i. e. by a rescaling of the frequency. Since $X^2 \sim (a + a^\dagger)^2$ the matrix elements V_{nk} will connect only states which differ from n up to ± 2, and the contributions at ± 1 drop out.

For concreteness, consider the change to the ground state, i. e. set $n = 0$. The relevant matrix elements are

$$\begin{aligned} V_{00} &= \frac{m\omega^2}{2}\langle 0|X^2|0\rangle = \frac{\hbar\omega}{4} \\ V_{20} &= \frac{m\omega^2}{2}\langle 2|X^2|0\rangle = \frac{\hbar\omega}{2\sqrt{2}}. \end{aligned}$$

If a state with $n \geq 2$ would be considered, there would be a third non-trivial matrix element $V_{n-2,n}$. This is one of the reasons why looking at the ground state simplifies the calculation.

This yields to leading order in λ

$$\Delta_0^{(1)} = V_{00} = \frac{\hbar\omega}{4} \tag{9.24}$$

$$|0^{(1)}\rangle = \sum_{k\neq 0} \frac{V_{k0}}{E_0^{(0)} - E_k^{(0)}} |k^{(0)}\rangle = \frac{1}{4\sqrt{2}}|2\rangle \tag{9.25}$$

$$Z_n^{(1)} = 0. \tag{9.26}$$

The wave-function normalization is still one, as, due to (9.24), it is intrinsically of order λ^2. That there is no contribution from the first excited state is because the perturbation does not change that the Hamilton operator is invariant under the parity symmetry of section 8.3.1, and thus also the perturbed states must respect this symmetry. That follows immediately as the selection rules of section 8.3.1 are actually the reason for the vanishing of odd V_{nk}, and thus enforce the corresponding symmetry order by order in the perturbative series. Thus, this amounts to an exact proof that such a symmetry is not broken. Conversely, this implies that a symmetry which is not broken by either H_0 nor by V will remain unbroken order by order in perturbation theory.

To compare to the exact result, note that the stationary wave-functions (6.18) depend on ω only through the characteristic length a_h (6.10). Thus

$$\begin{aligned} \frac{\hbar\omega\sqrt{1+\lambda}}{2} &= \frac{\hbar\omega}{2}\left(1 + \frac{\lambda}{2} + \mathcal{O}(\lambda^2)\right) \\ \psi_0(x) &= \frac{1}{\pi^{\frac{1}{4}}\sqrt{a_h}} e^{-\frac{x^2}{2a_h^2}} = \left(\frac{m\omega\sqrt{1+\lambda}}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{x^2 m\omega\sqrt{1+\lambda}}{2\hbar}} \\ &= \frac{1}{\pi^{\frac{1}{4}}\sqrt{a_h}} e^{-\frac{x^2}{2a_h}} + \lambda \frac{1}{\pi^{\frac{1}{4}}\sqrt{a_h}}\left(\frac{1}{8} - \frac{x^2}{4a_h^2}\right) e^{-\frac{x^2}{2a_h^2}} + \mathcal{O}(\lambda^2) \\ &= \psi_0(x) - \lambda\frac{1}{4\sqrt{2}}\psi_2(x) + \mathcal{O}(\lambda^2), \end{aligned}$$

and thus in agreement with (9.24-9.25), including the absence of change in the wave-function renormalization.

Next is to test this on a non-trivial example. Consider to this end once more the harmonic oscillator. Add as a perturbation a term $\alpha(a^\dagger a^2 + (a^\dagger)^2 a)$, which is Hermitian for α real. As has been seen, the decisive quantity are the matrix elements (9.19). They can be determined in the operator language of section 6.2.2.

$$\begin{aligned}\langle k|V|n\rangle &= \alpha\langle k|\left(\theta(n-1)(n-1)\sqrt{n}|n-1\rangle + (n+2)\sqrt{n+1}|n+1\rangle\right)\\ &= \alpha\left(\theta(n-1)(n-1)\sqrt{n}\delta_{k,n-1} + (n+2)\sqrt{n+1}\delta_{k,n+1}\right).\end{aligned}$$

The disturbance therefore only connects states with quantum numbers differing by ± 1

For simplicity, start with $n = 0$. Then to order λ

$$\begin{aligned}\Delta_0^{(1)} &= 0\\ |n^{(1)}\rangle &= -\frac{2\alpha}{\hbar\omega}|1\rangle\\ Z_0^{(1)} &= 0,\end{aligned}$$

i. e. the state gets some admixture of the first level, but no shift in energy. Note that this implies that the perturbation explicitly breaks parity. Note that Z_n is still one, because at this level any change in normalization is of order λ^2, and thus negligible.

At order λ^2

$$\begin{aligned}\Delta_0^{(2)} &= -\frac{4|\alpha|^2}{\hbar\omega}\\ |n^{(2)}\rangle &= |\alpha|^2\sum_{k\neq 0, l\neq 0}\frac{2\delta_{l,1}\left(\theta(k-1)(k-1)\sqrt{k}\delta_{k,l-1} + (k+2)\sqrt{k+1}\delta_{k,l+1}\right)}{\left(\frac{\hbar\omega}{2} - E_k^{(0)}\right)\left(\frac{\hbar\omega}{2} - E_l^{(0)}\right)}|k^{(0)}\rangle\\ &= -|\alpha|^2\sum_{k\neq 0}\frac{2\left(\theta(k-1)(k-1)\sqrt{k}\delta_{k,0} + (k+2)\sqrt{k+1}\delta_{k,2}\right)}{\hbar\omega\left(\frac{\hbar\omega}{2} - E_k^{(0)}\right)}|k^{(0)}\rangle\\ &= \frac{2\sqrt{3}|\alpha|^2}{(\hbar\omega)^2}|2\rangle\\ Z_n^{(2)} &= \frac{4|\alpha|^2}{(\hbar\omega)^2}\end{aligned}$$

and thus now both the energy and the wave-function renormalization start to decrease, while the ket gets now a contribution from the second excited state of the unperturbed oscillator. Thus, more kets start to factor in. Because the potential does not connect same states, the second term in (9.21) does not contribute. Note how the contribution are not

only in higher order in λ, but also higher order in $|\alpha|^2/(\hbar\omega)^2$. The factors of α are not surprising, as this is just a prefactor of the potential, and thus contributes like λ. The energy denominator is because of the even level spacing of the harmonic oscillator, and thus a feature of this particular system, and not generic.

9.1.2 Including degeneracies

9.1.2.1 Degenerate perturbation theory

The absence of degeneracies was important in the construction of Φ_n (9.8). If there would be degeneracies, it would no longer be possible to get away with removing just a single state. However, as chapter 7 has shown degeneracies are ubiquitous, especially in three dimensions. It is therefore necessary to take care of them. There is also a less obvious reason why special care is needed: Generically, there is no reason to expect that a degeneracy will not be broken by a perturbation. But then, it is necessary to uniquely identify in the limit $\lambda \to 0$ which of the degenerate energy levels is connected with which of the perturbed states. Therefore, it will be necessary to deal with the degenerate states explicitly, rather than just with the energy levels as in case of section 9.1.1. This implies that if the eigenkets of the unperturbed system are labeled by the energy E and eigenvalues n of another operator A the perturbed kets can no longer be eigenkets of A anymore, as otherwise the degeneracy would be kept. This implies $[V, A] \neq 0$. Though, of course, such a situation may occur in special cases (also called accidental), as well as that the degeneracy is not, or only partly, broken by the perturbation. All of this has to be taken care of by a general formalism.

The way to go is to enhance (9.8) in such a way as no singularities can arise. For this it is necessary to somehow subtract off all states, which could lead to a vanishing denominator. However, such a suitable basis may actually not be the basis into which the limit $\lambda \to 0$ evolves. This needs to be taken care of.

To make all of this explicit, assume that the degenerate kets into which the perturbed kets evolve for $\lambda \to 0$ are $|E, l\rangle^{(0)}$. If another basis $|E, m\rangle^{(0)}$ is more suitable for the following,

$$|E, l\rangle^{(0)} = \sum_m {}^{(0)}\langle E, m\, |E, l\rangle^{(0)} |E, m\rangle^{(0)}$$

will necessarily hold. Similarly to (9.8) define

$$\begin{aligned} P_0 &= \sum_m |E, m\rangle^{(0)}\, {}^{(0)}\langle E, m| \\ P_1 &= 1 - P_0 \end{aligned}$$

as projectors on the subspace belonging to the fixed energy $E^{(0)}$ and to its complement, respectively. Following again the same strategy as before to separate the two parts rewrite the Schrödinger equation as

$$0 = (E - H_0 - \lambda V)|E, l\rangle = (E - E^{(0)} - \lambda V)P_0|E, l\rangle + (E - H_0 - \lambda V)P_1|E, l\rangle \quad (9.27)$$

where it was used that P_0 is a projector on the unperturbed space, and therefore H_0 can be evaluated on it.

Because the eigenstates of the unperturbed Hamilton operator are an orthonormal basis $P_0P_1 = P_1P_0 = 0$. Because the projector projects on subspaces at fixed energy of H_0 it also follows that $[P_{0,1}, H_0] = 0$. This can be used to decompose (9.27) into two equations by acting with $P_{0/1}$ on it. This yields

$$(E - E^{(0)} - \lambda P_0 V)P_0|E, l\rangle - \lambda P_0 V P_1|E, l\rangle = 0 \quad (9.28)$$

$$-\lambda P_1 V P_0|E, l\rangle + (E - H_0 - \lambda P_1 V)P_1|E, l\rangle = 0. \quad (9.29)$$

Because P_1 projects out of the degenerate subspace, (9.29) is formally solved by

$$P_1|E, l\rangle = P_1 \frac{\lambda}{E - H_0 - \lambda P_1 V P_1} P_1 V P_0|E, l\rangle \quad (9.30)$$

Inserting the same kind of expansion (9.14) yields, to leading order in λ,

$$P_1|E, l\rangle^{(1)} = \sum_{k \neq \{l\}} \frac{V_{kl}}{E_l^{(0)} - E_k^{(0)}} |E, k\rangle^{(0)}, \quad (9.31)$$

where V_{kl} is defined as in (9.19), and $E_l^{(0)}$ is the (identical) energy for the degenerate subspace onto which P_0 projects.

Inserting (9.30) into (9.28) yields a formal solution of (9.28),

$$\left(E - E^{(0)} - \lambda P_0 V P_0 - \lambda^2 P_0 V P_1 \frac{1}{E - H_0 - \lambda V} P_1 V P_0\right) P_0|E, l\rangle = 0. \quad (9.32)$$

Expanding to order λ yields

$$(E - E^{(0)} - \lambda P_0 V P_0)P_0|E, l\rangle^{(0)} = 0 \quad (9.33)$$

Applying another state ${}^{(0)}\langle E, l|$ from the left, this turns into a homogeneous linear equation. But for such an equation to have a solution

$$\det\left(v - 1(E - E_l^{(0)})\right) = 0 \quad (9.34)$$

where $v =^{(0)} \langle E, m|V|E, m'\rangle^{(0)}$ is the matrix generated from $P_0 V P_0$ projected with $|E, l\rangle^{(0)}$ from both sides. Thus, the roots (9.34), or likewise the eigenvalues v_i of the hermitian

matrix v, determine the energy shifts $\Delta^{(1)}_{E,l} = E - E^{(0)}_l$ for each of the original degenerate states. Only if v is, after diagonalization, proportional to the unit matrix, the degeneracies will not be lifted by the perturbation. In fact, after diagonalization

$$\Delta^{(1)}_{E,l} =^{(0)} \langle E, l|V|E, l\rangle^{(0)},$$

but this would require to know the correct basis $|E, l\rangle^{(0)}$ beforehand. This also shows that to leading order the perturbation will only act inside the degenerate subspace and (generally) break the degeneracies. Only at higher order, due to (9.30) and (9.32) other degenerate subspaces will start to contribute to the shifts in energies. Note that is not the case for the kets, as (9.31) shows. This can be understood that the shift in energy is given by the matrix elements of the perturbation in the perturbed kets, and thus such a combined effect needs to be of order λ^2 to reach outside the degenerate subspace, and thus does not contribute to lowest order.

What is still missing is the effect on $P_0|E, l\rangle^{(0)}$, as (9.31) only yields effects outside the degenerate subspace. This is determined again by (9.32). Because there is a term which combines a λ^2 as a factor and a $1/\lambda$ because of the inverted operator, it is necessary to carefully write all terms first to order λ^2, even if finally only the terms at order λ are kept. This is a generic feature for such non-linear equations. Doing so yields

$$\left(E - E^{(0)} - \lambda P_0 V P_0 - \lambda^2 P_0 V P_1 \frac{1}{E^{(0)} - H_0} P_1 V P_0\right) P_0|E, l\rangle^{(0)} = 0.$$

Expanding again the kets according to (9.14), solving for $P_0|E, l\rangle^{(0)}$, and assuming that the degeneracy is completely broken, i. e. all $v_i \neq 0$, yields

$$P_0|E, l_i\rangle^{(1)} = \sum_{i\neq j} \frac{P_0|E, l_j\rangle^{(0)}}{v_j - v_i} \sum_{k\neq\{l\}}^{(0)} \langle E, l_j|V|E, k\rangle \frac{1}{E^{(0)}_l - E^{(0)}_k} \langle E, k|V|E, l_i\rangle^{(0)}, \tag{9.35}$$

and the sum of (9.31) and (9.35) gives the total change to an eigenstate in the degenerate subspace. Thus, (9.31) represents the amount that other kets spill into this subspace, which at this order is independent of the sublevel, because the breaking of degeneracy is itself a perturbative effect, as is the breaking of degeneracies inside the subspace. Conversely, the spilling in at higher order will then depend on the level inside the original subspace. If the degeneracy is not fully broken, things will again be a bit more involved, as then the unbroken subspaces need to be dealt with individually.

Concerning the normalization, using the same conditions as in the non-degenerate case yields, for the same reason, that at order λ the wave-function renormalization constant will not be changed. However, this then also implies that there is no change of overlap in

the equation (9.33) with the splitted states at second order in λ^2, and thus

$$\Delta^{(2)}_{E,l} = \sum_{k\neq\{l\}} \frac{|V_{kl}|^2}{E^{(0)}_l - E^{(0)}_k}$$

quite similar to the non-degenerate case (9.20). This is the statement that at first order the energy shifts are determined by the degeneracy-breaking effect, while at second order the spilling over determines, in a sense, an overall shift of the levels.

Note that this procedure needs to be repeated for every degenerate subspace. Thus, by and large, the procedure for degenerate perturbation theory is much more involved as before. Because the cross-talk between different degenerate subspaces becomes relevant at higher orders, the shifts in the kets becomes much more involved already at the next order in λ, and thus, in contrast to section 9.1.1, all formulas become quite involved, and the logistics becomes challenging. This will not be treated here, already due to space-time restrictions.

However, if the perturbation is already fully broken by the first order, there is a substantial simplification. Then the first order result can be regarded as a system H_0, and to obtain the second order the non-degenerate perturbation theory of section 9.1.1 can be applied as if it was applied at first order to a non-degenerate system. Because the degeneracy-broken kets form still an orthogonal base system at leading order, this will effectively require to not sum over the original degenerate subspaces.

As the procedure is technically involved, it is helpful, before going to an example, to repeat the pertinent points:

1) Select a degenerate subspace, determine the matrix v and solve its eigenproblem

2) The eigenvalues are the first-order shifts. The eigenvectors are the kets to which the perturbed kets will evolve in the limit $\lambda \to 0$, i. e. the needed basis $|E, l\rangle^{(0)}$

3) Construct the perturbed kets to first order using (9.31) and (9.35)

Now everything is set up for an example.

9.1.2.2 An example

Consider the hydrogen atom of section 7.4. A possible external perturbation can be done by an electric field. If the field is homogeneous of strength ϵ in z direction, it can be shown that the corresponding potential is given by $V = -z\epsilon$, and the electric charge e plays the role of λ. The $n = 1$ level is not degenerate, but the state $n = 2$ is. The admissible values for l are zero and one, and thus the state is four-fold degenerate. The matrix v is given by

$$v_{ll_3l'l'_3} = -\epsilon\delta_{ll'+1}\delta_{l_30}\delta_{l'_30}\langle 200|V|210\rangle$$

because the hydrogen eigenstates are parity invariant, and then the selection rules of section 8.3.1 require this structure. Evaluating the remaining matrix element yields $\langle 2l_3 0|V|2l_3' 0\rangle = 3a_B\epsilon$. Because it is real, the matrix is actually symmetric. However, the perturbation does not lift the degeneracy entirely. The energy levels for $l_3 \neq 0$ will thus not change. For the altered states at $l_3 = 0$, the eigenvalue calculation yields

$$\begin{aligned} E_{1,2} &= E_{2l0}^{(0)} \pm 3ea_B\epsilon \\ E_{3,4} &= E_{2l\neq 0}^{(0)} \end{aligned}$$

and the levels to which the system evolves without perturbations are

$$\begin{aligned} |2\{1,2\}\rangle^{(0)} &= \frac{1}{\sqrt{2}}\left(|200\rangle \pm |210\rangle\right) \\ |2\{3,4\}\rangle^{(0)} &= |20 \pm 1\rangle. \end{aligned}$$

This is known as the linear Stark effect. There are many more cases, of similar structure, which play an important role in the spectroscopy of the hydrogen atom.

9.2 Variational approach

9.2.1 Setup

The perturbative approach is based on the idea that there is a small quantity, and the kets and eigenvalues depend analytically on them, i. e. they can be expanded into a Taylor series. As discussed above, the latter is not necessarily the case. Also, not every system offers a small quantity on which a perturbative approach can be based on. Thus, in this section and the next section 9.3 alternative approximate treatments will be discussed. Of course, there are more.

