Lecture Script

Quantum Mechanics 101

Prof. Dr. Anja Metelmann

KIT Sommersemester 2023

Karlsruher Institut für Technologie

Contents

1 The Schrödinger equation **5**

1.1 Wave-particle duality . . . 5

1.1.1 What is light? . . . 5

1.1.2 And what about matter? [1] . . . 7

1.1.3 A brief reminder of stochastic principles [2] . . . 8

1.2 The path integral perspective . . . 11

1.2.1 The principle of least action . . . 11

1.2.2 Feynman's path integral [3] . . . 13

1.2.3 The free particle [3] . . . 16

1.2.4 Deducting the Schrödinger equation [3] . . . 18

1.3 The time-dependent Schrödinger Equation . . . 21

1.3.1 Probability and continuity [4, 5] . . . 21

1.3.2 Solution for a free particle in 1D [4] . . . 22

1.4 The stationary Schrödinger Equation . . . 23

1.4.1 Separation ansatz [4, 5] . . . 23

1.4.2 Example: plane waves [4] . . . 23

1.4.3 Example: constant potential [4] . . . 24

1.4.4 Example: particle in a box [4, 5] . . . 25

2 Basic elements of quantum mechanics **27**

2.1 Position and momentum in quantum mechanics . . . 27

2.1.1 Correspondence principle [4] . . . 27

2.1.2 Wave function in position and momentum space . . . 29

2.2 The Hilbert space . . . 30

2.2.1 Superposition of waves: Fourier series [6] . . . 30

2.2.2 Complex vector space . . . 31

2.2.3 Dirac notation . . . 32

2.2.4 Expectation values and hermitian operators . . . 34

2.2.5 The uncertainty principle . . . 35

2.2.6 Some more formal aspects of the Hilbert space [4, 7] . . . 36

2.2.7 Continuous basis: improper vectors [7, 8] . . . 40

2.2.8 Discrete basis: proper vectors and matrix notation . . . 42

2.2.9 Linear Operators . . . 45

2.3 Prominent examples of quantum systems . . . 50

2.3.1 The atom as a two-level system . . . 50

2.3.2 Multi-level systems . . . 52
2.3.3 The quantum harmonic oscillator . . . 55

3 Perturbation theory & time evolution **59**
3.1 Time-independent perturbation theory [5, 9] . . . 59
3.1.1 Non-degenerate perturbation theory . . . 61
3.1.2 The nonlinear oscillator . . . 63
3.1.3 Degenerate perturbation theory . . . 66
3.2 Time-evolution in quantum mechanics . . . 68
3.2.1 The Schrödinger picture . . . 68
3.2.2 The Heisenberg picture . . . 69
3.2.3 The interaction picture . . . 70
3.2.4 The interaction frame . . . 72
3.2.5 Example: Single photon measurement . . . 74

4 Orbitals, spins and rotations **77**
4.1 Angular momentum operators . . . 77
4.1.1 General properties of angular momentum operators [5, 9] . . . 77
4.1.2 Orbital angular momentum operator . . . 81
4.1.3 Spin angular momentum operator . . . 85
4.1.4 Clebsch-Gordan coefficients . . . 86
4.1.5 The Wigner-Eckart theorem . . . 91

5 Particle in central potentials and fields **95**
5.1 Particle in a central potential . . . 95
5.1.1 Separation ansatz . . . 95
5.1.2 The hydrogen atom [5, 10] . . . 97
5.1.3 Fine structure in hydrogen [10] . . . 99
5.2 Particles in fields . . . 105
5.2.1 Precession of a spin in a magnetic field [5] . . . 105
5.2.2 An atom in a magnetic field . . . 106

6 Applications & fundamentals **109**
6.1 Measurement principles . . . 109
6.1.1 The measurement chain [15] . . . 109
6.1.2 Example: Stern-Gerlach experiment . . . 110
6.1.3 Irreversibility and Heisenberg's cut . . . 112
6.1.4 Projective measurement, systems and meters . . . 113

Chapter 1

The Schrödinger equation

1.1 Wave-particle duality

In this subsection we are going to discuss on of the basic consequences of quantum mechanics: the wave-particle duality. Meaning that in the quantum realm we cannot longer think of light as waves and matter consisting of particles, but rather that they can be both. This duality had to be developed in the early 20th century to explain experiments which where no longer described by classical physics.

1.1.1 What is light?

Young's double-slit experiment [1].

In contrast to what one might assume today, the question if light is a wave or a particle was historically argued for centuries (dating back to Greek philosophers multiple centuries before BC). It took till 1801 to most convincingly confirm the wave behavior of visible light when Thomas Young performed the first double-slit experiment. This experiment consists of a bright source of light, an intermediate screen with two very thin identical slits ($n = 1, 2$), and a viewing screen, see Fig.1.1. If only one slit is open then intensity of light on the viewing screen is maximum on the straight line path and falls off in either direction. We define the amplitude of the light field as a complex quantity $\alpha_n \in \mathbb{C}$, and the corresponding intensity profile on the viewing screen is proportional to the absolute value of this field amplitude

$$I_n \propto |\alpha_n|^2 , \tag{1.1}$$

which result in a uniform intensity distribution on the screen around a maximum in case either slit is open, see Fig.1.1 on the left. However, if both slits are open then the intensity oscillates according to the familiar interference pattern predicted by wave theory:

$$I_{12} \propto |\alpha_1 + \alpha_2|^2 = \left[|\alpha_1| e^{i\varphi_1} + |\alpha_2| e^{i\varphi_2}\right]^2 = |\alpha_1|^2 + |\alpha_2|^2 + 2|\alpha_1||\alpha_2| \cos(\varphi_1 - \varphi_2), \tag{1.2}$$

with $\varphi_{1,2}$ as the phase of the respective light amplitude. These facts can be very convincingly explained, both qualitatively and quantitatively, by positing that light travels in waves. The interference process creates a pattern which is distinct from just adding both intensities.

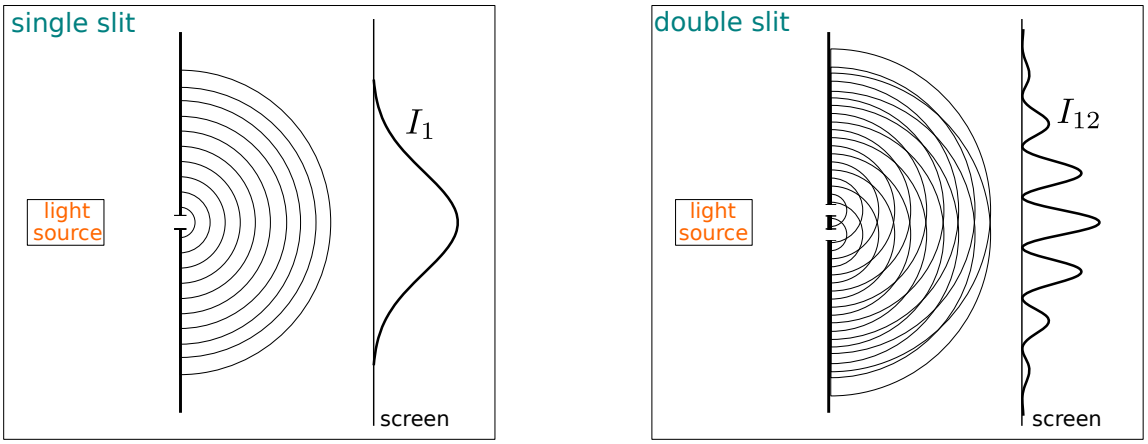

Figure 1.1: Young's double slit experiment.

Black body radiation and photo effect. [4]

The concept of light being a wave was questioned again in the context of black body radiation. A black body is here an idealized physical body that absorbs all incident electromagnetic radiation. The radiation emitted by a black body in thermal equilibrium with its environment is called black-body radiation. In the 19th century Kirchhoff has already postulated that a black body radiation strength depends on the temperature and the wavelength (or frequency), but not on the shape of the black body. The actual form of the this dependencies was unclear at first. Close before the end of the 19th century, Wilhelm Wien postulated the relation

$$B_\lambda^W(T) = b\,\frac{4c}{\lambda^5}\,e^{-a\frac{c}{T\lambda}}, \qquad a, b = const., \tag{1.3}$$

for the spectral radiance, i.e., the power emitted per area, direction and wavelength. c denotes here the speed of light and T the temperature. The proposed form was in analogy of Maxwell's velocity distribution for a gas. Wien's law explained nicely experiments performed till the mid-1900, however, there were small deviations for higher temperatures. Around 1900 Otto Lummer and colleagues performed refined experiments for higher temperatures and larger wavelengths, making the discrepancy with Wien's law more apparent. Crucially, for larger wavelengths the spectral radiance followed the Rayleigh-Jeans law

$$B_\lambda^{RJ}(T) = \frac{2c}{\lambda^4}k_B T, \tag{1.4}$$

with the Boltzman constant k_B. However, the issue here clearly arises for small wavelength as $B_\lambda^{RJ}(T)$ diverges for $\lambda \to 0$ (which is not the case for Wien's law). The latter is referred to as the ultraviolet catastrophe, as it arises at wavelengths in the ultraviolet region of the electromagnetic spectrum. Max Planck tried to unite both laws and deduced Planck's law

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5}\,\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}, \qquad h = 6.626 \times 10^{-34}\ \text{Js}, \tag{1.5}$$

with Planck's constant h (note one often uses $\hbar = h/2\pi$). To arrive at this expression Planck had to make some unusual assumptions, namely that electromagnetic radiation can only be absorbed or emitted in discrete packages of energy $E = h\nu$ (with frequency $\nu = c/\lambda$). This was the beginning of the development of a new theory: quantum mechanics.

At first, the postulation that light exchanges energy in discrete quanta was more considered as a theoretical trick of Max Planck to explain successfully the experimental results. Albert Einstein used this trick to explain another experimental observation at that time – the photo effect. It was observed that shining light on a metal can release free electrons out of the surface (Hertz, 1887). If light is a wave the intensity of the light field should determine the resulting kinetic energy of the electrons E_{kin}. However, this was not the observation, in contrast the frequency of the light was identified as the determining factor. Einstein postulated that an electron with mass m and velocity v has the kinetic energy

$$E_{\text{kin}} = \frac{1}{2}mv^2 = h\nu - W, \tag{1.6}$$

with W (work function) being the required energy to remove the electron from the metal. Hence, the kinetic energy of the electrons is determined by the quanta of energy of the light field – the photons. With the postulate that the energy of photons themselves is quantized, Einstein was able to explain the photo effect successfully. Taking now the expression for the total energy of a relativistic particle with zero mass (as a photon is) we obtain

$$E = pc = h\nu \quad \Rightarrow \quad p = h\frac{\nu}{c} = \frac{h}{\lambda}, \tag{1.7}$$

which is also called the de Broglie relation. This expression connects a typical characteristic of light, the wavelength, with a typical characteristic of a particle, the momentum. More generally, re-expressing the de Broglie relation with the wave vector $\mathbf{k}$ with absolute value $|\mathbf{k}| = 2\pi/\lambda$, one obtains the so-called dispersion relation $\mathbf{p} = \hbar\mathbf{k}$.

So if light does indeed behave like a particle, what would happen if we repeat the double-slit experiment with a light source ejecting single photons? From the particle nature we would expect that the measured intensities (or probabilities for detecting a photon in screen distance) are additive $I_{12} = I_1 + I_2$. However, this is not what one observes, interference experiments like the basic double-split setup will not reveal the particle nature of light.

1.1.2 And what about matter? [1]

Repeating the double slit experiments from above with electrons we expect that if either split is open the probability distribution of corresponding to either slits is similar to the intensity distribution in case for the photons, i.e., $P_n \sim I_n$. However, by considering that electrons are particles we indeed expect that by opening up both splits, we have just the resulting distribution is going to be the sum $P_{12} = P_1 + P_2$. Crucially, we find that this is not the case – the distribution takes the similar form as in the light case, $P_{12} \sim I_{12}$. So electrons behave like a wave not like a particle.

Quantum mechanics provides a way to reconcile both the wave and particle nature of electrons. Let us sketch how it might address the situation described above. Quantum mechanics introduces the notion of the complex amplitude $\psi_1(x) \in \mathbb{C}$ with which the electron goes through slit 1 and hits point x on the viewing screen. This amplitude is called a probability, which we will later identify with the wave function of the electron. The probability that the electron is actually detected at x is the square of the magnitude of this complex number: $P_1(x) = |\psi_1(x)|^2$.

Similarly, let $\psi_2(x)$ be the amplitude if only slit 2 is open. $P_2(x) = |\psi_2(x)|^2$. Now when both slits are open, the amplitude with which the photon hits point x on the screen is just the sum of the amplitudes over the two ways of getting there: $\psi_1(x) + \psi_2(x)$. As before the probability that the photon is detected at x is the squared magnitude of this amplitude:

$$P_n(x) = |\psi_n(x)|^2, \ \rightarrow \ P_{12}(x) = |\psi_1(x) + \psi_2(x)|^2. \tag{1.8}$$

The two complex numbers $\psi_1(x)$ and $\psi_2(x)$ can cancel each other out to produce destructive interference, or reinforce each other to produce constructive interference or anything in between. Thus electrons (or more generally particles) can behave as waves, and, vice versa, we know from the photo effect that light can behave as particles - the photons.
In 1923 Louis de Broglie proposed that all particles, not only electrons, are associated with waves, and that the frequency and wave number of the wave are given by the same relations we found above for photons.

1.1.3 A brief reminder of stochastic principles [2]

We learned in the last section that quantum mechanics works with probability amplitudes. Hence, we have to ensure that we understand some basic notions of probability theory and we start with a brief reminder of basic stochastic principles.

Classical randomness [2]. Consider the example of tossing a coin. There are two states of the coin, namely HEADS and TAILS, and since they are energetically equivalent, they appear with equal probability after a coin toss. In statistical mechanics individual states follow the 'equal a priori probability' postulate: All states with the same energy have equal weight or probability when computing averages. So the result of the toss is 'random'. It is worth going a little deeper to understand how the states (HEADS or TAILS) become random. To focus our thinking it is useful to do an experiment. Take a coin and flip it, always preparing it as HEADS. If you flip it at least 10-20 cm in the air, the result is random. If you flip it less than this height, the result is not random. Say you make N attempts, and the number of heads is n_h and the number of tails is n_t, for different flipping heights H. One finds that the ratio $(n_H - n_t)/N$ exponentially decreases by increasing the flipping height H, i.e., the coin toss is biased for small heights and the outcome depends on the initial state of the coin. In other words the coin toss is correlated with it's initial state (HEADS). There is a length scale over which those correlations persist. For our example, this scale is 2-3 cm, which we call ξ, the correlation length. The tailing of this bias follows $\propto e^{-H/\xi}$. Therefore we learn that true randomness requires $\xi/H \rightarrow 0$, which is analogous to requiring an infinite system in the thermodynamic limit.
In this classical example randomness can intuitively be understood as a lack of knowledge. We know Newton's laws and the fact that one would expect to be able to calculate the trajectory of the coin exactly, given some initial condition. It is true that given an initial configuration, and a precise force, we can calculate the trajectory of the coin if we neglect inhomogeneities in the air. However, your thumb cannot apply a precise force for flipping to within one correlation length of height (i.e., we lack the information about the force), and in practice computing the trajectory by using Newton's laws is impossible.

To note is, that quantum mechanical probabilities are not like classical probabilities. They are not a consequence of missing information. The accepted state-of-the art is that the probabilities are truly random, and there is no further information (so-called "hidden variables") that will make things un-random. The topic of hidden variables includes various theorems (such as Bell's theorem) and experimental results and will hopefully be revisited in later chapters of this script.

Random variables [2, 11]. To define a random variable x one needs to specify the range of values x can adopt, e.g. for the example of a tossed coin the values are HEADS and TAILS. Or for example, for casting a die, the sample space is $\{1, 2, 3, 4, 5, 6\}$. Each single outcome x_n is assigned a probability $p(x = x_n)$, where $p(x)$ is the probability distribution with

$$p(x) \geq 0, \quad \int dx p(x) = 1. \tag{1.9}$$

We interpret that $p(x)dx$ is the probability that the value of x lies between x and $x + dx$. The integral assumes that we have continuous values for x, but clearly we can have as well discrete outcomes. Taking that p_n is the probability of occurrence of x_n we have

$$p(x) = \sum_n p_n \delta(x - x_n) \quad \Rightarrow \quad \int dx p(x) = \sum_n p_n \int dx \delta(x - x_n) = \sum_n p_n = 1. \tag{1.10}$$

In physics probability distribution are of the visualized via an *ensemble*. This means that the probability distribution is replaced by a density distribution, where an arbitrary large number $\mathcal{N}$ of quantities is introduced, all having different values in a certain range, in such a way that the number of them having a value between x and $x + dx$ is $\mathcal{N}p(x)dx$.

Averages [2, 11]. The set of states and the probability distribution fully describe a random variable. However, to extract useful information for the description of the variable and it's distribution one requires as well other quantities. For example averages or expectation values, e.g. given a function $f(x)$ we have

$$\langle f(x) \rangle = \int dx f(x) p(x). \tag{1.11}$$

In particular we can introduce the n^{th} moment of x as

$$\mu_n = \langle x^n \rangle = \int dx \; x^n p(x) \tag{1.12}$$

which are a specific quantitative measure of the shape of a function, i.e., the first moment is the mean, the second moment is related to the width of the distribution, while the third moment is called the screwness (a measure of the asymmetry of the distribution). An important quantity is the variance, which can be interpreted as a measures how far a set of x is spread out from their average value:

$$\sigma^2 = \left\langle (x - \langle x \rangle)^2 \right\rangle = \left\langle x^2 \right\rangle - 2 \langle x \rangle^2 + \langle x \rangle^2 = \mu_2 - \mu_1^2 \tag{1.13}$$

with the standard deviation σ. A low standard deviation (or variance) indicates that the values tend to be close to the expectation value, while a high standard deviation indicates that the values are spread out over a wider range. Not all distributions have a variance, e.g. the Laury or Cauchy distribution have none.

Multivariate distributions [2, 11]. Besides distributions for single random variables one can as well define multivariate distributions involving multidimensional (or infinite dimensional) spaces. Let $x_1, x_2, ...x_r$ be random variables, one defines $p(x_1, x_2, ..., x_r)$ as the joint probability distribution. This is the probability that all the variables, $x_1, x_2, ..., x_r$ have exactly a specified value. To keep things compact we focus in the following on a bi-variate distribution, i.e., we consider two random variables x_1 and x_2 with joint probability distribution $p(x_1, x_2)$. It is useful to define the marginal probability as a partial integral,

$$p(x_1) = \int dx_2 p(x_1, x_2), \tag{1.14}$$

this is the probability that x_1 has a specified value, for any value of the remaining variable x_2. A conditional probability can be defined $p(x_1|x_2)$ which is the probability that x_1 has a specified value, given also a specified value of x_2, i.e., the probability for x_1 given the value x_2. Bayes Theorem asserts that

$$p(x_1, x_2) = p(x_1|x_2)p(x_2) = p(x_2|x_1)p(x_1). \tag{1.15}$$

With this we can introduce the concept of statistical independence. The random variable x_1 is statistically independent of x_2 if

$$p(x_1, x_2) = p(x_1)p(x_2), \quad p(x_1) = p(x_1|x_2) \tag{1.16}$$

that is, the joint probability factorizes and the distribution of x_1 is not affected by prescribing different values of x_2. The moments of the bi-variate distribution are defined as

$$\langle x_1^m x_2^n \rangle = \int dx_1 dx_2 \; x_1^m x_2^n \; p(x_1, x_2). \tag{1.17}$$

Of special interest here are the second moments. They can be combined in a 2×2 matrix called the covariance matrix $\mathbf{V}$, whose elements are defined as

$$V_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \,, \tag{1.18}$$

or explicitly written in matrix form

$$\mathbf{V} = \begin{pmatrix} \langle x_1^2 \rangle - \langle x_1 \rangle^2 & \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle \\ \langle x_2 x_1 \rangle - \langle x_2 \rangle \langle x_1 \rangle & \langle x_2^2 \rangle - \langle x_2 \rangle^2 \end{pmatrix} \tag{1.19}$$

from which we see that the diagonal elements are the individual variances of each variable. The off-diagonal elements are called covariances, if these are non-zero both variables are correlated, i.e., they define in some way the correlations as deviations from statistical independence. But if both variables are uncorrelated, i.e., the covariance is zero, is weaker than both variables being statistically independent. For the latter all moments have to factorize, i.e., $\langle x_1^m x_2^n \rangle = \langle x_1^m \rangle \langle x_2^n \rangle$ This all holds true as well for multivariate distributions for a set of random variables $x_1, x_2, ...x_r$.

1.2 The path integral perspective

In quantum mechanics, particles do have wave like properties and are described via a wave function, which is the solution to a particular wave equation: the Schrödinger equation. There is no way to derive this equation from first principles, however, there exist a rather elegant formalism from which the Schrödinger equation can be deducted: the path integral formalism. This formalism approaches quantum mechanics from the perspective that particles do not have trajectories, but rather take all paths simultaneously (in superposition). We will adopt this route in this script by trying to keep the mathematical complexities at a minimum, for further details please see the book by Feynmann and Hibbs [3].

1.2.1 The principle of least action

In classical mechanics we learned the three Newton's laws of motion, which describe the dynamics of a object under the influence of a force. The second of Newton's laws describes the change of motion under the influence of a force F. Taking for example a particle with mass m in a potential $V(x)$ the second law is captured in the relation

$$m\ddot{x} = -\frac{\partial V(x)}{\partial x} = F, \tag{1.20}$$

from which we can deduct the trajectory $x = x(t)$, i.e., the change of the position of the particle as a function of time. To derive the trajectory we also need to know the initial conditions of the particle: $x_0 = x(t_0)$ for the position, as well as $\dot{x}_0 = \dot{x}(t_0)$ for the velocity or momentum ($p = p(t) = m\dot{x}$). Taking for example a quadratic potential

$$V(x) = \frac{1}{2}m\omega^2 x^2 \quad \Rightarrow \quad m\ddot{x} = -m\omega^2 x, \tag{1.21}$$

a solution for this differential equation has the form

$$x(t) = X_1 \cos(\omega t) + X_2 \sin(\omega t), \tag{1.22}$$

where the coefficients $X_{1,2}$ are determined via the initial conditions

$$X_1 = \cos(\omega t_0)x_0 - \frac{\sin(\omega t_0)}{\omega}\dot{x}_0, \qquad X_2 = \sin(\omega t_0)x_0 + \frac{\cos(\omega t_0)}{\omega}\dot{x}_0. \tag{1.23}$$

There exists an alternative way to derive a trajectory in classical mechanics, something we will also refer to as the classical trajectory or equivalently the classical path. We stick with the example of a one-dimensional system involving a particle with mass m in a potential $V(x)$. In contrast to the Newtonian case above, we do not formulate a initial value problem, but a problem with boundary conditions. We consider the situation where the particle starts from a point A at t_0 and goes to a final point B at time T:

$$A: \ x(t_0) = x_A, \quad B: \ x(T) = x_B, \tag{1.24}$$

in principle we could imagine a infinite number of possible paths for the particle going from point A to point B, as illustrated in Fig.1.2. In classical mechanics we can deduct the actual

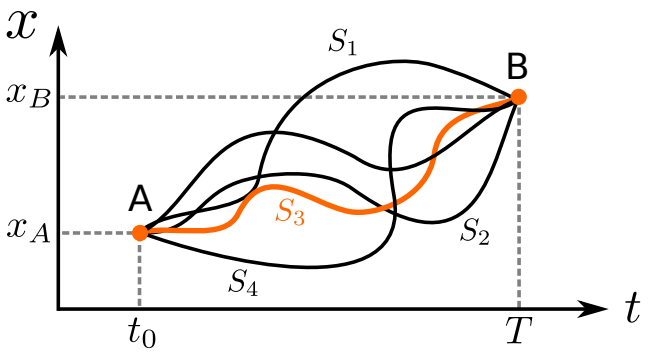

Figure 1.2: Multiple possible trajectories x between two fixed points A and B for a particle in a one-dimensional potential $V(x)$. All paths are weighted with an action S_n, a quantity which becomes maximal for the classical path, e.g. S_3.

trajectory out of all possible paths by using the *principle of least action*. Here we associate with each path x_n a quantity, the so-called action $S_n = S[x_n]$, and crucially, this quantity is extremal for the the classical trajectory. Note, the formulation *principle of least action* is a bit misleading here, as the action is not minimal but extremal, hence it is also referred to as stationary-action principle or Hamiltonian principle. The action of a path x for our particle is defined as

$$S[x] = \int_{t_0}^{T} dt\ \mathcal{L}(x, \dot{x}, t) = \int_{t_0}^{T} dt\ \left[\frac{1}{2}m\dot{x}^2 - V(x)\right], \tag{1.25}$$

where $\mathcal{L}(x, \dot{x}, t)$ is the Lagragian of the system, i.e., the difference of kinetic and potential energy. As mentioned above, the action becomes extremal for the classical trajectory. So let's remind us how we can find an extremum of an arbitrary function $f(x)$. At an extremum a function takes a maximal or minimal value (globally or locally), it can be obtained via setting the first derivative of the function to zero:

$$\begin{aligned}\frac{\partial f(x)}{\partial x} = 0, \quad \partial f(x) =& f(x + \partial x) - f(x) \\ =& f(x) + \frac{\partial f(x)}{\partial x}\partial x + \mathcal{O}\left[\partial x^2\right] - f(x) = \frac{\partial f(x)}{\partial x}\partial x + \mathcal{O}\left[\partial x^2\right],\end{aligned} \tag{1.26}$$

here we Taylor expanded the variation $\partial f(x)$ around small deviations δx and find that $\partial f(x) = 0$ up to first order in δx. We can transfer this to the action defined in Eq. (1.25) and state that the action becomes extremal if the variation (small change in the action)

$$\delta S = S(x + \delta x) - S(x) = 0 \tag{1.27}$$

vanishes up to first order in ∂x. In detail this gives us

$$\begin{aligned}\delta S =& \int_{t_0}^{T} dt \ \left[\frac{1}{2}m\left(\dot{x}+\delta\dot{x}\right)^2 - V(x+\delta x)\right] - S(x) \\ =& S(x) + \int_{t_0}^{T} dt \ \left[m\dot{x}\delta\dot{x} - \frac{\partial V(x)}{\partial x}\delta x\right] + \mathcal{O}\left[\delta x^2\right] - S(x) \\ =& m\dot{x}\delta x\Big|_{t_0}^{T} - \int_{t_0}^{T} dt \ \left[m\ddot{x} + \frac{\partial V(x)}{\partial x}\right]\delta x + \mathcal{O}\left[\delta x^2\right], \end{aligned} \tag{1.28}$$

where we performed a partial integration in the second step. We assumed that the boundary values of the position of our particle are fixed, hence $\delta x(t_0) = \delta x(T) = 0$ and the first term vanishes. We can conclude that the variation δS becomes extremal if

$$\delta S = 0 \quad \Rightarrow \quad m\ddot{x} = -\frac{\partial V(x)}{\partial x}, \tag{1.29}$$

which is nothing else than Newton's second law Eq (1.20). Meaning the classical path for which the action is extremal recovers the known Law of classical mechanics, we can also write this in a more general form [3]

$$\delta S = 0 \quad \Rightarrow \quad \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x} = 0, \tag{1.30}$$

which is nothing else than the Euler-Lagrange equation of classical mechanics. We can as well explicitly derive the action corresponding to the classical path, we will denote this action as S_{cl}. Taking for example a free particle, meaning that $V(x) = 0$. Here the principle of least action gives us the trajectory

$$m\ddot{x} = 0 \ \Rightarrow \ \dot{x} = const. = \frac{x_B - x_A}{T - t_0} \ \Rightarrow \ x = \frac{x_B - x_A}{T - t_0}t + x(t=0). \tag{1.31}$$

The corresponding classical action yields

$$S_{\text{cl}} = \frac{1}{2}m\int_{t_0}^{T} dt\ \dot{x}^2 = \frac{1}{2}m\int_{t_0}^{T} dt\ \left[\frac{x_B - x_A}{T - t_0}\right]^2 = \frac{1}{2}m\frac{(x_B - x_A)^2}{T - t_0}. \tag{1.32}$$

1.2.2 Feynman's path integral [3]

In quantum mechanics things are treated differently, as a particle can take every path from point A to point B. Each path is associated with a probability amplitude, and the question is how each trajectory contributes to the total amplitude. In contrast to the classical case, where only the path with extremal action contributes, all paths contribute. An they contribute equal amounts, but, crucially, different phases. The phase of each contribution is the action S in units of the quantum of action $\hbar$. Meaning that the probability $P(x_B, x_A)$ of going from the

point x_A at time t_0 to x_B at time T is the absolute square of an amplitude $\mathcal{K}(x_B, x_A)$ for going from A to B:

$$P(x_A, x_B) = |\mathcal{K}(x_B, x_A)|^2 , \quad \mathcal{K}(x_B, x_A) \propto \sum_{\text{all paths}} e^{\frac{i}{\hbar} S[x(t)]}, \tag{1.33}$$

where the amplitude is a sum over all paths. As mentioned above, the contribution of each path has a phase proportional to the action $S[x(t)]$. The action is calculated in the same way as for the classical case, see Eq. (1.25).

The classical limit. To understand more clearly why the definition of the probability amplitude is a meaningful formulation, we consider first the classical limit. It is not right away obvious why in the classical limit one specific path becomes most important. All paths contribute equally, only their phases vary. Importantly, the classical case can be recovered, when we first of all notice that in this limit all quantities, e.g. masses, times, dimensions, are so large that the classical action is enormous in relation to Planck's constant $\hbar$. Meaning that the contribution $s/\hbar$ is some very large angle, recall that we can use Euler's formula to express the phase in terms of real and imaginary part

$$e^{\frac{i}{\hbar} S[x(t)]} = \cos\left(\frac{S[x(t)]}{\hbar}\right) + i \sin\left(\frac{S[x(t)]}{\hbar}\right), \tag{1.34}$$

which tells us that the real and imaginary part can each be between plus and minus one. If we now move the path by a small amount of δx_n the change of the action is minimal on a classical level, but not when it set into relation to $\hbar$. The small change on the path will make generally enormous changes in the phase, and the cosine and sin will oscillate rapidly between plus and minus values. Thus, the total contribution will add up to zero; meaning that if one path makes a positive contribution, another infinitesimally close one (on the classical scale) makes an equal negative contribution. In total no contribution arises. Importantly, we learned from the principle of least action, that the classical path is an extremum, meaning that a small change δx will to first order produce no change in the action $S[x(t)]$ and hence, in the phase. Only the classical path, and paths very close in vicinity of it will contribute, while all other paths cancel out. In conclusion, in the classical limit only one trajectory will deliver the main contribution. So quiet naturally, the classical laws of motion arise in this limit $S[x(t)] \gg \hbar$. In the regime where quantum mechanics matters, e.g. for small masses, dimensions etc., the action S is small and may even be comparable to $\hbar$. Here we cannot pick out one particular trajectory and all paths have to be added together. The physical picture behind the path integral approach is that one can achieve a correct reformulation of quantum mechanics in terms of interfering classical paths.

The sum over paths. Qualitatively the idea of summing up all possible paths is quiet intuitive, however, the mathematical definition of such a summation has to be defined clearly. We will not do this here in detail, just sketch the general procedure. The basic idea is to discretize time and slice the continuous time interval from t_0 to T into a large set of discrete times:

$$(t_0, t_1, t_2, ..., t_{N-1}, T), \quad t_n = t_0 + n\epsilon, \quad \epsilon = \frac{T - t_0}{N}; \tag{1.35}$$

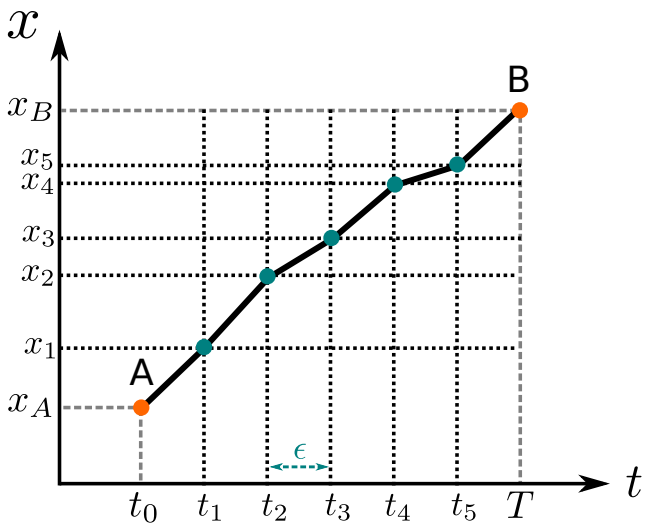

Figure 1.3: Constructing the summation over all paths.

accordingly we can define corresponding values for x

$$(x_A, x_1, x_2, ..., x_{N-1}, x_B), \quad x_n = x(t_n) \text{ with } x_A = x(t_0),\ x_B = x(T), \tag{1.36}$$

a single path can now be constructed connecting the selected point, see Fig.1.3. It is now possible to construct the sum over all paths by integrating over each x_n. We do not integrate over x_A to x_B as they are our boundary conditions and thus fixed. The path integral is then obtained via taking the limit $\epsilon \to 0$ or equivalently, $N \to \infty$. We can define the integration measure

$$\int \mathcal{D}x = \lim_{N\to\infty} \frac{1}{\mathcal{N}} \int \prod_{j=1}^{N-1} dx_j = \lim_{N\to\infty} \frac{1}{\mathcal{N}} \iiint \dots dx_1 dx_2 dx_3 \dots, \tag{1.37}$$

with $\mathcal{N}$ denoting a normalization factor which we define more clearly later. With this we can find a very compact formulation for the path integral

$$\mathcal{K}(x_B, x_A) = \int \mathcal{D}x\ e^{\frac{i}{\hbar}S[x(t)]}, \quad S[x(t)] = \int_{t_0}^{T} dt \mathcal{L}\left(\dot{x}, x, t\right). \tag{1.38}$$

here the action is a line integral taken over a trajectory passing through the points x_n with straight sections in between, see Fig. 1.3. Meaning the reverse approximation of the integral yields

$$\int_{t_0}^{T} dt \mathcal{L}\left(\dot{x}, x, t\right) \to \epsilon \sum_{j=0}^{N-1} \mathcal{L}\left(\frac{x_{j+1} - x_j}{\epsilon}, \frac{x_{j+1} + x_j}{2}\right), \tag{1.39}$$

note that in this definition we have the initial conditions appearing: $x_0 = x_A$ and $x_N = x_B$, due to the dependence of the Lagrangian on the velocity $\dot{x}$.

1.2.3 The free particle [3]

Let us now consider the example of a free particle, i.e., $V(x) = 0$. Based on our definition of the path integral in Eq. (1.38) we obtain

$$\mathcal{K}(x_B, x_A) = \int \mathcal{D}x \, \exp\left[i \frac{\epsilon m}{2\hbar} \sum_{j=0}^{N-1} \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 \right] = \int \mathcal{D}x \, \prod_{j=0}^{N-1} e^{i \frac{m}{2\hbar\epsilon}(x_{j+1} - x_j)^2}, \tag{1.40}$$

which represents a set of Gaussian integrals which we can exactly evaluate knowing that the integral of a Gaussian with a complex argument is (for real $a > 0$)

$$\int dx \, e^{iax^2} \, e^{ibx} = e^{-i\frac{b^2}{4a}} \sqrt{\frac{i\pi}{a}} \tag{1.41}$$

we can perform this integration step by step:

$$\begin{aligned} \int dx_1 \, e^{i\frac{m}{2\hbar\epsilon}(x_0 - x_1)^2} e^{i\frac{m}{2\hbar\epsilon}(x_2 - x_1)^2} &= \int dx_1 \, e^{i\frac{m}{\hbar\epsilon}\left[x_1^2 - x_1(x_0 + x_2)\right]} \, e^{i\frac{m}{2\hbar\epsilon}\left(x_0^2 + x_2^2\right)} \\ &= \sqrt{\frac{i\pi\hbar\epsilon}{m}} \, e^{-i\frac{m}{4\hbar\epsilon}(x_0 + x_2)^2} \, e^{i\frac{m}{2\hbar\epsilon}\left(x_0^2 + x_2^2\right)} = \sqrt{\frac{2\pi i\hbar\epsilon}{m}\frac{1}{2}} \, e^{i\frac{m}{2\hbar 2\epsilon}(x_0 - x_2)^2} \end{aligned} \tag{1.42}$$

now we can perform the next integration

$$\int dx_2 \;\; e^{i\frac{m}{4\hbar\epsilon}(x_0 - x_2)^2} e^{i\frac{m}{2\hbar\epsilon}(x_3 - x_2)^2} = \sqrt{\frac{2\pi i\hbar\epsilon}{m}\frac{2}{3}} \, e^{i\frac{m}{2\hbar 3\epsilon}(x_0 - x_3)^2} \tag{1.43}$$

form here we can see that this integration just continues on, crucially, the integral after n steps will have the form

$$\int dx_n \, e^{i\frac{m}{2\hbar n\epsilon}(x_0 - x_n)^2} e^{i\frac{m}{2\hbar\epsilon}(x_{n+1} - x_n)^2} = \sqrt{\frac{2\pi i\hbar\epsilon}{m}\frac{n}{n+1}} e^{i\frac{m}{2\hbar(n+1)\epsilon}(x_0 - x_{n+1})^2}. \tag{1.44}$$

hence, if we perform $N - 1$ steps we arrive at the path integral

$$\mathcal{K}(x_B, x_A) = \frac{1}{\mathcal{N}} \left[\frac{2\pi i\hbar\epsilon}{m} \right]^{\frac{N-1}{2}} \sqrt{\frac{(N-1)!}{N!}} e^{i\frac{m}{2\hbar N\epsilon}(x_0 - x_N)^2} = \frac{1}{\mathcal{N}} \left[\frac{2\pi i\hbar\epsilon}{m} \right]^{\frac{N-1}{2}} \frac{1}{\sqrt{N}} e^{i\frac{m}{2\hbar}\frac{(x_A - x_B)^2}{(T - t_0)}} \tag{1.45}$$

where we have used $\epsilon N = T - t_0$ is the second step. This is the probability amplitude for a free particle to move from point A to point B. Now we can try to obtain the normalization factor $\mathcal{N}$, the condition which has to be fulfilled is [12]:

$$\lim_{T \to t_0} \mathcal{K}(x_B, x_A) = \delta(x_B - x_A), \tag{1.46}$$

which is connected to the orthonormality in position and momentum space (we will learn more about that later on). Defining now $y = x_A - x_B$ we have

$$\int dy \, e^{i\frac{m}{2\hbar}\frac{y^2}{(T - t_0)}} = \sqrt{\frac{2\pi i\hbar(T - t_0)}{m}} \quad \Rightarrow \quad \lim_{T \to t_0} \sqrt{\frac{m}{2\pi i\hbar(T - t_0)}} e^{i\frac{m}{2\hbar}\frac{y^2}{(T - t_0)}} = \delta(y), \tag{1.47}$$

as expected a properly normalized Gaussian becomes a delta-function. Hence we know now the desired form of the prefactor in our transition amplitude and in combination with Eq. 1.45 we can deduct the normalization factor

$$\mathcal{N} = \sqrt{\frac{2\pi i\hbar(T-t_0)}{m}}\left[\frac{2\pi i\hbar\epsilon}{m}\right]^{\frac{N-1}{2}}\frac{1}{\sqrt{N}} = \sqrt{\frac{2\pi i\hbar\epsilon N}{mN}}\left[\frac{2\pi i\hbar\epsilon}{m}\right]^{\frac{N-1}{2}} = \left[\frac{2\pi i\hbar\epsilon}{m}\right]^{\frac{N}{2}}, \tag{1.48}$$

which is the normalization factor defined in the integration measure Eq. (1.37) for the free particle. General normalization of path integrals is a topic on it's own and is beyond the scope of this course.
Interestingly, with the path integral for the free particle we can already make some connection to the basis relations of quantum mechanics. For this we consider the case where the point A is the origin, i.e., $t_0 = 0$, $x_A = 0$ and $T = t$, $x_B = x$, which simplifies our expression to

$$\mathcal{K}(x,0) = \sqrt{\frac{m}{2i\pi\hbar t}}\, e^{i\frac{m}{2\hbar t}x^2}, \tag{1.49}$$

for a fixed time t the amplitude varies with distance. The farther one gets away from the origin the oscillations become more and more rapid. If x is so large the distance between nodes is nearly constant, at least for the next few oscillations. Hence, the amplitude behaves much like a sine wave of slowly varying wavelength. We can calculate this wavelength λ. For this if we change x by λ the phase of the amplitude should change by 2π, that is

$$2\pi = \frac{m}{2\hbar t}\left[(x+\lambda)^2 - x^2\right] = \frac{m}{2\hbar t}\left[2x\lambda + \lambda^2\right], \tag{1.50}$$

assuming now that $x \gg \lambda$ we can neglect the last term and obtain

$$2\pi = \frac{m}{\hbar t}x\lambda \quad \Rightarrow \quad \lambda = \frac{2\pi\hbar}{m\frac{x}{t}}. \tag{1.51}$$

From the classical point of view a particle moves from the origin to the point x in a time interval t and has the velocity $v = x/t$ corresponding to the momentum $p = mx/t$. From the quantum mechanical point of view, the probability amplitude varies in space with the wavelength

$$\lambda = \frac{2\pi\hbar}{m\frac{x}{t}} = \frac{h}{p}, \tag{1.52}$$

which is nothing else than the deBroglie relation we have discussed in the section before.
Next we can study the time-dependence of the kernel, while keeping the distance x fixed. Now both, phase and amplitude change with time t. Let's focus only on the change in phase and define the time $T_{2\pi}$ as the time to increase the phase by 2π. This means

$$2\pi = \frac{mx^2}{2\hbar}\left[\frac{1}{t} - \frac{1}{(t+T_{2\pi})}\right] = \frac{mx^2}{2\hbar}\left[\frac{T_{2\pi}}{t(t+T_{2\pi})}\right] = \frac{mx^2}{2t^2\hbar}\left[\frac{T_{2\pi}}{(1+T_{2\pi}/t)}\right]. \tag{1.53}$$

We now assume that $t \gg T_{2\pi}$ and use the definition for the angular frequency $\omega = 2\pi/T_{2\pi}$ and obtain

$$2\pi \backsimeq \frac{mx^2}{2t^2\hbar}\left[\frac{2\pi}{\omega}\right] \quad \Rightarrow \quad \omega\hbar \backsimeq \frac{m}{2}\left(\frac{x}{t}\right)^2 \tag{1.54}$$

associating the right hand side with the classical energy of the particle, we recover $E = h\nu$.

1.2.4 Deducting the Schrödinger equation [3]

Now we finally want to derive the Schrödinger equation on the basis of the path integral approach. The Schrödinger equation is a wave equation and describes the time-evolution of a wave function. So first we should connect the concept of a path integral to the notion of a wave function. So far we have derived the path integral for a particle moving from a specific point A to a specific point B. We associated the path integral with an amplitude $\mathcal{K}(x_B, x_A)$ providing us the probability for going from A to B. We can also loosen up this restriction a bit and ask the question what is the probability for finding a particle at a point x and a time t. Here we can introduce the notion of a wave function $\psi(x,t)$ as the amplitude to arrive at x at time t, associated with the probability $|\psi(x,t)|^2$. In the end we just have to change the notation here to connect the wave function to the path integral:

$$\mathcal{K}(x_B, x_A) = \mathcal{K}(x_B, t_B; x_A, t_A) = \psi(x_B, t), \tag{1.55}$$

as $\mathcal{K}(x_B, x_A)$ is the amplitude to arrive at point B. When the information of where the particle came from, i.e., point A, is not important we can simply use the wave function notation.
As we aim to derive an equation for the evolution of the wave function in time we have to first understand how we can model successive events within the path integral formalism. We can start with the simple example of two events occurring in succession: instead of moving directly from a point A to B we construct our path via another point C. The action along the path is additive, and thus we can straightforwardly express the amplitude from going from point A to point B via

$$S[x_B, x_A] = S[x_B, x_C] + S[x_C, x_A], \quad \mathcal{K}(x_B, x_A) = \int \mathcal{D}x \; e^{\frac{i}{\hbar}S[x_B, x_C] + \frac{i}{\hbar}S[x_C, x_A]}, \tag{1.56}$$

thus it is possible to split any path into two (or more) paths. Hence we can integrate now from A to C resulting in the amplitude $\mathcal{K}(x_C, x_A)$, followed by the integration from C to B, resulting in the amplitude $\mathcal{K}(x_B, x_C)$. The total amplitude is then obtained via performing the integral over x_C, i.e.,

$$\mathcal{K}(x_B, x_A) = \int dx_C \int \mathcal{D}x \; e^{\frac{i}{\hbar}S[x_B, x_C]} \mathcal{K}(x_C, x_A) = \int dx_C \; \mathcal{K}(x_B, x_C)\mathcal{K}(x_C, x_A). \tag{1.57}$$

We can conclude from this, that amplitudes for events occurring in succession in time multiply. We can as well split our path in more intervals and can deduct a general expression for the amplitude

$$\mathcal{K}(x_B, x_A) = \iint \cdots \iint \cdots dx_1 \cdots dx_i \cdots dx_{N-1} \; \mathcal{K}(x_B, x_{N-1}) \cdots \mathcal{K}(x_{j+1}, x_j) \cdots \mathcal{K}(x_1, x_A), \tag{1.58}$$

which is an alternative formulation for the probability amplitude. Crucially, we can extract a very useful relation from this

$$\mathcal{K}(x_{j+1}, x_j) = \frac{1}{\mathcal{A}} \exp\left[\epsilon \frac{i}{\hbar} \mathcal{L}\left(\frac{x_{j+1} - x_j}{\epsilon}, \frac{x_{j+1} + x_j}{2}\right)\right], \quad \mathcal{A} = \mathcal{N}^{\frac{1}{N}}, \tag{1.59}$$

which corresponds to the the amplitude for a particle to go between two points separated by an infinitesimal time interval ϵ (correct to first order in ϵ).

