

Quantum Mechanics

Uwe-Jens Wiese
Institute for Theoretical Physics
Bern University

November 19, 2010

Contents

1	Beyond Classical Physics	9
1.1	Introduction	9
1.2	Photons — the Quanta of Light	14
1.3	The Compton Effect	16
1.4	The Photoelectric Effect	17
1.5	Planck’s Formula for Black Body Radiation	18
1.6	Quantum States of Matter	22
1.7	The Double-Slit Experiment	23
1.8	Estimating the Atom	30
2	De Broglie Waves	35
2.1	Wave Packets	35
2.2	Expectation Values	38
2.3	Heisenberg’s Uncertainty Relation	39
2.4	Dispersion Relations	41
2.5	Spreading of a Gaussian Wave Packet	43
3	The Schrödinger Equation	47

3.1	From Wave Packets to the Schrödinger Equation	47
3.2	Conservation of Probability	48
3.3	The Time-Independent Schrödinger Equation	50
4	The Formal Structure of Quantum Mechanics	53
4.1	Wave Functions as Vectors in a Hilbert Space	53
4.2	Observables as Hermitean Operators	54
4.3	Eigenvalues and Eigenvectors	55
4.4	Completeness	56
4.5	Ideal Measurements	57
4.6	Simultaneous Measurability and Commutators	60
4.7	Time Evolution	62
5	Schrödinger Equation in One Dimension	65
5.1	A Particle in a Box	65
5.2	Parity	69
5.3	Simple Models for Atoms and Molecules	71
5.4	Shift Symmetry and Periodic Potentials	75
5.5	A Simple Model for Electrons in a Crystal	76
5.6	The Tunneling Effect	79
5.7	Reflection at a Potential Step	84
5.8	A Paradox	86
6	The Harmonic Oscillator	89
6.1	Solution of the Schrödinger Equation	89
6.2	Operator Formalism	91

<i>CONTENTS</i>	5
6.3 Coherent States and the Classical Limit	93
6.4 The Harmonic Oscillator in Two Dimensions	98
7 The Hydrogen Atom	103
7.1 Separation of the Center of Mass Motion	103
7.2 Angular Momentum	105
7.3 Solution of the Radial Equation	107
7.4 Relativistic corrections	109
8 EPR Paradox and Bell's Inequality	113
8.1 The Einstein-Podolsky-Rosen Paradox	113
8.2 The Quantum Mechanics of Spin Correlations	115
8.3 A Hidden Variable Theory	116
8.4 Bell's Inequality	117
8.5 Schrödinger's Cat	119
9 Abstract Form of Quantum Mechanics	121
9.1 Dirac's Bracket Notation	121
9.2 Unitary Time-Evolution Operator	126
9.3 Schrödinger versus Heisenberg Picture	127
9.4 Time-dependent Hamilton Operators	128
9.5 Dirac's Interaction picture	130
10 Approximation Methods	133
10.1 The Variational Approach	133
10.2 Time-independent Nondegenerate Perturbation Theory to First and Second Order	136

10.3	Time-independent Degenerate Perturbation Theory to First Order	139
10.4	The Hydrogen Atom in a Weak Electric Field	140
10.5	Time-independent Nondegenerate Perturbation Theory to All Orders	142
10.6	Time-independent Degenerate or Quasi-Degenerate Perturbation Theory to All Orders	145
11	Particle in an Electromagnetic Fields	149
11.1	The Classical Electromagnetic Field	149
11.2	Classical Motion of a Charged Particle in an External Electromag- netic Field	151
11.3	Gauge Invariant Form of the Schrödinger Equation	153
11.4	Magnetic Flux Tubes and the Aharonov-Bohm Effect	156
11.5	Flux Quantization for Monopoles and Superconductors	158
11.6	Charged Particle in a Constant Magnetic Field	160
11.7	The Quantum Hall Effect	162
11.8	The “Normal” Zeeman Effect	165
12	Coupling of Quantum Mechanical Angular Momenta	167
12.1	Quantum Mechanical Angular Momentum	167
12.2	Coupling of Spins	170
12.3	Coupling of Orbital Angular Momentum and Spin	173
13	Systems of Identical Particles	175
13.1	Bosons and Fermions	175
13.2	The Pauli Principle	178
13.3	The Helium Atom	180

13.4 Perturbation Theory for Ground and Excited States of the Helium Atom	183
14 Quantum Mechanics of Elastic Scattering	185
14.1 Differential Cross Section and Scattering Amplitude	185
14.2 Green Function for the Schrödinger Equation	187
14.3 The Lippmann-Schwinger Equation	188
14.4 Abstract Form of the Lippmann-Schwinger Equation	189

Chapter 1

Beyond Classical Physics

1.1 Introduction

In our everyday life we experience a world which seems to be governed by the laws of classical physics. The mechanics of massive bodies is determined by Newton's laws and the dynamics of the electromagnetic field follows Maxwell's equations. As we know, these two theories — nonrelativistic classical mechanics and classical electrodynamics — are not consistent with one another. While Maxwell's equations are invariant against Lorentz transformations, Newton's laws are only Galilei invariant. As first realized by Einstein, this means that nonrelativistic classical mechanics is incomplete. Once replaced by relativistic classical mechanics (special relativity) it becomes consistent with electromagnetism, which — as a theory for light — had relativity built in from the start. The fact that special relativity replaced nonrelativistic classical mechanics does not mean that Newton was wrong. In fact, his theory emerges from Einstein's in the limit $c \rightarrow \infty$, i.e. if light would travel with infinite speed. As far as our everyday experience is concerned this is practically the case, and hence for these purposes Newton's theory is sufficient. There is no need to convince a mechanical engineer to use Einstein's theory because her airplanes are not designed to reach speeds anywhere close to $c = 3.00 \times 10^8$ m/sec. Still, as physicists we know that nonrelativistic classical mechanics is not the whole story and relativity (both special and general) has totally altered our way of thinking about space and time. It is certainly a great challenge for a physics student to abstract from our everyday experience of flat space and absolute time and deal with Lorentz contraction, time dilation, or even curved space-time. At the end, it is Nature that forces us to do that

because experiments with objects moving almost with the speed of light are correctly described by relativity, but not by Newton's theory. The classical theories of special and general relativity and of electromagnetism are perfectly consistent with one another. At the beginning of the twentieth century some people thought that physics was a closed subject. Around that time a young student — Werner Heisenberg — was therefore advised to study mathematics instead of physics. He ignored this advice and got involved in the greatest scientific revolution that has ever taken place. Fortunately, he knew enough mathematics to formulate the rules of the new theory — quantum mechanics.

In fact, it was obvious already at the end of the nineteenth century that classical physics was headed for a crisis. The spectra of all kinds of substances consisted of discrete lines, indicating that atoms emit electromagnetic radiation at characteristic discrete frequencies. Classical physics offers no clue to why the frequency of the radiation is quantized. In fact, classically an electron orbiting around the atomic nucleus experiences a constant centripetal acceleration and would hence constantly emit radiation. This should soon use up all its energy and should lead to the electron collapsing into the atomic nucleus: classically the atom would be unstable.

The thermodynamics of light itself also was in bad shape. The classical Rayleigh-Jeans law predicts that the total energy radiated by a black body should be infinite. It required the genius of Max Planck to cure this complicated problem by postulating that matter emits light of frequency ν only in discrete energy units $E = h\nu$. Interestingly, it took a young man like Einstein to take this idea really seriously, and conclude that light actually consists of quanta — so-called photons. Not for relativity but for the explanation of the photoelectric effect he received the Nobel prize. Planck's constant $h = 6.63 \times 10^{-27}$ erg sec is a fundamental constant of Nature. It plays a similar role in quantum physics as c plays in relativity. In particular, it sets the scale at which the predictions of quantum mechanics differ significantly from those of classical physics. Again, the existence of quantum mechanics does not mean that classical mechanics is wrong. It is, however, incomplete and should not be applied to the microscopic quantum world. In fact, classical mechanics is entirely contained within quantum mechanics as the limit $h \rightarrow 0$, just like it is the $c \rightarrow \infty$ limit of special relativity.

Similarly, classical electrodynamics and special relativity can be extended to quantum electrodynamics — a so-called quantum field theory — by “switching on” h . Other forces like the weak and strong interactions are also described by quantum field theories which form the so-called standard model of particle physics. Hence, one might think that at the beginning of the twentyfirst century,

physics finally is complete. However, the force of gravity that is classically described by general relativity has thus far resisted quantization. In other words, again our theories don't quite fit together when we apply them to extreme conditions under which quantum effects of gravity become important. Ironically, today everything fits together only if we put Newton's gravitational constant $G = 0$. The "simplest" interaction — gravity — which got physics started, is now causing trouble. This is good news because it gives us something to think about for this century. In fact, there are promising ideas how to quantize gravity by using so-called string theories which might be a consistent way of letting $G \neq 0$, $c \neq \infty$ and $h \neq 0$ all at the same time.

Before one can get involved in an enterprise like string theory, one must, however, go about a hundred years back and understand what got people like Planck and Heisenberg excited in those days. Even more than understanding relativity, this is a tremendous challenge for a physics student because it requires to leave our everyday experience behind, and be ready to think in completely new categories. Like for Heisenberg, this will only be possible if we equip ourselves with some mathematical tools (like matrices, eigenvalues, eigenvectors, and partial differential equations). The rewards are plenty because this will open the door to the quantum world of atoms, molecules, atomic nuclei, and elementary particles, with phenomena that are totally different from anything in our macroscopic everyday experience.

Learning quantum mechanics can be a truly disturbing — but exciting — experience because the concept of quantum reality differs from that of classical physics. Classical physics is deterministic: given some initial conditions for position and velocity of a particle it makes definite predictions for the trajectory of the particle. In practice the predictive power of classical mechanics may be limited because we never know the initial conditions with absolute precision. This is especially important for chaotic systems. But still, the classical concept of reality implies that the initial conditions could, at least in principle, be determined exactly. In the microscopic quantum world this is not possible, even in principle. In fact, Heisenberg's uncertainty principle,

$$\Delta x \Delta p \geq \frac{h}{4\pi}, \quad (1.1.1)$$

states that position x and momentum p of a particle cannot be determined simultaneously with arbitrary precision. If the position is measured absolutely precisely ($\Delta x = 0$) the momentum will be completely uncertain ($\Delta p = \infty$) and vice versa. At best we can achieve $\Delta x \Delta p = h/4\pi$. In quantum reality position and momentum cannot be determined simultaneously. This immediately implies that the concept of a classical particle trajectory does not make sense at the

quantum level. After all, the trajectory just specifies x and p simultaneously as functions of time.

As Heisenberg's uncertainty principle indicates, quantum mechanics is not deterministic. In fact, subsequent measurements on identically prepared systems generally lead to different results. For example, a sample of identical uranium atoms shows different behavior of individual atoms. Some decay right in front of us, others may wait for years to decay, and we have no way of predicting when a particular atom will dissolve. Still, statistically we can predict how many atoms will decay on the average during a certain period of time. In fact, quantum mechanics only makes predictions of statistical nature and does not say what an individual particle "will do next". It only predicts the probability for a particular result of a measurement. Hence, only when we prepare identical experiments and repeat them over and over again, we can determine the statistical distribution of the results and thus test the predictions of quantum mechanics. The quantum concept of reality implies that a particle has no definite position or momentum before we measure it. During the measurement process the particle "throws the dice" and "chooses" its position from a given statistical distribution. According to the standard Copenhagen interpretation of quantum mechanics the position of a particle is not part of reality. Hence, it is not a failure of quantum mechanics that it does not predict it.

Starting with Einstein, some physicists have not been happy with the quantum concept of reality. (Einstein: "God does not play dice.") Since the quantum concept of reality contradicts our classical everyday experience one can construct various kinds of "paradoxa" involving half-dead animals like Schrödinger's cat. At the end they are all due to our difficulties to abstract from our usual concept of classical reality, and they get resolved when we accept the different concept of quantum reality. Some people, unwilling to do so, have suggested alternative theories, e.g. involving so-called hidden variables. This is equivalent to God telling every uranium atom deterministically when to decay, but such that for us it is indistinguishable from a statistical distribution. In any case, any experiment that has been performed so far is correctly described by quantum mechanics and the Copenhagen interpretation, even if it may contradict our classical concept of reality.

While it is fun to think about Schrödinger's cat, one can do so in a meaningful way only when one understands the basics of quantum mechanics. This requires some mathematical formalism. In fact, a beginner will lack intuition for the phenomena in the quantum world and it is the formalism that will help us to reach correct conclusions. Fortunately, we do not need a whole new mathematics

to do quantum physics. An essential feature of quantum mechanics is that it only predicts probabilities. For example, instead of predicting where a particle is (not an element of quantum reality), it predicts the probability density with which it will be found at a particular position. In effect, the particle position is uncertain and the particle is therefore smeared out. In this respect, it behaves like a wave (which is also defined simultaneously in different regions of space). In fact, in quantum mechanics the probability density to find a particle at position x at time t is given by the absolute square of its wave function $\Psi(x, t)$ which is in general complex. The simplest formulation of nonrelativistic quantum mechanics is Schrödinger's wave mechanics which is based on the Schrödinger equation

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x)\Psi(x, t). \quad (1.1.2)$$

Here $\hbar = h/2\pi = 1.055 \times 10^{-27}$ erg sec and m is the mass of a particle moving under the influence of a potential $V(x)$. The Schrödinger equation is a partial differential equation (just like Maxwell's equations or other wave equations) that describes the time evolution of the wave function. In fact, it is the quantum analog of Newton's equation. Since the Schrödinger equation is a differential equation, to a large extent the mathematics of quantum mechanics is that of differential equations.

A more abstract formulation of quantum mechanics is due to Dirac. He would write the above equation as

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H|\Psi(t)\rangle. \quad (1.1.3)$$

Here $|\Psi(t)\rangle$ is a state vector (which is equivalent to Schrödinger's wave function) in an abstract Hilbert space (a vector space with infinitely many dimensions). The object

$$H = \frac{p^2}{2m} + V(x) \quad (1.1.4)$$

is the so-called Hamilton operator describing kinetic and potential energy of the particle. The Hamilton operator can be viewed as a matrix (with infinitely many rows and columns) in the Hilbert space. Also the momentum p is an operator which can be represented as

$$p = \frac{\hbar}{i} \frac{\partial}{\partial x}, \quad (1.1.5)$$

and can again be viewed as a matrix in Hilbert space. It was Heisenberg who first formulated quantum mechanics as matrix mechanics. Then the various observables (like position x and momentum p) are represented by matrices which

not necessarily commute with each other. For example, in quantum mechanics one finds

$$px - xp = \frac{\hbar}{i}. \quad (1.1.6)$$

This is the origin of Heisenberg's uncertainty relation. Two observables cannot be measured simultaneously with arbitrary precision when the corresponding matrices (operators) do not commute. In the classical limit $\hbar \rightarrow 0$ we recover $px = xp$, such that we can then determine position and momentum simultaneously with arbitrary precision. For time-independent problems the Schrödinger equation reduces to

$$H|\Psi\rangle = E|\Psi\rangle, \quad (1.1.7)$$

which can be viewed as the eigenvalue equation of the Hamilton operator matrix H . The sometimes quantized energy E is a sometimes discrete eigenvalue and the state vector $|\Psi\rangle$ is the corresponding eigenvector. In the early days of quantum physics, when Heisenberg was looking for a permanent job, there were lots of arguments if his or Schödinger's formulation was correct. These ended when Schrödinger showed that both formulations are equivalent. Yet another equivalent formulation of quantum mechanics was developed later by Feynman. His path integral method is applicable to nonrelativistic quantum mechanics, but it shows its real strength only for relativistic quantum field theories.

We may now have a vague idea what quantum mechanics might be, but it will take a while before we fully understand the Schrödinger equation. However, it is time now to switch to a more quantitative approach to the problem. Hopefully you are ready to explore the exciting territory of the microscopic quantum world. Please, leave behind your concept of classical reality and grab your mathematical tool kit. It will help you understand an environment in which reality is probabilistic and where cats can be half alive and half dead.

1.2 Photons — the Quanta of Light

Historically the first theoretical formula of quantum physics was Planck's law for black body radiation. For example, the cosmic background radiation — a remnant of the big bang — is of the black body type. Planck was able to explain the energy distribution of black body radiation by assuming that light of angular frequency $\omega = 2\pi\nu$ is absorbed and emitted by matter only in discrete amounts of energy

$$E = \hbar\omega. \quad (1.2.1)$$

Nowadays, we would associate Planck's radiation formula with a rather advanced field of quantum physics. Strictly speaking, it belongs to the quantum statistical mechanics of the free electromagnetic field, i.e. it is part of quantum field theory — clearly a subject for graduate courses. It is amazing that theoretical quantum physics started with such an advanced topic. Still, we will try to understand Planck's radiation formula in what follows.

Einstein took Planck's idea more seriously than Planck himself and concluded that light indeed consists of quanta of energy $E = \hbar\omega$ which he called photons. In this way he was able to explain the so-called photoelectric effect. It was Hertz who discovered that light shining on a metal surface may knock out electrons from the metal. However, in contrast to classical physics, it requires a minimum frequency — not a minimum intensity — to knock out an electron. This is now understood because the minimum work W that is required to knock out an electron from the metal can only be provided by a photon of minimum frequency $\omega_0 = W/\hbar$. Photons with frequencies below the minimum value simply do not have enough energy. Also many photons of small energy cannot “collaborate” and combine their energies to knock out an electron because the electron is always hit by one photon at a time and none of them is energetic enough.

Collisions of photons and electrons do not only happen at the surface of metals to liberate bound electrons. Also free electrons may collide with photons. This is what is observed in the Compton effect. According to Maxwell's equations an electromagnetic wave of wave vector \vec{k} has angular frequency

$$\omega = |\vec{k}|c. \quad (1.2.2)$$

Recall that the wave length is given by $\lambda = 2\pi/|\vec{k}|$ such that $\lambda = c/\nu$. From special relativity we know that a massless particle such as a photon with momentum \vec{p} has energy

$$E = |\vec{p}|c. \quad (1.2.3)$$

Combining eqs.(1.2.1), (1.2.2) and (1.2.3) suggests that the momentum of a photon in an electromagnetic wave of wave vector \vec{k} is given by

$$\vec{p} = \hbar\vec{k}. \quad (1.2.4)$$

The Compton effect results from an elastic collision of electron and photon in which both energy and momentum are conserved. The Compton effect is correctly described when one uses the above quantum mechanical formula for the momentum of the photon.

It is common to refer to photons as particles, e.g. because they can collide with electrons almost like billiard balls. Then people get confused because light

may behave like waves or like particles. Books about quantum physics are full of remarks about particle-wave duality and about how one picture may be more appropriate than the other depending on the phenomenon in question. This confusion is entirely due to the fact that our human language is sometimes inappropriate to describe the quantum world. At the end the appropriate language is mathematics and its answers are always unique. For example, when we refer to the photon as a particle, we may think of it like a classical billiard ball. This is appropriate only as far as its energy and momentum are concerned. A classical billiard ball also has a definite position in space. Due to Heisenberg's uncertainty relation this is not the case for a photon of definite momentum. When we say that the photon is a particle we should hence not think of a classical billiard ball. Then what do we mean by a particle as opposed to a wave? Our human language distinguishes between the two, but in the mathematical language of quantum physics both are the same thing. It is a fact that quantum theory works with wave equations and objects never have simultaneously a well defined position and momentum. In this sense we are always dealing with waves. In our human language we sometimes like to call these objects particles as well. If you find this confusing, just don't worry about it. It has nothing to do with the real issue. Let mathematics be our guide in the quantum world: in quantum reality there is no distinction between particles and waves.

1.3 The Compton Effect

In the Compton effect we consider the elastic collision of a photon and an electron. In fact, Compton's original experiment was performed with electrons bound in the atoms of a metallic foil. This only complicates the issue and we ignore this here. The main lesson to be learned here is that a photon in an electromagnetic wave of angular frequency ω and wave vector \vec{k} carries energy $E = \hbar\omega$ and momentum $\vec{p} = \hbar\vec{k}$, and can exchange energy and momentum by colliding with other particles like electrons. When the energy of the photon changes in the collision, the frequency of the corresponding electromagnetic wave changes accordingly.

Let us consider an electron at rest. Then its energy is

$$E_e = mc^2, \tag{1.3.1}$$

where $m = 8.29 \times 10^{-7} \text{erg}/c^2$ is its rest mass and its momentum is $\vec{p}_e = 0$. Now we collide the electron with a photon of energy $E_p = \hbar\omega$ and momentum $\vec{p}_p = \hbar\vec{k}$. Of course, we must have $\omega = |\vec{k}|c$. After the collision the electron will no longer

be at rest, but will pick up a momentum \vec{p}'_e . Then its energy will be given by

$$E'_e = \sqrt{(mc^2)^2 + |\vec{p}'_e|^2 c^2}. \quad (1.3.2)$$

Also energy and momentum of the photon will change during the collision. The energy after the collision will be $E'_p = \hbar\omega'$ and the momentum will be $\vec{p}'_p = \hbar\vec{k}'$. The scattering angle θ of the photon is hence given by

$$\cos \theta = \frac{\vec{k} \cdot \vec{k}'}{|\vec{k}| |\vec{k}'|}. \quad (1.3.3)$$

Using energy and momentum conservation one can show that

$$\frac{1}{\omega'} - \frac{1}{\omega} = \frac{\hbar}{mc^2} (1 - \cos \theta). \quad (1.3.4)$$

This means that the frequency of the scattered photon changes depending on the scattering angle. This is clearly a quantum effect. In the classical limit $\hbar \rightarrow 0$ it disappears.

Let us consider two extreme cases for the photon scattering angle: $\theta = 0$ and $\theta = \pi$. For $\theta = 0$ the photon travels forward undisturbed, also leaving the electron unaffected. Then $\cos \theta = 1$ and hence $\omega' = \omega$. In this case the photon does not change its frequency. For $\theta = \pi$ the photon reverses its direction and changes its frequency to

$$\omega' = \frac{\omega}{1 + 2\hbar\omega/mc^2} < \omega. \quad (1.3.5)$$

The photon's frequency — and hence its energy — is smaller than before the collision because it has transferred energy to the electron.

1.4 The Photoelectric Effect

The photoelectric effect consists of the following observations first made by Hertz. When light is shining on a metal surface it may knock out electrons from the metal. The crucial observation is that light of too small frequency cannot knock out any electrons even if it is arbitrarily intense. Once a minimal frequency is exceeded, the number of electrons knocked out of the metal is proportional to the intensity of the light. The kinetic energy of the knocked out electrons was measured to rise linearly with the difference between the applied frequency and the minimal frequency. All these observations have a natural explanation in quantum physics, but cannot be understood using classical concepts.

Einstein explained the photoelectric effect as follows. Electrons are bound within the metal and it requires a minimal work W to knock out an electron from the metal's surface. Any excess energy will show up as kinetic energy of the knocked out electron. The energy necessary to knock out an electron is provided by a single photon in the incoming electromagnetic wave. As we have learned, such a photon has energy $E = \hbar\omega$ where ω is the angular frequency of the light wave. Only when ω exceeds a minimal frequency

$$\omega_0 = \frac{W}{\hbar}, \quad (1.4.1)$$

an electron can actually be knocked out. The excess energy shows up as kinetic energy

$$T = \hbar\omega - W = \hbar(\omega - \omega_0) \quad (1.4.2)$$

of the knocked out electron. This is exactly what was observed experimentally.

1.5 Planck's Formula for Black Body Radiation

Now that we have familiarized ourselves with the properties of photons, we may turn to their thermodynamics. Thermodynamics deals with systems of many particles that are in thermal equilibrium with each other and with a thermal bath. The energies of the individual states are statistically distributed, following a Boltzmann distribution for a given temperature T . The thermal statistical fluctuations are of a different nature than those related to quantum uncertainty. Thermal fluctuations are present also at the classical level, and e.g. reflect our inability to treat a system with a large number of particles exactly. Following the classical concept of reality, this is possible, at least in principle. In practice, it is, however, much more appropriate to use a classical statistical description. In the thermodynamics of photons, i.e. in quantum statistical mechanics, we deal with thermal and quantum fluctuations at the same time.

A system of photons in thermal equilibrium has been with us from the beginning of our Universe. Immediately after the big bang the energy density — and hence the temperature — was extremely high and all kinds of elementary particles (among them photons, electrons, and their anti-particles — positrons — as well as neutrinos and anti-neutrinos) have existed as an extremely hot gas filling all of space. These particles interacted with each other e.g. via Compton scattering. As the Universe expanded, the temperature decreased and electrons and positrons annihilated into a lot of photons. A very small fraction of the electrons (actually all the ones in the Universe today) exceeded the number of

positrons and thus survived annihilation. At this time — a few seconds after the big bang — no atom had ever been formed. As a consequence, there were no characteristic colors of selected spectral lines. This is what we mean when we talk about the cosmic photons as black body radiation. About 400000 years after the big bang the Universe had expanded and cooled so much that electrons and atomic nuclei could settle down to form neutral atoms. At that time the Universe became transparent. The photons that emerged from the mass extinction of electrons and positrons were left alone and are still floating through our Universe. Of course, in the last 14 billion years the Universe has expanded further and the cosmic photon gas has cooled down accordingly. Today the temperature of the cosmic background radiation is 2.735 K (degrees Kelvin). It is amazing that this temperature is to very high precision the same, no matter what corner of the Universe the photons come from. This was first explained by Alan Guth from MIT using the idea of the inflationary Universe.

How does one measure the temperature of a system of photons? The temperature is defined via the Boltzmann distribution, in our case by the intensity of radiation with a certain frequency. Hence, by measuring the photon spectrum one can determine the temperature. This is exactly what the antennae of the cosmic background explorer satellite WMAP have been doing. Equipped with the idea of Planck, let us now derive this spectrum theoretically. For simplicity we replace the Universe by a large box of spatial size $L \times L \times L$ with periodic boundary conditions. This is only a technical trick that will allow us to simplify the calculation. At the end we let $L \rightarrow \infty$. We will proceed in three steps. First we work classically and classify all possible modes of the electromagnetic field in the box. Then we switch to quantum mechanics and populate these modes with light quanta (photons). Finally, we do quantum statistical mechanics by summing over all quantum states using Boltzmann's distribution.

What are the modes of the electromagnetic field in an L^3 periodic box? First of all, we can classify them by their wave vector \vec{k} which is now restricted to discrete values

$$\vec{k} = \frac{2\pi}{L}\vec{m}, \quad m_i \in \mathbb{Z}. \quad (1.5.1)$$

The frequency of this mode is given by

$$\omega = |\vec{k}|c. \quad (1.5.2)$$

Each of the modes can exist in two polarization states.

Now we turn to quantum mechanics and populate the classical modes with photons. As we have learned, a mode of frequency ω can host photons of energy

$$E(\vec{k}) = \hbar\omega = \hbar|\vec{k}|c \quad (1.5.3)$$

only. Photons are so-called bosons. This means that an arbitrary number of them can occupy a single mode of the electromagnetic field. Electrons and neutrinos, for example, behave very differently. They are so-called fermions, i.e. at most one of them can occupy a single mode. All elementary particles we know are either fermions or bosons. We can completely classify a quantum state of the electromagnetic field by specifying the number of photons $n(\vec{k}) \in \{0, 1, 2, \dots\}$ occupying each mode (characterized by wave vector \vec{k} and polarization, which we suppress in our notation). It is important to note that it does not matter “which photon” occupies which mode. Individual photons are indistinguishable from each other, they are like perfect twins. Hence specifying their number per mode determines their state completely.