The idea of the variational approach is based on the observation that the full Hamilton operator and any (hermitian) part of it act in the same Hilbert space. Thus, either set of eigenkets are a full basis. It is thus necessarily possible to express any eigenket of the full Hamilton operator in terms of some other (part of) the Hamilton operator H_0. In fact, as long as it acts in the same Hilbert space, any basis can provide this. Thus, it is not even necessary that some H_0 has be solved exactly for this.

The variational approach uses this insight. Its basic mechanism is to use some (arbitrary) basis to build up systematically the eigenkets of an Hamilton operator H by adding successively more and more kets to a partial basis. Eventually, in the limit of including all kets, this will necessarily be exact. In contrast to perturbation theory, this therefore

does not need that there is some quantity on which the true solution depends analytically. The disadvantage is that for any subset of base vectors the found solution may be arbitrarily far away from the true solution. Thus, a bad choice of basis may render the method practically useless. There is, unfortunately, in many cases no possibility to systematically find a good basis, without solving the problem. Thus, the choice is usually based on an educated guess, and, more importantly, experience and, if available, experimental input. Fortunately, many systems can be tackled satisfactorily with experience and physical insight.

Also, there are a few powerful statements helping in the process.

The first, and arguably most powerful, is that if H is some Hamilton operator and $|0\rangle$ is such a first ket from the used basis, then

$$e_0 = \frac{\langle 0|H|0\rangle}{\langle 0|0\rangle} \geq E_0 \tag{9.36}$$

i. e. for any ket this expectation value is larger than the lowest energy of the Hamilton operator H. Thus, even with an arbitrarily bad choice of the first trial ket already a valid upper bound to the ground state energy is obtained. Likewise, when trying some set $\{|i\rangle\}$, it is guaranteed that $\min_i e_i \geq E_0$, allowing for reducing the upper bound by enlarging the basis. This also implies that the $|i\rangle$ with e_i lowest is the best approximation of the actual ground-state $|E_0\rangle$ of H among the $\{|i\rangle\}$.

This can be seen as follow. If $\{|E_i\rangle\}$ is the actual (unknown) eigenbasis of H (and any degeneracies are contained in the index) then

$$|0\rangle = \sum_i |E_i\rangle\langle E_i|0\rangle$$

necessarily holds. Rewrite the solution E_i of $H|E_i\rangle = E_i|E_i\rangle$ as $E_i = E_i - E_0 + E_0$ then

$$e_0 = \frac{\sum_i \langle 0|H|E_i\rangle\langle E_i|0\rangle}{\sum_j \langle 0|E_j\rangle\langle E_j|0\rangle} = \frac{\sum_i |\langle E_i|0\rangle|^2 E_i}{\sum_j |\langle E_j|0\rangle|^2} = \frac{\sum_{i>0} |\langle E_i|0\rangle|^2 (E_i - E_0)}{\sum_j |\langle E_j|0\rangle|^2} + E_0 \geq E_0,$$

where it was used that $E_{i>0} \geq E_0$ by construction. Note that only if $|0\rangle = |E_0\rangle$ can the equality hold. Conversely, this requires $\langle E_{i>0}|0\rangle$ to vanish. However, this is actually better than it looks. If $\langle E_{i>0}|0\rangle = \mathcal{O}(\epsilon)$, and ϵ is small, then the denominator behaves like $1+\epsilon^2$ while the nominator behaves like ϵ^2. Thus, $e_0 - E_0$ is of order $\mathcal{O}(\epsilon^2)$, and thus can already be pretty good. Of course, this cannot be checked, as this would require to know $|E_i\rangle$, making the variational analysis itself superfluous. Alternatively, this could be formalized by replacing $|0\rangle$ by $|E_0\rangle + \delta|0\rangle$, where δ is measuring the deviation of the ket from the true solution.

Unfortunately, this is still not sufficient to help in building a good choice of trial kets $\{|i\rangle\}$. But here again helps that (9.36) is formulating a minimization condition. For the true ground state this will reach a minimum. Thus, when parameterizing the trial ket by $|0, \lambda_k\rangle$ where the set $\{\lambda_k\}$ are parameters of the ket, then necessarily the optimal choice is

$$\partial_{\lambda_k} e_0 = 0$$

for all k. Because this minimizes e_0, and thus satisfies (9.36) as best as possible. Moreover, the second derivatives (and higher, if need be), can be used to assure that it is indeed a minimum.

If also higher state kets are searched for, the approach is to construct projectors, like (9.8), to create trial kets orthogonal to the best approximation of the ground-state ket. Afterwards, on this reduced Hilbert space, the same minimization procedure is applied. However, a bound like (9.36) is still only guaranteed with respect to the ground state energy, but not the next levels, as the projected-out part is not necessarily the ground state, and therefore the projected Hilbert space will usually contain still parts of the true ground state. This implies that e_i may fall below e_0. But if this happens, this can be used to improve $|0\rangle$, and then reorthogonalize again. The same then applies for even higher states $e_{i>1}$.

9.2.2 Examples

As an example, consider, e. g., the Coulomb potential of section 7.4. Bound states need to be exponentially or better localized. Thus, assuming as a trial state a wave-function $\psi(x) = e^{-\lambda_0 x}$ will immediately yield $\lambda_0 = Z/(a_B)$, with a_B the Bohr radius. Of course, for the harmonic oscillator of section 6.2.1, this will not work as well, as it is Gaussian localized. Here, again, it is seen that a good choice is helpful.

To see how higher states work, consider again the harmonic oscillator. Use as a trial wave function

$$\psi_0(x) = \frac{1}{\lambda_0^2 + x^2}.$$

A good choice is to have a trial wave-function without nodes. After all, ground state wave functions have no nodes, and this should be respected by the trial wave-function. The

necessary ingredients are

$$\begin{aligned}\langle 0|H|0\rangle &= \frac{\pi\hbar^2}{8m\lambda_0^5}+\frac{\pi m\omega^2}{4\lambda_0}\\ \langle 0|0\rangle &= \frac{\pi}{2\lambda_0^3}\\ e_0 &= \frac{\hbar^2}{4m\lambda_0^2}+\frac{m\omega^2}{2}\lambda_0^2.\end{aligned}$$

Taking the derivative yields

$$\partial_{\lambda_0} e_0 = -\frac{\hbar^2}{2m\lambda_0^3}+m\omega^2\beta_0.$$

The solution is $\lambda_0^2 = \hbar/(\sqrt{2}m\omega)$. Note that because a function which has, by construction, a minimum is calculated there must exist a solution for the λ_i. The resulting energy is

$$e_0 = \frac{1}{\sqrt{2}}\hbar\omega \approx 1.4E_0,$$

where the last statement used the known solution. Depending now on the situation, this may be considered a success (correct within a factor two) or a failure (40% off). The decision between success and failure often correlates with the available experimental precision and/or the results using other approximation methods and/or the availability of simulations.

9.3 WKB approximation

Another useful approximation method is based on the observation that many systems show only small deviations due to quantum effects compared to their classical behavior. Such systems are called quasiclassical. Depending on circumstances, this may either be because the system behaves intrinsically in this way or states are of interest, e. g. highly excited ones, which are already in the quasiclassical regime. This is usually a good approximation if the potential varies slowly on the scale of the de Broglie wave-length, i. e. the potential resolves the oscillations due to quantum effects poorly. As it will be seen, this turns out to be closely related to the action waves of classical mechanics, and thus is quite often used in the context of the path integral of section 11.3.

Under these conditions the WKB approximation, named after Wentzel, Kramers and Brillouin who developed it, can be used. Start again with a one-dimensional problem and write the wave-function as[4]

$$\psi(x) = F(x)e^{iS(x)} \tag{9.37}$$

[4]It is also not a coincidence that this seems similar to Bohm's formulation of section 11.5, especially (11.27), as this also centers on the classical aspects of wave-functions.

with two functions F and S. Assuming a stationary situation, and inserting it into Schrödinger's equation yields

$$\begin{aligned} \partial_x^2 F(x) - F(x)(\partial_x S(x))^2 + k(x)^2 F(x) &= 0 && (9.38) \\ 2(\partial_x F(x))(\partial_x S(x)) + F(x)\partial_x^2 S(x) &= 0 && (9.39) \\ k(x) &= \sqrt{2m(E - V(x))}. && (9.40) \end{aligned}$$

The quantity $k(x)$ reduces to the wave number in the limit of vanishing potential, i. e. free particle. It thus encodes the deviation of the wave-number from this quantity. Especially, in view of the preceding, the quantity $\lambda_B^2 \partial_x k(x)$ gives an estimate how classical the system behaves. The two equations (9.38) and (9.39) arise from the real and imaginary part of Schrödinger's equation. The latter necessarily has to vanish, as the eigenvalue is real. It is important to see that the potential, or $k(x)$, does not explicitly enter the equation. The other equation is the real part, and here the dynamics due to $V(x)$ enters explicitly. Equations (9.38-9.40) are yet an exact statement, and any one-dimensional, stationary problem can be written in this form.

As a next step, note that (9.39) is actually a total derivative,

$$\partial_x (F(x)^2 \partial_x S) = 0$$

which can be integrated to yield, still exactly,

$$F(x) = \frac{C}{\sqrt{\partial_x S(x)}}.$$

Thus, the real part of (9.37) is given, up to an integration constant C, which can be absorbed in the normalization, by its phase,

$$\psi(x) = \mathcal{N} \frac{1}{\sqrt{\partial_x S(x)}} e^{iS(x)}. \tag{9.41}$$

At first, this may seem like a very deep statement, as this applies still in general. However, ultimately this is only the statement that a single function ψ is written in terms of another single function S, which needs to satisfy that it solves Schrödinger's equation, which encodes an Hermitian operator. After all, that F and S should be real functions is suggested by (9.37), but has never been encoded mathematically in a strict way.

The next step is to perform the WKB approximation. The assumption was that $V(x)$ varies slowly with x. This implies that also the wave-function will be slowly varying, and thus both S and F. But slowly varying functions have even slower varying derivatives.

This will become even more pronounced for higher derivatives. Thus, the assumption is that the second derivatives in (9.38) can be neglected, yielding an equation for S only

$$(\partial_x S(x))^2 - k(x)^2 = 0. \tag{9.42}$$

This again assumes that F becoming zero is irrelevant. This is not an idle statement, as if F is zero, $\partial_x S$ needs to become infinite, and thus at nodes the WKB approximation breaks down. This is self-consistent with the initial assumption. Thus, under this initial assumption it is safe to proceed. But then (9.42) implies

$$\partial_x S = \pm k(x).$$

This yields for the free case $k(x) = k$ immediately the correct solution $S(x) = \pm kx$. In general

$$S(x) = \pm \int_{x_0}^{x} dx' k(x') = \pm \int_{x_0}^{x} dx' \sqrt{2m(E - V(x'))}. \tag{9.43}$$

The value of x_0 is the second integration constant.

The result (9.43) is familiar from classical physics. The same quantity $k(x)$ appears when calculating the time it takes to traverse a classical trajectory (though there the inverse is integrated), and it plays the central role in Fermat's principle of fastest arrival. Thus, (9.43) can be seen as solution with minimal phase. And this is once more related to the WKB approximation being 'almost classical', as the discussion of the path integral in section 11.3, and its approach to classical physics, will show.

An interesting twist is observed when the energy in (9.43) is chosen to be less than V, and thus in classically forbidden regions. This yields an imaginary S for these values of x, and thus, because of (9.41), to an exponential decay of the wave-function, as expected.

Because when tunneling through a classically forbidden region the wave-function usually does not show an oscillatory behavior, the WKB approximation should usually be not that bad. Thus, it is often useful to estimate tunneling rates, as this only requires to integrate the wave-function. This gives indeed often a reasonable good first approximation. It also immediately shows that tunneling is exponentially suppressed, if $E \ll V$, as then (9.43) becomes small.

It is possible to also improve the WKB approximation. In this case, the neglected contribution in (9.38) is also expanded in a power series around $k^2 F(x)$, and then added to $k(x)^2 F(x)$ as a series, thus altering (9.43).

9.4 Time-dependent perturbation theory

So far, all approximate solutions were based on stationary problems. Thus, problems like a wave packet incident on a potential or the transition of a particle between two states are not yet accessible in an approximate manner. The variational approach of section 9.2 and the WKB approximation of section 9.3 lend themselves relatively straightforwardly to an extension to time dependence. In the variational case it requires time-dependent trial states, and for the WKB approach it introduces a time dependence in the function $k(x,t)$ in (9.40).

The situation is more involved in the extension of perturbation theory of section 9.1. However, as this approach has different limitations than the other two, it is still worthwhile to understand the necessary modifications. In addition, in practice, perturbation theory is often the most useful, and successful, approximation method. Thus, also for reasons of practical relevance it is necessary.

9.4.1 Explicitly time-dependent problems

To do so, it is useful to first think a bit more about explicitly time-dependent problems.

To this end, it is useful to add to the Schrödinger and the Heisenberg picture of time development of section 5.6 a third one, which emphasizes the difference of H and H_0, the so-called interaction picture. For this to be useful, it will be necessary to define H_0 such that

$$H(t) = H_0 + V(t),$$

i. e. the unperturbed Hamilton operator does not depend explicitly on time. For the moment, the parameter λ will be included into the definition of $V(t)$. Then the stationary eigenstates of H_0 will remain a time-independent basis of the Hilbert space. This implies that time-evolution can now be more complicated than $\exp(-iHt/\hbar)$. Especially, eigenstates are generally no longer invariant under time-evolution. Thus, the central question will be the matrix elements $\langle i,t|j,0\rangle$ describing the transition from the state j to the state i during the time t. They can now be induced, especially, by $V(t)$.

Because this implies that the eigenvalues of H are no longer necessarily time-independent, they will no longer be considered to be energies. Rather the eigenvalues of the time-independent Hamilton operator H_0 will be called E_n, allowing also degeneracies until noted otherwise. Because the time-independent eigenstates of H_0 still form a basis any

time-dependent[5] state $|\alpha, t\rangle$ can be written as

$$|\alpha, t\rangle = \sum_n c_n(t) e^{-i\frac{E_n t}{\hbar}} |n\rangle, \tag{9.44}$$

with $H_0|n\rangle = E_n|n\rangle$. Note that the time-dependent coefficients have been arbitrarily split into a known phase factor, and some unknown function of t. If $H = H_0$ then the $c_n(t)$ do not depend on time. A time-dependence therefore encodes a deviation of the Hamilton operator from the unperturbed case. It is also true that the transition probabilities $|\langle m|\alpha, t\rangle|^2 = |c_m(t)|^2$ and they therefore encode how states move into different eigenstates over time.

Now (9.44) has splitted the time evolution of the state $|\alpha, t\rangle$ into two parts. A trivial part determined by H_0, and a non-trivial part $c_n(t)$, which is yet unknown. Thus, the aim must be to determine the $c_n(t)$, which will be done in a perturbative manner.

It is useful to formalize this splitting a bit more, which yields the so-called interaction picture. Given an H and an H_0, the interaction-picture kets are defined as

$$|\alpha, t\rangle_I = e^{\frac{iH_0 t}{\hbar}} |\alpha, t\rangle \tag{9.45}$$

i. e. the interaction picture state kets reverse the known time-evolution of a given Schrödinger-state space ket. In addition, the observables O in the interaction picture are defined as

$$O_I = e^{\frac{iH_0 t}{\hbar}} O e^{-\frac{iH_0 t}{\hbar}}. \tag{9.46}$$

Thus, the interaction picture is, in a sense, half-way between the Schrödinger picture and the Heisenberg picture, as it includes only the unperturbed time-dependence in the operators and kets. The most important thing is that any expectation value of an observable yields

$$\langle O\rangle_I = {}_I\langle\, \alpha, t|O_I|\alpha, t\rangle_I = \left\langle \alpha, t \left| e^{-\frac{iH_0 t}{\hbar}} e^{\frac{iH_0 t}{\hbar}} O e^{-\frac{iH_0 t}{\hbar}} e^{\frac{iH_0 t}{\hbar}} \right| \alpha, t \right\rangle = \langle \alpha, t|O|\alpha, t\rangle,$$

and thus coincides with the one obtained in the Schrödinger picture. Thus, this describes still the same physical object.

The time-dependence of (9.45) now implies

$$\begin{aligned} i\hbar\partial_t|\alpha, t\rangle_I &= i\hbar\partial_t \left(e^{\frac{iH_0 t}{\hbar}}|\alpha, t\rangle\right) = -H_0 e^{\frac{iH_0 t}{\hbar}}|\alpha, t\rangle + e^{\frac{iH_0 t}{\hbar}}(H_0 + \lambda V)|\alpha, t\rangle \\ &= e^{\frac{iH_0 t}{\hbar}} \lambda V e^{-\frac{iH_0 t}{\hbar}} e^{\frac{iH_0 t}{\hbar}}|\alpha, t\rangle = \lambda V_I|\alpha, t\rangle_I. \end{aligned} \tag{9.47}$$

Thus, the interaction-picture state ket (9.45) fulfills a Schrödinger equation where the Hamilton operator becomes the perturbation λV_I. Reversely, of course, if $\lambda V_I = 0$, the interaction-picture ket becomes stationary.

[5] Initial times equal to zero will be suppressed in the notation throughout.