Moreover, we can rewrite this with the wave function we defined above (ignoring the initial point A)

$$\psi(x_B, t_B) = \int dx_C \, \mathcal{K}(x_B, x_C)\psi(x_C, t_C). \tag{1.60}$$

This form is already familiar to us. From classical probability we now that the marginal probability for a bi-variate probability distribution $p(x_1, p_2)$ is obtained via

$$p(x_1) = \int dx_2 \, p(x_1, x_2) = \int dx_2 \, p(x_1|x_2)p(x_2) \tag{1.61}$$

which is the probability to obtain x_1 for any value of x_2, where we used Bayes theorem in the second step. This can be right away transferred to the expression $\psi(x_B, t_B)$ giving us the probability to arrive at point B for all values x_C.
In other words the wave function $\psi(x_B, t_B)$ satisfies the integral equation Eq. (1.60). The total amplitude or probability to arrive at B is obtained for all possible values of x_C. Hence, we can express all effects of the past history of a particle as a single function. To now derive the wave equation for this wave function we assume that the time t_B differs only slightly from t_C. Meaning we set

$$t_B = t + \epsilon, \quad t_C = t, \quad x_B = x \quad x_C = y, \tag{1.62}$$

and obtain for the integral function for the wave function

$$\psi(x, t+\epsilon) = \int dy \, \mathcal{K}(x, y) \, \psi(y, t) = \int dy \, \frac{1}{\mathcal{A}} \exp\left[\epsilon \frac{i}{\hbar} \mathcal{L}\left(\frac{x-y}{\epsilon}, \frac{x+y}{2}\right)\right] \, \psi(y, t) \tag{1.63}$$

where we have used Eq.(1.59) in the second step, i.e., approximating to first order in ϵ. We now consider the one-dimensional case of a particle moving in a potential $V(x)$ and obtain

$$\psi(x, t+\epsilon) = \int dy \, \frac{1}{\mathcal{A}} \exp\left[\frac{i}{\hbar}\frac{m}{2}\frac{(x-y)^2}{\epsilon}\right] \, \exp\left[-\epsilon \frac{i}{\hbar} V\left(\frac{x+y}{2}\right)\right] \, \psi(y, t). \tag{1.64}$$

We are integrating over all values of y, and the quantity $(x-y)^2/\epsilon$ appearing in the first term becomes rather large when $y \gg x$. Meaning that exponential will vary rapidly (similar argument as we did for the classical limit). As a consequence, assuming the factor oscillates rapidly the integral over y gives a very small value. Only when we are close to x and the exponential changes slowly, we will obtain a reasonable contribution from the integral. Hence, we can make a substitution $y = x + \eta$, and expect that only for small η we will get important contributions. This yields

$$\psi(x, t+\epsilon) = \int d\eta \, \frac{1}{\mathcal{A}} \exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right] \, \exp\left[-\epsilon \frac{i}{\hbar} V\left(x + \frac{\eta}{2}\right)\right] \, \psi(x+\eta, t), \tag{1.65}$$

we expect to obtain relevant contributions for η in the order of $\sqrt{\epsilon\hbar/m}$. We now expand ψ in a power series keeping only term of order ϵ or equivalently second order in η. We do this expansion step by step, for the left side we obtain

$$\psi(x, t+\epsilon) = \psi(x, t) + \epsilon \frac{\partial \psi(x, t)}{\partial t} + \mathcal{O}[\epsilon^2] \tag{1.66}$$

and on the right hand side

$$\psi(x+\eta,t) = \psi(x,t) + \eta\frac{\partial\psi(x,t)}{\partial x} + \frac{\eta^2}{2}\frac{\partial^2\psi(x,t)}{\partial^2 x} + \mathcal{O}[\eta^3], \tag{1.67}$$

and for the potential we can approximate

$$\exp\left[-\epsilon\frac{i}{\hbar}V\left(x+\frac{\eta}{2}\right)\right] = 1 - \epsilon\frac{i}{\hbar}V(x) + \mathcal{O}[\eta\epsilon]. \tag{1.68}$$

Plugging this all together we obtain to leading order in ϵ

$$\begin{aligned}\psi(x,t) + \epsilon\frac{\partial\psi(x,t)}{\partial t} =& \left\{\psi(x,t) - \epsilon\frac{i}{\hbar}V(x)\ \psi(x,t)\right\}\ \int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right] \\ &+ \left\{\frac{\partial\psi(x,t)}{\partial x} - \epsilon\frac{i}{\hbar}V(x)\frac{\partial\psi(x,t)}{\partial x}\right\}\int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right]\ \eta \\ &+ \frac{1}{2}\frac{\partial^2\psi(x,t)}{\partial^2 x}\ \int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right]\ \eta^2.\end{aligned} \tag{1.69}$$

Thus, we need a range of integrals:

$$\int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right] = \frac{1}{\mathcal{A}}\sqrt{\frac{2\pi i\hbar\epsilon}{m}} = 1 \tag{1.70}$$

here we used our previous definition for the normalization, in addition we have

$$\begin{aligned}&\int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right]\ \eta = 0 \\ &\int d\eta\ \frac{1}{\mathcal{A}}\exp\left[i\frac{m\eta^2}{2\hbar\epsilon}\right]\ \eta^2 = \frac{1}{\mathcal{A}}\sqrt{\frac{\pi}{4}},\left[\frac{2i\hbar\epsilon}{m}\right]^{\frac{3}{2}}\ =\ \frac{i\hbar\epsilon}{m},\end{aligned} \tag{1.71}$$

hence, we obtain

$$\psi(x,t) + \epsilon\frac{\partial\psi(x,t)}{\partial t} = \psi(x,t) - \epsilon\frac{i}{\hbar}V(x)\ \psi(x,t)\ + \frac{i\hbar\epsilon}{2m}\frac{\partial^2\psi(x,t)}{\partial^2 x}, \tag{1.72}$$

this will give us to order ϵ

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial^2 x} + V(x)\ \psi(x,t), \tag{1.73}$$

which finally is the Schrödinger equation we were aiming to derive. We can as well generalize this equation to the three-dimensional case:

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = \left[-\frac{\hbar^2\Delta}{2m} + V(\mathbf{x})\right]\Psi(\mathbf{x},t) \tag{1.74}$$

with the Laplace operator $\Delta = \partial^2/\partial^2 x + \partial^2/\partial^2 y + \partial^2/\partial^2 z$, which also sometimes written as the divergence of the gradient $\Delta = \nabla\cdot\nabla$ with $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ (also called *nabla* operator).

1.3 The time-dependent Schrödinger Equation

1.3.1 Probability and continuity [4, 5]

The Schrödinger equation Eq.(1.74) corresponds to the equation of motion for the wavefunction of a particle with mass m moving in the potential $V(\mathbf{x})$. The Schrödinger equation is a linear equation, thus $\lambda\psi(x,t)$ is a solution if $\psi(x,t)$ is a solution, where λ is a time- and coordinate-independent complex number. Moreover, we have $|\Psi(\mathbf{x},t)|^2 d\mathbf{x}$ as the probability to find the particle at time t in the tiny volume $d\mathbf{x} = dxdydz$. The probability to find the particle in a specific volume Ω is given as

$$P(\Omega) = \int_\Omega d\mathbf{x}\, |\Psi(\mathbf{x},t)|^2. \tag{1.75}$$

Hence, we can define a probability density

$$\rho(\mathbf{x},t) = |\Psi(\mathbf{x},t)|^2 = \Psi(\mathbf{x},t)\Psi^*(\mathbf{x},t), \tag{1.76}$$

and to have a valid probability density we need to ensure normalization:

$$\int d\mathbf{x}\, |\Psi(\mathbf{x},t)|^2 = 1. \tag{1.77}$$

where the integral is now over the total three-dimensional space. Note, we work here with a single particle, but one can straightforwardly expand this to N particles. As we are working with a probability density we can calculate averages in the same fashion as before, e.g. for the position we have

$$\langle \mathbf{x} \rangle = \int d\mathbf{x}\; \mathbf{x}\; \rho(\mathbf{x},t) = \int d\mathbf{x}\; \Psi^*(\mathbf{x},t)\mathbf{x}\; \Psi(\mathbf{x},t). \tag{1.78}$$

It is important to verify that the probability is conserved when we work with a probability density. A relevant check here is to derive a continuity equation. For this we consider again the simplest case of a free particle in 1 dimension. Starting from the equation of motion for the probability density:

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\rho(x,t) =& i\hbar\left(\frac{\partial\psi(x,t)}{\partial t}\right)\psi^*(x,t) + i\hbar\psi(x,t)\left(\frac{\partial\psi^*(x,t)}{\partial t}\right) \\
=& -\frac{\hbar^2}{2m}\left[\frac{\partial^2\psi(x,t)}{\partial^2 x}\psi^*(x,t) - \psi(x,t)\frac{\partial^2\psi^*(x,t)}{\partial^2 x}\right] \\
=& -\frac{\hbar^2}{2m}\left[\frac{\partial}{\partial x}\left(\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial x}\right) - \frac{\partial}{\partial x}\left(\psi(x,t)\frac{\partial\psi^*(x,t)}{\partial x}\right)\right] \\
=& -\frac{\hbar^2}{2m}\frac{\partial}{\partial x}\left[\psi^*(x,t)\frac{\partial\psi(x,t)}{\partial x} - \psi(x,t)\frac{\partial\psi^*(x,t)}{\partial x}\right] \equiv -i\hbar\,\frac{\partial}{\partial x}\,j(x,t),
\end{aligned} \tag{1.79}$$

where we have used $\partial_x[\psi\partial_x\psi^*] = \partial_x\psi\partial_x\psi^* + \psi\partial_x^2\psi^*$in the second step. Moreover, we have defined $j(x,t)$ denoting a probability current. We can now compactly write

$$\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}j(x,t) = 0 \tag{1.80}$$

which corresponds to a 1-dimensional continuity equation.

1.3.2 Solution for a free particle in 1D [4]

We now want to discuss the solution of the Schrödinger equation for the simple case of a free particle. We start from the initial value problem

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial^2 x}, \quad \psi(x,0)=\psi_0(x). \tag{1.81}$$

We can solve such an equation using the Fourier-transformation. We transform both sides of our dynamical equation to reciprocal space

$$i\hbar\frac{\partial}{\partial t}\int\frac{dk}{2\pi}\,\psi(k,t)\,e^{ikx} = -\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial^2 x}\int\frac{dk}{2\pi}\,\psi(k,t)\,e^{ikx} = -\frac{\hbar^2}{2m}\int\frac{dk}{2\pi}\,(ik)^2\psi(k,t)\,e^{ikx}, \tag{1.82}$$

which we can re-write as

$$\int\frac{dk}{2\pi}\,e^{ikx}\,\left[i\hbar\frac{\partial}{\partial t}\,\psi(k,t) - \frac{\hbar^2k^2}{2m}\,\psi(k,t)\right] = 0, \tag{1.83}$$

the expression in the brackets is a simple one-dimensional differential equation in time which we can solve straightforwardly.

$$\psi(k,t) = \psi_0(k)\,e^{-i\omega(k)t} = \psi_0(k)\,e^{-\frac{i}{\hbar}Et} \quad \omega(k) = \frac{\hbar k^2}{2m}, \tag{1.84}$$

where we have $\omega(k)$ as the dispersion relation, and used the deBroglie relation $p=\hbar k$ together with $E=\hbar\omega(k)=p^2/2m$, connecting the momentum of the particle and the wavevector. To obtain the wave function in position space we simply have to transform back

$$\psi(x,t) = \int\frac{dk}{2\pi}\,\psi_0(k)\,e^{-i\omega(k)t}\,e^{ixk} = \int\frac{dk}{2\pi}\,\psi_0(k)\,e^{-i\frac{\hbar k^2}{2m}t}\,e^{ixk}, \tag{1.85}$$

thus, to get the explicit expression we need the initial condition $\psi_0(k)$ in reciprocal space. Within the path integral formalism we normalized our probability amplitude going from a point A to B by considering the limit of going back to the origin $B\to A$. The latter limit is a delta function, hence, we take our initial condition as

$$\psi_0(x)=\delta(x) \quad \Rightarrow \quad \psi_0(k) = \int dx\,\psi_0(x)\,e^{-ixk} = \int dx\,\delta(x)\,e^{-ixk} = 1 \tag{1.86}$$

and transform our wave function back to position space

$$\psi(x,t) = \int\frac{dk}{2\pi}\,e^{-i\frac{\hbar k^2}{2m}t}\,e^{ixk} = \frac{1}{2\pi}\,e^{i\frac{mx^2}{2\hbar t}}\,\sqrt{\frac{2m\pi}{i\hbar t}} = \sqrt{\frac{m}{2i\pi\hbar t}}e^{i\frac{m}{2\hbar t}x^2} = \mathcal{K}(x,0), \tag{1.87}$$

where we have used the formula for the integral $\int dx\,e^{-iax^2}\,e^{ibx} = e^{i\frac{b^2}{4a}}\,\sqrt{\frac{\pi}{ia}}$ with $a=\frac{\hbar t}{2m}$ and $b=x$. Hence, we recover here the probability amplitude $\mathcal{K}(x,0)$, denoting the path starting at the origin to the point x at time t, which we derived in the last section.

1.4 The stationary Schrödinger Equation

1.4.1 Separation ansatz [4, 5]

In case of an arbitrary time-independent potential, the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{x},t) = \left[-\frac{\hbar^2\Delta}{2m} + V(\mathbf{x})\right]\Psi(\mathbf{x},t) \equiv \hat{H}\Psi(\mathbf{x},t), \tag{1.88}$$

can be simplified. Here we defined $\hat{H}$ as an operation acting on our wavefunction, we call this an operator the Hamiltonian operator. To indicate that this is not a number but an operator, we use the notation with the *hat*. We make the separation ansatz

$$\psi(\mathbf{x},t) = \phi(\mathbf{x})\ f(t), \tag{1.89}$$

where $\phi(\mathbf{x})$ describes now the spacial propagation and $f(t)$ the temporal evolution. We can insert this expression into the Schrödinger equation and obtain

$$i\hbar\phi(\mathbf{x})\ \frac{\partial}{\partial t}\ f(t) = \left[\hat{H}\phi(\mathbf{x})\right]\ f(t), \quad \Rightarrow \quad \frac{i\hbar}{f(t)}\ \frac{\partial f(t)}{\partial t} = \frac{\hat{H}\phi(\mathbf{x})}{\phi(\mathbf{x})} = const., \tag{1.90}$$

Since a purely time dependent function on the l.h.s. equals a purely space dependent function on the r.h.s., both can only be equal a time and space independent constant. We will identify this constant with E and can solve the time-dependent part of the equation right away

$$\frac{\partial f(t)}{\partial t} = -\frac{i}{\hbar}\ E\ f(t) \quad \Rightarrow \quad f(t) = e^{-\frac{i}{\hbar}Et}, \tag{1.91}$$

and thus the time dependence of the wave function is only a phase factor and the probability distribution $\rho(\mathbf{x},t) = |\Psi(\mathbf{x},t)|^2 = |\phi(\mathbf{x})|^2$ is independent of time.
The equation for the spacial evolution is called the stationary Schrödinger equation

$$\hat{H}\phi(\mathbf{x}) = E\ \phi(\mathbf{x}), \tag{1.92}$$

which is mathematically nothing else than an eigenvalue problem, with the constant E as the eigenvalue of the Hamilton operator. The solution to Eq.(1.92) are called stationary states and we will interpret E as the corresponding energy. We will find that stationary states do often not exist for arbitrary values of E, but that only specific eigenvalues E are possible. This is referred to as energy quantization.

1.4.2 Example: plane waves [4]

Let's start again with the simplest example of a free particle, the stationary Schrödinger equation Eq.(1.92) has two independent solutions

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi(x) = E\phi(x) \quad \Rightarrow \quad \phi_k(x) = e^{ikx}, \ \ k = \pm\frac{\sqrt{2mE}}{\hbar}, \tag{1.93}$$

for positive energy $E > 0$. The solution describes plane waves with fixed wavevector k and fixed impulse $p = \hbar k$. Note, for the case of plane waves of a free particle there is no meaningful physical solution for negative energies, i.e., the exponent would be real and the wavefunction would diverge for $x \to \pm\infty$.

One issue emerges here, the solution to the stationary Schrödinger equation cannot be normalized as required, i.e., the integral over the total x-axis $\int dx \phi_{\pm}^*(x)\phi_{\pm}(x) = \infty$. The practical solution one takes is that one considers a large but finite interval $[-L/2, L/2]$ instead of the whole x-axis, i.e. the wavefunction

$$\phi_k(x) = \frac{1}{\sqrt{L}}\, e^{ikx} \quad \Rightarrow \quad \int_{-\frac{L}{2}}^{\frac{L}{2}} dx\ |\phi_k(x)|^2 = \frac{1}{L} x \Big|_{-\frac{L}{2}}^{\frac{L}{2}} = 1, \tag{1.94}$$

is then normalized. In a next step, one can close the interval to a ring and demand periodic boundary conditions. For the wavefunction to be continuous these boundary conditions impose conditions on the values k:

$$\psi_k(L/2) = \psi_k(-L/2) \quad \Rightarrow \quad e^{ikL} = 1 \quad \Rightarrow \quad k_n = \frac{2\pi n}{L}, \ \ n = 0, \pm 1, \pm 2, ..., \tag{1.95}$$

thus, this leads to a quantization for the values of k and in turn to the quantization of the energy $E_n = \hbar^2 k_n^2/2m$. For every energy $E_n > 0$ exist two independent solutions with the wavevectors $\pm 2\pi n/L$. We say that the energy is two-fold degenerate. When L is very large the quantisation still holds, but the distance between wavevectors becomes smaller and smaller. In the relevant limit $L \to \infty$ one is left with a continuum of energy values as expected.

1.4.3 Example: constant potential [4]

Next we can consider the example of a particle in a constant potential $V(x) = V$. Here the solution for the stationary Schrödinger equation is only slightly modified

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi(x) = (E - V)\phi(x) \quad \Rightarrow \quad \phi_{\pm}(x) = e^{\pm ikx}, \ \ k = \frac{\sqrt{2m(E-V)}}{\hbar}, \tag{1.96}$$

the crucial aspect is now that the wavevector k can become complex for $E < V$, here we have to distinguish three cases:

1. For $E > V$, k is real and the wavefunctions $\phi_k(x)$ describe plane waves which run in positive and negative x-direction (depending on the sign of k). These solutions are called oscillatory solutions. We can also think of this as the classical permitted region, as classically a movement (a change in position) is only possible when the kinetic energy is larger than the potential energy.

2. For $E < V$, k is complex and the wavefunction has a real exponent. Here we are in the classical forbidden region, however, quantenmechanically a particle can have a finite probability to be in this region (tunnel effect).

3. For $E = V$, the second derivative of the wavefunction is zero and we obtain a linear form for the wavefunction $\phi(x) = a + bx$.

Overall, the solution of the stationary Schrödinger equation for $E \neq V$ is the superposition

$$\phi(x) = a_1 e^{+ikx} + a_2 e^{-ikx}, \quad a_{1,2} \in \mathbb{C}, \tag{1.97}$$

with the possibility of k becoming complex in the classical forbidden region.

1.4.4 Example: particle in a box [4, 5]

In this example we give the potential a bit more structure, we assume that there is a finite region, 'a box', where the particle experiences no potential. We stick to the one-dimensional case and give the box the length $L > 0$ and infinite height:

$$V(x) = \begin{cases} \infty, & -\infty < x \leq 0 \\ 0, & 0 < x < L \\ \infty, & L \leq x < \infty \end{cases}, \tag{1.98}$$

hence, the particle is trapped in the box and the infinite potential is only compatible with a vanishing wave function, we have

$$\phi(x) = \begin{cases} -\infty < x \leq 0, & \phi(x) = 0 \\ 0 < x < L, & \phi(x) = b_1 e^{+ikx} + b_2 e^{-ikx}, \;\; b_{1,2} \in \mathbb{C} \\ L \leq x < \infty, & \phi(x) = 0 \end{cases}, \tag{1.99}$$

here, in the potential free region we used our solution for the free particle, i.e. $k = \sqrt{2mE}/\hbar$. The requirement that wavefunction vanishes at the potential walls implies

$$\begin{aligned} &\phi(0) = 0 \quad \Rightarrow \quad b_1 + b_2 = 0 \quad \Rightarrow \quad b_2 = -b_1 \equiv \frac{1}{2} b = const., \\ &\phi(L) = 0 \quad \Rightarrow \quad \frac{1}{2} b \left(e^{+ikL} - e^{-ikL}\right) = ib \, \sin(kL) = 0 \quad \Rightarrow \quad kL = n\pi, \;\; n \in \mathbb{N}^+, \end{aligned} \tag{1.100}$$

thus the boundary conditions enforce the wave function in the box to be a sinus-function and crucially, the wavenumber can only take specific 'quantized' values k_n and the wavefunction inside the box becomes

$$0 < x < L \quad \Rightarrow \quad \phi(x) = b \, \sin\left(x k_n\right), \qquad k_n = n \frac{\pi}{L}. \tag{1.101}$$

Requiring that our wave function is normalized, we can determine the prefactor b

$$1 = \int_0^L dx \; |b|^2 \sin^2\left(x k_n\right) = |b|^2 \left[\frac{x}{2} - \frac{1}{4k_n} \sin\left(2x k_n\right)\right]_0^L = |b|^2 \frac{L}{2} \quad \Rightarrow \quad b = \sqrt{\frac{2}{L}} \, e^{i\varphi}, \tag{1.102}$$

thus the wavefunction in the box is now fully defined up to a phase factor φ.
The quantization of the wavevector implies as well a discretization of the energy

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 n^2 \pi^2}{2mL^2}, \;\; n \in \mathbb{N}^+, \tag{1.103}$$

as we have seen before for the example of the free particle. Crucially, in the case of plane waves the k values can take any value and posses a continuous spectrum. There we were using the box just as a practical trick, helping us to normalize the wave function. In contrast, for a particle in the box we obtain indeed a so-called discrete energy spectrum. Note, we have excluded the case $n = 0$, where the wave function would vanish for all values of x and is not normalizable. The lowest energy is $E_1 = \frac{\hbar^2 \pi^2}{2mL}$ and thus unequal zero! Meaning even though the potential vanishes in the box (and classically a particle at rest with energy $E = 0$ is allowed) this is not the case for the quantum solutions. E_1 is also called the ground state energy or the zero-point energy.

Chapter 2

Basic elements of quantum mechanics

2.1 Position and momentum in quantum mechanics

2.1.1 Correspondence principle [4]

In classical mechanics position and momentum are the central variables for describing the dynamics of particles. This section discusses how these variables are treated in the quantum regime. We have defined the probability density $\rho(x,t) = \psi^*(x,t)\psi(x,t)$ and could then calculate the expectation value of the position (for simplicity we stay in 1D)

$$\langle x \rangle = \int dx\ \psi^*(x,t)\ x\ \psi(x,t), \tag{2.1}$$

we can now postulate that the expectation value of the momentum behaves as expected from classical mechanics

$$\begin{aligned}\langle p \rangle = m\frac{\partial}{\partial t}\langle x \rangle =& m\int dx\ \left(\frac{\partial}{\partial t}\psi^*(x,t)\right)\ x\ \psi(x,t) + m\int dx\ \psi^*(x,t)\ x\ \left(\frac{\partial}{\partial t}\psi(x,t)\right) \\ =& \frac{\hbar}{i2}\int dx\ \left[\left(\frac{\partial^2\psi^*(x,t)}{\partial^2 x}\right)\ x\ \psi(x,t) - \psi^*(x,t)\ x\ \left(\frac{\partial^2\psi(x,t)}{\partial^2 x}\right)\right],\end{aligned} \tag{2.2}$$

we can now integrate the first term using partial integration and the assumption that the wavefunction vanishes for $x \to \pm\infty$:

$$\begin{aligned}\int dx\ \left(\frac{\partial^2\psi^*(x,t)}{\partial^2 x}\right)\ x\ \psi(x,t) =& -\int dx\ \left(\frac{\partial\psi^*(x,t)}{\partial x}\right)\ \frac{\partial}{\partial x}\left[x\ \psi(x,t)\right] \\ =& \int dx\ \ \psi^*(x,t)\ \frac{\partial^2}{\partial x^2}\left[x\ \psi(x,t)\right],\end{aligned} \tag{2.3}$$

and insert this into the expression for the expectation value of the momentum

$$\langle p \rangle = \frac{\hbar}{2i}\int dx\ \psi^*(x,t)\left[\ \frac{\partial^2}{\partial x^2}\left[x\ \psi(x,t)\right] - x\ \left(\frac{\partial^2\psi(x,t)}{\partial^2 x}\right)\right] = \int dx\ \psi^*(x,t)\left[\frac{\hbar}{i}\frac{\partial}{\partial x}\right]\psi(x,t), \tag{2.4}$$

hence, the expectation value of the momentum corresponds to the expectation value for the derivation operation. Generalizing this to the three-dimensional case we obtain

$$\langle \mathbf{p} \rangle = \frac{\hbar}{i} \langle \nabla \rangle . \tag{2.5}$$

In quantum mechanics, the physical quantity momentum is represented by the a derivative operation, while the physical quantity position is a multiplication operation. The so-called correspondence principle transfers the classical variables $\mathbf{x}$ and $\mathbf{p}$ into a quantum mechanical description, i.e., operators, it yields

$$\mathbf{x} \;\rightarrow\; \hat{\mathbf{x}}, \qquad \mathbf{p} \;\rightarrow\; \hat{\mathbf{p}} = \frac{\hbar}{i}\nabla, \tag{2.6}$$

we are using here the *hat*-notation, as we are working with operations and thus operators now. We can now revisit our Schrödinger equation derived in the former chapter

$$i\hbar \frac{\partial}{\partial t}\Psi(\mathbf{x},t) = \left[-\frac{\hbar^2 \Delta}{2m} + V(\mathbf{x})\right] \Psi(\mathbf{x},t) = \hat{H}\ \Psi(\mathbf{x},t), \tag{2.7}$$

and we can re-write the Hamiltonian operator as

$$\hat{H} = -\frac{\hbar\nabla \cdot \hbar\nabla}{2m} + V(\mathbf{x}) = \frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}(\hat{\mathbf{x}}). \tag{2.8}$$

We can now use this compact notation to write the evolution equation for the position as

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \langle \hat{\mathbf{x}} \rangle =& - \int d\mathbf{x}\ \left[\hat{H}\Psi^*(\mathbf{x},t)\right]\ \hat{\mathbf{x}}\ \Psi(\mathbf{x},t) + \int d\mathbf{x}\ \Psi^*(\mathbf{x},t)\ \hat{\mathbf{x}}\ \left[\hat{H}\Psi(\mathbf{x},t)\right] \\ =& - \int d\mathbf{x}\ \Psi^*(\mathbf{x},t) \left[\hat{H}\ \hat{\mathbf{x}}\ \Psi(\mathbf{x},t)\right] + \int d\mathbf{x}\ \Psi^*(\mathbf{x},t)\ \hat{\mathbf{x}}\ \left[\hat{H}\Psi(\mathbf{x},t)\right] \\ =& - \int d\mathbf{x}\ \Psi^*(\mathbf{x},t) \left[\hat{H}, \hat{\mathbf{x}}\right]\ \Psi(\mathbf{x},t) = - \left\langle \left[\hat{H}, \hat{\mathbf{x}}\right] \right\rangle , \end{aligned} \tag{2.9}$$

in the second step we performed an integration by parts as done in Eq.(2.3). We have defined here a very important element, the so-called commutator

$$\left[\hat{A}, \hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{2.10}$$

The commutator determines to what extend the order of the application of two operators matters. To evaluate commutators it is best to apply it to a wave function. For example taking the 1-dimensional case

$$[\hat{x}, \hat{p}]\,\psi(x,t) = \frac{\hbar}{i}\left(\hat{x}\frac{\partial}{\partial x}\psi(x,t) - \frac{\partial}{\partial x}\hat{x}\psi(x,t)\right) = \frac{\hbar}{i}\left(\hat{x}\frac{\partial}{\partial x}\psi(x,t) - \psi(x,t) - \hat{x}\frac{\partial}{\partial x}\psi(x,t)\right), \tag{2.11}$$

thus the commutator of position and momentum operators becomes

$$[\hat{x}, \hat{p}] = i\hbar, \tag{2.12}$$

with this we can recover again our earlier result

$$\frac{\partial}{\partial t}\langle \hat{\mathbf{x}} \rangle = \frac{i}{\hbar}\left\langle \left[\hat{H}, \hat{\mathbf{x}}\right] \right\rangle = \frac{i}{\hbar}\frac{1}{2m}\left\langle \left[\hat{\mathbf{p}}^2, \hat{\mathbf{x}}\right] \right\rangle = \frac{i}{\hbar}\frac{1}{2m}\left\langle \hat{\mathbf{p}}\,[\hat{\mathbf{p}}, \hat{\mathbf{x}}] + [\hat{\mathbf{p}}, \hat{\mathbf{x}}]\,\hat{\mathbf{p}} \right\rangle = \frac{\langle \hat{\mathbf{p}} \rangle}{m} \tag{2.13}$$

where we used the rule $[AB, C] = A[B, C] + [A, C]B$ for the commutator.

2.1.2 Wave function in position and momentum space

So far we have focused on the wavefunction as a function of position and time, i.e., in position space. We can as well choose a different perspective by chosen the momentum space. To do so we first need to remember that position space and reciprocal space (**k**-space) are connected via a Fourier transformation. Combining this with the de-Broglie relation $\mathbf{p} = \hbar\mathbf{k}$ we can make the connection between position and momentum space (1D for simplicity):

$$\psi(x,t) = \frac{1}{2\pi\hbar}\int dp\ \psi(p,t)\ e^{\frac{i}{\hbar}px}, \qquad \psi(p,t) = \int dx\ \psi(x,t)\ e^{-\frac{i}{\hbar}px}, \tag{2.14}$$

introducing the wave function in momentum space $\psi(p,t)$. Note, we treat here position and momentum again as scalars. This wavefunction has to be normalized, so we can check

$$\begin{aligned}\int dx\ |\psi(x,t)|^2 =& \int dx\ \frac{1}{2\pi\hbar}\int dp'\ \psi^*(p',t)\ e^{-\frac{i}{\hbar}p'x}\ \frac{1}{2\pi\hbar}\int dp\ \psi(p,t)\ e^{\frac{i}{\hbar}px}\\ =&\frac{1}{(2\pi\hbar)^2}\int dx\int dp'\int dp\ \psi^*(p',t)\ \psi(p,t)\ e^{\frac{i}{\hbar}(p-p')x}\\ =&\frac{1}{(2\pi\hbar)^2}\int dp'\int dp\ \psi^*(p',t)\ \psi(p,t)\ (2\pi\hbar)\delta(p-p') = \frac{1}{2\pi\hbar}\int dp\ |\psi(p,t)|^2.\end{aligned} \tag{2.15}$$

Thus we have for the normalization

$$\int dx\ |\psi(x,t)|^2 = \frac{1}{2\pi\hbar}\int dp\ |\psi(p,t)|^2 = 1, \tag{2.16}$$

we can interpret $|\psi(x,t)|^2$ as the probability density in position space and $|\psi(p,t)|^2/(2\pi\hbar)$ as the probability density in momentum space. Important to note is that in quantum mechanics a particle is described by its state, but we have the freedom to choose in which basis we represent this state, e.g. momentum or position basis. We can also re-define the correspondence principle discussed in the section before, as the latter was formulated in position space. In momentum space the momentum is a simple multiplication operation, while for the position we find

$$\begin{aligned}\langle x\rangle =& \int dx\ \psi^*(x,t)x\psi(x,t) = \frac{1}{(2\pi\hbar)^2}\int dx\int dp'\int dp\ \psi^*(p',t)\ e^{-\frac{i}{\hbar}p'x}x\psi(p,t)\ e^{\frac{i}{\hbar}px}\\ =&\frac{1}{(2\pi\hbar)^2}\int dx\int dp'\int dp\ \psi^*(p',t)\ \left(i\hbar\frac{\partial}{\partial p'}e^{-\frac{i}{\hbar}p'x}\right)\psi(p,t)\ e^{\frac{i}{\hbar}px}\\ =&-\frac{i\hbar}{(2\pi\hbar)^2}\int dx\int dp'\int dp\ \left(\frac{\partial}{\partial p'}\psi^*(p',t)\right)\psi(p,t)\ e^{\frac{i}{\hbar}(p-p')x}\\ =&-\frac{i\hbar}{2\pi\hbar}\int dp\ \left(\frac{\partial}{\partial p}\psi^*(p,t)\right)\psi(p,t) = \frac{1}{2\pi\hbar}\int dp\ \psi^*(p,t)\left(i\hbar\frac{\partial}{\partial p}\right)\psi(p,t),\end{aligned} \tag{2.17}$$

hence the position in momentum space becomes now the derivative operation, while the momentum is multiplication operation. With this the correspondence principle becomes

$$\begin{aligned}\text{position space:}\quad & \mathbf{x}\ \rightarrow\ \hat{\mathbf{x}}, && \mathbf{p}\ \rightarrow\ \hat{\mathbf{p}} = \frac{\hbar}{i}\nabla_x\\ \text{momentum space:}\quad & \mathbf{x}\ \rightarrow\ \hat{\mathbf{x}} = i\hbar\nabla_p, && \mathbf{p}\ \rightarrow\ \hat{\mathbf{p}}.\end{aligned} \tag{2.18}$$

2.2 The Hilbert space

2.2.1 Superposition of waves: Fourier series [6]

In Section 1.4.4 we considered a particle in a box with infinite walls, we found that the wave function exhibits oscillatory behaviour inside of the box with quantized values for the wavevector, i.e. the wavefunction yields (up to a phase)

$$\phi_n(x) = \sqrt{\frac{2}{L}} \sin\left(n\frac{\pi x}{L}\right), \quad n \in \mathbb{N}^+, \tag{2.19}$$

which was via solving the stationary Schrödinger equation. We know that the latter is a linear partial differential equation, thus, any linear combination of one solution is again a solution. We can write such a superposition as

$$\phi(x) = \sum_{n=1}^{\infty} b_n \sin\left(n\frac{\pi x}{L}\right) \tag{2.20}$$

which is nothing else than a Fourier sinus-series defined on the interval $[0, L]$. As a reminder, a Fourier series is an expansion of a periodic function into a sum of trigonometric functions. Under the assumption that we know $\phi(x)$ we can calculate the Fourier series coefficients via multiplication of $\phi(x)$ with one of the series elements (a sinus function) and integration:

$$\begin{aligned}
\int_0^L dx\ \phi(x)\ \sin\left(n'\frac{\pi x}{L}\right) =& \sum_{n=1}^{\infty} b_n \int_0^L dx\ \sin\left(n\frac{\pi x}{L}\right)\ \sin\left(n'\frac{\pi x}{L}\right) \\
=& \frac{1}{2}\sum_{n=1}^{\infty} b_n \int_0^L dx\ \left[\cos\left(\{n-n'\}\frac{\pi x}{L}\right) - \cos\left(\{n+n'\}\frac{\pi x}{L}\right)\right] \\
=& \frac{L}{2\pi}\sum_{n=1}^{\infty} b_n \left[\frac{\sin\left(\{n-n'\}\frac{\pi x}{L}\right)}{\{n-n'\}} - \frac{\sin\left(\{n+n'\}\frac{\pi x}{L}\right)}{\{n+n'\}}\right]_0^L \\
=& \frac{L}{2\pi}\sum_{n=1}^{\infty} b_n \frac{\sin\left(\{n-n'\}\pi\right)}{\{n-n'\}} = \frac{L}{2}\, b_n\, \delta_{n,n'},
\end{aligned} \tag{2.21}$$

in the last step we used the limit for $n \to n'$ and introduced the Kronecker symbol

$$\delta_{n,n'} = \begin{cases} 1, & n = n' \\ 0, & n \neq n' \end{cases}. \tag{2.22}$$

With this we have an explicit expression for the coefficients b_n

$$b_n = \frac{2}{L}\int_0^L dx\ \phi(x)\ \sin\left(n\frac{\pi x}{L}\right). \tag{2.23}$$

Slightly more general, we can express the periodic function $\psi(x) = \psi(x + 2L)$ as a complex Fourier series in the interval $[-L, L]$

$$\psi(x) = \sum_{n=-\infty}^{\infty} c_n \, e^{i\frac{n\pi}{L}x}, \quad c_n = \frac{1}{2L} \int_{-L}^{L} dx \, \psi(x) \, e^{-i\frac{n\pi}{L}x} \tag{2.24}$$

with complex coefficients c_n. One calls this representation of $\psi(x)$ as expansion in plane waves.

2.2.2 Complex vector space

The just introduced periodic complex functions ψ with $\psi(x) = \psi(x+2L)$ in an interval $[-L, L]$ can form a complex vector space $\mathcal{H}$. The basis of this vector space are the functions

$$e_n : \mathbb{R} \to \mathbb{C}, \quad e_n \equiv e^{i\frac{n\pi}{L}x}, \quad n = 0, \pm 1, \pm 2, ..., \tag{2.25}$$

and as we have an infinite number of n's, this basis and hence the vector space is infinite dimensional. Remember, a vector space is a linear space, an algebraic structure which is constructed from its elements, the vectors. Vectors can be be added or multiplied with scalars, and the resulting vector is still part of the vector space. The operations of vector addition and scalar multiplication must satisfy certain requirements, called vector axioms.

Definition: complex vector space
A set V of elements is called a complex vector space if for $x, y, z \in V$ and $\alpha, \beta \in \mathbb{C}$ the following axioms are valid

1. vector addition ($\oplus$)
   - $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ (associative property)
   - $x \oplus y = y \oplus x$ (commutative property)
   - existence of a neutral element: $0_V \in V$ with $x \oplus 0_V = 0_V \oplus x = x$
   - existence of an inverse element: $-x \in V$ with $x \oplus (-x) = (-x) \oplus x = 0_V$
2. scalar multiplication ($\odot$)
   - $\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y)$
   - $(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x)$
   - $(\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x)$
   - $1 \odot x = x$.

The vector space $\mathcal{H}$ defined with the basis vectors above above has to fulfill this definition and the vectors are the functions ψ. The Fourier coefficients c_n denote the components of the vectors in a orthonormal basis, and are obtained from the scalar product (inner product)

$$c_n = \langle e_n, \psi \rangle \equiv \frac{1}{2L} \int_{-L}^{L} dx \, e_n^* \psi(x). \tag{2.26}$$

In addition, the basis is orthonormal, i.e., we have

$$\langle e_n, e_m \rangle = \frac{1}{2L} \int_{-L}^{L} dx\; e_n^* e_m = \frac{1}{2L} \int_{-L}^{L} dx\; e^{-i\frac{\pi x}{L}(n-m)} = \delta_{n,m}, \tag{2.27}$$

and the basis is complete, which follows from $\sum_n e_n^* e_n = 1$ over all values of n. To note is that it is as well possible to choose a different set of basis functions. For the same interval $[-L, L]$ one could as well use real-valued base-functions utilizing the Fourier-series of sinus and cosine functions. For the example of the particle in the box we had the interval $[0, L]$ and the sinus basis functions as given in Eq. (2.19).

2.2.3 Dirac notation

In this section we introduce now a very important notation in quantum mechanics: the Dirac notation. We start again from the stationary Schrödinger equation,

$$\hat{H}\phi_n(x) = E_n \phi_n(x), \tag{2.28}$$

which corresponds to an eigenvalue problem, where $\phi_n(x)$ represents the eigenstate (stationary state) with eigenvalue E_n in the position basis. In the Dirac notation stationary state are denoted as Dirac-*kets* $|\phi_n\rangle$ and one re-expressed the stationary Schrödinger equation as

$$\hat{H}\,|\phi_n\rangle = E_n\,|\phi_n\rangle\,, \tag{2.29}$$

and $|\phi_n\rangle$ is defined as a vector in the complex vector space $\mathcal{H}$. For the example of a free particle we derived the wavefunction $\phi_n(x)$ as given in Eq.(2.19) and we discussed that a superposition of these waves is nothing else than a Fourier series. This translates to the Dirac formalism, where we can use the Fourier series to expand an arbitrary wave function in stationary eigenstates, which reads in Dirac notation

$$|\psi\rangle = \sum_n a_n\,|\phi_n\rangle\,, \quad \{|\phi_n\rangle\} \in \mathcal{H}, \quad \{a_n\} \in \mathbb{C}. \tag{2.30}$$

Here, we can as well introduce the dual vector so-called Dirac-*bra* $\langle\phi_n|$. Again we can expand an arbitrary wavefunction in the dual space as

$$\langle\psi| = \sum_n a_n^*\,\langle\phi_n|\,, \quad \{\langle\phi_n|\} \in \mathcal{H}^*, \quad \{a_n^*\} \in \mathbb{C}, \tag{2.31}$$

a vector which lives in the dual space $\mathcal{H}^*$. The Dirac notation is extremely practical as we can represent states in different bases in it. We will see that the choice of the basis will be connected to the measurement we like to perform, meaning that it if we measure in a certain basis it is very helpful to express the state in the same basis. And we are not limited to discrete bases, meaning if necessary we can also use a continuous representation to express our wavefunction. E.g., in the position basis $\{|x\rangle\}$ or the momentum basis $\{|p\rangle\}$ we have

$$|\psi\rangle = \int dx\; \psi(x)\,|x\rangle = \int dp\; \psi(p)\,|p\rangle\,, \tag{2.32}$$

where the sum over all coefficient has become an integral. To note is that a continuous basis is called *improper*, something we will discuss in more detail later on.

A crucial aspect of a vector space is the scalar product between two vectors which in Dirac notation becomes

$$\langle\phi|\psi\rangle = \sum_{n,m} a_n b_m^* \langle\phi_m|\phi_n\rangle = \sum_{n,m} a_n b_m^* \delta_{m,n} = \sum_n a_n b_n^* = \langle\psi|\phi\rangle^* \tag{2.33}$$

where we have employed that the basis vectors have to be orthogonal

$$\langle\phi_n|\phi_m\rangle = \delta_{n,m} \text{ (discrete basis)}, \quad \langle x|x'\rangle = \delta(x-x') \text{ (continuous basis)}. \tag{2.34}$$

The scalar product expressed in the position basis has the expected form

$$\langle\phi|\psi\rangle = \int dx' \int dx\ \phi^*(x')\psi(x)\langle x'|x\rangle = \int dx\ \phi^*(x)\psi(x). \tag{2.35}$$

We can now also straightforwardly determine the coefficient in our upper expansion of the wavefunction

$$\langle\phi_n|\psi\rangle = \sum_{n'} a_{n'} \langle\phi_n|\phi_{n'}\rangle = \sum_{n'} a_{n'}\delta_{n,n'} = a_n, \tag{2.36}$$

and we have

$$\langle\psi|\psi\rangle = \sum_{n,n'} a_n^* a_{n'} \langle\phi_n|\phi_{n'}\rangle = \sum_n |a_n|^2 = 1, \tag{2.37}$$

following from the normalization of the wavefunction. In addition we can connect the vector $|\phi_n\rangle$ to the wavefunction in position space by multiplying $|\phi_n\rangle$ from the right with the dual basis vector in the position basis:

$$\langle x|\phi_n\rangle = \int dx'\ \phi_n(x')\langle x'|x\rangle = \int dx'\ \phi_n(x')\ \delta(x-x') = \phi_n(x). \tag{2.38}$$

Besides the scalar product, also called inner product, we can as well define an outer product or dyadic product, which in contrast to the inner product, is not a scalar:

$$|\psi\rangle\langle\phi| = \sum_{n,m} a_n b_m^* |\phi_m\rangle\langle\phi_n|, \tag{2.39}$$

which leads us to a very practical expression, starting from the normalization

$$1 = \langle\psi|\psi\rangle = \sum_n |a_n|^2 = \sum_n \langle\phi_n|\psi\rangle^* \langle\phi_n|\psi\rangle = \sum_n \langle\psi|\phi_n\rangle\langle\phi_n|\psi\rangle = \langle\psi|\left(\sum_n |\phi_n\rangle\langle\phi_n|\right)|\psi\rangle, \tag{2.40}$$

we obtain the operator identity

$$1 = \sum_n |\phi_n\rangle\langle\phi_n| \equiv \hat{1}. \tag{2.41}$$

we will see that the outer product can represent operations and thus operators.

This identity is very useful, in calculation we will often introduce a $\hat{1}$ to achieve our respective goal. For example, and as already mentioned above, we can express a state $|\psi\rangle$ in the basis which suits our needs best. Taking a state $|\psi\rangle$ in the basis $\{\phi_n\}$ which we like to change to basis $\{\theta_n\}$ we do so via

$$|\psi\rangle = \sum_n a_n |\phi_n\rangle = \sum_n a_n \hat{1} |\phi_n\rangle = \sum_{n,m} a_n |\theta_m\rangle \langle\theta_m|\phi_n\rangle \equiv \sum_m b_m |\theta_m\rangle , \;\; b_m = \sum_n a_n \langle\theta_m|\phi_n\rangle , \tag{2.42}$$

here we were allowed to combine in the new coefficients b_m the sum over the complex coefficient a_n with the scalar product of the basis vectors, the latter is also just a complex number.

2.2.4 Expectation values and hermitian operators

A very nice aspect of the Dirac notation is that we can straightforwardly calculate expectation values. Taking for example the Hamiltonian operator from above we obtain for the expectation value in the stationary state or also the so-called eigenstate $|\phi_n\rangle$

$$\hat{H} |\phi_n\rangle = E_n |\phi_n\rangle \quad \Rightarrow \quad \left\langle \hat{H} \right\rangle = \langle\phi_n| \hat{H} |\phi_n\rangle = E_n \langle\phi_n|\phi_n\rangle = E_n, \tag{2.43}$$

meaning that the expectation value of the Hamilton operator in the eigenstate $|\phi_n\rangle$ is the eigenenergy E_n. We can generalize this eigenvalue problem and consider an arbitrary operator $\hat{O}$ which posses the eigenfunctions $|\varphi_n\rangle$ and the eigenvalues o_n:

$$\hat{O} |\varphi_n\rangle = o_n |\varphi_n\rangle \quad \Rightarrow \quad \left\langle \hat{O} \right\rangle = \langle\varphi_n| \hat{O} |\varphi_n\rangle = o_n. \tag{2.44}$$

When the operator represents a physical quantity, i.e., one which we can measure, we assume that the eigenvalue is real, just as for the example of the Hamilton operator where the energy has to be real. And crucially, all eigenvalues o_n have to be real, to check this we first have to define the adjoint operator $\hat{A}^\dagger$. Taking some arbitrary state $|\phi\rangle$ which is obtained after operating with the operators $\hat{A}$ on the state $|\psi\rangle$, the adjoint operator is defined as

$$|\phi\rangle = \hat{A} |\psi\rangle \quad \Rightarrow \quad \langle\phi| = \langle\psi| \hat{A}^\dagger, \tag{2.45}$$

thus $\hat{A}^\dagger$ acts on the dual space $\mathcal{H}^*$, just as $\hat{A}$ acts on $\mathcal{H}$. With this we can express the expectation value of $\hat{A}$ with respect to the state $\langle\psi|$

$$\langle\psi| \hat{A} |\psi\rangle = \langle\psi|\phi\rangle = \langle\phi|\psi\rangle^* = \langle\psi| \hat{A}^\dagger |\psi\rangle^* . \tag{2.46}$$

Thus, requiring that the operator $\hat{O}$ represents a physical quantity with real eigenvalues we can conclude that

$$\langle\varphi_n| \hat{O} |\varphi_n\rangle = o_n = o_n^* = \langle\varphi_n| \hat{O}^\dagger |\varphi_n\rangle \quad \Rightarrow \quad \hat{O} = \hat{O}^\dagger, \tag{2.47}$$

meaning that an operator with a real eigenvalue is self-adjoint, we call such operators hermitian operators. All the operators we have learned about so far are hermitian operators, e.g. the position operator $\hat{x}$, the momentum operator $\hat{p}$ as well as the Hamiltonian operator $\hat{H}$.