Now that we have classified all quantum states of the electromagnetic field by specifying the photon occupation numbers for all modes, we can turn to quantum statistical mechanics. Then we must evaluate the so-called partition function by summing over all states. Since the modes are completely independent of one another, the partition function factorizes into partition functions for each individual mode. Here we consider a single mode partition function

$$Z = \sum_{n(\vec{k})=0}^{\infty} \exp(-n(\vec{k})E(\vec{k})/k_B T) = \sum_{n(\vec{k})=0}^{\infty} \exp(-\beta n(\vec{k})\hbar|\vec{k}|c). \quad (1.5.4)$$

Each mode state is weighed by its Boltzmann factor $\exp(-n(\vec{k})E(\vec{k})/k_B T)$ which is determined by its total energy of photons $n(\vec{k})E(\vec{k})$ occupying that mode and by the temperature T . The Boltzmann constant $k_B = 1.38 \times 10^{-16}$ erg/K arises just for dimensional reasons. We have also introduced $\beta = 1/k_B T$. Now we make use of the well-known summation formula for a geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}. \quad (1.5.5)$$

Using $x = \exp(-\beta\hbar|\vec{k}|c)$ we obtain the partition function corresponding to the so-called Bose-Einstein statistics

$$Z = \frac{1}{1 - \exp(-\beta\hbar|\vec{k}|c)}. \quad (1.5.6)$$

We are interested in the statistical average of the energy in a particular mode,

which is given by

$$\begin{aligned}\langle n(\vec{k})E(\vec{k}) \rangle &= \frac{1}{Z} \sum_{n(\vec{k})=0}^{\infty} n(\vec{k})E(\vec{k}) \exp(-\beta n(\vec{k})E(\vec{k})) \\ &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} = \frac{\hbar |\vec{k}| c}{\exp(\beta \hbar |\vec{k}| c) - 1}.\end{aligned}\quad (1.5.7)$$

Finally, we are interested in the average total energy as a sum over all modes

$$\langle E \rangle = 2 \sum_{\vec{k}} \langle n(\vec{k})E(\vec{k}) \rangle \rightarrow 2 \left(\frac{L}{2\pi}\right)^3 \int d^3k \langle n(\vec{k})E(\vec{k}) \rangle. \quad (1.5.8)$$

Here a factor 2 arises due to the two polarization states. In the last step we have performed the infinite volume limit $L \rightarrow \infty$. Then the sum over discrete modes turns into an integral. It is no surprise that our result grows in proportion to the volume L^3 . We should simply consider the energy density $u = \langle E \rangle / L^3$. We can now perform the angular integration and also replace $|\vec{k}| = \omega/c$ to obtain

$$u = \frac{1}{\pi^2 c^3} \int_0^{\infty} d\omega \frac{\hbar \omega^3}{\exp(\beta \hbar \omega) - 1}. \quad (1.5.9)$$

Before we do the integral we read off the energy density per frequency unit for modes of a given angular frequency ω

$$\frac{du(\omega)}{d\omega} = \frac{1}{\pi^2 c^3} \frac{\hbar \omega^3}{\exp(\hbar \omega / k_B T) - 1}. \quad (1.5.10)$$

This is Planck's formula that was at the origin of quantum mechanics. If you have followed the above arguments: congratulations, you have just mastered a calculation in quantum field theory!

Since we have worked quite hard to produce this important result, let us discuss it in some detail. Let us first consider the classical limit $\hbar \rightarrow 0$. Then we obtain the classical Rayleigh-Jeans law

$$\frac{du(\omega)}{d\omega} = \frac{\omega^2 k_B T}{\pi^2 c^3}. \quad (1.5.11)$$

Integrating this over all frequencies gives a divergent result from the high frequency end of the spectrum. This is the so-called ultraviolet Jeans catastrophe. The classical thermodynamics of the electromagnetic field gives an unphysical result. Now we go back to Planck's quantum result and perform the integral over all frequencies ω . This gives the Stefan-Boltzmann law

$$u = \frac{\pi^2 k_B^4 T^4}{15 \hbar^3 c^3}, \quad (1.5.12)$$

which is not only finite, but also agrees with experiment. Again, in the classical limit the result would be divergent. It is interesting that for high temperatures and for low frequencies, i.e. for $k_B T \gg \hbar\omega$, Planck's formula also reduces to the classical result. Quantum effects become important only in the low-temperature and high-frequency regimes.

Now we can understand how WMAP measures the temperature of the cosmic background radiation. The energy density is measured for various frequencies and is then compared with Planck's formula which leads to a high precision determination of T . The WMAP data tell us a lot about how our Universe began. In fact, the early history of the Universe is encoded in the photons left over from the big bang. Sometimes one must understand the very small before one can understand the very large.

1.6 Quantum States of Matter

We have seen that at the quantum level the energy of a light wave is carried by photons which behave like particles. Similarly, in the quantum world objects that we like to think about as particles also behave like waves. Typical properties of classical electromagnetic waves are diffraction and interference. Similar experiments can also be performed with electrons, neutrons, or atoms. Quantum mechanically they all show typical wave behavior as e.g. observed in the experiments by Davisson and Germer. A famous Gedankenexperiment that illustrates the diffraction of matter waves is the double-slit experiment, in analogy to Young's experiment using light waves. De Broglie was first to suggest that not only photons but any particle of momentum \vec{p} has a wave vector \vec{k} associated with it. As for photons he postulated

$$\vec{p} = \hbar\vec{k}. \quad (1.6.1)$$

The de Broglie wave length of a particle is hence given by

$$\lambda = \frac{2\pi}{|\vec{k}|} = \frac{h}{|\vec{p}|}. \quad (1.6.2)$$

When one wants to detect the wave properties of matter experimentally, one must work with structures of the size of the wave length. In the classical limit $\hbar \rightarrow 0$ the wave length goes to zero. Then even on the smallest scales one would not detect any wave property. Of course, in the real world $\hbar \neq 0$ and wave properties are observed.

Another manifestation of quantum mechanics in the material world are the discrete spectra of atoms and molecules. They arise from transitions between quantized energy levels of the electrons surrounding the atomic nucleus. The Schrödinger equation predicts the existence of discrete energy values. For example, in the hydrogen atom (consisting of an electron and a much heavier proton that acts as the atomic nucleus) the allowed energy values in the discrete spectrum are given by

$$E_n = -\frac{me^4}{2\hbar^2 n^2}, \quad n \in \{1, 2, 3, \dots\}, \quad (1.6.3)$$

where m and e are mass and charge of the electron. When an atom is in an excited state with $n_1 > 1$ it can emit a photon and thus jump into a lower lying state with $n_2 < n_1$. The energy difference between the two states shows up as the energy (and hence the frequency) of the emitted photon, i.e.

$$E_1 - E_2 = \frac{me^4}{2\hbar^2} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right) = \hbar\omega. \quad (1.6.4)$$

This simple observation led Bohr to a basic understanding of atomic spectra. His model of the atom was ultimately replaced by the Schrödinger equation, but it played an important role in the early days of quantum physics.

1.7 The Double-Slit Experiment

Double-slit experiments were originally performed with electromagnetic waves. The interference patterns resulting from light passing through the slits in a screen reflect the wave properties of electromagnetic radiation. They also confirm that electromagnetic waves can be linearly superimposed. The fact that Maxwell's equations are linear in the electromagnetic field are a mathematical manifestation of the superposition principle. The double slit experiment has also been used to illustrate the wave properties of matter. Again, interference patterns emerge that point to a superposition principle for matter waves. Indeed, Schrödinger's wave equation is linear in the quantum mechanical wave function. Experimentally it is much harder to do double-slit experiments with matter waves than with light waves in the visible range of the spectrum. This is because the wave length of matter waves is usually a lot shorter. Hence, the slits that cause the diffraction pattern must be much more closely spaced than in experiments with light waves. Still, experiments with electron beams, for example, have been performed and they demonstrate clearly the wave properties of matter.

Let us first discuss the double-slit experiment for light waves that was originally performed by Young. We consider a plane electromagnetic wave propagating

in the x -direction with wave number k and angular frequency $\omega = kc$. The wave falls on an opaque planar screen oriented in the y - z -plane with two slits located at $x = z = 0$ and $y_1 = 0, y_2 = d$. Geometrical optics would predict that only two rays of light will make it through the slits, and the screen creates a sharp shadow behind it. This prediction is correct only when the wave length of the light is very small compared to the size of the holes in the screen. Here we are interested in the opposite limit of very small holes compared to the wave length. Then the predictions of geometrical optics fail completely and only Maxwell's equations predict the correct behavior. For matter waves the analog of geometrical optics is classical physics. Once the wave properties of matter become important, only the Schrödinger equation yields correct results. For light waves each of the slits in the screen acts like an antenna that radiates a circular wave. Let us restrict ourselves to the x - y -plane. Then the electric field of an antenna is given by

$$\vec{E}(r, t) = \frac{A}{r} \sin(\omega t - kr) \vec{e}_z, \quad (1.7.1)$$

where r is the distance from the antenna. Also we have assumed polarization in the z -direction. For an observer at large distance, the two antennae corresponding to the two slits have, to a first approximation, the same distance r . However, this approximation is not good enough for our purposes. Because we are interested in interference effects, we must be able to detect differences in the optical paths of the order of a wavelength. Still, for an observer in the far zone we can assume that all incoming light rays are practically parallel. The difference between path lengths between the circular waves emanating from the two slits is $d \sin \varphi$ where φ is the angle between the x -axis and the direction of observation behind the screen. When $d \sin \varphi = \lambda m$ is an integer multiple of the wavelength the two waves arrive at the observer in phase and we get a large intensity. To see this analytically, we add up the two electric fields

$$\begin{aligned} \vec{E}(r, \varphi, t) &= \vec{E}_1(r_1, t) + \vec{E}_2(r_2, t) \\ &= \frac{A}{r_1} \sin(\omega t - kr_1) + \frac{A}{r_2} \sin(\omega t - kr_2) \\ &= \frac{A}{r} \text{Im}[\exp i(\omega t - kr) + \exp i(\omega t - kr + \delta)] \vec{e}_z. \end{aligned} \quad (1.7.2)$$

Here we have used $r = r_1 \approx r_2$ in the amplitude of the waves. In the phase we need to be more precise. There we used

$$r_2 = r - d \sin \varphi, \quad (1.7.3)$$

and we have introduced the phase shift

$$\delta = kd \sin \varphi. \quad (1.7.4)$$

The complex notation is not really necessary but it makes life easier. We now write

$$\begin{aligned}\vec{E}(r, \varphi, t) &= \frac{A}{r} \text{Im}[\exp i(\omega t - kr + \delta/2)(\exp(-i\delta/2) + \exp(i\delta/2))] \vec{e}_z \\ &= \frac{A}{r} \sin(\omega t - kr + \delta/2) 2 \cos(\delta/2) \vec{e}_z.\end{aligned}\tag{1.7.5}$$

We are interested in the time averaged intensity of the total field. Hence, we square \vec{E} , average over t , and obtain for the intensity

$$I(\varphi) = 4I_0 \cos^2(\delta/2).\tag{1.7.6}$$

Here I_0 is the φ -independent intensity in the presence of just a single slit. What is the intensity observed at angle $\varphi = 0$? Naively, one might expect $2I_0$, but the value is indeed $4I_0$. What happens at the angles φ for which $d \sin \varphi = \lambda m$? Then

$$\delta = kd \sin \varphi = k\lambda m = 2\pi m.\tag{1.7.7}$$

The intensity is 2π -periodic in δ and hence the intensity is again $4I_0$. For these particular angles both waves are in phase, thus yielding a large intensity. The minima of the intensity, on the other hand, correspond to $\delta = \pi$.

Let us now turn to the diffraction and interference of matter waves. We can think of a beam of electrons incident on the same screen as before. We prepare an ensemble of electrons with momentum p_x in the x -direction and with $p_y = p_z = 0$. As we will understand later, this state is described by a quantum mechanical wave function which is again a plane wave. We will detect the diffracted electrons with a second detection screen a distance l behind the screen with the two slits, and parallel to it. We are interested in the probability to find electrons in certain positions at the detection screen at $x = l$. Again, we limit ourselves to $z = 0$ and ask about the probability as a function of y .

As for any problem in quantum mechanics, one can answer this question by solving the Schrödinger equation. Since we have not yet discussed the Schrödinger equation in sufficient detail, we will proceed differently. In fact, we will make a detour to so-called path integrals — a very interesting formulation of quantum mechanics that was discovered by Feynman. He noted that classical and quantum physics are more closely related than one might think based on the Schrödinger equation. In particular, Feynman realized that quantum physics is governed by the action of the corresponding classical problem. Let us introduce the action in the context of a particle of mass m moving under the influence of a potential $V(x)$ in one dimension. The action is a number associated with any possible path $x(t)$ from a point $x(0)$ at an initial time 0 to a point $x(T)$ at a final time T . The

path need not be the one that a classical particle will actually follow according to Newton's equation. Even a path that cannot be realized classically has an action associated with it. The action is defined as

$$S[x(t)] = \int_0^T dt \left[\frac{m}{2} \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) \right], \quad (1.7.8)$$

i.e. one simply integrates the difference of kinetic and potential energy over time. The action has a quite remarkable property. It is minimal (or at least extremal) for the classical path — the one that the particle takes following Newton's equation. All other (classically not realized) paths have a larger action than the actual classical path. This can be understood most easily when we discretize time into N small steps of size ε such that $N\varepsilon = T$ and $t_n = n\varepsilon$. At the end we will let $N \rightarrow \infty$ and $\varepsilon \rightarrow 0$ such that T remains fixed. The discretized version of the action takes the form

$$S[x(t)] = \sum_{n=0}^{N-1} \varepsilon \left[\frac{m}{2} \left(\frac{x(t_{n+1}) - x(t_n)}{\varepsilon} \right)^2 - V(x(t_n)) \right]. \quad (1.7.9)$$

Note that we have also replaced the derivative $dx(t)/dt$ by its discretized version. Now let us find the path that minimizes the action simply by varying $S[x(t)]$ with respect to all $x(t_n)$. Then we obtain

$$\frac{dS[x(t)]}{dx(t_n)} = \varepsilon \left[-m \frac{x(t_{n+1}) - x(t_n)}{\varepsilon^2} + m \frac{x(t_n) - x(t_{n-1})}{\varepsilon^2} - \frac{dV(x(t_n))}{dx(t_n)} \right] = 0. \quad (1.7.10)$$

We know that the gradient of the potential is just the force, i.e.

$$-\frac{dV(x)}{dx} = F(x), \quad (1.7.11)$$

such that the above equation takes the form

$$m \frac{x(t_{n+1}) - 2x(t_n) + x(t_{n-1}))}{\varepsilon^2} = F(x(t_n)). \quad (1.7.12)$$

The term on the left-hand side is a discretized second derivative. Taking the limit $\varepsilon \rightarrow 0$ we obtain

$$m \frac{d^2x(t)}{dt^2} = F(x(t)). \quad (1.7.13)$$

This is exactly Newton's equation that determines the classical path. Hence, we have shown that the classical path is a minimum (more generally an extremum) of the action. In fact, in modern formulations of classical mechanics the action plays a very important role.

Feynman discovered that the classical action also governs quantum mechanics. In fact, he found a new path integral formulation of quantum mechanics that is equivalent to Heisenberg's and Schrödinger's formulations, and that is often used in quantum field theories. In quantum mechanics, on the other hand, solving the Schrödinger equation is often easier than doing the path integral. Therefore, we will soon concentrate on Schrödinger's method. Still, to describe the diffraction of matter waves, the path integral is very well suited. Quantum mechanically we should not think of a particle tracing out a classical path. This follows already from Heisenberg's uncertainty relation which tells us that position and momentum of a particle are not simultaneously measurable with arbitrary precision. Feynman realized that the concept of a particle's path can still be useful when one wants to determine the probability for finding a particle that started at time $t = 0$ at position $\vec{x} = 0$ at a final time $t = T$ in a new position $\vec{x}(T)$. More precisely we are interested in the transition probability amplitude $\Psi(\vec{x}(T))$ — a complex number closely related to the quantum mechanical wave function — whose absolute value $|\Psi(\vec{x}(T))|^2$ defines a probability density, i.e. $|\Psi(\vec{x}(T))|^2 d^3x$ is the probability to find the particle at time T in a volume element d^3x around the position $\vec{x}(T)$. Recall that we have assumed that the particle started initially at time $t = 0$ at position $\vec{x} = 0$. Feynman's idea was that $\Psi(\vec{x}(T))$ gets contributions from all possible paths that connect $\vec{x} = 0$ with $\vec{x}(T)$, not just from the classical path that minimizes the action. The Feynman path integral represents $\Psi(\vec{x}(T))$ as an integral over all paths

$$\Psi(\vec{x}(T)) = \int \mathcal{D}\vec{x}(t) \exp\left(\frac{i}{\hbar} S[\vec{x}(t)]\right). \quad (1.7.14)$$

The essential observation is that the action $S[x(t)]$ of a path $x(t)$ determines its contribution to the total path integral. Defining precisely what we mean by the above path integral expression, and finally performing the path integral, goes beyond the scope of this course. What we will need for the calculation of electron diffraction is just the path integral for a free particle. In that case one finds

$$\Psi(\vec{x}(T)) = A \exp\left(\frac{i}{\hbar} S[\vec{x}_c(t)]\right). \quad (1.7.15)$$

Here $\vec{x}_c(t)$ is the classical path — the one determined by Newton's equation. Hence, for a free particle the whole path integral is proportional to the contribution of the classical path. The proportionality factor A is of no concern for us here.

What is the action of the classical path of a free particle moving from $\vec{x} = 0$ at time $t = 0$ to $\vec{x}(T)$ at time $t = T$? Newton's equation tells us that a free

particle moves along a straight line, and hence the classical path is

$$\vec{x}_c(t) = \frac{\vec{x}(T)}{T}t. \quad (1.7.16)$$

The derivative

$$\frac{d\vec{x}_c(t)}{dt} = \frac{\vec{x}(T)}{T}, \quad (1.7.17)$$

is just the constant velocity. The action of the classical path is thus given by

$$S[\vec{x}_c(t)] = \int_0^T dt \frac{m}{2} \frac{|\vec{x}(T)|^2}{T^2} = \frac{m|\vec{x}(T)|^2}{2T}, \quad (1.7.18)$$

and Feynman's transition probability amplitude is

$$\Psi(\vec{x}(T)) = A \exp\left(\frac{im|\vec{x}(T)|^2}{2\hbar T}\right). \quad (1.7.19)$$

The amplitude depends on the distance squared $|\vec{x}(T)|^2$ that the particle traveled during the time T .

Now we are ready to calculate the probability distribution of electrons at the detection screen. An electron detected at a position y may have originated either from the slit at $y_1 = d/2$ or from the slit at $y_2 = -d/2$, and we don't know through which slit it came. In the first case it has traveled a distance squared $|\vec{x}(T)|^2 = l^2 + (y - d/2)^2$, and in the second case the distance squared is $|\vec{x}(T)|^2 = l^2 + (y + d/2)^2$. Feynman's path integral method requires to add the contributions of the two paths such that the total transition probability amplitude is

$$\begin{aligned} \Psi(y) &= A \exp\left(\frac{im(l^2 + (y - d/2)^2)}{2\hbar T}\right) + A \exp\left(\frac{im(l^2 + (y + d/2)^2)}{2\hbar T}\right) \\ &= A \exp\left(\frac{im(l^2 + y^2 + d^2/4)}{2\hbar T}\right) \left[\exp\left(\frac{imy d}{2\hbar T}\right) + \exp\left(-\frac{imy d}{2\hbar T}\right) \right] \\ &= A \exp\left(\frac{im(l^2 + y^2 + d^2/4)}{2\hbar T}\right) 2 \cos\left(\frac{my d}{2\hbar T}\right). \end{aligned} \quad (1.7.20)$$

To obtain the probability density we take the absolute value squared

$$P(y) = |\Psi(y)|^2 = 4|A|^2 \cos^2\left(\frac{my d}{2\hbar T}\right) = 4P_0 \cos^2\left(\frac{my d}{2\hbar T}\right). \quad (1.7.21)$$

Here P_0 is the y -independent probability density that one finds in an experiment with just a single slit.

The above result is very similar to the one for diffraction of light waves. In particular, again we find the $\cos^2(\delta/2)$ distribution if we identify

$$\delta = \frac{myd}{\hbar T}. \quad (1.7.22)$$

In the case of light waves we had

$$\delta = kd \sin \varphi. \quad (1.7.23)$$

In the present case we have $y = l \tan \varphi \approx l \sin \varphi$ for small φ , such that we can identify

$$k = \frac{ml}{\hbar T}. \quad (1.7.24)$$

Here l/T is just the distance that the electron traveled in the x -direction divided by the time T . Multiplying this velocity by the mass we get the momentum component in the x -direction

$$p_x = \frac{ml}{T}. \quad (1.7.25)$$

Hence, we finally identify

$$p_x = \hbar k. \quad (1.7.26)$$

In other words, the diffraction pattern of electrons of momentum p_x is the same as the one of light waves with wave number $k = p_x/\hbar$. Then it is natural to associate that wave number also with the electrons. The corresponding wave length

$$\lambda = \frac{2\pi}{k} = \frac{2\pi\hbar}{p_x} = \frac{h}{p_x}, \quad (1.7.27)$$

is the so-called de Broglie wave length of a particle of momentum p_x . It is remarkable that we get the same relation as for photons. Generalizing to three dimensions we hence associate with any particle of momentum \vec{p} a wave with wave number vector \vec{k} such that

$$\vec{p} = \hbar \vec{k}. \quad (1.7.28)$$

It is important to note that in the above double slit experiments we did not know through which slit the particles (photons or electrons) actually went. In fact, we added contributions from both possibilities, and only that gave rise to interference patterns. Once we would figure out through which slit the particles went (for example, by putting a small detector) the interference pattern would disappear. This is a consequence of Heisenberg's uncertainty relation. At the quantum level the influence that a measuring device has on the observed phenomenon can in general not be made arbitrarily small. A small detector, for example, that determines through which slit a particle went would unavoidably change the particle's momentum, which then wipes out the interference effects.

1.8 Estimating the Atom

A series of experiments by Rutherford and his collaborators revealed that atoms consisting of negatively charged electrons and a positively charged atomic nucleus have the positive charge concentrated in a tiny region compared to the size of the atom. Consequently, as far as atomic physics is concerned, the atomic nucleus can essentially be considered point-like. Of course, the atomic nucleus itself consists of protons and neutrons, which in turn consist of quarks and gluons, but this is physics at much smaller length scales. In contrast to the positive charge, the negative charge of the electrons is distributed over the whole volume of the atom, i.e. over a region with a radius of about 10^{-10} m. The electron itself is point-like, as far as we can tell today. Hence, the distribution of negative charge over the atom is entirely due to the fact that electrons move inside the atom. The forces that govern the electron dynamics are of electromagnetic nature. The Coulomb force attracts the negatively charged electron to the positively charged atomic nucleus. As opposed to the charge, the mass of an atom is almost entirely concentrated in the nucleus, simply because protons and neutrons have masses roughly a factor 2000 bigger than the electron mass. The above picture of the atom that emerged from various experiments is inconsistent with classical physics. In fact, an electron that surrounds an atomic nucleus is constantly accelerated and should hence constantly emit electromagnetic radiation. This should lead to a fast energy loss, causing the electron to spiral into the nucleus. After this collapse the classical atom would not be bigger than the nucleus itself.

Based on the experimental observations of Rutherford, Bohr constructed a model of the atom that eliminated the classical collapse problem by postulating that electrons can move around the nucleus in certain discrete stable orbits without emitting electromagnetic radiation. These discrete orbits were singled out by certain quantization conditions. The energies of the discrete orbits are then also quantized. Atoms can absorb or emit electromagnetic radiation by undergoing transitions between the various quantized energy levels. The energy difference between the levels is then turned into the energy of a photon, and hence manifests itself in the characteristic spectral lines that are observed experimentally. In this way Bohr was able to describe the spectra of various atoms. Although in the early days of quantum mechanics Bohr's model of the atom played an important role, from a modern perspective Bohr's quantization method is to be considered semi-classical, or if we want semi-quantum. A full quantum description of the atom emerged only when one was able to derive atomic spectra from the Schrödinger equation.

It will take us until the end of the first part of this course to solve the

Schrödinger equation for the simplest atom — the hydrogen atom consisting of a single electron and a proton that plays the role of the atomic nucleus. Hence, we will now use Bohr's quantization method to get at least a first estimate of the atom. Let us consider hydrogen-like atoms with just a single electron of charge $-e$, but with Z protons in the atomic nucleus and hence with nuclear charge Ze . The Coulomb potential exerted on the electron is then given by

$$V(r) = -\frac{Ze^2}{r}, \quad (1.8.1)$$

where r is the distance of the electron from the atomic nucleus. Since the nucleus is so much heavier than the electron we can assume that it remains at rest in the center of mass frame of the whole atom. Further, let us assume that the electron moves around the nucleus classically in a circular orbit of radius r with angular frequency ω . Then its acceleration is given by $a = \omega^2 r$, and thus Newton's equation tells us that

$$m\omega^2 r = ma = F(r) = \frac{Ze^2}{r^2}. \quad (1.8.2)$$

Let us now use Feynman's path integral again. We have learned that a particle that started at $\vec{x} = 0$ at $t = 0$ has a transition probability amplitude $\Psi(\vec{x}(T))$ for arriving at a later time T at the new position $\vec{x}(T)$. This amplitude is given by a path integral, and for a free particle the amplitude is proportional to the contribution of the classical path, i.e.

$$\Psi(\vec{x}(T)) = A \exp\left(\frac{i}{\hbar} S[\vec{x}_c(t)]\right). \quad (1.8.3)$$

An electron that is bound inside an atom is certainly not free. Therefore the above equation can only be approximately correct. In fact, the assumption of the dominance of the classical path implies a semi-classical approximation. The motion of the electron on its circular orbit is periodic with period $T = 2\pi/\omega$. Hence, Feynman's transition amplitude for a complete period, with the particle returning to its initial position, will be approximately given by the action corresponding to the periodic path. As we will learn later, the phase of the quantum mechanical wave function of a stationary state with energy E changes by $\exp(-iET/\hbar)$ during the time T . Consistency with Feynman's transition amplitude therefore requires

$$\exp\left(\frac{i}{\hbar} S[\vec{x}_c(t)]\right) = \exp\left(-\frac{i}{\hbar} ET\right), \quad (1.8.4)$$

and hence

$$S[\vec{x}_c(t)] + ET = 2\pi n\hbar. \quad (1.8.5)$$

Here $n \in \{1, 2, 3, \dots\}$. This is the so-called Bohr-Sommerfeld quantization condition for stationary states with energy E .