Because the eigenstates $|n\rangle$ of H_0 still form a basis, any interaction-picture ket can still be written as

$$|\alpha, t\rangle_I = \sum_n c_{\alpha n}(t)|n\rangle \tag{9.48}$$

i. e. with stationary base kets, and all time-dependence in the coefficients $c_{\alpha n}(t)$. Rewriting (9.47) in terms of the components in this basis and inserting a one yields

$$i\hbar\partial_t\langle n|\alpha, t\rangle_I = \sum_m \langle n|V_I|m\rangle\langle m|\alpha, t\rangle_I.$$

Because the $|n\rangle$ are eigenkets of H_0

$$\langle n|V_I|m\rangle = V_{nm}(t)e^{\frac{i(E_n-E_m)t}{\hbar}} \tag{9.49}$$

where the dependence of V_{nm} on t is only the explicit time-dependence of the operator V, as the eigenstates of H_0 are time-independent. Inserting (9.48-9.49) into (9.47) yields a coupled set of ordinary differential equations for the $c_n(t)$

$$i\hbar d_t c_{\alpha n}(t) = \sum_m V_{nm}(t)e^{\frac{i(E_n-E_m)t}{\hbar}} c_{\alpha m}(t). \tag{9.50}$$

This is an equivalent formulation of the original problem. However, rather than to solve a partial differential equation in space and time, this is now a coupled set of ordinary differential equations, which is often simpler to solve. But this requires that the eigenspectrum of H_0 is exactly known, and often the expression (9.49) involves non-trivial integrals, so that (9.50) may not even be known in closed form.

9.4.2 Perturbative treatment of explicit time-dependence

Thus, an exact solution of (9.50) may still not be possible. This is where time-dependent perturbation theory is applied. There are several possible ways to do so.

The most direct approach is to expand

$$c_{\alpha n}(t) = \sum \lambda^m c_{\alpha n}^{(m)}(t)$$

with $c_{\alpha n}^{(0)}$ time independent. Inserting this into (9.50), and comparing order by order in λ then yields simpler differential equations for the $c_{\alpha n}^{(i)}$. This can then be solved explicitly, as this yields a recursion relation.

While this works, it is not the best practical choice to solve the problem. Rather, an approach based on the time-evolution operator U_I in the interaction picture, which satisfies

$$|\alpha, t\rangle_I = U_I(t, t_0)|\alpha, t_0\rangle_I$$

is a better starting point. In analogy to section 5.2 equation (9.45) implies

$$i\hbar\partial_t U_I(t,t_0) = V_I(t)U_I(t,t_0), \tag{9.51}$$

i. e. the time-evolution operator in the interaction picture satisfies the interaction-picture Schrödinger equation. This implies necessarily that $U_I(t_0,t_0) = 1$.

As in section (5.2), (9.51) can formally be solved by

$$U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt' V_I(t')U_I(t',t_0). \tag{9.52}$$

This operator equation is at first not much simpler than before, especially as V_I is involved. Now, reintroduce λ, and expand the operator U_I in a power series[6] in λ,

$$U_I(t,t_0) = \sum_n \lambda^n U_I^{(n)}(t,t_0),$$

and insert this into (9.52). This yields a recursion relation

$$\begin{aligned} U_I^{(0)}(t,t_0) &= 1 \\ U_I^{(n)}(t,t_0) &= -\frac{i}{\hbar}\int_{t_0}^{t} dt' \frac{1}{\lambda} V_I(t')U_I^{(n-1)}(t',t_0). \end{aligned} \tag{9.53}$$

Where the factor λ^{-1} comes from the fact that λ had been included in section 9.4.1 into the definition of V_I. Thus, the determination of U_I has been reduced to integrals. If matrix elements like (9.49) are only searched for, this becomes even ordinary integrals. Thus, this is a constructive solution.

In fact, this perturbative expansion can be resummed into the so-called Dyson series by inserting the expression (9.53) in (9.52), yielding

$$U_I(t,t_0) = 1 + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 ... \int_{t_0}^{t_{n-1}} dt_n V_I(t_1)...V_I(t_n), \tag{9.54}$$

where the ordering is important as the operators V_I do not necessarily commute at different times. This is just the Dyson series (5.7) encountered already in section 5.2, but in the interaction picture. Thus, this series has now been derived here, but in this process it has also become visible that this is formally a series expansion of the time evolution operator. Though this formula looks comparatively simple when considering the time-independent perturbation theory of section 9.1, it is in practice quite involved, and thus warrants some deeper investigations and examples.

[6] As in section 9.1, this is an assumption that this is possible. However, for an operator it is much harder to assess the validity of this assumption than for a function or a ket.

9.5 Applications of time-dependent perturbation theory

9.5.1 Transition rates

The primary application in quantum mechanics of (9.54) is the determination of transition probabilities, e. g. of (de)excitation phenomena like radiation of atoms. Such processes will now be used to illustrate the use of time-dependent perturbation theory.

Before embarking on calculating U_I for particular examples, it is useful to first understand with which aim this will be done. If U_I is known, any time-dependent ket $|\alpha, t\rangle$ can immediately be determined in terms of the (stationary) eigenkets of H_0,

$$|\alpha, t\rangle_I = U_I(t, t_0)|\alpha, t_0\rangle = \sum_n |n\rangle\langle n|U_I(t, t_0)|\alpha, t_0\rangle,$$

allowing to identify $c_n(t) = \langle n|U_I(t, t_0)|\alpha, t_0\rangle$ by comparison with (9.48). Using (9.46) implies

$$\langle n|U_I(t, t_0)|\alpha, t_0\rangle = \left\langle n \left| e^{\frac{iH_0 t}{\hbar}} U(t, t_0) e^{-\frac{iH_0 t}{\hbar}} \right| \alpha, t_0 \right\rangle.$$

Now select $|\alpha, t_0\rangle$ to be also an eigenstate $|m\rangle$ of the unperturbed Hamiltonian. After all, any state can be decomposed in this way. Then

$$\langle n|U_I(t, t_0)|m\rangle = e^{\frac{i(E_n t - E_m t_0)}{\hbar}} \langle n|U(t, t_0)|m\rangle.$$

Thus, up to (known) factors, the matrix elements of U_I coincide with the transition matrix elements $\langle n|U(t, t_0)|m\rangle$ of the time-evolution operator U. Especially, the transition rates between two fixed states

$$|\langle n|U_I(t, t_0)|m\rangle|^2 = |\langle n|U(t, t_0)|m\rangle|^2 \tag{9.55}$$

are equal.

Combining then (9.48) and (9.53), using a particular normalization scheme, yields a

perturbative expression for the[7] $c_n^{(k)}$,

$$
\begin{aligned}
c_{nm}^{(0)} &= \delta_{nm} \\
c_{nm}^{(1)} &= -\frac{i}{\hbar}\int_{t_0}^{t} dt' \langle n|V_I(t')|m\rangle = -\frac{i}{\hbar}\int_{t_0}^{t} dt' e^{i\omega_{nm}t'} V_n m(t') \\
c_{nm}^{(2)} &= \frac{1}{\hbar^2}\sum_k \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \langle n|V_I(t')|k\rangle\langle k|V_I(t'')|m\rangle \\
&= \frac{1}{\hbar^2}\sum_k \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' e^{i(\omega_{nk}t'+\omega_{km}t'')} V_{nk}(t')V_{km}(t'') \\
c_{nm}^{(l)} &= -\left(\frac{i}{\hbar}\right)^l \sum_{n_1\ldots n_{l-1}} \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t_{l-1}} dt_l e^{i\left(\omega_{nn_1}t_1+\sum_{k=1}^{l-2}\omega_{n_k n_{k+1}}t_{k+1}+\omega_{n_{l-1}m}t_l\right)} \times \\
&\times\ V_{nn_1}(t_1)V_{n_{l-1}m}(t_l)\Pi_{r=1}^{l-2} V_{n_r n_{r+1}}(t_r) \\
\omega_{nm} &= \frac{E_n - E_m}{\hbar}.
\end{aligned}
\tag{9.56}
$$

Inserting this into (9.55) then yields

$$
P_{nm}^{(l)}(t) = \left|\sum_k c_{nm}^{(k)}(t)\right|^2
$$

as an approximation of the transition probability. Note that if $n \neq m$ the order $c^{(0)}$ term drops out. Thus, at leading order no transition occurs, and when switching the perturbation continuously on the transition probability to a different level will start from zero. The derivative of this with respect to time is then the transition rate.

9.5.2 Switching on

As an example consider switching on a constant perturbation at time $t = 0$, i. e. $V(t) = \theta(t)V$, with V some arbitrary time-independent operator. Such a situation is typical if in an experiment at time $t = 0$ some additional effect is switched on.

Then the matrix elements take the form $V_{nm}(t) = \theta(t)V_{nm}$, and thus the time-dependence becomes independent of the details of the perturbation. This allows to explicitly evaluate (9.56) for an arbitrary time-independent perturbation V, yielding at first non-trivial order for $t_0 = 0$

$$
c_{nm}^{(1)}(t) = -\frac{i}{\hbar}V_{nm}\int_{t_0}^{t} e^{i\omega_{nm}t} = \frac{V_{nm}}{E_n - E_m}\left(1 - e^{i\omega_{nm}t}\right). \tag{9.57}
$$

[7]Note again the assumption that an exchange of summation and integration is possible.

There are a few interesting observations. The first is that the prefactor has the same structure as that of stationary perturbation theory in section 9.1, i. e. the ratio of the perturbation to the level splitting. Thus, similar as there both information are relevant. The second is that the transition probability becomes to leading order

$$|c^{(1)}_{n\neq m}|^2 = \frac{4|V_{nm}|^2}{|E_n - E_m|^2} \sin^2\left(\frac{(E_n - E_m)t}{2\hbar}\right). \tag{9.58}$$

Thus, the transitions gets larger with V_{nm}, but are temporally modulated by the level splitting.

If looking deeper into the structure of (9.58) a deep relation to the discussion on wave packets in section 5.3 is observed. The transition probability (9.58) will only differ appreciably from zero if

$$t \gtrsim \frac{2\pi\hbar}{|E_n - E_m|} \tag{9.59}$$

as $\sin^2$ is a peaked function. Setting the time t in (9.59) elapsed from $t = 0$ to Δt and $|E_n - E_m| = \Delta E$ yields

$$\Delta t \Delta E \gtrsim 2\pi\hbar \gtrsim \hbar, \tag{9.60}$$

the same expression as (5.12). Thus, similar as for the arguments in section 5.3 perturbations induce a behavior reminiscent of an uncertainty relation in time and energy. Thus, the transition probabilities after some time Δt behave such as if the energy of the particle at time $t = 0$ would have only been known within some range ΔE. Especially, small Δt need to be accompanied by large ΔE, thus showing again how for brief times large (relative) energies play a role. Especially, energies large compared to the actual energy of the particle.

On the other hand, assume that levels are nearby such that $E_n \approx E_m$ in (9.58). This yields

$$|c^{(1)}_{n=m}|^2 = \frac{1}{\hbar^2}|V_{nn}|^2 t^2 + \mathcal{O}((E_n - E_m)^2) \tag{9.61}$$

Thus, close-by energy levels are very quickly populated.

9.5.3 Atomic transitions and Fermi's golden rule

Consider now that $V(t)$ describes the radiation off an electron bound in an atom[8]. Then the radiation will be emitted at some time $t = 0$, and assume this effect to be otherwise

[8] A detailed description would require to model electromagnetic interactions in quantum mechanics. This will be left to the lecture in advanced quantum mechanics.

constant in time. If the electron changes to any of the close-by levels by this, the total probability to this order is

$$P_r = \sum_{m, E_n \approx E_m} |c_{nm}^{(1)}|^2. \tag{9.62}$$

Of course, the electron could also jump directly to a state farther away, and this will occur eventually. However, as long as t is small, the situation is that of (9.58), and thus the nearby states will dominate.

The expression (9.62) can always be rewritten as

$$P_r = \int dE_m \rho(E_m) |c_{nm}^{(1)}|^2 \tag{9.63}$$

where the density of states, or spectral density, $\rho(E)$ has been introduced. This is formally always possible. If the spectrum is discrete, the spectral density will just be a sum of Dirac-δ functions, with prefactors given by the degeneracy, while it can also be continuous if necessary. This also allows to specify the range of energies to be considered, as the spectral density can, e. g., also have θ-like contributions to cut out only a part of the spectrum. Thus (9.63) includes (9.62), but generalizes it.

Inserting $c_{nm}^{(1)}$ from (9.58) into (9.63) yields the probability to go from level n to another level of energy E

$$P_n(t) = 4 \int dE \rho(E) \frac{|V_{nE}|^2}{|E_n - E|^2} \sin^2 \left(\frac{(E_n - E)t}{2\hbar} \right)$$

Because inserting t/t and

$$\lim_{t \to \infty} \frac{1}{\pi} \frac{\sin^2(\alpha x)}{\alpha x^2} = \delta(x)$$

the total probability to eventually end up at some level is given by

$$P_n = \lim_{t \to \infty} P_n(t) = \frac{2\pi}{\hbar} |V_{nE_n}|^2 \rho(E_n) t.$$

In contrast to (9.61) this rises linear in t. The reason is that while it becomes very quickly likely to go somewhere, the number of possible places plays a role in asking where it is possible to go so quickly. Thus, the total possibility, factoring this into account, rises only linearly. The transition rate is thus constant in time, and the original states becomes linearly depleted.

If the density of states and V are slowly varying functions of E, this gives a transition rate between two nearby states as

$$w_{n \to m} = \frac{2\pi}{\hbar} |V_{nm}|^2 \rho(E_n), \tag{9.64}$$

which is known as Fermi's golden rule.

Following essentially the same arguments the next-to-leading order contribution, for similar assumptions, can be calculated and is found to be

$$w_{n\to m} = \frac{2\pi}{\hbar}\left|V_{nm} + \sum_k \frac{V_{nk}V_{km}}{E_n - E_k}\right|^2 \rho(E_n). \tag{9.65}$$

The structure is quite similar to second-order stationary perturbation theory of section 9.1. There are, however, a few interesting remarks. The most important is that the sum allows an interpretation of the transition as a two-step process, $n \to k \to m$. This looks like a sum over all possible histories, an idea taking up by the path integral in section 11.3. In addition, E_k can be now very different from E_n, though finally the energy needs to be of the same size as E_n. This realizes (9.59). In the brief intermediate stage energy conservation does not hold. Thus, it is said that the process proceeds by intermediate virtual states.

Care has to be taken in (9.65) if singularities occur because $V_{nk} \approx V_{km}$ is not compensating for vanishing $E_k - E_n \approx E_k - E_m$. This requires usually some regulation, which will be discussed later in a more general context.

9.5.4 The laser principle

Also the way how lasers operate can be described using the developed machinery. The basic idea of a laser is that somehow a system is created, in which a lot of excited states are populated, which then deexcite by interacting with an electromagnetic (standing) wave. If the deexcitation energy is the same as the energy of the photons in the electromagnetic field, this will add more photons to the field, and thus amplify it, without altering its frequency, giving a strong, monochromatic field. This can then be used to create a laser field or laser pulse.

The initial excitation can be created in many ways, from heating, electromagnetic radiation and electric currents. Assume that the energy levels so populated are described by H_0. The centrally important question is how it interacts with the electromagnetic wave to be amplified. As it is a wave, its time-dependent potential can be modeled as

$$V(t) = Ve^{i\omega t} + V^\dagger e^{-i\omega t},$$

where the second term is necessary to obtain an hermitian operator. Again, V will not be specified further, but could, e. g., model the spatial setup of the laser.

If suitably controlled, there will (essentially) only be a single excited state populated. Thus, the situation is the one for which Fermi's golden rule in section 9.5.3 has been

derived. Thus, at leading order

$$\begin{aligned} c_{nm}^{(1)} &= -\frac{i}{\hbar}\int_0^t dt' \left(V_{nm}e^{i(\omega+\omega_{nm})t'} + V_{nm}^\dagger e^{-i(\omega-\omega_{nm})t'}\right) \\ &= \frac{1}{\hbar}\left(\frac{1-e^{i(\omega+\omega_{nm})t}}{\omega+\omega_{nm}}V_{nm} - \frac{1-e^{-i(\omega-\omega_{nm})t}}{\omega-\omega_{nm}}V_{nm}^\dagger\right). \end{aligned} \tag{9.66}$$

This is quite similar to (9.57), except now two terms appear. Thus, following the same arguments contributions will only be significant if $\omega \pm \omega_{nm} \approx 0$. This may look at first like energy would not be conserved, and the system either gains or looses energy. This is indeed true. The reason is that the time-dependent potential has this feature already build in. The reason is that there is no mechanism in the Hamilton operator to sustain this field (or the fact that the system is in an excited state initially), it is external. This should not come as a surprise, as similar situations are already repeatedly encountered in classical physics. External time-dependent fields in general violate energy conservation.

Interpreting (9.66) two effects appear. Depending on ω_{nm} only either of both terms can be large. One of them corresponds to the situation that energy is taken from the field to put a particle into the excited state, while the other corresponds to move the excited particle into a different state, freeing energies. If thus the initial condition is that the excited state is occupied this will put energy into the field, and thus feed energy into the external field, amplifying the laser pulse.

Because of hermiticity $|V_{nm}|^2 = |V_{nm}^\dagger|^2$. Thus, the rates (9.64) satisfy

$$w_{n\to\pm\omega} = \frac{2\pi}{\hbar}|V_{nm}|^2\rho(E_n \pm \hbar\omega).$$

Hence, the difference in the transition rate is entirely due to the density of states. This result is summarized in the statement of detailed balance

$$\frac{w_{n\to-\omega}}{\rho(E_n-\hbar\omega)} = \frac{w_{n\to+\omega}}{\rho(E_n+\hbar\omega)}.$$

Thus, the emission and absorption process have rates only normalized by the availability of states.