2.2.5 The uncertainty principle

Now we will turn to a fundamental relation in quantum mechanics, which will tell us how accurate we can measure specific pairs of observables. For this we start with the expectation value of an observable denoted as operator $\hat{A}$ and its standard deviation $\hat{\sigma}_A$

$$\left\langle \hat{A} \right\rangle = \langle\psi| \hat{A} |\psi\rangle , \quad \langle \hat{\sigma}_A^2 \rangle = \left\langle \left(\hat{A} - \left\langle \hat{A} \right\rangle \right)^2 \right\rangle = \left\langle \hat{A}^2 - 2\hat{A} \left\langle \hat{A} \right\rangle + \left\langle \hat{A} \right\rangle^2 \right\rangle = \left\langle \hat{A}^2 \right\rangle - \left\langle \hat{A} \right\rangle^2 , \tag{2.48}$$

the square of the the standard deviation is called the variance. The expectation value tells us which average value we obtain if we perform a measurement in the basis $|\psi\rangle$ and the standard deviation will give us the uncertainty, meaning the variations of our measurement results around the expectation value. Uncertainty in this sense tells us how accurate we can measure an observable. Now consider two observables $\hat{A}$ and $\hat{B}$ with deviations $\hat{\sigma}_A$ and $\hat{\sigma}_B$. We already learned that the orders of operators is important, so considering the commutator of the standard deviations

$$\begin{aligned} [\hat{\sigma}_A, \hat{\sigma}_B] &= \left[\left(\hat{A} - \langle A \rangle \right), \left(\hat{B} - \langle B \rangle \right) \right] \\ &= \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \left(\hat{B} - \left\langle \hat{B} \right\rangle \right) - \left(\hat{B} - \left\langle \hat{B} \right\rangle \right) \left(\hat{A} - \left\langle \hat{A} \right\rangle \right) \\ &= \left[\hat{A}, \hat{B} \right] - \left[\left\langle \hat{A} \right\rangle, \hat{B} \right] - \left[\hat{A}, \left\langle \hat{B} \right\rangle \right] + \left[\left\langle \hat{A} \right\rangle, \left\langle \hat{B} \right\rangle \right] = \left[\hat{A}, \hat{B} \right] , \end{aligned} \tag{2.49}$$

hence, the commutator of the standard deviations equals the one of the commutators. We can analyse the expectation values of the individual products of the standard deviation, each product corresponds to a complex number:

$$\begin{aligned} \langle \hat{\sigma}_A \hat{\sigma}_B \rangle &\equiv x + iy \\ \langle \hat{\sigma}_B \hat{\sigma}_A \rangle &\equiv x - iy \end{aligned} \quad \Rightarrow \; | \langle \hat{\sigma}_A \hat{\sigma}_B \rangle |^2 = x^2 + y^2 \tag{2.50}$$

where $x, y \in \mathbb{R}$. On the other hand the expectation value of the commutators is

$$\langle [\hat{\sigma}_A, \hat{\sigma}_B] \rangle = 2iy \; \Rightarrow \; | \langle [\hat{\sigma}_A, \hat{\sigma}_B] \rangle |^2 = 4y^2, \tag{2.51}$$

from which we can conclude that

$$| \langle [\hat{\sigma}_A, \hat{\sigma}_B] \rangle |^2 \leq 4 | \langle \hat{\sigma}_A \hat{\sigma}_B \rangle |^2. \tag{2.52}$$

The right hand side is nothing else than the absolute value squared of the scalar product of two vectors, $\langle\psi| \hat{\sigma}_A$ and $\hat{\sigma}_B |\psi\rangle$. Which we can asses with the so-called Cauchy-Schwarz inequality:

$$\langle \alpha|\alpha\rangle \langle \beta|\beta\rangle \geq | \langle \alpha|\beta\rangle |^2, \tag{2.53}$$

for two vectors $|\alpha\rangle$ and $|\beta\rangle$. Which gives us

$$| \langle \hat{\sigma}_A \hat{\sigma}_B \rangle |^2 = | \left(\langle\psi| \hat{\sigma}_A \right) \left(\hat{\sigma}_B |\psi\rangle \right) |^2 \leq \left(\langle\psi| \hat{\sigma}_A \right) \left(\hat{\sigma}_A |\psi\rangle \right) \left(\langle\psi| \hat{\sigma}_B \right) \left(\hat{\sigma}_B |\psi\rangle \right) = \left\langle \hat{\sigma}_A^2 \right\rangle \left\langle \hat{\sigma}_B^2 \right\rangle . \tag{2.54}$$

From this follows

$$| \langle [\hat{\sigma}_A, \hat{\sigma}_B] \rangle |^2 \leq 4 | \langle \hat{\sigma}_A \hat{\sigma}_B \rangle |^2 \leq 4 \left\langle \hat{\sigma}_A^2 \right\rangle \left\langle \hat{\sigma}_B^2 \right\rangle , \tag{2.55}$$

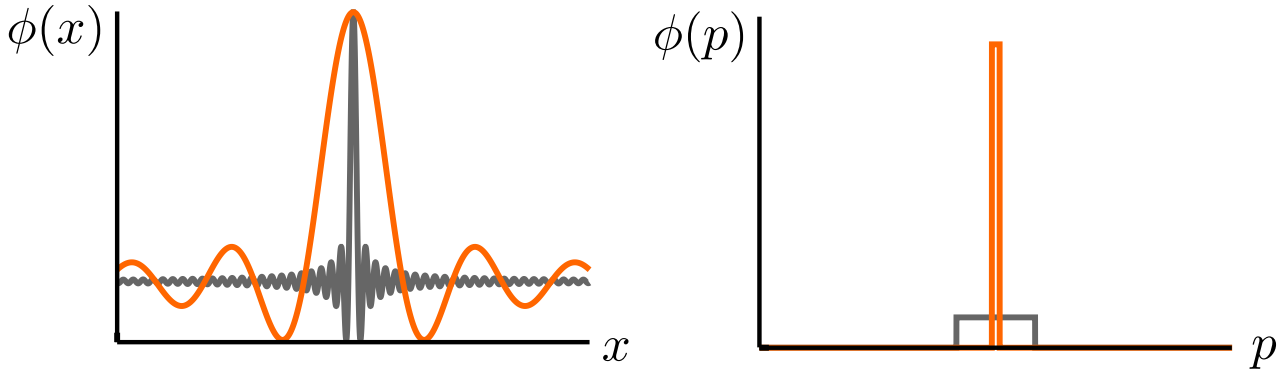

Figure 2.1: Position and momentum are related via Fourier transform, a sharp localization of the particle in space (grey), is related to a broader spread of the momentum. Reversely, a sharp momentum, leads to a broader spread of the wavefunction in position space. Here the position wavefunction is a sinc-function, whose Fourier transform is a rectangular pulse.

where the left hand side is equal to the commutator of the operators which we can now estimate

$$\langle \sigma_A^2 \rangle \langle \sigma_B^2 \rangle \geq \frac{1}{4} \left| \left\langle \left[\hat{A}, \hat{B} \right] \right\rangle \right|^2 \tag{2.56}$$

which is the general uncertainty relation. This relation tells us that two quantities who don't commute cannot be measured sharply at the same time. Having for example the situation of $\langle \hat{\sigma}_A \rangle \to 0$, i.e., a sharp measurement and thus determination of the observable $\hat{A}$, comes with $\langle \hat{\sigma}_B \rangle \to \infty$, a completely undetermined observable $\hat{B}$. Taking for example the non-commuting observables position and momentum, we have $[\hat{x}, \hat{p}] = i\hbar$ and thus

$$\langle \hat{\sigma}_x^2 \rangle \langle \hat{\sigma}_p^2 \rangle \equiv \Delta x^2 \Delta p^2 \geq \frac{1}{4} | \langle [\hat{x}, \hat{p}] \rangle |^2 \quad \Rightarrow \quad \Delta x \Delta p \geq \frac{\hbar}{2}, \tag{2.57}$$

also called the Heisenberg uncertainty relation for position and momentum. Meaning that a sharper momentum comes with enhanced spread in position space, see Fig.2.1 for an example. The uncertainty principle is a fundamental limitation, however, scientist have tried to find ways to avoid it [13]. A community which has been heavily invested in alternative measurement schemes is the one performing gravitational wave detection (to note is that this has been just recently achieved by the LIGO cooperation in 2016, which resulted in being awarded the Nobelprize in 2017). We will come back to this in a later section.

2.2.6 Some more formal aspects of the Hilbert space [4, 7]

A complex vector space has to fulfil the axioms of vector addition and scalar multiplication as discussed in the former section. In quantum mechanics we deal with a special kind of complex vector space the so-called Hilbert space. This space has to fulfill further properties, some of those we have already encountered, e.g. like the inner product of two vectors. In this section we will summarize in a more rigorous manner the requirements on that Hilbert space. Note, we will not go into the details of associated mathematical proofs, this is out of the scope of this course. However, in the script of Brandes [4] one can find further literature references for the mathematical details.

The Hilbert space $\mathcal{H}$ is defined as a set of elements which we call states or state vectors $|\psi\rangle$. $\mathcal{H}$ has the following properties:

1. **$\mathcal{H}$ is a complex, linear vector space.**
   The vector axioms as outlined in Sec.(2.2.2) are fulfilled. In Dirac notation, assuming the elements $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in \mathcal{H}$ and the complex numbers $c_{1,2} \in \mathbb{C}$, these axioms read
   (a) vector addition ($\oplus$)
   - $|\alpha\rangle \oplus (|\beta\rangle \oplus |\gamma\rangle) = (|\alpha\rangle \oplus |\beta\rangle) \oplus |\gamma\rangle \in \mathcal{H}$
   - $|\alpha\rangle \oplus |\beta\rangle = |\beta\rangle \oplus |\alpha\rangle \in \mathcal{H}$
   - existence of a neutral element: $|0\rangle_{\mathcal{H}} \in \mathcal{H}$ with $|\alpha\rangle \oplus |0\rangle_{\mathcal{H}} = |\alpha\rangle, \forall |\alpha\rangle \in \mathcal{H}$
   - existence of an inverse element: $|-\alpha\rangle_{\mathcal{H}} \in \mathcal{H}$ with $|\alpha\rangle \oplus |-\alpha\rangle_{\mathcal{H}} = |0\rangle_{\mathcal{H}}$

   (b) scalar multiplication ($\odot$)
   - $c_1 \odot (|\alpha\rangle \oplus |\beta\rangle) = (c_1 \odot |\alpha\rangle) \oplus (c_1 \odot |\beta\rangle) \in \mathcal{H}$
   - $(c_1 + c_2) \odot |\alpha\rangle = (c_1 \odot |\alpha\rangle) \oplus (c_2 \odot |\alpha\rangle) \in \mathcal{H}$,
   - $(c_1 \cdot c_2) \odot |\alpha\rangle = c_1 \odot (c_2 \odot |\alpha\rangle) \in \mathcal{H}$
   - $0 \odot |\alpha\rangle = |0\rangle_{\mathcal{H}}, \forall |\alpha\rangle \in \mathcal{H}$ and $c_1 |0\rangle_{\mathcal{H}} = |0\rangle_{\mathcal{H}}$

   Some comments on the notation, the lowest state of a system with discrete eigenstates is often denotes as $|0\rangle$, which is distinct from the neutral element $|0\rangle_{\mathcal{H}}$ (that is why we use the subindex). Moreover, the notation for the inverse element means here $|-\alpha\rangle_{\mathcal{H}} = -|\alpha\rangle$, therewith also defining the subtraction of two vectors $|\alpha\rangle - |\beta\rangle$. Used without the subindex, this notation can lead to confusion with the notation of certain states, for example coherent states where $|\alpha\rangle$ and $|-\alpha\rangle$ are states of opposite phase, and their addition is not zero. More generally, we will avoid the notation $c_1 |\alpha\rangle = |c_1\alpha\rangle$ as done in some textbooks, instead, and if necessary, we will define a new vector $|\alpha'\rangle = c_1 |\alpha\rangle$.

   A further aspect of a linear vector space is that a set of vectors is called **linearly independent** if there exists no non-trivial linear combination of the vectors that equals the zero vector. Meaning that a set of elements $|\alpha_1\rangle, |\alpha_2\rangle, ..., |\alpha_n\rangle$ is linear independent if

$$\sum_m c_m |\alpha_m\rangle = |0\rangle_{\mathcal{H}}, \tag{2.58}$$

   is only fulfilled for all $c_m = 0$. The set of basis vectors are linearly independent and span the vector space. All basis vectors have the same dimension. The dimension of $\mathcal{H}$ is the maximal number of linearly independent elements in $\mathcal{H}$.

2. **$\mathcal{H}$ is a complex inner product space.**
   A vector space with an inner product is promoted to an inner product space. A complex inner product space is also sometimes referred to as a unitary space. The inner product connects two complex vectors to a complex number. We already introduced the inner product of two states $|\alpha\rangle, |\beta\rangle \in \mathcal{H}$ in Dirac notation as

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^* \in \mathcal{C}, \tag{2.59}$$

   which has the following properties

- $\langle\alpha|\beta_1+\beta_2\rangle = \langle\alpha|\beta_1\rangle + \langle\alpha|\beta_2\rangle, \quad |\beta_{1,2}\rangle \in \mathcal{H}$
- with $|\beta'\rangle \equiv c\,|\beta\rangle$ and $|\alpha'\rangle \equiv c^*\,|\alpha\rangle \;\Rightarrow\; \langle\alpha|\beta'\rangle = c\,\langle\alpha|\beta\rangle = \langle\alpha'|\beta\rangle\,,\; c \in \mathbb{C}$
- $\langle\alpha|\alpha\rangle \geq 0\;\forall\,|\alpha\rangle \in \mathcal{H}$ and $\langle\alpha|\alpha\rangle = 0$ only for $|\alpha\rangle = |0\rangle_{\mathcal{H}}$

In addition, we have already learned that for two orthogonal vectors we have $\langle\alpha|\beta\rangle = 0$. Another crucial aspect is that the inner product automatically introduces a norm, which is defined as

$$||\alpha|| = \sqrt{\langle\alpha|\alpha\rangle}, \tag{2.60}$$

so a inner product space is as well a normed vector space. To note is that there exist as well vector spaces with alternative norms, however, here the inner product provides us with a norm which we can interpret as the length of the vector. Moreover, a vector is called normalized if the norm is $||\alpha|| = 1$. The distance, or distance function, between two vectors $|\alpha\rangle$ and $|\beta\rangle$ can as well be represented with help of the norm

$$d(\alpha,\beta) = ||\alpha-\beta|| = \sqrt{\langle\alpha-\beta|\alpha-\beta\rangle} = ||\beta-\alpha||. \tag{2.61}$$

3. **$\mathcal{H}$ is separable and complete.**
Up to this point we have learned about complex vector spaces with complex periodic functions defined on a finite interval Ω as basis vectors. Example functions are

$$\begin{aligned} e_n =& \sqrt{\frac{1}{2L}}e^{i\frac{n\pi}{L}x}, && \Omega = [-L,L] \in \mathbb{R}^1, \quad n = 0,\pm1,\pm2,... \\ e_n =& \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right), && \Omega = [0,L] \in \mathbb{R}^1, \qquad n = 1,2,3..., \end{aligned} \tag{2.62}$$

one important aspect is that the vector space spanned by these vectors is infinite dimensional, but so-called **countable infinite**. The latter means that if we have a set which can be put in one-to-one correspondence with the natural numbers and we can identify each element by this prescription. However, the actual dimension of $\mathcal{H}$ is provided by the physical problem at hand, and we can as well be confronted with the situation where we have a truly infinite dimensional space. The latter leads to two more requirements on the vector space: $\mathcal{H}$ has to be separable and complete. To understand what this means we have to remind us of how convergence of a sum can be defined, meaning that a sum over the elements of a series, e.g., a sum from integer n to ∞, converges to a fixed value. Taking for example a series over as set $\{|\alpha_n\rangle\}$, this series converges to α if

$$\lim_{n\to\infty} ||\alpha_n - \alpha|| = 0 \tag{2.63}$$

is fulfilled. Moreover, in mathematics one calls a subset $\mathcal{A}$ of a space $\mathcal{V}$ dense, if every element in $\mathcal{V}$ either belongs to $\mathcal{A}$ or is arbitrarily close to a member of $\mathcal{A}$. For example the rational numbers $\mathbb{Q}$ are a dense subset of the real numbers $\mathbb{R}$, as every real number either is a rational number p/q or approximates a rational number very closely. Taking all this, we call the Hilbert space $\mathcal{H}$ separable if there exists everywhere at least one

dense set of countable vectors $\{|\alpha_n\rangle\}$, meaning that there exist for every vector $|\psi\rangle$ at least one element $|\alpha_m\rangle \in \{|\alpha_n\rangle\}$ such that for arbitrary small $\epsilon > 0$ we have

$$|\psi\rangle \in \mathcal{H}: \quad ||\alpha_m - \psi|| < \epsilon. \tag{2.64}$$

In other words, we need to be able to represent every state in the Hilbert space $\mathcal{H}$ with a countable (infinite or finite) set of vectors. Important is that the set is dense, such the series converges to any element of $\mathcal{H}$. This means that every element of $\mathcal{H}$ can be uniquely specified with its coordinates in relation to an orthonormal basis.

We define a **complete orthonormal system** as a set $\mathcal{M}$ of orthonormal vectors obeying

$$\langle\alpha_n|\alpha_m\rangle = \delta_{n,m}, \quad |\alpha_{n,m}\rangle \in \mathcal{M} \subseteq \mathcal{H}, \tag{2.65}$$

which ensures that that the vectors are orthogonal to each other. Completeness means that there exist no other element in $\mathcal{H}$, which does not belong to $\mathcal{M}$, which is orthogonal to the elements in $\mathcal{M}$. Meaning that for a vector $|\beta\rangle$ which is not part of $\mathcal{M}$ we have

$$|\beta\rangle \in \mathcal{H} \nsubseteq \mathcal{M} \quad \Rightarrow \quad \langle\beta|\alpha_m\rangle \neq 0, \ \forall\, |\alpha_m\rangle \in \mathcal{M}. \tag{2.66}$$

The complete orthonormal system maximally contains countable infinite elements, in the turn the dimension of $\mathcal{H}$ is maximally countable infinite. Every vector $|\psi\rangle$ in the Hilbert space can be expanded in the complete orthonormal basis

$$|\psi\rangle = \sum_n a_n |\alpha_n\rangle; \quad a_n = \langle\alpha_n|\psi\rangle, \tag{2.67}$$

and we here call the inner product of a vector with a basis of the orthonormal system expansion (or Fourier) coefficient a_n in relation to the orthonormal system. One crucial aspect is the convergence of the expansion, meaning the convergence of

$$\sum_n |a_n|^2 = \langle\psi|\psi\rangle = ||\psi||^2, \tag{2.68}$$

which is not a sufficient condition. The conversion of the expansion could also lead to an element which is not part of $\mathcal{H}$, meaning we need another axiom besides separability, and this is completeness.

For this we have to remind us that a series $\{|\alpha_n\rangle\}$ is defined as a **Cauchy-series**, if for every $\epsilon > 0$ there exist an integer $N \in \mathbb{N}$ such that

$$||\alpha_n - \alpha_m|| < \epsilon, \quad \forall\, n, m > N, \tag{2.69}$$

meaning that a series converges if the distance between the elements become arbitrary small, and every strongly converging series is a Cauchy series. The Hilbert space is complete if every series $\{|\alpha_n\rangle\} \in \mathcal{H}$ converges to an element $|\alpha\rangle \in \mathcal{H}$. As a complete normed space Hilbert spaces are so-called Banach spaces. The latter are vector spaces which have a metric or norm which allows for the computation of the length of a vector, as well as the distance between two vectors. And a Banach space is complete as every Cauchy series of vectors converges to an element within the space.

In summary, the Hilbert space $\mathcal{H}$ is a complex inner product space, which is separable, complete and has a maximal dimension of countable infinity.

2.2.7 Continuous basis: improper vectors [7, 8]

We have learned that we can express the state of a system in a orthonormal and complete set of basis states. The choice of the basis states is guided by the physical problem at hand. A practical choice is to use the eigenfunctions of a hermitian operator, as these operators correspond to the observables, i.e., the measurable quantities. We can for example imagine to measure the position of an object, the eigenvectors in position space therefor can read

$$\hat{x}\,|x\rangle = x\,|x\rangle\,, \tag{2.70}$$

which are continuous eigenfunctions with continuous eigenvalues x, this posses an issue. We have seen, that the axioms of the Hilbert space are valid for a countable finite or infinite basis. Continuous states such as the position eigenfunction are thus formally not admitted, they are not countable and cannot be normalized. However, the measurement of position is a crucial aspect which we cannot simply exclude from our mathematical framework. Thus we have to find the right extension of the Hilbert space, to allow as well for continuous bases. To note is that for a continuous quantity α the inner product defined as a complex number in Eq.(2.59), as well as the expansion of a state in a discrete basis $\{|\alpha_n\rangle\}$ do not hold up. We cannot count a continuous value α in discrete steps. We can already guess that the inner product will have to be translated to a complex function $\phi(\alpha)$ and the sum over discrete indices has to be translated into an integral.

We have to extend our definition of proper vectors, i.e., which are countable infinite, and allow for so-called improper vectors. Mathematical this means that we let the improper states emerge out of the proper states via a limiting procedure. This is similar to the discussion for plane waves in Sec. 1.4.2, there we introduced artificial boundary conditions for the wave function. Similarly, we can confine a physical system in a finite volume V, which shall be large enough such that all relevant physical processes are not influenced by it. In a finite volume we have again quantization with discrete eigenvalues and eigenfunctions, this allows us to normalizes the eigenfunctions properly. In the limit $V \to \infty$ the distance between the eigenvalues approaches zero, and it becomes simpler to directly use a continuous basis.

We can sketch the limiting procedure as follows, in a discrete basis we have the countable proper vector $|\alpha_n\rangle$, with n being a positive integer. The inner product with an arbitrary state $|\psi\rangle$ is defined as $\langle\alpha_n|\psi\rangle$. Formally we can now re-write this vector as $|\alpha_{x,\Delta x}\rangle$, i.e., we replaced the integer n with the integer x and associated a $\Delta x = 1$ with it. The later is simply the distance between two integers x and $x' = x + 1$. For the inner product we also make this change in notation

$$\langle\alpha_n|\psi\rangle \quad \Leftrightarrow \quad \langle\alpha_{x,\Delta x}|\psi\rangle\,, \tag{2.71}$$

so the absolute value of the inner product is a real number which one can depict schematically for each integer value x, see Fig.2.2 left. Now we can extend this and assume that x can take arbitrary real numbers, and thus Δx can become arbitrary small, see Fig,2.2 right. We can now define a function $\psi(x)$ for the limit $\Delta x \to 0$:

$$\psi(x) = \lim_{\Delta x\to 0} \frac{1}{\sqrt{\Delta x}}\ \langle\alpha_{x,\Delta x}|\psi\rangle \equiv \langle\bar{\alpha}_x|\psi\rangle\,, \quad |\bar{\alpha}_x\rangle = \lim_{\Delta x\to 0} \frac{1}{\sqrt{\Delta x}}\ |\alpha_{x,\Delta x}\rangle\,, \tag{2.72}$$

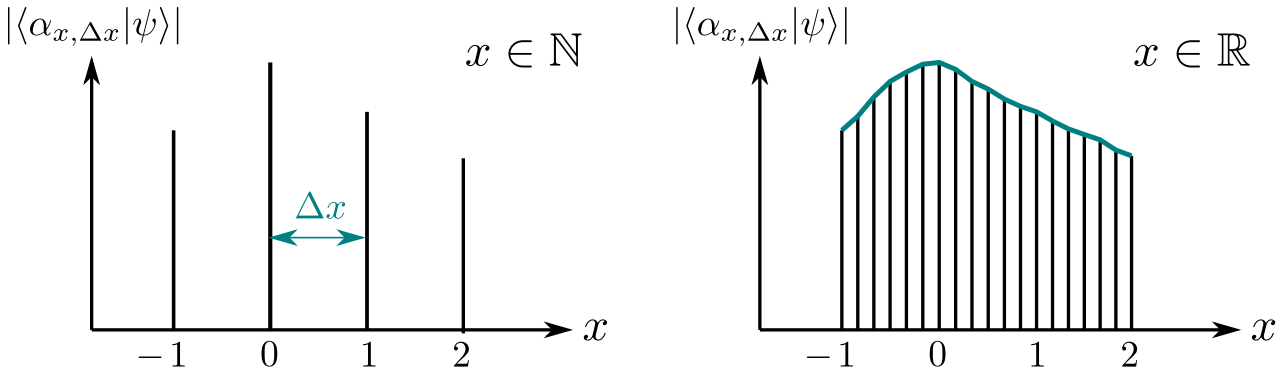

Figure 2.2: Transition from a discrete absolute value to a continuous one.

with the formal Dirac vector $|\bar{\alpha}_x\rangle$. Now we can expand our state in this new basis

$$|\psi\rangle = \lim_{\Delta x \to 0} \sum_x |\alpha_{x,\Delta x}\rangle \langle \alpha_{x,\Delta x}|\psi\rangle = \lim_{\Delta x \to 0} \sum_x |\bar{\alpha}_x\rangle \langle \bar{\alpha}_x|\psi\rangle \, \Delta x, \tag{2.73}$$

hence, as anticipated we can replace the sum with an integral and obtain

$$|\psi\rangle = \int dx \, |\bar{\alpha}_x\rangle \langle \bar{\alpha}_x|\psi\rangle \equiv \int dx \, |x\rangle \langle x|\psi\rangle \, . \tag{2.74}$$

This corresponds now to the expansion of a state $|\psi\rangle$ in the continuous basis $\{|\bar{\alpha}_x\rangle\} \equiv \{|x\rangle\}$. We can now multiply this with the dual vector $\langle x|$ and obtain

$$\langle x|\psi\rangle = \int dx' \, \langle x|x'\rangle \langle x'|\psi\rangle \, , \tag{2.75}$$

which can for arbitrary $|\psi\rangle$ only be fulfilled via

$$\langle x|x'\rangle = \delta(x - x') \quad \Rightarrow \quad \langle x|\psi\rangle = \int dx' \delta(x - x') \langle x'|\psi\rangle = \langle x|\psi\rangle \, . \tag{2.76}$$

This means that the improper Dirac vectors are 'normalized' to delta-functions, something we have already seen when we considered the probability amplitude for a free particle in the framework of the path integral approach. We identify $\psi(x) = \langle x|\psi\rangle$ as the wave function in position space, and the eigenfunctions of the position operators as

$$\phi_{x'}(x) = \langle x|x'\rangle = \delta(x - x'), \tag{2.77}$$

which are obviously not normalized, the norm is infinity. Hence, they are strictly no proper states in the Hilbert space, they are improper and can be interpreted as generalized states living in a so-called rigged Hilbert space. However, since the position is an observable quantity, the position operator has a complete set of eigenstates, and the set $\{|x\rangle\}$ are orthogonal vectors which span the state space. Being complete we have the identity

$$\hat{1} = \int dx \;\; |x\rangle \langle x| \, . \tag{2.78}$$

We already learned that we can represent the a state $|\psi\rangle$ in either position or momentum basis. We discussed that the connection between the wavefunction in position and momentum space is made via the Fourier transformation, we will now change slightly our definition for it

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp\ \psi(p) e^{\frac{i}{\hbar}px}, \quad \psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx\ \psi(x) e^{-\frac{i}{\hbar}px}. \tag{2.79}$$

This obviously still holds true and we can utilize it to discuss the properties of the momentum basis. First of all, the eigenfunction of the momentum operators in momentum space are

$$\hat{p}\,|p\rangle = p\,|p\rangle\,, \quad \phi_{p'}(p) = \langle p|p'\rangle = \delta(p-p'), \tag{2.80}$$

which we can express as well in the position space as

$$\phi_{p'}(x) = \langle x|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dp\ \phi_{p'}(p) e^{\frac{i}{\hbar}px} = \frac{1}{\sqrt{2\pi\hbar}} \int dp\ \delta(p-p') e^{\frac{i}{\hbar}px} = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}p'x}, \tag{2.81}$$

with the change of the definition of the Fourier transformations pre-factor we have ensured that $\langle x|p'\rangle = \langle p'|x\rangle^*$. With this we can show that the state $|\psi\rangle$ can be represented in both bases

$$|\psi\rangle = \int dp\ \psi(p)\,|p\rangle = \int dp\ \psi(p) \int dx\,|x\rangle\,\langle x|p\rangle = \int dx \int \frac{dp}{\sqrt{2\pi\hbar}} \psi(p) e^{\frac{i}{\hbar}px}\,|x\rangle = \int dx \psi(x)\,|x\rangle\,, \tag{2.82}$$

where we used in the second step the completeness relation for the position states, the set of states $\{|p\rangle\}$ is as well complete and thus has a equivalent completeness relation.

2.2.8 Discrete basis: proper vectors and matrix notation

Now we turn again to the realm of proper vectors as true elements of the Hilbert space. Meaning that we focus on a discrete set finite or countable infinite vectors $\{|\alpha_n\rangle\}$. There exist a range of bases to choose from, and the choice is guided by the physical problem at hand. We already have discussed the square integrable functions, examples for those defined on fixed intervals were given in Eq.(2.62). Formally they belong to the class of L_2 functions defined in the space $\Omega \in \mathbb{R}^n$, which posses the so-called L_2-norm

$$||f||_2 = \left[\int_\Omega d\mathbf{x}\ |f(\mathbf{x})|^2\right]^{\frac{1}{2}}, \tag{2.83}$$

important is that the integral of the absolute value squared is finite, and that we work now with a separable Hilbert space with a complete orthonormal systems. As we now work with a discrete basis we can actually start to work with a vector and matrix notation, i.e., if we have a state $|\psi\rangle$ we can represent it in a countable basis $\{|\phi_n\rangle\}$ with $n = 0, 1, 2, ..$ we can make the mapping

$$|\psi\rangle = \sum_n |\phi_n\rangle\,\langle\phi_n|\psi\rangle \longleftrightarrow \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix}, \quad \psi_n = \langle\phi_n|\psi\rangle\,, \tag{2.84}$$

and similarly for the dual vector

$$\langle\psi| = \sum_n \langle\psi|\phi_n\rangle \langle\phi_n| \longleftrightarrow \begin{pmatrix} \psi_0^* & \psi_1^* & \psi_2^* & \cdots \end{pmatrix}, \quad \psi_n^* = \langle\psi|\phi_n\rangle, \tag{2.85}$$

which is thus the complex transpose of the ket-vector. The elements of the vector are the so-called projections of the vector $|\psi\rangle$ onto the basis states $|\phi_n\rangle$. For simplicity regarding notation, let's assume we only have a two-dimensional vector $|\varphi\rangle$, i.e., $n = 0, 1$. The inner product of the corresponding state vectors becomes

$$\langle\varphi|\varphi\rangle = \sum_{n,n'=0}^{1} \psi_n^* \psi_{n'} \langle\phi_n|\phi_{n'}\rangle = \sum_{n=0}^{1} |\psi_n|^2 = \begin{pmatrix} \psi_0^* & \psi_1^* \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} = |\psi_0|^2 + |\psi_1|^2, \tag{2.86}$$

thus we have effectively two ways to calculate this inner product, we can just use the scalar product of two vectors, or we work with the bra-ket notation. We can as well calculate the outer product of the vector

$$|\varphi\rangle \langle\varphi| = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} \begin{pmatrix} \psi_0^* & \psi_1^* \end{pmatrix} = \begin{pmatrix} |\psi_0|^2 & \psi_0 \psi_1^* \\ \psi_0^* \psi_1 & |\psi_1|^2 \end{pmatrix}, \tag{2.87}$$

which results in a matrix. The diagonal elements here are the absolute value squared of our expansion coefficient, so we see that if we take the **trace of the matrix** (summation over all the diagonal components) we obtain

$$\mathrm{tr}\left(|\varphi\rangle \langle\varphi|\right) = |\psi_0|^2 + |\psi_1|^2 = \langle\varphi|\varphi\rangle, \tag{2.88}$$

i.e., we recover the inner product of the vectors. Alternatively we could have obtained this results via multiplying the outer products from both sides with the state vector:

$$\mathrm{tr}\left(|\varphi\rangle \langle\varphi|\right) = \sum_{n=0}^{1} \langle\phi_n| \left(|\varphi\rangle \langle\varphi|\right) |\phi_n\rangle = \sum_{n=0}^{1} \langle\phi_n|\varphi\rangle \langle\varphi|\phi_n\rangle = \sum_{n=0}^{1} |\psi_n|^2 = \langle\varphi|\varphi\rangle, \tag{2.89}$$

the trace of a matrix is an important quantity, it has the nice property that it is independent of the representation of the state. We can show this by considering another two-dimensional basis $\{|\theta_{0,1}\rangle\}$ and using the identity operation $\hat{1}$:

$$\begin{aligned}
\mathrm{tr}\left(|\varphi\rangle \langle\varphi|\right) &= \sum_{n=0}^{1} \langle\phi_n| \left(|\varphi\rangle \langle\varphi|\right) |\phi_n\rangle = \sum_{n=0}^{1} \langle\phi_n| \hat{1} \left(|\varphi\rangle \langle\varphi|\right) \hat{1} |\phi_n\rangle \\
&= \sum_{n,m,m'=0}^{1} \langle\phi_n|\theta_m\rangle \langle\theta_m|\varphi\rangle \langle\varphi|\theta_{m'}\rangle \langle\theta_{m'}|\phi_n\rangle \\
&= \sum_{n,m,m'=0}^{1} \langle\theta_{m'}|\phi_n\rangle \langle\phi_n|\theta_m\rangle \langle\theta_m|\varphi\rangle \langle\varphi|\theta_{m'}\rangle = \sum_{m,m'=0}^{1} \langle\theta_{m'}|\theta_m\rangle \langle\theta_m|\varphi\rangle \langle\varphi|\theta_{m'}\rangle \\
&= \sum_{m=0}^{1} \langle\theta_m|\varphi\rangle \langle\varphi|\theta_m\rangle .
\end{aligned} \tag{2.90}$$

hence, we obtain the same result in either basis, which is a characteristic we will re-use later.

Let us now turn to a concrete physical example for a discrete and two-dimensional system. From classical computation we know that information is encoded and processed in a binary fashion, i.e, one works with the logical states with the two possible values 0 and 1. Analogously one can define a so-called quantum bit (qubit), as the basic unit of quantum information. A qubit is a two-state quantum mechanical system and there exist a wide range of concrete realization, e.g., it could be the two electronic states of an atom, the polarization of a light field, the spin of a particle or an artificial atom realized in a superconducting circuit architecture. No matter which platform, the two-state quantum mechanical system can be represented with two orthonormal basis vectors:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.91}$$

Note, before we worked as well with the example of a two-dimensional system with state $|\varphi\rangle$, but we did not define the basis in more detail. Clearly, we can write the state $|\varphi\rangle$ in the just introduced basis $\{|0\rangle, |1\rangle\}$

$$|\varphi\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} = \psi_0 |0\rangle + \psi_1 |1\rangle, \tag{2.92}$$

which is a linear superposition of the basis states. For the complex coefficients we have

$$\langle\varphi|\varphi\rangle = |\psi_0|^2 + |\psi_1|^2 = 1, \tag{2.93}$$

to ensure normalization of the state. The value of the coefficients can be obtained via performing a measurement in the right basis. E.g, making a measurement in the basis $\{|0\rangle, |1\rangle\}$ gives 0 with probability $|\psi_0|^2$ and 1 with probability $|\psi_1|^2$. However, these do not have to be the best bases for the measurement. Assuming that the qubit is prepared in the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |0\rangle + \frac{e^{i\theta}}{\sqrt{2}} |1\rangle, \tag{2.94}$$

a measurement in the standard basis will yield the result 0 and 1 with the same probability 1/2. But is tells us nothing about the phase θ. To obtain any information about this phase we have to change the basis in which we measure. The so called phase estimation can be performed in the basis

$$|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle). \tag{2.95}$$

We can now rewrite the state of the qubit in this basis

$$\begin{aligned} |\Psi\rangle =& \frac{1}{\sqrt{2}} |0\rangle + \frac{e^{i\theta}}{\sqrt{2}} |1\rangle = \frac{1}{2} (|+\rangle + |-\rangle) + \frac{e^{i\theta}}{2} (|+\rangle - |-\rangle) \\ =& \frac{1}{2} \left(1 + e^{i\theta}\right) |+\rangle + \frac{1}{2} \left(1 - e^{i\theta}\right) |-\rangle. \end{aligned} \tag{2.96}$$

Using the Euler relation $e^{i\theta} = \cos\theta + i \sin\theta$ we obtain for the respective measurement probabilities

$$\begin{aligned} |+\rangle &\to \frac{1}{4} \left[(1 + \cos\theta)^2 + \sin^2\theta\right] = \frac{1}{2} [1 + \cos\theta] = \cos^2 \frac{\theta}{2}, \\ |-\rangle &\to \frac{1}{4} \left[(1 - \cos\theta)^2 + \sin^2\theta\right] = \frac{1}{2} [1 - \cos\theta] = \sin^2 \frac{\theta}{2}, \end{aligned} \tag{2.97}$$

thus, a measurement in the basis $\{|+\rangle, |-\rangle\}$ reveals information about the phase θ.

2.2.9 Linear Operators

We have already discussed how hermitian operators can represent physical quantities. Now we want to investigate in more detail how operations mediated by general operators manipulate a quantum state. Taking a quantum state $|\psi\rangle \in \mathcal{H}$ and acting with the operator $\hat{A}$ results in the state $|\phi\rangle \in \mathcal{H}$:

$$\hat{A}\,|\psi\rangle = |\phi\rangle\,, \tag{2.98}$$

in other words, the operator assigns a wave function (vector) to another wave function (vector). The expectation value of the operators with respect to a state $|\psi\rangle$ is a complex number

$$\langle\psi|\,\hat{A}\,|\psi\rangle = \langle\psi|\phi\rangle = \sum_{n,m} a_n a_m^*\,\langle\psi_m|\,\hat{A}\,|\psi_n\rangle\,, \tag{2.99}$$

with $a_{n,m} = \langle\psi_{n,m}|\phi\rangle$ as usual. To express the operator in matrix notation we start from

$$\hat{A} = \hat{1}\hat{A}\hat{1} = \sum_{n,n'} |\psi_n\rangle\,\langle\psi_n|\,\hat{A}\,|\psi_{n'}\rangle\,\langle\psi_{n'}| \equiv \sum_{n,n'} A_{n,n'}\,|\psi_n\rangle\,\langle\psi_{n'}| \tag{2.100}$$

which translates into the matrix notation

$$\hat{A} \longleftrightarrow \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n'} & \cdots \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n'} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n'} & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix}, \quad A_{n,n'} = \langle\psi_n|\,\hat{A}\,|\psi_{n'}\rangle\,, \tag{2.101}$$

when the Hilbert space is N-dimensional the matrix is a quadratic matrix of dimension $N \times N$, and for a countable infinite Hilbert space the matrix formally possesses infinite matrix elements. Taking an orthonormal basis $\{|\psi_n\rangle\}$ we can make an expansion in this basis

$$|\phi\rangle = \hat{A}\,|\psi\rangle = \sum_n \hat{A}\,|\psi_n\rangle\,a_n = \sum_n b_n\,|\psi_n\rangle\,, \quad a_n = \langle\psi_n|\psi\rangle\,, \tag{2.102}$$

with the coefficients

$$b_n = \langle\psi_n|\phi\rangle = \sum_{n'} \langle\psi_n|\,\hat{A}\,|\psi_{n'}\rangle\,a_{n'} = \sum_{n'} A_{n,n'}a_{n'} \tag{2.103}$$

which we can re-write in matrix notation

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n'} & \cdots \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n'} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n'} & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix} \tag{2.104}$$

hence, the new state vector arises via the multiplication of the original state vector with the operator matrix.

In case that $|\psi\rangle$ is an eigenstate of the operator $\hat{A}$ we have $\hat{A}|\psi\rangle = A|\psi\rangle$ with $A \in \mathbb{C}$ (real for hermitian operators). And thus we have $b_n = Aa_n$ meaning that the state vector after acting with the operator is parallel to the original state vector. We have the eigenvalue problem

$$\begin{pmatrix} A_{1,1}-A & A_{1,2} & \cdots & A_{1,n'} & \cdots \\ A_{2,1} & A_{2,2}-A & \cdots & A_{2,n'} & \cdots \\ \vdots & \vdots & & \vdots & \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n'}-A & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}, \tag{2.105}$$

as expected. The set of all eigenvalues of a system is also called a spectrum of the operator. If multiple linear independent eigenstates have the same eigenvalue A one calls the latter degenerate. And the grade of degeneracy is the maximal number of linear independent eigenstates belonging to the same eigenvalue.
We now want to introduce a number of special operators we will encounter in this course:

1. **Adjoint operator** $\hat{A}^\dagger$ is an element we have already introduced when we talked about hermitian operators, for the latter we had $\hat{O} = \hat{O}^\dagger$, i.e., an hermitian operator is self-adjoint (if the operator is bounded,see [4]). The definition of the adjoint operator is

$$\begin{aligned} \hat{A}|\psi\rangle = |\phi\rangle \quad &\Rightarrow \quad \langle\psi|\hat{A}^\dagger = \langle\phi|, \\ \hat{A}^\dagger|\psi\rangle = |\phi'\rangle \quad &\Rightarrow \quad \langle\psi|\hat{A} = \langle\phi'|, \end{aligned} \tag{2.106}$$

hence $\hat{A}^\dagger$ act in the dual Hilbert space $\mathcal{H}^*$ as the operator $\hat{A}$ in $\mathcal{H}$, and acting with $\hat{A}^\dagger$ on a state in the Hilbert space $\mathcal{H}$ is as acting with $\hat{A}$ in $\mathcal{H}^*$. There exist two important properties: the adjoined of the adjoined is again the operator itself, and the adjoint of a product of two operators $\hat{A}$ and $\hat{B}$ changes ordering, i.e.,

$$\left(\hat{A}^\dagger\right)^\dagger = \hat{A}, \quad \left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger. \tag{2.107}$$

Thus, the product of two hermitian operators is hermitian itself only if the two operators commute. Moreover, we can as well express the adjoint operator in the matrix notation

$$\hat{A}^\dagger = \left(\sum_{n,n'} A_{n,n'}|\psi_n\rangle\langle\psi_{n'}|\right)^\dagger = \sum_{n,n'} A^*_{n',n}|\psi_{n'}\rangle\langle\psi_n| \tag{2.108}$$

hence, adjoint means taking the transpose and complex conjugate, with the upper matrix representation for the operator $\hat{A}$ this means for the adjoint matrix

$$\hat{A}^\dagger \longleftrightarrow \begin{pmatrix} A^*_{1,1} & A^*_{2,1} & \cdots & A^*_{n',1} & \cdots \\ A^*_{1,2} & A^*_{2,2} & \cdots & A^*_{n',2} & \cdots \\ \vdots & \vdots & & \vdots & \\ A^*_{1,n} & A^*_{2,n} & \cdots & A^*_{n',n} & \cdots \\ \vdots & \vdots & & \vdots & \end{pmatrix}, \tag{2.109}$$

and for the the expansion of the dual vector $|\phi\rangle$ we have

$$\langle\phi| = \langle\psi|\hat{A}^\dagger = \sum_n b^*_n\langle\psi_n|, \quad b^*_n = \langle\phi|\psi_n\rangle = \sum_{n'} A_{n',n}a^*_{n'} \tag{2.110}$$

which we can again write the multiplication of a vector and a matrix.

2. **Projection operator** $\hat{\mathcal{P}}$ corresponds to a diagonal dyadic product of the state $|\theta\rangle$ with itself, and if the state is normalized $||\theta|| = 1$, the operator projects a state $|\psi\rangle$ exactly into this state $|\theta\rangle$:

$$\hat{\mathcal{P}} = |\theta\rangle \langle\theta| , \quad \hat{\mathcal{P}} |\psi\rangle = |\theta\rangle \langle\theta|\psi\rangle = c_\theta |\theta\rangle , \; c_\theta \in \mathbb{C}. \tag{2.111}$$

The projection operator is hermitian and so-called idempotent, meaning that

$$\hat{\mathcal{P}}^\dagger = (|\theta\rangle \langle\theta|)^\dagger = \hat{\mathcal{P}}, \quad \hat{\mathcal{P}}^2 = |\theta\rangle \langle\theta|\theta\rangle \langle\theta| = |\theta\rangle \langle\theta| = \hat{\mathcal{P}}, \tag{2.112}$$

respectively. The projection operator can be used to represent a measurement in the basis $|\theta\rangle$, this kind of measurement is also called projective measurement or von Neumann measurement. This also already hints towards the so-called measurement postulate, which tells us how much information about a quantum state we can access. We are free to choose any basis for our measurement and then check in which basis state the system is in. However, we will not know exactly in which state $|\psi\rangle$ the system actually was before the measurement, but we do know that after the measurement it is in the measured basis state.

3. **Unitary operator** $\hat{U}$ mediates a transformation which leave the scalar product unchanged, meaning that a so-called unitary transformation does not alter the physics. Taking two arbitrary vectors $|\psi\rangle$ and $|\phi\rangle$ undergoing the transformation

$$|\psi'\rangle = \hat{U} |\psi\rangle , \quad |\phi'\rangle = \hat{U} |\phi\rangle \quad \Rightarrow \quad \langle\phi'| = \langle\phi| \hat{U}^\dagger, \tag{2.113}$$

we obtain for the scalar product between the transformed vectors

$$\langle\phi'|\psi'\rangle = \langle\phi| \hat{U}^\dagger \hat{U} |\psi\rangle \stackrel{!}{=} \langle\phi|\psi\rangle \tag{2.114}$$

leading to the requirement

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger \hat{U} = \hat{1}, \quad \hat{U} = \hat{U}^{-1}, \tag{2.115}$$

which denoted the condition for a unitary operator (matrix) and introduces the notion of an inverse operator $\hat{U}^{-1}$. A unitary transformation of a system can be very practical, as changing into a different *frame* of reference can simplify the (analytical) treatment of a system. Besides the state vectors one can as well transform an operator

$$\hat{A}' = \hat{U} \hat{A} \hat{U}^\dagger, \quad \hat{A} = \hat{U}^\dagger \hat{A}' \hat{U} = \hat{U}^\dagger \hat{U} \hat{A} \hat{U}^\dagger \hat{U} = \hat{A}. \tag{2.116}$$

An simple example for a unitary transformation involves a hermitian operators $\hat{O} = \hat{O}^\dagger$:

$$\hat{U} = e^{+i\hat{O}}, \;\; \hat{U}^\dagger = e^{-i\hat{O}}, \;\; \Rightarrow \;\; \hat{U}\hat{U}^\dagger = \hat{1}, \tag{2.117}$$

applying this transformation to an arbitrary state $|\psi\rangle$ yields

$$|\psi'\rangle = \hat{U} |\psi\rangle = e^{+i\hat{O}} |\psi\rangle = \sum_{n=0}^{\infty} i^n \frac{\hat{O}^n}{n!} |\psi\rangle , \tag{2.118}$$

corresponding to a successive application of the operators $\hat{O}$ on the state $|\psi\rangle$, which can have or have not a closed expression. For example, if we consider the case that we have an eigenstate $|\psi_o\rangle$ of the operators $\hat{O}$ with eigenvalue O, i.e., $\hat{O}|\psi_o\rangle = O|\psi_o\rangle$, we obtain

$$\begin{aligned}|\psi_o'\rangle =& e^{+i\hat{O}}|\psi_o\rangle = \sum_{n=0}^{\infty} i^n \frac{\hat{O}^n}{n!}|\psi_o\rangle = |\psi_o\rangle + i\hat{O}|\psi_o\rangle + \frac{i^2}{2!}\hat{O}\hat{O}|\psi_o\rangle + ... \\ =& |\psi_o\rangle + iO|\psi_o\rangle + \frac{i^2}{2!}O^2|\psi_o\rangle + ... = e^{+iO}|\psi_o\rangle, \end{aligned} \tag{2.119}$$

thus, after applying the unitary transformation, the resulting state inherits simply a phase factor corresponding to the respective eigenvalue. The general transformation for an operator $\hat{A}$ becomes

$$\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger = e^{+i\hat{O}}\hat{A}e^{-i\hat{O}} = \hat{A} + i\left[\hat{O},\hat{A}\right] + \frac{(i)^2}{2!}\left[\hat{O},\left[\hat{O},\hat{A}\right]\right] + ... \tag{2.120}$$

corresponding to the Baker–Campbell–Hausdorff formula for two operators. Again this expansion can be for some cases exact, for others one has to truncate the series (perturbation theory). In case the operator $\hat{A}$ and $\hat{O}$ commute we have $\hat{A}' = \hat{A}$, i.e., the transformation does not affect the operator $\hat{A}$.