Let us use the Bohr-Sommerfeld quantization condition to estimate the energies of stationary states of hydrogen-like atoms. First, we need to compute the action of the periodic circular path of the electron. The kinetic energy on that path is $\frac{1}{2}m\omega^2r^2$ and the potential energy is $-Ze^2/r$, which are both time-independent for a circular path. Hence, the action for a complete period — as the time integral of the difference between kinetic and potential energy — is simply given by

$$S[\vec{x}_c(t)] = T\left(\frac{1}{2}m\omega^2r^2 + \frac{Ze^2}{r}\right). \quad (1.8.6)$$

The total energy, on the other hand, has the form

$$E = \frac{1}{2}m\omega^2r^2 - \frac{Ze^2}{r}. \quad (1.8.7)$$

Then the quantization condition takes the form

$$S[\vec{x}_c(t)]/T + E = m\omega^2r^2 = 2\pi n\hbar/T = n\hbar\omega, \quad (1.8.8)$$

and therefore

$$m\omega r^2 = n\hbar. \quad (1.8.9)$$

This is nothing but the angular momentum, more precisely its component perpendicular to the plane of the circular path, which is hence quantized in integer multiples of \hbar . Using this together with eq.(1.8.2) one obtains

$$r = \frac{n^2\hbar^2}{Ze^2m} = \frac{n^2\hbar}{Z\alpha mc}. \quad (1.8.10)$$

Here we have introduced the fine-structure constant

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.04}, \quad (1.8.11)$$

which characterizes the strength of the electromagnetic interaction. One notices that the radius of the electronic orbit grows in proportion to the quantum number n squared. The orbit of smallest radius corresponds to $n = 1$. In the hydrogen atom, i.e. for $Z = 1$, this orbit has a radius

$$r = \frac{\hbar}{\alpha mc} = 5.3 \times 10^{-9} \text{cm}. \quad (1.8.12)$$

This so-called Bohr radius sets the scale for the typical size of atoms. We also note that the radius of the smallest orbit is proportional to $1/Z$. Thus in heavy atoms electrons can come closer to the atomic nucleus, simply because there is

a stronger Coulomb force due to the larger number of protons in the nucleus. Using eq.(1.8.7) together with eq.(1.8.2) and eq.(1.8.10) one obtains

$$E = -\frac{mZ^2e^4}{2n^2\hbar^2} = -\frac{mc^2(Z\alpha)^2}{2n^2}. \quad (1.8.13)$$

These are the correct quantized energy values of the hydrogen spectrum. Although our calculation was only semi-classical, we get the full quantum answer — the same that one gets from Schrödinger's equation. This is a peculiarity of the hydrogen atom and will in general not be the case. The state of lowest energy of the hydrogen atom — its so-called ground state — corresponds to $n = 1$. If we want to ionize the hydrogen atom out of its ground state, i.e. if we want to strip off the electron from the proton, we thus need to invest the energy

$$E = \frac{me^4}{2\hbar^2} = \frac{mc^2\alpha^2}{2} = 13.6\text{eV}. \quad (1.8.14)$$

Similarly, if we want to raise the energy of the hydrogen atom from its ground state with $n = 1$ to its first excited state with $n = 2$ we must provide the energy difference

$$E = \frac{mc^2\alpha^2}{2}\left(1 - \frac{1}{4}\right) = 10.2\text{eV}. \quad (1.8.15)$$

A hydrogen atom in its first excited state will, after a while, emit a photon and return into its ground state. The energy of the emitted photon is then given by the above expression, and hence its angular frequency is

$$\omega = \frac{mc^2\alpha^2}{2\hbar}\left(1 - \frac{1}{4}\right). \quad (1.8.16)$$

The corresponding wave length then is

$$\lambda = \frac{2\pi c}{\omega} = \frac{16\pi\hbar}{3\alpha^2 mc} = 1.2 \times 10^{-5}\text{cm}, \quad (1.8.17)$$

which is indeed seen in the ultraviolet part of the hydrogen spectrum.

Chapter 2

De Broglie Waves

2.1 Wave Packets

We have seen that light, which is classically described by wave equations, at the quantum level also displays particle features. On the other hand, matter, which is classically described as particles, shows wave behavior at the small scales of the quantum world. In fact, in quantum physics we should think of particle and wave as the same thing. In particular, the classical particle picture of an object with definite position and momentum always breaks down at very short distances. Schrödinger's formulation of quantum mechanics uses the wave function $\Psi(x, t)$ to describe the probability amplitude to find a particle at position x at time t . The Schrödinger equation is the wave equation for $\Psi(x, t)$. Here we discuss the quantum description of free nonrelativistic particles by de Broglie waves.

We have seen that quantum mechanically particles with momentum $p = \hbar k$ behave like waves with wave number k . The corresponding wave function given by

$$\Psi(x) = A \exp(ikx), \quad (2.1.1)$$

describes a particle with definite momentum. When we form the probability density we find

$$|\Psi(x)|^2 = |A|^2, \quad (2.1.2)$$

which is independent of the position x . Hence, all positions are equally probable and it is completely uncertain where the particle is. This is consistent with Heisenberg's uncertainty relation. To describe a particle that is localized in a particular region of space, one needs to go beyond a simple plane wave. In

general one can superimpose plane waves with different k values to construct a wave packet

$$\Psi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k) \exp(ikx). \quad (2.1.3)$$

Here $\Psi(k)$ is the amplitude of the plane wave with wave number k . The above equation describes a Fourier transformation and $\Psi(k)$ is called the Fourier transform of the wave function $\Psi(x)$. Note that $\Psi(k)$ and $\Psi(x)$ are two different functions. They are denoted by the same letter Ψ just because they are closely related via the Fourier transform.

The Fourier transform is a generalization of a Fourier series which can be used to describe periodic functions. Let us assume for a moment that $\Psi(x)$ is defined in the interval $[-L/2, L/2]$ and that it is periodically extended outside that interval. Then we can write it as a Fourier series

$$\Psi(x) = \frac{1}{L} \sum_k \Psi_k \exp(ikx). \quad (2.1.4)$$

Here Ψ_k is the amplitude of the mode with wave number k . Due to the periodicity requirement the wave lengths of the various modes must now be an integer fraction of L , and therefore the k values are restricted to

$$k = \frac{2\pi}{L}m, \quad m \in \mathbb{Z}. \quad (2.1.5)$$

The separation of two modes $dk = 2\pi/L$ becomes infinitesimal in the large volume limit $L \rightarrow \infty$. In this limit the sum over modes turns into an integral

$$\Psi(x) \rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi_k \exp(ikx). \quad (2.1.6)$$

When we identify $\Psi(k) = \Psi_k$ we recover the above expression for a Fourier transform. We can use our knowledge of the Fourier series to derive further relations for Fourier transforms. For example, we know that the amplitudes in a Fourier series are obtained as

$$\Psi_k = \int_{-L/2}^{L/2} dx \Psi(x) \exp(-ikx). \quad (2.1.7)$$

Let us convince ourselves that this is indeed correct by inserting eq.(2.1.4) in this expression

$$\begin{aligned} \Psi_k &= \int_{-L/2}^{L/2} dx \frac{1}{L} \sum_{k'} \Psi_{k'} \exp(ik'x) \exp(-ikx) \\ &= \sum_{k'} \Psi_{k'} \frac{1}{L} \int_{-L/2}^{L/2} dx \exp(i(k' - k)x) = \sum_{k'} \Psi_{k'} \delta_{k,k'}. \end{aligned} \quad (2.1.8)$$

Here we have identified the Kronecker δ -function

$$\delta_{k',k} = \frac{1}{L} \int_{-L/2}^{L/2} dx \exp(i(k' - k)x), \quad (2.1.9)$$

which is 1 for $k' = k$ and 0 otherwise. Then indeed the above sum returns Ψ_k . Let us now take the $L \rightarrow \infty$ limit of eq.(2.1.7) in which we get

$$\Psi_k \rightarrow \int_{-\infty}^{\infty} dx \Psi(x) \exp(-ikx). \quad (2.1.10)$$

Again identifying $\Psi(k) = \Psi_k$ we find the expression for an inverse Fourier transform

$$\Psi(k) = \int_{-\infty}^{\infty} dx \Psi(x) \exp(-ikx). \quad (2.1.11)$$

Let us check the validity of this expression in the same way we just checked eq.(2.1.7). We simply insert eq.(2.1.3) into eq.(2.1.11)

$$\begin{aligned} \Psi(k) &= \int_{-\infty}^{\infty} dx \frac{1}{2\pi} \int_{-\infty}^{\infty} dk' \Psi(k') \exp(ik'x) \exp(-ikx) \\ &= \int_{-\infty}^{\infty} dk' \Psi(k') \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(i(k' - k)x) \\ &= \int_{-\infty}^{\infty} dk' \Psi(k') \delta(k' - k). \end{aligned} \quad (2.1.12)$$

Here we have identified

$$\delta(k' - k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(i(k' - k)x), \quad (2.1.13)$$

which is known as the Dirac δ -function. It is the continuum analog of the discrete Kronecker δ -function. The Dirac δ -function has the remarkable property that

$$\int_{-\infty}^{\infty} dk' \Psi(k') \delta(k' - k) = \Psi(k). \quad (2.1.14)$$

This already follows from the fact that the limit $L \rightarrow \infty$ led to eq.(2.1.11).

As we have learned, $|\Psi(x)|^2$ is the probability density to find the particle at position x . Consequently, $|\Psi(x)|^2 dx$ is the probability to find it in an interval dx around the point x . The integral

$$\int_{-\infty}^{\infty} dx |\Psi(x)|^2 = 1 \quad (2.1.15)$$

therefore is the probability to find the particle anywhere in space. This total probability should be normalized to one. The normalization condition for $\Psi(x)$ implies a normalization condition on $\Psi(k)$ which can be derived simply by inserting eq.(2.1.3) into eq.(2.1.15) such that

$$\begin{aligned} \int_{-\infty}^{\infty} dx \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k)^* \exp(-ikx) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk' \Psi(k') \exp(ik'x) &= \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \Psi(k)^* \Psi(k') \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(i(k' - k)x) &= \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k)^* \Psi(k). \end{aligned} \quad (2.1.16)$$

We have again identified the Dirac δ -function and hence obtain for the normalization condition in momentum space

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk |\Psi(k)|^2 = 1. \quad (2.1.17)$$

It is important that the wave function in momentum space is properly normalized to 1 only if the integral of $|\Psi(k)|^2$ is divided by 2π . Hence, $|\Psi(k)|^2/2\pi$ is the probability density to find our particle with wave number k (or equivalently with momentum $p = \hbar k$).

2.2 Expectation Values

Since $|\Psi(x)|^2 dx$ is the probability to find our particle in an interval dx around the position x we can compute the expectation value of the position, i.e. the average over a large number of position measurements, as

$$\langle x \rangle = \int_{-\infty}^{\infty} dx x |\Psi(x)|^2. \quad (2.2.1)$$

Similarly we obtain

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} dx x^2 |\Psi(x)|^2. \quad (2.2.2)$$

The variance Δx of the position is given by

$$\begin{aligned} (\Delta x)^2 &= \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle \\ &= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 = \langle x^2 \rangle - \langle x \rangle^2. \end{aligned} \quad (2.2.3)$$

Similarly, one defines expectation values of momenta (or k values) as

$$\langle k \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk k |\Psi(k)|^2, \quad \langle k^2 \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk k^2 |\Psi(k)|^2, \quad (2.2.4)$$

and the variance of the momentum is given by

$$\Delta k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}. \quad (2.2.5)$$

One can also derive $\langle k \rangle$ directly from $\Psi(x)$. This follows from

$$\begin{aligned} & \int_{-\infty}^{\infty} dx \Psi(x)^* (-i\partial_x) \Psi(x) = \\ & \int_{-\infty}^{\infty} dx \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k)^* \exp(-ikx) (-i\partial_x) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk' \Psi(k') \exp(ik'x) = \\ & \int_{-\infty}^{\infty} dx \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k)^* \exp(-ikx) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk' \Psi(k') k' \exp(ik'x) = \\ & \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \Psi(k)^* \Psi(k') k' \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp(i(k' - k)x) = \\ & \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \Psi(k)^* \Psi(k') k' \delta(k' - k) \\ & \frac{1}{2\pi} \int_{-\infty}^{\infty} dk k |\Psi(k)|^2 = \langle k \rangle. \end{aligned} \quad (2.2.6)$$

Here ∂_x is the derivative with respect to x . Similarly, one finds

$$\int_{-\infty}^{\infty} dx \Psi(x)^* (-i\partial_x)^2 \Psi(x) = \langle k^2 \rangle. \quad (2.2.7)$$

In fact, in quantum mechanics observables like x and k are generally described by operators, and $-i\partial_x$ is the operator that represents k in coordinate space. Similarly, the operator $i\partial_k$ is the operator that describes x in momentum space.

2.3 Heisenberg's Uncertainty Relation

We are now in a position to state the Heisenberg uncertainty relation. It takes the form

$$\Delta x \Delta k \geq \frac{1}{2}, \quad (2.3.1)$$

or equivalently $\Delta x \Delta p \geq \hbar/2$. Gaussian wave packets have a minimal uncertainty product, i.e. for them $\Delta x \Delta k = 1/2$. The corresponding wave function takes the form

$$\Psi(x) = A \exp\left(-\frac{x^2}{2a^2}\right), \quad (2.3.2)$$

where A follows from the normalization condition. For a Gaussian wave packet one finds $\langle x \rangle = 0$ and $\langle x^2 \rangle = a^2/2$ such that $\Delta x = a/\sqrt{2}$. Performing the Fourier transform to momentum space one finds

$$\Psi(k) = B \exp\left(-\frac{a^2 k^2}{2}\right), \quad (2.3.3)$$

where B again follows from the normalization condition. In momentum space we find $\langle k \rangle = 0$ and $\langle k^2 \rangle = 1/2a^2$ such that $\Delta k = 1/\sqrt{2}a$. Hence, for a Gaussian wave packet indeed $\Delta x \Delta k = 1/2$.

Let us now prove Heisenberg's relation for general wave packets. For that purpose we construct the function

$$\Phi = \partial_x \Psi + \alpha x \Psi + \beta \Psi. \quad (2.3.4)$$

Here we choose α to be real, while β may be complex. The idea of the proof is to evaluate the integral

$$I = \int_{-\infty}^{\infty} dx |\Phi|^2 \geq 0, \quad (2.3.5)$$

which is positive by construction, and then vary α and β to find the minimum value of that integral. The resulting relation will turn out to be Heisenberg's inequality. One obtains

$$\begin{aligned} I &= \int_{-\infty}^{\infty} dx (\partial_x \Psi^* + \alpha x \Psi^* + \beta^* \Psi^*)(\partial_x \Psi + \alpha x \Psi + \beta \Psi) \\ &= \int_{-\infty}^{\infty} dx (\partial_x \Psi^* \partial_x \Psi + \partial_x \Psi^* \alpha x \Psi + \partial_x \Psi^* \beta \Psi + \alpha x \Psi^* \partial_x \Psi \\ &\quad + \alpha x \Psi^* \alpha x \Psi + \alpha x \Psi^* \beta \Psi + \beta^* \Psi^* \partial_x \Psi + \beta^* \Psi^* \alpha x \Psi + \beta^* \Psi^* \beta \Psi) \\ &= \int_{-\infty}^{\infty} dx (-\Psi^* \partial_x^2 \Psi - \alpha |\Psi|^2 - \Psi^* \alpha x \partial_x \Psi - \Psi^* \beta \partial_x \Psi + \alpha x \Psi^* \partial_x \Psi \\ &\quad + \alpha^2 x^2 |\Psi|^2 + \alpha \beta x |\Psi|^2 + \beta^* \Psi^* \partial_x \Psi + \alpha \beta^* x |\Psi|^2 + |\beta|^2 |\Psi|^2) \\ &= \langle k^2 \rangle - \alpha + 2\beta_i \langle k \rangle + \alpha^2 \langle x^2 \rangle + 2\alpha \beta_r \langle x \rangle + \beta_r^2 + \beta_i^2. \end{aligned} \quad (2.3.6)$$

Here $\beta = \beta_r + i\beta_i$. Now we search for the minimum of the integral

$$\begin{aligned} \frac{dI}{d\alpha} &= -1 + 2\alpha \langle x^2 \rangle + 2\beta_r \langle x \rangle = 0, \\ \frac{dI}{d\beta_r} &= 2\alpha \langle x \rangle + 2\beta_r = 0, \\ \frac{dI}{d\beta_i} &= 2\langle k \rangle + 2\beta_i = 0. \end{aligned} \quad (2.3.7)$$

These equations imply

$$\alpha = \frac{1}{2(\Delta x)^2}, \quad \beta_r = -\frac{\langle x \rangle}{2(\Delta x)^2}, \quad \beta_i = -\langle k \rangle. \quad (2.3.8)$$

Inserting this back into eq.(2.3.6) yields

$$I = (\Delta k)^2 - \frac{1}{4(\Delta x)^2} \geq 0, \quad (2.3.9)$$

and hence

$$\Delta x \Delta k \geq \frac{1}{2}. \quad (2.3.10)$$

From the above derivation it follows that Heisenberg's inequality can be saturated, i.e. $\Delta x \Delta k = 1/2$ can be obtained, only if $\Phi = 0$. For the Gaussian wave packet

$$\Psi = A \exp\left(-\frac{x^2}{2a^2}\right), \quad (2.3.11)$$

we had $\langle x \rangle = 0$, $\langle k \rangle = 0$ and $\Delta x = a/\sqrt{2}$. Hence, in that case $\alpha = 1/a^2$ and $\beta_r = \beta_i = 0$ and indeed

$$\Phi = \partial_x \Psi + \alpha x \Psi + \beta \Psi = -\frac{x}{a^2} \Psi + \frac{x}{a^2} \Psi = 0. \quad (2.3.12)$$

2.4 Dispersion Relations

Classically, the electromagnetic field is described by Maxwell's wave equations which are consistent with special relativity. As a consequence, an electromagnetic plane wave in vacuum with space-time dependence $\exp(i(\vec{k} \cdot \vec{x} - \omega t))$ obeys the dispersion relation

$$\omega = |\vec{k}|c. \quad (2.4.1)$$

Using the quantum physics relations

$$E = \hbar\omega, \quad \vec{p} = \hbar\vec{k}, \quad (2.4.2)$$

for photons, this is indeed equivalent to the relativistic energy-momentum relation

$$E = |\vec{p}|c \quad (2.4.3)$$

for massless particles. For massive particles with rest mass m the corresponding relation takes the form

$$E = \sqrt{(mc^2)^2 + (|\vec{p}|c)^2}. \quad (2.4.4)$$

Here we will restrict ourselves to the nonrelativistic limit of small momenta or large masses, such that

$$E = mc^2 + \frac{|\vec{p}|^2}{2m}. \quad (2.4.5)$$

The rest energy mc^2 represents a constant (momentum-independent) energy shift, which in nonrelativistic physics is usually absorbed in a redefinition of the energy, and hence

$$E = \frac{|\vec{p}|^2}{2m}. \quad (2.4.6)$$

De Broglie was first to postulate the general validity of eq.(2.4.2) not just for photons but for any free massive particle as well. This immediately implies that for a free nonrelativistic particle the dispersion relation reads

$$\omega = \frac{|\vec{p}|^2}{2m\hbar} = \frac{\hbar|\vec{k}|^2}{2m}. \quad (2.4.7)$$

Indeed, in Schrödinger's wave mechanics a free particle with momentum \vec{p} is described by the wave function

$$\Psi(\vec{x}, t) = A \exp(i(\vec{k} \cdot \vec{x} - \frac{\hbar|\vec{k}|^2}{2m}t)). \quad (2.4.8)$$

The time evolution of a 3-d wave packet is then determined by

$$\begin{aligned} \Psi(\vec{x}, t) &= \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k}) \exp(i(\vec{k} \cdot \vec{x} - \frac{\hbar|\vec{k}|^2}{2m}t)) \\ &= \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k}, t) \exp(i\vec{k} \cdot \vec{x}). \end{aligned} \quad (2.4.9)$$

The time-dependent momentum space wave function takes the form

$$\Psi(\vec{k}, t) = \Psi(\vec{k}) \exp(-i\frac{\hbar|\vec{k}|^2}{2m}t). \quad (2.4.10)$$

In particular, the probability density in momentum space

$$\frac{1}{2\pi} |\Psi(\vec{k}, t)|^2 = \frac{1}{2\pi} |\Psi(\vec{k})|^2, \quad (2.4.11)$$

is time-independent. This is typical for the propagation of a free particle. Its momentum is a conserved quantity, and therefore the probability distribution of momenta also does not change with time. The position of the particle, on the other hand, is obviously not conserved, since it is moving with a certain velocity. Let us now calculate the time evolution of a 1-d wave function in coordinate

space, first approximately for a wave packet that has a narrow distribution $\Psi(k)$ in momentum space centered around a wave number k_0 . In that case we can approximate the dispersion relation by

$$\omega(k) = \omega(k_0) + \frac{d\omega}{dk}(k_0)(k - k_0) + \mathcal{O}((k - k_0)^2). \quad (2.4.12)$$

Neglecting the quadratic terms we find

$$\begin{aligned} \Psi(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k) \exp(i(kx - \omega(k)t)) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k) \exp(i(kx - (\omega(k_0) + \frac{d\omega}{dk}(k_0)(k - k_0))t)) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \Psi(k) \exp(ik(x - \frac{d\omega}{dk}(k_0)t)) \\ &\quad \times \exp(-i(\omega(k_0) - \frac{d\omega}{dk}(k_0)k_0)t) \\ &= \Psi(x - \frac{d\omega}{dk}(k_0)t, 0) \exp(-i(\omega(k_0) - \frac{d\omega}{dk}(k_0)k_0)t). \end{aligned} \quad (2.4.13)$$

Up to a phase factor the wave function is just the one at time $t = 0$ shifted in space by $\frac{d\omega}{dk}(k_0)t$, i.e. the wave packet is moving with the so-called group velocity

$$\frac{d\omega}{dk}(k_0) = \frac{\hbar k_0}{m}, \quad (2.4.14)$$

which indeed is the velocity of a nonrelativistic particle with momentum $\hbar k_0$. Still, it is important to note that the above calculation is only approximate because we have neglected quadratic terms in the dispersion relation. As a consequence, our wave packet traveled undistorted like an electromagnetic wave in vacuum. Once the quadratic terms are included, the wave packet will also change its shape because the different wave number components travel with different velocities. This is more like an electromagnetic wave traveling in a medium. Unfortunately, the spreading of a quantum mechanical wave packet is difficult to work out for a general wave packet. Therefore we restrict ourselves to Gaussian wave packets in the following.

2.5 Spreading of a Gaussian Wave Packet

Let us consider a properly normalized Gaussian wave packet

$$\Psi(k) = \sqrt{2a\sqrt{\pi}} \exp\left(-\frac{a^2(k - k_0)^2}{2}\right), \quad (2.5.1)$$

centered around k_0 in momentum space. The corresponding wave function in coordinate space is given by

$$\begin{aligned}
\Psi(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \sqrt{2a\sqrt{\pi}} \exp\left(-\frac{a^2(k-k_0)^2}{2}\right) \exp\left(i(kx - \frac{\hbar k^2}{2m}t)\right) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \sqrt{2a\sqrt{\pi}} \exp\left(-\frac{a^2 k^2}{2}\right) \\
&\quad \times \exp\left(i((k+k_0)x - \frac{\hbar(k+k_0)^2}{2m}t)\right) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \sqrt{2a\sqrt{\pi}} \exp\left(-\left(a^2 + i\frac{\hbar t}{m}\right)\frac{k^2}{2}\right) \exp\left(ik\left(x - \frac{\hbar k_0}{m}t\right)\right) \\
&\quad \times \exp\left(i(k_0 x - \frac{\hbar k_0^2}{2m}t)\right). \tag{2.5.2}
\end{aligned}$$

At this point we use a result from before

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk \sqrt{2a\sqrt{\pi}} \exp\left(-\frac{a^2 k^2}{2}\right) \exp(ikx) = \frac{1}{\sqrt{a\sqrt{\pi}}} \exp\left(-\frac{x^2}{2a^2}\right). \tag{2.5.3}$$

We can therefore read off

$$\Psi(x, t) = \sqrt{\frac{a}{a'^2\sqrt{\pi}}} \exp\left(-\frac{x'^2}{2a'^2}\right) \exp\left(i(k_0 x - \frac{\hbar k_0^2}{2m}t)\right), \tag{2.5.4}$$

where $x' = x - \hbar k_0 t/m$, i.e. the particle moves with velocity $\hbar k_0/m$ and

$$a'^2 = a^2 + i\frac{\hbar t}{m} \Rightarrow \frac{1}{a'^2} = \frac{a^2 - i\hbar t/m}{a^4 + (\hbar t/m)^2}. \tag{2.5.5}$$

Inserting this in the above result and squaring the wave function we obtain

$$|\Psi(x, t)|^2 = \frac{a}{\sqrt{a^4 + (\hbar t/m)^2}\sqrt{\pi}} \exp\left(-\frac{(x - \hbar k_0 t/m)^2 a^2}{a^4 + (\hbar t/m)^2}\right). \tag{2.5.6}$$

The width of the wave packet in coordinate space increases with time according to

$$a(t) = \sqrt{a^2 + \left(\frac{\hbar t}{ma}\right)^2}. \tag{2.5.7}$$

At $t = 0$ we have $a(t) = a$ and at very large times we have $a(t) = \hbar t/ma$. The speed of spreading of the wave packet is \hbar/ma , i.e. a packet that was very localized initially (had a small a) will spread very fast. This is understandable because a spatially localized wave packet (with a small $\Delta x = a/\sqrt{2}$) necessarily has a large spread $\Delta k = 1/a\sqrt{2}$ in momentum space. Hence, the various wave number

contributions to the packet travel with very different speeds and consequently the wave packet spreads out in coordinate space. One should not think of the spread Δx as the size of our quantum mechanical particle. Indeed, we have assumed the particle to be point-like. The increasing spread $\Delta x(t) = a(t)/\sqrt{2}$ simply tells us that with time we know less and less about where the particle is. This is simply because it has traveled with a velocity that is picked from the given Gaussian probability distribution.

Chapter 3

The Schrödinger Equation

3.1 From Wave Packets to the Schrödinger Equation

In the last chapter we have learned how to deal with wave functions describing free non-relativistic particles. In particular, these are entirely determined by an initial distribution $\Psi(\vec{k})$ in momentum space. The time evolution of the wave function in 3-d coordinate space is then given by

$$\Psi(\vec{x}, t) = \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k}) \exp\left(i\left(\vec{k} \cdot \vec{x} - \frac{\hbar|\vec{k}|^2}{2M}t\right)\right). \quad (3.1.1)$$

Writing

$$\Psi(\vec{x}, t) = \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k}, t) \exp(i\vec{k} \cdot \vec{x}), \quad (3.1.2)$$

we can thus identify

$$\Psi(\vec{k}, t) = \Psi(\vec{k}) \exp\left(-i\frac{\hbar|\vec{k}|^2}{2M}t\right), \quad (3.1.3)$$

such that $|\Psi(\vec{k}, t)|^2 = |\Psi(\vec{k})|^2$ is time-independent. Next we look at the space and time dependence of $\Psi(\vec{x}, t)$ by taking the following derivatives

$$\begin{aligned} i\hbar\partial_t\Psi(\vec{x}, t) &= \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k}) \frac{\hbar^2|\vec{k}|^2}{2M} \exp\left(i\left(\vec{k} \cdot \vec{x} - \frac{\hbar|\vec{k}|^2}{2M}t\right)\right), \\ \Delta\Psi(\vec{x}, t) &= \frac{1}{(2\pi)^3} \int d^3k \Psi(\vec{k})(i\vec{k})^2 \exp\left(i\left(\vec{k} \cdot \vec{x} - \frac{\hbar|\vec{k}|^2}{2M}t\right)\right). \end{aligned} \quad (3.1.4)$$

Here Δ is the Laplacian. One thus finds

$$i\hbar\partial_t\Psi(\vec{x}, t) = -\frac{\hbar^2}{2M}\Delta\Psi(\vec{x}, t), \quad (3.1.5)$$

which is the Schrödinger equation for a free particle. By our construction it is clear that the Schrödinger equation resulted just from the non-relativistic dispersion relation $E = |\vec{p}|^2/2M$. Indeed

$$-\hbar^2\Delta = (-i\hbar\vec{\nabla})^2, \quad (3.1.6)$$

is nothing but the square of the momentum operator

$$\vec{p} = -i\hbar\vec{\nabla} = -i\hbar(\partial_x, \partial_y, \partial_z). \quad (3.1.7)$$

Now it is straightforward to include interactions of the particle with an external potential $V(\vec{x})$. The energy of the classical problem then consists of kinetic and potential energy

$$E = \frac{|\vec{p}|^2}{2M} + V(\vec{x}). \quad (3.1.8)$$

Hence, we should include $V(\vec{x})$ also in the Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{x}, t) = -\frac{\hbar^2}{2M}\Delta\Psi(\vec{x}, t) + V(\vec{x})\Psi(\vec{x}, t). \quad (3.1.9)$$

In fact, this is the equation that Schrödinger wrote down originally. In quantum mechanics the total energy is represented by the operator

$$H = -\frac{\hbar^2}{2M}\Delta + V(\vec{x}), \quad (3.1.10)$$

which is called the Hamilton operator or Hamiltonian. The kinetic part of the Hamiltonian acts on the wave function by taking its second derivative. The potential part, on the other hand, just multiplies the wave function with $V(\vec{x})$. The Schrödinger equation is the wave equation for the quantum mechanical wave function, i.e. it determines the time evolution of $\Psi(\vec{x}, t)$ once we specify an initial distribution $\Psi(\vec{x}, 0)$. Hence, Schrödinger's equation plays the same role in quantum mechanics that Maxwell's equations play in electromagnetism.