While detailed balance has now been derived in a particular context, it actually occurs in many cases. It is a quite general principle, though not something which is guaranteed to hold in any situation.

9.5.5 Decays

A quite important quantum mechanical effect are decays. Decays have already been briefly mentioned in section 6.6. There, they were modeled as a potential well bounded by a

potential step, and the particle could tunnel to infinity. Modeling this effect required to find a superposition of eigenfunctions which described a localized particle inside the well at initial time $t = 0$, and the times before that. While this is a good microscopical modeling of a decay process, it is in practice cumbersome for two reasons. It first needs a separate solution for any initial state, and thus for any initial energy. The second is that it requires a full solution, which is usually not possible.

Time-dependent perturbation theory allows for a more generic modeling. It is based on the idea, similar to the laser case in section 9.5.4, that initially an eigenstate of H_0 is occupied, and thus the system is described by an eigenket of H_0. As H_0 is assumed to be stationary, such a state would never change. Hence, it is necessary to introduce a perturbation which initiates the transition.

So far, all effects involved a sudden switch-on of the perturbation. This is usually not a good approximation for decays, as then immediately a high transition rate will commence. Rather, a better description is by a so-called adiabatic switching on of the perturbation. This assumes that at $t \to -\infty$ no perturbation was present, and it is then switched on exponentially slowly with time. The question will then be to determine the probability for the system to deexcite into some lower state at time t.

This describes, e. g., an atom, where the electron is not in the ground state, but in an excited state. A full description of this effect would involve the photon to be radiated off. This requires an adequate description of a photon, which requires (relativistic) quantum electrodynamics, and is thus beyond the scope of this lecture. Rather, it will be assumed that the influence of the radiated off photon can be described in some effective potential

$$V(t) = e^{\eta t} V,$$

which is adiabatically switched on. Note that this perturbation will become arbitrarily large as $t \to \infty$. Thus, this will not be a good model for late times. But this does not play a role at some (early) finite time t. What happens later has no bearing to what happens at the time t. The only requirement will thus be that t is small enough compared to η^{-1} that (leading-order) time-dependent perturbation theory is adequate.

At leading order,

$$c_{nm}^{(1)}(t) = -\frac{i}{\hbar} V_{ni} \int_{-\infty}^{t} e^{\eta t'} e^{i\omega_{nm} t'} dt' = -\frac{i}{\hbar} V_{nm} \frac{e^{(\eta + i\omega_{nm})t}}{\eta + i\omega_{nm}}$$

gives a transition probability

$$|c_{nm}^{(1)}|^2 = \frac{|V_{nm}|^2}{\hbar^2} \frac{e^{2\eta t}}{\eta^2 + \omega_{nm}^2}$$

and transition rate

$$d_t|c_{nm}^{(1)}|^2 = \frac{2|V_{nm}|^2}{\hbar^2}\frac{\eta e^{2\eta t}}{\eta^2+\omega_{nm}^2}.$$

If the perturbation is arbitrarily weak, $\eta \to 0$, this yields

$$\lim_{\eta\to 0} d_t|c_{nm}^{(1)}|^2 = \frac{2\pi}{\hbar}|V_{ni}|^2\delta(E_n - E_m). \tag{9.67}$$

Thus, for an arbitrary weak perturbation a transition is only possible if the energies exactly match. This is accordance with (9.60), as then the time needed becomes infinity, and thus the line width ΔE needs to become zero. Note also that (9.67) has again the form of Fermi's golden rule (9.64) as the transition is between two discrete states.

This leads for the first three orders to the results for $n \approx m$

$$\begin{aligned}
c_{n\approx m}^{(0)} &= 1 \\
c_{n\approx m}^{(1)} &= -\frac{1}{\hbar\eta}\delta_{nm}V_{nn}e^{\eta t} \\
c_{n\approx m}^{(2)} &= -\frac{1}{\hbar^2\eta^2}|V_{nn}|^2 e^{2\eta t} - \frac{ie^{2\eta t}}{2\hbar\eta}\sum_{k\neq n,m}\frac{|V_{nk}|^2}{E_n - E_k + i\hbar\eta}.
\end{aligned}$$

This looks quite similar to the time-independent perturbation theory of section 9.1. Note, however, the appearance of the $i\eta\hbar$ term in the denominator of the second-order contribution. It will play an important role now.

Consider the normalized rate, the ratio $d_t c_{nm}/c_{nm}$ up to this order. Since ultimately η will be send to zero, it is acceptable to approximate here $e^{\eta t} \approx e^{2\eta t} \approx 1$. This yields

$$\frac{d_t c_{n\approx m}}{c_{n\approx m}} = -\frac{i}{\hbar}V_{nn} - \frac{i}{\hbar}\sum_{k\neq n,m}\frac{|V_{nk}|^2}{E_n - E_k + i\hbar\eta}. \tag{9.68}$$

This normalized rate is independent of the time. Thus, this equation looks formally like the differential equation for the exponential function, though with a complex time parameter. The solution is

$$c_{n\approx m}(t) = e^{-\frac{i\Omega_{n\approx m}t}{\hbar}}.$$

If Ω is indeed complex, the rate will have both a decaying and an oscillatory dependence. But because of section 9.4.1 the c_{nm} are just the matrix elements of the time-evolution operator in the interaction picture,

$$|n,t\rangle_I = e^{-\frac{i\Omega_{n\approx m}t}{\hbar}}|n,0\rangle_I,$$

where it was used that the time-evolution operator in the current approximation is essentially diagonal. Moreover, this implies in the Schrödinger picture

$$|n,t\rangle = e^{-\frac{i\Omega_{n\approx m}t}{\hbar} - \frac{iE_n^{(0)}t}{\hbar}}|n,0\rangle,$$

and thus, independently of the actual value of Ω, always an oscillatory time-dependence is superimposed. This is equivalent to say that $E_n = E_n^{(0)} + \Omega_{n\approx m}$, and thus the perturbed energy levels are shifted by Ω, where the value of Ω is given by (9.68).

However, formally it is still necessary to take the limit $\eta \to 0$, as η still appears in (9.68). To do so, note that

$$\lim_{\eta\to 0} \frac{1}{x + i\eta} = P\frac{1}{x} - i\pi\delta(x), \tag{9.69}$$

where P is the principal value, a distribution formally defined as

$$P\frac{1}{x} = \lim_{\epsilon\to 0}\left(\int_{-\infty}^{-\epsilon} \frac{dx}{x} + \int_{\epsilon}^{\infty} \frac{dx}{x}\right).$$

This should be understood also as acting on something, just like the Dirac-δ function itself. Thus, the limit (9.69) rewrites the expression as a difference of two distributions.

Applying this to $\Omega_{n\approx m}$ from (9.68) implies an effect only at second order,

$$\Re\Omega^{(2)}_{n\approx m} = P\sum_{k\neq n,m} \frac{|V_{nk}|^2}{E_n - E_k} \tag{9.70}$$

$$\Im\Omega^{(2)}_{n\approx m} = -\pi\sum_{k\neq n,m} |V_{nk}|^2\delta(E_n - E_k). \tag{9.71}$$

Up to the principle part, the real part is just the same shift as in time-independent perturbation theory. More interesting is the imaginary part. It has the same structures as Fermi's golden rule (9.64). Thus, the summed transition rates

$$\sum_{k\neq n,m} w_{n\approx m\to k} = \frac{2\pi}{\hbar}\sum_{k\neq n,m} |V_{nk}|^2\delta(E_n - E_m) = -\frac{2}{\hbar}\Im\Omega^{(2)}_{n\approx m} = \frac{\Gamma_{n\approx m}}{\hbar}$$

describe the total rate of depletion of the state, and thus how quickly it decays. This gives rise to the decay width $\Gamma_{n\approx m}$, or lifetime $\tau_{n\approx m} = \hbar/\Gamma_{n\approx m}$, which describes by

$$|c_{n\approx m}|^2 = e^{-\frac{\Gamma_{n\approx m}t}{\hbar}}$$

the rate how quickly the state decays into nearby states.

Note that because

$$|c_{n\approx m}|^2 + \sum_{k\neq m} |c_{n\approx k}|^2 = 1 - \frac{\Gamma_{n\approx m}t}{\hbar} + \sum_{n\neq k} w_{n\approx m\to k} = 1$$

this does not violate conservation of probability. The amplitude to remain in the state n decreases, but this is compensated for by appearing in some other state.

The name width for Γ is related to the Fourier transformation of the c_i, the Fourier transformed is

$$\int dt e^{-\frac{i\Omega_{n\approx m}t}{\hbar}-\frac{iE_n^{(0)}t}{\hbar}} e^{\frac{iEt}{\hbar}} \approx \frac{1}{\left(E-(E_n^{(0)}+\Re\Omega_{n\approx m})\right)^2+\frac{\Gamma_{n\approx m}^2}{4}}. \tag{9.72}$$

Thus, the peak at $E_n^{(0)}+\Re\Omega_{n\approx m}$ is broadened to a structure of width $\Gamma_{n\approx m/2}$. The width itself satisfies

$$\Delta\tau_{n\approx m}\Delta\Gamma_{n\approx m}=\hbar$$

in analogy to (9.60). Thus, the time to deplete the state to nearby states is of order the inverse decay width, which localizes the process in Fourier space.

This structure is very generic, and reappears in many situations in quantum physics. The form (9.72), called Breit-Wigner shape, is typical: A resonance at some energy, broaden by a width. The 'some energy' is usually some characteristic scale, be it a bound-state energy, or also the harmonic frequencies at $\pm\hbar\omega$ of section 9.5.4. What, however, was important is that the transition was to nearby states, and that the decay rate was small compared to the energy, and thus perturbation theory applicable. If either of these conditions is not met, not even the functional form (9.72) should be expected.

Chapter 10

Quantum systems

The lecture so far has operated essentially on the description of a single particle, with the exception of section 6.2.3. In this it is similar to classical mechanics, where the understanding of single particles comes before multiparticle systems, especially before thermodynamics as a consequence of the statistical physics of classical mechanics with many particles.

Multiparticle quantum systems are much more involved than classical systems of single particles. This has will become very evident in the description of Bell's inequalities in section 11.2, which emphasizes the concept of entanglement to be discussed. This will be aggravated when taking into account angular momenta as well as the bosons and fermions of section 8.4. However, this will go far beyond the scope of this lecture, and will be taken up in the lectures on advanced quantum physics and quantum statistics. Here, rather, the most basic composition of quantum systems, and the emerging concepts, will be briefly discussed.

10.1 Multiparticle systems

10.1.1 Non-interacting case

In the following only the case of spinless particles will be considered, to reduce the complexity arising. Consider then a situation like in section 6.2.3, i. e. a system with multiple particles, which do not interact with each other. This is the first step before considering interacting systems.

In this case, every particle i is described by its Hamilton operator H_i, yielding energy eigenlevels E_{α_i} and eigenstates $|\alpha_i\rangle$, where Greek letters will be used to numerate the states of the particles, which are counted by Latin indices. The multi-particle system is

then a tensor product of the individual Hilbert spaces

$$|\{\alpha_k\}\rangle = |\alpha_1, \alpha_2, ..., \alpha_n\rangle = \odot_{i=1}^n |\alpha_i\rangle.$$

The total Hamilton operator is then given by

$$H = \sum_{i=1}^n H_i \odot 1_i,$$

where 1_i is a unit operator acting in all subspaces but the ith. This operator will be suppressed below. Any operator acting only on particle k acts then also only in the subspace of this particle, e. g.

$$\begin{aligned} H_k|\{\alpha_i\}\rangle &= \left(\odot_{i=1}^{k-1}|\alpha_i\rangle\right)\left(H_k|\alpha_k\rangle\right)\left(\odot_{j=k+1}^n|\alpha_j\rangle\right) = \left(\odot_{i=1}^{k-1}|\alpha_i\rangle\right)\left(E_{\alpha_k}|\alpha_k\rangle\right)\left(\odot_{j=k+1}^n|\alpha_j\rangle\right) \\ &= E_{\alpha_k}|\{\alpha_i\}\rangle, \end{aligned}$$

and likewise for any other operator.

Especially, when obtaining a wave-function, this is a matrix element with the eigenstate $|\{x_k\}\rangle$ of the position operator $X = \sum X_k$, yielding

$$\langle\{x_k\}|\{\alpha_l\}\rangle = \Pi_k^n \langle x_k|\alpha_k\rangle = \Pi_k^n \psi_n(x_k)$$

i. e. the full wave-function is just the product of the individual wave-functions. Likewise, any expectation values are sums

$$\langle\{\beta_k\}|O|\{\alpha_l\}\rangle = \left\langle \{\beta_k\} \left| \sum_r^n O_r \right| \{\alpha_l\} \right\rangle = \sum_k^n \langle\beta_k|O_k|\alpha_k\rangle,$$

or, in terms of wave-functions (for operators being diagonal in position space)

$$\langle\{\beta_k\}|O|\{\alpha_l\}\rangle = \int dx_1 ... \int dx_n \psi_{\beta_1}(x_1)^* ... \psi_{\beta_n}(x_n)^* \left(\sum_k^n O_k(x_k)\right) \psi_{\alpha_1}(x_1) ... \psi_{\alpha_n} x_n.$$

Thus, if $|\{\alpha_k\}\rangle$ is an eigenstate of some operator O, the eigenvalue is just the sum of eigenvalues of the individual particles.

In section 6.2.3, this manifested by the fact that the total energy was just the sum of the individual energies, counted by the individual quantum numbers.

Thus, in total everything is essentially added, and nothing happens.

10.1.2 Interactions

Things become more interesting if it is allowed that operators connect different particles.

As an example consider two harmonic oscillators. Without interactions, the eigenstates of the Hamilton operator are $|n_1, n_2\rangle$, signified by the numbers of oscillation quanta n_1 and n_2. Assume that $\omega_1 = \omega_2 = \omega$ for simplicity, then

$$H|n_1, n_2\rangle = \hbar\omega(n_1 + n_2 + 1)|n_1, n_2\rangle$$

are the full solutions of the eigenvalue problem. Given the creation and annihilation operators $a_i^\dagger$ and a_i, it is now possible to couple both one-particle Hilbert spaces by operators describing an interaction. Consider the action of, say, $O = a_1 a_2^\dagger$,

$$O|n_1, n_2\rangle = (a_1|n_1\rangle)\odot(a_2^\dagger|n_2\rangle) = \sqrt{n_1}|n_1{-}1\rangle\odot\sqrt{n_2+1}|n_2{+}1\rangle = \sqrt{n_1(n_2+1)}|n_1{-}1, n_2{+}1\rangle$$

i. e. the operator O transfers one oscillation quantum between both oscillators. Thus, such operators connecting both Hilbert spaces are defined by their action on the individual Hilbert spaces describing the individual particles. This is not surprising. After all, operators can be mapped to (infinite-dimensional) matrices, kets to vectors, and thus any multiparticle Hilbert space is just a tensor product of two vectorspaces. And thus there are block-diagonal matrices, like the Hamilton operator of this example, which do not mix different subspaces, and not-block-diagonal matrices like the one representing O, which do mix these subspaces.

Likewise, having two arbitrary kets and a two-particle operator, e. g. $(\vec{R}_1\vec{R}_2)^k$, an expectation value becomes

$$\langle 1, 2|(\vec{R}_1\vec{R}_2)^k|1, 2\rangle = \int d^3\vec{r}_1 d^3\vec{r}_2 \psi_1(\vec{r}_1)^* \psi_2(\vec{r})^* (\vec{r}_1\vec{r}_2)^k \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)$$

which no longer decomposes into a product of integrals in general, e. g. for $k = 1/2$.

Consider now the case that the Hamilton operator itself contains such operators. Start again with the harmonic oscillator, and consider

$$H = \alpha(a_1^\dagger a_2 + a_2^\dagger a_1)$$

Such an Hamilton operator is no longer decomposable into one-particle operators. Thus, eigenstates will no longer be tensor products of one-particle states. However, it is still possible to write any state as

$$|\beta\rangle = \sum_{n,k} c_{nk}|n, k\rangle$$

and from such general vectors certain values of c_{nk} will create eigenstates. Thus

$$\begin{aligned} H|\beta\rangle &= \alpha \sum_{n,k} c_{nk}(a_1^\dagger a_2 + a_2^\dagger a_1)|n,k\rangle \\ &= \alpha \sum_{n,k} c_{nk}\left(\sqrt{k(n+1)}|n+1,k-1\rangle + \sqrt{n(k+1)}|n-1,k+1\rangle\right) \\ &= \alpha \sum_{n,k} \left(\sqrt{n(k+1)}c_{n-1k+1} + c_{n+1k-1}\sqrt{k(n+1)}\right)|n,k\rangle \end{aligned}$$

yields a condition

$$\sqrt{n(k+1)}c_{n-1k+1} + c_{n+1k-1}\sqrt{k(n+1)} = c_{nk}$$

which needs to be satisfied for eigenstates of the Hamilton operator, which can then be solved as usual.

However, a wave-function

$$\langle x_1, x_2|\beta\rangle = \sum_{n,k} c_{nk}\langle x_1, x_2|1,2\rangle$$

is now an involved function. Resumming the series in the sense of a Taylor series can produce functions which do no longer decompose as a product $\psi_1(x_1)\psi_2(x_2)$, but are genuine two-argument wave-functions $\psi(x_1, x_2)$.

Likewise, the eigenfunctions of a Hamilton operator

$$H = \frac{P_1^2}{2m} + \frac{P_2^2}{2m} + V(X_1, X_2)$$

with a potential involving the position operators X_1 and X_2 in a non-trivial way will have in general eigenwavefunctions $\psi(x_1, x_2)$, which can no longer be written as a simple decomposition in a product of wave functions $\psi(x_1)$ and $\psi(x_2)$. It should be noted that, just like not any one-particle wave-function can be written as a Taylor series, not every two-body wave-function can be written in terms of a Taylor series.