4. **The density operator** $\hat{\rho}$ is an crucial object to represent the state of a quantum system. As we discussed above, the state of a quantum system can be modeled as a vector in complex vector space, labeled $|\psi\rangle$ in Dirac notation. Describing a quantum state by its density matrix is a fully general alternative formalism to describing a quantum state, and in addition, it is also possible to use the density matrix in case a system is in a statistical mixture as outlined below. The density matrix is the quantum-mechanical analogue to a phase-space probability measure (probability distribution of position and momentum) in classical statistical mechanics. For a pure state the density matrix is defined as the outer product of the state vector:

$$\hat{\rho} = |\psi\rangle\langle\psi|, \quad \hat{\rho} = \hat{\rho}^\dagger, \tag{2.121}$$

thus the density matrix is a hermitian operator. The trace of the matrix

$$\mathrm{tr}\left[\hat{\rho}\right] = \mathrm{tr}\left[|\psi\rangle\langle\psi|\right] = \langle\psi|\psi\rangle = \sum_{n,m} c_n c_m^* \langle m|n\rangle = \sum_n |c_n|^2 = 1. \tag{2.122}$$

where we used that the trace of the outer product is the inner product. We also have the property that $\hat{\rho} \geq 0$, which means that all of its eigenvalues are non-negative. Moreover, for pure states the trace of the density matrix squared yields

$$\mathrm{tr}\left[\hat{\rho}^2\right] = \mathrm{tr}\left[|\psi\rangle\langle\psi|\,|\psi\rangle\langle\psi|\right] = \langle\psi|\psi\rangle^2 = 1. \tag{2.123}$$

Crucially, a system can as well be in an ensemble of states $|\psi_j\rangle$, which do not have to be orthogonal (meaning they are not basis states), and each state is associated with a probability p_j. The density matrix for a so-called mixed state becomes

$$\hat{\rho} = \sum_j p_j |\psi_j\rangle\langle\psi_j|, \tag{2.124}$$

which obeys still the following properties:

$$\hat{\rho} = \hat{\rho}^\dagger, \quad \sum_j p_j = 1, \quad \mathrm{tr}\left[\hat{\rho}\right] = 1, \quad \hat{\rho} \geq 0, \tag{2.125}$$

thus the density matrix for a mixed state is hermitian and positive definite. However, one quality does not longer hold for mixed states:

$$\mathrm{tr}\left[\hat{\rho}^2\right] = \mathrm{tr}\left[\left(\sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right|\right)\left(\sum_k p_k \left|\psi_k\right\rangle \left\langle\psi_k\right|\right)\right] = \sum_{j,k} p_j p_k \left|\left\langle\psi_j|\psi_k\right\rangle\right|^2 \leq 1, \tag{2.126}$$

which is then also a simple measure if the state is pure or not, i.e., a system is in a combination of states states $|\psi_j\rangle$ is also called an impure state. Only in the case when the density matrix consists of only a single state, i.e., $p_j = \delta_{j,j'}$, we have a pure state ($\hat{\rho}^2 = 1$), which is as well a rank-1 projection operator. The density matrix is the most general representation of a quantum state. It is a positive matrix, i.e., all eigenvalues are non-negative (but can be zero). The expectation value of a physical quantity does becomes

$$\left\langle \hat{A} \right\rangle = \sum_j p_j \left\langle\psi_j\right| \hat{A} \left|\psi_j\right\rangle = \mathrm{tr}\left[\hat{A}\left(\sum_j p_j \left|\psi_j\right\rangle \left\langle\psi_j\right|\right)\right] \equiv \mathrm{tr}\left[\hat{A}\hat{\rho}\right], \quad \sum_j p_j = 1, \tag{2.127}$$

thus the density matrix $\hat{\rho}$ is enough to calculate expectation value of any operator.

Note that a mixed state is not the same as a superposition. For a mixed state the system is in either states with probability p_m, while if the system is in a superposition it is in neither of the states. For example consider the superposition state $|\psi_{12}\rangle = \alpha_1 |\psi_1\rangle + \alpha_2 |\psi_2\rangle$, the corresponding density operator reads

$$\hat{\rho}_{12} = \left|\psi_{12}\right\rangle \left\langle\psi_{12}\right| = \sum_{j=1,2} \left|\alpha_j\right|^2 \left|\psi_j\right\rangle \left\langle\psi_j\right| + \alpha_1\alpha_2^* \left|\psi_1\right\rangle \left\langle\psi_2\right| + \alpha_1^*\alpha_2 \left|\psi_2\right\rangle \left\langle\psi_1\right|, \tag{2.128}$$

the first two terms would appear for a bi-bipartite mixed state, however, the remaining terms make clear that it isn't a mixed state, but rather a superposition of both states.

As an example we can discuss the density matrix of a two-level system. Let's assume that we have the ground $|0\rangle$ and excited state $|1\rangle$ prepared with equal probability of 1/2. We can associate this with a density operator

$$\hat{\rho} = \frac{1}{2}\left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\left|1\right\rangle \left\langle 1\right| = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \tag{2.129}$$

obviously the trace is equal to 1 and we obtain for the trace of the density matrix squared

$$\mathrm{tr}\left[\hat{\rho}^2\right] = \mathrm{tr}\left[\left(\frac{1}{2}\left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\left|1\right\rangle \left\langle 1\right|\right)\left(\frac{1}{2}\left|0\right\rangle \left\langle 0\right| + \frac{1}{2}\left|1\right\rangle \left\langle 1\right|\right)\right] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \tag{2.130}$$

as expected the obtained value is smaller than one as we have a mixed state. This is actually the smallest value this quantity can take. This becomes clear if we consider the

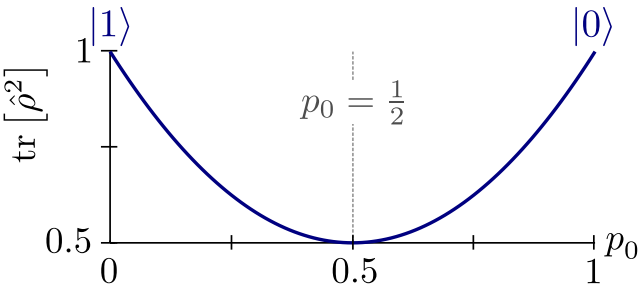

Figure 2.3: Example of a two level system in the state $\hat{\rho}$, the trace of the density matrix squared has it's minimum at $p_0 = 1/2$, here the state is maximally mixed.

general case of having the probabilities p_0 and p_1 to prepare $|0\rangle$ and $|1\rangle$ respectively, the density operator becomes

$$\hat{\rho} = \begin{pmatrix} p_0 & 0 \\ 0 & p_1 \end{pmatrix}, \quad \mathrm{tr}\left[\hat{\rho}^2\right] = p_0^2 + p_1^2 = p_0^2 + (p_0 - 1)^2, \tag{2.131}$$

this expression takes its minimum value for $p_0 = 1/2$, see Fig.2.3. We can also change the basis in which the state is prepared. Having the situation of preparing $|+\rangle$ and $|-\rangle$ with equal probabilities is captured as

$$\hat{\rho} = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \tag{2.132}$$

which is just the same density matrix as for the standard basis! So we are faced here with the irritating situation that two different mixed ensembles are described by the same density operator. Let us face it: There are many different ways of preparing the same density operator. And once we arrive at a given density operator, there is no way to reconstruct the mixed ensemble that can be held responsible for the density operator. In retrospect, there always would have been infinitely many other ways of preparing the same density operator (unless it is a pure state).

2.3 Prominent examples of quantum systems

2.3.1 The atom as a two-level system

We have already encountered the most simple quantum system consisting only of two distinguishable quantum states. An example can be the ground and excited state of an atom which we can label as $|g\rangle$ and $|e\rangle$ respectively. The Hamiltonian for such a system

$$\hat{H} = \hbar\frac{\Omega}{2}\left(|e\rangle\langle e| - |g\rangle\langle g|\right) = \hbar\frac{\Omega}{2}\,\hat{\sigma}_z, \tag{2.133}$$

with Ω as the atomic transition frequency and we have introduced the Pauli-spin operator $\hat{\sigma}_z$. These **Pauli-spin operators** $\hat{\sigma}$ are crucial for modeling a two-state system, they are defined as follows

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.134}$$

from which we see that they are hermitian, and that the eigenvalues of each Pauli-spin operator takes the values ± 1. They obey the commutation relation

$$[\hat{\sigma}_n, \hat{\sigma}_m] = 2i\varepsilon_{nmk}\, \hat{\sigma}_k \tag{2.135}$$

where we have introduced the Levi-Civita symbol $\epsilon_{n,m,k}$, which is defined as

$$\varepsilon_{nmk} = \begin{cases} +1 & \text{if } (n,m,k) \text{ is } (x,y,z),(y,z,x), \text{ or } (z,x,y), \\ -1 & \text{if } (n,m,k) \text{ is } (z,y,x),(x,z,y), \text{ or } (y,x,z), \\ 0 & \text{if } n=m, \text{ or } m=k, \text{ or } k=n, \end{cases} \tag{2.136}$$

changing sign depending on the order of permutation. We can straightforwardly see how the ground and excited state are denoted in vector notation

$$|g\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \;\Rightarrow\; |g\rangle\langle g| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad |e\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \;\Rightarrow\; |e\rangle\langle e| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.137}$$

to note is here that this is a distinct definition from the one used in quantum information, where the logical zero $|0\rangle$ would be modeled with the vector $(1,0)^T$. This notation here is used from the physics community, and two avoid confusion we will use the labels g, e and not $0, 1$ here. We can define the operators

$$\begin{aligned} \hat{\sigma}_+ =& \frac{1}{2}(\hat{\sigma}_x + i\hat{\sigma}_y) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{i}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = |e\rangle\langle g|, \\ \hat{\sigma}_- =& \frac{1}{2}(\hat{\sigma}_x - i\hat{\sigma}_y) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \frac{i}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = |g\rangle\langle e|, \end{aligned} \tag{2.138}$$

which are called raising and lowering operators as they mediate transitions between the levels

$$\hat{\sigma}_+ |g\rangle = |e\rangle\langle g|g\rangle = |e\rangle, \quad \hat{\sigma}_+ |e\rangle = 0 \quad \hat{\sigma}_- |e\rangle = |g\rangle\langle e|e\rangle = |g\rangle, \quad \hat{\sigma}_- |g\rangle = 0. \tag{2.139}$$

These operators are not hermitian operators, but are hermitian conjugate, meaning that $\hat{\sigma}_+^\dagger = \hat{\sigma}_-$. We can now also identify how the remaining Pauli-spin operators operate on a state, therefor we express them in therms of the lowering and raising operators

$$\hat{\sigma}_x = \hat{\sigma}_+ + \hat{\sigma}_- = |e\rangle\langle g| + |g\rangle\langle e| \tag{2.140}$$

$$\hat{\sigma}_y = -i(\hat{\sigma}_+ - \hat{\sigma}_-) = -i(|e\rangle\langle g| - |g\rangle\langle e|). \tag{2.141}$$

Taking now that we have prepared an atom in a superposition state of the form

$$|\psi\rangle = \alpha|g\rangle + \beta|e\rangle, \quad \alpha, \beta \in \mathbb{C}, \;\; |\alpha|^2 + |\beta|^2 = 1, \tag{2.142}$$

we can operate for example with the Pauli-x operator

$$\begin{aligned} |\psi'\rangle =& \hat{\sigma}_x |\psi\rangle = (|e\rangle\langle g| + |g\rangle\langle e|)(\alpha|g\rangle + \beta|e\rangle) \\ =& \alpha|e\rangle\langle g|g\rangle + \beta|e\rangle\langle g|e\rangle + \alpha|g\rangle\langle e|g\rangle + \beta|g\rangle\langle e|e\rangle \\ =& \alpha|e\rangle + \beta|g\rangle, \end{aligned} \tag{2.143}$$

thus, the resulting state has swapped eigenstates $|e,g\rangle$, similarly, the Pauli-y operator mediates such a swapping and changes the phase.To note is, that in quantum information processing the Pauli operators mediate fundamental single qubit gates. For example, the x-gate corresponds to the logical NOT gate for classical computers.

2.3.2 Multi-level systems

Clearly, discreet quantum systems can have more than two levels and here we can introduce the creation and annihilation operators of so-called Fock states. These Fock states are quantum states with a well defined number of excitations, e.g., the Fock state $|n\rangle$ contains n quanta or excitations. These operators mediate transitions between energy levels just as in the case of lowering and raising operators for the two-level system. This formulation of quantum mechanics is also referred to as second quantization and will be more intensively discussed in the second course of quantum mechanics. The creation and annihilation operators act on the normalized Fock states $|n\rangle$ as follows

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\,, \quad \hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \tag{2.144}$$

thus, the annihilation operator $\hat{a}$ destroys an excitation, while the creation operator $\hat{a}^\dagger$ creates an excitation. The operators fulfill so-called bosonic commutation relations $[\hat{a}, \hat{a}^\dagger] = 1$. Bosonic systems are for example photons (particles of light) or phonons (particles of sound), however, for example creation and annihilation operators for electrons have a different commutation relation (electrons are so-called fermions). We will clarify more the distinction between bosons and fermions later in this course. To note is that the expectation value of the annhilation operator with respect to a Fock state $|n\rangle$ yields

$$\langle\hat{a}\rangle = \langle n|\,\hat{a}\,|n\rangle = \sqrt{n}\,\langle n|n-1\rangle = \sqrt{n+1}\,\langle n+1|n\rangle = 0, \tag{2.145}$$

and thus vanishes, this hold as well true for the creation operator. Moreover, it follows as well that $\langle\hat{a}^m\rangle = 0$ holds for all $m = 1, 2, ...$ This means that the Fock states are no eigenstates of the creation and annhilation operators. An operator with a finite expectation value is the so-called number operator $\hat{n} = \hat{a}^\dagger\hat{a}$, as $|n\rangle$ is an eigenstate of $\hat{n}$:

$$\hat{n}\,|n\rangle = \hat{a}^\dagger\hat{a}\,|n\rangle = \sqrt{n}\hat{a}^\dagger\,|n-1\rangle = n\,|n\rangle\,, \quad \langle\hat{n}\rangle = \langle n|\,\hat{n}\,|n\rangle = n, \tag{2.146}$$

meaning that the number operator acting on the state reveals how many excitations are in the system (n). Important is that acting with the annhilation operator or the number operator on the ground state $|0\rangle$ we have

$$\hat{a}\,|0\rangle = 0, \quad \hat{n}\,|0\rangle = 0, \tag{2.147}$$

as one cannot destroy a non-existing excitation, in other words, it implies that we cannot generate lower and lower eigenstates with energies unbounded from below. In contrast, when we use the creation operator

$$\begin{aligned}
\hat{a}^\dagger\,|0\rangle &= |1\rangle\,, \\
\hat{a}^\dagger\hat{a}^\dagger\,|0\rangle &= \hat{a}^\dagger\,|1\rangle = \sqrt{2}\,|2\rangle\,, \\
\hat{a}^\dagger\hat{a}^\dagger\hat{a}^\dagger\,|0\rangle &= \hat{a}^\dagger\hat{a}^\dagger\,|1\rangle = \sqrt{2}\hat{a}^\dagger\,|2\rangle = \sqrt{2}\sqrt{3}\,|3\rangle\,, \\
&\vdots \\
\left(\hat{a}^\dagger\right)^n|0\rangle &= \sqrt{n!}\,|n\rangle \quad \Rightarrow \quad |n\rangle = \frac{\left(\hat{a}^\dagger\right)^n}{\sqrt{n!}}\,|0\rangle
\end{aligned} \tag{2.148}$$

thus we can create the state $|n\rangle$ via operating n-times with the creation operator on the ground state. This expression can be helpful when dealing with Fock-states in general.

Obviously the creation and annhilation operators are non-hermitian operators and thus no observables. But we can define Hermitian operators as linear combinations of $\hat{a}$ and $\hat{a}^\dagger$

$$\hat{X}_1 = \frac{1}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right), \quad \hat{X}_2 = \frac{-i}{\sqrt{2}}\left(\hat{a} - \hat{a}^\dagger\right), \tag{2.149}$$

which obey the commutation relation

$$\left[\hat{X}_1, \hat{X}_2\right] = \frac{-i}{2}\left[\left(\hat{a} + \hat{a}^\dagger\right), \left(\hat{a} - \hat{a}^\dagger\right)\right] = \frac{i}{2}\left[\hat{a}, \hat{a}^\dagger\right] - \frac{i}{2}\left[\hat{a}^\dagger, \hat{a}\right] = i, \tag{2.150}$$

and are thus non-commuting, we also call them orthogonal. As these are now hermitian operators we can check if they are observables by calculating the expectation value of $\hat{X}_1$ with respect to some arbitrary (Fock) state $|n\rangle$:

$$\left\langle \hat{X}_1 \right\rangle = \langle n| \hat{X}_1 |n\rangle = \frac{1}{\sqrt{2}} \langle n| \left(\hat{a} + \hat{a}^\dagger\right) |n\rangle = \frac{1}{\sqrt{2}} \left(\sqrt{n} \langle n|n-1\rangle + \sqrt{n+1} \langle n|n+1\rangle\right) = 0, \tag{2.151}$$

thus again we have a vanishing expectation value for the operator in this basis. However, we can interpret the expectation value as the mean value obtained in a measurement, and for the hermitian operators $\hat{X}_{1,2}$ the average value in this basis vanishes. In contrast, the variance of these operators does not vanish, e.g. for the $\hat{X}_1$ operator we obtain

$$\begin{aligned}\left\langle \left(\hat{X}_1 - \left\langle \hat{X}_1 \right\rangle\right)^2 \right\rangle = \left\langle \hat{X}_1^2 \right\rangle &= \frac{1}{2} \langle n| \left(\hat{a} + \hat{a}^\dagger\right)^2 |n\rangle = \frac{1}{2} \langle n| \left(\hat{a}^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}^{\dagger 2}\right) |n\rangle \\ &= \frac{1}{2} \langle n| \left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) |n\rangle = \frac{1}{2} \langle n| \left(1 + 2\hat{a}^\dagger\hat{a}\right) |n\rangle = n + \frac{1}{2},\end{aligned} \tag{2.152}$$

which is finite even for zero excitations! This is connected to the so-called vacuum fluctuations in quantum mechanics which we will discuss in more detail at a later point in this course.
We can now ask the question what would happen if we measure the system (and thus the hermitian operators $\hat{X}_{1,2}$) in an eigenstate of the creation and annhilation operators? Or even do these eigenstates exist? For this we would have the eigenvalue equation

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle \tag{2.153}$$

with the new eigenstates $|\alpha\rangle$ and the corresponding complex eigenvalue α. An ansatz for these eigenstates is to express them in the Fock basis

$$|\alpha\rangle = \sum_{n=0}^{\infty} c_n |n\rangle, \tag{2.154}$$

and we can determine the coefficients c_n in an iterative fashion

$$\hat{a} |\alpha\rangle = \sum_{n=0}^{\infty} c_n \hat{a} |n\rangle = \sum_{n=0}^{\infty} c_n \sqrt{n} |n-1\rangle = \alpha |\alpha\rangle = \alpha \sum_{n=0}^{\infty} c_n |n\rangle, \tag{2.155}$$

for this to hold true we have the relation

$$c_{n+1}\sqrt{n+1} = \alpha c_n \quad \Rightarrow \quad c_{n+1} = \frac{\alpha}{\sqrt{n+1}} c_n, \quad c_n : \; c_0, c_1 = \frac{\alpha}{\sqrt{1}}, c_2 = \frac{\alpha^2}{\sqrt{2}}, ... \tag{2.156}$$

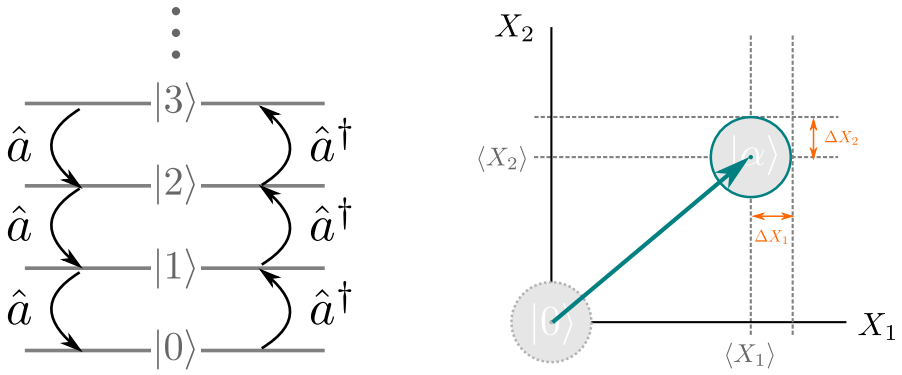

Figure 2.4: Left: Latter operators $\hat{a}$ and $\hat{a}^\dagger$ raising or lowering a state by one excitation. Right: phase space representation of a coherent state, which is a displaced vacuum state.

resulting in the expression for the eigenstate of $\hat{a}$ in the Fock basis as

$$|\alpha\rangle = c_0 \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,, \tag{2.157}$$

what remains is to define the coefficient c_0 via norm of the vector

$$||\alpha||^2 = |c_0|^2 \sum_{n,n'=0}^{\infty} \frac{\alpha^{*n}}{\sqrt{n!}} \frac{\alpha^{n'}}{\sqrt{n'!}} \langle n|n'\rangle = |c_0|^2 \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = |c_0|^2 e^{|\alpha|^2}, \tag{2.158}$$

with we can now write the eigenstate of the annhilation operators as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,, \tag{2.159}$$

which is called a coherent state or also Glauber state, which are so-called displaced vacuum states, and we can define a displacement operator preparing these states out of vacuum:

$$|\alpha\rangle = \mathcal{D}(\alpha) |0\rangle \,, \quad \mathcal{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}, \tag{2.160}$$

which can be straightforwardly verified (using Baker-Hausdorff formula and Eq.(2.148))

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} e^{\alpha\hat{a}^\dagger} e^{-\alpha^*\hat{a}} |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\left(\alpha\hat{a}^\dagger\right)^n}{n!} |0\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,. \tag{2.161}$$

To note is that the coherent states $|\alpha\rangle \in \mathcal{H}$ are the eigenstates of the annhilation operator $\hat{a}$, while there are no eigenstates for the creation operator in $\mathcal{H}$ (but we have $\langle\alpha|\,\hat{a}^\dagger = \langle\alpha|\,\alpha^*$ in the dual space $\mathcal{H}^*$). A special aspect of coherent states is that they are not exactly orthogonal states:

$$|\,\langle\beta|\alpha\rangle\,| = e^{-\frac{1}{2}|\alpha|^2} e^{-\frac{1}{2}|\beta|^2} \left| \sum_{n=0}^{\infty} \frac{\alpha^n \beta^{*n}}{n!} \right| = e^{-\frac{1}{2}|\alpha|^2} e^{-\frac{1}{2}|\beta|^2} \left| e^{\alpha\beta^*} \right| = e^{-\frac{1}{2}|\alpha-\beta|^2}, \tag{2.162}$$

thus only for large enough distance between the eigenstates they become orthogonal. Nevertheless the representation in coherent states is especially important in the field of quantum optics, as they refer to one of the most common states of the quantized electromagnetic field. Moreover they are specific states of the quantum harmonic oscillator which we will discuss in the next section.

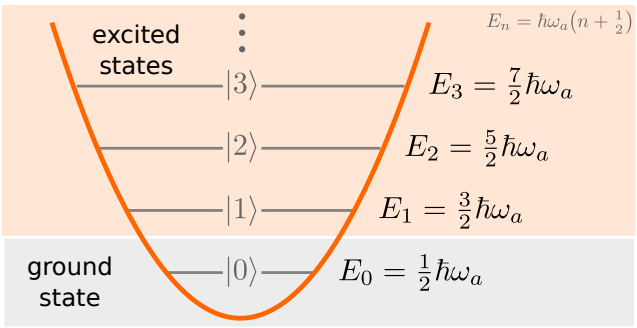

Figure 2.5: Discrete and equidistance energy levels of the quantum harmonic oscillator.

2.3.3 The quantum harmonic oscillator

As an example of a multi-level system we can consider the quantum version of an harmonic oscillator. The corresponding Hamiltonian can straightforwardly deducted via the correspondence principle:

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{m\omega_a^2}{2}\hat{x}^2, \tag{2.163}$$

where m denotes the mass of the oscillator and ω_a it's frequency. This harmonic oscillator model is used for all kinds of harmonic oscillations, one can use it to model electromagnetic modes (photons), vibrational modes or sound waves (phonons). We change now into the basis of creation and annihilation operators via

$$\hat{x} = x_{\mathrm{ZPF}}\left(\hat{a} + \hat{a}^\dagger\right), \quad \hat{p} = -im\omega_a x_{\mathrm{ZPF}}\left(\hat{a} - \hat{a}^\dagger\right), \quad x_{\mathrm{ZPF}} = \sqrt{\frac{\hbar}{2m\omega_a}}, \tag{2.164}$$

here x_{ZPF} denotes the zero-point fluctuation amplitude of the oscillator, i.e., it tells us about the spread of the coordinate in the ground state and set's the natural length scale of the oscillator. Note, the operators $\hat{x}$ and $\hat{p}$ are hermitian as expected as these are measurable quantities or observables. In contrast, the operators $\hat{a}, \hat{a}^\dagger$ are non-hermitian and hence, cannot be measured directly. The Hamiltonian in this new basis takes a rather simple form

$$\hat{\mathcal{H}} = \frac{\hbar\omega_a}{4}\left[\left(\hat{a}^\dagger - \hat{a}\right)\left(\hat{a} - \hat{a}^\dagger\right) + \left(\hat{a} + \hat{a}^\dagger\right)\left(\hat{a} + \hat{a}^\dagger\right)\right] = \frac{\hbar\omega_a}{2}\left[\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger\right] = \hbar\omega_a\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \tag{2.165}$$

in the last step we used the commutation relations for bosonic operators: $[\hat{a}, \hat{a}^\dagger] = 1$. We can determine the discrete energies of the oscillator as we already know that the Fock states are the eigenstates of the number operator $\hat{n} = \hat{a}^\dagger\hat{a}$:

$$E_n = \langle n|\,\hat{H}\,|n\rangle = \langle n|\,\hbar\omega_a\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|n\rangle = \hbar\omega_a\left(n + \frac{1}{2}\right), \tag{2.166}$$

and the distance between two energy levels

$$\Delta E = E_{n+1} - E_n = \hbar\omega_a\left(n + \frac{3}{2}\right) - \hbar\omega_a\left(n + \frac{1}{2}\right) = \hbar\omega_a, \tag{2.167}$$

is independent of the value of n, hence the energy levels are equidistant, see Fig.2.5. Moreover, we find that even for zero excitations in the system, i.e., $n = 0$ aka the ground state, we have the zero-point energy of the oscillator. This is a true quantum feature and is a consequence of the uncertainty principle. To see this we can determine the variances of the position and momentum operator in the Fock basis, with $\langle\hat{x}\rangle = \langle\hat{p}\rangle = 0$ we obtain

$$\begin{aligned}\Delta p^2 &= \langle\hat{p}^2\rangle = -m^2\omega_a^2 x_{\mathrm{ZPF}}^2 \left\langle \left(\hat{a} - \hat{a}^\dagger\right)^2 \right\rangle = \frac{\hbar m\omega_a}{2}(2n+1),\\ \Delta x^2 &= \langle\hat{x}^2\rangle = x_{\mathrm{ZPF}}^2 \left\langle \left(\hat{a} + \hat{a}^\dagger\right)^2 \right\rangle = \frac{\hbar}{2m\omega_a}(2n+1),\end{aligned} \tag{2.168}$$

which we insert into the uncertainty principle for momentum and position:

$$\Delta x\Delta p \geq \frac{\hbar}{2} \quad \Rightarrow \quad \frac{\hbar}{2}(2n+1) \geq \frac{\hbar}{2}, \tag{2.169}$$

which is always fulfilled for finite values of n, and the equal sign corresponds to the case $n = 0$, the ground state of the oscillator. For the ground state the variance of position and momentum becomes minimal, thus one calls the ground state a **minimum uncertainty state**. Interestingly, repeating this calculation for the coherent states gives the same results and hence, coherent states are minimum uncertainty states.
We can now as well determine the wave function in position space for the harmonic oscillator, for this we start by expressing our creation and annhilation operators in terms of the position and momentum operators

$$\hat{a} = \frac{1}{2x_{\mathrm{ZPF}}}\left(\hat{x} + \frac{i}{m\omega_a}\hat{p}\right), \qquad \hat{a}^\dagger = \frac{1}{2x_{\mathrm{ZPF}}}\left(\hat{x} - \frac{i}{m\omega_a}\hat{p}\right) \qquad x_{\mathrm{ZPF}} = \sqrt{\frac{\hbar}{2m\omega_a}}. \tag{2.170}$$

We have discussed that the annhilation operator acting on the ground state gives us zero, hence multiplying from the left with an eigenvector of position space $|x\rangle$ has to vanish as well:

$$\hat{a}\,|0\rangle = 0 \quad \Rightarrow \quad \langle x|\,\hat{a}\,|0\rangle = 0, \tag{2.171}$$

with this as a starting point we can use the completeness relation for the position eigenstates, as well as the annhilation operator expressed with $\hat{x}$ and $\hat{p}$, and obtain

$$\begin{aligned}0 =& \int dx'\ \langle x|\,\hat{a}\,|x'\rangle\,\langle x'|0\rangle = \frac{1}{2x_{\mathrm{ZPF}}}\int dx'\ \left(\langle x|\,\hat{x}\,|x'\rangle + \frac{i}{m\omega_a}\langle x|\,\hat{p}\,|x'\rangle\right)\psi_0(x')\\ =& \frac{1}{2x_{\mathrm{ZPF}}}\left[x\ \psi_0(x) + \frac{i}{m\omega_a}\int dx'\ \langle x|\,\hat{p}\,|x'\rangle\,\psi_0(x')\right]\end{aligned} \tag{2.172}$$

where we used $\langle x|x'\rangle = \delta(x - x')$ for the first term, and we need in addition

$$\langle x|\,\hat{p}\,|x'\rangle = \int dp'\ \langle x|\,\hat{p}\,|p'\rangle\,\langle p'|x'\rangle = \frac{1}{2\pi\hbar}\int dp'\ p'\ e^{\frac{i}{\hbar}p'(x-x')} = \frac{1}{2\pi i}\frac{d}{dx}\int dp'\ e^{\frac{i}{\hbar}p'(x-x')} \tag{2.173}$$

where we have used again the completeness relation for the momentum eigenstates, the remaining integral is simply a delta function and which gives us

$$0 = \frac{1}{2x_{\mathrm{ZPF}}}\left[x\ \psi_0(x) + \frac{i}{m\omega_a}\int dx'\ \frac{\hbar}{i}\frac{d}{dx}\ \delta(x - x')\psi_0(x')\right] = \frac{1}{2x_{\mathrm{ZPF}}}\left[x + \frac{\hbar}{m\omega_a}\frac{d}{dx}\right]\psi_0(x), \tag{2.174}$$

with this we have obtained a differential equation for the ground state wave function, the solution to this equation is a Gaussian

$$\frac{\partial}{\partial x}\psi_0(x) = -x\ \psi_0(x) \ \ \Rightarrow \ \ \psi_0(x) = \mathcal{A}_0\, e^{-\frac{m\omega_a}{2\hbar}x^2}, \quad \mathcal{A}_0 = \left(\frac{m\omega_a}{\hbar\pi}\right)^{\frac{1}{4}}, \tag{2.175}$$

where the pre-factor $\mathcal{A}_0$ originates from the normalization of the wave function. In as similar fashion we can obtain the wave function for arbitrary excitation n

$$\begin{aligned}\psi_n(x) = \langle x|n\rangle &= \frac{1}{\sqrt{n!}}\langle x|\left(\hat{a}^\dagger\right)^n|0\rangle = \frac{1}{(2x_{\rm ZPF})^n}\ \frac{1}{\sqrt{n!}}\left(x - \frac{\hbar}{m\omega_a}\frac{\partial}{\partial x}\right)^n \psi_0(x) \\ &= \frac{1}{(2x_{\rm ZPF})^n}\ \frac{1}{\sqrt{n!}}\mathcal{A}_0\left(x - \frac{\hbar}{m\omega_a}\frac{\partial}{\partial x}\right)^n e^{-\frac{m\omega_a}{2\hbar}x^2} \equiv \mathcal{A}_n H_n(\chi)e^{-\frac{\chi^2}{2}}, \quad \chi = \sqrt{\frac{m\omega_a}{\hbar}}x,\end{aligned} \tag{2.176}$$

introducing the so-called Hermite polynomials, which are functions of the dimensionless quantity χ, and defined as

$$H_n(\chi) = e^{\frac{\chi^2}{2}}\left(\chi - \frac{\partial}{\partial\chi}\right)^n e^{-\frac{\chi^2}{2}}, \tag{2.177}$$

where the lowest order functions are

$$H_0(\chi) = 1, \quad H_1(\chi) = 2\chi, \quad H_2(\chi) = 4\chi^2 - 2, \ \ ... \tag{2.178}$$

The Hermite polynomials have the property

$$\int d\chi\ H_n(\chi)H_m(\chi)e^{-\chi^2} = \sqrt{\pi}2^n n!\ \delta_{n,m}, \tag{2.179}$$

which we can use to normalize the wave functions

$$\int dx\ |\psi_n(x)|^2 = \sqrt{\frac{\hbar}{m\omega_a}}\ \mathcal{A}_n^2\ \int d\chi H_n^2(\chi)e^{-\chi^2} = \sqrt{\frac{\hbar\pi}{m\omega_a}}\ \mathcal{A}_n^2\ 2^n n! \ \ \Rightarrow \ \ \mathcal{A}_n = \frac{2^{-\frac{n}{2}}}{\sqrt{n!}}\left(\frac{m\omega_a}{\hbar\pi}\right)^{\frac{1}{4}}. \tag{2.180}$$

The wave function in position space could have as well been obtained via solving the stationary Schrödinger equation. The latter derivation is a bit pedestrian, while the usage of the second quantization approach involving the creation and annhilation operators gave as straightaway access to the energy spectrum of the harmonic oscillator. Things become more involved if we start to include some anharmonicity, e.g. taking the so-called quantized Duffing oscillator we have the Hamiltonian with a quartic nonlinearity

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{m\omega_a^2}{2}\hat{x}^2 - g\frac{m^2\omega_a^2}{\hbar}\hat{x}^4 \backsimeq \hbar\omega_a'\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) - g\frac{3\hbar}{2}(\hat{a}^\dagger\hat{a})^2, \ \ \omega_a' = \omega_a - \frac{3}{2}g, \tag{2.181}$$

here, in the last step, we made an approximation to estimate the change in the energy spectrum, keeping only the square of the number operator (while accounting for a change in the frequency), which gives us access to the approximated eigenenergies of the system:

$$E_n = \hbar\omega_a'\left(n + \frac{1}{2}\right) - g\frac{3\hbar}{2}n^2 \ \ \Rightarrow \ \ \Delta E_n = \hbar\omega_a - \hbar 3g(n+1), \tag{2.182}$$

now the energy spectrum is not longer equidistant, i.e., the distance between levels decreases for larger values of n. However, the exact energy spectrum is inaccessible, but we can utilize perturbation theory for a proper approximation, which is the topic of our next chapter.

Chapter 3

Perturbation theory & time evolution

3.1 Time-independent perturbation theory [5, 9]

An exact solution of a quantum mechanical problem exists only in rather few cases. In many situations is it therefore important to use approximate methods to gain qualitative insight into the properties of a physical system. Progress can be made if the Hamiltonian can be split into two contributions

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{3.1}$$

with $\hat{H}_0$ as the so-called *unperturbed* part of the Hamiltonian, which is the part of the system we can straightforwardly solve as an eigenvalue problem. For example, for the quantum anharmonic oscillator discussed at the end of the last chapter we could have the decomposition

$$\hat{H}_0 = \hbar\omega_a \left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\right), \quad \hat{V} = -g\frac{\hbar}{4}(\hat{a} + \hat{a}^\dagger)^4, \tag{3.2}$$

treating the full nonlinear term as a perturbation. For a perturbative ansatz one has to assume that the perturbation $\hat{V}$ is small (what that means will be defined further below). Perturbation theory comes in two flavours: degenerate and non-degenerate, which here refers to the occurrence of state vectors corresponding to one eigenvalues. Non-degenerate means that all eigenvalues are distinct, i.e., each eigenstate is associated with one eigenvalue. Moreover, one makes a distinction between time-dependent and time-independent perturbation theory, in the course of this lecture we will only focus on the latter, meaning that we have a time-independent perturbation $\hat{V}$. The starting point for our perturbation theory, is that the unperturbed part of the problem at hand can be solved. Meaning that we have a solution for the stationary Schrödinger equation of the form

$$\hat{H}_0 \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(0)}\right\rangle, \quad \left\langle\psi_n^{(0)}|\psi_m^{(0)}\right\rangle = \delta_{m,n}, \tag{3.3}$$

corresponding to the solution in zeroth order. Clearly, the aim is it go beyond this and find approximate solutions for the Schrödinger equation for the full system

$$\hat{H} \left|\psi_n\right\rangle = E_n \left|\psi_n\right\rangle, \quad \left\langle\psi_n|\psi_m\right\rangle = \delta_{m,n}, \tag{3.4}$$

for this we make the ansatz

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}, \tag{3.5}$$

with $0 \leq \lambda \leq 1$ as a dimensionless variable with which we can track the order of our perturbative expansion. The latter we apply to the eigenvalues and eigenvectors

$$\begin{aligned} E_n =& E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots, \\ |\psi_n\rangle =& \left|\psi_n^{(0)}\right\rangle + \lambda \left|\psi_n^{(1)}\right\rangle + \lambda^2 \left|\psi_n^{(2)}\right\rangle + \dots, \end{aligned} \tag{3.6}$$

to note is, that no matter where we truncate the series expansion, the coefficients $E_n^{(m)}$ have to be real as they approximate the eigenvalue of an hermitian operator, i.e., the one of the total system Hamiltonian. We insert the expansions for the eigenstates and eigenvalues into the left side of Schrödinger equation Eq.(3.4) and obtain

$$\begin{aligned} \hat{H}\left|\psi_n\right\rangle =& \left(\hat{H}_0 + \lambda \hat{V}\right)\left(\left|\psi_n^{(0)}\right\rangle + \lambda \left|\psi_n^{(1)}\right\rangle + \lambda^2 \left|\psi_n^{(2)}\right\rangle + \dots\right) \\ =& \hat{H}_0 \left|\psi_n^{(0)}\right\rangle + \lambda \left[\hat{H}_0 \left|\psi_n^{(1)}\right\rangle + \hat{V}\left|\psi_n^{(0)}\right\rangle\right] + \lambda^2 \left[\hat{H}_0 \left|\psi_n^{(2)}\right\rangle + \hat{V}\left|\psi_n^{(1)}\right\rangle\right] + \dots, \end{aligned} \tag{3.7}$$

and the same procedure for the right side of Schrödinger equation yields

$$\begin{aligned} E_n\left|\psi_n\right\rangle =& \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots\right)\left(\left|\psi_n^{(0)}\right\rangle + \lambda \left|\psi_n^{(1)}\right\rangle + \lambda^2 \left|\psi_n^{(2)}\right\rangle + \dots\right) \\ =& E_n^{(0)} \left|\psi_n^{(0)}\right\rangle + \lambda \left[E_n^{(0)} \left|\psi_n^{(1)}\right\rangle + E_n^{(1)} \left|\psi_n^{(0)}\right\rangle\right] \\ &+ \lambda^2 \left[E_n^{(0)} \left|\psi_n^{(2)}\right\rangle + E_n^{(1)} \left|\psi_n^{(1)}\right\rangle + E_n^{(2)} \left|\psi_n^{(0)}\right\rangle\right] + \dots, \end{aligned} \tag{3.8}$$

this can now be combined together and be sorted in powers of λ

$$\begin{aligned} \text{0th order:} \quad & \hat{H}_0 \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(0)}\right\rangle, \\ \text{1st order:} \quad & \hat{H}_0 \left|\psi_n^{(1)}\right\rangle + \hat{V}\left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(1)}\right\rangle + E_n^{(1)} \left|\psi_n^{(0)}\right\rangle, \\ \text{2nd order:} \quad & \hat{H}_0 \left|\psi_n^{(2)}\right\rangle + \hat{V}\left|\psi_n^{(1)}\right\rangle = E_n^{(0)} \left|\psi_n^{(2)}\right\rangle + E_n^{(1)} \left|\psi_n^{(1)}\right\rangle + E_n^{(2)} \left|\psi_n^{(0)}\right\rangle, \\ & \vdots \\ m\text{th order:} \quad & \hat{H}_0 \left|\psi_n^{(m)}\right\rangle + \hat{V}\left|\psi_n^{(m-1)}\right\rangle = E_n^{(0)} \left|\psi_n^{(m)}\right\rangle + E_n^{(1)} \left|\psi_n^{(m-1)}\right\rangle + E_n^{(2)} \left|\psi_n^{(m-2)}\right\rangle + \dots \\ & \qquad = \sum_{k=0}^{m} E_n^{(k)} \left|\psi_n^{(m-k)}\right\rangle. \end{aligned} \tag{3.9}$$

An important aspect is the normalization of the state vector, above we have stated that the state vectors for the 0th order as well as the state vector resulting from the full solution have to have unity norm. Taking now our expansion we find

$$\begin{aligned} 1 = \langle\psi_n|\psi_n\rangle =& \left\langle\psi_n^{(0)}|\psi_n^{(0)}\right\rangle + \lambda\left[\left\langle\psi_n^{(0)}|\psi_n^{(1)}\right\rangle + \left\langle\psi_n^{(1)}|\psi_n^{(0)}\right\rangle\right] \\ &+ \lambda^2\left[\left\langle\psi_n^{(0)}|\psi_n^{(2)}\right\rangle + \left\langle\psi_n^{(2)}|\psi_n^{(0)}\right\rangle + \left\langle\psi_n^{(1)}|\psi_n^{(1)}\right\rangle\right] + \dots, \end{aligned} \tag{3.10}$$

thus, we have to enforce that for each order $m \geq 1$ we fulfill the following condition

$$m\text{th order:} \quad \sum_{k=0}^{m} \left\langle\psi_n^{(m-k)}|\psi_n^{(k)}\right\rangle = 0, \tag{3.11}$$

resulting in a system of m equations which have to be fulfilled, e.g. from the first order equation we find $\left\langle\psi_n^{(0)}|\psi_n^{(1)}\right\rangle = -\left\langle\psi_n^{(1)}|\psi_n^{(0)}\right\rangle$ which can be fulfilled for a purely complex scalar product of the two vectors or alternatively, if the scalar product of the vectors is zero.