3.2 Conservation of Probability

We have learned that $\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2$ is the probability density to find the particle at position \vec{x} at time t . For this interpretation of the wave function to

be consistent we must demand that the wave function is normalized, i.e.

$$\int d^3x \rho(\vec{x}, t) = \int d^3x |\Psi(\vec{x}, t)|^2 = 1, \quad (3.2.1)$$

simply because it should always be possible to find the particle somewhere in space. In particular, this should be so at any moment in time. In other words, the total probability should be conserved and should be equal to 1 at all times. Let us calculate the change of the probability density with time

$$\begin{aligned} i\hbar\partial_t\rho(\vec{x}, t) &= i\hbar\partial_t|\Psi(\vec{x}, t)|^2 = i\hbar\Psi(\vec{x}, t)\partial_t\Psi(\vec{x}, t)^* + i\hbar\Psi(\vec{x}, t)^*\partial_t\Psi(\vec{x}, t) \\ &= -\Psi(\vec{x}, t) \left[-\frac{\hbar^2}{2M}\Delta + V(\vec{x}) \right] \Psi(\vec{x}, t)^* + \Psi(\vec{x}, t)^* \left[-\frac{\hbar^2}{2M}\Delta + V(\vec{x}) \right] \Psi(\vec{x}, t) \\ &= \frac{\hbar^2}{2M} [\Psi(\vec{x}, t)\Delta\Psi(\vec{x}, t)^* - \Psi(\vec{x}, t)^*\Delta\Psi(\vec{x}, t)] = -i\hbar\vec{\nabla} \cdot \vec{j}(\vec{x}, t). \end{aligned} \quad (3.2.2)$$

Here we have introduced the probability current density

$$\vec{j}(\vec{x}, t) = \frac{i\hbar}{2M} \left[\Psi(\vec{x}, t)\vec{\nabla}\Psi(\vec{x}, t)^* - \Psi(\vec{x}, t)^*\vec{\nabla}\Psi(\vec{x}, t) \right]. \quad (3.2.3)$$

Indeed, taking the divergence of the current, we obtain

$$\begin{aligned} -i\hbar\vec{\nabla} \cdot \vec{j}(\vec{x}, t) &= \frac{\hbar^2}{2M} \vec{\nabla} \cdot \left[\Psi(\vec{x}, t)\vec{\nabla}\Psi(\vec{x}, t)^* - \Psi(\vec{x}, t)^*\vec{\nabla}\Psi(\vec{x}, t) \right] \\ &= \frac{\hbar^2}{2M} \left[\vec{\nabla}\Psi(\vec{x}, t) \cdot \vec{\nabla}\Psi(\vec{x}, t)^* + \Psi(\vec{x}, t)\Delta\Psi(\vec{x}, t)^* \right. \\ &\quad \left. - \vec{\nabla}\Psi(\vec{x}, t)^* \cdot \vec{\nabla}\Psi(\vec{x}, t) - \Psi(\vec{x}, t)^*\Delta\Psi(\vec{x}, t) \right] \\ &= \frac{\hbar^2}{2M} [\Psi(\vec{x}, t)\Delta\Psi(\vec{x}, t)^* - \Psi(\vec{x}, t)^*\Delta\Psi(\vec{x}, t)]. \end{aligned} \quad (3.2.4)$$

By construction we thus have

$$\partial_t\rho(\vec{x}, t) + \vec{\nabla} \cdot \vec{j}(\vec{x}, t) = 0. \quad (3.2.5)$$

This is a continuity equation that guarantees probability conservation. It relates the probability density $\rho(\vec{x}, t)$ to the probability current density $\vec{j}(\vec{x}, t)$ in the same way in which electric charge and current density are related in electromagnetism. In that case, charge conservation follows from Maxwell's equations. Here probability conservation follows from the Schrödinger equation.

Let us consider the probability to find the particle in a region G of space. Its rate of change with time is given by

$$\partial_t \int_G d^3x |\Psi(\vec{x}, t)|^2 = \int_G d^3x \partial_t\rho(\vec{x}, t) = - \int_{\partial G} d^2\vec{s} \cdot \vec{j}(\vec{x}, t). \quad (3.2.6)$$

Here we have used the Gauss law. The continuity equation tells us that the probability to find the particle in a region G can change only if there is a probability current flowing through the boundary ∂G of that region. Assuming that the probability current density $\vec{j}(\vec{x}, t)$ vanishes at spatial infinity immediately implies that the total probability remains constant.

3.3 The Time-Independent Schrödinger Equation

We have seen that the Schrödinger equation governs the time evolution of the quantum mechanical wave function. In many cases, however, one is interested in stationary solutions of the Schrödinger equation, for example, in the stationary states of an atom. Also it turns out that once one has found all stationary solutions of the Schrödinger equation, one can use them to construct a general time-dependent solution. In a stationary state the probability density $|\Psi(\vec{x}, t)|^2 = |\Psi(\vec{x})|^2$ itself is time-independent. This does not mean that $\Psi(\vec{x}, t)$ is time-independent. In fact, it may still have a time-dependent phase. Hence we make the ansatz

$$\Psi(\vec{x}, t) = \Psi(\vec{x}) \exp(i\varphi(t)). \quad (3.3.1)$$

Inserting this in the Schrödinger equation implies

$$-\hbar\partial_t\varphi(t)\Psi(\vec{x})\exp(i\varphi(t)) = \left[-\frac{\hbar^2}{2M}\Delta + V(\vec{x})\right]\Psi(\vec{x})\exp(i\varphi(t)). \quad (3.3.2)$$

Dividing this equation by $\exp(i\varphi(t))$ one finds that the right-hand side becomes time-independent. Hence, the left-hand side should also be time-independent and thus

$$-\hbar\partial_t\varphi(t) = E, \quad (3.3.3)$$

where E is a constant. Integrating the above equation one obtains

$$\varphi(t) = -\frac{1}{\hbar}Et, \quad (3.3.4)$$

and thus

$$\Psi(\vec{x}, t) = \Psi(\vec{x}) \exp\left(-\frac{i}{\hbar}Et\right). \quad (3.3.5)$$

The remaining equation for the time-independent wave function $\Psi(\vec{x})$ takes the form

$$\left[-\frac{\hbar^2}{2M}\Delta + V(\vec{x})\right]\Psi(\vec{x}) = E\Psi(\vec{x}). \quad (3.3.6)$$

This is the so-called time-independent Schrödinger equation. Identifying the Hamilton operator

$$H = -\frac{\hbar^2}{2M}\Delta + V(\vec{x}), \quad (3.3.7)$$

we can also write it as

$$H\Psi(\vec{x}) = E\Psi(\vec{x}). \quad (3.3.8)$$

As we will understand soon, this can be viewed as an eigenvalue problem for the operator H with the eigenvalue E and the eigenvector $\Psi(\vec{x})$. We can now calculate the expectation value of the Hamiltonian (the operator that describes the total energy). Using the time-independent Schrödinger equation we find

$$\int d^3x \Psi(\vec{x}, t)^* H\Psi(\vec{x}, t) = \int d^3x \Psi(\vec{x})^* H\Psi(\vec{x}) = \int d^3x E|\Psi(\vec{x})|^2 = E. \quad (3.3.9)$$

This clarifies the meaning of the constant E . It is nothing but the total energy.

Chapter 4

The Formal Structure of Quantum Mechanics

4.1 Wave Functions as Vectors in a Hilbert Space

As we have seen in the example of a free particle on a discretized circle with N points, the quantum mechanical wave function resembles an N -component vector, and the Hamilton operator takes the form of an $N \times N$ matrix. In the continuum limit $N \rightarrow \infty$ the dimension of the corresponding vector space diverges, and we end up in a so-called Hilbert space. Let us consider the space of all normalizable wave functions $\Psi(\vec{x})$. We can consider them as vectors in the Hilbert space, i.e. we can add them up and multiply them with complex factors such that, for example, $\lambda\Psi(\vec{x}) + \mu\Phi(\vec{x})$ is also an element of the Hilbert space. The scalar product of ordinary 3-component vectors $\langle a|b \rangle = \vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + a_3b_3$ has an analog in the Hilbert space. The scalar product of two wave functions is defined as

$$\langle \Phi|\Psi \rangle = \int d^3x \Phi(\vec{x})^* \Psi(\vec{x}). \quad (4.1.1)$$

For ordinary 3-component vectors the scalar product induces the norm $\|a\|^2 = \langle a|a \rangle = \vec{a} \cdot \vec{a} = a_1^2 + a_2^2 + a_3^2$ while for wave functions

$$\|\Psi\|^2 = \langle \Psi|\Psi \rangle = \int d^3x \Psi(\vec{x})^* \Psi(\vec{x}) = \int d^3x |\Psi(\vec{x})|^2. \quad (4.1.2)$$

The physical interpretation of the wave function as a probability amplitude implied the normalization condition $\|\Psi\|^2 = 1$, i.e. physical wave functions are

unit vectors in the Hilbert space. Two ordinary vectors are orthogonal when $\langle a|b\rangle = \vec{a} \cdot \vec{b} = 0$. Similarly, two wave functions are orthogonal when

$$\langle \Phi|\Psi\rangle = \int d^3x \Phi(\vec{x})^* \Psi(\vec{x}) = 0. \quad (4.1.3)$$

4.2 Observables as Hermitean Operators

In quantum mechanics physical observable quantities (so-called observables) are represented by operators. For example, we have seen that the momentum is represented by the derivative operator (times $-i\hbar$), while the energy is represented by the Hamilton operator. Also the angular momentum and other physical observables are represented by operators. As we have seen, we can think of these operators as matrices with infinitely many rows and columns acting on the infinite dimensional vectors in the Hilbert space. A general operator A that acts on a wave function $\Psi(\vec{x})$ generates a new vector $A\Psi(\vec{x})$ in the Hilbert space. We can, for example, evaluate the scalar product of that vector with any other wave function $\Phi(\vec{x})$ and obtain

$$\langle \Phi|A\Psi\rangle = \int d^3x \Phi(\vec{x})^* A\Psi(\vec{x}). \quad (4.2.1)$$

For example, the expectation value of an operator A is given by

$$\langle A\rangle = \langle \Psi|A\Psi\rangle = \int d^3x \Psi(\vec{x})^* A\Psi(\vec{x}). \quad (4.2.2)$$

We have also seen that eigenvalue problems play a central role in quantum mechanics. For example, the time-independent Schrödinger equation is the eigenvalue problem of the Hamilton operator. As we will soon learn, the eigenvalues of an operator are the only possible results of measurements of the corresponding physical observable. Since physical observables are real numbers, they are represented by operators with real eigenvalues. These operators are called Hermitean. To formally define Hermiticity we first need to define the Hermitean conjugate (or adjoint) A^\dagger of a general (not necessarily Hermitean) operator A . The operator A^\dagger is defined such that

$$\langle A^\dagger\Phi|\Psi\rangle = \langle \Phi|A\Psi\rangle, \quad (4.2.3)$$

for all wave functions $\Phi(\vec{x})$ and $\Psi(\vec{x})$. This means

$$\int d^3x (A^\dagger\Phi(\vec{x}))^* \Psi(\vec{x}) = \int d^3x \Phi(\vec{x})^* A\Psi(\vec{x}). \quad (4.2.4)$$

An operator is called Hermitean (or self-adjoint) if

$$A^\dagger = A, \quad (4.2.5)$$

which means

$$\int d^3x (A\Phi(\vec{x}))^*\Psi(\vec{x}) = \int d^3x \Phi(\vec{x})^*A\Psi(\vec{x}), \quad (4.2.6)$$

for all $\Phi(\vec{x})$ and $\Psi(\vec{x})$. For the momentum operator we have

$$\begin{aligned} \int d^3x (-i\hbar\vec{\nabla}\Phi(\vec{x}))^*\Psi(\vec{x}) &= \int d^3x i\hbar\vec{\nabla}\Phi(\vec{x})^*\Psi(\vec{x}) \\ &= \int d^3x \Phi(\vec{x})^*(-i\hbar\vec{\nabla})\Psi(\vec{x}), \end{aligned} \quad (4.2.7)$$

which shows that it is indeed Hermitean. Here we have used partial integration. There are no boundary terms because a normalizable wave function necessarily vanishes at infinity. It is obvious that a Hermitean operator always has real expectation values because

$$\langle A \rangle = \langle \Psi|A\Psi \rangle = \langle A^\dagger\Psi|\Psi \rangle = \langle A\Psi|\Psi \rangle = \int d^3x (A\Psi(\vec{x}))^*\Psi(\vec{x}) = \langle A \rangle^*. \quad (4.2.8)$$

4.3 Eigenvalues and Eigenvectors

Let us derive some facts about the eigenvalues and eigenvectors of Hermitean operators. We consider the eigenvalue problem of a Hermitean operator A

$$A\chi_n(\vec{x}) = a_n\chi_n(\vec{x}). \quad (4.3.1)$$

Here the a_n are the eigenvalues and the $\chi_n(\vec{x})$ are the corresponding eigenvectors (sometimes also called eigenfunctions). For example, A could be the Hamilton operator. Then the a_n would be the allowed energy values and $\chi_n(\vec{x})$ would be the corresponding time-independent part of the wave function. Let us consider the scalar product

$$\langle \chi_m|A\chi_n \rangle = \int d^3x \chi_m(\vec{x})^*A\chi_n(\vec{x}) = a_n \int d^3x \chi_m(\vec{x})^*\chi_n(\vec{x}) = a_n\langle \chi_m|\chi_n \rangle. \quad (4.3.2)$$

On the other hand, because A is Hermitean, we also have

$$\langle \chi_m|A\chi_n \rangle = \langle A\chi_m|\chi_n \rangle = \int d^3x (A\chi_m(\vec{x}))^*\chi_n(\vec{x}) = a_m^*\langle \chi_m|\chi_n \rangle. \quad (4.3.3)$$

Let us first consider the case $m = n$. Then comparing the above equations implies $a_n = a_n^*$, i.e. the eigenvalues of a Hermitean operator are always real. Now consider the case $m \neq n$ and use $a_m^* = a_m$. Then

$$(a_n - a_m)\langle\chi_m|\chi_n\rangle = 0. \quad (4.3.4)$$

Assuming that $a_n \neq a_m$ (non-degenerate eigenvalues) this implies

$$\langle\chi_m|\chi_n\rangle = 0, \quad (4.3.5)$$

i.e. the eigenvectors corresponding to two non-degenerate eigenvalues are orthogonal. For degenerate eigenvalues (i.e. for $a_n = a_m$) the corresponding eigenvectors are not always automatically orthogonal, but one can always find a set of orthogonal eigenvectors. Normalizing the eigenvectors one hence has the orthonormality relation

$$\langle\chi_m|\chi_n\rangle = \delta_{mn}. \quad (4.3.6)$$

4.4 Completeness

The eigenvectors of a Hermitean operator form an orthogonal basis of the Hilbert space. This basis is complete, i.e. every wave function in the Hilbert space can be written as a linear combination of eigenvectors

$$\Psi(\vec{x}) = \sum_n c_n \chi_n(\vec{x}). \quad (4.4.1)$$

The expansion coefficients are given by

$$c_n = \langle\chi_n|\Psi\rangle. \quad (4.4.2)$$

An equivalent way to express the completeness of the eigenvectors is

$$\sum_n \chi_n(\vec{x})^* \chi_n(\vec{x}') = \delta(\vec{x} - \vec{x}'), \quad (4.4.3)$$

where

$$\delta(\vec{x} - \vec{x}') = \delta(x_1 - x'_1)\delta(x_2 - x'_2)\delta(x_3 - x'_3) \quad (4.4.4)$$

is the 3-dimensional Dirac δ -function. Indeed one then finds

$$\begin{aligned} \sum_n c_n \chi_n(\vec{x}) &= \sum_n \langle\chi_n|\Psi\rangle \chi_n(\vec{x}) = \sum_n \int d^3x' \chi_n(\vec{x}')^* \Psi(\vec{x}') \chi_n(\vec{x}) \\ &= \int d^3x' \Psi(\vec{x}') \delta(\vec{x} - \vec{x}') = \Psi(\vec{x}). \end{aligned} \quad (4.4.5)$$

The normalization condition for the wave function implies

$$\begin{aligned}
\int d^3x |\Psi(\vec{x})|^2 &= \int d^3x \left(\sum_m c_m \chi_m(\vec{x}) \right)^* \sum_n c_n \chi_n(\vec{x}) \\
&= \int d^3x \sum_m c_m^* \chi_m(\vec{x})^* \sum_n c_n \chi_n(\vec{x}) \\
&= \sum_{m,n} c_m^* c_n \langle \chi_m | \chi_n \rangle = \sum_{m,n} c_m^* c_n \delta_{mn} = \sum_n |c_n|^2 = 1.
\end{aligned} \tag{4.4.6}$$

Let us now consider the expectation value of a Hermitean operator A in the state described by the above wave function $\Psi(\vec{x})$

$$\begin{aligned}
\langle A \rangle &= \int d^3x \Psi(\vec{x})^* A \Psi(\vec{x}) \\
&= \int d^3x \left(\sum_m c_m \chi_m(\vec{x}) \right)^* A \sum_n c_n \chi_n(\vec{x}) \\
&= \int d^3x \sum_m c_m^* \chi_m(\vec{x})^* \sum_n c_n A \chi_n(\vec{x}) \\
&= \int d^3x \sum_m c_m^* \chi_m(\vec{x})^* \sum_n c_n a_n \chi_n(\vec{x}) \\
&= \sum_{m,n} c_m^* c_n a_n \langle \chi_m | \chi_n \rangle = \sum_{m,n} c_m^* c_n a_n \delta_{mn} = \sum_n |c_n|^2 a_n.
\end{aligned} \tag{4.4.7}$$

The expectation value of A gets contributions from each eigenvalue with the weight $|c_n|^2$ that determines the contribution of the eigenvector $\chi_n(\vec{x})$ to the wave function $\Psi(\vec{x})$.

4.5 Ideal Measurements

An important part of the formal structure of quantum mechanics concerns the interpretations of the various objects in the theory, as well as their relation with real physical phenomena. Quantum mechanics makes predictions about the wave function, which contains information about probabilities. Such a prediction can be tested only by performing many measurements of identically prepared systems, and averaging over the results. The result of a specific single measurement is not predicted by the theory, only the probability to find a certain result is. Still, of course every single measurement gives a definite answer. Although the

outcome of a measurement is uncertain until it is actually performed, after the measurement we know exactly what has happened. In other words, by doing the measurement we generally increase our knowledge about at least certain aspects of the quantum system. This has drastic consequences for the state of this system. In classical physics the influence of a measurement device on an observed phenomenon is assumed to be arbitrarily small, at least in principle. In quantum physics, on the other hand, this is not possible. A measurement always influences the state of the physical system after the measurement. In a real measurement the interaction of the quantum system and the measuring device may be very complicated. Therefore we want to discuss ideal measurements, whose outcome is directly related to the objects in the theory that we have discussed before. Let us consider the following relations from before. First, we have an expansion of the physical state

$$\Psi(\vec{x}) = \sum_n c_n \chi_n(\vec{x}), \quad (4.5.1)$$

in terms of the eigenvectors $\chi_n(\vec{x})$ of an operator A describing an observable that we want to measure. The expansion coefficients are given by

$$c_n = \langle \chi_n | \Psi \rangle, \quad (4.5.2)$$

and they are normalized by

$$\sum_n |c_n|^2 = 1. \quad (4.5.3)$$

The expectation value of our observable — i.e. the average over the results of a large number of measurements — is given by

$$\langle A \rangle = \sum_n |c_n|^2 a_n, \quad (4.5.4)$$

where the a_n are the eigenvalues of the operator A . These equations suggest the following interpretation. The possible outcomes of a measurement of the observable A are just its eigenvalues a_n . The probability to find a particular a_n is given by $|c_n|^2$, which determines the contribution of the corresponding eigenvector $\chi_n(\vec{x})$ to the wave function $\Psi(\vec{x})$ that describes the state of the physical system. Note that the interpretation of the $|c_n|^2$ as probabilities for the results of measurements is consistent because they are correctly normalized and indeed summing over all n with the appropriate probability gives the correct expectation value $\langle A \rangle$.

It is important that the theory can be tested only by performing many measurements of identically prepared systems. This means before each measurement we have to restart our system from the same initial conditions. We cannot simply

perform a second measurement of the same system that we have just measured before because the first measurement will have influenced the state of that system. In general the state of the system after the first measurement is not described by $\Psi(\vec{x})$ any longer. Then what is the state after the measurement? Let us assume that the actual measurement of the observable A gave the result a_n (one of the eigenvalues of A). Then after the measurement we know exactly that the system is in a state in which the observable A has the value a_n . What was a possibility before the measurement, has been turned into reality by the measurement itself. We may say that our decision to perform the measurement has forced the quantum system to “make up its mind” and decide which a_n it wants to have. The system “picks” a value a_n from the allowed set of eigenvalues with the appropriate probability $|c_n|^2$. Once the value a_n has been measured the new state of the system is $\chi_n(\vec{x})$. In this state the only possible result of a measurement of A is indeed a_n . Would we decide to measure A again immediately after the first measurement, we would get the same result a_n with probability one. If we wait a while until we perform the second measurement, the system evolves from the new initial state $\chi_n(\vec{x})$ following the Schrödinger equation. In particular, it will no longer be in the state $\Psi(\vec{x})$. The possible results of a second measurement hence depend on the outcome of the first measurement.

The above concept of an ideal measurement has led to many discussions almost of philosophical nature. To define an ideal measurement we have made a clear cut between the quantum system we are interested in, and the measuring device. For example, the decision to perform the measurement has nothing to do with the quantum system itself, and is not influenced by it in any way. The Schrödinger equation that governs the time evolution of the quantum system does not know when we will decide to do the next measurement. It only describes the system between measurements. When a measurement is performed the ordinary time evolution is interrupted and the wave function $\Psi(\vec{x})$ is replaced by the eigenvector $\chi_n(\vec{x})$ corresponding to the eigenvalue a_n that was obtained in the actual measurement. In a real experiment it may not always be clear where one should make the cut between the quantum system and the observer of the phenomenon. In particular, should we not think of ourselves as part of a big quantum system — called the Universe? Then aren't our “decisions” to perform a particular measurement just consequences of solving the Schrödinger equation for the wave function of the Universe? If so, how should we interpret that wave function? After all, in the past our Universe has behaved in a specific way. Who has decided to make the “measurements” that forced the Universe to “make up its mind” how to behave? These questions show that the above interpretation of the formal structure of quantum mechanics does not allow us to think of quantum mechanics as a theory of everything including the observer. Fortunately, when we

observe concrete phenomena at microscopic scales it is not very important where we draw the line between the quantum system and the observer or the measuring device. Only when the microscopic physics has macroscopic consequences like in the thought experiments involving Schrödinger's cat, which may be in a quantum state of half alive and half dead, we are dragged into the above philosophical discussions. These discussions are certainly interesting and often very disturbing, and they may even mean that quantum mechanics is not the whole story when the physics of everything (the whole Universe) is concerned. From a practical point of view, however, the above assumptions about ideal measurements are not unreasonable, and the above interpretation of the formalism of quantum mechanics has led to a correct description of all experiments performed at the microscopic level.

4.6 Simultaneous Measurability and Commutators

As we have learned earlier, in quantum mechanics position and momentum of a particle cannot be measured simultaneously with arbitrary precision. This follows already from Heisenberg's uncertainty relation. Now that we know more about measurements we can ask which observables can be measured simultaneously with arbitrary precision. According to our discussion of ideal measurements this would require that the operators describing the two observables have the same set of eigenvectors. Then after a measurement of one observable that reduces the wave function to a definite eigenstate, a measurement of the other observable will also give a unique answer (not just a set of possible answers with various probabilities). The quantity that decides about simultaneous measurability of two observables is the so-called commutator of the corresponding Hermitean operators A and B

$$[A, B] = AB - BA. \quad (4.6.1)$$

The two observables are simultaneously measurable with arbitrary precision if

$$[A, B] = 0, \quad (4.6.2)$$

i.e. if the order in which the operators are applied does not matter. Let us prove that this is indeed equivalent to A and B having the same set of eigenvectors. We assume that the eigenvectors are the same and write the two eigenvalue problems as

$$A\chi_n(\vec{x}) = a_n\chi_n(\vec{x}), \quad B\chi_n(\vec{x}) = b_n\chi_n(\vec{x}). \quad (4.6.3)$$

Then indeed

$$\begin{aligned} [A, B]\chi_n(\vec{x}) &= (AB - BA)\chi_n(\vec{x}) = Ab_n\chi_n(\vec{x}) - Ba_n\chi_n(\vec{x}) \\ &= (a_nb_n - b_na_n)\chi_n(\vec{x}) = 0. \end{aligned} \quad (4.6.4)$$

Now let us proceed in the other direction by assuming $[A, B] = 0$. Then we want to show that A and B indeed have the same set of eigenvectors. First, we write very generally

$$A\chi_n(\vec{x}) = a_n\chi_n(\vec{x}), \quad B\phi_n(\vec{x}) = b_n\phi_n(\vec{x}). \quad (4.6.5)$$

Then

$$BA\phi_n(\vec{x}) = AB\phi_n(\vec{x}) = b_nA\phi_n(\vec{x}), \quad (4.6.6)$$

which implies that $A\phi_n(\vec{x})$ is an eigenvector of B with eigenvalue b_n . Assuming non-degenerate eigenvalues this implies

$$A\phi_n(\vec{x}) = \lambda\phi_n(\vec{x}), \quad (4.6.7)$$

i.e. up to a factor λ the vector $A\phi_n(\vec{x})$ is one of the $\phi_m(\vec{x})$. Then λ is an eigenvalue of A — say a_n — and hence indeed

$$\phi_m(\vec{x}) = \chi_n(\vec{x}). \quad (4.6.8)$$

Let us evaluate the commutator for position and momentum operators

$$[x_i, p_j]\Psi(\vec{x}) = -i\hbar(x_i\partial_j - \partial_jx_i)\Psi(\vec{x}) = i\hbar\delta_{ij}\Psi(\vec{x}), \quad (4.6.9)$$

such that

$$[x_i, p_j] = i\hbar\delta_{ij}. \quad (4.6.10)$$

This is consistent with Heisenberg's uncertainty relation. The commutator does not vanish, and hence position and momentum of a particle (or more precisely their components in the same direction) cannot be measured simultaneously with arbitrary precision. Can energy and momentum always be measured simultaneously? To answer that question we evaluate

$$[\vec{p}, H]\Psi(\vec{x}) = \left[\vec{p}, \frac{\vec{p}^2}{2m} + V(\vec{x})\right]\Psi(\vec{x}) = -i\hbar[\vec{\nabla}, V(\vec{x})]\Psi(\vec{x}) = -i\hbar(\vec{\nabla}V(\vec{x}))\Psi(\vec{x}), \quad (4.6.11)$$

and we obtain

$$[\vec{p}, H] = -i\hbar\vec{\nabla}V(\vec{x}) = i\hbar\vec{F}(\vec{x}). \quad (4.6.12)$$

Here $\vec{F}(\vec{x})$ is the force acting on the particle. Only if the force vanishes — i.e. only for a free particle — energy and momentum can be simultaneously measured with arbitrary precision.