The best-known example would be the two-electron atom, as an extension of section 7.4,

$$H = \frac{P_1^2}{2m} + \frac{P_2^2}{2m} - \frac{2\gamma}{|\vec{R}_1|} - \frac{2\gamma}{|\vec{R}_2|} + \frac{\gamma}{|\vec{R}_1 - \vec{R}_2|}$$

which, however, is already so involved that it can no longer be solved exactly. Here, all but the last term are one-particle operators, acting only on either particle. Only the last term, the electron-electron interaction, is a genuine two-particle interaction, as it cannot be written as a sum of two one-particle operators.

This also highlights that two (or more) particle operators describe interactions between particles. Hence, one-particle operators describe the action of an external influence on a

particle, while n-particle operators describe the interaction between n particles. This also implies that any fundamental law, which, by definition, has no exterior, will generally involve n-body operators.

A general treatment of such many-particle systems is beyond the current scope, but will be dominating topics in later lectures.

10.2 Ensembles and the density operator

An important formalism to discuss multiparticle systems are so-called density operators. This is not only important for the multiparticle systems of section 10.1, but also for measurements. A measurement on a single particle yields a particular outcome, with a certain probability. Thus, the value of, say, the energy is fixed to be an eigenvalue of the Hamilton operator. Expectation values of the Hamilton operator, on the other hand, describe the average energy obtained when doing many measurements on identically prepared states. Thus expectation values are obtained from ensembles of (identically prepared) systems/particles.

Consider again the Stern-Gerlach experiment of section 2.1. The distinction between the ensemble of particles and the probability to find a single value of the polarization did not seem to matter there. However, this was because the particles were distributed with the same probabilities as the number of outcomes, 50% each. What would happen with a partially polarized beam? This could not so easily be covered.

To understand the problem more clearly, consider the ket

$$|\alpha\rangle = c_+|+\rangle + c_-|-\rangle \tag{10.1}$$

as a superposition of kets with positive and negative polarization. Measurements show that this gives a probability $|c_+|^2$ to have positive polarization and $|c_-|^2$ to have negative polarization. But this is exactly the same as a ket entirely polarized along some axis not parallel to the polarization measurement. Thus (10.1) does not describe a randomly polarized particle, but one with a definite polarization, even though the measurement in a different direction can yield different results with different probabilities. The same will be true for any other values of $c_\pm$. How can then a genuine random ensemble be described?

This would need to realize a way how to have with some probability w_+ the ket $|+\rangle$ and with probability w_- the ket $|-\rangle$, satisfying $\sum_i w_i = 1$. For total randomness $w_\pm = 1/2$ would be required. Thus, it would not only be necessary to have a ket, encoding the randomness of measuring a single particle, but an ensemble of kets from which to draw a ket.

Where is now the difference to setting $c_\pm = 1/2$? The difference is that $c_\pm$ can have a relative phase, while the probabilities $w_\pm$ are real. Thus, an ensemble described by $w_\pm$ can have no interference, and is thus called an incoherent mixture. However, a ket with $c_\pm$ describes a definite state, and is thus a coherent mixture of the two kets $|\pm\rangle$. Especially, though numerically $w_\pm^2 = |c_\pm|^2$ may hold, this is not an identity, as it describes two different quantities. Any equality would be a consequence of the dynamics, and not a-priori true.

To distinguish the two cases, a situation where any $w_i = 1$ is called a pure ensemble, as all kets are necessarily equal. If this does not hold true, not all kets in the ensemble are the same, and it is called a mixed ensemble. If all w_i are the same, the ensemble is called (completely) random.

To pinpoint the difference again: Given a Stern-Gerlach filter, there is one orientation of (10.1), where no particle passes. This describes a pure ensemble. On the other hand, if in a completely random ensemble all values of $c_\pm$ in (10.1) appear with equal probabilities. Thus for any orientation of the Stern-Gerlach filter the same number of particles (but not the same particles) will pass. Thus, the values of $w_\pm$ should be considered to describe the probability distribution form which the $c_\pm$ are drawn. Note that this implies that any mixed ensemble can be considered to be a mixture of pure ensembles.

It is now left to find a way how to mathematically implement these ideas. The key will be the probabilities w_i, which satisfy

$$\sum_i^n w_i = 1, \tag{10.2}$$

where the number n gives the different possibilities present in the ensemble. This may be very different from the possible number of states. E. g., in a two state system, the ensemble may be characterized by $w_1 = 1/4$ of kets $|1\rangle$, $w_2 = 1/4$ of kets $|2\rangle$, and $w_3 = 1/2$ of kets $|1\rangle - 2|2\rangle$. In fact, the w_i may even become a continuum even for a system with discrete states, in which case (10.2) will become an integral over a probability distribution, rather than discrete probabilities. For simplicity, the kets drawn with probability w_i will be denoted by $|i\rangle$ in the following.

The ensemble average for any operator will therefore be given by a double average. Given that the eigenbasis of an operator A is $\{|a_k\rangle\}$, its expectation value in an ensemble characterized by the probabilities w_i is given by

$$\langle\langle A\rangle\rangle = \sum_i w_i \langle i|A|i\rangle = \sum_{i,k} w_i a_k |\langle a_k|i\rangle|^2, \tag{10.3}$$

and the double brackets $\langle\langle\rangle\rangle$ serve here to distinguish this from the ordinary expectation

value. Thus, the ensemble expectation value is determined by the average of the ordinary expectation values $\langle i|A|i\rangle$ in a fixed state.

While (10.3) certainly achieves the aim of giving a calculational tool to describe ensembles, it is in practice not yet convenient. The reason is that (10.3) still combines in a non-trivial way the ensemble kets $|i\rangle$ and the observable, and therefore needs recalculations for every observable. Consider instead some arbitrary basis $\{|b_k\rangle\}$. Then

$$\begin{aligned}\langle\langle A\rangle\rangle &= \sum_{ijk} w_i\langle i|b_j\rangle\langle b_j|A|b_k\rangle\langle b_k|i\rangle = \sum_{jk}\left(\sum_i w_i\langle b_k|i\rangle\langle i|b_j\rangle\right)\langle b_j|A|b_k\rangle \\ &= \sum_{jk}\langle b_k|\rho|b_j\rangle\langle b_j|A|b_k\rangle = \sum_k\langle b_k|\rho A|b_k\rangle \qquad (10.4)\end{aligned}$$

in which the density operator

$$\rho = \sum_i w_i|i\rangle\langle i| \qquad (10.5)$$

has been defined. Thus, the ensemble average can be rewritten in terms of a sum of ordinary expectation values, but not of the operator A, but of the product of operators ρA. Because (10.4) is just the sum of the diagonal elements of the matrix elements of ρA, this is also written as

$$\langle\langle A\rangle\rangle = \mathrm{tr}(\rho A),$$

i. e. formally as a trace[1]. The matrix elements of the density operator $\rho_{kj} = \langle b_k|\rho|b_j\rangle$ form the density matrix.

The density operator is, by construction, hermitian. Also

$$\mathrm{tr}\rho = \sum_{ij} w_i\langle b_j|i\rangle\langle i|b_j\rangle = \sum_i w_i\langle i|i\rangle = \sum_i w_i = 1$$

holds by construction. For a pure ensemble, $w_i = \delta_{ij}$ for some state j. Then $\rho = |j\rangle\langle j|$, but then

$$\rho^2 = |j\rangle\langle j|j\rangle\langle j| = |j\rangle\langle j| = \rho$$

and thus $\mathrm{tr}\rho^2 = 1$. It is less obvious, but can be shown, that for any none-pure ensembles $0 < \mathrm{tr}\rho^2 < 1$. Thus, this test can be used to check whether a given density operator describes a pure ensemble or not.

Note that the density matrix can itself be a continuous function, e. g.

$$\rho(x,y) = \langle x|\rho|y\rangle = \sum_i w_i\langle x|i\rangle\langle i|y\rangle = \sum_i w_i\psi_i(x)^*\psi_i(y) \qquad (10.6)$$

[1] In a finite-dimensional situation it is indeed a trace.

without being an integral over w_i. Once more, the set $\{i\}$ is in no relation to the basis of the kets. Note, however, that the diagonal elements of (10.6), $\rho(x,x)$, have a direct interpretation as the probability to find a particle of the ensemble at x, as this is just the weighted sum of all such probability densities. This is one of the reason for the name of the density operator.

As an example, consider again the Stern-Gerlach case. A completely polarized beam, and thus a pure ensemble, say in z+direction, has as density operator and matrix

$$\rho = |+\rangle\langle+| \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{10.7}$$

which evidently exemplify the properties of the density operator listed above. Likewise, a totally random ensemble is

$$\rho = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| \to \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \tag{10.8}$$

with $\mathrm{tr}\rho^2 = 1/2$. A mixed ensemble of 1/3 in $|+\rangle + |-\rangle$ and 2/3 in $|+\rangle - |-\rangle$ yields

$$\begin{aligned} \rho &= \frac{1}{3}(|+\rangle + |-\rangle)(\langle+| + \langle-|) + \frac{2}{3}(|+\rangle - |-\rangle)(\langle+| - \langle-|) \\ &= |+\rangle\langle+| + |-\rangle\langle-| - \frac{1}{3}|+\rangle\langle-| - \frac{1}{3}|-\rangle\langle-| \to \begin{pmatrix} 1 & -\frac{1}{3} \\ -\frac{1}{3} & 1 \end{pmatrix} \end{aligned}$$

with $\mathrm{tr}\rho^2 = 4/9$ after suitable normalization of the ensemble kets. The ensemble averages for $S_3 = \hbar\mathrm{diag}(1,-1)/2$ are then, using the matrix formulation, $\hbar/2$, 0, and 0, respectively. The latter would probably not be expected easily from a non-homogeneous mixture, but can now be straightforwardly determined.

While it is certainly possible to describe the time-evolution of an ensemble by using time-dependent states, there is a simpler procedure. Because any state obeys the Schrödinger equation, a density operator build from time-dependent states satisfies

$$i\hbar\partial_t\rho = \sum_i w_i(H|i,t\rangle\langle t,i| - |i,t\rangle\langle i,t|H) = -[\rho, H].$$

This is not (5.15), which is not surprising. After all, the time-dependence has been determined in the Schrödinger picture by introducing time-dependent states. Note that the w_i have been left time-independent. The reason is that otherwise the ensemble composition, which is determined not dynamically but by initial conditions, would change over time. Though such an external influence could be described, this is not the aim here.

It remains to understand what the observable is that ρ describes, as it is an Hermitian operator. However, this will be most transparent not by evaluating the expectation values of ρ directly, but rather some derived quantity.

To this end, consider again (10.7) and (10.8). It appears that a pure ensemble has a density matrix $\rho_{ij} = \delta_{ij}\delta_{ik}$, i. e. a single diagonal element, while a random ensemble has $\rho_{ij} = \delta_{ij}/\dim\rho$. Note for this that, because ρ is an hermitian operator, it can always be diagonalized.

Now define the quantity

$$\sigma = -\mathrm{tr}(\rho\ln\rho) = -\sum_i (S^{-1}\rho S)_{ii}\ln(S^{-1}\rho S)_{ii},$$

where in the second step S diagonalizes ρ. But then $(S^{-1}\rho S)_{ii}$ are just the eigenvalues of ρ, and hence σ is defined by the eigenvalues of ρ. This also gives a definition of the logarithm of the operator. Now, for pure ensembles

$$\sigma = 0$$

as either the eigenvalue is zero or the logarithm of the eigenvalue is zero. For the mixed ensemble

$$\sigma = -\sum_i^{\dim\rho} \frac{1}{\dim\rho}\ln\frac{1}{\dim\rho} = \ln\dim\rho.$$

It can be shown that this is actually the maximum value σ can obtain. This is due to the normalization condition $\sum w_i = 1$, but this will not be detailed here. Thus, σ is a measure of (dis)order. The pure ensemble has perfect order, and thus the lowest disorder, as everything is the same. The random ensemble has maximum disorder, as everything is as different as it can be, and therefore has the largest disorder. Thus, σ measures the amount of disorder. Thus, the observable associated with the density operator is a measure of disorder.

In classical statistical mechanics the amount of disorder is measured by the entropy S. It will be possible to show in quantum-statistical mechanics that one possibility to obtain the entropy will be to set $S = k_B\sigma$, where k_B is Boltzmann's constant. Thus, the observable associated to the density operator will ultimately become the entropy if the system is in thermodynamic equilibrium. This makes sense, as a thermodynamic system is just an ensemble of particles.

10.3 Entanglement

Finally, something needs to be said on the concept of entanglement, which will again come up in section 11.2. Consider a state of two particles of the Stern-Gerlach experiment,

$$|s\rangle = \frac{1}{\sqrt{2}}\left(|-+\rangle - |+-\rangle\right) = \frac{1}{\sqrt{2}}\left(|-\rangle \odot |+\rangle - |+\rangle \odot |-\rangle\right). \tag{10.9}$$

Though this state is build from the one particle states $|+\rangle$ and $|-\rangle$, it cannot be written as a product state

$$|s\rangle \neq (a|+\rangle + b|-\rangle) \odot (c|+\rangle + d|-\rangle),$$

as this would require at the same time a and c to be zero or non-zero simultaneously. This fact is called entanglement: Any multiparticle state, which cannot be written as a simple product state of one-particle states, is entangled. Thus, the characteristic feature of entangled states is that correlations between different particles cannot be neglected. Likewise, states which can be written as a product state, and thus essentially two independent particles, are not entangled.

Entanglement is an orthogonal feature to being an ensemble. After all, making up an ensemble entirely from (10.10) would be pure, while one made of 50% (10.10) and 50% of the not entangled state $|++\rangle$ would be mixed, and a purely random state made out of equal contributions of $|++\rangle$, $|\pm\mp\rangle$, and $|--\rangle$ would have again no entangled states present.

The important difference between entangled and not entangled states is that for a measurement in a not entangled state, say $\psi_1(x_1)\psi_2(x_2)$, any measurement of one-particle operators $O(x_1, x_2) = O_1(x_1) + O_2(x_2)$ automatically decomposes into a sum of single-particle measurements,

$$\langle O_1 + O_2\rangle = \int dx_1 \psi_1(x_1)^* O_1(x_1)\psi_1(x_1) + \int dx_2 \psi_2(x_2)^* O_2(x_2)\psi_2(x_2) = \langle 1|O|1\rangle + \langle 2|O_2|2\rangle \tag{10.10}$$

while this is not the case for an entangled state $\psi(x_1, x_2) \neq \psi_1(x_1)\psi_2(x_2)$,

$$\langle O_1 + O_2\rangle = \int dx_1 \int dx_2 \psi(x_1, x_2)^* O_1(x_1)\psi(x_1, x_2) + \int dx_1 \int dx_2 \psi(x_1, x_2)^* O_2(x_2)\psi(x_1, x_2). \tag{10.11}$$

Thus, in the presence of entanglement the system can no longer be decomposed into its subsystems, even when measuring only one-particle observables. The other particles always make their presence felt.

This has a wide range of consequences, which can actually sometimes be even features and advantages, and even be technologically useful. However, this is beyond the scope of this lecture. It will, however, give an interesting twist in section 11.4.

Chapter 11

Alternative and interpretation views on quantum physics

After having now a good understanding of the mathematical structure and the basic phenomenology, it is possible to return to the understanding of quantum physics.

Probably the most fundamental problem in understanding quantum mechanics is related to its probabilistic nature. This was one of the fundamental postulates in chapter 3. At the same time, the other postulate, that physics is encoded in states, is much less controversial. In fact, the original state space of classical mechanics is already quite close in spirit, and the formulation in symplectic geometry also. Thus, it is the probabilistic nature on which the following will focus.

It is important to note that, despite all efforts, so far no final interpretational answer has been given to what this all means. Not for lack of effort. On the other hand, the experimental confirmation of many of the more weird features of quantum physics makes it very hard, in fact almost impossible, to get rid of quantum physics and return to a deterministic view. However, as quantum effects are only noticeable on scales not accessible to us directly, it is very hard to make any statement what 'actually' happens. This precludes an intuitive way of understanding of quantum physics. Thus the question is justified, if it is just our ignorance of day-to-day experience which makes quantum physics strange, or if 'strange' is a fundamental statement.

Many of these questions reach (far) into the realm of philosophy of science, and ask deep questions about what physics actually is. It is impossible to do justice to such questions here, and this is necessarily left to lectures on philosophy of physics. The aim here is primarily to provide different view points which allow every one to have a start for making up their own mind[1].

[1] A strongly recommended reading on these subjects is the introduction of chapter 7 of "The message

11.1 Schrödinger's cat and decoherence

The one central question is, of course, if nature is really non-deterministic. This is especially amplified by the idea of Schrödinger's cat. If a state is not observed by an observer, what is it? Really an indeterministic superposition of all possibilities? Does it then have an independent reality at all? And should this then not also apply to even macroscopic objects? Perhaps even to an unobserved observer? Or, alternatively, does it just escape us what is going on, and the probabilistic description is just a helpless, but efficient, way of capturing the underlying actual behavior of nature?

One of the most common interpretations, the Copenhagen interpretation, is indeed to assume that nature is nondeterministic, and thus quantum theory captures this feature faithfully. However, in its original formulation, it made a hard cut between (microscopic) quantum physics and (macroscopic) classical physics. In this view, measurement lifted a quantum object into the classical world, which is determined by classical physics. It thus postulated two different regimes of physics, which do not mix.