3.1.1 Non-degenerate perturbation theory

First we will focus on the case of distinct eigenvalues and eigenvectors for the unperturbed system (0th order solution). Note, the latter non-degenerate condition is not necessary for the perturbed system. First of all we want to determine the corrections to the unperturbed energies, for this we multiply the mth order Schrödinger equation from the left with the 0th order vector $\left\langle \psi_n^{(0)} \right|$ we obtain

$$\left\langle \psi_n^{(0)} \right| \hat{H}_0 - E_n^{(0)} \left| \psi_n^{(m)} \right\rangle + \left\langle \psi_n^{(0)} \right| \hat{V} \left| \psi_n^{(m-1)} \right\rangle = \sum_{k=1}^{m-1} E_n^{(k)} \left\langle \psi_n^{(0)} | \psi_n^{(m-k)} \right\rangle + E_n^{(m)}, \tag{3.12}$$

where we have pulled out of the summation the cases $k = 0$ and $k = m$, and used the fact that the unperturbed states are orthogonal. The first term on the left side vanishes, which becomes clear if we consider the multiplication with another vector $\left\langle \psi_{n'}^{(0)} \right|$ and use the completeness relation for the 0th order state vectors:

$$\begin{aligned} &\sum_{n''} \left\langle \psi_{n'}^{(0)} \right| \hat{H}_0 \left| \psi_{n''}^{(0)} \right\rangle \left\langle \psi_{n''}^{(0)} | \psi_n^{(m)} \right\rangle - E_n^{(0)} \left\langle \psi_{n'}^{(0)} | \psi_n^{(m)} \right\rangle \\ &= \sum_{n''} E_{n''}^{(0)} \left\langle \psi_{n'}^{(0)} | \psi_{n''}^{(0)} \right\rangle \left\langle \psi_{n''}^{(0)} | \psi_n^{(m)} \right\rangle - E_n^{(0)} \left\langle \psi_{n'}^{(0)} | \psi_n^{(m)} \right\rangle = (E_{n'}^{(0)} - E_n^{(0)}) \left\langle \psi_{n'}^{(0)} | \psi_n^{(m)} \right\rangle, \end{aligned} \tag{3.13}$$

which vanishes for the case $n = n'$. With this we obtain for the energy corrections

$$E_n^{(m)} = \left\langle \psi_n^{(0)} \right| \hat{V} \left| \psi_n^{(m-1)} \right\rangle - \sum_{k=1}^{m-1} E_n^{(k)} \left\langle \psi_n^{(0)} | \psi_n^{(m-k)} \right\rangle \tag{3.14}$$

all quantities on the right side are lower in order than the right and hence, expected to be known. Meaning that knowing all corrections up to the $m-1$th order allows us to calculate the mth order. Note, some textbooks include already on this stage a normalization condition which removes the second term in the expression for $E_n^{(m)}$. However, we will keep it general for the moment and enforce normalization at a later stage. Second, we want to derive the corrections for the state vector, for this we multiply the mth order Schrödinger equation from the left with $\left\langle \psi_{n'}^{(0)} \right|$, use Eq.(3.13) and obtain

$$(E_{n'}^{(0)} - E_n^{(0)}) \left\langle \psi_{n'}^{(0)} | \psi_n^{(m)} \right\rangle + \left\langle \psi_{n'}^{(0)} \right| \hat{V} \left| \psi_n^{(m-1)} \right\rangle = \sum_{k=1}^{m-1} E_n^{(k)} \left\langle \psi_{n'}^{(0)} | \psi_n^{(m-k)} \right\rangle + E_n^{(m)} \delta_{n,n'}. \tag{3.15}$$

Here the assumption that the eigenstates of the unperturbed problem are non-degenerate becomes important. In that case we have for the energy difference $E_{n'}^{(0)} - E_n^{(0)} \neq 0$ as long as $n' \neq n$, hence, we can divide our expression by it and obtain

$$\left\langle \psi_{n'}^{(0)} | \psi_n^{(m)} \right\rangle \Big|_{n' \neq n} = \sum_{k=1}^{m-1} \frac{E_n^{(k)} \left\langle \psi_{n'}^{(0)} | \psi_n^{(m-k)} \right\rangle}{(E_{n'}^{(0)} - E_n^{(0)})} - \frac{\left\langle \psi_{n'}^{(0)} \right| \hat{V} \left| \psi_n^{(m-1)} \right\rangle}{(E_{n'}^{(0)} - E_n^{(0)})}. \tag{3.16}$$

Utilizing the completeness relation for the 0th order states we obtain for the state corrections

$$\begin{aligned}\left|\psi_n^{(m)}\right\rangle &= \sum_{n'} \left|\psi_{n'}^{(0)}\right\rangle \left\langle\psi_{n'}^{(0)}|\psi_n^{(m)}\right\rangle \\ &= \sum_{n'(n'\neq n)} \left[\sum_{k=1}^{m-1} \frac{E_n^{(k)} \left\langle\psi_{n'}^{(0)}|\psi_n^{(m-k)}\right\rangle}{(E_{n'}^{(0)} - E_n^{(0)})} - \frac{\left\langle\psi_{n'}^{(0)}\right| \hat{V} \left|\psi_n^{(m-1)}\right\rangle}{(E_{n'}^{(0)} - E_n^{(0)})}\right] \left|\psi_{n'}^{(0)}\right\rangle + \left\langle\psi_n^{(0)}|\psi_n^{(m)}\right\rangle \left|\psi_n^{(0)}\right\rangle,\end{aligned} \tag{3.17}$$

note, the last term appears here as we have not yet included a proper normalization (which is distinct from some quantum mechanics textbooks). The relevant normalization condition in Eq.(3.11) can be reformulated using the completeness relation for the unperturbed system

$$\sum_{n'=0}^{m} \sum_{k=0}^{m} \left\langle\psi_n^{(m-k)}|\psi_{n'}^{(0)}\right\rangle \left\langle\psi_{n'}^{(0)}|\psi_n^{(k)}\right\rangle = \left\langle\psi_n^{(m)}|\psi_n^{(0)}\right\rangle + \left\langle\psi_n^{(0)}|\psi_n^{(m)}\right\rangle + \sum_{k=1}^{m-1} \left\langle\psi_n^{(m-k)}|\psi_n^{(k)}\right\rangle = 0, \tag{3.18}$$

which can be compactly expressed as

$$\mathrm{Re}\left[\left\langle\psi_n^{(0)}|\psi_n^{(m)}\right\rangle\right] = -\frac{1}{2} \sum_{k=1}^{m-1} \left\langle\psi_n^{(m-k)}|\psi_n^{(k)}\right\rangle. \tag{3.19}$$

The perturbative method thus requires to solve for the corrections to the eigenenergies and eigenvectors in a successive manner going up to the desired order. The aim is to express all corrections in terms of the known eigenvalues and eigenvectors of the unperturbed system. It is not always clear up to which order one has to go for a specific problem, some problems can also not be solved using a perturbative method. In addition, things become rather cumbersome if one goes to higher orders. Thus, we will only discuss briefly the first two orders.
Clearly, the 0th and lowest order is simple the unperturbed problem, i.e. the case $\lambda = 0$. Next in line is the **first order**, for which we obtain the energy correction term

$$E_n^{(1)} = \left\langle\psi_n^{(0)}\right| \hat{V} \left|\psi_n^{(0)}\right\rangle \equiv V_{nn}, \tag{3.20}$$

where we introduced the abbreviation V_{nn} for the expectation value of the perturbation with respect to the unperturbed state. Next, we find that normalization enforces

$$\mathrm{Re}\left[\left\langle\psi_n^{(0)}|\psi_n^{(1)}\right\rangle\right] = 0 \quad \Rightarrow \quad \left\langle\psi_n^{(0)}|\psi_n^{(1)}\right\rangle \equiv 0, \tag{3.21}$$

demanding that the scalar product of the corrected vector with the unperturbed vector vanished is a common choice for the normalization. The correction for the state vector becomes

$$\left|\psi_n^{(1)}\right\rangle = -\sum_{n'(n'\neq n)} \frac{V_{n'n}}{E_{n'n}} \left|\psi_{n'}^{(0)}\right\rangle, \quad E_{n'n} = E_{n'}^{(0)} - E_n^{(0)}, \tag{3.22}$$

which includes the normalization condition and also clearly fulfills it, as the summation excludes the case $n = n'$. The norm of the first order correction vector becomes

$$\left\langle\psi_n^{(1)}|\psi_n^{(1)}\right\rangle = \sum_{n'(n'\neq n)} \frac{|V_{n'n}|^2}{E_{n'n}^2}. \tag{3.23}$$

Keeping the obtained normalization for from the first order in mind, we can now move onto the **second order** corrections. For the energy we find

$$E_n^{(2)} = \left\langle \psi_n^{(0)} \right| \hat{V} \left| \psi_n^{(1)} \right\rangle = - \sum_{n'(n' \neq n)} \frac{|V_{n'n}|^2}{E_{n'n}}, \tag{3.24}$$

where we used our result for the first order correction Eq.(3.22) in the second step. To note is that the second order energy correction is negative. One can show, at least for the ground state energy, that the first order energy correction $E_1^{(0)}$, together with the unperturbed value, is already the upper bound for the ground state energy. Consequently the second order energy correction has to be negative, see [9] for details. The normalization for the second order has to fulfill

$$\mathrm{Re}\left[\left\langle \psi_n^{(0)} | \psi_n^{(2)} \right\rangle\right] = \left| \left\langle \psi_n^{(0)} | \psi_n^{(2)} \right\rangle \right| \cos\varphi_{02} = -\frac{1}{2} \left\langle \psi_n^{(1)} | \psi_n^{(1)} \right\rangle = -\frac{1}{2} \sum_{n'(n' \neq n)} \frac{|V_{n'n}|^2}{E_{n'n}^2}. \tag{3.25}$$

with φ_{02} as the phase of the scalar product of the unperturbed vector and the second order vector. We have the freedom to choose an overall phase for a time-independent quantum system, hence, we simply set $\varphi_{02} = 0$, i.e., we make the scalar product real valued. With this we obtain for the second order correction to the state vector

$$\begin{aligned} \left| \psi_n^{(2)} \right\rangle &= \sum_{n'(n' \neq n)} \left[\frac{E_n^{(1)} \left\langle \psi_{n'}^{(0)} | \psi_n^{(1)} \right\rangle}{E_{n'n}} - \frac{\left\langle \psi_{n'}^{(0)} \right| \hat{V} \left| \psi_n^{(1)} \right\rangle}{E_{n'n}} \right] \left| \psi_{n'}^{(0)} \right\rangle + \left\langle \psi_n^{(0)} | \psi_n^{(2)} \right\rangle \left| \psi_n^{(0)} \right\rangle \\ &= \sum_{n',n''(n',n'' \neq n)} \left[-\frac{V_{nn} V_{n'n}}{E_{n'n}^2} \delta_{n',n''} + \frac{V_{n'n''} V_{n''n}}{E_{n'n} E_{n''n}} \right] \left| \psi_{n'}^{(0)} \right\rangle - \frac{1}{2} \sum_{n'(n' \neq n)} \frac{|V_{n'n}|^2}{E_{n'n}^2} \left| \psi_n^{(0)} \right\rangle. \end{aligned} \tag{3.26}$$

As mentioned above, it is not directly possible to estimate the needed order for the perturbative series and make an assessment of the quality of the approximation. Form the second order correction to the state vector we can expect that the second order is sufficient when $|V_{n'n}| \ll |E_{nn'}|$, i.e., the perturbation has to be small and the energy distance large.

3.1.2 The nonlinear oscillator

Now we want to discuss a concrete example, the nonlinear oscillator which is described by the unperturbed Hamiltonian $\hat{H}_0$ and perturbation $\hat{V}$:

$$\hat{H}_0 = \hbar\omega_a \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \quad \hat{V} = -g \frac{\hbar}{4} (\hat{a} + \hat{a}^\dagger)^4. \tag{3.27}$$

We can use the commutation relation for the bosonic operators to express the perturbation as

$$\hat{V} = -g \frac{\hbar}{4} \left(\hat{a}^4 + \hat{a}^{\dagger 4} + 4\hat{a}^3 \hat{a}^\dagger + 4\hat{a}\hat{a}^{\dagger 3} - 6\hat{a}^2 - 6\hat{a}^{\dagger 2} + 6 \left[(\hat{a}^\dagger \hat{a})^2 + \hat{a}^\dagger \hat{a} + \frac{1}{2} \right] \right), \tag{3.28}$$

hence, the first two terms describe the annhilation and creation of four excitations, while the next four terms denote add or remove two excitations in the system. The remaining terms

correspond to the approximate case we discussed at the end of the last chapter. For the latter part we could solve the eigensystem as the perturbation contains only the number operator, her the number of excitations is conserved. We start with the unperturbed case, recovering the expected result:

$$\hat{H}_0 \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(0)}\right\rangle \quad \Rightarrow \quad \left|\psi_n^{(0)}\right\rangle = |n\rangle\,, \quad E_n^{(0)} = \hbar\omega_a \left(n + \frac{1}{2}\right), \tag{3.29}$$

thus our eigenstates are simply the Fock-states $|n\rangle$ and we have $E_{n',n} = \hbar\omega_a(n'-n)$ for the energy difference. Next, we can calculate the first order correction to the energy eigenvalues using Eq.(3.20) and obtain

$$E_n^{(1)} = -\,g\frac{\hbar}{4}\,\langle n|\,(\hat{a}+\hat{a}^\dagger)^4\,|n\rangle = -g\frac{3\hbar}{2}\,\langle n|\,(\hat{a}^\dagger\hat{a})^2 + \hat{a}^\dagger\hat{a} + \frac{1}{2}\,|n\rangle = -g\frac{3\hbar}{2}\left(n^2+n+\frac{1}{2}\right), \tag{3.30}$$

here only the excitation number conserving part of the perturbation is relevant, as we work with orthogonal eigenstates and $\langle n|m\rangle = \delta_{n,m}$. We see that our earlier result for the approximated eigenenergies given in Eq.(2.182) equals the energies $E_n = E_n^{(0)} + E_n^{(1)}$ up to the first order. Using now the first order correction to the state vector in Eq.(3.22) we have

$$\left|\psi_n^{(1)}\right\rangle = -\sum_{n'(n\neq n')} \frac{V_{n'n}}{E_{n'n}} \left|\psi_{n'}^{(0)}\right\rangle = -\frac{g}{4\omega_a} \sum_{n'(n\neq n')} \frac{\langle n'|\,(\hat{a}+\hat{a}^\dagger)^4\,|n\rangle}{(n'-n)}\,|n'\rangle\,, \tag{3.31}$$

the case $n = n'$ is excluded, thus, only the excitation non-conserving terms of the perturbation are important. We require the off-diagonal matrix elements of these terms

$$\begin{aligned}
&\langle n'|\left(\hat{a}^4 + \hat{a}^{\dagger 4} + 4\hat{a}^3\hat{a}^\dagger + 4\hat{a}\hat{a}^{\dagger 3} - 6\hat{a}^2 - 6\hat{a}^{\dagger 2}\right)|n\rangle \\
&= \sqrt{n(n-1)(n-2)(n-3)}\;\delta_{n',n-4} + \sqrt{(n+1)(n+2)(n+3)(n+4)}\;\delta_{n',n+4} \\
&\quad + 4\sqrt{n(n-1)}\left[n-\frac{1}{2}\right]\delta_{n',n-2} + 4\sqrt{(n+1)(n+2)}\left[n+\frac{3}{2}\right]\delta_{n',n+2},
\end{aligned} \tag{3.32}$$

the terms arising due to the annhilation of four (two) excitations only contribute for $n' \geq 4(2)$. We can take this expressions now and obtain for the first order state correction

$$\begin{aligned}
\left|\psi_n^{(1)}\right\rangle =& \frac{g}{4\omega_a}\left\{\frac{1}{4}\sqrt{\frac{n!}{(n-4)!}}\;|n-4\rangle - \frac{1}{4}\sqrt{\frac{(n+4)!}{n!}}\;|n+4\rangle\right. \\
&\left. + 2\sqrt{\frac{n!}{(n-2)!}}\left[n-\frac{1}{2}\right]|n-2\rangle - 2\sqrt{\frac{(n+2)!}{n!}}\left[n+\frac{3}{2}\right]\;|n+2\rangle\right\},
\end{aligned} \tag{3.33}$$

from which we see that we obey the normalization condition $\left\langle n|\psi_n^{(1)}\right\rangle = 0$ as desired. We can already foresee that the expressions for the corrections will become more and more cumbersome. The last explicit part for general n we will discuss here is the second order energy

correction. Taking the result we obtained for the first order state correction we have

$$\begin{aligned}
E_n^{(2)} = \langle n|\,\hat{V}\,|\psi_n^{(1)}\rangle = -\hbar\frac{g^2}{16\omega_a}\Bigg\{&\frac{1}{4}\sqrt{\frac{n!}{(n-4)!}}\;\langle n|\,\hat{a}^{\dagger 4}\,|n-4\rangle - \frac{1}{4}\sqrt{\frac{(n+4)!}{n!}}\;\langle n|\,\hat{a}^4\,|n+4\rangle \\
&+ 8\sqrt{\frac{n!}{(n-2)!}}\left[n-\frac{1}{2}\right]\langle n|\left(\hat{a}\hat{a}^{\dagger 3} - \frac{3}{2}\hat{a}^{\dagger 2}\right)|n-2\rangle \\
&- 8\sqrt{\frac{(n+2)!}{n!}}\left[n+\frac{3}{2}\right]\;\langle n|\left(\hat{a}^3\hat{a}^\dagger - \frac{3}{2}\hat{a}^2\right)|n+2\rangle\Bigg\} \\
= -\hbar\frac{g^2}{16\omega_a}\Bigg\{&\frac{1}{4}\frac{n!}{(n-4)!} - \frac{1}{4}\frac{(n+4)!}{n!} + 8\frac{n!}{(n-2)!}\left[n-\frac{1}{2}\right]^2 - 8\frac{(n+2)!}{n!}\left[n+\frac{3}{2}\right]^2\Bigg\} \\
=\hbar\frac{g^2}{8\omega_a}&\Big\{34n^3+51n^2+59n+21\Big\},
\end{aligned} \tag{3.34}$$

again, the excitation conserving perturbation terms did not play a role and we selectively could choose which perturbative parts can act on $\left|\psi_n^{(1)}\right\rangle$ and give a contribution. For example, for the first term in the second order correction the element $\langle n|\,\hat{a}^4\,|n-4\rangle$ is non-vanishing, as only the creation operator applied four times can recover the eigenvector $|n\rangle$. To avoid further complicated calculations, we are going to focus on the second order state correction to the ground state. The first order correction becomes for $n=0$

$$E_0^{(1)} = -g\frac{3\hbar}{4},\quad \left|\psi_0^{(1)}\right\rangle = -\frac{3}{4}\frac{g}{\omega_a}\left\{\frac{1}{\sqrt{6}}\;|4\rangle + \sqrt{2}\;|2\rangle\right\}\quad\Rightarrow\quad \left\langle\psi_0^{(1)}|\psi_0^{(1)}\right\rangle = \frac{39}{32}\frac{g^2}{\omega_a^2}, \tag{3.35}$$

and thus a superposition of two even Fock-states. From Eq.(3.26) we find

$$\begin{aligned}
\left|\psi_0^{(2)}\right\rangle &= \sum_{n'(n'\neq 0)}\left[\frac{E_0^{(1)}\left\langle\psi_{n'}^{(0)}|\psi_0^{(1)}\right\rangle}{E_{n'0}} - \frac{\left\langle\psi_{n'}^{(0)}\right|\hat{V}\left|\psi_0^{(1)}\right\rangle}{E_{n'0}}\right]\left|\psi_{n'}^{(0)}\right\rangle - \frac{1}{2}\left\langle\psi_0^{(1)}|\psi_0^{(1)}\right\rangle\left|\psi_0^{(0)}\right\rangle \\
&= -\frac{1}{\hbar\omega_a}\sum_{n'(n'\neq 0)}\frac{1}{n'}\left[\frac{3}{4}g\hbar\left\langle n'|\psi_0^{(1)}\right\rangle + \langle n'|\,\hat{V}\left|\psi_0^{(1)}\right\rangle\right]|n'\rangle - \frac{39}{64}\frac{g^2}{\omega_a^2}|0\rangle\,,
\end{aligned} \tag{3.36}$$

thus we need

$$\left\langle n'|\psi_0^{(1)}\right\rangle = -\frac{3}{4}\frac{g}{\omega_a}\left\{\frac{1}{\sqrt{6}}\;\langle n'|4\rangle + \sqrt{2}\;\langle n'|2\rangle\right\} = -\frac{3}{4}\frac{g}{\omega_a}\left\{\frac{1}{\sqrt{6}}\;\delta_{n',4} + \sqrt{2}\;\delta_{n',2}\right\} \tag{3.37}$$

and the element

$$\begin{aligned}
\langle n'|\,\hat{V}\left|\psi_0^{(1)}\right\rangle =&\frac{3\hbar}{16\sqrt{6}}\frac{g^2}{\omega_a}\,\langle n'|\,\hat{a}^4 + \hat{a}^{\dagger 4} + 4\hat{a}^3\hat{a}^\dagger + 4\hat{a}\hat{a}^{\dagger 3} - 6\hat{a}^2 - 6\hat{a}^{\dagger 2} + 6\left[(\hat{a}^\dagger\hat{a})^2 + \hat{a}^\dagger\hat{a} + \frac{1}{2}\right]|4\rangle \\
&+ \frac{3\hbar\sqrt{2}}{16}\frac{g^2}{\omega_a}\,\langle n'|\,\hat{a}^4 + \hat{a}^{\dagger 4} + 4\hat{a}^3\hat{a}^\dagger + 4\hat{a}\hat{a}^{\dagger 3} - 6\hat{a}^2 - 6\hat{a}^{\dagger 2} + 6\left[(\hat{a}^\dagger\hat{a})^2 + \hat{a}^\dagger\hat{a} + \frac{1}{2}\right]|2\rangle \\
=&\frac{3\hbar}{16}\frac{g^2}{\omega_a}\left[59\sqrt{2}\delta_{n',2} + \sqrt{6}\left[48+\frac{1}{2}\right]\delta_{n',4} + 34\sqrt{5}\delta_{n',6} + 2\sqrt{70}\delta_{n',8}\right].
\end{aligned} \tag{3.38}$$

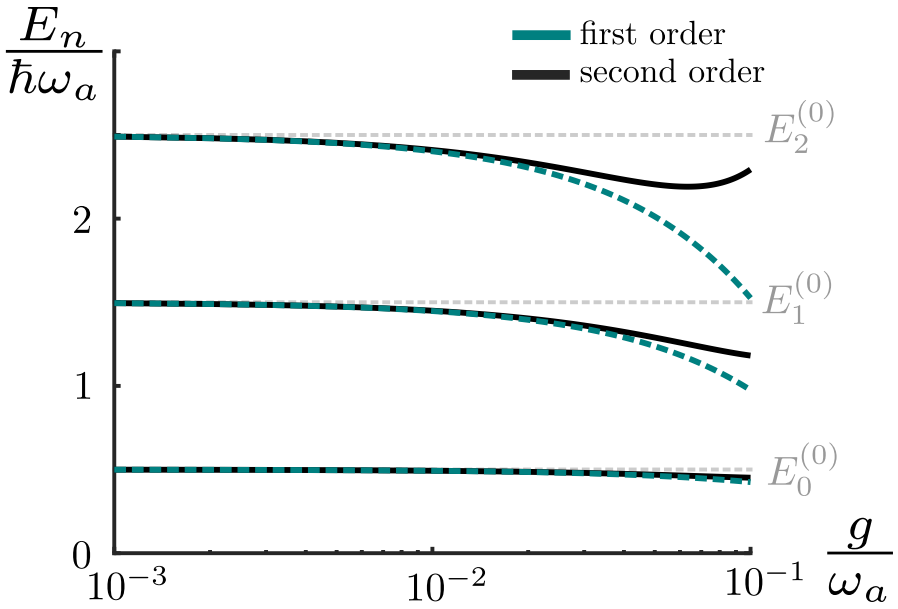

Figure 3.1: First three energy levels of the nonlinear oscillator as a function of g/ω_a including first and second order energy correction.

Inserting this into our second order state correction we obtain

$$\left|\psi_0^{(2)}\right\rangle = -\frac{3}{64}\frac{g^2}{\omega_a^2}\left[13\,|0\rangle + 112\sqrt{2}\;\langle 2| - 48\sqrt{6}\;\langle 4| + \frac{68}{3}\sqrt{5}\,|6\rangle + \sqrt{70}\,|8\rangle\right], \tag{3.39}$$

and thus we can approximate the energy and state vector for the system up to second order

$$\begin{aligned}\frac{E_0}{\hbar\omega_a} \approx& \frac{1}{2} - \frac{3}{4}\frac{g}{\omega_a} + \frac{21}{8}\frac{g^2}{\omega_a^2}\\ |\psi_0\rangle \approx& |0\rangle - \frac{3}{4}\frac{g}{\omega_a}\left(\sqrt{2}\,|2\rangle + \frac{1}{\sqrt{6}}\,|4\rangle\right)\\ &- \frac{3}{64}\frac{g^2}{\omega_a^2}\left[13\,|0\rangle + 112\sqrt{2}\;\langle 2| - 48\sqrt{6}\;\langle 4| + \frac{68}{3}\sqrt{5}\,|6\rangle + \sqrt{70}\,|8\rangle\right],\end{aligned} \tag{3.40}$$

where the state is a superposition of even Fock states, crucially, for this approximation to hold we need the ratio $g/\omega_a \ll 1$. Figure 3.1 depicts the first three eigenenergies as a function of the nonlinear coefficient strength g/ω_a. For very small values g/ω_a the first order correction (dashed cyan line) approximated the energies well. However for larger values we observe a clear discrepancy with the second order correction, the latter is more pronounced for higher values of n.

3.1.3 Degenerate perturbation theory

Now we like to briefly discuss the case of degenerate eigenenergies of the unperturbed case, i.e., m state vectors $\left\{\left|\psi_n^{(0)}\right\rangle\right\}_m$ have the same eigenenergy $E_n^{(0)}$., i.e.. m denotes the degree of

degeneracy. We use the the label α to mark the state vectors with same eigenenergies which all obey the Schrödinger equation

$$\hat{H}_0 \left|\psi_{n\alpha}^{(0)}\right\rangle = E_n^{(0)} \left|\psi_{n\alpha}^{(0)}\right\rangle, \quad \alpha = 1, 2, ..., m. \tag{3.41}$$

Any linear combination of the state vectors $\left|\psi_{n\alpha}^{(0)}\right\rangle$ is again a solution, hence it is not clear which ones are the correct states of 0th order. The task is to find them, for this we now start from the expansion in Eq.(3.9), and consider only the first order

$$\hat{H}_0 \left|\psi_n^{(1)}\right\rangle + \hat{V} \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(1)}\right\rangle + E_n^{(1)} \left|\psi_n^{(0)}\right\rangle, \tag{3.42}$$

where is unclear what states $\left|\psi_n^{(0)}\right\rangle$ are the right ones. However, to obtain these we can express the state as a linear combination of the eigenvectors belonging to the eigenvalue $E_n^{(0)}$:

$$\left|\psi_n^{(0)}\right\rangle = \sum_{\alpha=1}^{m} c_\alpha \left|\psi_{n\alpha}^{(0)}\right\rangle, \tag{3.43}$$

If we assume that this state is the right one, then the crucial point is to determine the coefficients c_α. We insert this ansatz into Eq.(3.42) and multiply from the left with the vector $\left\langle\psi_{n\beta}^{(0)}\right|$ and obtain

$$\left\langle\psi_{n\beta}^{(0)}\right| \left(\hat{H}_0 - E_n^{(0)}\right) \left|\psi_n^{(1)}\right\rangle + \sum_{\alpha=1}^{m} c_\alpha \left\langle\psi_{n\beta}^{(0)}\right| \left(\hat{V} - E_n^{(1)}\right) \left|\psi_{n\alpha}^{(0)}\right\rangle = 0, \tag{3.44}$$

The first term vanishes as operating with $\hat{H}_0$ onto the 0th order state vector to the right just gives the corresponding eigenvalue $E_n^{(0)}$. The remaining term can be used to determine the coefficients c_α. To illustrate the procedure we will focus on the case where $m = 2$ (a generalization is straightforward, see [9]). We define the perturbation matrix

$$\mathcal{S}_n \equiv \begin{pmatrix} H_n^{1,1} - E_n^{(1)} & H_n^{1,2} \\ H_n^{2,1} & H_n^{2,2} - E_n^{(1)} \end{pmatrix}, \quad H_n^{\beta,\alpha} = \left\langle\psi_{n\beta}^{(0)}\right| \hat{V} \left|\psi_{n\alpha}^{(0)}\right\rangle. \tag{3.45}$$

The desired coefficients $c_{1,2}$ are obtained from the linear equation system for

$$\mathcal{S}_n \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.46}$$

Such a linear equation has only non-trivial solutions for the case of a vanishing determinant of the perturbation matrix (singular matrix condition), i.e.,

$$\det\left[\mathcal{S}_n\right] = \left(H_n^{1,1} - E_n^{(1)}\right)\left(H_n^{2,2} - E_n^{(1)}\right) - |H_n^{1,2}|^2 = 0, \tag{3.47}$$

which is a polynomial of second order for the energy corrections of the first order, which we can solve straightforwardly and obtain the two solutions

$$E_{nx}^{(1)} = \frac{1}{2}\left(H_n^{1,1} + H_n^{2,2}\right) \pm \frac{1}{2}\sqrt{\left(H_n^{1,1} - H_n^{2,2}\right)^2 + 4|H_n^{1,2}|^2} \equiv \frac{1}{2}\left(H_n^{1,1} + H_n^{2,2} \pm \Delta E\right), \quad x = 1, 2. \tag{3.48}$$

Assuming that the $\Delta E > 0$ the degeneracy of the unperturbed system has been lifted, i.e., $\Delta E = 0$ requires that the two states do not mix and the perturbation has identical matrix elements. We will focus on the case where the perturbation is lifted, then we can solve for the coefficients $c_{\alpha x}$ (there are two possible energies now, so we have to determine two coefficients per solution). The linear equation system obtained from the perturbation matrix is now linear dependent, thus, we need to choose one solution for

$$\frac{1}{2}\left[\left(H_n^{1,1} - H_n^{2,2}\right) \mp \Delta E\right] c_{1x} = - H_n^{1,2} c_{2x}, \tag{3.49}$$

and combine it with the normalization condition $|c_{1x}|^2 + |c_{2x}|^2 = 1$. Taking Eq.(3.43) the correct states of the zeroth order are then the linear combinations

$$\left|\psi_n^{(0)}\right\rangle = c_{1x}\left|\psi_{1x}^{(0)}\right\rangle + c_{2x}\left|\psi_{2x}^{(0)}\right\rangle. \tag{3.50}$$

This procedure is straightforwardly expanded to systems with a higher degree of degeneracy, see [9] for a more detailed discussion. We will come back to this perturbation theory in a latter chapter when we discuss the so-called Stark-effect appearing in an atom exposed to an electric field.

3.2 Time-evolution in quantum mechanics

So far we have mainly focused on the stationary Schrodinger equation. Clearly, this is not the full story as quantum systems evolve as well in time. For describing the dynamics of quantum systems one utilizes so-called *pictures*, in which one decides if one evolves the states and/or the operators in time. Which picture to use depends on the problem at hand, i.e., if one is interested in the evolution of the state or the operator. In this sub-chapter we will introduce the three pictures used, namely the Schrödinger, Heisenberg or interaction picture. Crucially, no matter which picture one chooses, the physics stays the same.

3.2.1 The Schrödinger picture

In the Schrödinger picture the state are time-dependent and the operators are time-independent. The evolution of a quantum state in time follows from the solution of the Schrödinger equation

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \quad \Rightarrow \quad |\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle, \quad \hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}t}, \tag{3.51}$$

here $\hat{U}(t)$ denotes the unitary time-evolution operator, unitary means that it fulfills $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = 1$. Note, we have set the initial time here to be $t_0 = 0$ to simplify our expressions, and assumed that the Hamiltonian is time-independent. More generally the time-evolution operator has to fulfill the equation of motion:

$$\frac{d}{dt}\hat{U}(t)|\psi(0)\rangle = -\frac{i}{\hbar}\hat{H}\hat{U}(t)|\psi(0)\rangle \quad \Rightarrow \quad \frac{d}{dt}\hat{U}(t) = -\frac{i}{\hbar}\hat{H}\hat{U}(t), \tag{3.52}$$

which follows form the fact that $|\psi(0)\rangle$ is a constant and the equation of motion has to be fulfilled for any constant $|\psi(0)\rangle$. If the Hamiltonian is time-dependent the solution for the time-evolution operators becomes

$$\hat{U}(t) = \mathcal{T} \exp\left[-\frac{i}{\hbar}\int_0^t \hat{H}(t')dt'\right], \tag{3.53}$$

here the time-ordering operator is only necessary if the Hamiltonian's at different times do not commute, i.e., $\left[\hat{H}(t), \hat{H}(t')\right] \neq 0$. For the case that they do commute one can drop the time-ordering operator, in what follows we will stick to the later case.
Let us discuss a brief example, taking the case of the harmonic oscillator, the time-evolution operator of this system reads

$$\hat{H} = \hbar\omega_a\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \quad \Rightarrow \quad \hat{U}(t) = e^{-i\omega_a\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)}. \tag{3.54}$$

Assuming now that the initial state of the system is the Fock-state $|n\rangle$

$$|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle = e^{-i\omega_a\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)t}\,|n\rangle = e^{-i\frac{\omega_a}{2}t}\sum_{m=0}^{\infty}(-i\omega_a t)^m\,\frac{\left(\hat{a}^\dagger\hat{a}\right)^m}{m!}\,|n\rangle = e^{-i\omega_a(n+\frac{1}{2})t}\,|n\rangle\,, \tag{3.55}$$

and thus the time-evolution is simply an oscillation with the frequency ω_a. Important to note is that the scalar product of the vector with itself is not affected:

$$\langle\psi(t)|\psi(t)\rangle = e^{-i\omega_a(n+\frac{1}{2})t}e^{+i\omega_a(n+\frac{1}{2})t}\,\langle n|n\rangle = 1 \tag{3.56}$$

this matches what we already learned of unitary transformations, i.e., the norm of the states is conserved. Thus, the time-evolution dictated by quantum mechanics is unitary. For the chosen example of an harmonic oscillator we recover the case discussed in chapter 2, in Eq.(2.119) we derived the transformed eigenstate for a unitary $\hat{U} = e^{i\hat{O}}$ applied to the eigenstate of $\hat{O}$. Here we have $\hat{O} = \hat{n}$ with the Fock states as eigenstates.

3.2.2 The Heisenberg picture

Alternatively, we can work in a picture where the time dynamics is not on the states of a quantum system. In the so-called Heisenberg picture the states don't change, while the operators are time-dependent and their dynamics is described via

$$\frac{d}{dt}\hat{A}_H(t) = +\frac{i}{\hbar}\left[\hat{H}, \hat{A}_H(t)\right] + \left[\frac{d\hat{A}_S(t)}{dt}\right]_H \quad \Rightarrow \quad \hat{A}_H(t) = \hat{U}^\dagger(t)\hat{A}_S(t)\hat{U}(t), \tag{3.57}$$

where the explicit time-derivative of the operator $\hat{A}$ is only relevant if it is explicitly time-dependent. The subscript H denotes the transformation of the derivative in the Schrödinger

picture into the Heisenberg picture. It is straightforward to show that the solution for $\hat{A}_H(t)$ given above fulfills the equations of motion:

$$\begin{aligned}\frac{d}{dt}\left[\hat{U}^\dagger(t)\hat{A}_S(t)\hat{U}(t)\right] &= \left[\frac{d}{dt}\hat{U}^\dagger(t)\right]\hat{A}_S(t)\hat{U}(t) + \hat{U}^\dagger(t)\left[\frac{d\hat{A}_S(t))}{dt}\right]\hat{U}(t) + \hat{U}^\dagger(t)\hat{A}_S(t)\frac{d}{dt}\hat{U}(t)\\ &= +\frac{i}{\hbar}\hat{U}^\dagger(t)\hat{H}\hat{A}_S(t)\hat{U}(t) + \hat{U}^\dagger(t)\left[\frac{d\hat{A}_S(t)}{dt}\right]\hat{U}(t) - \frac{i}{\hbar}\hat{U}^\dagger(t)\hat{A}_S(t)\hat{H}\hat{U}(t)\\ &= +\frac{i}{\hbar}\hat{U}^\dagger(t)\hat{H}\hat{U}(t)\hat{A}_H(t) + \hat{U}^\dagger(t)\left[\frac{d\hat{A}_S(t)}{dt}\right]\hat{U}(t) - \frac{i}{\hbar}\hat{A}_H(t)\hat{U}^\dagger(t)\hat{H}\hat{U}(t)\\ &= +\frac{i}{\hbar}\left[\hat{H}_H, \hat{A}_H(t)\right] + \left[\frac{d\hat{A}_S(t)}{dt}\right]_H, \end{aligned} \tag{3.58}$$

where we used that the time-evolution operator obeys the same equation of motion as it does in the Schrödinger picture. If the time-evolution operator commutes with the Hamiltonian we have $\hat{H}_H = \hat{H}$. The Heisenberg picture is equivalent to the Schrödinger picture, which becomes clear if we consider the expectation value of the operator $\hat{A}_H$

$$\left\langle \hat{A}_H(t)\right\rangle = \langle\psi(0)|\,\hat{U}^\dagger(t)\hat{A}_S(t)\hat{U}(t)\,|\psi(0)\rangle = \mathrm{tr}\left\{\hat{A}_S(t)\hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t)\right\} = \mathrm{tr}\left\{\hat{A}_S(t)\hat{\rho}(t)\right\}. \tag{3.59}$$

To note is, that in general an operator is time-independent in the Schrödinger picture, which leaves the question how one deals with an explicit time-dependence of $\hat{A}_S(t)$. One work around is to split the time-evolution for the states and the operators, i.e., one can capture the time-evolution in a unitary transformation and work with a modified time-evolution operator, which we will discuss in the next sections.
Before this let us consider again an the example of an harmonic oscillator. Here we do not have an explicit time-dependence, and the equation of motion of the annhilation operator becomes

$$\frac{d}{dt}\hat{a} = i\omega_a\left[\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \hat{a}\right] = -i\omega_a\hat{a} \quad \Rightarrow \quad \hat{a}(t) = \hat{a}(0)e^{-i\omega_a t}, \tag{3.60}$$

again the dynamics in this case are oscillations with frequency ω_a. To note is that the number operator will not show any evolution in time, it commutes with the Hamiltonian and thus the time-evolution operator. It is a constant of motion, matching the fact that the number of excitations in this system does not change with time, it is conserved.

3.2.3 The interaction picture

Now we will discuss the last of the three pictures, this so-called interaction or Dirac picture. Here, we decompose the Hamiltonian in two parts: $\hat{H} = \hat{H}_0 + \hat{V}(t)$, where $V(t)$ could be time-dependent and $\hat{H}_0$ contains the 'boring' free evolution which we assume to be time-independent. The time-evolution operator can then be decomposed into two parts:

$$\hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}_0 t}\,\hat{U}_{\mathrm{IP}}(t), \qquad \frac{d}{dt}\hat{U}_{\mathrm{IP}}(t) = -\frac{i}{\hbar}\hat{V}_{\mathrm{IP}}(t)\hat{U}_{\mathrm{IP}}(t) \quad \Rightarrow \quad \hat{U}_{\mathrm{IP}}(t) = \mathcal{T}\, e^{-\frac{i}{\hbar}\int_0^t \hat{V}_{\mathrm{IP}}(t')dt'}, \tag{3.61}$$

with $\hat{V}_{\rm IP}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t}$. Note, we often call $\hat{H}_0$ the free evolution of the system, while $\hat{V}(t)$ is referred to as perturbation or interaction. The decomposed time-evolution operator obeys the same equation of motion as the undecomposed one:

$$\begin{aligned}\frac{d}{dt}\hat{U}(t) =& -\frac{i}{\hbar}\hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\,\hat{U}_{\rm IP}(t) + e^{-\frac{i}{\hbar}\hat{H}_0 t}\,\frac{d}{dt}\hat{U}_{\rm IP}(t)\\ =& -\frac{i}{\hbar}\hat{H}_0 e^{-i\hat{H}_0 t}\,\hat{U}_{\rm IP}(t) - \frac{i}{\hbar}e^{-\frac{i}{\hbar}\hat{H}_0 t}\,\hat{V}_{\rm IP}(t)\hat{U}_{\rm IP}(t)\\ =& -\frac{i}{\hbar}\hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\,\hat{U}_{\rm IP}(t) - \frac{i}{\hbar}e^{-i\hat{H}_0 t}\,e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}_{\rm IP}(t) = -\frac{i}{\hbar}\hat{H}\hat{U}(t).\end{aligned} \tag{3.62}$$

The interaction picture lies between the Heisenberg and the Schrödinger picture. States and operators become time-dependent, but evolve with different Hamiltonians. The not explicitly time-independent operators evolve with the unperturbed Hamiltonian

$$\frac{d}{dt}\hat{A}_{\rm IP}(t) = +\frac{i}{\hbar}\left[\hat{H}_0, \hat{A}_{\rm IP}(t)\right] \quad\Rightarrow\quad \hat{A}_{\rm IP}(t) = e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{A}_S(0)e^{-\frac{i}{\hbar}\hat{H}_0 t}, \tag{3.63}$$

connecting this to the Heisenberg picture we have to add the dynamics mediated by $\hat{U}_{\rm IP}$

$$\hat{A}_{\rm H}(t) = \hat{U}^\dagger_{\rm IP}(t)\hat{A}_{\rm IP}(t)\hat{U}_{\rm IP}(t) = \hat{U}^\dagger_{\rm IP}(t)e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{A}_S(0)e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}_{\rm IP}(t) = \hat{U}^\dagger(t)\hat{A}_S(0)\hat{U}(t). \tag{3.64}$$

In contrast, the states evolve with the $\hat{U}_{\rm IP}$ and thus with the transformed *interaction* $\hat{V}_{\rm IP}(t)$

$$\frac{d}{dt}\left|\psi_{\rm IP}(t)\right\rangle = -\frac{i}{\hbar}\hat{V}_{\rm IP}(t)\left|\psi_{\rm IP}(t)\right\rangle \quad\Rightarrow\quad \left|\psi_{\rm IP}(t)\right\rangle = \hat{U}_{\rm IP}(t)\left|\psi_{\rm IP}(0)\right\rangle = \hat{U}_{\rm IP}(t)\left|\psi(0)\right\rangle, \tag{3.65}$$

connecting this to the Schrödinger picture we have to add the free dynamics

$$\left|\psi_{\rm IP}(t)\right\rangle = e^{+\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle \quad\Rightarrow\quad \left|\psi(t)\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi_{\rm IP}(t)\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}_{\rm IP}(t)\left|\psi(0)\right\rangle = \hat{U}(t)\left|\psi(0)\right\rangle. \tag{3.66}$$

Thus, all three pictures are identical and can be transformed into each other. They all contain the complete dynamics associated with the Hamiltonian $\hat{H}$.
Coming back to our example of the harmonic oscillator, we just extend this case a bit including some time-dependent driving term acting on the oscillator. The Hamiltonian then reads

$$\hat{H} = \hbar\left\{\omega_a\left[\hat{a}^\dagger\hat{a} + \frac{1}{2}\right] + \alpha(t)\hat{a}^\dagger + \alpha^*(t)\hat{a}\right\} \equiv \hat{H}_0 + \hat{V}(t). \tag{3.67}$$

So first we can calculate the dynamics of the operators with respect to $\hat{H}_0$,

$$\begin{aligned}\hat{a}_{\rm IP}(t) =& e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{a}_S(0)e^{-\frac{i}{\hbar}\hat{H}_0 t} = e^{+i\omega_a\hat{a}^\dagger\hat{a}t}\hat{a}e^{-i\omega_a\hat{a}^\dagger\hat{a}t}\\ =& \hat{a} + i\omega_a\left[\hat{a}^\dagger\hat{a}, \hat{a}\right] + \frac{1}{2!}\left(i\omega_a\right)^2\left[\hat{a}^\dagger\hat{a}, \left[\hat{a}^\dagger\hat{a}, \hat{a}\right]\right] + ... = \hat{a}e^{-i\omega_a t},\end{aligned} \tag{3.68}$$

were we used the Baker-Campell-Haussdorff formula and obtained the same time dynamics for the annhilation operator as for the case of the Harmonic oscillator in the Heisenberg picture. The states will evolve with the modified perturbation:

$$\hat{V}_{\rm IP}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t} = \hbar\left\{\alpha(t)\hat{a}^\dagger e^{+i\omega_a t} + \alpha^*(t)\hat{a}e^{-i\omega_a t}\right\}. \tag{3.69}$$

To evaluate this further we have to specify what kind of form has the driving $\alpha(t)$, we simply assume here a periodic modulation with amplitude $\bar{\alpha} \in \mathbb{R}$:

$$\alpha(t) = \bar{\alpha}\cos(\omega_a t) = \frac{1}{2}\bar{\alpha}\left[e^{+i\omega_a t} + e^{-i\omega_a t}\right], \tag{3.70}$$

with this we can determine the time-evolution operator (no time-ordering operator required)

$$\begin{aligned}\hat{U}_{\rm IP}(t) =& \exp\left\{-\frac{i}{\hbar}\int_0^t dt'\ \hat{V}_{\rm IP}(t')\right\} = \exp\left\{-\frac{i\bar{\alpha}}{2}\int_0^t dt'\ \left(\hat{a}^\dagger\left[e^{+i2\omega_a t'}+1\right] + h.c.\right)\right\}\\ =& \exp\left\{-\frac{i\bar{\alpha}}{2}\left(\hat{a}^\dagger\left[\frac{1}{i2\omega_a}\left[e^{+i2\omega_a t}-1\right]+t\right]+h.c.\right)\right\} \approx e^{-\frac{i\bar{\alpha}}{2}\left(\hat{a}^\dagger+\hat{a}\right)t},\end{aligned} \tag{3.71}$$

in the last step we made the approximation that $\omega_a \gg 1$, which physically means that these are fast-oscillating terms with can be neglected under a so-called rotating wave approximation. What remains resembles the form of the displacement operator in Eq.(2.160), introduced when we talked about coherent states. This is no coincidence, as for an initial state $|0\rangle$, the drive described by $\hat{V}$ is producing a displaced vacuum state and hence a coherent state.