4.7 Time Evolution

As we have just discussed, a measurement has an effect on the time evolution of a quantum system. In fact, immediately after the measurement it is in the eigenstate of the observable that corresponds to the eigenvalue that was actually measured. Between measurements the time evolution of a quantum system is determined by the time-dependent Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{x},t) = H\Psi(\vec{x},t). \quad (4.7.1)$$

We will now see that it is indeed sufficient to solve the time-independent Schrödinger equation

$$H\chi_n(\vec{x}) = E_n\chi_n(\vec{x}). \quad (4.7.2)$$

The information about the time-dependence can be extracted from there. In the above equation the $\chi_n(\vec{x})$ are eigenvectors of the Hamilton operator, i.e. wave functions of stationary states, and the E_n are the corresponding energy eigenvalues. Since H is Hermitean, its eigenvectors form a complete basis of orthonormal vectors in the Hilbert space. Consequently, we can write the time-dependent wave function

$$\Psi(\vec{x},t) = \sum_n c_n(t)\chi_n(\vec{x}), \quad (4.7.3)$$

as a linear combination of eigenvectors. Now, however, the coefficients $c_n(t)$ will be time-dependent. To figure out the time-dependence we plug the above expression into the time-dependent Schrödinger equation and obtain

$$i\hbar\sum_n \partial_t c_n(t)\chi_n(\vec{x}) = \sum_n c_n(t)H\chi_n(\vec{x}) = \sum_n c_n(t)E_n\chi_n(\vec{x}). \quad (4.7.4)$$

Since the eigenvectors are orthogonal to one another, we can identify their coefficients in a linear combination and conclude that

$$i\hbar\partial_t c_n(t) = c_n(t)E_n. \quad (4.7.5)$$

This implies

$$c_n(t) = c_n(0)\exp\left(-\frac{i}{\hbar}E_n t\right), \quad (4.7.6)$$

and hence

$$\Psi(\vec{x},t) = \sum_n c_n(0)\exp\left(-\frac{i}{\hbar}E_n t\right)\chi_n(\vec{x}). \quad (4.7.7)$$

Still, we need to determine $c_n(0)$. At $t = 0$ the above equation reads

$$\Psi(\vec{x}, 0) = \sum_n c_n(0) \chi_n(\vec{x}), \quad (4.7.8)$$

which implies

$$c_n(0) = \langle \chi_n | \Psi(0) \rangle. \quad (4.7.9)$$

In fact, solving the time-dependent Schrödinger equation requires to specify an initial condition for the wave function. We have to pick $\Psi(\vec{x}, 0)$ and then the Schrödinger equation determines $\Psi(\vec{x}, t)$ for any later time t . All we have to do is to decompose the initial wave function into a linear combination of eigenvectors of the Hamilton operator, and then

$$\Psi(\vec{x}, t) = \sum_n \langle \chi_n | \Psi(0) \rangle \exp\left(-\frac{i}{\hbar} E_n t\right) \chi_n(\vec{x}). \quad (4.7.10)$$

Chapter 5

Schrödinger Equation in One Dimension

5.1 A Particle in a Box

We now know enough quantum mechanics to go ahead and apply it to physical problems. In particular, we know how to interpret our results correctly, and how to relate them to actual measurements performed on real physical systems. Doing quantum mechanics essentially means solving the Schrödinger equation. As we have seen, for time-independent Hamilton operators the solution of the full time-dependent Schrödinger equation follows from that of the time-independent Schrödinger equation. Hence we will now concentrate on that equation. For general potentials it can be very difficult to solve the time-independent Schrödinger equation. In fact, in general one will not be able to find an analytic solution, and hence one must then use some approximate numerical method. Still, there are some potentials for which the Schrödinger equation can be solved analytically, among them several examples in one dimension. Obviously, it is much simpler to work in one rather than three dimensions. This is the main reason for us to study one dimensional problems in this chapter. We will learn techniques to solve the Schrödinger equation that will be useful also in three dimensions.

One of the simplest problems in quantum mechanics concerns a particle in a 1-dimensional box with perfectly reflecting walls. We can realize this situation by switching on a potential that is infinite everywhere outside the interval $[0, a]$, and zero in that interval. This has the effect that the wave function must vanish outside the interval, and has to go the zero at the two walls. The time-independent

Schrödinger equation inside the box takes the form

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} = E\Psi(x), \quad (5.1.1)$$

which must be solved with the boundary condition

$$\Psi(0) = \Psi(a) = 0. \quad (5.1.2)$$

We make the ansatz

$$\Psi(x) = A \sin(kx). \quad (5.1.3)$$

Then the boundary condition implies

$$k = \frac{\pi n}{a}, \quad (5.1.4)$$

and the Schrödinger equation yields

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2ma^2}. \quad (5.1.5)$$

As a consequence of the infinite potential on both sides of the interval the energy spectrum is entirely discrete. Also, the ground state energy

$$E = \frac{\hbar^2 \pi^2}{2ma^2} \quad (5.1.6)$$

— i.e. the lowest possible energy — is nonzero. This means that the particle in the box cannot be at rest. It always has a nonzero kinetic energy. This is a genuine quantum effect. Classically, there would be nothing wrong with a particle being at rest in a box, and indeed in the classical limit $\hbar \rightarrow 0$ the ground state energy goes to zero.

The fact that the ground state energy is nonzero already follows from the Heisenberg uncertainty relation

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (5.1.7)$$

For the particle in the box $\langle p \rangle = 0$, and hence $\Delta p = \sqrt{\langle p^2 \rangle}$. Also the spread of the wave function in x can at most be $a/2$ such that $\Delta x \leq a/2$. Therefore the kinetic energy of the particle is restricted by

$$\left\langle \frac{p^2}{2m} \right\rangle \geq \frac{\hbar^2}{2ma^2}, \quad (5.1.8)$$

which is consistent with the actual ground state energy.

Let us also use the semi-classical Bohr-Sommerfeld quantization for this problem. The quantization condition has the form

$$S[x_c(t)] + ET = 2\pi n\hbar, \quad (5.1.9)$$

where $S[x_c(t)]$ is the action of a periodic classical path with period T . The action is the time integral of the difference between kinetic and potential energy. Since there is no potential inside the box, the action is simply the integrated kinetic energy, which at the same time is the integrated total energy, which is conserved (time-independent). Therefore in this case

$$S[x_c(t)] = ET, \quad (5.1.10)$$

and thus

$$ET = \pi n\hbar. \quad (5.1.11)$$

Inside the box the particle moves with constant speed v bouncing back and forth between the two walls. The period of a corresponding classical path is

$$T = \frac{2a}{v}, \quad (5.1.12)$$

and the energy is

$$E = \frac{mv^2}{2}. \quad (5.1.13)$$

The quantization condition thus implies

$$ET = mav = \pi n\hbar \Rightarrow v = \frac{\pi n\hbar}{ma}, \quad (5.1.14)$$

and hence

$$E = \frac{mv^2}{2} = \frac{\hbar^2 \pi^2 n^2}{2ma^2}, \quad (5.1.15)$$

which is indeed the exact spectrum that we also obtained from the Schrödinger equation. As we will see, this is not always the case.

Let us now modify the above problem and make it more interesting. We still consider an infinite potential for all negative x , and a zero potential in the interval $[0, a]$, but we will let the potential have a finite value $V_0 > 0$ for $x > a$. Then already the semi-classical analysis get modified. In fact, a classical path with an energy larger than V_0 will no longer be confined to the interior of the box, but will simply go to infinity. Such a path is not periodic, and can therefore not be used in the Bohr-Sommerfeld quantization procedure. Consequently, the predicted energy spectrum now is

$$E = \frac{\hbar^2 \pi^2 n^2}{2ma^2}, \quad (5.1.16)$$

with the additional condition $E < V_0$. This means that the allowed values of n are limited to

$$n < \frac{\sqrt{2ma^2V_0}}{\hbar\pi}, \quad (5.1.17)$$

and only a finite set of discrete energy values exists. Let us see what the Schrödinger equation gives for the modified problem. First, let us again assume that $E < V_0$. Still, the wave function has to vanish at $x = 0$. However, it can now extend into the classically forbidden region $x > a$. In that region the equation takes the form

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} + V_0\Psi(x) = E\Psi(x), \quad (5.1.18)$$

such that

$$\frac{d^2\Psi(x)}{dx^2} = \frac{2m(V_0 - E)}{\hbar^2}\Psi(x). \quad (5.1.19)$$

We make the ansatz

$$\Psi(x) = B \exp(-\kappa x), \quad (5.1.20)$$

and obtain

$$\kappa = \pm \frac{\sqrt{2m(V_0 - E)}}{\hbar}. \quad (5.1.21)$$

Only the positive solution is physical because it corresponds to an exponentially decaying normalizable wave function. The other solution is exponentially rising, and thus not normalizable. Inside the box the solution of the Schrödinger equation still takes the form

$$\Psi(x) = A \sin(kx), \quad (5.1.22)$$

but now the wave function need not vanish at $x = a$. Still, both the wave function and its first derivative must be continuous at $x = a$. This implies

$$A \sin(ka) = B \exp(-\kappa a), \quad Ak \cos(ka) = -B\kappa \exp(-\kappa a), \quad (5.1.23)$$

and thus

$$\kappa = -k \cot(ka). \quad (5.1.24)$$

Also we can use the relations for k and κ and obtain

$$k^2 + \kappa^2 = \frac{2mE}{\hbar^2} + \frac{2m(V_0 - E)}{\hbar^2} = \frac{2mV_0}{\hbar^2}. \quad (5.1.25)$$

This equation describes a circle of radius $\sqrt{2mV_0}/\hbar$ in the k - κ -plane, which suggests a graphical method to solve for the energy spectrum. In fact, plotting

$\kappa = -k \cot(ka)$ as well as the circle yields the quantized energies as the intersections of the two curves. The condition for at least n intersections to exist is

$$\frac{\sqrt{2mV_0}}{\hbar} > \frac{\pi}{a} \left(n - \frac{1}{2} \right). \quad (5.1.26)$$

For $n = 1$ this implies

$$V_0 > \frac{\hbar^2 \pi^2}{8ma^2}. \quad (5.1.27)$$

The potential must be larger than this critical value in order to support at least one discrete energy level. A particle in that state is localized essentially inside the box, although it has a nonzero probability to be at $x > a$. Therefore, such a state is called a bound state. The above condition for the existence of at least one bound state is not identical with the one that follows from the semi-classical method. In that case we found at least one bound state if

$$V_0 > \frac{\hbar^2 \pi^2}{2ma^2}. \quad (5.1.28)$$

This shows that Bohr-Sommerfeld quantization does not always give the exact answer. Still, it gives the right answer in the limit of large quantum numbers n . For example, semi-classical quantization predicts that in order to support n bound states the strength of the potential must be

$$V_0 > \frac{\hbar^2 \pi^2 n^2}{2ma^2}, \quad (5.1.29)$$

which is consistent with the exact result of eq.(5.1.26) obtained from the Schrödinger equation for large n .

5.2 Parity

Symmetries are always very important in a physical problem. In fact, using a symmetry will often simplify the solution considerably. When we are dealing with potentials that are reflection symmetric, i.e. when

$$V(-x) = V(x), \quad (5.2.1)$$

the corresponding symmetry is called parity. The parity operator P acts on a wave function as

$$P\Psi(x) = \Psi(-x). \quad (5.2.2)$$

For reflection symmetric potentials the parity operator commutes with the Hamilton operator because

$$\begin{aligned}
 [P, H]\Psi(x) &= P\left[-\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + V(x)\Psi(x)\right] - \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\Psi(-x) \\
 &= -\frac{\hbar^2}{2m}\frac{d^2\Psi(-x)}{dx^2} + V(-x)\Psi(-x) \\
 &\quad + \frac{\hbar^2}{2m}\frac{d^2\Psi(-x)}{dx^2} - V(x)\Psi(-x) = 0.
 \end{aligned} \tag{5.2.3}$$

Since $[P, H] = 0$ both P and H have the same eigenvectors. Let us consider the eigenvalue problem for the parity operator

$$P\Psi(x) = \lambda\Psi(x). \tag{5.2.4}$$

Acting with the parity operator twice implies

$$P^2\Psi(x) = \lambda^2\Psi(x). \tag{5.2.5}$$

On the other hand

$$P^2\Psi(x) = P\Psi(-x) = \Psi(x), \tag{5.2.6}$$

such that $\lambda = \pm 1$, and thus

$$P\Psi(x) = \pm\Psi(x). \tag{5.2.7}$$

The eigenvectors with eigenvalue 1 are even functions

$$\Psi(-x) = \Psi(x), \tag{5.2.8}$$

while the eigenvectors with eigenvalue -1 are odd

$$\Psi(-x) = -\Psi(x). \tag{5.2.9}$$

Since the Hamilton operator has the same eigenvectors as P , we can assume that the wave functions of stationary states are either even or odd for reflection symmetric potentials. Making an appropriate ansatz for the wave function will simplify the solution of the Schrödinger equation. Of course, even without noticing the symmetry one will get the same answer, but in a more complicated way. It is important to note that symmetry of the potential ($V(-x) = V(x)$) does not imply symmetry of the wave function in the sense of $\Psi(-x) = \Psi(x)$. In fact, half of the solutions have $\Psi(-x) = -\Psi(x)$. This is understandable because only $|\Psi(x)|^2$ has the physical meaning of probability density, and in both cases $|\Psi(-x)|^2 = |\Psi(x)|^2$.

5.3 Simple Models for Atoms and Molecules

While real physics takes place in three dimensions, one dimensional models often shed some light on the physical mechanisms responsible for a certain physical phenomenon. Here we discuss a very simple 1-d model for atoms and molecules, that certainly oversimplifies the real 3-d physics, but still captures some essential properties of chemical binding. The model uses so-called δ -function potentials, which are easy to handle analytically. The δ -function potential models a positively charged ion that attracts a negatively charged electron, thus forming an electrically neutral atom. Of course, the interaction between real ions and electrons is via the long-ranged Coulomb potential, not via an ultra short-ranged δ -function potential. In this respect the model is certainly unrealistic. Still, as we will see, it has some truth in it.

Let us first write down the potential as

$$V(x) = -V_0 a \delta(x). \quad (5.3.1)$$

Here $V_0 > 0$ such that the potential is attractive. The parameter a has the dimension of a length to compensate for the dimension of the δ -function. As a consequence, V_0 has the dimension of an energy. The fact that the δ -function has the dimension of an inverse length already follows from its most important property

$$\int_{-\infty}^{\infty} dx f(x) \delta(x) = f(0), \quad (5.3.2)$$

which holds for any smooth function $f(x)$. Let us now write the Schrödinger equation with the above δ -function potential

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} - V_0 a \delta(x) \Psi(x) = E \Psi(x). \quad (5.3.3)$$

Integrating this equation over the interval $[-\epsilon, \epsilon]$ and using

$$\int_{-\epsilon}^{\epsilon} dx f(x) \delta(x) = f(0), \quad (5.3.4)$$

yields

$$-\frac{\hbar^2}{2m} \left[\frac{d\Psi(\epsilon)}{dx} - \frac{d\Psi(-\epsilon)}{dx} \right] = V_0 a \Psi(0). \quad (5.3.5)$$

Here we have also taken the limit $\epsilon \rightarrow 0$. The above equation implies a discontinuity in the first derivative of the wave function, that is proportional to the strength of the potential and to the value of the wave function at the location of the δ -function.

How can we solve the above Schrödinger equation? Because we are interested in binding of model electrons to model ions we restrict ourselves to bound state solutions with $E < 0$. For $x \neq 0$ the δ -function vanishes, and thus we are then dealing with a free particle Schrödinger equation. Its solutions with $E < 0$ are exponentially rising or falling. The rising solution is unphysical because it cannot be properly normalized and hence

$$\Psi(x) = A \exp(-\kappa|x|), \quad (5.3.6)$$

with

$$E = -\frac{\hbar^2 \kappa^2}{2m}. \quad (5.3.7)$$

We have used parity symmetry to motivate the above ansatz. Then by construction the wave function is automatically continuous at $x = 0$. However, we still must impose the appropriate discontinuity of the derivative of the wave function. We find

$$-\frac{\hbar^2}{2m} \left[\frac{d\Psi(\epsilon)}{dx} - \frac{d\Psi(-\epsilon)}{dx} \right] = \frac{\hbar^2}{2m} 2A\kappa = V_0 a \Psi(0) = V_0 a A \Rightarrow \kappa = \frac{V_0 a m}{\hbar^2}, \quad (5.3.8)$$

and hence

$$E = -\frac{V_0^2 a^2 m}{2\hbar^2}. \quad (5.3.9)$$

In our model this is the binding energy of an atom consisting of electron and ion (represented by the δ -function potential). There are also solutions with positive energy, but we are not interested in them at the moment.

Instead, we now want to make a model for a pair of ions that share a single electron. It is important to note that we do not consider two electrons yet. Up to now we have always dealt with the Schrödinger equation for a single particle, and we will continue to do so until we will talk about the hydrogen atom. We will see that the model electron that is shared by two model ions induces an attractive interaction (chemical binding) of two atoms. Hence, we are now considering a pair of δ -functions at a distance R such that

$$V(x) = -V_0 a \left[\delta\left(x - \frac{R}{2}\right) + \delta\left(x + \frac{R}{2}\right) \right]. \quad (5.3.10)$$

Again, parity symmetry allows us to consider even and odd wave functions separately. Since we are interested in $E < 0$, and especially in the ground state, we first consider even wave functions. We know that away from the points $x = \pm R/2$ we are dealing with the free particle Schrödinger equation. For $|x| > R/2$ normalizability again singles out the exponentially decaying solution, i.e. then

$$\Psi(x) = A \exp(-\kappa(R)|x|). \quad (5.3.11)$$

We explicitly account for the fact that now $\kappa(R)$ will depend on R . In particular, it will in general be different from the κ of a single δ -function potential. Of course, then also the energy

$$E(R) = -\frac{\hbar^2 \kappa(R)^2}{2m}, \quad (5.3.12)$$

will be a function of the separation R . For $|x| < R/2$ normalizability does not impose any restrictions, i.e. both exponentially rising and falling solutions are allowed. The even combination of the two is a hyperbolic cosine, such that for $|x| < R/2$

$$\Psi(x) = B \cosh(\kappa(R)x). \quad (5.3.13)$$

Note that $\kappa(R)$ has the same value as for $|x| > R/2$ because it is related to the energy just as before. To find the physical solution we first require continuity of the wave function at $x = R/2$. Parity symmetry then guarantees that the wave function is continuous also at $x = -R/2$. We demand

$$A \exp(-\kappa(R)R/2) = B \cosh(\kappa(R)R/2). \quad (5.3.14)$$

Also we must impose the appropriate discontinuity of the derivative of the wave function

$$\begin{aligned} & -\frac{\hbar^2}{2m} \left[\frac{d\Psi(R/2 + \epsilon)}{dx} - \frac{d\Psi(R/2 - \epsilon)}{dx} \right] \\ &= \frac{\hbar^2}{2m} \left[B\kappa(R) \sinh\left(\frac{\kappa(R)R}{2}\right) + A\kappa(R) \exp\left(-\frac{\kappa(R)R}{2}\right) \right] \\ &= V_0 a \Psi(0) = V_0 a A \exp\left(-\frac{\kappa(R)R}{2}\right), \end{aligned} \quad (5.3.15)$$

and hence

$$B \sinh\left(\frac{\kappa(R)R}{2}\right) = \left[\frac{2V_0 a m}{\hbar^2 \kappa(R)} - 1 \right] A \exp\left(-\frac{\kappa(R)R}{2}\right). \quad (5.3.16)$$

Combining the two conditions implies

$$\tanh\left(\frac{\kappa(R)R}{2}\right) = \frac{2V_0 a m}{\hbar^2 \kappa(R)} - 1. \quad (5.3.17)$$

Using a graphical method it is easy to see that this equation always has a solution. Furthermore, using $\tanh x \leq 1$ for $x > 0$ one obtains

$$\frac{2V_0 a m}{\hbar^2 \kappa(R)} - 1 \leq 1 \Rightarrow \kappa(R) \geq \frac{V_0 a m}{\hbar^2} = \kappa, \quad (5.3.18)$$

i.e. the value $\kappa(R)$ for two δ -functions at a distance R is always bigger than the value κ for a single δ -function. Hence

$$E(R) \leq E, \quad (5.3.19)$$

i.e. the energy of an electron in the field of two ions a distance R apart, is smaller than the binding energy of a single atom. In other words, it is energetically favorable for the two ions to share the electron. This is indeed the mechanism responsible for chemical binding. Of course, real molecules are more complicated than the one in our model, for example, because the direct Coulomb interaction between the two ions leads to a repulsive contribution as well. Although we have ignored this effect in our model, it still captures an essential quantum aspect of chemistry. Sharing an electron, in other words delocalizing it and thus spreading its wave function over a larger region in space, allows a molecule to have less energy than two separated atoms.

We have not yet discussed the parity odd solutions in the double δ -function potential. The appropriate ansatz is

$$\Psi(x) = \pm A \exp(-\kappa(R)x), \quad (5.3.20)$$

for $|x| \geq R/2$ and

$$\Psi(x) = B \sinh(\kappa(R)x), \quad (5.3.21)$$

for $|x| \leq R/2$. The matching conditions then take the form

$$A \exp(-\kappa(R)R/2) = B \sinh\left(\frac{\kappa(R)R}{2}\right), \quad (5.3.22)$$

and

$$B \cosh\left(\frac{\kappa(R)R}{2}\right) = \left[\frac{2V_0am}{\hbar^2\kappa(R)} - 1\right]A \exp\left(-\frac{\kappa(R)R}{2}\right), \quad (5.3.23)$$

such that now

$$\tanh\left(\frac{\kappa(R)R}{2}\right) = \left(\frac{2V_0am}{\hbar^2\kappa(R)} - 1\right)^{-1}. \quad (5.3.24)$$

Again, one can use a graphical method to solve this equation. A solution exists only if

$$\frac{R}{2} \geq \frac{\hbar^2}{2V_0am} \Rightarrow V_0 \geq \frac{\hbar^2}{amR}. \quad (5.3.25)$$

This condition is satisfied if the potential is strong enough, or if the two δ -functions are separated by a large distance. For too weak potentials no bound state exists besides the ground state. Again, we can use $\tanh x \leq 1$, such that now

$$\left(\frac{2V_0am}{\hbar^2\kappa(R)} - 1\right)^{-1} \leq 1 \Rightarrow \kappa(R) \leq \frac{V_0am}{\hbar^2} = \kappa, \quad (5.3.26)$$

i.e. in this energetically excited state the energy of the molecule is larger than that of two separated atoms. This is understandable because an odd wave function goes through zero between the ions, and consequently the electron is then more localized than in the case of two separated atoms. In chemistry this would correspond to an anti-binding orbital.