Even though such a hard cut is conceptually possible, it turns out in practice that it becomes somewhat arbitrary, especially if multiple particles are involved, where to make it. An alternative is decoherence, which provides a smooth transition from microscopic quantum physics and large scale classical physics, very much in line with the special case of the Ehrenfest relations of section 5.7.

In this setting the fact that Schrödinger's cat is not noted in every-day experience is explained by a decoherence effect. The randomness emerges from the postulate on measurements. But what are measurements? Measurements are actually an interaction with a (large) system. Thus, to keep a system in a superposition it needs to be isolated from interactions which 'measure' in any way a quantity, because otherwise the state is thrown into a fixed eigenstate. This is also observed experimentally: As long as it is possible to isolate a state from a larger system, the quantum effects can be preserved even on macroscopic dimensions. Still, usual dispersions are microscopic.

Conversely, if a state is actually within our everyday experience, it needs to interact with us and our surrounding. By this, it is repeatedly measured, throwing it into eigenstates. Thus, the state is decoherent, as rather than keeping things like self-interference over long-distances, it becomes definite almost constantly. Thus, this decoherence creates essentially classical states, which then, due to the Ehrenfest relations, also behave classically. This interprets the emergence of classical physics as a continuous interaction process of a state with the surrounding system, and therefore carefully avoids any need for a hard

of quantum science", Lect. Notes. Phys. 899, Springer. Just to get a feeling that the situation is even worse.

separation of classical and quantum physics. Rather, a smooth emergence of classical physics from quantum physics is possible. This seamless embedding and agreement with experiment has led to making decoherence, rather than the Copenhagen interpretation, to the mainstream interpretation of classical quantum physics.

However, when taking the stance that quantum mechanics is only describing nature, rather than embodying it, this may make the interpretation irrelevant, but this changes little on an operational level. In fact, a possible viewpoint is that quantum mechanics is just a suitable parametrization of nature, but that human beings are not able to comprehend what is actually going on. This gives up on an interpretation. Another possibility is to require that we need a full understanding of consciousness, and how quantum effects trigger understanding and perception, in our mind before a true interpretation emerges. Neither of this options are currently logically excluded.

Finally, of course, one could hope that all of this is just a result of an imperfect understanding of the underlying physics, and that the real theory is actually deterministic. This, however, can be addressed, and essentially refuted, by Bell's inequalities to be treated next.

11.2 Bell's inequalities

One of the feelings one can get is that the reason quantum physics is probabilistic is possibly because we have so far an incomplete knowledge of what is going on. That is, there are some unknown, thus hidden, bits of information, which we are missing. A first argument against this are Bell's inequalities.

For this, it is easiest to consider a simple system. Consider a state which forms a product of two of the states of the Stern-Gerlach experiment of section 3.5. A possible state which can be constructed is

$$|s\rangle = \frac{1}{\sqrt{2}}\left(|+-\rangle - |-+\rangle\right) \tag{11.1}$$

where the $+$ and $-$ indicate in which direction the two particles are polarized with respect to the z-direction. Note that this is the same state as (10.10), and thus an entangled state. It is in fact this possibility for entanglement, which cannot be reproduced with such hidden information.

If a measurement of the polarization of the first particle is performed, then according to chapter 3 there is a 50% chance to find either polarization. This is achieved by a projection operator, $|\pm\cdot\rangle\langle\cdot\pm|$, where $\cdot$ means that the second particle is not affected. But then the

state $|s\rangle$ is afterwards either $|-+\rangle$ or $|+-\rangle$, and thus any subsequent measurement of the polarization of the second particle is already uniquely fixed.

The important point is that the state $|s\rangle$ was written without any reference to the position of the particle. Thus, the position does not matter[2], even if they are separated over an arbitrary distance. Thus, even if the distance is so large that there cannot be any reasonable interaction between both states, the statement holds true, and the measurement of one particle ascertains the measurement of the other one[3]. This is what entanglement is really about. In fact, after measurement of the first particle it is possible to predict with certainty what the measurement of the second particle will yield.

However, the measurement is not restricted to measure the polarization in z direction. Using the relation of section 3.5, the state (11.1) can be rewritten as

$$|s\rangle = \frac{1}{\sqrt{2}}\left(|x-x+\rangle - |x+x-\rangle\right), \tag{11.2}$$

i. e. it is rewritten in a basis for the polarization along the x-directions. If now the polarization of the first particle is measured in the z-direction, any measurement of the polarization of the second particle in the x-direction will still yield both polarizations with equal probability. But if now the first particle's polarization is measured in x direction then the measurement of the second particle's polarization becomes once more predetermined, as then the same argument holds as for (11.1)! Thus, even if both particles are arbitrarily separated, whatever is done to particle one matters to what behavior particle two exhibits. In fact, if particle one is not measured at all, measurements on particle two in either direction will be completely random.

Just following the measurement postulate to the letter, this is not surprising. Accepting this postulate leaves no other alternative to the outcome. However, because it is a postulate, it does not explain how this comes about. It merely arbitrates that this has to happen.

This non-locality of quantum physics is still one of the most opposed features of it. One approach, the Einstein-Podolski-Rosen (EPR) argument (or sometimes called paradox), was attempting to highlight this non-locality. How it does this will be shown now.

This is shown by Bell's inequalities. This works as follows. Presuppose that there is no such non-locality. Then somehow an information must be available to each particle how to behave. This needs to be stored somewhere, which is called hidden variables. The

[2]In practice, the simplest way to create such a situation is to obtain two particles from a common source like a decay of a particle, which ensures this state. Both particles are then allowed to travel far from the point of creation.

[3]In fact, quantum cryptography is based on this feature.

question is, is there any testable prediction in which a genuine quantum system and such a system of hidden variables would differ in an experimentally accessible way?

To establish such a hidden-variable approach, it is first necessary to somehow get the measurement of a single particle reproduced, as this is experimentally well established: Picking a random particle, a measurement of the various polarizations will be randomly distributed. To this end, the hidden variable approach postulates the existence of particles which are identical except that their polarization will be of a certain, but fixed, type. Since there is no other difference, the different particles will only be identified by the measurement of polarization, and if the pool of particles has the correct distribution the measurement of individual particles will yield no difference to ordinary quantum mechanics.

Now, to reproduce the preceding result with two particles this implies that the particle pairs must be produced in such a way that the relative amount of pairs of particles created recreate the pattern discussed in quantum mechanics above. This is necessary, as such measurements have been performed and were in agreement with the quantum prediction. This is still possible. It can be achieved by supposing that there a four particle types, called $z+x-$, $z+x+$, $z-x+$, and $z-x+$. After fixing the first particle, the second particle is uniquely identified by the resulting measurement. The only thing left to do is to ascertain that the production rates corresponds to the observed pattern.

The reason is that for two non-commuting quantities it can be generally shown that it is still possible to create such hidden variables. It fails when trying to get three non-commuting quantities.

Consider now a setup where three, not necessarily orthogonal, polarization measurements can be made along directions $\vec{a}$, $\vec{b}$, and $\vec{c}$. Again, in every direction there are only the two possible outcomes $+$ and $-$. There must therefore be $2^3 = 8$ different types of particles to ascertain any possible outcome for all three measurements. The production of these eight particle types will be in fractions $N_{abc} \geq 0$ each. This can now be used to deduce inequalities, which any measurement should obey.

Since it will be the aim to construct something where different predictions between the hidden variables idea and quantum mechanics can be found, it is useful to start with hindsight with inequalities which are known to lead to different predictions. To this end, note that

$$(N_{+-+} + N_{+--}) \leq (N_{+-+} + N_{+--}) + (N_{++-} + N_{--+}) \tag{11.3}$$

trivially. At the same time the probability P_{12} to measure the first particle in $\vec{a}$ with

positive polarization and the second particle in $\vec{b}$ with positive polarization is

$$P_{\vec{a}+\vec{b}-} = \frac{N_{+-+} + N_{+--}}{\sum N_{abc}} \tag{11.4}$$

as then the first particle is needed to have negative polarization in $\vec{b}$ direction. Likewise

$$\begin{aligned} P_{\vec{a}+\vec{c}-} &= \frac{N_{++-} + N_{+--}}{\sum N_{abc}} \\ P_{\vec{c}+\vec{b}-} &= \frac{N_{+-+} + N_{--+}}{\sum N_{abc}}. \end{aligned}$$

Dividing (11.3) by $\sum N_{abc}$ it becomes an inequality for these probabilities

$$P_{\vec{a}+\vec{b}-} \leq P_{\vec{a}+\vec{c}-} + P_{\vec{c}+\vec{b}-}. \tag{11.5}$$

This is Bell's inequality, and a hard prediction of hidden variables. If experiment would not satisfy this inequality, at the very least this version of hidden variables is ruled out. However, as only very general statements have been used, it requires very elaborate constructions to evade any such experimental result.

Doing such an experiment turns outs to be very challenging, and only within the last decades or so first results, which were systematically and statistically under reasonable, though not yet fully satisfying, control have become available. They yield a violation of Bell's inequality (11.5).

While this is evidence against hidden variables, this alone is not sufficient to support quantum mechanics. For this the corresponding prediction of quantum mechanics needs to be in agreement with the experimental outcome. It is thus time to derive a prediction for quantum mechanics for the same system.

The problem in doing so is now to first derive the quantum-mechanical prediction if the two vectors $\vec{a}$ and $\vec{b}$ are not orthogonal. This can be determined by a suitable decomposition of the eigenket in eigenkets in the different directions. This will not be done here in detail. Eventually, the probability to find a state polarized with one sign in $\vec{a}$ direction to be polarized in $\vec{b}$ direction with the same sign is

$$P_{||\vec{a}\vec{b}} = \cos^2\left(\frac{\theta_{\vec{a}\vec{b}}}{2}\right), \tag{11.6}$$

$\theta_{\vec{a}\vec{b}}$ is the angle between both directions. As required, this is one if $\theta_{\vec{a}\vec{b}} = 0$ and zero if $\theta_{\vec{a}\vec{b}} = \pi$, and varies continuously between both values. The probability to be not so polarized is then $1 - P_{||\vec{a}\vec{b}}$. Thus, the probability $P_{\vec{a}+\vec{b}+}$ (11.4) is in the quantum mechanics case

$$P_{\vec{a}+\vec{b}-} = \frac{1}{2} \times \sin^2\left(\frac{\theta_{\vec{a}\vec{b}}}{2}\right),$$

as in the first step the chance are 50-50 and in the second step the opposite of (11.6) needs to be measured. All other probabilities are determined in the same way, and only replace the angles.

Now, consider the case $\theta_{\vec{a}\vec{b}} = \pi/2$ and $\theta_{\vec{a}\vec{c}} = \theta_{\vec{c}\vec{b}} = \pi/4$. Then

$$\frac{1}{2} = P_{\vec{a}+\vec{b}-} \geq 0.292 \approx P_{\vec{a}+\vec{c}-} + P_{\vec{c}+\vec{b}-}$$

and thus Bell's inequality (11.5) is violated by quantum mechanics. This gives quantum mechanics at least a chance to be consistent with experiment, in contrast to hidden variables. To gain really support form the experiments requires now that the individual values for the $P_{..}$ agree with the experiment. This was also found, within the experimental uncertainty, giving further support for quantum physics to be the correct theory of microscopic physics.

Once one accepts, or even becomes acquainted, with these non-localities in quantum physics, the immediate question arises if it is possible to use this for instantaneous communication. Unfortunately, this is not even without special relativity possible. Though the polarization of the second particle is fixed once the one of the first particle is measured, if correlated axes are used, this does not help. Because it is entirely random what happens to particle one, the outcome for the second particle is also looking purely random, as long as the result on particle one has not been communicated by a different channel. As the same also applies to the different channel, this ultimately leads to the result that this cannot be used for fast communication[4].

It is tempting to ask what happens once relativity forbids arbitrarily fast interactions. Can then entanglement only be maintained over timelike distances? The answer is no. The reason is that relativity only forbids the possibility to transport energy/information over space-like distances, but allows still for correlations to be present. As to make use out of the result it is necessary to communicate between two measurement places, and this is again restricted by the speed of light. Thus, also relativistic quantum physics violates Bell's inequalities in a way consistent with experiments.

[4]Therefore the use in quantum cryptography is not the speed at which things happen, but rather the perfect correlation, which is destroyed by any intervening manipulation of data. Note that this entanglement is also at the heart of the complications of quantum teleportation, which works by transferring correlations rather than matter. Still, also this allows not fast communication, but consistent communication using different objects than the original particles.

11.3 The path integral

11.3.1 The propagator

As Bell's inequalities show, it becomes very involved to think of the measurement process as a feature of a particle, as really the operators themselves dominate the description. At the same time, the measurement postulate identifies any possible outcome of a quantum system as a sequence of measurements interspersed with time evolution from one measurement to another, and thus as an individual history of an actual measurement. Thus, quantum mechanics seems to be rather about histories than about particles.

Following up on this idea, it is possible to construct a different, but equivalent, formulation of quantum physics. This is the idea of the path integral, which shifts the perspective of quantum mechanics from expectation values and matrix elements of operators at instances of time to quantum physics being a superposition of all possibilities a system can realize, and then making probability statements about which history actually happens. This view also explains directly phenomena like the self-interference of the double-slit experiment of section 7.2.3. In contrast to this approach the construction so far, based on operators and the correspondence principle, is often called canonical quantization.

As both formulations turn out to be mathematically equivalent, neither is really 'the' picture of nature, just two facets of the same thing. Either can be postulated and the other derived, so it is a matter of taste, to the best of our current knowledge, which should be seen as fundamental. At the same time they are technically very distinct, and problems simple in one approach become horribly complicated in the other, and vice versa. Which to use in practice is therefore problem-dependent. As a rule of thumb, quantum-mechanical problems are usually easier to deal with using canonical quantization. The path integral shows it full strength in the case of (relativistic) quantum field theories, like quantum electro dynamics, and its generalizations.

While it is possible to postulate the path integral and then show how canonical quantization is implied by it, it is probably better, given the previous course of this lecture, to chose the other way around, and construct the path integral from canonical quantization. This also shows constructively how both approaches are equivalent.

The primary concept of the path integral is instead of the ket concept the notion that a particle undergoes all possible histories, and which history is realized, in the sense of a sequence of measurements in the form of (11.22), emerges then as an interference effect. Thus, the start will be to encode all such histories. In canonical quantization this is given by the time-evolution of some initial state ket $|\alpha, t_0\rangle$, where α will summarily denote anything else uniquely identifying the state at time t_0. Following chapter 5, this implies

that any possible state at a later time t is given by[5]

$$|\alpha, t_0, t\rangle = e^{-i\frac{H(t-t_0)}{\hbar}}|\alpha, t_0\rangle.$$

Its wave function is then given by

$$\psi_\alpha(x,t) = \langle x|\alpha, t_0, t\rangle = \left\langle x \left| e^{-i\frac{H(t-t_0)}{\hbar}} \right| \alpha, t_0 \right\rangle. \tag{11.7}$$

This connects the wave-function at some time t to an initial ket at time t_0. An interesting question occurs here: Is it possible to connect the wave-function at time t to the wave-function at time t_0, rather than to the ket?

This is indeed possible. To do so, insert into (11.7) a one with the eigenkets of the Hamilton operator[6] and one more one based on the position operator[7]

$$\begin{aligned}
\psi_\alpha(x,t) &= \sum_E e^{-i\frac{E(t-t_0)}{\hbar}}\langle x|E\rangle\langle E|\alpha, t_0\rangle = \int dy \sum_E \langle x|E\rangle\langle E|y\rangle e^{-i\frac{E(t-t_0)}{\hbar}}\langle y|\alpha, t_0\rangle \\
&= \int dy K(x,t,y,t_0)\psi(y,t_0) && (11.8)\\
K(x,t,y,t_0) &= \sum_E \langle x|E\rangle\langle E|y\rangle e^{-i\frac{E(t-t_0)}{\hbar}}. && (11.9)
\end{aligned}$$

Thus, the wave-functions at two different times are related by an integration over all of space with an integral kernel K constructed from the properties of the Hamilton operator only: Its eigenspectrum E and its eigenwavefunctions $\langle x|E\rangle = \psi_E(x)$. This quantity K is therefore called the propagator, because it propagates the wave-function from some time t_0 and y to some other time t at position x. The final wave-function is then obtained as the sum over all such propagated wave functions.

This has an interesting take on causality[8]: An undisturbed system evolves in time in a way which is entirely predictable from knowledge of the wave-function at some time t_0 in a way entirely determined by the Hamilton operator. This can be seen as a quantum version of causality. The wave-function in the future is uniquely determined by the past. The only part where this changes is when a measurement occurs. However, the reason is, in a sense, that a measurement is not part of the Hamilton operator, and therefore missing. If, as in section 11.1, measurements are taken to be only an idealization of a

[5]This can be generalized to cases where the time-evolution operator is not just the exponential of the Hamilton operator, but this leads right now to unnecessary complications.

[6]If there are degeneracies this requires to sum also over the degenerate states at fixed energy. This is suppressed in the following.

[7]For simplicity, they are taken to be discrete, but that can be changed in a straightforward way.

[8]Of course, once special relativity is included, this requires some amendments, but leads ultimately to a quite similar realization of causality.

measurement operator not taken sufficiently into account by the Hamilton operator, and thus a decoherence effect, this is alleviated. Then, the apparent disturbance of causality is reduced to a lack of details in the Hamilton operator. At any rate, equation (11.8) will still describe perfectly anything which happens between measurements.