3.2.4 The interaction frame

Now we consider a different treatment of the time-evolution under an Hamiltonian $\hat{H}$, it's called the interaction or rotating frame. It's mainly used in quantum optics, and a lot of the literature calls it as well an interaction picture - but it is not! As we will discuss in what follows, the interaction frame is not really a picture. If one is operating with a Hamiltonian which is time-dependent, i.e., $\hat{H} = \hat{H}(t)$, we saw that the time-evolution operator is still unitary, but includes now time(-ordered) integrals. The Hamiltonian for a system which is exposed to external driving is for example time-dependent. However, some time-dependent Hamiltonians can become time-independent in the right frame. For example it can be favorable to decompose the Hamiltonian in two parts: $\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$ and to go into an **interaction frame** with respect to some *free* Hamiltonian $\hat{H}_0$. One then neglects the 'boring dynamics' associated the the free Hamiltonian and works with the modified time evolution operator

$$\hat{U}_{\rm IF}(t) = e^{+\frac{i}{\hbar}\hat{H}_0 t}\,\hat{U}(t) = \hat{U}_{\rm IP}(t). \tag{3.72}$$

Neglecting the 'boring' dynamics means that the time-evolution operator $\hat{U}_{\rm IF}(t)$ does not longer obey the same equation of motion as the full time-evolution operator $\hat{U}(t)$, in contrast, it obeys the equation of motion of the time-evolution operator in the interaction picture (most likely the origin for all the confusion)

$$\begin{aligned}\frac{d}{dt}\hat{U}_{\rm IF}(t) =& \frac{d}{dt}\left[e^{+\frac{i}{\hbar}\hat{H}_0 t}\,\hat{U}(t)\right] = \frac{i}{\hbar}\hat{H}_0\hat{U}_{\rm IF}(t) + e^{+\frac{i}{\hbar}\hat{H}_0 t}\left[\frac{d}{dt}\hat{U}(t)\right]\\ =& \frac{i}{\hbar}\hat{H}_0\hat{U}_{\rm IF}(t) - \frac{i}{\hbar}e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{H}\hat{U}(t)\\ =& \frac{i}{\hbar}\hat{H}_0\hat{U}_{\rm IF}(t) - \frac{i}{\hbar}e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_0\hat{U}(t) - \frac{i}{\hbar}e^{+\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)\hat{U}(t)\\ =& -\frac{i}{\hbar}\hat{V}_{\rm IF}(t)\hat{U}_{\rm IF}(t), \quad \hat{V}_{\rm IF}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}(t)e^{-\frac{i}{\hbar}\hat{H}_0 t}.\end{aligned} \tag{3.73}$$

In the interaction frame one can again decide if one wants to work in the Schrödinger picture or the Heisenberg picture. The state evolution in in this interaction frame becomes

$$\frac{d}{dt}\left|\psi_{\rm IF}(t)\right\rangle = -\frac{i}{\hbar}\hat{V}_{\rm IF}(t)\left|\psi_{\rm IF}(t)\right\rangle \quad \Rightarrow \quad \left|\psi_{\rm IF}(t)\right\rangle = \hat{U}_{\rm IF}(t)\left|\psi_{\rm IF}(0)\right\rangle = \hat{U}_{\rm IF}(t)\left|\psi(0)\right\rangle, \tag{3.74}$$

thus the state evolve in the same fashion as in the interaction picture, which makes sense, as in the interaction picture we neglected as well the evolution due to the free Hamiltonian for the state. One can straightforwardly transform the state in the interaction frame back into the original frame

$$\left|\psi(t)\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi_{\rm IF}(t)\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}_{\rm IF}(t)\left|\psi(0)\right\rangle = \hat{U}(t)\left|\psi(0)\right\rangle. \tag{3.75}$$

For the Heisenberg picture things are slightly different, instead of evolving via the free Hamiltonian in the interaction picture, they now evolve as well with the perturbation. We obtain the evolutions of the operators in the interaction frame as

$$\hat{A}_{\rm IF}(t) = \hat{U}^{\dagger}_{\rm IF}(t)\hat{A}_{\rm IF}(0)\hat{U}_{\rm IF}(t) \quad \Rightarrow \quad \frac{d}{dt}\hat{A}_{\rm IF}(t) = +\frac{i}{\hbar}\left[\hat{V}_{\rm IF}(t), \hat{A}_{\rm IF}(t)\right]. \tag{3.76}$$

thus we ignore the evolution via the free Hamiltonian here, so that the operator $\hat{A}_{\rm IF}(0)$ is time-independent in this frame.
We can also consider a more general notion for the interaction frame, only assuming here that the Hamiltonians at different times commute. For this we define a new state vector $|\Phi(t)\rangle$ as

$$|\Phi(t)\rangle \equiv \hat{P}(t)|\psi(t)\rangle, \quad \text{with } |\psi(t)\rangle = \hat{U}(t)|\psi(t_0)\rangle, \tag{3.77}$$

and leave the operator $\hat{P}(t)$ undefined for now, it corresponds to the rotation operator into the interaction frame. The time dependence of the new state vector $|\Phi(t)\rangle$ yields

$$\begin{aligned}\partial_t|\Phi(t)\rangle =& \partial_t\left(\hat{P}(t)|\psi(t)\rangle\right) = i\left[\partial_t\hat{P}(t)\right]|\psi(t)\rangle + \hat{P}(t)\partial_t|\psi(t)\rangle \\ =& \left\{\left[\partial_t\hat{P}(t)\right]\hat{P}^{\dagger}(t) - \frac{i}{\hbar}\hat{P}(t)\hat{H}(t)\hat{P}^{\dagger}(t)\right\}|\Phi(t)\rangle.\end{aligned} \tag{3.78}$$

We set $\hat{P}(t) = \mathbf{P}^{\dagger}$ which gives for the modified density matrix $\hat{\rho}(t) = |\Phi(t)\rangle\langle\Phi(t)|$ the dynamics

$$\partial_t\hat{\rho} = -\frac{i}{\hbar}\left[\mathbf{P}^{\dagger}\hat{H}(t)\mathbf{P}, \hat{\rho}\right] + \left[\partial_t\mathbf{P}^{\dagger}\right]\mathbf{P}\hat{\rho} + \hat{\rho}\mathbf{P}^{\dagger}\left[\partial_t\mathbf{P}\right]. \tag{3.79}$$

Taking the free evolution as our rotation operator we can connect this to our former result

$$\mathbf{P} = e^{-\frac{i}{\hbar}\hat{H}_0 t} \quad \Rightarrow \quad \partial_t\hat{\rho} = -\frac{i}{\hbar}\left[e^{+\frac{i}{\hbar}\hat{H}_0 t}\left(\hat{H}_0 + V(t)\right)e^{-\frac{i}{\hbar}\hat{H}_0 t}, \hat{\rho}\right] + \frac{i}{\hbar}\left[\hat{H}_0, \hat{\rho}\right] = -\frac{i}{\hbar}\left[\hat{V}_{\rm IF}(t), \hat{\rho}\right], \tag{3.80}$$

hence, the *boring* dynamics drops out as before. The nice aspect of this general form is that we can as well deal with rotation operators with more complex time-dependencies. The time-derivative for unitary operators of the form $\mathbf{P} = e^{X}$ and $\mathbf{P}^{\dagger} = e^{-X}$ is obtained via

$$\mathbf{P}^{\dagger}\left[\partial_t\mathbf{P}\right] = -\left[\partial_t\mathbf{P}^{\dagger}\right]\mathbf{P} = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}\left(\mathbf{ad}_X\right)^k\frac{dX}{dt}, \quad \mathbf{ad}_X(Y) = [X, Y]. \tag{3.81}$$

In this case we work with the modified Hamiltonian

$$\hat{H}'(t) = \mathbf{P}^{\dagger}\hat{H}(t)\mathbf{P} - i\mathbf{P}^{\dagger}\left[\partial_t\mathbf{P}\right]. \tag{3.82}$$

3.2.5 Example: Single photon measurement

In 1999 the Haroche group in Paris performed an experiment where they measured a single photon [14]. The challenge here was to detect the single photon without destroying it. The experimental setup comprises a standard cavity-quantum-electrodynamic (cavity-QED) setting: an atom is coupled to the electromagnetic field in a cavity. The basic Hamiltonian of such a system reads

$$\hat{H}/\hbar = \omega_c \hat{a}^\dagger \hat{a} + \omega_g \left|g\right\rangle \left\langle g\right| + \omega_e \left|e\right\rangle \left\langle e\right| + \frac{i\Omega}{2} \left(\left|e\right\rangle \left\langle g\right| + \left|g\right\rangle \left\langle e\right|\right) \left(\hat{a} - \hat{a}^\dagger\right) \tag{3.83}$$

where $\hat{a}$ denotes the cavity photon annihilation operator, $|g\rangle$ and $|e\rangle$ are the ground and excited state of the atom respectively. The first free terms describe the free energy of the system, e.g. ω_c is the resonant frequency of the cavity mode. The last term results the interaction between the dipole of the atom and the electromagnetic field, also called the Rabi-model and Ω is called the single-photon Rabi frequency. The operator $|e\rangle\langle g|$ is the creation operator of the atomic excitation, it takes the atom from the ground state to the excited state, while the operator $|g\rangle\langle e|$ initiates the reverse process.
We now move into an interaction frame with respect to the free Hamiltonian

$$\hat{H}_0/\hbar = \omega_c \hat{a}^\dagger \hat{a} + \omega_g \left|g\right\rangle \left\langle g\right| + (\omega_g + \omega_c) \left|e\right\rangle \left\langle e\right|, \tag{3.84}$$

i.e., we apply a unitary transformation of the form $\hat{U}(t) = e^{-i/\hbar \hat{H}_0 t}$. If we apply this transformation to the ground and excited states of the atom we obtain

$$\begin{aligned}
\hat{U}(t)\left|g\right\rangle =& e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|g\right\rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\hat{H}_0 t\right)^n \left|g\right\rangle = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-i\omega_g t\right)^n \left|g\right\rangle = e^{-i\omega_g t}\left|g\right\rangle, \\
\hat{U}(t)\left|e\right\rangle =& e^{-i(\omega_g+\omega_c)t}\left|e\right\rangle,
\end{aligned} \tag{3.85}$$

which would mean that we are in an interaction frame in the Schrödinger picture, i.e., the states are time-dependent while the operators are time-independent. However, it is also possible to go into an interaction frame in the Heisenberg picture, here we obtain for the atomic operators:

$$\begin{aligned}
\hat{U}^\dagger(t)\left|e\right\rangle\left\langle g\right|\hat{U}(t) &= e^{+\frac{i}{\hbar}\hat{H}_0 t}\left|e\right\rangle\left\langle g\right| e^{-\frac{i}{\hbar}\hat{H}_0 t} \\
&= \left|e\right\rangle\left\langle g\right| + \left(\frac{i}{\hbar}t\right)\left[\hat{H}_0, \left|e\right\rangle\left\langle g\right|\right] + \frac{1}{2!}\left(\frac{i}{\hbar}t\right)^2 \left[\hat{H}_0, \left[\hat{H}_0, \left|e\right\rangle\left\langle g\right|\right]\right] + \ldots
\end{aligned} \tag{3.86}$$

hence we need the commutator

$$\left[\hat{H}_0, \left|e\right\rangle\left\langle g\right|\right] = (\omega_g + \omega_c)\left|e\right\rangle\left\langle g\right| - \omega_g \left|e\right\rangle\left\langle g\right| = \omega_c \left|e\right\rangle\left\langle g\right|, \tag{3.87}$$

inserting this into the upper equation we obtain

$$\hat{U}^\dagger(t)\left|e\right\rangle\left\langle g\right|\hat{U}(t) = \left\{1 + (i\omega_c t) + \frac{1}{2!}(i\omega_c t)^2 + \ldots\right\}\left|e\right\rangle\left\langle g\right| = \sum_{n=0}^{\infty}\frac{1}{n!}(i\omega_c t)^n \left|e\right\rangle\left\langle g\right| = e^{i\omega_c t}\left|e\right\rangle\left\langle g\right|. \tag{3.88}$$

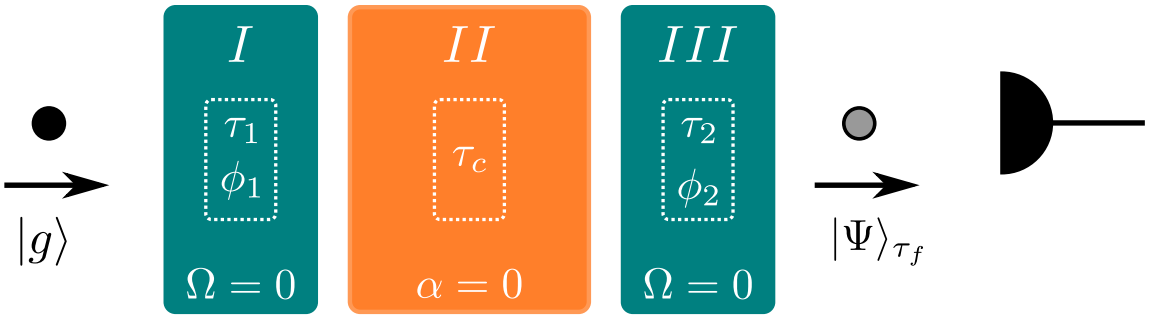

Figure 3.2: Sketch of the single photon measurement from the Haroche group [14].

The same procedure can be performed for the photonic operator $\hat{a} \to \hat{a}e^{-i\omega_c t}$. In this Heisenberg picture the operators are time-dependent and the Hamiltonian in this interaction frame becomes

$$\hat{H}' = \hat{U}^\dagger(t)\hat{H}\hat{U}(t) - \hat{H}_0 = \hbar\Delta\,|e\rangle\langle e| + \frac{i\Omega\hbar}{2}\left(|e\rangle\langle g|\,e^{i\omega_c t} + |g\rangle\langle e|\,e^{-i\omega_c t}\right)\left(\hat{a}e^{-i\omega_c t} - \hat{a}^\dagger e^{+i\omega_c t}\right), \tag{3.89}$$

with the detuning $\Delta = (\omega_e - \omega_g) - \omega_c$ between the two-level system and the cavity resonance frequency. In what follows we set this detuning to be zero, $\Delta = 0$. This interaction frame has the advantage that we can identify the resonant interactions, e.g., the process $|e\rangle\langle g|\,\hat{a}^\dagger$ which leads to the creation of an excitation of the cavity and the atom simultaneously is rather unlikely. In this frame such a process oscillates fast with $2\omega_c$ and is neglected under a rotating wave approximation. This yields the well-known Jaynes-Cummings Hamiltonian

$$\hat{H}' = \frac{i\Omega\hbar}{2}\left(|e\rangle\langle g|\,\hat{a} - |g\rangle\langle e|\,\hat{a}^\dagger\right), \tag{3.90}$$

this interaction describes a swapping of excitations between the atom and the photonic field. Assuming that one has only one excitation in the system - either in the cavity mode or stored in the excited state of the atom, the relevant states are $|g\rangle|1\rangle = |g,1\rangle$ and $|e\rangle|0\rangle = |e,0\rangle$, acting with the Hamiltonian on these states shows nicely this swapping behavior

$$\hat{H}'\,|g,1\rangle = +\frac{i\hbar\Omega}{2}\,|e,0\rangle\,, \qquad \hat{H}'\,|e,0\rangle = -\frac{i\hbar\Omega}{2}\,|g,1\rangle\,. \tag{3.91}$$

We assume now that the swapping interaction is on for a time interval τ. If the system is in the state $|g,1\rangle$ the final state after the time τ is

$$\begin{aligned} e^{-\frac{i}{\hbar}\hat{H}'\tau}\,|g,1\rangle &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{\hbar}\hat{H}'\tau\right)^n|g,1\rangle \\ &= \left\{1 - \frac{1}{2!}\left(\frac{\Omega\tau}{2}\right)^2 + \frac{1}{4!}\left(\frac{\Omega\tau}{2}\right)^4 - \ldots\right\}|g,1\rangle \\ &\quad + \left\{\left(\frac{\Omega\tau}{2}\right) - \frac{1}{3!}\left(\frac{\Omega\tau}{2}\right)^3 + \frac{1}{5!}\left(\frac{\Omega\tau}{2}\right)^5 - \ldots\right\}|e,0\rangle \\ &= \cos\left(\frac{\Omega\tau}{2}\right)|g,1\rangle + \sin\left(\frac{\Omega\tau}{2}\right)|e,0\rangle\,. \end{aligned} \tag{3.92}$$

After a time-interval $\tau = \frac{2\pi}{\Omega}$ the atom is again in the ground state:

$$\tau = \frac{2\pi}{\Omega}: \quad |g,1\rangle \to -|g,1\rangle \quad (|g,0\rangle \to +|g,0\rangle) \tag{3.93}$$

but the state has accumulated an overall phase. This is the important correlation between the atom and the electromagnetic field. Preparing the photon in a superposition state of the vacuum and one photon, the interaction with the atom for the time $\tau = \frac{2\pi}{\Omega}$ changes it to

$$|\Psi\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes |g\rangle \;\to\; (\alpha|0\rangle - \beta|1\rangle) \otimes |g\rangle\,, \tag{3.94}$$

hence it does not change the atom, but a sign flip occurs only if a photon present. This was exploited to perform the measurement of a photon without destroying it [14]. The idea is to use the atom as the object performing the measurement on the photon, the so-called meter, and after the atom has interacted via the swapping interaction with the photon one performs the actually measurement on the atom. However, measuring the atom in either state does not give information about the photon so far. Here the experimentalist used an auxiliary level $|f\rangle$, which interacts with an external microwave pulse in a similar manner as the cavity photon, with the distinction that the pulse is treated classically, i.e., the Hamiltonian in the rotated frame has the form

$$\hat{H}' = \frac{i\Omega\hbar}{2}\left(|e\rangle\langle g|\,\hat{a} - |g\rangle\langle e|\,\hat{a}^\dagger\right) + i\hbar\alpha\left(|f\rangle\langle g|\,e^{+i\phi} - |g\rangle\langle f|\,e^{-i\phi}\right). \tag{3.95}$$

Important is, that the microwave pulse prepares a superposition of the ground state and the auxiliary state: $|g\rangle \to (|g\rangle + |f\rangle)/\sqrt{2}$. After the atom has interacted with the cavity field , i.e., acting with $\hat{H}'$ for time τ, another microwave pulse is applied, which reverses the action of the first pulse: $|g\rangle \to (|g\rangle - |f\rangle)/\sqrt{2}$ and $|f\rangle \to (|g\rangle + |f\rangle)/\sqrt{2}$. The whole protocol goes as follows, see also Fig.3.2,

$$\begin{aligned}
(\alpha|0\rangle + \beta|1\rangle)|g\rangle \;&\to\; \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle)(|f\rangle + |g\rangle) \\
&\to\; \frac{1}{\sqrt{2}}(\alpha|0\rangle(|f\rangle + |g\rangle) + \beta|1\rangle(|f\rangle - |g\rangle)) \\
&\to\; \alpha|0\rangle|g\rangle + \beta|1\rangle|f\rangle\,,
\end{aligned} \tag{3.96}$$

measuring now the atom in the ground state gives us the the desired information about the photon: $|\alpha|^2$ results from the measurement in the ground state, and $|\beta|^2 = 1 - |\alpha|^2$ follows. Important is that if one photon is in the cavity, i.e., $\alpha = 0$ and thus atom and photon are in a product state, a measurement of the atom in the ground state will not destroy the photon, while one still has detected the photon.

Chapter 4

Orbitals, spins and rotations

4.1 Angular momentum operators

In the last chapter we have learned about the time-evolution of quantum systems. We introduced the time-evolution operator $\hat{U}(t) = e^{-\frac{i}{\hbar}\hat{H}t}$, which is a unitary operator, mediating a unitary transformation of the system and thus leaves the scalar product and also the norm of the state vector unchanged. One can also call this a translation operator, as it translates our state from time t to a time $t' = t + \delta t$. Such translation operators are not only defined regarding time, one can define unitary operations which are translations in space:

$$\hat{U}_{\mathbf{x}} = e^{-\frac{i}{\hbar}\,\delta\mathbf{x}\hat{\mathbf{p}}}, \quad \hat{U}_{\phi} = e^{-\frac{i}{\hbar}\delta\phi\hat{\mathbf{L}}}, \tag{4.1}$$

where the first is a spacial transformation taking a state from point $\mathbf{x}$ to a point $\mathbf{x} + \delta\mathbf{x}$. The second operator mediates a rotation around a small angle $\delta\phi$. The spacial translation is involving the momentum operator $\hat{\mathbf{p}}$, while the rotation is mediated by the angular momentum operator $\hat{\mathbf{L}}$, which is the focus of this chapter. Crucially, the fact that these are unitary transformation is connected to conservation laws for the momentum and the angular momentum operator, see [5] for further details. The latter establish a relation between the invariance of a physical system with respect to rotations and spacial translations.

4.1.1 General properties of angular momentum operators [5, 9]

The angular momentum operator is the analog of the momentum operator for quantum systems involving a rotational symmetry. These operators are defined as the generator of rotations on any wave function, and come in three flavours: the orbital momentum operator denoted $\hat{\mathbf{L}}$, the spin angular momentum operator denotes as $\hat{\mathbf{S}}$, and the total angular momentum operator labeled as $\hat{\mathbf{J}}$. From an mathematical perspective they all belong to the same (Lie)-group, namely $SU(2)$. Important is to first understand the algebra operators

$$\left[\hat{J}_n, \hat{J}_m\right] = i\hbar\epsilon_{nmk}\hat{J}_k, \quad \{k, n, m\} = x, y, z, \tag{4.2}$$

here we have again the Levi-Civita symbol as defined in Eq.(2.136). Figure 4.1 depicts a sketch which is helpful to remember the cyclic rules of the Levi-Civita symbol. Moreover, $\hat{\mathbf{J}}$ and thus, the components $\hat{J}_n$, are hermitian operators.

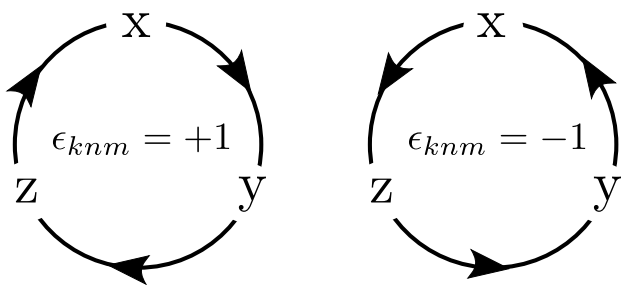

Figure 4.1: Illustration of the cyclic permutation rules mediated by the Levi-Civita symbol. Orders of indices which go clockwise, e.g., x, y, x or z, x, y, have $\epsilon_{nmk} = 1$, the latter changes its sign for a counterclockwise order. If two or more indices are the same $\epsilon_{nmk} = 0$.

The components of the angular momentum operator do not commute, thus, we cannot measure two (or all three) simultaneously. This also makes sense as the order on how rotations are performed matters even classically, i.e., rotations in different directions do not commute. This also means that these operators do not have common, or so-called simultaneous eigenstates, i.e., states were we can say what is the value of all components of the angular momentum operator. Thus, we can only determine one component $\hat{J}_n$ with certainty. This is straightforward to verify, assuming there would be a shared eigenstate $|\psi\rangle$, it would fulfill three equations

$$\hat{J}_k |\psi\rangle = \lambda_k |\psi\rangle , \quad k = x, y, z, \tag{4.3}$$

if we now take the case for the z-component

$$\hat{J}_z |\psi\rangle = \lambda_z |\psi\rangle = -i\frac{i}{\hbar}\left[\hat{J}_x, \hat{J}_y\right] |\psi\rangle = -\frac{i}{\hbar}\left[\lambda_x\lambda_y - \lambda_y\lambda_x\right] |\psi\rangle = 0, \tag{4.4}$$

and thus $\lambda_z = 0$, and one can analogously check that only $\lambda_k = 0$ is a solution for all three eigenvalue equations. This means that there exists no non-trivial solution and thus no simultaneous eigenstates. The question is if there is any other operator which commutes with the angular momentum components? We can already guess, that such an operator should be a rotational invariant operator. It turns out there is, the magnitude of a vector is invariant under rotation, so we can take the magnitude squared

$$\hat{\mathbf{J}}^2 = \hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z, \tag{4.5}$$

and check if this operator commutes with the components of the angular momentum operator

$$\begin{aligned}\left[\hat{\mathbf{J}}^2, \hat{J}_z\right] &= \left[\hat{J}_x\hat{J}_x + \hat{J}_y\hat{J}_y + \hat{J}_z\hat{J}_z, \hat{J}_z\right] = \left[\hat{J}_x\hat{J}_x, \hat{J}_z\right] + \left[\hat{J}_y\hat{J}_y, \hat{J}_z\right] \\ &= \left[\hat{J}_x, \hat{J}_z\right]\hat{J}_x + \hat{J}_x\left[\hat{J}_x, \hat{J}_z\right] + \left[\hat{J}_y, \hat{J}_z\right]\hat{J}_y + \hat{J}_y\left[\hat{J}_y, \hat{J}_z\right] \\ &= i\hbar\left(-\hat{J}_y\hat{J}_x - \hat{J}_x\hat{J}_y + \hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x\right) = 0,\end{aligned} \tag{4.6}$$

crucially, this holds for all components:

$$\left[\hat{\mathbf{J}}^2, \hat{J}_m\right] = 0. \tag{4.7}$$

With this we have two observables one can determine at the same time. It is possible to show within linear algebra that two commuting hermitian operators have a simultaneous eigenstate, for details please see [4]. Next it will be useful to define as well raising and a lowering operators

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y, \qquad \left(\hat{J}_\pm\right)^\dagger = \hat{J}_\mp, \tag{4.8}$$

which are now non-hermitian operators, and follows the same logic as we already discussed for creation and annhilation operators for Fock states, i.e., they will take us up and down a ladder of states. Which ladder structure we have at hand here will be clarified below. The commutation relations of these operators are

$$\left[\hat{J}_z, \hat{J}_\pm\right] = \pm\hbar\hat{J}_\pm, \qquad \left[\hat{\mathbf{J}}^2, \hat{J}_\pm\right] = 0, \qquad \left[\hat{J}_+, \hat{J}_-\right] = 2\hbar\hat{J}_z \tag{4.9}$$

where the proof of this is left for the homework. In addition we have the useful relation

$$\hat{\mathbf{J}}^2 = \hat{J}_\mp\hat{J}_\pm + \hat{J}_z^2 \pm \hbar J_z, \tag{4.10}$$

which can be shown by simply taking the product of $\hat{J}_\mp$ and using the commutation relation for the components Eq.(4.3). As we have now $\hat{\mathbf{J}}^2$ commuting with the components of $\hat{\mathbf{J}}$, we can now derive the simultaneous eigenstates of $\hat{\mathbf{J}}^2$ with one component $\hat{J}_n$. In general one takes here the $\hat{J}_z$ component, and we can start with the ansatz

$$\begin{aligned}\hat{J}_z\left|\alpha_m m\right\rangle &= \hbar m\left|\alpha_m m\right\rangle,\\ \hat{\mathbf{J}}^2\left|\alpha_m m\right\rangle &= \hbar^2\alpha_m^2\left|\alpha_m m\right\rangle,\end{aligned} \tag{4.11}$$

with $|\alpha_m m\rangle$ as the simultaneous eigenstate, and $\hbar m(\hbar^2\alpha_m^2)$ as the eigenvalue for $\hat{J}_z(\hat{\mathbf{J}}^2)$. The subindex m on the eigenvalue for $\hat{\mathbf{J}}^2$ accounts for a possible dependence of the two eigenvalues. From here we can first show that operating on the state with the raising and lowering operator results in an eigenstate with the same eigenvalue:

$$\hat{\mathbf{J}}^2\left(\hat{J}_\pm\left|\alpha_m m\right\rangle\right) = \hat{J}_\pm\hat{\mathbf{J}}^2\left|\alpha_m m\right\rangle = \hbar^2\alpha_m^2\left(\hat{J}_\pm\left|\alpha_m m\right\rangle\right), \tag{4.12}$$

where we used the commutation relation in Eq.(4.9). Thus the state $\hat{J}_\pm|\alpha_m m\rangle$ is as well an eigenstate of $\hat{\mathbf{J}}^2$. Moreover, using again the commutation relations in Eq.(4.9) we find

$$\hat{J}_z\left(\hat{J}_\pm\left|\alpha_m m\right\rangle\right) = \left(\hat{J}_\pm\hat{J}_z \pm \hbar\hat{J}_\pm\right)\left|\alpha_m m\right\rangle = \hbar\left(m\pm 1\right)\left(\hat{J}_\pm\left|\alpha_m m\right\rangle\right), \tag{4.13}$$

so acting with $\hat{J}_\pm$ on the state $|\alpha_m m\rangle$ raises or lowers the eigenvalue of $\hat{J}_z$ by $\hbar$. Hence, we can we can define the state after acting of the lowering and raising operators as

$$\hat{J}_\pm\left|\alpha_m m\right\rangle = \varepsilon^\pm_{\alpha_m,m}\left|\alpha_m m\pm 1\right\rangle, \tag{4.14}$$

with normalization factors $\varepsilon_m^\pm$ to be determined at a later stage. As before for the ladder operators we can now generate a state $|\alpha_m m\pm n\rangle$ by applying the lowering/raising operator n-times. We can check if moving up the the m-ladder changes the eigenvalue problem for $\hat{\mathbf{J}}^2$

$$\hat{\mathbf{J}}^2\left|\alpha_m m\pm n\right\rangle = \hat{\mathbf{J}}^2\frac{\hat{J}_\pm^n}{\widetilde{\varepsilon}}\left|\alpha_m m\right\rangle = \frac{\hat{J}_\pm^n}{\widetilde{\varepsilon}}\hat{\mathbf{J}}^2\left|\alpha_m m\right\rangle = \hbar^2\alpha_m^2\frac{\hat{J}_\pm^n}{\widetilde{\varepsilon}}\left|\alpha_m m\right\rangle = \hbar^2\alpha_m^2\left|\alpha_{m\pm n} m\pm n\right\rangle, \tag{4.15}$$

here we used that $\hat{\mathbf{J}}^2$ commutes with $\hat{J}_\pm$ and did not further define the normalization factor $\widetilde{\varepsilon}$. Crucially, the eigenvalue of $\hat{\mathbf{J}}^2$ does not change if the state is moved up and down the m-ladder. Since m is arbitrary it follows that α_m^2 is independent of m and thus we can drop now the subindex $\alpha_m^2 = \alpha^2$. The question is how many states are along the m-ladder, for this we can calculate the expectation value of $\hat{\mathbf{J}}^2$

$$\langle \alpha\ m|\,\hat{\mathbf{J}}^2\,|\alpha\ m\rangle = \langle \alpha\ m|\left(\hat{J}_x^2+\hat{J}_y^2+\hat{J}_z^2\right)|\alpha\ m\rangle = \left\langle \hat{J}_x^2\right\rangle + \left\langle \hat{J}_y^2\right\rangle + \hbar^2 m^2 = \hbar^2\alpha^2. \tag{4.16}$$

$\hat{J}_{x,y}$ are hermitian operators and have real eigenvalues, thus the expectation value of $\hat{J}_{x,y}^2$ can never be negative. Hence it follows that

$$\hbar^2\alpha^2 - \hbar^2 m^2 \geq 0 \quad \Rightarrow \quad \alpha^2 \geq m^2. \tag{4.17}$$

Thus, for a given value $\alpha > 0$, the possible values of m fall between $-\alpha$ and $+\alpha$. Note, we can always assume $\alpha > 0$ as only α^2 is relevant. This means that our m-ladder is bounded from below and above, and that we have extremal values m_{max} and m_{min} resulting in the conditions

$$\hat{J}_+\,|\alpha\ m_{max}\rangle = 0 \quad \text{and} \quad \hat{J}_-\,|\alpha\ m_{min}\rangle = 0, \tag{4.18}$$

using now Eqs.(4.10) we obtain

$$\begin{aligned}
\hat{\mathbf{J}}^2\,|\alpha\ m_{max}\rangle &= \left(\hat{J}_-\hat{J}_+ + \hat{J}_z^2 + \hbar J_z\right)|\alpha\ m_{max}\rangle = \hbar^2 m_{max}\,(m_{max}+1)\,|\alpha\ m_{max}\rangle = \hbar^2\alpha^2\,|\alpha\ m_{max}\rangle\,,\\
\hat{\mathbf{J}}^2\,|\alpha\ m_{min}\rangle &= \left(\hat{J}_+\hat{J}_- + \hat{J}_z^2 - \hbar J_z\right)|\alpha\ m_{min}\rangle = \hbar^2 m_{min}\,(m_{min}-1)\,|\alpha\ m_{min}\rangle = \hbar^2\alpha^2\,|\alpha\ m_{min}\rangle\,,
\end{aligned} \tag{4.19}$$

from which follows that

$$m_{max}\,(m_{max}+1) = m_{min}\,(m_{min}-1) \quad \Rightarrow \quad m_{max}^{(1)} = -m_{min},\ m_{max}^{(2)} = m_{min}-1, \tag{4.20}$$

where the second solution cannot be true as we have $m_{max} \geq m_{min}$, thus the relevant solution is $m_{max} = -m_{min} \equiv j$. With this definition we also have determined $\alpha^2 = j(j+1)$. And we have a sequence of values for m ranging from $-j$ to j, which is symmetric around 0. So which values can j take? Operating multiple times with the raising operator takes us from the state associated with $m_{min} = -j$ up the ladder, i.e., we have (ignoring the normalisation coefficients)

$$\left(\hat{J}_+\right)^n|\alpha\ -j\rangle \Rightarrow \left(\hat{J}_+\right)^{n-1}|\alpha\ -j+1\rangle \Rightarrow \left(\hat{J}_+\right)^{n-2}|\alpha\ -j+2\rangle \Rightarrow ... \Rightarrow |\alpha\ j\rangle\,, \tag{4.21}$$

hence, to go up the ladder starting from the state $|\alpha\ -j\rangle$ to the $|\alpha\ j\rangle$ we have to apply n-times the raising operator, which correspond to discrete steps up the latter and thus $n = m_{max} - m_{min} = 2j = 0, 1, 2, ...$ This can only be true for j taking the values

$$j = \begin{cases} 0, 1, 2, ... & \Rightarrow \quad m = \quad 0, \pm 1, \pm 2, ..., \pm j, \\ \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, ... & \Rightarrow \quad m = \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, ..., \pm j, \end{cases} \tag{4.22}$$

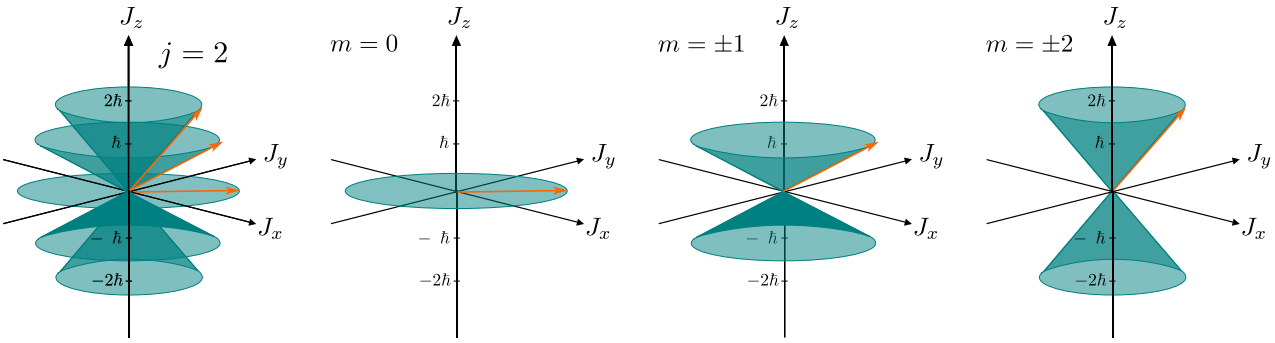

Figure 4.2: Angular momentum operators cannot be drawn as vectors as they are operators, however one make a virtual interpretation of it. Here we depict the case with quantum number $l = 2$, i.e., $m = \pm 2, \pm 1, 0$. The *state vector* is determined in the z direction, but not in the x, y direction and thus can be any vector in the circle (for $m = 0$) or along the cone ($m > 0$). Crucially, the length of the vector is constant, for $l = 1$ it is $\sqrt{\langle \mathbf{J}^2 \rangle} = \hbar\sqrt{6}$.

thus m is an integer if j is one, and an odd integer $\times \frac{1}{2}$ if j is it as well. Taking this all together we have for the eigenstates and eigenvalues for the angular momentum operators

$$\begin{aligned} \hat{J}_z \left| j\ m \right\rangle &= \hbar m \left| j\ m \right\rangle , \\ \hat{\mathbf{J}}^2 \left| j\ m \right\rangle &= \hbar^2 j(j+1) \left| j\ m \right\rangle , \quad m = -j, ..., j. \end{aligned} \tag{4.23}$$

What is left to determine is the coefficients $\varepsilon_m^{\pm}$ in Eq.(4.14), i.e., the normalization factor for the state generated when acting with the lowering or raising operator on the state $|j, m\rangle$:

$$\hat{J}_+ \left| j\ m \right\rangle = \varepsilon_{j,m}^{+} \left| j\ m+1 \right\rangle , \quad \hat{J}_- \left| j\ m \right\rangle = \varepsilon_{j,m}^{-} \left| j\ m-1 \right\rangle , \tag{4.24}$$

the scalar product of the state $\hat{J}_+ \left| j\ m \right\rangle$ with itself is

$$\begin{aligned} |\varepsilon_{j,m}^{\pm}|^2 = &\left\langle j\ m | \hat{J}_{\mp} \hat{J}_{\pm} | j\ m \right\rangle = \left\langle j\ m | \hat{\mathbf{J}}^2 - \hat{J}_z^2 \mp \hbar J_z | j\ m \right\rangle \\ = &\hbar^2 \left[j(j+1) - m(m \pm 1) \right] , \end{aligned} \tag{4.25}$$

from which follows that the coefficients are given as (up to a phase)

$$\varepsilon_{j,m}^{\pm} = \hbar\sqrt{j(j+1) - m(m \pm 1)}, \tag{4.26}$$

where $\varepsilon_{j,\pm j}^{\pm}$ vanish as desired.

4.1.2 Orbital angular momentum operator

The former section discussed the general properties of angular momentum operators. Now we turn to the concrete example of the orbital angular momentum operator. Where the definition follows from the classical expression under the correspondence principle

$$\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}, \tag{4.27}$$

thus the components of the orbital angular momentum operator are

$$\begin{aligned}\hat{L}_x =& \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right),\\ \hat{L}_y =& \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right),\\ \hat{L}_z =& \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right),\end{aligned} \tag{4.28}$$

which are hermitian operators and thus observables. To note is that their is no ambiguity in the definition in terms of the ordering of the operator, e.g. it holds that $\hat{y}\hat{p}_z = \hat{p}_z\hat{y}$ as these operators do commute. Taking the commutation relation for the position and momentum operator $[\hat{x}_n, \hat{p}_k] = i\hbar\delta_{n,k}$ we confirm for the commutator

$$\left[\hat{L}_x, \hat{L}_y\right] = \left[\left(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right), \left(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right)\right] = \left[\hat{p}_z, \hat{z}\right]\hat{y}\hat{p}_x + \left[\hat{z}, \hat{p}_z\right]\hat{x}\hat{p}_y = i\hbar\hat{L}_z, \tag{4.29}$$

analogous the results for the other commutators. In general one uses the quantum number l instead of j and we have for the eigenvalue problem

$$\begin{aligned}&\hat{L}_z\left|l\ m\right\rangle = \hbar m\left|l\ m\right\rangle,\\ &\hat{\mathbf{L}}^2\left|l\ m\right\rangle = \hbar^2 l(l+1)\left|l\ m\right\rangle, \quad m = -l, ..., l.\end{aligned} \tag{4.30}$$

Next we like to transfer the orbital angular momentum operator into position space, which is done as usual by multiplying from the left with the bra-vector in position space $\langle\mathbf{x}|$

$$\psi_{lm}(\mathbf{x}) = \langle\mathbf{x}|l\ m\rangle, \tag{4.31}$$

with this our eigenvalue problem becomes

$$\begin{aligned}\langle\mathbf{x}|\hat{L}_z\left|l\ m\right\rangle &= \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi_{lm}(\mathbf{x}) = \hbar m\ \psi_{lm}(\mathbf{x}),\\ \langle\mathbf{x}|\hat{\mathbf{L}}^2\left|l\ m\right\rangle &= -\hbar^2\left(\mathbf{x}\times\nabla\right)^2\psi_{lm}(\mathbf{x}) = \hbar^2 l(l+1)\ \psi_{lm}(\mathbf{x}),\end{aligned} \tag{4.32}$$

we now aim to translate the expressions of the angular momentum operator into spherical coordinates, the prescription for the transformation for the coordinates is

$$\begin{aligned}x &= r\sin(\theta)\cos(\phi), & r^2 &= x^2 + y^2 + z^2,\\ y &= r\sin(\theta)\sin(\phi), & \tan(\phi) &= \tfrac{y}{x},\\ z &= r\cos(\theta), & \cos(\theta) &= \tfrac{z}{r},\end{aligned}$$

and for the derivatives we find

$$\begin{aligned}\frac{\partial}{\partial x} =& \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi} + \frac{\partial r}{\partial x}\frac{\partial}{\partial r} = \frac{\cos(\theta)\cos(\phi)}{r}\frac{\partial}{\partial\theta} - \frac{\sin(\phi)}{r\sin(\theta)}\frac{\partial}{\partial\phi} + \sin(\theta)\cos(\phi)\frac{\partial}{\partial r},\\ \frac{\partial}{\partial y} =& \frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial y}\frac{\partial}{\partial\phi} + \frac{\partial r}{\partial y}\frac{\partial}{\partial r}, = \frac{\cos(\theta)\sin(\phi)}{r}\frac{\partial}{\partial\theta} + \frac{\cos(\phi)}{r\sin(\theta)}\frac{\partial}{\partial\phi} + \sin(\theta)\sin(\phi)\frac{\partial}{\partial r},\\ \frac{\partial}{\partial z} =& \frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial z}\frac{\partial}{\partial\phi} + \frac{\partial r}{\partial z}\frac{\partial}{\partial r} = -\frac{\sin(\theta)}{r}\frac{\partial}{\partial\theta} + \cos(\theta)\frac{\partial}{\partial r},\end{aligned} \tag{4.33}$$

which gives us for the components of the angular momentum operator in spherical coordinates

$$\begin{aligned}\hat{L}_x =& \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = i\hbar\left[\sin(\phi)\frac{\partial}{\partial\theta} + \cot(\theta)\cos(\phi)\frac{\partial}{\partial\phi}\right],\\ \hat{L}_y =& \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) = i\hbar\left[\cot(\theta)\sin(\phi)\frac{\partial}{\partial\phi} - \cos(\phi)\frac{\partial}{\partial\theta}\right],\\ \hat{L}_z =& \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = \frac{\hbar}{i}\frac{\partial}{\partial\phi},\end{aligned} \tag{4.34}$$

and the magnitude squared becomes

$$\hat{\mathbf{L}}^2 = -\hbar^2\left[\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\sin(\theta)\frac{\partial}{\partial\theta} + \frac{1}{\sin^2(\theta)}\frac{\partial^2}{\partial^2\phi}\right]. \tag{4.35}$$

To note is that, while inserting the expression for the components into the definition of $\hat{\mathbf{L}}^2$ one has to remember that is an operator acting on a state, thus when simplifying this expression one has to account for this. Finally, the lowering and raising operator become

$$\begin{aligned}\hat{L}_+ &= +\hbar e^{+i\phi}\left[\frac{\partial}{\partial\theta} + i\cot(\theta)\frac{\partial}{\partial\phi}\right],\\ \hat{L}_- &= -\hbar e^{-i\phi}\left[\frac{\partial}{\partial\theta} - i\cot(\theta)\frac{\partial}{\partial\phi}\right],\end{aligned} \tag{4.36}$$

to note is here that these operators are still hermitian conjugate of each other, for this one has to keep in mind that the adjoint operator of the derivative comes in with a minus sign. The eigenvalue problem becomes in spherical coordinates:

$$\begin{aligned}m\ \psi_{lm}(\mathbf{x}) =& -i\frac{\partial}{\partial\phi}\psi_{lm}(\mathbf{x}),\\ l(l+1)\ \psi_{lm}(\mathbf{x}) =& -\frac{1}{\sin^2(\theta)}\left[\sin(\theta)\frac{\partial}{\partial\theta}\sin(\theta)\frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial^2\phi}\right]\psi_{lm}(\mathbf{x}),\end{aligned} \tag{4.37}$$

from these equation we see that the radial component of the wave function will be completely undetermined by this equation system. The orbital angular momentum operators only act on the angular part of the wavefunction. We choose the separation ansatz

$$\psi_{lm}(\mathbf{x}) = R(r)Y_{lm}(\theta,\phi), \tag{4.38}$$

inserting this into the equation for the eigenvalue problem, the radial part of the wavefunction drops out. Taking the form of Eq.(4.37), we can make another separation ansatz for the angular part of the wavefunction

$$Y_{lm}(\theta,\phi) = \Phi_m(\phi)\Theta_{lm}(\theta). \tag{4.39}$$

With the first equation in Eq.(4.37) we find the solution

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}}\,e^{im\phi}, \tag{4.40}$$

with normalization coming from integration over the whole interval $\phi \in [0, 2\pi]$. As a wave-function (or state) has to be unique we need $\Phi_m(\phi) = \Phi_m(\phi + 2\pi)$ and thus m to be an integer, as well as l. So up to here we have

$$Y_{lm}(\theta, \phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} \Theta_{lm}(\theta), \quad l = 0, 1, 2, .. \; ; \; m = -l, ..., +l, \tag{4.41}$$

which we can insert into the second equation of Eq.(4.37) and obtain

$$l(l+1) \; \Theta_{lm}(\theta) = \; - \frac{1}{\sin^2(\theta)} \left[\sin(\theta) \frac{\partial}{\partial \theta} \sin(\theta) \frac{\partial}{\partial \theta} - m^2 \right] \Theta_{lm}(\theta), \tag{4.42}$$

we can rearrange this differential equation by applying the substitution

$$z = \cos(\theta) \quad \Rightarrow \quad \sin(\theta) \frac{\partial}{\partial \theta} = (z^2 - 1) \frac{\partial}{\partial z}, \tag{4.43}$$

and obtain

$$\left[\, l(l+1) - \frac{m^2}{1 - z^2} \; + \frac{\partial}{\partial z} (1 - z^2) \frac{\partial}{\partial z} \, \right] \Theta_{lm}(z) = 0, \tag{4.44}$$

this is a so-called generalized Legendre equation, a very well studied equation in mathematical physics, for example also appearing in connection to multipol moments of electrical charge densities. Important is that the solutions to the generalized Legendre equation are known, they are called **associate Legendre polynomials**:

$$\begin{aligned} P_l^m(z) =& (-1)^m \; (1 - z^2)^{\frac{m}{2}} \frac{d^m}{dz^m} P_l(z), \quad m \geq 0, \\ P_l^{-m}(z) =& (-1)^m \; \frac{(l-m)!}{(l+m)!} P_l^m(z), \quad m < 0, \end{aligned} \tag{4.45}$$

with the **Legendre polynomials**:

$$P_l(z) = \frac{1}{2^l l!} \frac{d^l}{dz^l} \; (z^2 - 1)^l \tag{4.46}$$

which are solutions to the ordinary Legendre equation, see [9]. The $P_l(z)$ are polynomial of grade l and the form a complete orthonormal system in the interval $[-1, +1]$. However, they are not normalized to unity

$$\int_{-1}^{1} dz \; P_l(z) P_k(z) = \frac{2}{2l+1} \delta_{l,k}, \tag{4.47}$$

but are orthogonal which translates as well to the associated Legendre polynomials

$$\int_{-1}^{1} dz \; P_l^m(z) P_k^m(z) = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{l,k}. \tag{4.48}$$

With this we can now determine the eigenfunctions for the angular momentum operators which have the form

$$Y_{lm}(\theta,\phi) = c_{l,m}\ \frac{1}{\sqrt{2\pi}} P_l^m(\cos(\theta))\ e^{im\phi}, \tag{4.49}$$

with the normalization coefficients $c_{n,m}$ left to determine:

$$\int_0^{2\pi} d\phi \int_{-1}^{1} d\cos(\theta)\ Y^*_{l'm'}(\theta,\phi) Y_{lm}(\theta,\phi) = \frac{1}{2\pi} c^*_{l',m'} c_{l,m} \int_0^{2\pi} d\phi \int_{-1}^{1} dz\ P_{l'}^{m'}(z) P_l^m(z)\ e^{i(m-m')\phi}$$

$$= c^*_{l',m} c_{l,m} \delta_{m,m'} \int_{-1}^{1} dz\ P_{l'}^m(z) P_l^m(z) = |c_{l,m}|^2 \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{l,l'} \delta_{m,m'}, \tag{4.50}$$

and we obtain the final result

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos(\theta))\ e^{im\phi}, \quad Y_{l-m}(\theta,\phi) = (-1)^m Y^*_{lm}(\theta,\phi), \tag{4.51}$$

which are the so-called spherical harmonics, who form a complete system form a complete set of orthogonal functions and thus an orthonormal basis, each function defined on the surface of a sphere can be written as a sum of these spherical harmonics.

4.1.3 Spin angular momentum operator

In 1922 Otto Stern and Walther Gerlach performed an experiment in which a beam of silver was send through an inhomogeneous magnetic field which were then detected on a glass screen. For the case that the atoms have a magnetic moment μ, they experience a force when they pass through the magnetic field. For example if the atoms fly in x-direction and the magnetic field has only one non-zero component in the z-direction $\mathbf{B} = B(z)\mathbf{e}_z$, one would classically expect that the atoms experience a force in the z-direction $F = \mu_z \partial B(z)/\partial z$. Hence, classically one would expect the observation that the atoms accumulate at one spot on the glass screen. However, the atoms accumulate not on one, but on two spots! The splitting of the atomic beam in two directions can, and thus the two possible values for the magnetic moment of the atoms, is ultimately traced back to the individual electrons. Furthermore, the effect also exists for a single electron.
To put this in another way, if the atoms (or electrons) would have a random spatial distribution of the magnetic moment while passing through the magnetic field in z-direction, one would classically expect a continuous distribution on the glass screen along the z-axis. This is not the case, as one observes two distributions, hence the magnetic moment takes only two possible values and thus is quantized. The magnetic moment of a charged particle is directly connected to the angular momentum via the gyromagnetic ratio. Hence, particles such as electrons posses an intrinsic and quantized angular momentum – the so-called spin.
The spin angular momentum is denoted as $\hat{\mathbf{S}}$ and follows the algebra introduced in the former sections. In terms of notation one uses $j = s$ for the spin. In contrast to the orbital angular

momentum operator it can also have half-integer values. This leads to a fundamentally important classification of particles: particles with half-integer spins are called **fermions**, while integer spin particles are called **bosons**. Fermions are for example electrons or neutrinos, they follow the so-called Fermi-Dirac statistic and obey the Pauli-exclusion principle, stating that two or more identical particles cannot occupy the same quantum state. The latter is not the case for bosons, which follow the Bose-Einstein statistics, examples are photons, phonons or – a very famous one, the Higgs-particle.
We can start by discussing the consequences of the spin for the case of electrons. Electrons have a spin of $s = \frac{1}{2}$ and are thus fermions. The components of the spin angular momentum operator can be expressed with the help of the Pauli-spin operators:

$$\hat{S}_n = \frac{\hbar}{2}\hat{\sigma}_n, \tag{4.52}$$

with the definitions (same as given in Eq.(2.134))

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.53}$$

the operator $\hat{\mathbf{S}}^2$ takes the form

$$\hat{\mathbf{S}}^2 = \frac{\hbar^2}{4}\left[\hat{\sigma}_x^2 + \hat{\sigma}_y^2 + \hat{\sigma}_z^2\right] = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.54}$$

which confirms our spin values as the eigenvalue for the magnitude squared is $\hbar^2 s(s+1)$. There exists different conventions on how to label the eigenstates $|s\ m\rangle$

$$\left|\frac{1}{2}\ \frac{1}{2}\right\rangle = |\uparrow\rangle = |+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \left|\frac{1}{2}\ -\frac{1}{2}\right\rangle = |\downarrow\rangle = |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.55}$$

one also calls the eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ spin-up and spin-down state respectively. To account for the spin as an intrinsic property of the electron, one has to extend the description of the state vector accounting for the spin. One can define a so-called spinor

$$|\psi(\mathbf{x})\rangle = \begin{pmatrix} \psi_\uparrow(\mathbf{x}) \\ \psi_\downarrow(\mathbf{x}) \end{pmatrix} \tag{4.56}$$

to account for this extra degree of freedom on top of the spacial degrees. As mentioned above, angular momentum is directly connected to the magnetic moment. For the electron we have

$$\hat{\mu} = -\mu_B \hat{S}, \quad \mu_B = \frac{e\hbar}{2mc} = 9.2740154 \times 10^{-24}\frac{\mathrm{J}}{\mathrm{T}}, \tag{4.57}$$

with the so-called Bohr magneton as the gyromagnetic ratio.

4.1.4 Clebsch-Gordan coefficients

A particle can both have a spin and an orbital angular momentum, e.g. taking for the example an atom with one electron moving in the spherical potential of the nucleus. For the latter case

exist as well a so-called spin-orbit interaction, which is a relativistic interaction leading to the shift in an electrons' atomic energy levels. Leaving this aside for the moment, we can also think of the example of two particles moving on a sphere with angular momentum $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$. In either cases we have to understand how to deal with the addition of angular momentum operators is conducted.