5.4 Shift Symmetry and Periodic Potentials

We have seen that parity considerations simplify calculations involving potentials with a reflection symmetry. In crystals a periodic structure of ions generates a periodic potential for the electrons

$$V(x + a) = V(x), \quad (5.4.1)$$

which is symmetric under translations by multiples of the lattice spacing a . In the next section we will study a simple model for electrons in a crystal, and the shift symmetry will simplify the solution of that model. The analog of the parity operator P in a translation invariant problem is the shift operator

$$T = \exp\left(\frac{i}{\hbar}pa\right), \quad (5.4.2)$$

where p is the momentum operator. The shift operator acts on a wave function as

$$T\Psi(x) = \Psi(x + a). \quad (5.4.3)$$

Let us consider the commutator of the shift operator with the Hamilton operator

$$\begin{aligned} [T, H]\Psi(x) &= T\left[-\frac{\hbar^2}{2m}\frac{d^2\Psi(x)}{dx^2} + V(x)\Psi(x)\right] - \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\Psi(x + a) \\ &= -\frac{\hbar^2}{2m}\frac{d^2\Psi(x + a)}{dx^2} + V(x + a)\Psi(x + a) \\ &+ \frac{\hbar^2}{2m}\frac{d^2\Psi(x + a)}{dx^2} + V(x)\Psi(x + a) = 0. \end{aligned} \quad (5.4.4)$$

Since $[T, H] = 0$ both T and H have the same eigenvectors. The eigenvalue problem for the shift operator takes the form

$$T\Psi(x) = \lambda\Psi(x). \quad (5.4.5)$$

The Hermitean conjugate of the shift operator is

$$T^\dagger = \exp\left(-\frac{i}{\hbar}pa\right), \quad (5.4.6)$$

which describes a shift backwards, i.e. $T^\dagger\Psi(x) = \Psi(x - a)$. The eigenvalue problem of T^\dagger takes the form

$$T^\dagger\Psi(x) = \lambda^*\Psi(x), \quad (5.4.7)$$

with the same λ and $\Psi(x)$ as in the eigenvalue problem of T . Acting with $T^\dagger T$ implies

$$T^\dagger T \Psi(x) = T^\dagger \lambda \Psi(x) = \lambda \lambda^* \Psi(x) = |\lambda|^2 \Psi(x). \quad (5.4.8)$$

On the other hand

$$T^\dagger T \Psi(x) = T^\dagger \Psi(x+a) = \Psi(x), \quad (5.4.9)$$

such that $|\lambda|^2 = 1$ and hence

$$\lambda = \exp(iqa), \quad (5.4.10)$$

is a phase factor. Here the phase was arbitrarily written as qa . One obtains

$$T \Psi(x) = \Psi(x+a) = \exp(iqa) \Psi(x), \quad (5.4.11)$$

i.e. the eigenvectors of T are periodic up to a phase factor. It is important to note that $\Psi(x)$ itself is not periodic. Only the probability density is periodic, i.e.

$$|\Psi(x+a)|^2 = |\Psi(x)|^2. \quad (5.4.12)$$

5.5 A Simple Model for Electrons in a Crystal

We have seen that a δ -function potential provides a simple model for an ion in an atom or molecule. The same model can also describe the ions in a crystal with lattice spacing a , which generate a periodic potential

$$V(x) = -V_0 a \sum_{n \in \mathbb{Z}} \delta(x - na), \quad (5.5.1)$$

for an electron. Let us look for positive energy solutions of the Schrödinger equation with this potential. We make an ansatz for the wave function in each interval $[na, (n+1)a]$ separately. For $x \in [na, (n+1)a]$ we write

$$\Psi(x) = A_n \exp(ik(x - na)) + B_n \exp(-ik(x - na)). \quad (5.5.2)$$

The energy is then given by

$$E = \frac{\hbar^2 k^2}{2m}. \quad (5.5.3)$$

By definition, for $x \in [(n+1)a, (n+2)a]$ the wave function is given by

$$\Psi(x) = A_{n+1} \exp(ik(x - (n+1)a)) + B_{n+1} \exp(-ik(x - (n+1)a)). \quad (5.5.4)$$

Therefore for $x \in [na, (n+1)a]$ we have

$$\Psi(x+a) = A_{n+1} \exp(ik(x - na)) + B_{n+1} \exp(-ik(x - na)). \quad (5.5.5)$$

Since we are dealing with a periodic potential, the Hamilton operator commutes with the shift operator. Hence, the eigenvectors of H can simultaneously be chosen as eigenvectors of T . As such they obey

$$\Psi(x+a) = \exp(ika)\Psi(x). \quad (5.5.6)$$

Comparing this with eq.(5.5.2) and eq.(5.5.5) yields

$$A_{n+1} = \exp(ika)A_n, \quad B_{n+1} = \exp(ika)B_n. \quad (5.5.7)$$

Next we impose continuity and discontinuity conditions for the wave function and its first derivative. The continuity condition at $x = (n+1)a$ takes the form

$$\begin{aligned} A_{n+1} + B_{n+1} &= A_n \exp(ika) + B_n \exp(-ika) \Rightarrow \\ A_n [\exp(ika) - \exp(-ika)] + B_n [\exp(ika) - \exp(-ika)] &= 0. \end{aligned} \quad (5.5.8)$$

The discontinuity equation reads

$$\begin{aligned} &-\frac{\hbar^2}{2m} [ikA_{n+1} - ikB_{n+1} - ikA_n \exp(ika) + ikB_n \exp(-ika)] \\ &= V_0a [A_n \exp(ika) + B_n \exp(-ika)] \Rightarrow \\ &A_n [\exp(ika) - \exp(-ika) + \frac{2mV_0a}{\hbar^2 ik} \exp(ika)] \\ &+ B_n [-\exp(ika) + \exp(-ika) + \frac{2mV_0a}{\hbar^2 ik} \exp(-ika)] = 0. \end{aligned} \quad (5.5.9)$$

The two conditions can be summarized in matrix form

$$\begin{pmatrix} [e^{ika} - e^{-ika}] & [e^{ika} - e^{-ika}] \\ [e^{ika} + e^{-ika}(\frac{2mV_0a}{\hbar^2 ik} - 1)] & [-e^{ika} + e^{-ika}(\frac{2mV_0a}{\hbar^2 ik} + 1)] \end{pmatrix} \begin{pmatrix} A_n \\ B_n \end{pmatrix} = 0. \quad (5.5.10)$$

We can read this as an eigenvalue problem with eigenvalue zero. A nonzero solution for A_n and B_n hence only exists if the above matrix has a zero eigenvalue. This is the case only if the determinant of the matrix vanishes, i.e. if

$$\begin{aligned} &[\exp(ika) - \exp(-ika)][-\exp(ika) + \exp(-ika)(\frac{2mV_0a}{\hbar^2 ik} + 1)] \\ &- [\exp(ika) - \exp(-ika)][\exp(ika) + \exp(-ika)(\frac{2mV_0a}{\hbar^2 ik} - 1)] \\ &= 4 \exp(ika) [-\cos(ka) + \cos(ka) - \frac{mV_0a}{\hbar^2 k} \sin(ka)] = 0, \end{aligned} \quad (5.5.11)$$

and thus

$$\cos(ka) = \cos(ka) - \frac{mV_0a}{\hbar^2 k} \sin(ka). \quad (5.5.12)$$

A graphical method reveals that there are bands of allowed energies separated by forbidden energy regions. This follows when we plot the function $\cos(ka) - (mV_0a/\hbar^2k)\sin(ka)$ noticing that $-1 \leq \cos(ka) \leq 1$. For example, when ka is slightly less than π we have $\cos(ka) \approx -1$ and $\sin(ka) < 0$ such that $\cos(ka) - (mV_0a/\hbar^2k)\sin(ka) < -1$. In that case no solution to the above equation exists, and the corresponding energy is forbidden.

This simple model sheds some light on what happens in real 3-d crystals. In fact, also there finite bands of allowed energy states exist. In a conductor the energy levels of an allowed band are not all occupied by electrons. Therefore an externally applied electric field can raise the energy of the electrons by accelerating them within the crystal, thus inducing an electric current. In an insulator, on the other hand, an allowed energy band is completely filled. In that case the external field cannot increase the energy of the electrons unless it provides enough energy to lift them to the next allowed energy band. Hence a small external field does not induce an electric current. A large field, however, may cause the insulator to break down, and become a conductor.

5.6 The Tunneling Effect

Tunneling is a genuine quantum phenomenon that allows a particle to penetrate a classically forbidden region of large potential energy. This effect plays a role in many branches of physics. For example, the radioactive α -particle decay of an atomic nucleus can be described as a tunneling phenomenon. Also the tunneling of electrons between separate metal surfaces has important applications in condensed matter physics, for example in the tunneling electron microscope. Here we consider tunneling through various potential barriers, starting with a particle confined to a box, with a potential barrier dividing the box into two halves. The corresponding potential is infinite for $|x| > b$, for $|x| < a < b$ we have $V(x) = V_0 > 0$ and $V(x) = 0$ for $a \leq |x| \leq b$. We are interested in the ground state and in the first excited state in this potential. Since the problem is parity symmetric, we look for even and odd solutions separately. For the even solution we make the ansatz

$$\Psi_0(x) = A \sin(k(x - b)), \quad (5.6.1)$$

for $x \in [a, b]$ and $\Psi_0(x) = -A \sin(k(x + b))$ for $x \in [-b, -a]$, as well as

$$\Psi_0(x) = B \cosh(\kappa x), \quad (5.6.2)$$

for $x \in [-a, a]$. The energy is given by

$$E_0 = \frac{\hbar^2 k^2}{2m} = V_0 - \frac{\hbar^2 \kappa^2}{2m}. \quad (5.6.3)$$

Continuity of the wave function at $x = a$ implies

$$A \sin(k(a - b)) = B \cosh(\kappa a), \quad (5.6.4)$$

and continuity of its derivative corresponds to

$$Ak \cos(k(a - b)) = B\kappa \sinh(\kappa a). \quad (5.6.5)$$

Dividing the two equations implies

$$k \cot(k(a - b)) = \kappa \tanh(\kappa a). \quad (5.6.6)$$

In the limit $V_0 \rightarrow \infty$ we have $\kappa \rightarrow \infty$ and thus $k = \pi/(b - a)$. For the odd solution we write

$$\Psi_1(x) = A \sin(k(x - b)), \quad (5.6.7)$$

for $x \in [a, b]$ and $\Psi_1(x) = A \sin(k(x + b))$ for $x \in [-b, -a]$, as well as

$$\Psi_1(x) = B \sinh(\kappa x), \quad (5.6.8)$$

for $x \in [-a, a]$. Continuity of the wave function at $x = a$ now implies

$$A \sin(k(a - b)) = B \sinh(\kappa a), \quad (5.6.9)$$

and continuity of its derivative corresponds to

$$Ak \cos(k(a - b)) = B\kappa \cosh(\kappa a). \quad (5.6.10)$$

Dividing the two equations yields

$$k \cot(k(a - b)) = \kappa \coth(\kappa a). \quad (5.6.11)$$

Again, for $V_0 \rightarrow \infty$ we have $k = \pi/(b - a)$. In this limit the ground state and the first excited state become degenerate. At finite but large V_0 there is a small energy splitting $E_1 - E_0$ between the two states.

Let us now prepare an initial wave function (at $t = 0$) that is entirely localized in the right hand side of the box, i.e.

$$\Psi(x, 0) = A \sin\left(\frac{\pi}{a}(x - b)\right), \quad (5.6.12)$$

for $x \in [a, b]$ and zero otherwise. This wave function can be written as a superposition of the ground state and first excited state wave functions

$$\Psi(x, 0) = \frac{1}{\sqrt{2}}[\Psi_0(x) + \Psi_1(x)]. \quad (5.6.13)$$

Hence, the time evolution of the full wave function is

$$\Psi(x, t) = \frac{1}{\sqrt{2}}[\Psi_0(x) \exp(-\frac{i}{\hbar}E_0t) + \Psi_1(x) \exp(-\frac{i}{\hbar}E_1t)]. \quad (5.6.14)$$

After the tunneling time

$$t_T = \frac{\hbar\pi}{E_1 - E_0}, \quad (5.6.15)$$

the wave function takes the form

$$\begin{aligned} \Psi(x, t_T) &= \frac{1}{\sqrt{2}} \exp(-\frac{i}{\hbar}E_0t) [\Psi_0(x) + \Psi_1(x) \exp(-\frac{i}{\hbar}(E_1 - E_0)t_T)] \\ &= \frac{1}{\sqrt{2}} \exp(-\frac{i}{\hbar}E_0t) [\Psi_0(x) - \Psi_1(x)], \end{aligned} \quad (5.6.16)$$

which is entirely localized in the left hand side of the box. The particle has penetrated the potential barrier, although classically it would not have had enough energy to do so. Indeed, in the classical limit $\hbar \rightarrow 0$ the tunneling time diverges,

i.e. then tunneling does not happen. In the above problem tunneling is a periodic process, i.e. at time $2t_T$ the particle will be back in the right hand side of the box. It should be mentioned that this is not exactly true. In the above calculation we have taken the ground state and first excited state wave functions in the limit $V_0 \rightarrow \infty$, but we have still maintained a small energy splitting, i.e. as far as the energies were concerned we have assumed that V_0 is large but finite. Then there would also be a very small contribution of the higher excited states that we have neglected here.

Let us now discuss tunneling through a square well potential without putting the particle in a box. Then the potential is $V(x) = V_0 > 0$ for $0 \leq x \leq a$, and zero otherwise. Again, we are interested in energies $E = \hbar^2 k^2 / 2m < V_0$, i.e. classically the particle would not have enough energy to go over the barrier. We prepare a state in which the particle has momentum k at $x \rightarrow -\infty$. Then the incoming probability wave will be partially reflected back to $-\infty$, and partly it will tunnel through the barrier and then go to $+\infty$. The corresponding ansatz for the wave function is

$$\Psi(x) = \exp(ikx) + R \exp(-ikx), \quad (5.6.17)$$

for $x \leq 0$, and

$$\Psi(x) = T \exp(ikx), \quad (5.6.18)$$

for $x \geq a$. Here T is the tunneling amplitude and $|T|^2$ is the tunneling rate. Inside the barrier the wave function takes the form

$$\Psi(x) = A \exp(\kappa x) + B \exp(-\kappa x). \quad (5.6.19)$$

Continuity of the wave function at $x = 0$ implies

$$1 + R = A + B, \quad (5.6.20)$$

and continuity of its derivative yields

$$ik(1 - R) = \kappa(A - B), \quad (5.6.21)$$

which implies

$$2A = 1 + R + \frac{ik}{\kappa}(1 - R), \quad 2B = 1 + R - \frac{ik}{\kappa}(1 - R). \quad (5.6.22)$$

Similarly, at $x = a$ we obtain

$$T \exp(ika) = A \exp(\kappa a) + B \exp(-\kappa a), \quad (5.6.23)$$

as well as

$$ikT \exp(ika) = A\kappa \exp(\kappa a) - B\kappa \exp(-\kappa a), \quad (5.6.24)$$

and thus

$$\begin{aligned} T \exp(ika) \left(1 + \frac{ik}{\kappa}\right) &= 2A \exp(\kappa a) = \left[1 + \frac{ik}{\kappa} + R\left(1 - \frac{ik}{\kappa}\right)\right] \exp(\kappa a), \\ T \exp(ika) \left(1 - \frac{ik}{\kappa}\right) &= 2B \exp(-\kappa a) = \left[1 - \frac{ik}{\kappa} + R\left(1 + \frac{ik}{\kappa}\right)\right] \exp(-\kappa a). \end{aligned}$$

Eliminating R from these equations we obtain

$$T \exp(ika) \left[\left(1 + \frac{ik}{\kappa}\right)^2 \exp(-\kappa a) - \left(1 - \frac{ik}{\kappa}\right)^2 \exp(\kappa a)\right] = \left(1 + \frac{ik}{\kappa}\right)^2 - \left(1 - \frac{ik}{\kappa}\right)^2, \quad (5.6.25)$$

which results in

$$T = \exp(-ika) \frac{2ik\kappa}{(k^2 - \kappa^2) \sinh(\kappa a) + 2ik\kappa \cosh(\kappa a)}. \quad (5.6.26)$$

The tunneling rate then takes the form

$$|T|^2 = \frac{4k^2\kappa^2}{(k^2 + \kappa^2)^2 \sinh^2(\kappa a) + 4k^2\kappa^2}. \quad (5.6.27)$$

In the limit of a large barrier ($\kappa a \rightarrow \infty$) the tunneling rate is exponentially suppressed

$$|T|^2 = \frac{4k^2\kappa^2}{(k^2 + \kappa^2)^2} \exp(-2\kappa a). \quad (5.6.28)$$

The most important term is the exponential suppression factor. The prefactors are small corrections to this effect, and hence we can write approximately

$$|T|^2 = \exp(-2\kappa a) = \exp\left(-2a \sqrt{\frac{2m}{\hbar^2} [V_0 - E]}\right). \quad (5.6.29)$$

In practical applications tunneling barriers in general do not have square well shapes. Therefore, let us now turn to tunneling through an arbitrary potential barrier. Then we will not be able to solve the Schrödinger equation completely. Still, we can perform a semi-classical approximation (similar to Bohr-Sommerfeld quantization) and obtain a closed expression for the tunneling rate. Of course, in contrast to Bohr-Sommerfeld quantization, for a tunneling process no corresponding classical path exists, simply because classically tunneling is forbidden. Still, as we will see, classical mechanics at purely imaginary “time” will provide us with a tunneling path. The concept of imaginary or so-called Euclidean time plays an important role in the modern treatment of quantum field theories, which

simplifies the mathematics of path integrals. Also here imaginary time is just a mathematical trick to derive an expression for the tunneling rate in a simple way. We should keep in mind that in real time there is no classical path that describes a tunneling process. Let us write the tunneling rate as

$$T = \exp\left(\frac{i}{\hbar}(S[x(t)] + Et_T)\right). \quad (5.6.30)$$

This expression follows from Feynman's path integral in the same way as Bohr-Sommerfeld quantization. In that case the action was to be evaluated for a classical path $x_c(t)$ with period T . Here a tunneling path $x(t)$ occurs instead, and the period T is replaced by the time t_T . Note also that in this calculation T represents the tunneling amplitude. How is the tunneling path determined? Let us just write energy conservation, and then let mathematics be our guide

$$E = \frac{m}{2}\left(\frac{dx(t)}{dt}\right)^2 + V(x(t)). \quad (5.6.31)$$

This classical equation has no physically meaningful solution because now $E < V(x(t))$, and hence formally the kinetic energy would be negative. Mathematically we can make sense of

$$\frac{m}{2}\left(\frac{dx(t)}{dt}\right)^2 = E - V(x(t)) < 0, \quad (5.6.32)$$

if we replace real time t by imaginary "time" it . Then the above equation takes the form

$$-\frac{m}{2}\left(\frac{dx(t)}{d(it)}\right)^2 = E - V(x(t)) < 0, \quad (5.6.33)$$

which is perfectly consistent. The effect of using imaginary time is the same as changing $E - V(x(t))$, which is negative, into $V(x(t)) - E$. Hence, the tunneling path in imaginary time would be the classical path in an inverted potential. Next we write

$$S[x(t)] + Et_T = \int_0^{t_T} dt \left[\frac{m}{2} \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) + E \right] = \int_0^{t_T} dt 2[E - V(x(t))]. \quad (5.6.34)$$

Here we have just used the energy conservation equation from above. Next we turn the integral over t into an integral over x between the classical turning points $x(0)$ and $x(t_T)$ at which $V(x) = E$. Then

$$S[x(t)] + Et_T = \int_{x(0)}^{x(t_T)} dx \frac{1}{dx(t)/dt} 2[E - V(x(t))]. \quad (5.6.35)$$

Using

$$\frac{dx(t)}{dt} = \sqrt{\frac{2}{m}[E - V(x(t))]} = i\sqrt{\frac{2}{m}[V(x(t)) - E]}, \quad (5.6.36)$$

we thus obtain

$$T = \exp\left(\frac{i}{\hbar}(S[x(t)] + Et_T)\right) = \exp\left(-\frac{1}{\hbar} \int_{x(0)}^{x(t_T)} dx \sqrt{2m[V(x) - E]}\right). \quad (5.6.37)$$

For the square well potential from before the classical turning points are $x(0) = 0$ and $x(t_T) = a$. Hence, the tunneling rate then takes the form

$$|T|^2 = \exp\left(-\frac{2}{\hbar} \int_{x(0)}^{x(t_T)} dx \sqrt{2m[V(x) - E]}\right) = \exp\left(-\frac{2a}{\hbar} \sqrt{2m[V_0 - E]}\right), \quad (5.6.38)$$

in complete agreement with our previous result.

5.7 Reflection at a Potential Step

We have seen that quantum mechanics differs from classical mechanics because a quantum particle can tunnel through a classically forbidden region. A similar non-classical behavior occurs when a particle is reflected from a potential step. Let us consider a potential which is zero for $x \leq 0$, and constant and repulsive for $x \geq 0$, i.e. then $V(x) = V_0 > 0$. Let us consider energies $E > V_0$ that are sufficient to go over the step classically. Then classically a particle that is coming from $-\infty$ would just slow down at $x = 0$ and continue to move to $+\infty$ with a reduced kinetic energy. Certainly, classically there would be no reflection in this situation. Still, quantum mechanically there is a probability for the particle to be reflected although it has enough energy to continue to $+\infty$. We make the following ansatz for the wave function at $x \leq 0$

$$\Psi(x) = \exp(ikx) + R \exp(-ikx), \quad (5.7.1)$$

while for $x \geq 0$ we write

$$\Psi(x) = T \exp(iqx). \quad (5.7.2)$$

We have not included a wave $\exp(-iqx)$ at $x \geq 0$ because we are describing an experiment in which a particle is coming in only from the left. The energy is given by

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 q^2}{2m} + V_0. \quad (5.7.3)$$

Next we impose continuity conditions both for the wave function and its first derivative

$$1 + R = T, \quad ik(1 - R) = iqT \Rightarrow 1 - R = \frac{q}{k}T. \quad (5.7.4)$$

Solving these equations yields

$$T = \frac{2k}{k+q}, \quad R = \frac{k-q}{k+q}. \quad (5.7.5)$$

The reflection rate $|R|^2$ vanishes only when $k = q$. This happens either if there is no potential ($V_0 = 0$) or if the energy goes to infinity. Otherwise, there is a probability for the particle to be reflected back to $-\infty$.

Let us look at the above problem from the point of view of probability conservation. Probability conservation is encoded in the continuity equation

$$\partial_t \rho(\vec{x}, t) + \vec{\nabla} \cdot \vec{j}(\vec{x}, t) = 0, \quad (5.7.6)$$

where

$$\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2, \quad (5.7.7)$$

is the probability density, and

$$\vec{j}(\vec{x}, t) = \frac{\hbar}{2im} [\Psi(\vec{x}, t)^* \vec{\nabla} \Psi(\vec{x}, t) - \Psi(\vec{x}, t) \vec{\nabla} \Psi(\vec{x}, t)^*], \quad (5.7.8)$$

is the probability current density. In the above time-independent 1-dimensional problem the continuity equation reduces to

$$\partial_x j(x) = 0, \quad (5.7.9)$$

i.e. the current

$$j(x) = \frac{\hbar}{2im} [\Psi(x)^* \partial_x \Psi(x) - \Psi(x) \partial_x \Psi(x)^*], \quad (5.7.10)$$

should then be constant (independent of x). Let us calculate the current for the problem from before. First we consider $x \leq 0$ and we obtain

$$\begin{aligned} j(x) &= \frac{\hbar k}{2m} [(\exp(-ikx) + R^* \exp(ikx))(\exp(ikx) - R \exp(-ikx)) \\ &\quad - (\exp(ikx) + R \exp(-ikx))(-\exp(-ikx) + R^* \exp(ikx))] \\ &= \frac{\hbar k}{m} [1 - |R|^2]. \end{aligned} \quad (5.7.11)$$

For $x \geq 0$, on the other hand, we find

$$j(x) = \frac{\hbar q}{2m} [T^* \exp(-ikx) T \exp(ikx) - T \exp(ikx) T^* \exp(-ikx)] = \frac{\hbar q}{m} |T|^2. \quad (5.7.12)$$

Indeed, both currents are independent of x . In addition, they are equal (as they should) because

$$k[1 - |R|^2] = k[1 - (\frac{k - q}{k + q})^2] = \frac{4k^2q}{(k + q)^2} = q|T|^2. \quad (5.7.13)$$

5.8 A Paradox

We have seen that quantum mechanically particles behave differently than in classical physics. For example, a classical billiard will never tunnel through a region of potential energy higher than its total energy, and it will never get reflected from a region of potential energy lower than its total energy. Let us go back to the problem with the potential step from before. However, now we want the particle to step down, in other words now $V_0 < 0$. Certainly, a classical billiard ball incident from $-\infty$ would enter the region of decreased potential energy, thereby increasing its kinetic energy, and would then move to $+\infty$. In particular, it would not be reflected back to $-\infty$. Let us see what happens quantum mechanically. In the previous calculation with the potential step we have never used the fact that $V_0 > 0$. Therefore our result is equally valid for $V_0 < 0$. Hence, again

$$T = \frac{2k}{k + q}, \quad R = \frac{k - q}{k + q}. \quad (5.8.1)$$

Let us take the limit of a very deep potential step ($V_0 \rightarrow -\infty$). Then $q \rightarrow \infty$ and thus

$$T \rightarrow 0, \quad R \rightarrow -1, \quad (5.8.2)$$

such that the transmission rate is zero and the reflection rate is $|R|^2 = 1$. This is the exact opposite of what happens classically. The quantum particle stops at the top of the potential step and turns back, instead of entering the region of very low potential energy, as a classical particle would certainly do. One may want to call this the “quantum height anxiety paradox” because the quantum particle steps back into the region of high potential energy, while a classical billiard ball would happily jump down the cliff. It is obvious that something is wrong in this discussion. Quantum mechanics may give different results than classical physics, but the results should not differ as much as in this paradox. In particular, in the classical limit $\hbar \rightarrow 0$ quantum mechanics should give the same result as classical mechanics. This is certainly not the case in our paradox. In fact, our result for R does not even contain \hbar . So taking the classical limit does not change our answer: the particle would still turn around instead of going forward.

The above paradox gets resolved when we consider a potential that interpolates more smoothly between the regions of large and small potential energies. As we will see, the abrupt step function change in the potential is responsible for the paradox, and we should hence be careful in general when we use step function potentials. Still, they are the simplest potentials for which the Schrödinger equation can be solved in closed form, and they therefore are quite useful for learning quantum mechanics. To understand why the step function is special, let us smear out the step over some length scale a by writing the potential as

$$V(x) = V_0 f(x/a), \quad (5.8.3)$$

where f is a smooth function that interpolates between $f(-\infty) = 0$ and $f(+\infty) = -1$, which in the limit $a \rightarrow 0$ approaches the step function. The Schrödinger equation with the smooth potential takes the form

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} + V_0 f(x/a) \Psi(x) = E \Psi(x). \quad (5.8.4)$$

Dividing the equation by V_0 and introducing the dimensionless variable

$$y = \sqrt{2mV_0/\hbar^2} x, \quad (5.8.5)$$

one obtains

$$-\frac{d^2 \Psi(y)}{dy^2} + f(y/\sqrt{2ma^2V_0/\hbar^2}) \Psi(y) = \frac{E}{V_0} \Psi(y). \quad (5.8.6)$$

This rescaled form of the Schrödinger equation shows that the reflection rate $|R|^2$ can depend only on E/V_0 and on the combination $2ma^2V_0/\hbar^2$. In particular, now the limit $V_0 \rightarrow \infty$ is equivalent to the classical limit $\hbar \rightarrow 0$ as well as to the limit $a \rightarrow \infty$ of a very wide potential step. This is exactly what one would expect. The paradox arises only in the case of a true step function, which has $a = 0$. Only then \hbar disappears from the scaled equations, and we do not get a meaningful result in the classical limit. In fact, an arbitrarily small amount of smearing of the potential step is sufficient to eliminate the paradox.