Before continuing, it is worthwhile to note two features of the propagator:

$$\begin{aligned} K(x,t,y,t) &= \delta(x-y) && (11.10)\\ i\hbar\partial_t K(x,t,y,t_0) &= \sum_E \langle x|E|E\rangle\langle E|y\rangle e^{-i\frac{E(t-t_0)}{\hbar}} && (11.11)\\ &= \sum_E \langle x|H(P,X)|E\rangle\langle E|y\rangle = e^{-i\frac{E(t-t_0)}{\hbar}} H(x,\partial_x)K(x,t,y,t_0). \end{aligned}$$

These statements imply that the propagator is the Green's function of the Schrödinger equation in position space,

$$\begin{aligned} \left(-\frac{\hbar^2}{2m}\partial_x^2 + V(x) - i\hbar\partial_t\right) K(x,t,y,t_0) &= -i\hbar\delta(x-y)\delta(t-t_0)\\ K(x,t,y,t_0>t) &= 0. \end{aligned}$$

Note that this implies that the propagator is only defined for forward evaluation in time.

Interpreting, (11.10) is necessary to ensure that in (11.8) at coinciding times the equation is consistent. Then (11.11) follows from (11.9), as $\langle x|E\rangle \exp(iEt/\hbar)$ is an eigenstate of the Hamilton operator, and thus (11.9) is just a linear superposition of all such eigenstates. This leads to an interesting realization. Rewriting the exponential in (11.9) as a time-development operator, the inserted one in the energy eigenstates can be removed, to yield

$$K(x,t,y,t_0) = \sum_E \left\langle x \left| e^{-\frac{iHt}{\hbar}} \right| E \right\rangle \left\langle E \left| e^{\frac{iHt_0}{\hbar}} \right| y \right\rangle = \left\langle x \left| e^{-i\frac{H(t-t_0)}{\hbar}} \right| y \right\rangle. \qquad (11.12)$$

Thus, the propagator is the matrix element of the time-evolution operator in position eigenstates. This gives also one more possibility of how to view the contents of (11.8). The propagator describes how a particle localized at some point y at time t_0 is localized at point x at time t. A general wave-function encodes the probability with which a particle is localized at a fixed time somewhere in space. Hence, (11.8) yields how the localization probability of a particle in the past evolves into its localization probability in the present (or future) as a weighted average over the individual probabilities to move from some point in the past to another point in the presence weighted by how probable the particle was there in the past. Thus, in a sense, (11.8) is a weighted average over all possibilities how the particle can move from the places the wave function encodes.

This can be made even more explicit. In the sense of the Schrödinger picture of section 5.6 the quantity

$$|y, t_0\rangle = e^{\frac{iHt_0}{\hbar}}|y\rangle$$

can be seen as the time-dependent eigenstate of the position operator. The propagator is then

$$K(x, t, y, t_0) = \langle x, t|y, t_0\rangle \tag{11.13}$$

and thus nothing but the transition amplitude of the time-dependent position eigenstates of the particle.

To fill out the results, it is useful to consider an explicit example. The simplest is, of course, the free particle. In this case the eigenkets $|E\rangle$ are also eigenkets of the momentum operator. Using (4.13), this implies

$$K(x, t, y, t_0) = \theta(t - t_0)\int dp e^{\frac{ip(x-y)}{\hbar} - \frac{ip^2(t-t_0)}{2m\hbar}}.$$

Note that the wave-functions have not been normalized. Thus, using this propagator will destroy the normalization in (11.8), but this can always be remedied by dividing by a suitable normalization factor. Thus, this will not be considered to be an issue. The integrals can be performed by quadratic extension, and the final propagator is

$$K(x, t, y, t_0) = \sqrt{\frac{m}{2\pi i\hbar(t - t_0)}}e^{\frac{im(x-y)^2}{2\hbar(t-t_0)}}\theta(t - t_0). \tag{11.14}$$

This looks very much like the wave packet of section 5.1. This is not a coincidence, as the wave packet described a particle which starts localized from a point and then smears out over time, and thus is exactly what the propagator describes. However, this identification will not be so simple for a general potential, and thus should be taken as a confirmation rather than as a rule.

The propagator not only implements causality. It also contains all the relevant information on the spectrum in a very convenient way. That it is actually convenient can be seen by directly extracting it. Consider the integral

$$C(t, t_0) = \int dx K(x, t, x, t_0) = \sum_E e^{-\frac{iE(t-t_0)}{\hbar}}\int dx|\langle x|E\rangle|^2 = \sum_E e^{-\frac{iE(t-t_0)}{\hbar}}.$$

Setting $x = y$ means that the propagator now describes how a particle propagates back to its point of origin. The integral is then an average over all such acts of backpropagation. The function $C(t, t_0)$ describes this, and is given by a sum of phases. Thus, it describes the correlation between the same positions at different times. It is therefore called a correlator.

This is not yet so helpful, but two ways make it more useful. One is the observation that

$$C(-it,-it_0)=\sum_E e^{-\frac{E(t-t_0)}{\hbar}} \Rightarrow \lim_{t-t_0\to\infty} C(-it,-it_0)=e^{-\frac{E_0(t-t_0)}{\hbar}}.$$

Thus, the analytically continued function C is a sum of exponentials, with decay rates given by the spectrum of the Hamilton operator. At long imaginary times only the ground state E_0 dominates. This provides an often used technical tool to determine the ground state energy. Probably even more useful is the Laplace transform,

$$C(E)=-\frac{i}{\hbar}\int_0^\infty dt' C(t,0)e^{\frac{iE't}{\hbar}}=-\frac{i}{\hbar}\int_0^\infty dt'\sum_E e^{\frac{i(E'-E)t}{\hbar}}.$$

This integral cannot be directly integrated. However, by analytically continuing it to $E+i\epsilon$ with ϵ infinitesimally, this is possible. Because every term is still bounded, by an exponential, the sum is absolutely convergent, and summation and integration can be exchanged. This leads finally to

$$C(E)=\lim_{\epsilon\to 0} C(E+i\epsilon)=\sum_{E'}\frac{1}{E-E'}.$$

Thus, the function $C(E)$ has poles at the eigenenergies of the Hamilton operator. Thus, finding the poles of $C(E)$ is equivalent to solving the Schrödinger equation. Though this may not seem to be an advantage at this time, in more general settings this is often the simpler problem.

11.3.2 Formulating the path integral

The investigation of the propagator has already yielded a rich set of insights into the structure of quantum physics. This suggests that it should be possible to develop a view of quantum mechanics which sets the propagator at the center of the dynamical formulation, rather than the Schrödinger equation. This is indeed possible. To see how it emerges, a suitable starting point is (11.13).

This result has so far been seen as a transition amplitude from one point y to a different point x over time. But it is possible to view this differently. As the position operator is a hermitian operator, its eigenstates form necessarily a basis. This is still true for the time-dependent position eigenkets $|x,t\rangle$. But then the propagator just describes the matrix elements for a transformation of basis $\{|y,t_0\rangle\}$ to the basis $\{|x,t\rangle\}$. As this transformation is mediated by the time-evolution operator, this can be seen as that time-evolution is a base transformation. This is, not coincidental, reminiscent of how the Hamilton function creates time evolution as a flow along trajectories in phase space in classical mechanics.

Because at every fixed time the position operator has a basis, it follows for the propagator that

$$\begin{aligned} K(x_2,t_2,x_0,t_0) &= \langle x_2,t_2|x_0,t_0\rangle = \int dx_1 \langle x_2,t_2|x_1,t_1\rangle\langle x_1,t_1|x_0,t_0\rangle \\ &= \int dx_1 K(x_2,t_2,x_1,t_1)K(x_1,t_1,x_0,t_0), \end{aligned} \tag{11.15}$$

provided $t_2 \geq t_1 \geq t_0$. Thus, the propagator can be composed into the propagators at intermediate time steps. That could be repeated arbitrarily often by further splittings $t_1 \geq t_{n-1} \geq ... \geq t_1 \geq t_0$. For simplicity, consider now equal splittings of time, i. e. time intervals

$$\Delta t = \frac{t_n - t_0}{n}$$

with $t_{i+1} = t_i + \Delta t$. Thus

$$\begin{aligned} K(x_n,t_n,x_0,t_0) &= \int dx_{n-1}...dx_1 K(x_n,t_n,x_{n-1},t_n-\Delta t) \times \\ &\times K(x_{n-1},t_n-\Delta t,x_{n-2},t_n-2\Delta t)...K(x_2,t_0+2\Delta t,x_1,t_0+\Delta t) \times \\ &\times K(x_1,t_0+\Delta t,x_0,t_0). \end{aligned} \tag{11.16}$$

Close scrutiny reveals the meaning of this expression: It is the statement that time evolution proceeds by summing over all possible paths how a particle can move from t_0 to t_n, without the path being restrained by anything, not even by being differentiable. They just need to be continuous. In this sense, in quantum physics time evolution is an average over all possible paths, weighted by the propagator over all intermediate steps. This is very contrary to how classical physics operates, where there is a unique path, the geodesic in phase space, which is chosen among all the twice continuously differentiable paths possible to connect initial and final positions at initial and final times. Quantum mechanics is thus the extreme opposite: All paths, even the non-differentiable ones, contribute, and their weighted average creates the evolution of the particle from the initial situation to the final one.

There are several noteworthy observations to be made at this point. The first is that in arriving in (11.16) any notion of a wave-function is gone. So is any explicit reference to the Schrödinger equation or the Hamilton operator. Especially, no notion of kets appears right now anymore. Thus, if there would be a computational way of getting the propagator else wise, these concepts would no longer be needed. The second is that thus the same conceptual stage has been arrived as when passing from the Newtonian formulation to the integral formulation of Hamilton's principle in classical mechanics: No longer the dynamical equation but the trajectories play the central role. The third is the obvious

question: How can this at all create classical physics? It seems to be impossible that taken non-differentiable paths can at all connect to any classical idea of the trajectory of a particle. Nonetheless, so far only derivations have been made, and nothing added or changed in the postulates of quantum physics. Thus, this is still the same framework which gave the Ehrenfest relations of section 5.7, the coherent states of section 7.6.3, and generically the emergence of classical features at scales where $\hbar$ no longer plays a role. Thus, there must be some way how the classical limit emerges once more out of this quagmire.

All of this occurs in terms of the idea of the path integral formulation. This will now be accomplished by making a new postulate, which replaces Schrödinger's equation. Afterwards, it will be shown how Schrödinger's equation reemerges from the postulate. At the end, this will also lead to a new way of thinking on the fundamental postulates of chapter 3, and will require rephrasing them appropriately.

This will be a postulate, and thus there will be no way of deriving it. However, it can be motivated by the apparent mismatch to classical physics. For this, it should be recalled that classical paths are paths which minimize the action

$$S = \int dtL,$$

where L is the Lagrange function of the system. Thus, the classical path which connects two times t_i and t_{i+1} will minimize

$$S(x_{i+1}, t_{i+1}, x_i, t_i) = \int_{t_i}^{t_{i+1}} dtL$$

where the arguments of the action indicate that initial and final conditions are needed to fully specify a path classically. Quantum-mechanical now all paths should be averaged. It has been seen, e. g. for the coherent states of section 7.6.3, that classical properties appear as a superposition. Thus, a starting point is to require

$$\langle x_{i+1}, t_{i+1} | x_i, t_i \rangle = \sum_{\{x\}} e^{\frac{i}{\hbar} S(x_{i+1}, t_{i+1}, x_i, t_i)}, \tag{11.17}$$

where still it needs to be specified how to perform the sum over all possible paths $\{x\}$. However, this gives already a good idea of how classical physics could emerge: Because the classical path minimizes[9] the action, its contribution oscillates least. If the actions of all non-classical paths is large compared to $\hbar$, they could interfere away. In fact, this

[9]Note that in general in classical mechanics only extremalization is required. For any relevant physical system at the quantum level, however, it turns out that only systems described by minimization are relevant.

is how it will work in the end, and this actually turns out to be equivalent to the WKB approximation of section 9.3. But now, it is first necessary to give (11.17) a more concrete mathematical definition.

To do so, the following postulated procedure is used. The action is no longer considered to be a functional of a single path. Rather, the action is now considered to be just the quantum of action needed to go in a straight line from x_i to x_{i+1} in the, infinitesimally small, time interval Δt. Assuming for simplicity the simplest version of a conservative action[10] then

$$\begin{aligned} \langle x_{i+1}, t_i + \Delta t | x_i, t_i \rangle &= w(\Delta t) e^{\frac{i}{\hbar} S(i+1,i)} \\ S(i+1,i) &= \int_{t_i}^{t_i+\Delta t} dt \left(\frac{m(d_t x)^2}{2} - V(x) \right) \\ &= \Delta t \left(\frac{m}{2} \left(\frac{x_{i+1} - x_i}{\Delta t} \right)^2 - V\left(\frac{x_{i+1} + x_i}{2} \right) \right) + \mathcal{O}(\Delta t^2), \end{aligned} \tag{11.18}$$

i. e. the action is approximated as one term of the Riemann sum[11]. There is also a, not yet determined, normalization factor $w(\Delta t)$ introduced, which is assumed to be independent of the potential. Because of this, it can be determined by comparison of (11.18) with the free case (11.14), yielding

$$w(\Delta t) = \sqrt{\frac{m}{2\pi i \hbar \Delta t}}.$$

It is also already comforting that the expression (11.14) for the free particle coincides with (11.18) in the case of infinitesimal Δt. This weight-factor also ensures that in the limit $\Delta t \to 0$ (11.18) correctly reduces to $\delta(x_{i+1} - x_i)$.

Putting everything together, this allows to write the propagator as limiting prescription for a fixed time interval $T = n\Delta t$ and fixed endpoints $x_{n\to\infty}$ and x_0 at time t_0 and $t_{n\to\infty} = t_0 + T$

$$\langle x_n, t_n | x_0, t_0 \rangle = \lim_{n\to\infty} \left(\frac{m}{2\pi i \hbar \Delta t} \right)^{\frac{n}{2}} \int dx_{n-1} \int dx_{n-2} ... \int dx_2 \int dx_1 \Pi_{i=0}^{n} e^{\frac{iS(i+1,i)}{\hbar}} \tag{11.19}$$

To abbreviate this construction, this is used to define the mathematical operation of path integration as

$$\langle z, t | y, t_0 \rangle = \int_y^z \mathcal{D}x(t) e^{\frac{i}{\hbar} \int_{t_0}^{t} dt L(x(t), d_t x(t))}$$

[10] The procedure can be expanded to more involved ones, though this alters nothing conceptually.

[11] Once more, in general there are more technicalities involved, but without changing fundamentally the concepts. This also applies when it comes to other discretizations instead of the mid-point one adopted here.

yielding the path integral. Again, this is a postulate. It describes that the propagator is obtained by averaging over all possible continuous, though not necessarily continuously differentiable, paths, weighted by a phase given by the action. This also implies that this prescription no longer involves any operators, as the paths are ordinary functions, as is the action. This is thus an operator-free formulation of quantum physics.

11.3.3 Equivalence to canonical quantization

However, so far it is a postulate. To show that it is indeed equivalent to the formulation using the Schrödinger equation can be achieved by showing that the so obtained propagator still satisfies (11.11), i. e. the Schrödinger equation. Based on the principle that the same equations have the same solutions this implies then that this is an alternative formulation of quantum physics. Again, this will be done for the simplest version of an action as in (11.18).

To do so note that (11.19) by construction still implies the composition property (11.15) of the propagator, and thus

$$\begin{aligned}\langle z, t+\Delta t|y, t_0\rangle &= \int dx \langle z, t+\Delta t|x, t\rangle\langle x, t|y, t_0\rangle \\ &= \int dx \sqrt{\frac{m}{2\pi i\hbar\Delta t}} e^{\frac{i\Delta t}{\hbar}\left(\frac{m}{2}\left(\frac{z-x}{\Delta t}\right)^2 - V\left(\frac{z+x}{2}\right)\right)} \langle x, t|y, t_0\rangle. \end{aligned} \tag{11.20}$$

Because the path integral is a limiting procedure and because it satisfies (11.10) the propagator is a differentiable function in its arguments, even though it is an integral over non-differentiable paths. Furthermore, because

$$\lim_{\epsilon\to 0} \int dx \sqrt{\frac{1}{\pi i\epsilon}} e^{\frac{ix^2}{\epsilon}} = \delta(\epsilon) \tag{11.21}$$

the integral will be dominated by the region $x \approx z$ for sufficiently small Δt. Thus (11.20) can be expanded in both $z - x$ and Δt, yielding

$$\begin{aligned} & \langle z, t|y, t_0\rangle + \Delta t\partial_t\langle z, t|y, t_0\rangle + \mathcal{O}(\Delta t^2) \\ = & \sqrt{\frac{m}{2\pi i\hbar\Delta t}} \int dx e^{\frac{i\Delta t m}{2\hbar}\left(\frac{z-x}{\Delta t}\right)^2} \left(1 - \frac{iV\left(\frac{z+x}{2}\right)\Delta t}{\hbar} + \mathcal{O}(\Delta t^2)\right) \times \\ & \times \left(\langle z, t|y, t_0\rangle + \frac{(z-x)^2}{2}\partial_z^2\langle z, t|x_1, t_1\rangle + \mathcal{O}((x-z)^2)\right), \end{aligned}$$

where a term linear in $x-z$ in the last expansion is dropped, as the integration is symmetric with respect to it.

Because of (11.21) the 1 in the expansion of the potential term and the first term in the expansion of the propagator will yield $\langle z,t|y,t_0\rangle$. This cancels the corresponding term on the right-hand side. Dropping all higher-order terms, approximating $V(x+z)\approx V(z)$, and performing the integration, which is now a Gaussian multiplied by x^2, leaves

$$\Delta t\partial_t\langle z,t|y,t_0\rangle = \frac{i\Delta t\hbar}{2m}\partial_x^2\langle z,t|x_1,t_1\rangle - \frac{i\Delta t}{\hbar}V(z)\langle z,t|x_1,t_1\rangle.$$

But this is nothing but (11.11), and thus also the propagator obtained from the path integral satisfies the Schrödinger equation, establishing that both approaches are indeed identical.