We start from two angular momentum operators $\hat{\mathbf{J}}_{1,2}$. Each of these vector operators has three components and is a so-called rank one Cartesian tensor operator

$$\hat{\mathbf{J}}_n = \begin{pmatrix} \hat{J}_{x,n} \\ \hat{J}_{y,n} \\ \hat{J}_{z,n} \end{pmatrix}, \quad n = 1, 2. \tag{4.58}$$

The components individually obey the standard commutation relations in Eq.(4.2). The angular momentum operators have their individual eigenstates

$$|j_1 m_1\rangle \quad \text{and} \quad |j_2 m_2\rangle\,, \tag{4.59}$$

with j_n as the angular momentum quantum number and m_n is the angular momentum projection onto the z-axis. We define V_n as the space each tensor operator acts upon, the dimension of the space is $2j_n + 1$ (corresponding to the number for the values of m). Remember for example the spin angular momentum operator for the electron, there we have $s = \frac{1}{2}$ and thus the dimension of the corresponding space is 2. As we are now considering two tensor operators, we can define combined operators acting on the tensor product space $V_{12} = V_1 \otimes V_2$ which has the dimension $(2j_1 + 1)(2j_2 + 1)$. Remember how tensor products work, e.g. for the example of two 2-dimensional matrices:

$$\begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix} \otimes \begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} = \begin{bmatrix} x_{1,1}\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} & x_{1,2}\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} \\ x_{2,1}\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} & x_{2,2}\begin{bmatrix} y_{1,1} & y_{1,2} \\ y_{2,1} & y_{2,2} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} x_{1,1}y_{1,1} & x_{1,1}y_{1,2} & x_{1,2}y_{1,1} & x_{1,2}y_{1,2} \\ x_{1,1}y_{2,1} & x_{1,1}y_{2,2} & x_{1,2}y_{2,1} & x_{1,2}y_{2,2} \\ x_{2,1}y_{1,1} & x_{2,1}y_{1,2} & x_{2,2}y_{1,1} & x_{2,2}y_{1,2} \\ x_{2,1}y_{2,1} & x_{2,1}y_{2,2} & x_{2,2}y_{2,1} & x_{2,2}y_{2,2} \end{bmatrix}$$

where the tensor product is a 4-dimensional matrix. The tensor product space V_{12} can be expressed in the uncoupled basis as

$$|j_1 m_1 j_2 m_2\rangle \equiv |j_1 m_1\rangle \otimes |j_2 m_2\rangle\,, \tag{4.60}$$

the individual angular momentum operators act in this space as follows

$$\begin{aligned} (\hat{\mathbf{J}}_1 \otimes \mathbb{1})\,|j_1 m_1 j_2 m_2\rangle &\equiv \hat{\mathbf{J}}_1\,|j_1 m_1\rangle \otimes |j_2 m_2\rangle\,, \\ (\mathbb{1} \otimes \hat{\mathbf{J}}_2)\,|j_1 m_1 j_2 m_2\rangle &\equiv\ |j_1 m_1\rangle \otimes \hat{\mathbf{J}}_2\,|j_2 m_2\rangle\,, \end{aligned} \tag{4.61}$$

and we define the combined angular momentum operator as a coproduct

$$\hat{\mathbf{J}} \equiv \hat{\mathbf{J}}_1 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{J}}_2 = \left[\hat{J}_x, \hat{J}_y, \hat{J}_z\right]^T, \tag{4.62}$$

the components of this operator still have the form of the commutation relations as given in Eq.(4.2). The action of the combined or total angular angular momentum operator constitutes a representation of the SU(2) Lie algebra, but a so-called reducible one. The aim is now to reduce the reducible representation into irreducible parts. In other words, we aim to define new eigenstates for the total angular momentum operator via an expansion in the uncoupled tensor product basis. The coefficients associated with this expansion are the so-called Clebsch-Gordan coefficients. Note, for this section the corresponding Wikipedia article was helpful ↗ .

A first thing to note is that the states defined in Eq.(4.60) are eigenstates of the z-component of the total angular momentum operator, we have

$$\hat{J}_z \left|j_1 m_1 j_2 m_2\right\rangle = \left(\hat{J}_{z,1} \otimes \mathbb{1} + \mathbb{1} \otimes \hat{J}_{z,2}\right) \left|j_1 m_1 j_2 m_2\right\rangle = \hbar(m_1 + m_2) \left|j_1 m_1 j_2 m_2\right\rangle . \tag{4.63}$$

However, we can check if they are as well eigenstates of the magnitude squared operator

$$\hat{\mathbf{J}}^2 \left|j_1 m_1 j_2 m_2\right\rangle = \left(\hat{\mathbf{J}}_1^2 \otimes \mathbb{1} + \mathbb{1} \otimes \hat{\mathbf{J}}_2^2 + 2\hat{\mathbf{J}}_1 \otimes \hat{\mathbf{J}}_2\right) \left|j_1 m_1 j_2 m_2\right\rangle , \tag{4.64}$$

and find that due to the last term these states are no eigenstates for the operator. This means that one cannot perform a measurement in this basis and obtain information about both operators. To be able to extract more information out of the system we need to find other eigenstates. For this we first need to find the right operators which commute and thus have a simultaneous eigenstate. This is straightforwardly done:

$$\left[\hat{\mathbf{J}}^2, \hat{J}_z\right] = \left[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_1^2\right] = \left[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_2^2\right] = 0, \tag{4.65}$$

and thus we have a set of four hermitian operators associated with four quantum numbers

$$\left\{\hat{\mathbf{J}}_1^2, \hat{\mathbf{J}}_2^2, \hat{\mathbf{J}}^2, \hat{J}_z\right\} \quad \longrightarrow \quad \left|j_1 j_2 J M\right\rangle . \tag{4.66}$$

the expression in the ket vector notation is the new simultaneous eigenstate with the four quantum numbers j_1, j_2, J, M. The hermitian operators obey the eigenvalue equation

$$\begin{aligned}
\hat{\mathbf{J}}_n^2 \left|j_1 j_2 J M\right\rangle &= \hbar^2 j_n(j_n+1) \left|j_1 j_2 J M\right\rangle , \\
\hat{\mathbf{J}}^2 \left|j_1 j_2 J M\right\rangle &= \hbar^2 J(J+1) \left|j_1 j_2 J M\right\rangle , \\
\hat{J}_z \left|j_1 j_2 J M\right\rangle &= \hbar M \left|j_1 j_2 J M\right\rangle ,
\end{aligned} \tag{4.67}$$

with the bounded M-ladder $-J \leq M \leq J$ as before, meaning we can as well define raising and lowering operators for the total angular momentum

$$\hat{J}_\pm \left|j_1 j_2 J M\right\rangle = \hbar\sqrt{J(J+1) - M(M \pm 1)} \ \left|j_1 j_2 J M\right\rangle . \tag{4.68}$$

The aim is now to express the simultaneous eigenstate with help of the tensor product states defined in Eq.(4.60). The later forms an orthonormal basis, and thus we can expand the simultaneous eigenstate in it:

$$\left|j_1 j_2 J M\right\rangle = \sum_{m_1, m_2} \sum_{j_1', j_2'} \left|j_1' m_1 j_2' m_2\right\rangle \left\langle j_1' m_1 j_2' m_2 | j_1 j_2 J M\right\rangle . \tag{4.69}$$

First we can show that the summation over $j_{1,2}'$ will vanish, for this we start from

$$0 = \left\langle j_1' m_1 j_2' m_2 | (\hat{\mathbf{J}}_n^2 - \hat{\mathbf{J}}_n^2) | j_1 j_2 J M \right\rangle = \hbar^2 \left[j_n'(j_n'+1) - j_n(j_n+1)\right] \left\langle j_1' m_1 j_2' m_2 | j_1 j_2 J M\right\rangle , \tag{4.70}$$

where we omitted the tensor $\otimes$ notation for simplicity (and will do so in what follows). Important is that, as we have an hermitian operator $\hat{\mathbf{J}}_n^2$, the eigenvalues and quantum numbers are positive, hence the summation vanishes for $j_n' \neq j_n$ and we set $j_n' = j_n$ and remove the summation, leaving us with the expansion

$$|j_1 j_2 JM\rangle = \sum_{m_1,m_2} |j_1 m_1 j_2 m_2\rangle \, C^{JM}_{j_1 m_1, j_2 m_2}, \quad C^{JM}_{j_1 m_1, j_2 m_2} \equiv \langle j_1 m_1 j_2 m_2 | j_1 j_2 JM\rangle \,, \tag{4.71}$$

with $C^{JM}_{j_1 m_1, j_2 m_2}$ as the so-called Clebsch-Gordan coefficients. The absolute value of this quantity measures the probability that for two particles with total angular momentum j_1 and j_1 at fixed J and M one finds one particle with $\left\langle \hat{J}_{1,z} \right\rangle = \hbar m_1$ and the other with $\left\langle \hat{J}_{2,z} \right\rangle = \hbar m_2$ [5]. We are now left with determining the Clebsch-Gordan coefficients and with understanding which values J and M can take. For the later we start from

$$\begin{aligned} 0 =& \left\langle j_1 m_1 j_2 m_2 | (\hat{J}_z - \hat{J}_z) | j_1 j_2 JM \right\rangle = \left\langle j_1 m_1 j_2 m_2 | (\hat{J}_{z,1} + \hat{J}_{z,2} - \hat{J}_z) | j_1 j_2 JM \right\rangle \\ =& \hbar \left[m_1 + m_2 - M \right] \langle j_1 m_1 j_2' m_2 | j_1 j_2 JM \rangle \,, \end{aligned} \tag{4.72}$$

resulting in the condition

$$M = m_1 + m_2, \tag{4.73}$$

which will simply the summation over $m_{1,2}$, i.e., it gives us selection criteria which m_1 and m_2 are relevant. What remains is the knowing about which possible values J can take. Deriving the condition leaves for J is a bit cumbersome, and we omit it here, for details see [9]. The important result is the relation

$$|j_1 - j_2| \leq J \leq j_1 + j_2, \tag{4.74}$$

providing us with a lower and upper bound for our quantum number J. With this we have all relevant ingredients, however, one can as well make more arguments about which Glebsch-Gordan coefficients have to vanish due to symmetry, for details see [9].
We will conclude this section by discussing a concrete example. We consider two angular momentum operators j_1 and j_2 with the values

$$j_1 = 1; \quad j_2 = \frac{1}{2} \quad \Rightarrow \quad J = \frac{1}{2}, \frac{3}{2}. \tag{4.75}$$

To make our notation a bit more compact we use the fact that we have now fixed values for $j_{1,2}$ and remove it from our expansion expression

$$|JM\rangle = \sum_{m_1,m_2} |m_1\rangle_1 \, |m_2\rangle_2 \; C^{JM}_{m_1,m_2}, \quad |m_n\rangle_n \equiv |j_n m_n\rangle \,, \tag{4.76}$$

the subindex on the ket-vectors shall help us identify the right state vector of the two subspaces. In what follows we will need the action of the raising and lowering operators

$$\hat{J}_{\pm,n} \, |j_n m_n\rangle = \varepsilon^{\pm}_{j_n,m_n} \, |j_n (m_n \pm 1)\rangle \,, \quad \varepsilon^{\pm}_{j_n,m_n} = \hbar \sqrt{j_n(j_n+1) - m_n(m_n \pm 1)}. \tag{4.77}$$

The quantum number J can take two possible values, we choose here $J = \frac{1}{2}$, which means that the quantum number of the z-components is $M = \pm\frac{1}{2}$. We will focus on the case with the positive sign. From the values of $j_{1,2}$ we know which values $m_{1,2}$ take.

$$j_1 = 1 \quad \Rightarrow \quad m_1 = -1, 0, 1; \qquad j_2 = \frac{1}{2} \quad \Rightarrow \quad m_2 = -\frac{1}{2}, \frac{1}{2}. \tag{4.78}$$

From the condition $M = m_1 + m_2$ we can select which $m_{1,2}$ are relevant:

$$\left|\frac{1}{2}\frac{1}{2}\right\rangle = |0\rangle_1 \left|\frac{1}{2}\right\rangle_2 C^+_{0,\frac{1}{2}} + |1\rangle_1 \left|-\frac{1}{2}\right\rangle_2 C^+_{1,-\frac{1}{2}}, \qquad C^+_{n,m} \equiv C^{\frac{1}{2}\frac{1}{2}}_{n,m}. \tag{4.79}$$

We now use the fact that $M = \frac{1}{2}$ is the upper bound for the M-ladder, meaning that we cannot raise the state to higher M and have the condition

$$\begin{aligned}
\hat{J}_+ \left|\frac{1}{2}\frac{1}{2}\right\rangle = 0 =& \left(\hat{J}_{+,1} + \hat{J}_{+,2}\right)\left[|0\rangle_1 \left|\frac{1}{2}\right\rangle_2 C^+_{0,\frac{1}{2}} + |1\rangle_1 \left|-\frac{1}{2}\right\rangle_2 C^+_{1,-\frac{1}{2}}\right] \\
=& \hat{J}_{+,1}|0\rangle_1 \left|\frac{1}{2}\right\rangle_2 C^+_{0,\frac{1}{2}} + |1\rangle_1 \hat{J}_{+,2}\left|-\frac{1}{2}\right\rangle_2 C^+_{1,-\frac{1}{2}} \\
=& \hbar\left(\sqrt{2}\,|1\rangle_1 \left|\frac{1}{2}\right\rangle_2 C^+_{0,\frac{1}{2}} + |1\rangle_1 \left|\frac{1}{2}\right\rangle_2 C^+_{1,-\frac{1}{2}}\right) = \hbar\left(\sqrt{2}\, C^+_{0,\frac{1}{2}} + C^+_{1,-\frac{1}{2}}\right)|1\rangle_1 \left|\frac{1}{2}\right\rangle_2
\end{aligned} \tag{4.80}$$

in the second line we have used that the states with quantum number $m_1 = 1$ and $m_2 = 1/2$ are as well the upper bound for their respective ladder of states. In addition we used Eq.(4.77) to determine the normalization coefficients when operating with the individual raising operators on the states. From all this we find the condition

$$C^+_{0,\frac{1}{2}} = -\frac{1}{\sqrt{2}}\, C^+_{1,-\frac{1}{2}}. \tag{4.81}$$

On top of this condition, normalization provides us with the relation

$$\left|C^+_{0,\frac{1}{2}}\right|^2 + \left|C^+_{1,-\frac{1}{2}}\right|^2 = 1 \quad \Rightarrow \quad \left|C^+_{1,-\frac{1}{2}}\right|^2 = \frac{2}{3}, \tag{4.82}$$

and with this we have the result for the Glebsch-Gordon coefficients and the state

$$C^+_{0,\frac{1}{2}} = -\sqrt{\frac{1}{3}}, \quad C^+_{1,-\frac{1}{2}} = \sqrt{\frac{2}{3}} \quad \Rightarrow \quad \left|\frac{1}{2}\frac{1}{2}\right\rangle = \sqrt{\frac{2}{3}}\,|1\rangle_1 \left|-\frac{1}{2}\right\rangle_2 - \sqrt{\frac{1}{3}}\,|0\rangle_1 \left|\frac{1}{2}\right\rangle_2. \tag{4.83}$$

The state associated with $M = -1/2$ follows from the action of the lowering operator

$$\begin{aligned}
\hat{J}_- \left|\frac{1}{2}\frac{1}{2}\right\rangle =& \hbar\left|\frac{1}{2} - \frac{1}{2}\right\rangle = \hat{J}_{-,1}\left[\sqrt{\frac{2}{3}}\,|1\rangle_1 \left|-\frac{1}{2}\right\rangle_2 - \sqrt{\frac{1}{3}}\,|0\rangle_1 \left|\frac{1}{2}\right\rangle_2\right] - \sqrt{\frac{1}{3}}\,|0\rangle_1\, \hat{J}_{-,2}\left|\frac{1}{2}\right\rangle_2 \\
=& \hbar\left[\sqrt{\frac{4}{3}}\,|0\rangle_1 \left|-\frac{1}{2}\right\rangle_2 - \sqrt{\frac{2}{3}}\,|-1\rangle_1 \left|\frac{1}{2}\right\rangle_2\right] - \hbar\sqrt{\frac{1}{3}}\,|0\rangle_1 \left|-\frac{1}{2}\right\rangle_2
\end{aligned} \tag{4.84}$$

and we obtain

$$\left|\frac{1}{2} - \frac{1}{2}\right\rangle = \sqrt{\frac{1}{3}}\,|0\rangle_1 \left|-\frac{1}{2}\right\rangle_2 - \sqrt{\frac{2}{3}}\,|-1\rangle_1 \left|\frac{1}{2}\right\rangle_2. \tag{4.85}$$

4.1.5 The Wigner-Eckart theorem

In this section we will further discuss the topic rotations in quantum mechanics, this is an important, but rather extensive topic and in this course we will only discuss selected elements. For further reading the script by R. Littlejohn ↗ is recommended. Parts of this section are as well adapted from their script, i.e., chapter 18 and 20.
We already mentioned at the beginning of this chapter that the applications of the angular momentum operator on a state represents a rotation. A proper definition of the rotation operator, which is a unitary operator, is given as

$$\hat{U}[\varphi, \hat{n}] = e^{-\frac{i}{\hbar}\hat{n}\varphi\hat{J}}, \tag{4.86}$$

where $\hat{n}$ denotes the rotation axis and φ the rotation angle. One can define so-called Wigner D-matrices, which appear when we rotate a basis vector in the angular momentum basis, i.e., using the completeness relation we find

$$\hat{U}[\varphi, \hat{n}] \left|jm\right\rangle = \sum_{j',m'} \left|j'm'\right\rangle \left\langle j'm'|\hat{U}[\varphi, \hat{n}]|jm\right\rangle \equiv \sum_{m'} \left|jm'\right\rangle D^{j}_{m'm}, \tag{4.87}$$

with $D^{j}_{m'm}$ as the elements of the Wigner D-matrix. The summation over j' vanishes, as the associated operator, the magnitude squared of the angular momentum operator, commutes with the components of the angular momentum operator and thus with the rotation operator. The consequence is that if we rotate the state for a fixed value j we obtain a linear combination of basis vectors with the same quantum number j, where the Wigner D-matrix provides us the expansion coefficients. There is no mixing of different values of j. Moreover, states can as well be invariant under rotations, so called rotational invariant states, a trivial example is the $j = 0$ state where all components of the angular momentum operator are zero.
Next we can consider the effect of the rotation operator on the tensor product space $V_{12} = V_1 \otimes V_2$ which we have introduced in the last section when we considered the combined space on which the two angular momentum operators $\hat{J}_{1,2}$ act upon. Taking that the total angular momentum operator is now a combination of those, i.e., $\hat{J} = \hat{J}_1 + \hat{J}_2$,

$$\begin{aligned}\hat{U}[\varphi, \hat{n}] \left|j_1 m_1 j_2 m_2\right\rangle =& e^{-\frac{i}{\hbar}\hat{n}\varphi(\hat{J}_1+\hat{J}_2)} \left|j_1 m_1 j_2 m_2\right\rangle = \left[e^{-\frac{i}{\hbar}\hat{n}\varphi\hat{J}_1} \left|j_1 m_1\right\rangle\right] \left[e^{-\frac{i}{\hbar}\hat{n}\varphi\hat{J}_2} \left|j_2 m_2\right\rangle\right] \\ =& \sum_{m_1' m_2'} \left|j_1 m_1' j_2 m_2'\right\rangle D^{j_1}_{m_1' m_1} D^{j_2}_{m_2' m_2}, \end{aligned} \tag{4.88}$$

as the individual operators commute we can apply the rotation one each state separately.
One action of the rotation operator on the state is referred to as active rotation, while there also exists the notion of passive rotation where the operators instead of the state are rotated. Hence we should pose the question how operators transform under rotations. We will keep this discussion rather short, for details please see ↗ . Rotations are unitary transformation, so we can start from what we already know about such transformation. The expectation value of an operator $\hat{O}$ should not change now matter what we rotate:

$$\left\langle jm\right| \hat{U}^{\dagger}[\varphi, \hat{n}] \hat{O} \hat{U}[\varphi, \hat{n}] \left|jm\right\rangle \equiv \left\langle jm\right| \hat{O}_{-\varphi} \left|jm\right\rangle \quad \Rightarrow \quad \hat{O}_{\varphi} = \hat{U}[\varphi, \hat{n}] \hat{O} \hat{U}[\varphi, \hat{n}], \tag{4.89}$$

where $\hat{O}_{-\varphi}$ denotes the operator rotated in the opposite direction than the state was rotated in. And the transpose rotation operator simply corresponds to an rotation in the opposite direction. We want to focus now on a specific class of operators: the spherical tensor operators: a spherical tensor operator is denoted as $T^{(j)}_m$ and is a collection of $(2j+1)$ operators which transform under rotations in the same way as spherical harmonics do. The spherical harmonics are the functions $Y_{lm}(\theta,\phi)$, which we got to know as the eigenfunctions of the angular momentum operator. For example, accounting as well for some radial dependence, we can define $\widetilde{Y}_{lm}(\mathbf{x})$ for $l=1$ are

$$\begin{aligned}
\widetilde{Y}_{1-1}(\mathbf{x}) =& rY_{1-1}(\theta,\phi) = \; r\sqrt{\frac{3}{8\pi}}\sin(\theta)e^{-i\varphi} = \sqrt{\frac{3}{8\pi}}\,(x-iy)\,,\\
\widetilde{Y}_{10}(\mathbf{x}) =& \, rY_{10}(\theta,\phi) = r\sqrt{\frac{3}{4\pi}}\cos(\theta) = \sqrt{\frac{3}{4\pi}}z,\\
\widetilde{Y}_{11}(\mathbf{x}) =& \, rY_{11}(\theta,\phi) = -r\sqrt{\frac{3}{8\pi}}\sin(\theta)e^{i\varphi} = -\sqrt{\frac{3}{8\pi}}\,(x+iy)\,,
\end{aligned} \tag{4.90}$$

and thus a rank-1 operator can be expressed as a linear combination of these three spherical harmonics, e.g., a rank-1 operator is a vector with m components, and we have

$$\begin{aligned}
x =& \sqrt{\frac{2\pi}{3}}\left[\widetilde{Y}_{1-1}(\mathbf{x}) - \widetilde{Y}_{11}(\mathbf{x})\right] \equiv \frac{1}{\sqrt{2}}\left[T^{(1)}_{-1} - T^{(1)}_{1}\right],\\
y =& i\sqrt{\frac{2\pi}{3}}\left[\widetilde{Y}_{1-1}(\mathbf{x}) + \widetilde{Y}_{11}(\mathbf{x})\right] \equiv \frac{i}{\sqrt{2}}\left[T^{(1)}_{-1} + T^{(1)}_{1}\right],\\
z =& \sqrt{\frac{4\pi}{3}}\widetilde{Y}_{10}(\mathbf{x}) = T^{(1)}_0,
\end{aligned} \tag{4.91}$$

here we defined the three components of the spherical rank-1 tensor as

$$T^{(1)}_m = \sqrt{\frac{4\pi}{3}}\widetilde{Y}_{1m}(\mathbf{x}), \qquad m=-1,0,1. \tag{4.92}$$

So how do spherical harmonics transform under rotations? The spherical harmonics are the eigenstates of the angular momentum operator projected onto the position basis

$$\psi_{lm}(\mathbf{x}) = \langle \mathbf{x}|lm\rangle = R(r)\; Y_{lm}(\theta,\phi), \tag{4.93}$$

here we account for a dependence on the magnitude of $\mathbf{x}$, however, the important part is the angular part $Y_{lm}(\theta,\phi)$, i.e., where the eigenstates are projected onto the unit length basis. Taking the state rotation $\psi_{lm}(\mathbf{x}) \xrightarrow{\varphi} \psi_{lm}(\mathbf{x}')$ in the basis of the spherical harmonics we obtain

$$\psi_{lm}(\mathbf{x}') = \langle \mathbf{x}'|lm\rangle = \left\langle \mathbf{x}|\hat{U}^\dagger[\varphi,\hat{n}]|lm\right\rangle = \sum_{l',m'}\langle \mathbf{x}|l'm'\rangle\left\langle l'm'|\hat{U}^\dagger[\varphi,\hat{n}]|lm\right\rangle = \sum_{m'}\psi_{lm'}(\mathbf{x})D^{l*}_{m,m'}, \tag{4.94}$$

where we have a again the Wigner D-matrix elements appearing. This is a rotation in the opposite direction than as defined in Eq. (4.87), this becomes clear when we rotate in the

opposite direction, i.e., we have $\psi_{lm}(\mathbf{x}) \xrightarrow{-\varphi} \psi_{lm}(\mathbf{x}'')$ and obtain

$$\psi_{lm}(\mathbf{x}'') = \langle \mathbf{x}''|lm\rangle = \left\langle \mathbf{x}|\hat{U}^\dagger[-\varphi,\hat{n}]|lm\right\rangle = \left\langle \mathbf{x}|\hat{U}[\varphi,\hat{n}]|lm\right\rangle = \sum_{m'} \psi_{lm'}(\mathbf{x}) D^l_{m',m}. \tag{4.95}$$

Let us now define the spherical tensor operator $T^{(j)}_m$ of rank j and component m more deeply. These tensor operators are a set of operators that transform in a special way under rotations, meaning that if we apply a rotation onto the tensor operator the rank of the operator does not change, i.e., we have

$$\hat{U}[\varphi,\hat{n}] T^{(j)}_m \hat{U}^\dagger[\varphi,\hat{n}] = \sum_{m'} T^{(j)}_{m'} D^j_{m'm}, \tag{4.96}$$

and thus one obtains simply a linear combination of tensors with the same rank. Hence, they have transformation properties similar to those of the spherical harmonics and thus the state vectors with the same quantum numbers.
An important relation for such tensors is provided by the so-called Wigner-Eckart theorem which is based on an analysis on how operators transform under rotation. It has many practical applications, e.g., one can use it to determine quickly the selection rules for the matrix element that follow from rotational invariance. We will directly start with the explicit expression of the Wigner-Eckart theorem

$$\langle j_1 m_1|\, T^{(j)}_m \,|j_2 m_2\rangle = \langle j_2 m_2 jm|j_2 j j_1 m_1\rangle \;\left\langle j_1||T^{(j)}||j_2\right\rangle = C^{j_1 m_1}_{j_2 m_2, jm} \;\left\langle j_1||T^{(j)}||j_2\right\rangle \tag{4.97}$$

with the Clebsch-Gordon coefficients defined in Eq.(4.77). Let us define the different objects in this equation, first of all, we have on the left hand side the $T^{(j)}_m$ as a the spherical tensor operator of rank j and component m. Spherical tensor operators are expressed in a spherical basis. In case one start with a tensor operator in a different basis one first has to express the operator in the form of spherical tensor operators before one can apply the Wigner-Eckart theorem. Hence, the left hand side of Eq.(4.97) is the matrix element of spherical tensor operator evaluated between two states of given angular momentum with quantum numbers j_1, m_1 and j_2, m_2. On the right hand side, we first have the Clebsch-Gordon coefficient associated with the addition of angular momentum operators with quantum numbers j_2, m_2 and j, m, followed by the so-called reduced matrix element. The later is independent of the quantum numbers $m_{1,2}$ and m, and the notation with the double lines is a special notation used in the Wigner-Eckart theorem, we will see that this is simply a proportionality factor. The main point of the theorem is that we can express the matrix elements of our tensor operator associated with all possible combinations of the quantum numbers $m_{1,2}$ and m simply by the product of the corresponding Clebsch-Gordon coefficient times this proportionality factor (and there exist tables where one can extract the Clebsch-Gordon coefficient from). Hence, we only have to calculate the proportionality factor once and all corresponding matrix elements follow by multiplication with the corresponding Clebsch-Gordon coefficient.
Why is this practical? It is for example an important element for calculating possible transitions among levels in an atom exposed to an electric field. We will talk about the physical aspect of this in more detail in a later chapter. For the moment we stick to the simple picture of asking for the transition between two levels of the same energy, but we only know we start out in the a state with angular momentum quantum number $j_2 = 2$ and ask which transitions

to a state with angular momentum $j_1 = 1$ are possible. We assume that the electric field is applied in the z-direction, i.e., we have $\mathbf{E} = E_z \mathbf{e}_z$, which means that we need to calculate the transition matrix elements

$$d_{m_1,m_2} = E_z \langle 1m_1|z|2m_2\rangle \,, \tag{4.98}$$

in principle these are three possibilities for $m_1 = -1, 0, 1$ and five possible values for $m_2 = -2, -1, 0, 1, 2$, meaning we have $3 \times 5 = 15$ possible transition matrix elements. The nice thing is, with the Wigner-Eckart theorem we only have to calculate one of this elements to obtain the right proportionality factor. As it is not important which element we take, we can just start with the simplest example: $m_{1,2} = 0$. The first think we need is the to express our rank-1 tensor component z in spherical coordinates:

$$z = \sqrt{\frac{4\pi}{3}} \widetilde{Y}_{10}(\mathbf{x}) = T_0^{(1)} = r\cos\theta, \tag{4.99}$$

then we can evaluate the transition matrix element for $m_{1,2} = 0$ in spherical coordinates, meaning we represent the corresponding state with spherical harmonics (setting for now $r = 1$)

$$\begin{aligned} d_{0,0} =& E_z \int_0^{2\pi}\int_{-1}^{+1} d\cos\theta \;\; Y_{10}^*(\theta,\phi)\cos\theta\, Y_{20}(\theta,\phi) = E_z \frac{\sqrt{15}}{8\pi} \int_0^{2\pi}\int_{-1}^{+1} d\cos\theta \;\; \cos^2\theta \;\; \left(3\cos^2\theta - 1\right) \\ =& E_z \frac{\sqrt{15}}{4} \int_{-1}^{+1} dz \;\; \left(3z^4 - z^2\right) = E_z \frac{\sqrt{15}}{4} \left[\frac{3}{5}z^5 - \frac{1}{3}z^3\right]_{-1}^{1} = \frac{2E_z}{\sqrt{15}}. \end{aligned} \tag{4.100}$$

Taking that we have a rank-1 tensor operator in spherical coordinates we can apply the Wigner-Eckhart theorem

$$\langle 1m_1|\, T_0^{(1)}\, |2m_2\rangle = C^{1m_1}_{2m_2,1m} \;\; \langle 1||T^{(1)}||2\rangle = \frac{1}{E_z} d_{m_1,m_2}, \tag{4.101}$$

we have the result for the case that $m_{1,2} = 0$ (and thus $m = 0$) and obtain the operator

$$C^{10}_{20,10} \;\; \langle 1||T^{(1)}||2\rangle = \frac{2}{\sqrt{15}}, \;\; \Rightarrow \;\; \langle 1||T^{(1)}||2\rangle = \frac{1}{C^{10}_{20,10}} \frac{2}{\sqrt{15}} = -\sqrt{\frac{5}{2}}\frac{2}{\sqrt{15}} = -\sqrt{\frac{2}{3}}, \tag{4.102}$$

thus we have found our proportionality factor and can express the matrix elements as

$$\langle 1m_1|\, T_0^{(1)}\, |2m_2\rangle = -\sqrt{\frac{2}{3}} C^{1m_1}_{2m_2,1m}, \tag{4.103}$$

knowing now simply the Clebsch-Gordon coefficient we can determine all 15 matrix elements and thus transition probability. For example the transition $m_1 = 1$ and $m_2 = 2$ is not possible as the corresponding Clebsch-Gordon coefficient is zero. Moreover, as we can construct as well the components x and y of the rank-1 tensors, the proportionality factor stays the same and one can as well determine the transition elements for the case of having an electric field in the x and/or y direction. In total this means that with calculating the proportionality factor once, meaning solving one integral, on can obtain solutions for $3 \times 15 = 45$ transition matrix elements without solving 45 integrals! This is the power of the Wigner-Eckart theorem. Note this example was adapted from ↗ .

Chapter 5

Particle in central potentials and fields

5.1 Particle in a central potential

5.1.1 Separation ansatz

We consider a particle in a central potential, meaning a particle which is forced on a circular orbit. The general Hamiltonian for a particle of mass m in such a potential is given as

$$\hat{H} = -\frac{\hbar^2 \Delta}{2m} + V(\mathbf{x}), \tag{5.1}$$

where the force field is described by a potential $V(\mathbf{x}) = V(r)$, which depends only on the radius r. Such a situation appears for example in a hydrogen atom, in which an electron experience a Coulomb force due to the proton in the core. As we have a spherical potentials we want to transform the Hamiltonian into spherical coordinates, meaning the Laplace operator

$$\Delta = \frac{\partial^2}{\partial^2 r} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{\hbar^2 r^2}\,\hat{\mathbf{L}}^2 \tag{5.2}$$

with the magnitude squared of the angular momentum operator

$$\hat{\mathbf{L}}^2 = -\hbar^2\left[\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\sin(\theta)\frac{\partial}{\partial\theta} + \frac{1}{\sin^2(\theta)}\frac{\partial^2}{\partial^2\phi}\right]. \tag{5.3}$$

For the moment we will ignore the spin of the particle and work with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial^2 r} + \frac{2}{r}\frac{\partial}{\partial r}\right] + \frac{\hat{\mathbf{L}}^2}{2mr^2} + V(r) \equiv \frac{\hat{p}_r^2}{2m} + \frac{\hat{\mathbf{L}}^2}{2mr^2} + V(r), \quad \hat{p}_r \equiv -i\hbar\frac{1}{r}\frac{\partial}{\partial r}r \tag{5.4}$$

with the so-called radial impulse operator $\hat{p}_r$. Crucially, the Hamiltonian possesses rotational symmetry, i.e., it commutes with all rotation operators. This means that the Hamiltonian commutes with the operators $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ which commute themselves and thus we can have a simultaneous eigenstate for the three operators. The Schrödinger equation reads

$$\left[\frac{\hat{p}_r^2}{2m} + \frac{\hat{\mathbf{L}}^2}{2mr^2} + V(r)\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{5.5}$$

and we already know the simultaneous eigenstates of the angular momentum operators to be the spherical harmonics, hence, we can make the separation ansatz

$$\psi(\mathbf{r}) = R(r)\, Y_{lm}(\theta, \phi) \tag{5.6}$$

decomposing the radial and the angular part of the wave function. Inserting this into the Schrödinger equation gives us

$$\frac{r^2}{R(r)}\hat{p}_r^2 R(r) + 2mr^2\left[V(r) - E\right] = -\frac{1}{Y_{lm}(\theta,\phi)}\hat{\mathbf{L}}^2\, Y_{lm}(\theta,\phi) = const., \tag{5.7}$$

here sorted the equation into an radial and the angular part, where a purely radial dependent function on the left hand side equals a purely angular dependent function on the right side, meaning the can only be equal to a radial and angular independent constant (we used a similar logic to derive the time-independent Schrödinger equation in Sec. 1.4). We can bow separate the Schrödinger equation into two parts, first into one equation for the angular components

$$\hat{\mathbf{L}}^2\, Y_{lm}(\theta,\phi) = -const.\, Y_{lm}(\theta,\phi), \tag{5.8}$$

which corresponds to an eigenvalue problem which we already know, this is simply the eigenvalue equation for the orbital angular momentum operator

$$\hat{\mathbf{L}}^2\, Y_{lm}(\theta,\phi) = -\hbar^2 l(l+1)\, Y_{lm}(\theta,\phi), \tag{5.9}$$

which provides us with the constant, and thus the radial equation reads

$$\left[\frac{1}{2m}\hat{p}_r^2 + V_{\text{eff}}(r)\right] R(r) = E\, R(r), \qquad V_{\text{eff}}(r) = V(r) + \frac{\hbar^2}{2mr^2}l(l+1). \tag{5.10}$$

The effective potential $V_{\text{eff}}(r)$ is the sum of the centrifugal potential (second term) and the actual spherical potential we have not defined yet in more detail. The solution to the radial Schrödinger equation can be found once the latter has been determined. There exists an alternative form of the radial equation if one makes a substitution

$$u(r) = rR(r),, \qquad \frac{\partial^2}{\partial r^2}u(r) = r\left(\frac{\partial^2}{\partial^2 r} + \frac{2}{r}\frac{\partial}{\partial r}\right)R(r) = -r\frac{\hat{p}_r^2}{\hbar^2}R(r) \tag{5.11}$$

and hence multiplying the radial equation with r gives us

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + V_{\text{eff}}(r)\right] u(r) = E\, u(r), \tag{5.12}$$

which has a similar form as the one-dimensional Schrödinger equation, with the modification due to the centrifugal potential and the restriction that the radius is positive, i.e., $0 \leq r < \infty$. Due to the centrifugal potential the Schrödinger equation is parameterized by the quantum number l which takes the values $l = 0, 1, ...$ In some way this means that there is a different radial equation for each values of l and the energy eigenvalues for a given value of l may be either discrete or continuous. For example, for potential $V(r)$ which go to zero for $r \to \infty$ there will always be a continuous energy spectrum for $E > 0$, with $E = 0$ as a so-called continuum threshold. However, for the latter example there may or may not exist a discrete spectrum of of equivalently bound states, see .

5.1.2 The hydrogen atom [5, 10]

Now we want to turn to a concrete example for a particle in a central potential: the hydrogen atom. We will consider the most simple case where we only include the Coulomb potential and neglect corrections due to relativity and spin. Under these assumptions the effective potential becomes

$$V(r) = -\frac{1}{4\pi\epsilon_0}\,\frac{Ze^2}{r} \quad \Rightarrow \quad V_{\text{eff}}(r) = -\frac{1}{4\pi\epsilon_0}\,\frac{Ze^2}{r} + \frac{\hbar^2}{2mr^2}l(l+1) \tag{5.13}$$

with e as the electron, m denotes the reduced mass of the electron-nucleus system, and Z as the number of protons, which for the hydrogen is $Z = 1$ but can take different values for *hydrogen-like* atoms. The latter are atoms with a nucleus of some charge Z and a single electron, e.g. a singly ionized helium or double-ionized lithium. Now one prefers to work in atomic units which modifies the dimension of distance and energy scaling with

$$\begin{aligned} \text{Bohr radius: } a_0 &= \frac{4\pi\epsilon_0\hbar^2}{me^2} \approx 0.5 \times 10^{-10} m \\ \text{Rydberg constant: } R_0 &= \frac{\hbar^2}{2ma_0^2} = \frac{e^2}{8\pi\epsilon_0 a_0} \approx 13.6\ eV \end{aligned} \tag{5.14}$$

setting now $x = rZ/a_0$, and $E_R = E/(R_0 Z^2)$ we obtain for the radial equation

$$\left[-\frac{\partial^2}{\partial^2 x} - \frac{2}{x} + \frac{1}{x^2}l(l+1)\right] u(x) = E_R\, u(x), \tag{5.15}$$

with the dimensionless variables x and E_R. Bound states and a discrete spectrum will be obtained for $E_R < 0$, thus one introduced a new parameter

$$n = \frac{1}{\sqrt{-E_R}}, \tag{5.16}$$

and a modified radial variable

$$\rho = \frac{2x}{n}, \tag{5.17}$$

inserting this into the radial equation leaves us with the differential equation

$$\left[\frac{\partial^2}{\partial^2\rho} + \frac{n}{\rho} - \frac{1}{\rho^2}l(l+1) - \frac{1}{4}\right] u(\rho) = 0. \tag{5.18}$$

To find a solution we first consider the desired behaviour for large values of ρ, here we can reduce the equation to

$$\left[\frac{\partial^2}{\partial^2\rho} - \frac{1}{4}\right] u(r) = 0 \quad \Rightarrow \quad u(\rho) \propto e^{\pm\frac{\rho}{2}}, \tag{5.19}$$

where we discard the positive sign as it is inconsistent with a bound state. Analogously we can estimate the behavior for small values of ρ from the equation

$$\left[\frac{\partial^2}{\partial^2\rho} - \frac{1}{\rho^2}l(l+1)\right] u(\rho) = 0 \quad \Rightarrow \quad u(\rho) \propto a_1\,\rho^{-l} + a_2\,\rho^{l+1}, \tag{5.20}$$

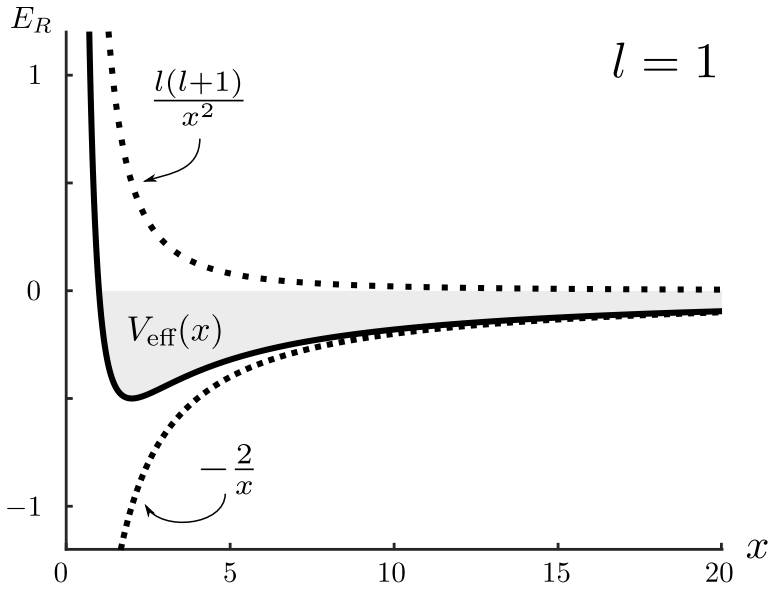

Figure 5.1: Effective potential for the hydrogen atom and $l = 1$. The positive centrifugal potential dominates for small x, while the negative Coulomb potential dominates for large x. The competition between the two produces a well in the effective potential, as shown in the figure, and we can expect bound states (grey shaded area). In fact, the well supports an infinite number of bound states, for all values of l. This is because of the long tail, which dies off only as $1/x$. One can see from the figure that there are also unbound states with energies in the range $0 < E_R < \infty$, taken from [10].

where we will set $a_1 = 0$ to ensure a vanishing at the origin. With all this we can make an ansatz for the whole radial function:

$$u(\rho) = e^{-\frac{\rho}{2}}\, \rho^{l+1}\, f(\rho), \tag{5.21}$$

which we can insert into the full radial differential equation and obtain after some algebra

$$\left[\rho \frac{\partial^2}{\partial^2 \rho} + [2l + 2 - \rho]\, \frac{\partial}{\partial \rho} + [n - l - 1]\right]\; f(\rho) = 0. \tag{5.22}$$

To solve this equation we could start with a power law ansatz

$$f(\rho) = \sum_{k=0}^{\infty} C_k \rho^k \quad \Rightarrow \quad C_{k+1} = \frac{[l + k + 1 - n]}{[2l + 2 + k]\,(k + 1)} C_k, \tag{5.23}$$

where the recursion relation was obtained by inserting the ansatz into the upper differential equation. For the situation that this power series does not terminate we would have a scaling of $C_{k+1} \backsimeq 1/kC_k$ for large values of k. The latter would lead to $f(\rho) \backsimeq e^{\rho}$, which would diverge for large values of ρ and overwhelm the factor $e^{-\rho/2}$ in the full solution. Such solutions are not normalizable and not acceptable as bound states. But we find that if n is an integer $\geq l + 1$, then the power series solution terminates. Meaning we require a maximal value of k for which all C_k with $k > k_{\max}$ vanish, this boils down to the condition

$$(l + 1 + k_{\max}) - n = 0, \tag{5.24}$$

in this case the wave function is normalizable, and we have a bound state solution. The integer n is called the principal quantum number. We see that there are an infinite number of bound states for each value of l. For given l, n takes on the values

$$n = l+1, l+2, ..., \quad n^2 = -\frac{1}{E_R} \quad \Rightarrow \quad E_n = -\frac{R_0 Z^2}{n^2}, \tag{5.25}$$

and we have a discrete energy spectrum which depends only on the principle quantum number $n \geq 1$, but not on the value of the angular momentum quantum number l, an effect typically referred to as accidental degeneracy. For given l we have a degeneracy $2l+1$, reflecting the different values m can take. As we have $n \geq l+1$ the degree of the degeneracy becomes

$$g_n = \sum_{l=0}^{n-1}(2l+1) = 2\sum_{l=1}^{n-1} l + n = n(n-1) + n = n^2. \tag{5.26}$$

The solutions $f(\rho)$ to the differential equation Eq.(5.22) are the so-called associated Laguerre polynomials, see ↗ for more details, and have the order of $n-l-1$. With all this we now have the solution for the radial equation as

$$u_{nl}(r) = \frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{(n+1)!}}\rho^{l+1}e^{-\frac{\rho}{2}}L^{2l+1}_{n-l-1}(\rho), \quad \rho = \frac{2Zr}{na_0}, \quad \int_0^\infty dr\; u_{nl}(r)u^*_{n'l'}(r) = \delta_{n',n}\delta_{l',l} \tag{5.27}$$

which are normalized, and incorporates the Laguerre polynomials

$$L^{2l+1}_{n-l-1}(\rho) = \sum_{k=0}^{n-l-1}(-1)^k \frac{(n+l)!}{k!(2l+1+k)!(n-l-1-k)!}\rho^k. \tag{5.28}$$

As seen at the top of Fig. 5.2, the angular momentum quantum numbers $l = 0, 1, 2, 3, ...$ are associated with a series of code letters, s, p, d, f , etc. After f the code letters follow the letters of the alphabet, g, h, i, etc. These have only a historical significance, but they are used universally in spectroscopic notation so one must know them and how to translate them into angular momentum values [10].