Chapter 6

The Harmonic Oscillator

6.1 Solution of the Schrödinger Equation

The harmonic oscillator plays an important role in many areas of physics. It describes small oscillations of a physical system around an equilibrium configuration. For example, the ions in a crystal oscillate around their equilibrium positions, and thus for small oscillations represent a coupled system of harmonic oscillators. In quantum mechanics the energy of such oscillations is quantized, and the corresponding quanta are called phonons. Phonons play an important role, for example, in superconductors, because they provide a mechanism that binds electrons into Cooper pairs, which can undergo so-called Bose-Einstein condensation and therefore carry electric current without resistance. Also photons — the quantized oscillations of the electromagnetic field — are directly related to harmonic oscillators. In fact, one can view the electromagnetic field as a system of coupled oscillators, one for each space point. Here we consider the quantum mechanics of a single harmonic oscillator first in one dimension. A classical realization of this situation is, for example, a particle attached to a spring with spring constant k . The potential energy of such an oscillator is

$$V(x) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2x^2. \quad (6.1.1)$$

Here ω is the angular frequency of the oscillator. Solving the Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} + \frac{1}{2}m\omega^2x^2\Psi(x) = E\Psi(x), \quad (6.1.2)$$

for the harmonic oscillator is more complicated than for the potentials that we have previously discussed because now we are no longer just patching together

solutions of the free particle Schrödinger equation in various regions. Let us divide the equation by $\hbar\omega$ and use the rescaled dimensionless position variable

$$y = \alpha x, \quad \alpha = \sqrt{\frac{m\omega}{\hbar}}. \quad (6.1.3)$$

Then the Schrödinger equation takes the form

$$-\frac{1}{2} \frac{d^2\Psi(y)}{dy^2} + \frac{1}{2} y^2 \Psi(y) = \epsilon \Psi(y), \quad (6.1.4)$$

where $\epsilon = E/\hbar\omega$ is a rescaled dimensionless energy variable. Let us make the following ansatz for the wave function

$$\Psi(y) = \varphi(y) \exp\left(-\frac{1}{2}y^2\right). \quad (6.1.5)$$

Inserting this in the Schrödinger equation yields the following equation for $\varphi(y)$

$$-\frac{1}{2} \frac{d^2\varphi(y)}{dy^2} + y \frac{d\varphi(y)}{dy} = \left(\epsilon - \frac{1}{2}\right) \varphi(y). \quad (6.1.6)$$

Since this problem has a parity symmetry, we look separately for even and odd solutions. The above equation has an even solution $\varphi(y) = 1$ with $\epsilon = 1/2$. The corresponding wave function

$$\Psi_0(x) = A \exp\left(-\frac{1}{2}\alpha^2 x^2\right), \quad (6.1.7)$$

with energy $E = \hbar\omega/2$ represents the ground state of the harmonic oscillator. It should be noted that, unlike a classical oscillator, the quantum oscillator cannot have zero energy. In fact, this would imply that the particle is at rest at $x = 0$, which is in contradiction with the Heisenberg uncertainty principle. Next we look for the first excited state, which must be odd, by putting $\varphi(y) = y$ and we realize that this is a solution for $\epsilon = 3/2$. This solution corresponds to the wave function

$$\Psi_1(x) = Bx \exp\left(-\frac{1}{2}\alpha^2 x^2\right), \quad (6.1.8)$$

with energy $E = 3\hbar\omega/2$. The second excited state is again even, and follows from the ansatz $\varphi(y) = y^2 + c$. Inserting this in the above equation one obtains $\epsilon = 5/2$ and $c = -1/2$, such that

$$\Psi_2(x) = C(2\alpha^2 x^2 - 1) \exp\left(-\frac{1}{2}\alpha^2 x^2\right), \quad (6.1.9)$$

with energy $E = 5\hbar\omega/2$. One can construct all higher excited states $\Psi_n(x)$ by writing $\varphi(y)$ as a polynomial of degree n , and inserting it in the equation for $\varphi(y)$. The resulting polynomials are the so-called Hermite polynomials, and the resulting energy is

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right). \quad (6.1.10)$$

6.2 Operator Formalism

An alternative and more elegant method to solve the harmonic oscillator problem uses an operator formalism that we will now discuss. The Hamilton operator for the harmonic oscillator takes the form

$$\frac{H}{\hbar\omega} = -\frac{1}{2\alpha^2} \frac{d^2}{dx^2} + \frac{1}{2}\alpha^2 x^2. \quad (6.2.1)$$

Let us introduce the operators

$$a = \frac{1}{\sqrt{2}}\left(\alpha x + \frac{1}{\alpha} \frac{d}{dx}\right), \quad a^\dagger = \frac{1}{\sqrt{2}}\left(\alpha x - \frac{1}{\alpha} \frac{d}{dx}\right), \quad (6.2.2)$$

which are Hermitean conjugates of one another. Acting with them on a wave function one obtains

$$\begin{aligned} a^\dagger a \Psi(x) &= \frac{1}{2}\left(\alpha x - \frac{1}{\alpha} \frac{d}{dx}\right)\left(\alpha x \Psi(x) + \frac{1}{\alpha} \frac{d\Psi(x)}{dx}\right) \\ &= -\frac{1}{2\alpha^2} \frac{d^2\Psi(x)}{dx^2} + \frac{1}{2}\alpha^2 x^2 \Psi(x) - \frac{1}{2}\Psi(x) \\ &= \left(\frac{H}{\hbar\omega} - \frac{1}{2}\right)\Psi(x), \end{aligned} \quad (6.2.3)$$

such that

$$\frac{H}{\hbar\omega} = a^\dagger a + \frac{1}{2}. \quad (6.2.4)$$

We can solve the eigenvalue problem of H very elegantly, just by working out various commutation relations. First of all

$$\begin{aligned} [a, a^\dagger]\Psi(x) &= (aa^\dagger - a^\dagger a)\Psi(x) \\ &= \frac{1}{2}\left(\alpha x + \frac{1}{\alpha} \frac{d}{dx}\right)\left(\alpha x \Psi(x) - \frac{1}{\alpha} \frac{d\Psi(x)}{dx}\right) \\ &\quad - \frac{1}{2}\left(\alpha x - \frac{1}{\alpha} \frac{d}{dx}\right)\left(\alpha x \Psi(x) + \frac{1}{\alpha} \frac{d\Psi(x)}{dx}\right) = \Psi(x), \end{aligned} \quad (6.2.5)$$

such that

$$[a, a^\dagger] = 1. \quad (6.2.6)$$

Also we need the commutation relations with the Hamilton operator

$$\begin{aligned} \left[\frac{H}{\hbar\omega}, a\right] &= \left[a^\dagger a + \frac{1}{2}, a\right] = [a^\dagger, a]a = -a, \\ \left[\frac{H}{\hbar\omega}, a^\dagger\right] &= \left[a^\dagger a + \frac{1}{2}, a^\dagger\right] = a^\dagger[a, a^\dagger] = a^\dagger. \end{aligned} \quad (6.2.7)$$

We will now prove that a^\dagger and a act as raising and lowering operators between the various eigenstates of the harmonic oscillator. First, we act with a^\dagger on an eigenstate $\Psi_n(x)$ and obtain a new state

$$\Phi(x) = a^\dagger \Psi_n(x), \quad (6.2.8)$$

which, as we will now show, is also an eigenstate of H . For this purpose we evaluate

$$\begin{aligned} \frac{H}{\hbar\omega} \Phi(x) &= \frac{H}{\hbar\omega} a^\dagger \Psi_n(x) = ([\frac{H}{\hbar\omega}, a^\dagger] + a^\dagger \frac{H}{\hbar\omega}) \Psi_n(x) \\ &= (\varepsilon_n + 1) a^\dagger \Psi_n(x) = (\varepsilon_n + 1) \Phi(x). \end{aligned} \quad (6.2.9)$$

Thus, $\Phi(x)$ is an eigenstate of H with eigenvalue $\hbar\omega(\varepsilon_n + 1)$. This means it must be proportional to $\Psi_{n+1}(x)$, i.e.

$$\Phi(x) = \lambda \Psi_{n+1}(x). \quad (6.2.10)$$

To determine the proportionality constant λ we compute the normalization of $\Phi(x)$ as

$$\langle \Phi | \Phi \rangle = \langle a^\dagger \Psi_n | a^\dagger \Psi_n \rangle = \langle a a^\dagger \Psi_n | \Psi_n \rangle = \langle (a^\dagger a + 1) \Psi_n | \Psi_n \rangle = n + 1. \quad (6.2.11)$$

On the other hand

$$\langle \Phi | \Phi \rangle = |\lambda|^2 \langle \Psi_{n+1} | \Psi_{n+1} \rangle = |\lambda|^2, \quad (6.2.12)$$

such that $\lambda = \sqrt{n+1}$ and thus

$$a^\dagger \Psi_n(x) = \sqrt{n+1} \Psi_{n+1}(x). \quad (6.2.13)$$

Iterating this equation one can construct the n -th excited state by acting with the raising operator a^\dagger on the ground state n times, i.e.

$$\Psi_n(x) = \frac{1}{\sqrt{n!}} (a^\dagger)^n \Psi_0(x). \quad (6.2.14)$$

Next we act with the lowering operator a on an eigenstate $\Psi_n(x)$ and obtain

$$\Phi(x) = a \Psi_n(x), \quad (6.2.15)$$

which again is an eigenstate of H because now

$$\begin{aligned} \frac{H}{\hbar\omega} \Phi(x) &= \frac{H}{\hbar\omega} a \Psi_n(x) = ([\frac{H}{\hbar\omega}, a] + a \frac{H}{\hbar\omega}) \Psi_n(x) \\ &= (\varepsilon_n - 1) a \Psi_n(x) = (\varepsilon_n - 1) \Phi(x). \end{aligned} \quad (6.2.16)$$

Thus, $\Phi(x)$ is an eigenstate of H with eigenvalue $\hbar\omega(\varepsilon_n - 1)$, and hence it is proportional to $\Psi_{n-1}(x)$, i.e. now

$$\Phi(x) = \lambda\Psi_{n-1}(x). \quad (6.2.17)$$

The constant λ again follows from

$$\langle\Phi|\Phi\rangle = \langle a\Psi_n|a\Psi_n\rangle = \langle a^\dagger a\Psi_n|\Psi_n\rangle = n, \quad (6.2.18)$$

and

$$\langle\Phi|\Phi\rangle = |\lambda|^2\langle\Psi_{n-1}|\Psi_{n-1}\rangle = |\lambda|^2, \quad (6.2.19)$$

such that now $\lambda = \sqrt{n}$ and

$$a\Psi_n(x) = \sqrt{n}\Psi_{n-1}(x). \quad (6.2.20)$$

Indeed a acts as a lowering operator. When a acts on the ground state the above equation implies

$$a\Psi_0(x) = 0, \quad (6.2.21)$$

which is consistent because there is no eigenstate below the ground state. The operator $a^\dagger a$ is called the number operator because

$$a^\dagger a\Psi_n(x) = a^\dagger\sqrt{n}\Psi_{n-1}(x) = n\Psi_n(x), \quad (6.2.22)$$

i.e. it measures the number of energy quanta in the state $\Psi_n(x)$.

6.3 Coherent States and the Classical Limit

Classically a harmonic oscillator changes its position and momentum periodically such that

$$x(t) = A \cos(\omega t), \quad p(t) = m \frac{dx(t)}{dt} = -mA\omega \sin(\omega t). \quad (6.3.1)$$

The eigenstates of the quantum harmonic oscillator, on the other hand, have expectation values of x and p , which are zero independent of time. Hence, the question arises how the classical oscillations can emerge from the quantum theory in the limit $\hbar \rightarrow 0$. Obviously, the states that resemble classical oscillatory behavior cannot be the energy eigenstates. Instead, we will discuss so-called coherent states, which are linear combinations of energy eigenstates, and which indeed have the expected behavior in the classical limit. The coherent states are eigenstates of the lowering operator a , i.e.

$$a\Phi_\lambda(x) = \lambda\Phi_\lambda(x), \quad (6.3.2)$$

For the eigenvalue λ we have the corresponding eigenstate $\Phi_\lambda(x)$. Since a is not a Hermitian operator, the eigenvalue λ will in general be complex. Let us now construct the coherent state as a linear combination of energy eigenstates

$$\Phi_\lambda(x) = \sum_n c_n(\lambda) \Psi_n(x). \quad (6.3.3)$$

The eigenvalue equation takes the form

$$a\Phi_\lambda(x) = \sum_n c_n(\lambda) a\Psi_n(x) = \sum_n c_n(\lambda) \sqrt{n} \Psi_{n-1}(x) = \lambda \sum_n c_n(\lambda) \Psi_n(x). \quad (6.3.4)$$

Hence, we can read off a recursion relation

$$c_n(\lambda) = \frac{\lambda}{\sqrt{n}} c_{n-1}(\lambda). \quad (6.3.5)$$

Iterating the recursion relation we find

$$c_n(\lambda) = \frac{\lambda^n}{\sqrt{n!}} c_0(\lambda). \quad (6.3.6)$$

We still need to normalize the state $\Phi_\lambda(x)$. The normalization condition takes the form

$$\langle \Phi_\lambda | \Phi_\lambda \rangle = \sum_n |c_n(\lambda)|^2 = \sum_n \frac{|\lambda|^{2n}}{n!} |c_0(\lambda)|^2 = \exp(|\lambda|^2) |c_0(\lambda)|^2 = 1, \quad (6.3.7)$$

such that

$$c_n(\lambda) = \frac{\lambda^n}{\sqrt{n!}} \exp\left(-\frac{1}{2}|\lambda|^2\right). \quad (6.3.8)$$

By construction it is clear that the coherent state is not an energy eigenstate. Still we can compute the expectation value of the energy

$$\begin{aligned} \langle \Phi_\lambda | H \Phi_\lambda \rangle &= \langle \Phi_\lambda | \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \sum_n c_n(\lambda) \Psi_n \rangle = \hbar\omega \sum_n \left(n + \frac{1}{2} \right) |c_n(\lambda)|^2 \\ &= \hbar\omega \sum_n \left(n + \frac{1}{2} \right) \frac{|\lambda|^{2n}}{n!} \exp(-|\lambda|^2) = \hbar\omega \left(|\lambda|^2 + \frac{1}{2} \right). \end{aligned} \quad (6.3.9)$$

We see explicitly that we are not in an energy eigenstate when we calculate the variance of the energy

$$\begin{aligned}
\langle H^2 \rangle &= \langle \Phi_\lambda | \hbar^2 \omega^2 (a^\dagger a + \frac{1}{2})^2 \Phi_\lambda \rangle = \hbar^2 \omega^2 \langle \Phi_\lambda | (a^\dagger a a^\dagger a + a^\dagger a + \frac{1}{4}) \Phi_\lambda \rangle \\
&= \hbar^2 \omega^2 \langle \Phi_\lambda | (a^\dagger ([a, a^\dagger] + a^\dagger a) a + a^\dagger a + \frac{1}{4}) \Phi_\lambda \rangle \\
&= \hbar^2 \omega^2 \langle \Phi_\lambda | ((a^\dagger)^2 a^2 + 2a^\dagger a + \frac{1}{4}) \Phi_\lambda \rangle \\
&= \hbar^2 \omega^2 (\langle a^2 \Phi_\lambda | a^2 \Phi_\lambda \rangle + 2\langle a \Phi_\lambda | a \Phi_\lambda \rangle + \frac{1}{4}) \\
&= \hbar^2 \omega^2 (|\lambda|^4 + 2|\lambda|^2 + \frac{1}{4}), \tag{6.3.10}
\end{aligned}$$

and thus

$$\Delta H = \sqrt{\langle H^2 \rangle - \langle H \rangle^2} = \hbar \omega |\lambda|. \tag{6.3.11}$$

Only for $\lambda = 0$ there is no fluctuation in the energy, and then indeed the coherent state reduces to the oscillator ground state, which of course has a sharp energy.

Let us now consider the following position and momentum expectation values

$$\begin{aligned}
\langle x \rangle &= \langle \Phi_\lambda | \frac{1}{\sqrt{2}\alpha} [a + a^\dagger] \Phi_\lambda \rangle = \frac{1}{\sqrt{2}\alpha} [\langle \Phi_\lambda | a \Phi_\lambda \rangle + \langle a \Phi_\lambda | \Phi_\lambda \rangle] \\
&= \frac{1}{\sqrt{2}\alpha} [\lambda + \lambda^*], \\
\langle p \rangle &= \langle \Phi_\lambda | \frac{\alpha \hbar}{\sqrt{2}i} [a - a^\dagger] \Phi_\lambda \rangle = \frac{\alpha \hbar}{\sqrt{2}i} [\langle \Phi_\lambda | a \Phi_\lambda \rangle - \langle a \Phi_\lambda | \Phi_\lambda \rangle] \\
&= \frac{\alpha \hbar}{\sqrt{2}i} [\lambda - \lambda^*], \\
\langle x^2 \rangle &= \langle \Phi_\lambda | \frac{1}{2\alpha^2} [a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2] \Phi_\lambda \rangle \\
&= \frac{1}{2\alpha^2} \langle \Phi_\lambda | [a^2 + 2a^\dagger a + 1 + (a^\dagger)^2] \Phi_\lambda \rangle \\
&= \frac{1}{2\alpha^2} [\langle \Phi_\lambda | a^2 \Phi_\lambda \rangle + 2\langle a \Phi_\lambda | a \Phi_\lambda \rangle + \langle \Phi_\lambda | \Phi_\lambda \rangle + \langle a^2 \Phi_\lambda | \Phi_\lambda \rangle] \\
&= \frac{1}{2\alpha^2} [\lambda^2 + 2|\lambda|^2 + 1 + \lambda^{*2}], \\
\langle p^2 \rangle &= \langle \Phi_\lambda | -\frac{\alpha^2 \hbar^2}{2} [a^2 - aa^\dagger - a^\dagger a + (a^\dagger)^2] \Phi_\lambda \rangle \\
&= -\frac{\alpha^2 \hbar^2}{2} \langle \Phi_\lambda | [a^2 - 2a^\dagger a - 1 + (a^\dagger)^2] \Phi_\lambda \rangle \\
&= -\frac{\alpha^2 \hbar^2}{2} [\langle \Phi_\lambda | a^2 \Phi_\lambda \rangle - 2\langle a \Phi_\lambda | a \Phi_\lambda \rangle - \langle \Phi_\lambda | \Phi_\lambda \rangle + \langle a^2 \Phi_\lambda | \Phi_\lambda \rangle] \\
&= -\frac{\alpha^2 \hbar^2}{2} [\lambda^2 - 2|\lambda|^2 - 1 + \lambda^{*2}].
\end{aligned}$$

Thus, we obtain

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{1}{2\alpha^2}, \quad (\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 = \frac{\alpha^2 \hbar^2}{2}, \quad (6.3.12)$$

such that

$$\Delta x \Delta p = \frac{\hbar}{2}, \quad (6.3.13)$$

i.e. the coherent state saturates Heisenberg's inequality. This is a first indication that this state is as classical as possible, it has a minimal uncertainty product. From the proof of the Heisenberg uncertainty relation of chapter 3 it follows that the inequality can be saturated for a state $\Psi(x)$ only if

$$\frac{d}{dx} \Psi(x) + \frac{1}{2(\Delta x)^2} x \Psi(x) - \left(\frac{\langle x \rangle}{2(\Delta x)^2} + \frac{i}{\hbar} \langle p \rangle \right) \Psi(x) = 0. \quad (6.3.14)$$

Inserting the above results for the coherent state this equation takes the form

$$\frac{d}{dx}\Phi_\lambda(x) + \alpha^2 x\Phi_\lambda(x) - \sqrt{2}\alpha\lambda\Phi_\lambda(x) = 0. \quad (6.3.15)$$

Using

$$a = \frac{1}{\sqrt{2}}\left(\alpha x + \frac{1}{\alpha} \frac{d}{dx}\right), \quad (6.3.16)$$

we indeed identify this as the eigenvalue equation for a .

As we learned earlier, a Gaussian wave packet saturates the Heisenberg uncertainty relation. As we now know, the same is true for an oscillator coherent state, and as we will see next, the two are actually the same thing. In fact, let us consider a shifted Gaussian wave packet that also has an average momentum $\langle p \rangle$

$$\Phi(x) = A \exp\left(-\frac{1}{2}\alpha^2(x - \langle x \rangle)^2\right) \exp\left(\frac{i}{\hbar}\langle p \rangle x\right). \quad (6.3.17)$$

Such a wave packet spreads during its time evolution if we are dealing with free particles. Here, however, we have a harmonic oscillator potential. As we will see later, the wave packet then does not spread. Instead it oscillates back and forth without changing its shape. Let us first convince ourselves that the above wave packet is indeed equal to one of our coherent states. For that purpose we act on it with the lowering operator

$$\begin{aligned} a\Phi(x) &= \frac{1}{\sqrt{2}}\left(\alpha x + \frac{1}{\alpha} \frac{d}{dx}\right)\Phi(x) \\ &= \frac{1}{\sqrt{2}}\left(\alpha x - \alpha(x - \langle x \rangle) + \frac{i}{\alpha\hbar}\langle p \rangle\right)\Phi(x) \\ &= \frac{1}{\sqrt{2}}\left(\alpha\langle x \rangle + \frac{i}{\alpha\hbar}\langle p \rangle\right)\Phi(x) = \lambda\Phi(x). \end{aligned} \quad (6.3.18)$$

We see that the Gaussian wave packet $\Phi(x)$ is indeed an eigenstate of the lowering operator, and hence equal to the coherent state with eigenvalue

$$\lambda = \frac{1}{\sqrt{2}}\left(\alpha\langle x \rangle + \frac{i}{\alpha\hbar}\langle p \rangle\right). \quad (6.3.19)$$

In order to understand that the Gaussian wave packet equivalent to the coherent state does not spread, we now investigate its time evolution. We know that each energy eigenstate $\Psi_n(x)$ simply picks up the phase $\exp(-iE_n t/\hbar)$ during its

time evolution. Hence we obtain

$$\begin{aligned}
\Phi_\lambda(x, t) &= \sum_n c_n(\lambda) \Psi_n(x) \exp\left(-\frac{i}{\hbar} E_n t\right) \\
&= \sum_n \frac{\lambda^n}{\sqrt{n!}} \exp\left(-\frac{1}{2} |\lambda|^2\right) \Psi_n(x) \exp\left(-i\omega\left(n + \frac{1}{2}\right)t\right) \\
&= \exp\left(-i\frac{\omega}{2}t\right) \sum_n \frac{[\lambda \exp(-i\omega t)]^n}{\sqrt{n!}} \Psi_n(x) \\
&= \exp\left(-i\frac{\omega}{2}t\right) \Phi_{\lambda(t)}(x).
\end{aligned} \tag{6.3.20}$$

We see that up to an irrelevant global phase $\exp(-i\frac{\omega}{2}t)$ the time evolution only manifests itself in a time-dependent eigenvalue

$$\lambda(t) = \lambda \exp(-i\omega t). \tag{6.3.21}$$

In particular, a coherent state remains coherent, i.e. the corresponding Gaussian wave packet does not change its shape (in coordinate space). However, it oscillates back and forth exactly like a classical harmonic oscillator. This becomes clear when we remind ourselves that the real part of λ is proportional to $\langle x \rangle$, while the imaginary part is proportional to $\langle p \rangle$. Since $\lambda(t)$ has a periodic behavior described by $\exp(-i\omega t)$, the same is true for its real and imaginary parts. Consequently, both average position and average momentum of the Gaussian wave packet change periodically, but its shape remains unchanged. This is as close as one can get to a quasi-classical behavior of the quantum oscillator.

6.4 The Harmonic Oscillator in Two Dimensions

Let us now turn to the harmonic oscillator in more than one dimension. Here we consider two dimensions, but it will become clear that we could go to three (and even higher) dimensions in exactly the same way. In two dimensions the harmonic oscillator potential takes the form

$$V(\vec{r}) = V(x, y) = \frac{1}{2} m \omega^2 (x^2 + y^2). \tag{6.4.1}$$

We see that the potential separates into an x -dependent and a y -dependent piece. This is a special property of the harmonic oscillator, which will allow us to reduce the harmonic oscillator problem in higher dimensions to the one-dimensional problem. Alternatively we can write the potential in polar coordinates

$$V(\vec{r}) = V(r, \varphi) = \frac{1}{2} m \omega^2 r^2, \tag{6.4.2}$$

where

$$x = r \cos \varphi, \quad y = r \sin \varphi. \quad (6.4.3)$$

It is obvious that the problem is rotation invariant because the potential is φ -independent. This symmetry leads to the conservation of angular momentum, both classically and at the quantum level. In quantum mechanics symmetries often lead to degeneracies in the spectrum. In fact, as we will see, the harmonic oscillator spectrum is highly degenerate in more than one dimension. We can understand part of the degeneracy based on rotational invariance. However, there is more degeneracy in the harmonic oscillator spectrum than one would expect based on rotation invariance. In fact, there are other less obvious symmetries in the problem. These symmetries are sometimes called accidental because it was for some time not well understood where the symmetry came from. Now we know that the symmetry reflects a special property of the harmonic oscillator force law. At the classical level, the symmetry implies that all orbits in a harmonic oscillator potential are closed curves in space. This is a very special property of the harmonic oscillator. In fact, a generic rotation invariant potential $V(r)$ will not have closed classical orbits. Another important problem with closed classical orbits is given by the Coulomb potential. In that case all bound classical orbits are closed, and indeed we will see later that the discrete spectrum of the hydrogen atom again has more degeneracy than one would expect based on rotation symmetry alone.

Let us first look at the classical solutions of the two dimensional harmonic oscillator. The most general solution of Newton's equation is

$$x(t) = A \cos(\omega t), \quad y(t) = B \cos(\omega t - \varphi_0), \quad (6.4.4)$$

which parameterizes an ellipse. After one period $T = 2\pi/\omega$ both $x(t)$ and $y(t)$ return to their original positions and the classical orbit closes. In fact, the particle returns to its original position with its original momentum because

$$p_x(t) = -MA\omega \sin(\omega t), \quad p_y(t) = -MB\omega \sin(\omega t - \varphi_0), \quad (6.4.5)$$

are also periodic functions. Consequently, the whole motion repeats itself after one period. Let us now consider the angular momentum

$$\begin{aligned} L_z &= xp_y - yp_x \\ &= MAB\omega [\sin(\omega t) \cos(\omega t - \varphi_0) - \cos(\omega t) \sin(\omega t - \varphi_0)] \\ &= MAB\omega \sin \varphi_0, \end{aligned} \quad (6.4.6)$$

which is time-independent and thus conserved. We will see later that the conservation of angular momentum also manifests itself in quantum mechanics.

We are now ready to solve the Schrödinger equation of the two dimensional harmonic oscillator. It takes the form

$$-\frac{\hbar^2}{2M}\Delta\Psi(\vec{r}) + V(\vec{r})\Psi(\vec{r}) = E\Psi(\vec{r}), \quad (6.4.7)$$

or equivalently in Cartesian coordinates

$$-\frac{\hbar^2}{2M}[\partial_x^2\Psi(x, y) + \partial_y^2\Psi(x, y)] + \frac{1}{2}M\omega^2(x^2 + y^2)\Psi(x, y) = E\Psi(x, y). \quad (6.4.8)$$

Since the potential separates into an x -dependent and a y -dependent piece it is natural to try a separation of variables also for the wave function. This motivates the ansatz

$$\Psi(x, y) = \phi(x)\chi(y). \quad (6.4.9)$$

Inserting this in the Schrödinger equation and dividing by $\chi(y)$ implies

$$-\frac{\hbar^2}{2M}[\partial_x^2\phi(x) + \frac{\phi(x)}{\chi(y)}\partial_y^2\chi(y)] + \frac{1}{2}M\omega^2(x^2 + y^2)\phi(x) = E\phi(x), \quad (6.4.10)$$

or

$$-\frac{\hbar^2}{2M}\partial_x^2\phi(x) + \frac{1}{2}M\omega^2x^2\phi(x) - E\phi(x) = \frac{\hbar^2}{2M}\frac{\phi(x)}{\chi(y)}\partial_y^2\chi(y) - \frac{1}{2}M\omega^2y^2\phi(x). \quad (6.4.11)$$

The left hand side of this equation is y -independent, while the right-hand side seems to be y -dependent. However, if the two sides are equal, also the right hand side must be y -independent. Calling the function on the left-hand side $f(x)$ we can write

$$\frac{\hbar^2}{2M}\frac{\phi(x)}{\chi(y)}\partial_y^2\chi(y) - \frac{1}{2}M\omega^2y^2\phi(x) = -f(x), \quad (6.4.12)$$

such that

$$-\frac{\hbar^2}{2M}\partial_y^2\chi(y) + \frac{1}{2}M\omega^2y^2\chi(y) = \frac{f(x)}{\phi(x)}\chi(y), \quad (6.4.13)$$

Now the left hand side is x -independent, and therefore the same must be true for the right hand side. This implies that

$$\frac{f(x)}{\phi(x)} = E_y, \quad (6.4.14)$$

is some constant. Then the above equation is just the Schrödinger equation of a one dimensional harmonic oscillator with energy E_y

$$-\frac{\hbar^2}{2M}\partial_y^2\chi(y) + \frac{1}{2}M\omega^2y^2\chi(y) = E_y\chi(y). \quad (6.4.15)$$

Inserting this back into the original Schrödinger equation we also obtain

$$-\frac{\hbar^2}{2M}\partial_x^2\phi(x) + \frac{1}{2}M\omega^2x^2\phi(x) = E_x\phi(x), \quad (6.4.16)$$

with

$$E = E_x + E_y. \quad (6.4.17)$$

This shows that the two dimensional (and in fact any higher dimensional) harmonic oscillator problem is equivalent to a set of one dimensional oscillator problems. This allows us to write down the solutions of the two dimensional oscillator problem as

$$\Psi(x, y) = \Psi_{n_x}(x)\Psi_{n_y}(y), \quad (6.4.18)$$

where $\Psi_{n_x}(x)$ and $\Psi_{n_y}(y)$ are eigenfunctions of one dimensional harmonic oscillators. The corresponding energy eigenvalue is

$$E_n = E_x + E_y = \hbar\omega(n_x + n_y + 1) = \hbar\omega(n + 1). \quad (6.4.19)$$

The n -th excited state is $(n + 1)$ -fold degenerate because there are $n + 1$ ways of writing the integer n as a sum of non-negative integers n_x and n_y . This high degeneracy is only partly due to the rotational symmetry of the problem. The rest of the degeneracy is related to the symmetry that ensures that all classical orbits of the oscillator are closed curves.