This leaves to establish how observables O can be obtained. In fact, this is done using that the position operator is hermitian, and its eigenkets therefore form a basis. Thus, everything of interest resides in the ordinary functions

$$O(z,t,y,t_0) = \langle z,t|O|y,t_0\rangle.$$

Consider the simpler case that the operators are local, i. e. $O(z,t,y,t_0) = \delta(z-y)\delta(t-t_0)O(z,t)$. Now, though tedious, it is found by inserting back the whole construction that

$$O(z,t,y,t_0) = \int_y^z \mathcal{D}x(t)O(x(t),t)e^{\frac{i}{\hbar}\int_{t_0}^t dtL(x(t),d_tx(t))}$$

which can then be calculated, at least in principle, by reinserting the prescription (11.19). However, this implies that the path integral is over ordinary functions $O(x(t),t)$ of the paths, and not over operators. Operators are indeed eliminated entirely from this language, and only ordinary functions appear.

For non-diagonal operators, this is actually much more involved, but this is rarely happening. This also indicates that in ordinary quantum mechanics the path integral, while giving a very interesting alternative interpretation of quantum effects, is not convenient for actual problems. This changes radically in quantum statistics or relativistic quantum physics, but they will only be encountered in the coming lectures.

A severe problem also in the path integral remain measurements. Just as with the Schrödinger equation they cannot be implemented in a continuous fashion. Thus, measurements in the path integral formalism take the form that the propagator describes the evolution of the system up to the time of measurement. Then, the system is thrown in an eigenstate of the observable, and starting from this state the propagator describes the further time-evolution. I. e., there is again a repetition of propagation with measurement intermissions changing discontinuously the state.

As a final remark, it is now more natural to think of the WKB approximation in the context of the path-integral of section 11.3. In that case, all paths are rewritten as

$x_c(t) + \delta x(t)$, with x_c the classical path, and then the exponential is expanded in powers of δ. The ensuing integral is then merely polynomial integrals, which can be performed exactly. This assumes, however, that the summation in the series expansion and the path integral are interchangeable, which is only true if the series uniformly converges, which is not always guaranteed.

11.4 The measurement process revisited

From the previous discussion it becomes clear that a central element of the quantum nature is the measurement process. It is therefore interesting to have a more closer look at this very strange postulate of chapter 3. The fundamental insight of Bell's inequality of section 11.2 is that the probabilities in quantum physics are not working in the same way as in ordinary probability theory. After all, in ordinary probability theory probabilities are consequences of combinatorial counting and shares in possibilities. This is the idea behind hidden variables, and this failed (except if there are only two information/measurements involved).

It is here necessary to appreciate one of the fundamental features of the measurement postulates: Information is lost. The state will be in an eigenstate of the observable after the measurement, and anything else, e. g. the overlaps with other eigenstates, does no longer exist. But this happens actually as a time-evolution: There is a before the measurement, and there is an after the measurement. The first thing necessary is here to really understand the difference between measurement and time-evolution. Consider again the situation described by (3.4) and (3.5). What really happens is that the measurement occurs after the third filter. What happens before this is actually part of the time-evolution of the beam, and should be treated as such.

One of the central problems arising in the discussion of the measurement process is that an observable needs to interact with the state to be measured. Thus, the measurement itself becomes part of the system, and a disentanglement of the environment and the state is no longer strictly possible. However, in many cases the effects becomes exponentially damped, so it is often possible to not consider everything, but some reduced subset only, if only a 'feasible' precision is required. Still that there is an 'outside' is important, as it allows to dissipate quantities away from the system to be measured[12].

To really pass into the difference between quantum systems and classical systems it is important to note what has been said about observables in chapter 3: The difference between quantum mechanical observables A_i and their classical analogous a_i is that a_i

[12]Relativistic theories allow for more possibilities, but this is beyond the scope of this lecture.

are functions of trajectories in phase space, and thus necessarily commute. The operators A_i, however, do not necessarily commute anymore. In particular, two classical objects are described by two independent, and non-intersecting, phase space trajectories, while the same state in quantum physics is described by a common state, which allows for superpositions. Especially, entanglement is only possible in the quantum case.

Especially, for any classical theories the trajectories $x(t)$ are differentiable functions. Thus, any measurement $a_i(x(t))$ are uniquely known, and any product of two or more consecutive measurements is given by $a_i(x(t_1))...a_j(x(t_n))$. Especially, any ordering is the same. In particular, given any state $x(t)$, it is classically time-evolved again by the Hamilton function in phase space. In particular, any phase space trajectory, say for a single particle in one dimension, evolves as

$$d\vec{v} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \partial_q H \\ \partial_p H \end{pmatrix} dt$$

A measurement is then given by some function

$$a_i(x(t)) = a_i \left(\int_0^t \frac{d\vec{v}}{dt} dt \right)$$

and an arbitrary sequence of measurements is given by

$$a_i(x(t_1))...a_j(x(t_n)) = a_i \left(\int_0^{t_1} \frac{d\vec{v}}{dt} dt \right) ...a_i \left(\int_0^{t_n} \frac{d\vec{v}}{dt} dt \right)$$

and thus all measurements do not influence each other. In quantum physics, however, this is not true, as a sequence of measurements with a final result is described by

$$\langle i|\Pi_i U(t_n, t_{n-1})\Pi_j U(t_{n-1}, t_{n-2})...\Pi_k U(t_1, 0)|0\rangle \tag{11.22}$$

where, in line with chapter 3, the Π are projection operators. Even if the choice of projection operators is not random, the order here matters, as in general $[\Pi, U] \neq 0$, as now time evolution and measurements do not commute.

This problem is enforced by the fact that the same statement holds true with the outside. Any element Π' of the outside world may satisfy $[\Pi'_m, \Pi_n] \neq 0$, and thus it is not strictly possible to separate the outside from the observed system. Thus, in addition to the problems with the idealized measurement process (11.22) there are further problems when an actual measurement should be described.

In the end, this boils down to the problem of entanglement of section 10.3. If it is aimed for that there is no artificial outside of the system, the experiment and the system must be described simultaneously by the same Hamilton operator. But than, as was seen, only

entangled systems can transfer information. If they are not, any expectation values are necessarily blind to the other subsystem, see (10.10). But the price to be paid is that they also cannot avoid a back-flow of information, as the wave-functions cannot be separated in (10.11). For this not any size of couplings in the operators are alone relevant, but also the overlaps of the wave-functions. And the wave-functions themselves are a solution to the Schrödinger equation including the experiment itself. This has therefore very direct consequences.

This also implies a generic problem for optimizing the experiment to minimize this overlap: This is only possible if it is already known what the system to be investigated is, and how it works. Which makes the purpose relatively mute.

Of course, it allows to study the principle mechanism. E. g. it would be possible to organize operators and the system such that the wave-functions indeed show the property to evolve into something very well localized by the measurement process, thereby modelling the collapse of the wave-function by interaction with the experiment.

It is useful to consider an explicit example. Consider two particles, a massive one to be probed and a light test-particle, interacting only if they meet. The Hamilton operator is then[13]

$$H = \frac{P^2}{2m} + V(X) + \frac{P_t^2}{2m_t} + g\delta(X - X_t),$$

where the index t refers to the test particle, and thus $m_t \ll m$. It will be considered that the expectation value of momentum of the heavy body will be small compared to the one of the test particle. In this case, there will be no appreciable wave function of the light particle overlapping with the heavy one outside of the interaction region. If the wave-function of the heavy particle at $g = 0$ is then given by $\psi(x, t)$, this implies that

$$\psi(x, x_t, t) = \mathcal{N}\left(\theta(x - x_t)\left(e^{ipx} + re^{ip(2x-x_t)}\right) + \theta(x - x_t)te^{ipx}\right)e^{-i\frac{p_t^2 t}{2m_t}}\psi(x, t) \qquad (11.23)$$

is a fairly good approximation. The terms then correspond to an incoming light particle, modelled as a plane wave, which experiences a scattering similar to the δ-function potential of section 6.4 when encountering the particle to be probed. This leads to either reflection or transmission of it.

Introducing the reduced mass

$$\mu = \frac{mm_t}{m + m_t} \approx m$$

[13] One can replace the δ-interaction with any, sufficiently short-ranged, interaction, without any substantial changes in the following.

the reflection coefficient and transmission coefficient are given by

$$\begin{aligned} t &= \frac{1}{1+\frac{ig\mu}{p}} \\ r &= -\frac{\frac{ig\mu}{p}}{1+\frac{ig\mu}{p}}, \end{aligned}$$

respectively. This can be shown by inserting (11.23) into the Schrödinger equation, and using the approximations made.

It is important to note that, due to the transmitted term, the wave-function (11.23) does not separate into two one-particle wave functions. Thus, there is entanglement. To see this consider the density operator (10.5). There is only one state contributing, (11.23), with $w_1 = 0$. The interesting question is now to see how the experiment disturbed the heavy particle, not so much what the light particle is up to. The relevant matrix elements of the density matrix are thus the matrix elements in the one-particle basis $|x,t\rangle$. These are given by

$$\begin{aligned} \langle x,t|\rho|x',t'\rangle &= \int dx_t \psi(x,x_t,t)^*\psi(x',x_t,t)dx = \psi(x,t)^*\psi(x',t)d(x-x') \quad (11.24) \\ d(x-x') &= \frac{1}{1+\left(\frac{g\mu}{p}\right)^2}\left(1+\left(\frac{g\mu}{p}\right)^2 e^{ip(x-x')}\cos(p(x-x'))\right). \end{aligned}$$

Note that in this subspace the density operator is not a product of two wave-functions, and thus the state is not pure. Now, the modulus squared of d is a position peaked at certain distances. But it is still non-zero everywhere. Thus, the system is still in a quantum state.

In particular, if one measures the position,

$$\mathrm{tr}(X\rho) = \int dx' x' \psi(x,t)^*\psi(x,t)d(x-x') \quad (11.25)$$

this will pick up preferentially the peaks of d. But d is a relatively broad function, with peaks at $n\pi/p$, with n integer, and it is nowhere zero. Thus, the particle is not localized.

But any genuine experiment is not made from a single particle. Any real experiment involves many particles. Now model this by allowing N identical interactions with such light particles with the heavy particle. Since the light particles do not interact, the procedure is the same, and the full wave-function is hence given by repeated scattering. Removing then the light particles by repeated applications of (11.24) yields d^N instead of d in (11.25). But d^N is very strongly peaked around $\pi n/p$. Thus, the heavy particle is localized around discrete points. Hence, while the wave function is still extended, the particle has on average very few possibilities.

To get rid of this effect, chose the momenta of the particles slightly different and incommensurate[14]. Then the product $d_{p_1} \times ... \times d_{p_N}$ will be peaked only at $x - x' = 0$. Thus, the massive body entangled, by means of the density matrix, with all the light bodies from the experiment, is now actually localized. This gives now an explicit example of how decoherence proceeds by entangling with the system. Thus, decoherence, and emergence of a classical behavior, can be viewed as an effect due to the entanglement of the experiment made from a macroscopic number of particles, and the system under observation. It needs to be understood that it is really that only the expectation values localize, while the wave-function remains spread-out always, and never collapses.

But this implies that the measurement postulate is just an efficient way of encoding the entanglement between the system and the experiment. While this has been done now with a very special setup, the features generalize, provided that the interaction remain sufficiently short-ranged. Otherwise, quantum effects appear also at long distances. In addition, this can be made time-dependent, to show how classical paths emerge.

Note, however, that this explanation of the measurement process ultimately requires that the human brain is itself included into the description, as its particles are also entangled. This requires a physical description of the mind, something so far out of reach.

11.5 Bohm's formulation

Another alternative to the Schrödinger-Heisenberg formalism and the path-integral formulation and their interpretation is Bohm's formulation. Mathematically, it is essentially given again by Schrödinger's equation. However, given a wave-function Ψ, a trajectory $q(t)$ is defined to be a solution of

$$md_t q = \hbar \partial_q \Im \ln \Psi(q,t). \tag{11.26}$$

It is in this context important that this is not considered to be an equation to be coupled to the Schrödinger equation. Rather, it is necessary to solve the Schrödinger equation and then to solve this equation. Therefore, the wave-function is also called a pilot wave or guiding field for the trajectory $q(t)$. Since (11.26) is solved by a trajectory, this is interpreted to describe a path of a particle. This particle is always well localized, and there is no dispersion for it. Thus, it is a classical particle. This is associated to be the real particle. This can also be seen as the requirement that the trajectory $q(t)$ is selected among all admissible trajectories according to a probability distribution $|\Psi(q,t)|^2$. Interpreting this statement in terms of classical statistics, this implies that the randomness in quantum

[14] In fact, two different values are enough.

physics emerges as an equilibration phenomenon among classical trajectories, where the equilibration procedure is described by Schrödinger's equation.

This becomes even more clear, if a wave-function is written as

$$\Psi(x,t) = R(x,t)e^{i\frac{W(x,t)}{\hbar}}. \tag{11.27}$$

Then $\rho(x,t) = |\Psi(x,t)|^2 = R(x,t)^2$ and Schrödinger's equation can be decomposed in its real part and imaginary part, yielding two equations

$$\begin{aligned} -\partial_t \rho &= \frac{1}{m}\partial_x \left(\rho \partial_x W\right) \\ -\partial_t W &= \frac{1}{2m}(\partial_x W)^2 + V(x) - \frac{\hbar^2}{2m}\frac{\partial_x^2 \rho}{\rho}. \end{aligned}$$

Comparing to Hamilton-Jaocbi theory, this is a description of action waves W supplemented by an additional potential $\rho^{-1}\partial_x^2\rho$, which is coupled to the behavior of the action wave. Thus, essentially by introducing an additional, field-like, degree of freedom quantum physics can be seen as an extension of classical mechanics, and describing a classically moving particle.

Because the approach is seen as an equilibration procedure this explains the need for randomness and expectation values. Since picking a single particle out of an equilibrating system is discontinuously destroying equilibration yields the usual measurement prescription of quantum physics, pushing the particle into a particular trajectory.

This interpretation is consistent, and especially produces results in exact agreement with the Schrödinger equation. However, it can therefore not be distinguished experimentally from either the Schrödinger-Heisenberg version or from the path-integral version. Thus, the trajectories and classical particles are, in a sense, superfluous. They are not needed to explain anything, as (11.26) does not backcouple to the Schrödinger equation, and therefore cannot alter anything.

Therefore, when basing a theory on the concept of Occam's razor of having only what is necessary, Bohm's version is not suitable. If this criterion is not used, it is equally valid. However, extending this idea to relativistic theories is at best complicated, and it is not entirely clear if possible and if at all without being either again just more without different predictions. There are also attempts to actually backcouple (11.26) to Schrödinger's equation, and thus establish situations where a difference would arise. This has also not been successful yet, especially as no deviation of experiment and quantum physics within its realm of applicability has been observed.

11.6 And what does it all mean...?

This still does not exhaust all possibilities of how one can view quantum physics. There are many other alternative views, most of them trying to yield randomness not by a genuine probability structure, but rather to be, like in the case of Bohm's formulation, generated by some statistical effects of ensembles. In the end, without any observable discrepancies they all are just another reformulation of the Schrödinger-Heisenberg version, either without additional concepts like the path integral, or with some hidden but practically irrelevant concepts like Bohm's version.

Thus, while giving a different interpretation of how these things emerge, observationally everything is thrown back to the effects leading to the postulates in chapter 3. In the end, it thus remains to interpret these observations. Of these the measurement is the one really giving the problems.

There are, in principle, three major ways of interpreting them.

One is the so-called many-worlds (or Everett) interpretation. In this case, any measurement leads to a multiplication of the universe, and every possibility is realized in one of them. Thus, time-evolution becomes a permanent branching of universe histories. Each of them is as real as each other, but disconnect from each other, and thus not observable. Thus, all possibilities are real.

Another is that there is indeed a hard cut, and just one option is realized (the Copenhagen version), and the system is really not deterministic, and microscopic and macroscopic world are thoroughly disconnected.

The last one is that decoherence somehow creates a smooth transition, with one possibility discussed in section 11.4.

In versions one and three, there is no need for a hard cut between quantum physics and classical physics. It is here in the very nature of the non-localities, leading also to Bell's inequalities in section 11.2, which imply that classicalization has something to do with making a, in principle not allowed, cut between observer and system. Quantum physics, as especially the formulation using the path integral in section 11.3 emphasizes, is a global concept, and does not allow for hard cuts between parts of a system without loosing information.

In the end, what would really be required to topple, in the sense of Occam's razor to use the simplest explanation, quantum physics would be a genuine, reliable, reproducible, experimental results which blatantly disagrees with quantum effects, and requires a deterministic explanation. Otherwise, any additional discrepancies will only serve in extending quantum physics with its probabilistic realization. As of the time of this writing, there is not even a hint of disagreement. It remains to be seen, whether this will be so also in the

future. At any rate, a deterministic alternative will be very intricate, as to be able to come not into conflict with Bell's inequalities of section 11.2. And, of course, there is always the possibility that a third option arises due to experiment, beyond both possibilities.

In the end, the questions arising here have also a philosophical dimension, and touch upon questions like free will (not possible in a deterministic world, but perhaps in a quantum world) or the nature of reality. While relevant to physics, these questions are far beyond the scope of this lecture.

Index