5.1.3 Fine structure in hydrogen [10]

In the last section we considered in principle a simplified model of an hydrogen atom, accounting only for the Coulomb potential, a simple electrostatic model. However, as usual, life is more complex and in this section we will briefly discuss the the effects of relativity and spin on the dynamics of the electron. Both these effects are of the same order of magnitude, and must be treated together in any realistic treatment of the atomic structure. As we have avoided a relativistic treatment in this lecture so far, we will treat the effects within a nominally nonrelativistic framework by the inclusion of extra terms in the Schrödinger equation. Meaning that we include the effects as a perturbation onto the electrostatic model, resulting in the so-called fine structure. The fine structure terms have the effect of enlarging the Hilbert space by the

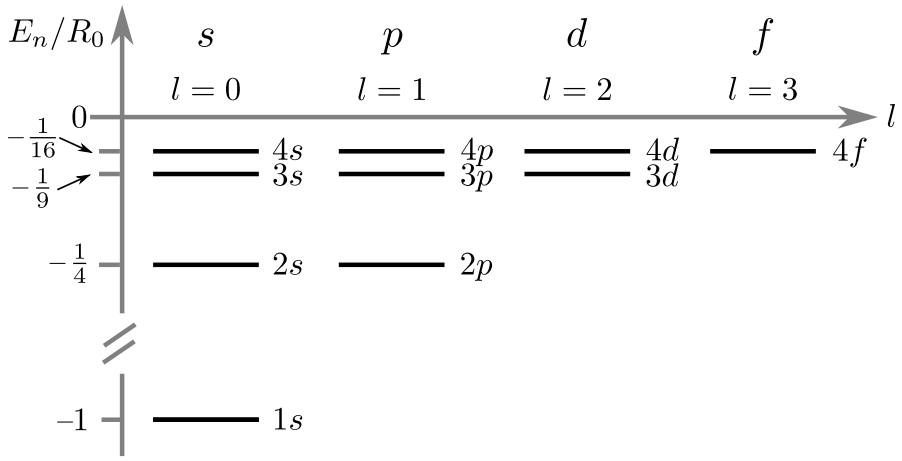

Figure 5.2: In central force problems the radial wave equation is parameterized by l, so we get a different set of energy levels for each l. It is often convenient to plot these on a diagram in which the levels for each value of l are arranged in a column labeled by l. Effectively this is a plot with energy on the vertical axis and angular momentum on the horizontal axis, from [10].

inclusion of the spin degrees of freedom, introducing new quantum numbers, and shifting and splitting the energy levels of the electrostatic model.
In principle one can describe the atomic systems accounting for relativistic effects using the so-called Dirac equation, the latter is solvable for the Coulomb-potential. However, as we did exclude relativity in this course, we start right away from an Hamiltonian which is obtained from a perturbation expansion of the Dirac equation around the non-relativistic limit. The resulting Hamiltonian for the atomic system reads

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{pert}}, \quad \hat{H}_{\text{pert}} = \hat{H}_{\text{Rkin}} + \hat{H}_{\text{SO}} + \hat{H}_D, \tag{5.29}$$

here, $\hat{H}_0$ denotes the Hamiltonian of the electrostatic model for which we derived the exact solution in the last section, i.e., an electron in the electrostatic potential

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + V(r), \quad V(r) = -\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r}. \tag{5.30}$$

The second term is the relativistic energy correction

$$\hat{H}_{\text{Rkin}} = -\frac{\hat{\mathbf{p}}^4}{8m^3c^2}, \tag{5.31}$$

obtained by simply expanding the energy of a relativistic particle (rest mass plus kinetic energy) for $\mathbf{p} \ll mc$. The third term is the so-called spin-orbit coupling

$$\hat{H}_{\text{SO}} = -\frac{e}{2m^2c^2}\frac{1}{r}\frac{\partial\Phi(r)}{\partial r}\,\hat{\mathbf{L}}\,\hat{\mathbf{S}}, \quad \hat{\mathbf{S}}_n = \frac{\hbar}{2}\hat{\sigma}_n \tag{5.32}$$

accounting for the interaction of the angular momentum and spin of the electron, here $\frac{\partial\Phi}{\partial r}$ arises from the gradient of the electric potential $\Phi(\mathbf{r})$ due to the positively charged nucleus. The spin-orbit coupling is a relativistic effect and can be understood as the electromagnetic

interaction between the electron's magnetic moment, its orbital motion, and the electrostatic field of the positively charged nucleus. One can as well describe this interaction starting from $\hat{H}_{\rm SO} \sim \mu_B \hat{\mathbf{S}}\ \mathbf{B}$ describing an electron in a magnetic field $\mathbf{B}$. In the case of hydrogen, this magnetic field can be understood as due to the positive charge of the nucleus, which from the electron's rest frame appears to be orbiting the electron, creating a current loop.
Th last term in our equation for the atomic system is the so-called Darwin term

$$\hat{H}_D = -\frac{e\hbar^2}{8m^2c^2}\Delta\Phi(r), \tag{5.33}$$

which is connected to the so-called *Zitterbewegung*, i.e., fast oscillator motion of a particle described by a relativistic wave equation. The latter will only provide a correction for the s-orbitals.
For hydrogen, or hydrogen-like atoms, we have the electric potential of the positively charged nucleus

$$\Phi(r) = \frac{1}{4\pi\epsilon_0}\ \frac{Ze}{r}, \quad \frac{1}{r}\frac{\partial\phi(r)}{\partial r} = -\frac{1}{4\pi\epsilon_0}\ \frac{Ze}{r^3}, \quad \Delta\Phi(r) = -\frac{Ze}{\epsilon_0}\ \delta(r) \tag{5.34}$$

which gives us for the spin-orbit coupling and Darwin term

$$\hat{H}_{\rm SO} = \frac{1}{4\pi\epsilon_0}\ \frac{Ze^2}{2m^2c^2}\ \frac{1}{r^3}\ \hat{\mathbf{L}}\ \hat{\mathbf{S}}, \quad \hat{H}_D = \frac{Ze^2\hbar^2}{8\epsilon_0 m^2c^2}\ \delta(r), \tag{5.35}$$

where we have absorbed all pre-factor into $\alpha_{{\rm SO},D}$. We now want to determine the correction due to the effects of relativity and spin. For this we need time-independent degenerate perturbation theory. We start with the unperturbed system, here we have the eigenenergies and states

$$E_n = -\frac{R_0 Z^2}{n^2}, \quad |nlm_l\rangle \otimes |sm_s\rangle = |nlm_l m_s\rangle\,, \tag{5.36}$$

where we do account for the spin of the electron, as the spin-orbit term explicitly involves the spin, so we must enlarge our Hilbert space to include the spin degrees of freedom. Without the spin the eigenvalues are n^2 degenerate, but with the two possible states for the spin the degeneracy doubles to $2n^2$. The states $|nlm_l m_s\rangle$ is the product state, and all of these states form a eigenbasis of the commuting observables $\left\{\hat{H}_0, \hat{\mathbf{L}}^2, \hat{L}_z, \hat{S}_z\right\}$ corresponding to the quantum numbers $\{n, l, m_l, m_s\}$. We suppress the quantum number for the spin as it is constant, i.e., $s = 1/2$ and $\hat{\mathbf{S}}^2 = \hbar^2 s(s+1) = 3/4\hbar^2$.
Since the unperturbed energy levels are degenerate, we must think in terms of degenerate perturbation theory, in which the shifts in the energy levels are the eigenvalues of the matrix of the perturbing Hamiltonian in the eigenspaces of the unperturbed system. This means that we need the perturbation matrix elements

$$\langle nlm_l m_s|\, \hat{H}_{\rm pert}\, |nl'm_l'm_s'\rangle\,, \tag{5.37}$$

the index n is the same on both sides of the matrix element because it labels an unperturbed eigenspace, while the remaining indices are allowed to be different, since these label the basis states inside the unperturbed eigenspace. It is advantageous for perturbation theory to use

an eigenbasis of a complete set of commuting observables in which as many as possible of the observables commute with the perturbing Hamiltonian. So we look for a simultaneous eigenbasis involving as well the joint angular momentum operator $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. A set of commuting hermitian operators is $\left\{\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2, \hat{\mathbf{J}}^2, \hat{J}_z\right\}$ with quantum numbers $\{l, s, j, m_j\}$. So we can define a new joint basis which we can express as an expansion in the uncoupled basis

$$|nljm_j\rangle = \sum_{m_l,m_s} |nlm_lm_s\rangle \langle nlm_lm_s|nljm_j\rangle = \sum_{m_l,m_s} C^{jm_j}_{lm_l,\frac{1}{2},m_s} |nlm_lm_s\rangle , \tag{5.38}$$

with the already familiar Clebsch-Gordon coefficients. Important is, that all fine structure terms are commuting with the chosen set of commuting observables. This will make the evaluation of the perturbation matrix much simpler as the matrix will be diagonal right away, and thus providing us with it's eigenvalues without the need of diagonalization. To see this we start with the relativistic energy correction term, which is a purely orbital operator, and commutes for example with $\hat{J}_z$ and we have

$$\langle nljm_j| \hat{J}_z\hat{H}_{\text{Rkin}} - \hat{H}_{\text{Rkin}}\hat{J}_z \left|nl'j'm'_j\right\rangle = 0 \tag{5.39}$$

and thus operating with $\hat{J}_z$ to the left and the right we obtain

$$(m_j - m'_j) \langle nljm_j| \hat{H}_{\text{Rkin}} \left|nl'j'm'_j\right\rangle = 0 \tag{5.40}$$

so either we have $m_j = m'_j$ or the matrix element vanishes. This holds analogously for all other operators and hence we need to only evaluate the diagonal elements

$$\langle nljm_j| \hat{H}_{\text{Rkin}} |nljm_j\rangle = \sum_{m_l,m'_l} C^{jm_j*}_{lm_l,\frac{1}{2},m_s} C^{jm_j}_{lm'_l,\frac{1}{2},m_s} \langle nlm_l| \hat{H}_{\text{Rkin}} |nlm'_l\rangle , \tag{5.41}$$

where we already used that the relativistic energy correction term is independent of the spin. Moreover, the relativistic energy correction term involves a scalar and the Wigner-Eckart theorem for a scalar, i.e., a rank-0 tensor operator is

$$\langle j_1m_1| T^{(0)}_0 |j_2m_2\rangle = \langle j_2m_200|(j_20)j_1m_1\rangle \ \left\langle j_1||T^{(0)}||j_2\right\rangle = \delta_{j_1,j_2}\delta_{m_1,m_2} \ \left\langle j_1||T^{(j)}||j_1\right\rangle , \tag{5.42}$$

which follows from the conditions for angular momentum addition

$$|j - j_2| \leq j_1 \leq j + j_2 \ \xrightarrow{j=0} \ j_1 = j_2, \quad m + m_2 = m_1 \ \xrightarrow{m=0} \ m_1 = m_2. \tag{5.43}$$

This means that the Wigner-Eckart theorem for scalar operators is equal to δ_{m_1,m_2} times a quantity that is independent of magnetic quantum numbers, meaning we can freely choose a convenient set of magnetic quantum numbers. Choosing $m_l = m'_l = 0$ we have

$$\langle nljm_j| \hat{H}_{\text{Rkin}} |nljm_j\rangle = C^{jm_j*}_{l0,\frac{1}{2},m_s} C^{jm_j}_{l0,\frac{1}{2},m_s} \langle nl0| \hat{H}_{\text{Rkin}} |nl0\rangle = \langle nl0| \hat{H}_{\text{Rkin}} |nl0\rangle , \tag{5.44}$$

where we used in the last step that the orthogonality of the Clebsch-Gordon coefficients. So we see that the resulting correction term will only depend on the quantum numbers n and l.

The remaining step is to evaluate the matrix element

$$\begin{aligned} E^{(1)}_{\text{Rkin}} = \langle nl0|\, \hat{H}_{\text{Rkin}}\, |nl0\rangle &= -\frac{1}{2mc^2}\, \langle nl0|\, \frac{\hat{\mathbf{p}}^4}{4m^2}\, |nl0\rangle = -\frac{1}{2mc^2}\, \langle nl0|\, \left(\hat{H}_0 - V(r)\right)^2 |nl0\rangle \\ &= -\frac{1}{2mc^2}\, \langle nl0|\, \left(\hat{H}_0^2 - V(r)\hat{H}_0 - \hat{H}_0 V(r) + V(r)^2\right) |nl0\rangle \\ &= -\frac{\alpha^2}{4R_0}\, \langle nl0|\, \left(E_n^2 - 2E_n V(r) + V^2(r)\right) |nl0\rangle\,, \quad \alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137}, \end{aligned} \tag{5.45}$$

for the final steps we require the expectation values for the potential terms, we will omit the detailed calculation and just give the results

$$\langle V(r)\rangle = -\frac{1}{4\pi\epsilon_0}\, \frac{Z^2e^2}{a_0 n^2}, \quad \left\langle V^2(r)\right\rangle = \frac{1}{16\pi^2\epsilon_0^2}\, \frac{Z^4e^4}{a_0^2 n^3 (l+\frac{1}{2})}, \tag{5.46}$$

which gives us in total for the energy corrections

$$E^{(1)}_{\text{Rkin}} = -\frac{\alpha^2}{4R_0}\left(E_n^2 + 4E_n \frac{Z^2R_0}{n^2} + \frac{4Z^4R_0^2}{n^3(l+\frac{1}{2})}\right) = -\frac{\alpha^2 E_n^2}{R_0}\left(\frac{n}{(l+\frac{1}{2})} - \frac{3}{4}\right). \tag{5.47}$$

Next we can consider the corrections due to the spin-orbit coupling, i.e., the correction due to the term

$$\hat{H}_{\text{SO}} = \frac{1}{4\pi\epsilon_0}\, \frac{Ze^2}{2m^2c^2}\, \frac{1}{r^3}\, \hat{\mathbf{L}}\, \hat{\mathbf{S}} = \frac{1}{8\pi\epsilon_0}\, \frac{Ze^2}{2m^2c^2}\, \frac{1}{r^3}\, \left(\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2\right), \tag{5.48}$$

as before we evaluate this term in the simultaneous eigenbasis

$$E^{(1)}_{\text{SO}} = \langle nljm_j|\, \hat{H}_{\text{SO}}\, |nljm_j\rangle = \frac{1}{8\pi\epsilon_0}\, \frac{Ze^2\hbar^2}{2m^2c^2}\, [j(j+1) - l(l+1) - s(s+1)]\, \langle nl0|\, \frac{1}{r^3}\, |nl0\rangle \tag{5.49}$$

where we have used the eigenvalues of the respective operators and as the remaining matrix element is again the expectation value of the a scalar (and thus a rank-zero tensor), we could use again Wigner-Eckart theorem as before. We know that $s = 1/2$ and thus we have $j = l+1/2$ and $j = l-1/2$. For the case were $l = 0$ we have $j = 1/2$ and the energy correction term vanishes for this case. What remains is to determine the expectation value

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3 n^3}\, \frac{1}{l(l+\frac{1}{2})(l+1)}, \tag{5.50}$$

which would diverge for $l = 0$, however this is exactly the point were the expectation value of $\hat{L}\hat{S}$ vanishes. Thus the energy correction due to the spin-orbit interaction for $l \neq 0$ can be summarized as

$$E^{(1)}_{\text{SO}} = \frac{\alpha^2 E_n^2}{2R_0}\, \frac{n}{l(l+\frac{1}{2})(l+1)}\, \begin{cases} l, & j = l + \frac{1}{2} \\ -[l+1], & j = l - \frac{1}{2} \end{cases}. \tag{5.51}$$

Now remains only the Darwin correction

$$\hat{H}_D = \frac{Ze^2\hbar^2}{8\epsilon_0 m^2 c^2}\,\delta(r), \tag{5.52}$$

here we have again a scalar operator and the energy correction reduces to the element

$$E_D^{(1)} = \langle nl0|\,\hat{H}_{\rm SO}\,|nl0\rangle = \frac{Ze^2\hbar^2}{8\epsilon_0 m^2 c^2}\,\langle nl0|\,\delta(r)\,|nl0\rangle = \frac{Ze^2\hbar^2}{8\epsilon_0 m^2 c^2}|\psi_{n,l}(0)|^2, \tag{5.53}$$

thus we need the wavefunction at the origin, which is given as

$$\psi_{n,l}(0) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{na_0}\right)^{\frac{3}{2}}\,\delta_{l,0} \tag{5.54}$$

The Darwin term affects only the s orbitals. This is because the wave function of an electron with $l > 0$ vanishes at the origin, and the delta function has no effect. Thus, we obtain the final correction due to the Darwin term

$$E_D^{(1)} = \frac{\alpha^2 E_n^2}{R_0}\,n\,\delta_{l,0}. \tag{5.55}$$

Let us first evaluate the case for the fine structure were $l = 0$ and we only have the Darwin term and the relativistic energy correction:

$$l = 0: \qquad E_{\rm FS}^{(1)} = -\frac{\alpha^2 E_n^2}{R_0} n\left[\left(2 - \frac{3}{4n}\right) - 1\right] = -\frac{\alpha^2 E_n^2}{R_0} n\left[1 - \frac{3}{4n}\right]. \tag{5.56}$$

Considering now the case $l \neq 0$ we find a very similar expression for both values of the j, i.e., $j = l \pm 1/2$. The expression becomes

$$E_{\rm FS}^{(1)} = -\frac{\alpha^2 E_n^2}{R_0}\,n\,\left[\frac{1}{\left(j + \frac{1}{2}\right)} - \frac{3}{4n}\right], \tag{5.57}$$

which does not depend on the angular quantum number l. This correction term includes as well the case $l = 0$! As we know that for $l = 0$ we can only have $j = \frac{1}{2}$ and thus the upper expression for the energy corrections still holds. So in total, the resulting expression for the fine structure is rather simple, it depends only on the quantum numbers j and n. Meaning that for every fixed principal quantum number n and joint angular momentum quantum number j exist different $l = j \pm 1/2$ where $l = 0, 1, ..., n-1$.
As mentioned above the hydrogen atom can also be solved accounting for the full nonrelativistic model, i.e., the Dirac equation is solvable. Interestingly, the exact Dirac energy levels also only depend on j and n. Moreover The energy shifts in the fine structure are negative for all values of n and j, so fine structure effects depress all energy levels. However, smaller values of j are more strongly depressed, so the total energy (unperturbed plus fine structure) is an increasing function of j. Since the unperturbed levels did not depend on j, fine structure effects have partially resolved the degeneracy in the unperturbed levels. The fact that the levels still do not depend on l means that the hydrogen atom, even including relativistic corrections, still has some extra symmetry that goes beyond rotational invariance (in particular, there is still a degeneracy between states of opposite parity), from [10].

5.2 Particles in fields

5.2.1 Precession of a spin in a magnetic field [5]

We already discussed that the spin was discovered when silver particles were send through an inhomogeneous magnetic field. We now want to briefly discuss what happens when an electron is send through a magnetic field **B**. We will focus here only on the spin degree of freedom and ignore any spacial dependence. The effective spin Hamiltonian becomes

$$\hat{H} = \mu_B \hat{\sigma} \mathbf{B}, \tag{5.58}$$

with the already introduced Bohr magneton μ_B. We assume now that the magnetic field is pointing into the z-direction, for this case the Hamiltonian simplifies to

$$\hat{H} = \mu_B B_z \hat{\sigma}_z = \mu_B B_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{5.59}$$

which allows us to read of the eigenvalues of the system

$$E_\pm = \pm \mu_B B_z \equiv \pm \frac{1}{2} \hbar \omega_L, \quad \omega_L = \frac{2\mu_B B_z}{\hbar} = \frac{eB_z}{mc}, \tag{5.60}$$

with the so-called Larmor frequency ω_L. To account for the spin as an intrinsic property of the electron, we had defined a so-called spinor

$$|\psi(\mathbf{x})\rangle = \begin{pmatrix} \psi_\uparrow(\mathbf{x}) \\ \psi_\downarrow(\mathbf{x}) \end{pmatrix} \tag{5.61}$$

to account for this extra degree of freedom on top of the spacial degrees. Neglecting the spacial dependence we obtain the following equations of motion just for the spin degrees of freedom

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi_\uparrow &= + \mu_B B_z \psi_\uparrow = + \frac{\hbar}{2} \omega_L \psi_\uparrow, \\ i\hbar \frac{\partial}{\partial t} \psi_\downarrow &= - \mu_B B_z \psi_\downarrow = - \frac{\hbar}{2} \omega_L \psi_\downarrow, \end{aligned} \tag{5.62}$$

which we can straightforwardly solve

$$\psi_\uparrow(t) = \psi_\uparrow(0) e^{+\frac{i}{2}\omega_L t}, \quad \psi_\downarrow(t) = \psi_\downarrow(0) e^{-\frac{i}{2}\omega_L t}, \tag{5.63}$$

hence we observe an oscillation of the spinor components. This will translate into a precession of the spin with the Larmor frequency. To see this we can calculate an expectation value of one of our spin components

$$\begin{aligned} \left\langle \hat{S}_x \right\rangle =& \begin{pmatrix} \psi_\uparrow(t) \\ \psi_\downarrow(t) \end{pmatrix}^\dagger \hat{S}_x \begin{pmatrix} \psi_\uparrow(t) \\ \psi_\downarrow(t) \end{pmatrix} = \frac{\hbar}{2} \left(\psi_\uparrow^*(t), \psi_\downarrow^*(t) \right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi_\uparrow(t) \\ \psi_\downarrow(t) \end{pmatrix} \\ =& \frac{\hbar}{2} \left(\psi_\uparrow^*(t) \psi_\downarrow(t) + \psi_\downarrow^*(t) \psi_\uparrow(t) \right) = \frac{\hbar}{2} \left(\psi_\uparrow^*(0) \psi_\downarrow(0) e^{-i\omega_L t} + \psi_\downarrow^*(0) \psi_\uparrow(0) e^{+i\omega_L t} \right). \end{aligned} \tag{5.64}$$

To ensure normalization of the state we need to have the usual condition fulfilled

$$|\psi_\uparrow(0)|^2 + |\psi_\downarrow(0)|^2 = 1, \tag{5.65}$$

so we can choose a valid ansatz

$$\psi_\uparrow(0) = \cos\alpha\, e^{i\varphi}, \quad \psi_\downarrow(0) = \sin\alpha, \tag{5.66}$$

inserting this into the expression for the expectation value we have

$$\left\langle \hat{S}_x \right\rangle = \frac{\hbar}{2}\sin\alpha\cos\alpha\left(e^{-i\varphi}e^{-i\omega_L t} + e^{i\varphi}e^{+i\omega_L t}\right) = \mathcal{A}\cos(\omega_L t + \varphi), \tag{5.67}$$

with $\mathcal{A} = \hbar\sin\alpha\cos\alpha$. And in a similar fashion we have

$$\left\langle \hat{S}_y \right\rangle = \mathcal{A}\sin(\omega_L t + \varphi). \tag{5.68}$$

While for the z-components we obtain

$$\left\langle \hat{S}_z \right\rangle = \frac{\hbar}{2}\left(|\psi_\uparrow(t)|^2 - |\psi_\downarrow(t)|^2\right) = \frac{\hbar}{2}\left(|\psi_\uparrow(0)|^2 - |\psi_\downarrow(0)|^2\right) = \frac{\hbar}{2}\cos(2\alpha). \tag{5.69}$$

And thus the z-component is time-independent! Which is exactly the direction of the magnetic field, i.e., in a constant magnetic field, the spin precesses about the field direction at the Larmor frequency.

5.2.2 An atom in a magnetic field

We consider a hydrogen atom inside a homogeneous magnetic field. The Hamiltonian is obtained by the correspondence principle, i.e., via quantization of the the classical Hamiltonian for a particle in an external magnetic field $\mathbf{B}$. In addition we include the spin degrees of freedom. In total, we obtain

$$\hat{H} = \frac{1}{2m}\left(\hat{\mathbf{p}} + \frac{e}{c}\mathbf{A}(\hat{\mathbf{r}})\right)^2 + V(r) + \frac{e}{mc}\hat{\mathbf{S}}\mathbf{B} \tag{5.70}$$

with $V(r)$ as the central force potential and $\mathbf{A}(\hat{\mathbf{r}})$ is the vector potential. We neglect here the fine structure contribution for simplicity. We know that the magnetic and electric fields are unaffected if there is a gauge transformation and that we are free to choose a suitable gauge, here we take

$$\mathbf{A}(\hat{\mathbf{r}}) = \frac{1}{2}\mathbf{B}\times\hat{\mathbf{r}} \tag{5.71}$$

which corresponds to the Coulomb gauge such that $\nabla\mathbf{A} = 0$ and consequently we have

$$\mathbf{A}\cdot\hat{\mathbf{p}} = \hat{\mathbf{p}}\cdot\mathbf{A}, \tag{5.72}$$

so the cross terms in the expansion of the kinetic energy can be written in either order, i.e., the momentum operator and the vector potential commute. We can now expand the Hamiltonian

$$\hat{H} = \hat{H}_0 + \frac{e}{mc}\mathbf{A}(\hat{\mathbf{r}})\hat{\mathbf{p}} + \frac{e^2}{2mc^2}\mathbf{A}(\hat{\mathbf{r}})^2 + \frac{e}{mc}\hat{\mathbf{S}}\mathbf{B}, \tag{5.73}$$

with $\hat{H}_0$ denoting our electrostatic model without the magnetic field, as before we can consider this as the unperturbed system. With our choice of gauge we can rewrite the cross term as

$$\mathbf{A}(\hat{\mathbf{r}})\hat{\mathbf{p}} = \frac{1}{2}\left(\mathbf{B}\times\hat{\mathbf{r}}\right)\cdot\hat{\mathbf{p}} = \frac{1}{2}\mathbf{B}\left(\hat{\mathbf{r}}\times\hat{\mathbf{p}}\right) = \frac{1}{2}\mathbf{B}\hat{\mathbf{L}}, \tag{5.74}$$

with this our Hamiltonian becomes

$$\hat{H} = \hat{H}_0 + \frac{e}{2mc}\mathbf{B}\left[\hat{\mathbf{L}} + 2\hat{\mathbf{S}}\right] + \frac{e^2}{8mc^2}\left(\mathbf{B}\times\hat{\mathbf{r}}\right)^2, \tag{5.75}$$

the middle term is called the Zeemann Hamiltonian, the last term is small compared to the other terms and is in general dropped. Hence, we now want to include the Zeemann term as an perturbation on the system. We can do this again using degenerate perturbation theory, as before we want to finds the right basis in which we work. A good eigenbasis is found when the perturbation Hamiltonian commutes with the associated observables. For this situation the simultaneous eigenbasis used for our fine structure calculation is not optimal, as the Zeemann term will not commute with $\hat{\mathbf{J}}$. We rather work with the product basis, since the operators $\{L^2, L_z, S_z\}$ of the complete set that define the uncoupled basis all commute with the Zeemann term. We are going to work with a uniform magnetic field $\mathbf{B} = B_z\hat{z}$ and obtain for the energy corrections due to the magnetic field

$$\Delta E = \frac{e}{2mc}B_z\left\langle nlm_lm_s\right|\left[\hat{L}_z + 2\hat{S}_z\right]\left|nlm_lm_s\right\rangle = \mu_B B_z(m_l + 2m_s), \tag{5.76}$$

hence, we obtain a lifting of the degeneracy towards the quantum numbers $m_{l,s}$, the reason why this quantum number is also called the magnetic quantum number.

Chapter 6

Applications & fundamentals

6.1 Measurement principles

6.1.1 The measurement chain [15]

In classical mechanics an ideal measurement can determine all values of the system variables with full certainty. One can obtain so-called complete knowledge over a system and one can determine the outcome of future ideal measurements. Clearly, in a realistic experimental setup this holds now longer true as for example the noise/error of a measurement apparatus comes into play. However, in principle measurement and probability do not play a major role in the formulation of classical mechanics. In contrast, in quantum measurements one can obtain for some systems maximal knowledge, but it will not determine the outcome of subsequent measurements. One can only give probability distribution for measurement outcomes. Hence, a state of maximal knowledge is not a state of complete knowledge as in classical mechanics.
The standard understanding of measurements in the quantum regime involves the collapse of the wave function. Meaning that the state of a system, which is before the measurement described by an superposition of eigenstates, is projects/reduces into a single eigenstate by the measurement. And after the measurement the system is considered to be in this single eigenstate. But this so-called projective measurement is not the end of the story - and can even be considered to be inadequate for the description of quantum measurements. One fundamental reason here is that in an experiment one does not really measure the actual system of interest. For example in a quantum optical setup a quantum system (for example an atom) interacts with the electromagnetic field, and the latter is the measured quantity. In other words the observer performs a measurement of the environment and not on the state of the system itself.
This raises another question, about when the measurement is actually made. Sticking with the upper example of optical read-out, the photons leaving the system are injected in a photo detector, which in turn produces an electric current, which is then displayed on a screen and processed by the observer's eye. Thus we actually have a chain of systems, which is also called a von Neumann chain. One has to decide where to cut this chain and apply the projective postulate (Heisenberg's cut). For the discussed example, the projective measurement of the electromagnetic field seems to be reasonable. The photons contain the information of the system of interest, i.e., their respective quantum states are correlated, thus with a measurement

of the photons we gather knowledge about the system (without measuring it directly). Crucial is here as well, that the photons entering the photomultiplier are not longer interacting with the system of interest.
The arguments raised so far, show that we need a more general formulation of a quantum measurement, which for example accounts for the dynamics of a so-called meter, i.e., a second quantum system which is measured projectively and coupled to the system of interest. The following section are partly following the online lecture series by F. Marquardt on Quantum foundations ↗ .

6.1.2 Example: Stern-Gerlach experiment

We want to revisit the Stern-Gerlach experiment which we briefly discussed when we introduced the spin angular momentum operator. In this experiment a beam of spin 1/2-particles is injected in x-direction into an apparatus with a strong magnetic field gradient transverse to the beam. The field exerts a deflecting force on the magnetic moments of the moving particles, proportional to the magnetic field gradient and dependent on the spin of the particle. For an inhomogeneous magnetic field in z-direction we have the force $F_\pm = \pm\mu_B \partial B/\partial z$ acting on the particle, and if it was a spin-up (spin-down) particle it deflects the particle upwards (downwards). The Hamiltonian for the system becomes

$$\hat{H} = \frac{\hat{p}_z^2}{2m} - \mu_B \hat{\sigma}_z B_z(z) = \frac{\hat{p}_z^2}{2m} - \mu_B \hat{\sigma}_z \left(B_z + \hat{z}\frac{\partial B}{\partial z} \right), \tag{6.1}$$

where we assumed that the spacial dependent magnetic field has a linear dependence on the position. In addition we focus only on the motion of the particle in z-direction, as this is the only component influenced by the magnetic field. The latter interaction has two components, one constant term which we already analysed in the last chapter, associated with the precession of the spin, the second part is the crucial one as it describes the force acting on the particle depending on it's spin state. To determine the evolution of the system's state, we first determine the initial state, i.e., before the interaction is active. Here we can assume that we have a product state of the wavefunction in real space and a spin state which is a simple superposition of the up and down state:

$$|\Psi(t=0)\rangle = (\alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle) \otimes |\phi(\mathbf{x})\rangle\,, \tag{6.2}$$

to note is, that as this is a product state we have not an entangled state. Luckily, the Hamiltonian only contains the $\hat{\sigma}_z$ operator, and thus exactly the operator we want to measure, a so-called QND measurement. The time-evolution of the state becomes

$$\begin{aligned} |\Psi(t)\rangle =& e^{-\frac{i}{\hbar}\hat{H}t}\,|\Psi(0)\rangle = \alpha e^{-\frac{i}{\hbar}\hat{H}t}\,(|\uparrow\rangle \otimes |\phi(\mathbf{x})\rangle) + \beta e^{-\frac{i}{\hbar}\hat{H}t}\,(|\downarrow\rangle \otimes |\phi(\mathbf{x})\rangle) \\ =& \alpha\,|\uparrow\rangle \otimes e^{-\frac{i}{\hbar}\hat{H}_\uparrow t}\,|\phi(\mathbf{x})\rangle + \beta\,|\downarrow\rangle \otimes e^{-\frac{i}{\hbar}\hat{H}_\downarrow t}\,|\phi(\mathbf{x})\rangle \end{aligned} \tag{6.3}$$

here we used that the spin-states are the eigenstates of the $\hat{\sigma}_z$ and the spacial wavefunctions evolve under the Hamiltonian

$$\hat{H}_{\uparrow,\downarrow} = \frac{\hat{p}_z^2}{2m} \mp \mu_B \left(B_z + \hat{z}\frac{\partial B}{\partial z} \right), \tag{6.4}$$

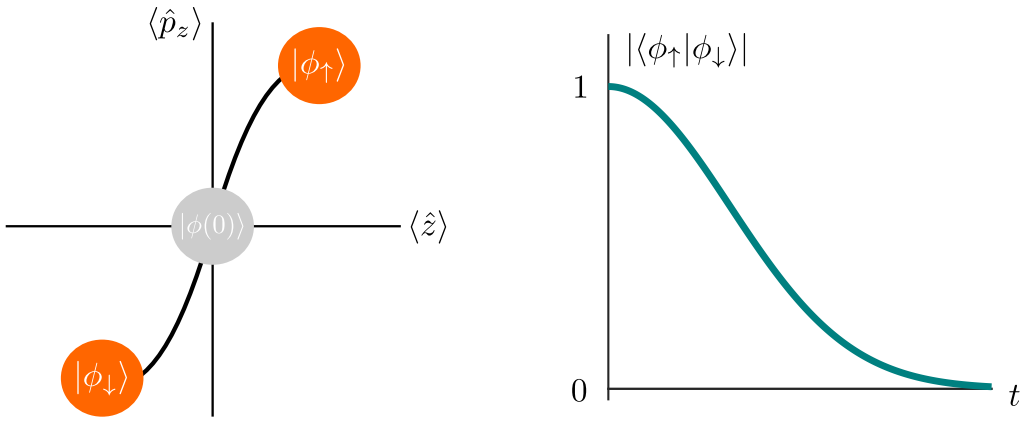

Figure 6.1: LEFT: phase space representation of the two pointer states, the initial states evolves upwards/downwards depending on the orientation of the spin. RIGHT: Overlap of the pointer states decreases with time, i.e., the states become orthogonal.

this means we can write the state as

$$|\Psi(t)\rangle = \alpha\,|\uparrow\rangle \otimes |\phi_\uparrow(t)\rangle + \beta\,|\downarrow\rangle \otimes |\phi_\downarrow(t)\rangle\,, \tag{6.5}$$

this is now an entangled state! As it cannot be written in the basis of the product states as before. The states $\left|\phi_{\uparrow/\downarrow}(t)\right\rangle$ are so-called pointer states, as they describe the spacial wavefunction evolving under depending on the orientation of the spin. The position is here considered as part of the measurement apparatus as in the end the position of the particle is collected and will provide the information if the particle was in spin up-or a spin down state (they 'point to the measurement result').
We will omit deriving the explicit expressions for the pointer states here, however, assuming that the initial spacial state was a wavepacket described by a Gaussian profile it will keep this form and is displaced in phase-space depending on the orientation of the spin as sketched in the phase space plot in Fig. 6.1. Over time the wavepackets get further displaced from the origin, and the distance between the wavefunction for spin-up and spin-up grows. The latter means that their overlap $|\,\langle\phi_\uparrow|\phi_\downarrow\rangle\,|$ becomes smaller and smaller, see sketch on the right in Fig. 6.1. Once the overlap is zero both states become clearly distinguishable and one can read-off the respective spin state. To note is that once the particle leaves the magnetic field region, the overlap will not change anymore. In this situation the two states evolve under the same Hamiltonian and thus under the same time-evolution operator, i.e., assuming the magnetic region ends at t_0 we find for the overlap

$$|\,\langle\phi_\uparrow(t)|\phi_\downarrow(t)\rangle\,| = \left\langle \phi_\uparrow(t_0)|\hat{U}^\dagger(t-t_0)\hat{U}(t-t_0)|\phi_\downarrow(t_0)\right\rangle = \langle\phi_\uparrow(t_0)|\phi_\downarrow(t_0)\rangle\,, \tag{6.6}$$

where we used the properties of a unitary operator. From this we see that the overlap stays constant if the measurement would end here. However, what happened if we do not end it here, but instead try to reverse the measurement? Such protocols are indeed possible for certain experiments and are called the quantum eraser. For the Stern-Gerlach experiment such a protocol goes as follows. After leaving the first apparatus the particle enters another apparatus twice the length of the first apparatus and with a gradient in the opposite direction, meaning

that the force is reversed. The particle will be slow down till the middle of the apparatus and then get accelerated into the opposite direction. Once the particle leaves this second apparatus it will continue to have a velocity into the $z-$ direction and hence the entanglement is still alive. To fully remove the effect of the experiment one has to use another apparatus similar to the first one to remove it. Only after this full stage one has fully reversed or erased the measurement. This is the quantum eraser experiment, which undoes the entanglement and leaves the system back into a product state. So depending on the intend of the experimentalist, it is not clear when the measurement ended.

6.1.3 Irreversibility and Heisenberg's cut

So we have to introduce a crucial aspect of measurements, and this is irreversibility. Defining that the measurement ends once it is practically impossible to reverse it. As an example we can take the detecting of a photon by a photomultiplier, the single photon hits a metal plate and an electron is ejected from the plate. The latter process is already hard to reverse, however, a single electron is as well hard to detect. We need to amplify the signal, and for the example of the photomultiplier, the electron is accelerated to a strongly positively charged plate where it ejects more electrons. The latter process is repeated multiple times, generating a detectable electronic current. Now the process can clearly be seen as no longer reversible. In a way we can understand the point of irreversibility as a point where we have an enormous and untrackable amount of additional degrees of freedom.
We want to quantify a bit more when are we allowed to make the quantum to classical cut and say that the measurement has been performed, i.e. the wave function has collapsed. In the initial formulation of the measurement principle (so-called Copenhagen interpretation), the collapse on the wave function was simply postulated, something we want to avoid here. In other words, when can we safely change our description from a quantum mechanical one to a classical one. Sticking with Stern-Gerlach example, we ask the question when we are allow to replace the superposition

$$|\Psi\rangle = \alpha\,|\uparrow\rangle \otimes |\phi_\uparrow\rangle + \beta\,|\downarrow\rangle \otimes |\phi_\downarrow\rangle \equiv \alpha\,|\psi_\uparrow\rangle + \beta\,|\psi_\downarrow\rangle\,, \tag{6.7}$$

with an *incoherent mixture* and assume that obtaining the state $|\psi_\uparrow\rangle$ with probability α^2 and the state $|\psi_\downarrow\rangle$ with probability β^2. While we still want to get correct predictions for all future measurements of all relevant observables. Meaning that for all relevant observables $\hat{O}$ we have

$$\left\langle \Psi(t)|\hat{O}|\Psi(t)\right\rangle \stackrel{!}{=} |\alpha|^2 \left\langle \psi_\uparrow(t)|\hat{O}|\psi_\uparrow(t)\right\rangle + |\beta|^2 \left\langle \psi_\downarrow(t)|\hat{O}|\psi_\downarrow(t)\right\rangle, \tag{6.8}$$

for the time t and any future time. Crucially, this expression is an approximation as we have neglected the interference term which would appear for a superposition:

$$\alpha\beta^* \left\langle \psi_\downarrow\right| \hat{O}\left|\psi_\uparrow\right\rangle + \alpha^*\beta \left\langle \psi_\uparrow\right| \hat{O}\left|\psi_\downarrow\right\rangle \quad \to \quad 0, \tag{6.9}$$

i.e., we have no overlap of the pointer states and consider them orthogonal. Here we can we with most certainty know that we are in either states. Hence, when we reach that point in time we can make the cut. To note is that some observables are excluded, e.g. for highly non-local many particle observables that act on system and measurement apparatus one cannot

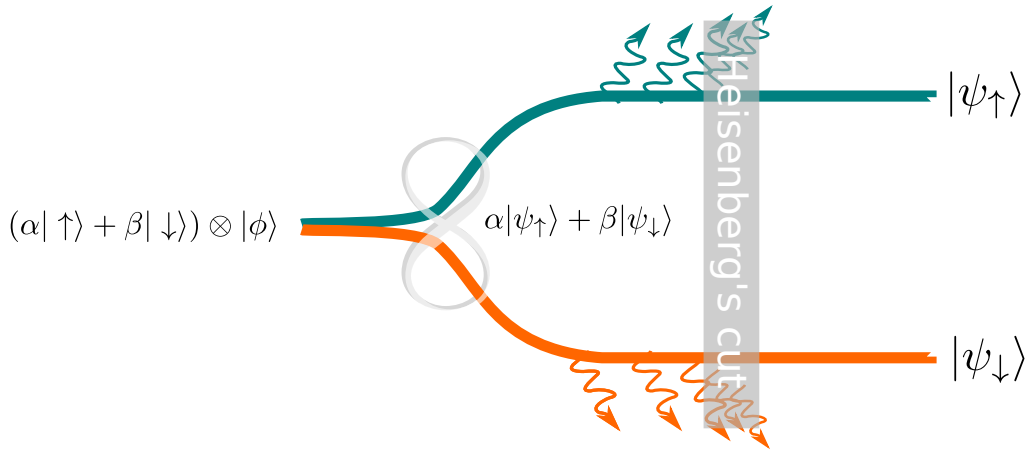

Figure 6.2: Evolution from a product state of system and apparatus into an entangled superposition state. We identify each branch with a single wave function, and the overlap of the wavefunctions of each branch decreases, i.e., the single states become orthogonal. As this can be in principle be reversed (quantum eraser), the interaction with many degrees of freedom allows us to make the cut from quantum to a classical description, i.e. we can state that a measurement has been performed and the wave function has collapsed. Important is that the cut is not a physical event, but rather a change in description. After the cut we have two possible solutions with associated probabilities $|\alpha|^2$ and $|\beta|^2$

safely make a cut, see ↗ for further discussion. Another crucial aspect is that the pointer states remain orthogonal, and one case were this is ensured is when the system-apparatus interaction is turned off. We can again taker the discussion for the Stern-Gerlach experiment as an example. When the particle leaves the detector, and the overlap of the wavefunction has reduced to zero, the overlap also stays zero. However, for the case that the system and-apparatus interaction cannot be turned off, one has to exclude certain observables as mentioned above and in addition, ensure that there is no recurrence i.e., irreversibility is in place (many degrees of freedom, and one cannot undo the entanglement).

6.1.4 Projective measurement, systems and meters

Next, we want to introduce some formal notions of how to technically describe measurements. Starting the with the standard von Neumann measurement, which can be described via

$$\hat{\Pi}_n = |n\rangle \langle n| , \qquad \hat{\Pi}_n \hat{\Pi}_{n'} = \delta_{n,n'} \hat{\Pi}_n, \tag{6.10}$$

which projects the system state onto the basis state $|n\rangle$. The second condition tells us that the projection operators are orthonormal. After the measurement has been performed the probability to find the system in state $|n\rangle$ is

$$p(n) = \langle n|\hat{\rho}|n\rangle = \mathrm{tr}\left[\hat{\Pi}_n \hat{\rho}\right]. \tag{6.11}$$

And the conditional state (a-posteriori) becomes

$$\hat{\rho}'_c = \frac{\hat{\Pi}_n \hat{\rho} \hat{\Pi}_n}{\mathrm{tr}\left[\hat{\Pi}_n \hat{\rho}\right]}. \tag{6.12}$$

this conditional state of the system is the state depending on the measurement, assuming that one performs the measurement and also takes a look at the result. In contrast, if one ignores the result, the system is in a mixed state. The so-called unconditioned state:

$$\hat{\rho}' = \sum_n \frac{\hat{\Pi}_n \hat{\rho} \hat{\Pi}_n}{\mathrm{tr}\left[\hat{\Pi}_n \hat{\rho}\right]} \, p(n) = \sum_n \hat{\Pi}_n \hat{\rho} \hat{\Pi}_n \tag{6.13}$$

As the unconditioned state is mixed, it is in general less pure than the a-priori state. This is different from classical undisturbed measurements, there the unconditioned a-posteriori state is identical to the a-priori state. As we will become more clear later on, the conditioned state it is directly linked to the measurement of the system - it is 'conditioned on it'. It denotes the observer's actual state of knowledge resulting from recording the outcomes of the measurement. While the unconditioned state describes the state of knowledge if we make the measurement but throw away the information of the measurement record. It is therefore the result of averaging over all the possible final states resulting from the measurement history. By loosing the information of the the actual measurement record, the unconditioned density matrix is not unique, i.e., different measurements can lead to the same kind of unconditioned density matrix. Projective measurement are only a sub-class of quantum measurements, as mentioned earlier, in general we do not measure the system directly. The system of interest couples to a so-called meter (or probe) and after the meter has interacted with the target system a projective measurement of the meter is performed. To model this approach more general we consider a target system in the initial state $|\phi(t_0)\rangle$ coupled to a second quantum system, the meter, which is initial in state $|\theta(t_0)\rangle$. The initial combined state simply is unentangled and yields

$$|\Psi(t_0)\rangle = |\theta(t_0)\rangle \, |\phi(t_0)\rangle \, . \tag{6.14}$$

The system and the meter interact which each other for a time T and their interaction is described via the unitary evolution operator $\hat{U}(t+T)$. The combined system-meter state after the time interval T becomes

$$|\Psi(t_0 + T)\rangle = \hat{U}(T) \, |\theta(t_0)\rangle \, |\phi(t_0)\rangle \tag{6.15}$$

which can in general no longer be factorized as the meter and the system are entangled. Performing now a projective measurement for a short time-interval $\delta T = T' - T$, assuming that the evolution of the system and the meter over this time-interval is negligible. The measurement is performed on the meter and described by the projector $\Pi_n = |n\rangle \langle n| \otimes \mathbb{1}$. Here n denotes the value of the observed quantity N and $\{|n\rangle\}$ form an orthonormal basis for the meter space. Thus the final state becomes

$$|\Psi(t_0 + T')\rangle = \frac{1}{\sqrt{p(n)}} \, |n\rangle \langle n| \, \hat{U}(T) \, |\theta(t_0)\rangle \, |\phi(t_0)\rangle \tag{6.16}$$

with the probability of measuring the value n for N given as

$$p(n) = \langle\phi(t_0)| \langle\theta(t_0)| \hat{U}^\dagger(T) \ (|n\rangle \langle n| \otimes \mathbb{1}) \ \hat{U}(T) |\theta(t_0)\rangle |\phi(t_0)\rangle . \tag{6.17}$$

which leads to the conditioned state

$$|\Psi(t_0 + T')\rangle = \frac{1}{\sqrt{p(n)}} \ |n\rangle \hat{M}_n |\phi(t_0)\rangle , \quad \hat{M}_n \equiv \langle n| \hat{U}(T) |\theta(t_0)\rangle . \tag{6.18}$$

The measurement on the meter disentangles the system and the meter and so thus the final state can be written as

$$\hat{\rho}_c(t_0 + T') = |n\rangle \langle n| \otimes \frac{1}{p(n)} \ \hat{M}_n \hat{\rho}(t_0) \hat{M}_n^\dagger, \tag{6.19}$$

where the system which was initially in state $\rho(t_0) = |\phi(t_0)\rangle \langle\phi(t_0)|$ is now in the final state described by the last term. The measurement operator $\hat{M}_n$ acts only on the system Hilbert space and the probabilities can be written as

$$p(n) = \langle\phi(t_0)| \hat{M}_n^\dagger \hat{M}_n |\phi(t_0)\rangle . \tag{6.20}$$

The measurement operator is not further defined yet, it's actual form depends on the performed measurement (given the value N with probability n). As the meter and the system are not entangled anymore, one neglects the dynamics of the meter and focuses on the measurement operators and effects E_n which are probability operators defined as

$$|\Psi(t_0 + T')\rangle = \frac{1}{\sqrt{p(n)}} \hat{M}_n |\phi(t_0)\rangle , \quad \hat{E}_n = \hat{M}_n^\dagger \hat{M}_n, \quad \sum_n \hat{E}_n = \mathbb{1} \tag{6.21}$$

the last restriction that the sum over all effects has to equal to the unity matrix (dimension of the system Hilbert space size) is the only restriction here. The set of all effects constitutes a *effect-valued measure* or *probability-operator-valued measure* (POM) on the space of results n (which does not have to the eigenvalue of an observable anymore).

Bibliography

[1] Umesh V. Vazirani. *Qubits, Quantum Mechanics and Computers Script*. UoC Berkeley, 2012.

[2] Martin Grant. *Advanced Statistical Mechanics Script*. McGill University, 2009.

[3] R.P. Feynman and A.R. Hibbs. *Quantum Mechanics and Path Integrals*. McGraw-Hill Book Company, 1965.

[4] Tobias Brandes. *Quantenmechanik*. TU Berlin, 2009.

[5] Joerg Schmalian. *Quantenmechanik*. Karlsruhe Institute of Technology, 2021.

[6] Tobias Brandes. *Mathematische Methoden*. TU Berlin, 2015.

[7] Wolfgang Nolting. *Grundkurs Theoretische Physik: 5.1 Quantenmechanik*. Friedrich Vieweg & Sohn, 1997.

[8] Torsten Fliessbach. *Quantenmechanik*. Springer Spektrum, 2018.

[9] Wolfgang Nolting. *Grundkurs Theoretische Physik: 5.2 Quantenmechanik*. Friedrich Vieweg & Sohn, 1997.

[10] Robert Littlejohn. *Quantum Mechanics*. University of California, Berkeley, 2018/19.

[11] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North Holland, 2007.

[12] J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley Longman, 1994.

[13] Carlton M. Caves, Kip S. Thorne, Ronald W. P. Drever, Vernon D. Sandberg, and Mark Zimmermann. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. i. issues of principle. *Rev. Mod. Phys.*, 52:341–392, Apr 1980.

[14] G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J. M. Raimond, and S. Haroche. Seeing a single photon without destroying it. *Nature*, 400:239–242, 1999.

[15] H. M. Wiseman and G.J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, 2009.