Chapter 7

The Hydrogen Atom

7.1 Separation of the Center of Mass Motion

The hydrogen atom is the simplest atom consisting of a single electron of charge $-e$ and a proton of charge $+e$ that represents the atomic nucleus. The two particles are bound together via the attractive Coulomb interaction. In the following we want to discuss a slight generalization of the hydrogen problem. So-called hydrogen-like atoms consist again of a single electron, but may have a more complicated atomic nucleus consisting of Z protons (and hence with charge $+Ze$) and $A - Z$ neutrons. A neutron has almost the same mass as a proton, and thus the mass of the atomic nucleus is then given by $M_A \approx AM_p$, where M_p is the mass of the proton. As far as atomic physics is concerned we can consider the nucleus as a single particle of mass M_A and charge $+Ze$. Hence hydrogen-like atoms can be viewed as two-particle systems consisting of an electron and an atomic nucleus. So far we have only considered single particles moving under the influence of an external potential. Now we are dealing with two particles influenced by their mutual Coulomb attraction. We do not consider any external forces in this case, which implies that the system is translation invariant. This invariance leads to the conservation of the total momentum. We can then go to the center of mass frame, and reduce the two-particle problem formally to the problem of a single particle moving in an external potential.

The two-particle Schrödinger equation for hydrogen-like atoms takes the form

$$\left[-\frac{\hbar^2}{2M_e}\Delta_e - \frac{\hbar^2}{2M_A}\Delta_A + V(\vec{r}_e - \vec{r}_A)\right]\Psi(\vec{r}_e, \vec{r}_A) = E\Psi(\vec{r}_e, \vec{r}_A), \quad (7.1.1)$$

where M_e is the electron mass, M_A is the mass of the nucleus, and $V(\vec{r})$ is the Coulomb potential depending of the separation $\vec{r} = \vec{r}_e - \vec{r}_A$ of the two particles. The Laplace operators Δ_e and Δ_A represent second derivatives with respect to the position of the electron and the nucleus respectively. Although we now have a two-particle Schrödinger equation we still have a single wave function $\Psi(\vec{r}_e, \vec{r}_A)$, which, however, now depends on the positions of both particles. $\Psi(\vec{r}_e, \vec{r}_A)$ specifies the probability amplitude to find the electron at position \vec{r}_e and simultaneously the nucleus at position \vec{r}_A . The proper normalization of the wave function therefore is

$$\int d^3r_e d^3r_A |\Psi(\vec{r}_e, \vec{r}_A)|^2 = 1. \quad (7.1.2)$$

Next we introduce the center of mass coordinate

$$\vec{R} = \frac{M_e \vec{r}_e + M_A \vec{r}_A}{M_e + M_A}. \quad (7.1.3)$$

The kinetic energy of the two particles can then be expressed as

$$-\frac{\hbar^2}{2M_e} \Delta_e - \frac{\hbar^2}{2M_A} \Delta_A = -\frac{\hbar^2}{2M} \Delta_R - \frac{\hbar^2}{2\mu} \Delta_r, \quad (7.1.4)$$

where $M = M_e + M_A$ is the total mass, and

$$\mu = \frac{M_e M_A}{M_e + M_A} \quad (7.1.5)$$

is the so-called reduced mass. Since the atomic nucleus is several thousand times heavier than an electron, the reduced mass is very close to the electron mass. The total Hamilton operator now takes the form

$$H = -\frac{\hbar^2}{2M} \Delta_R - \frac{\hbar^2}{2\mu} \Delta_r + V(\vec{r}). \quad (7.1.6)$$

The first term represents the kinetic energy of the center of mass motion and depends only on \vec{R} . The second and third term represent kinetic and potential energy of the relative motion and depend only on \vec{r} . This separation of the Hamilton operator suggests a separation ansatz for the wave function

$$\Psi(\vec{r}, \vec{R}) = \psi(\vec{r}) \Phi(\vec{R}). \quad (7.1.7)$$

Inserting this ansatz in the two-particle Schrödinger equation implies

$$-\frac{\hbar^2}{2M} \Delta_R \Phi(\vec{R}) = E_R \Phi(\vec{R}), \quad (7.1.8)$$

for the center of mass motion. This is the equation for a free particle of mass M . Its solutions are plane waves

$$\Phi(\vec{R}) = \exp\left(\frac{i}{\hbar}\vec{R} \cdot \vec{P}\right), \quad (7.1.9)$$

characterized by the total momentum \vec{P} . From now on we will go to the center of mass frame in which $\vec{P} = 0$. Then the energy of the center of mass motion $E_R = P^2/2M$ also vanishes. The remaining equation for the relative motion takes the form

$$-\frac{\hbar^2}{2\mu}\Delta_r\psi(\vec{r}) + V(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}). \quad (7.1.10)$$

Formally, this is the Schrödinger equation for a single particle of the reduced mass μ interacting with a potential $V(\vec{r})$. By separating the center of mass motion we have reduced the original two-particle problem to a single-particle problem.

7.2 Angular Momentum

Up to now we have not used the fact that the Coulomb potential only depends on the distance between the electron and the atomic nucleus. The corresponding rotational invariance implies that the angular momentum is a conserved quantity. This will allow us to also separate the angular variables, and to reduce the problem further to a 1-dimensional radial Schrödinger equation. In spherical coordinates r , θ and φ the Laplace operator takes the form

$$\Delta_r = \partial_r^2 + \frac{2}{r}\partial_r + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{r^2\sin^2\theta}\partial_\varphi^2. \quad (7.2.1)$$

Using the definitions of the angular momentum operator $\vec{L} = \vec{r} \times \vec{p}$ and the relative momentum operator $\vec{p} = \hbar\vec{\nabla}_r/i$ one can show that

$$\vec{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{\sin^2\theta}\partial_\varphi^2\right], \quad (7.2.2)$$

which implies

$$\Delta_r = \partial_r^2 + \frac{2}{r}\partial_r - \frac{\vec{L}^2}{\hbar^2r^2}. \quad (7.2.3)$$

Therefore the Hamilton operator for the relative motion takes the form

$$H = -\frac{\hbar^2}{2\mu}\left(\partial_r^2 + \frac{2}{r}\partial_r\right) + V_{eff}(r), \quad (7.2.4)$$

where

$$V_{eff}(r) = V(r) + \frac{\vec{L}^2}{2\mu r^2} \quad (7.2.5)$$

is an effective potential consisting of the original Coulomb potential $V(r)$ and a so-called centrifugal barrier, which corresponds to a repulsive potential for states with non-zero angular momentum. Since the effective potential is rotation invariant (independent of the angles θ and φ) the above Hamilton operator commutes with the angular momentum operator, i.e. $[H, \vec{L}] = 0$. As we have seen earlier, the various components of the angular momentum vector, however, do not commute with each other. Still, they all commute with the angular momentum squared, i.e. $[L_i, \vec{L}^2] = 0$. We can now select L_z , \vec{L}^2 and H as a set of mutually commuting operators. This means that we can diagonalize these three operators simultaneously. In particular, we can choose the eigenfunctions of H to be simultaneously eigenfunctions of L_z and \vec{L}^2 . The eigenfunctions of $L_z = -i\hbar\partial_\varphi$ are simply $\exp(im\varphi)$ with $m \in \mathbb{Z}$ as we have already seen for a particle on a circle. The eigenvalue of this state is the z -component of angular momentum $m\hbar$. The simultaneous eigenfunctions of L_z and \vec{L}^2 are known as spherical harmonics $Y_{lm}(\theta, \varphi)$. One can show that

$$\vec{L}^2 Y_{lm}(\theta, \varphi) = \hbar^2 l(l+1) Y_{lm}(\theta, \varphi), \quad L_z Y_{lm}(\theta, \varphi) = \hbar m Y_{lm}(\theta, \varphi). \quad (7.2.6)$$

Here $l \in \mathbb{Z}$ and the possible values for m are given by $m \in \{-l, -l+1, \dots, l\}$. The simplest spherical harmonics are given by

$$\begin{aligned} Y_{00}(\theta, \varphi) &= \frac{1}{\sqrt{4\pi}}, \\ Y_{10}(\theta, \varphi) &= \sqrt{\frac{3}{4\pi}} \cos \theta, \\ Y_{1\pm 1}(\theta, \varphi) &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta \exp(\pm i\varphi). \end{aligned}$$

It is beyond the scope of this course to discuss angular momentum in more detail. For the moment we should accept the above result without a detailed proof. The proof is not difficult, just time consuming. The states with angular momentum $l = 0$ are sometimes also called s-states, states with $l = 1$ are called p-states, and states with $l = 2$ are so-called d-states. The p-states are 3-fold, and the d-states are 5-fold degenerate because there are $2l + 1$ possible m -values for a given angular momentum l . Since the Hamilton operator commutes with the angular momentum vector, we can write its eigenfunctions as

$$\psi(\vec{r}) = R(r) Y_{lm}(\theta, \varphi). \quad (7.2.7)$$

Inserting this in the Schrödinger equation for the relative motion implies

$$\left[-\frac{\hbar^2}{2\mu}(\partial_r^2 + \frac{2}{r}\partial_r) + V(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2}\right]R(r) = ER(r). \quad (7.2.8)$$

This is a 1-dimensional Schrödinger equation for the radial motion. Note, however, that the angular momentum l enters the equation via the centrifugal barrier term. Note also that the z -component m of the angular momentum does not appear in the equation. This again shows that all states with angular momentum l are $(2l+1)$ -fold degenerate because there are that many spherical harmonics $Y_{lm}(\theta, \varphi)$ all with the same eigenvalue $l(l+1)$.

7.3 Solution of the Radial Equation

Having reduced the problem to a 1-dimensional radial Schrödinger equation, we are now ready to solve that equation. Up to this point we have only used the fact that the potential $V(r)$ is rotation invariant. From now on we will use the explicit Coulomb form

$$V(r) = -\frac{Ze^2}{r}. \quad (7.3.1)$$

First we introduce the dimensionless variable

$$\rho = \sqrt{\frac{8\mu|E|}{\hbar^2}}r. \quad (7.3.2)$$

Then the radial equation takes the form

$$\left[\partial_\rho^2 + \frac{2}{\rho}\partial_\rho - \frac{l(l+1)}{\rho^2} + \frac{\lambda}{\rho} - \frac{1}{4}\right]R(\rho) = 0, \quad (7.3.3)$$

where

$$\lambda = \frac{Ze^2}{\hbar^2} \sqrt{\frac{\mu}{2|E|}}. \quad (7.3.4)$$

First we consider the large distance limit $\rho \rightarrow \infty$. Then

$$\left[\partial_\rho^2 - \frac{1}{4}\right]R(\rho) = 0, \quad (7.3.5)$$

such that $R(\rho) \sim \exp(-\rho/2)$. This motivates the following ansatz

$$R(\rho) = G(\rho) \exp(-\rho/2). \quad (7.3.6)$$

Inserting this ansatz in the above equation yields

$$[\partial_\rho^2 + (\frac{2}{\rho} - 1)\partial_\rho - \frac{l(l+1)}{\rho^2} + \frac{\lambda-1}{\rho}]G(\rho) = 0. \quad (7.3.7)$$

At very small ρ we thus have

$$[\partial_\rho^2 + \frac{2}{\rho}\partial_\rho - \frac{l(l+1)}{\rho^2}]G(\rho) = 0, \quad (7.3.8)$$

which is solved by $G(\rho) \sim \rho^l$. Therefore we also explicitly separate off this short distance behavior and write

$$G(\rho) = H(\rho)\rho^l. \quad (7.3.9)$$

Inserting this in the above equation for $G(\rho)$ we obtain

$$[\partial_\rho^2 + (\frac{2l+2}{\rho} - 1)\partial_\rho + \frac{\lambda-1-l}{\rho}]H(\rho) = 0. \quad (7.3.10)$$

We see that a constant $H(\rho) = a_0$ solves the equation provided that $\lambda = l + 1$. The corresponding wave function has no radial node, and thus corresponds to the state of lowest energy for a given value of l . However, we can also make the ansatz $H(\rho) = a_0 + a_1\rho$. Then

$$(\frac{2l+2}{\rho} - 1)a_1 + \frac{\lambda-1-l}{\rho}(a_0 + a_1\rho) = 0, \quad (7.3.11)$$

such that $\lambda = l+2$ and $(2l+2)a_1 + a_0 = 0$. This solution has one zero of the radial wave function, and thus corresponds to the first excited state for a given value of l . Similarly, one can find solutions with n_r zeros of the radial wave function. In that case

$$\lambda = n_r + l + 1 = n \in Z. \quad (7.3.12)$$

This implies for the quantized energy values

$$E = -\frac{Z^2 e^4 \mu}{2\hbar^2 n^2}. \quad (7.3.13)$$

Based on rotation invariance one would expect that a state with angular momentum l is $(2l+1)$ -fold degenerate. However, we can realize a given n with various n_r and hence with various l . For example, for the first excited states with $n = 2$ we can have $n_r = 0, l = 1$ or $n_r = 1, l = 0$. This means that an s-state and a p-state have the same energy (something one would not expect based on rotation invariance alone). The s-state is 1-fold degenerate, while the p-state is 3-fold degenerate, and hence the first excited state of hydrogen-like atoms is in fact 4-fold degenerate. Generally, the state with quantum number n is n^2 -fold degenerate. This high degeneracy, which one would not expect based on rotation invariance alone, is due to the fact that all bound classical orbits in the Coulomb potential are closed curves (ellipses).

7.4 Relativistic corrections

At this point we have reached our goal of understanding quantitatively the simplest atom — hydrogen. We have gained some insight in the quantum world, which is indeed consistent with experiments. Still, the nonrelativistic $1/r$ potential problem does not exactly represent the real hydrogen atom. There are small relativistic corrections, which we want to mention at the end of the course. The Schrödinger equation provides a nonrelativistic description, while we know that Nature is relativistic. The correct relativistic equation for the electron is the so-called Dirac equation, which follows from the Hamilton operator

$$H = \vec{\alpha} \cdot \vec{p}c + \beta M_e c^2 + V(r). \quad (7.4.1)$$

Here

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad (7.4.2)$$

are 4×4 matrices with

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (7.4.3)$$

as 2×2 submatrices. The Dirac Hamilton operator acts on a 4-component wave function

$$\Psi(\vec{r}) = \begin{pmatrix} \Psi_1(\vec{r}) \\ \Psi_2(\vec{r}) \\ \Psi_3(\vec{r}) \\ \Psi_4(\vec{r}) \end{pmatrix}, \quad (7.4.4)$$

a so-called Dirac spinor. The upper components $\Psi_1(\vec{r})$ and $\Psi_2(\vec{r})$ describe electrons, and the lower components $\Psi_3(\vec{r})$ and $\Psi_4(\vec{r})$ describe positrons — the antiparticles of the electron. The fact that we need a 2-component wave function to describe the electron is related to a property called spin, which represents the intrinsic angular momentum of the electron, which is independent of its orbital motion. In fact, the above Pauli matrices $\vec{\sigma}$ are related to the spin

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma}. \quad (7.4.5)$$

Indeed, it is straightforward to show that the components of \vec{S} obey angular momentum commutation relations. For example

$$\begin{aligned} [S_x, S_y] &= \frac{\hbar^2}{4} \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \\ &= \frac{\hbar^2}{4} \begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} = i\hbar S_z. \end{aligned} \quad (7.4.6)$$

Also

$$\vec{S}^2 = \frac{\hbar^2}{4} \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^2 + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2 \right) = \frac{3\hbar^2}{4} I. \quad (7.4.7)$$

Writing this as $s(s+1)\hbar^2$ we identify $s = 1/2$, i.e. the Pauli matrices describe an angular momentum $1/2$ with z -components $S_z = m\hbar$ and $m = \pm 1/2$. This additional intrinsic quantum number of the electron leads to a doubling of the whole hydrogen spectrum. In fact, each state that we have discussed previously can now occur with spin up ($m = 1/2$) and with spin down ($m = -1/2$). This effect survives even in the extreme nonrelativistic limit. Hence the degeneracy of hydrogen states with quantum number n is indeed $2n^2$.

Spin is not only an extra label on the electron. It also influences the dynamics via relativistic effects. The most important of these is due to the so-called spin-orbit coupling. When one reduces the Dirac equation to the upper electron components one finds an extra contribution to the potential which has the form

$$V_{so}(r) = \frac{Ze^2}{2M_e^2 c^2 r^3} \vec{S} \cdot \vec{L}, \quad (7.4.8)$$

involving the scalar product of spin and orbital angular momentum. The s -states are unaffected by this term because they have $l = 0$, but p -states are energetically shifted due to the spin-orbit coupling. With the above term in the Hamilton operator, the orbital angular momentum no longer commutes with the Hamiltonian. Still, the total angular momentum

$$\vec{J} = \vec{S} + \vec{L} \quad (7.4.9)$$

does commute with H reflecting the fact that the problem is still rotation invariant. The total angular momentum \vec{J} also obeys angular momentum commutation relations, and hence the eigenvalues of \vec{J}^2 are given by $\hbar^2 j(j+1)$. For a given value l of the orbital angular momentum one can construct states with $j = l \pm 1/2$. We can now write

$$\vec{J}^2 = \vec{S}^2 + 2\vec{S} \cdot \vec{L} + \vec{L}^2, \quad (7.4.10)$$

such that

$$\vec{S} \cdot \vec{L} = \frac{\hbar^2}{2} (j(j+1) - l(l+1) - \frac{3}{4}). \quad (7.4.11)$$

Consequently, the state with $j = l + 1/2$ has $\vec{S} \cdot \vec{L} = \hbar^2 l/2$ and is shifted to higher energies, while the state with $j = l - 1/2$ has $\vec{S} \cdot \vec{L} = -\hbar^2(l+1)/2$ such that its energy is lowered. The actual spin-orbit energy shifts in the hydrogen atom are very small, but experimentally clearly verified.

One can solve the Dirac equation for hydrogen-like atoms just like we have solved the Schrödinger equation. The resulting energy spectrum is given by

$$E = M_e c^2 \left(1 + \frac{\alpha^2}{\left[n + \sqrt{(j + 1/2)^2 - \alpha^2} \right]^2} \right)^{1/2}. \quad (7.4.12)$$

Here $\alpha = e^2/\hbar c$ is the fine-structure constant and again $n \in \mathbb{Z}$. Even in this spectrum there is more degeneracy than one would expect based on rotation invariance alone. Still, due to spin-orbit couplings the degeneracy is now greatly reduced compared to the nonrelativistic calculation.

In the real hydrogen atom there are even more subtle effects that the above Dirac Hamilton operator does not contain. For example, the atomic nucleus also has a spin, which interacts with the spin of the electron. This gives rise to extremely small (but observable) energy shifts known as the hyperfine structure of the hydrogen spectrum. Finally, a full quantum treatment of the hydrogen atom must also quantize the electromagnetic field that mediates the Coulomb interaction between electron and proton. Then we are entering the subject of Quantum Electrodynamics (QED) which is far beyond the scope of this course. Still, it is worth mentioning that a tiny energy shift — the so-called Lamb shift — is an observable consequence of QED in the hydrogen spectrum. Also this effect finally lifts all the degeneracies that we cannot understand based on rotation invariance alone. In the real hydrogen atom with all its tiny relativistic corrections the classical orbit of an electron would finally no longer be closed.

Chapter 8

EPR Paradox and Bell's Inequality

8.1 The Einstein-Podolsky-Rosen Paradox

By now we have a solid knowledge of quantum mechanics. We have understood the formal structure of the theory, the statistical interpretation of the wave function, and its relation to actual measurements. In particular, we have learned how to do quantum mechanics by solving the Schrödinger equations for several systems of physical interest. Hopefully, we have gained some confidence in our ability to use quantum mechanics for describing concrete physical phenomena, and perhaps we have also gained some intuition for what quantum mechanics does. Although we have learned how quantum mechanics works, have we understood what it really means? As Feynman once said: “I think I can safely say that nobody understands quantum mechanics.” So, we should not feel too bad if we don't either. In order to understand better what it is that is so hard to understand, we will now discuss the Einstein-Podolsky-Rosen paradox.

Einstein was convinced that quantum mechanics cannot be the whole story. In particular, the probabilistic nature of quantum reality was unacceptable for him. Of course, he had to acknowledge that quantum mechanics correctly described Nature at the microscopic scale, but he believed that it must be incomplete. Instead of assuming that there is nothing more to know about a quantum system than its wave function, he believed that there are elements of reality that quantum mechanics simply does not “know” about. Following Einstein, people have attempted to describe such elements of reality by so-called hidden variables.

These attempts ended when Bell showed that local theories of hidden variables in inconsistent with quantum mechanics.

In order to argue that quantum mechanics is incomplete, in 1935 Einstein, Podolsky, and Rosen discussed the following “paradox”. Consider the decay of a spinless particle, e.g. a neutral pion at rest, into two particles with spin 1/2, e.g. an electron and a positron. After the decay, the electron is moving in one direction, and, as a consequence of momentum conservation, the positron is moving in the opposite direction. Angular momentum conservation implies that the spin of the electron plus the spin of the positron add up to the spin zero of the neutral pion. Hence, either the electron has spin up and the positron has spin down, or the electron has spin down and the positron has spin up. The spin wave function takes the form

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \quad (8.1.1)$$

We have used Dirac's bracket notation which will be formally introduced in the next section. The total spin operator

$$\vec{S} = \vec{S}_e + \vec{S}_p, \quad (8.1.2)$$

is the sum of the spin $\vec{S}_e = \frac{\hbar}{2}\vec{\sigma}_e$ of the electron and the spin $\vec{S}_p = \frac{\hbar}{2}\vec{\sigma}_p$ of the positron. We have

$$\sigma_{e3}|\uparrow\downarrow\rangle = |\uparrow\downarrow\rangle, \quad \sigma_{p3}|\uparrow\downarrow\rangle = -|\uparrow\downarrow\rangle, \quad \sigma_{e3}|\downarrow\uparrow\rangle = -|\downarrow\uparrow\rangle, \quad \sigma_{p3}|\downarrow\uparrow\rangle = |\downarrow\uparrow\rangle, \quad (8.1.3)$$

and hence

$$\sigma_3|\Psi\rangle = (\sigma_{e3} + \sigma_{p3})\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = 0, \quad (8.1.4)$$

i.e. the total spin indeed vanishes.

Let us imagine that the electron travels a large distance until it reaches a detector that measures its spin. Simultaneously, the positron travels a large distance in the opposite direction until it reaches another detector that measures its spin. If the electron spin is measured to be up, the positron spin must be measured to be down, and vice versa. Experiments of this kind have actually been done, and indeed the measured spins always point in opposite directions. What did Einstein find paradoxical about this situation? According to the standard interpretation of quantum mechanics, the outcome of the measurement of the electron spin is uncertain until it is actually being performed. In particular, both spin up and spin down are equally probable. Only the measurement itself forces the electron “to make up its mind” and choose one of the two results from the given statistical distribution. Once the electron spin has been measured,

the result of the measurement of the positron spin is completely determined. However, the positron is by now so far away that no light signal can communicate the result of the measurement of the electron spin before the measurement of the positron spin is actually being performed. Still, the measured positron spin is certainly opposite to the one of the electron. Einstein believed that immediately after the pion decay, the electron and positron are equipped with a hidden variable that determines if they have spin up or down. Hence, before the electron spin is actually measured, the result of the measurement is already pre-determined by the hidden variable. In particular, in order to ensure that the positron has opposite spin, no action need to travel faster than light. Since the hidden variable is not part of quantum mechanics, Einstein, Podolsky, and Rosen argued that quantum mechanics is incomplete.

8.2 The Quantum Mechanics of Spin Correlations

In 1964 Bell showed that local hidden variable theories are inconsistent with quantum mechanics. Since the predictions of quantum mechanics are indeed in agreement with experiment, this means that Nature is fundamentally nonlocal. No wonder that Einstein found that unacceptable. Still, there seems to be no way around this. In order to appreciate what Bell argued, we first need to understand what quantum mechanics has to say about spin correlations. Bell suggested to consider measurements of the electron and positron spins along different quantization axes defined by the unit-vectors \vec{a} and \vec{b} . He then considered the expectation value of the product of the two measured spin values

$$P(\vec{a}, \vec{b}) = \langle (\vec{\sigma}_e \cdot \vec{a})(\vec{\sigma}_p \cdot \vec{b}) \rangle. \quad (8.2.1)$$

Let us evaluate this expectation value for the state $|\Psi\rangle$ from above. First, we consider

$$\langle \sigma_{e3} \sigma_{p3} \rangle = -1. \quad (8.2.2)$$

Similarly, one obtains

$$\langle \sigma_{e1} \sigma_{p1} \rangle = \langle \sigma_{e2} \sigma_{p2} \rangle = -1. \quad (8.2.3)$$

Next, we consider

$$\langle \sigma_{e1} \sigma_{p3} \rangle = \left\langle \frac{1}{2} (\sigma_{e+} + \sigma_{e-}) \sigma_{p3} \right\rangle = 0. \quad (8.2.4)$$

All these results can be summarized as

$$\langle \sigma_{ei} \sigma_{pj} \rangle = -\delta_{ij}, \quad (8.2.5)$$

which implies

$$P(\vec{a}, \vec{b}) = -\vec{a} \cdot \vec{b}. \quad (8.2.6)$$

In particular, if the two quantization axes are parallel ($\vec{a} = \vec{b}$), the product of the two measured spins is always -1 . If the axes are antiparallel ($\vec{a} = -\vec{b}$), on the other hand, the product is 1 .

8.3 A Hidden Variable Theory

Let us consider a concrete example of a hidden variable theory. As we will see, this theory is inconsistent with the above results of quantum mechanics, which in turn are consistent with experiment. Although the theory is hence ruled out by experiment, it nicely illustrates the ideas behind hidden variable theories. This hidden variable theory is classical but still probabilistic. It assumes that the quantum spins of the electron and the positron are determined by corresponding classical spin vectors $\pm\vec{\lambda}$ of unit length, which are a concrete realization of a hidden variable. The hidden variable theory assumes that in the moment of the decay the electron is endowed with a classical hidden spin variable $\vec{\lambda}$ which points in a random direction, while the positron is endowed with the classical spin $-\vec{\lambda}$. In this model, the measurement of the spin of the electron along the quantization axis \vec{a} gives

$$s_e(\vec{a}, \vec{\lambda}) = \text{sign}(\vec{a} \cdot \vec{\lambda}) = \pm 1, \quad (8.3.1)$$

while the measurement of the spin of the positron along the quantization axis \vec{b} gives

$$s_p(\vec{b}, \vec{\lambda}) = -\text{sign}(\vec{b} \cdot \vec{\lambda}) = \pm 1. \quad (8.3.2)$$

In other words, the quantum spin is $+1$ if the classical spin has an angle less than $\pi/2$ with the quantization axis, and -1 otherwise. It is typical of a hidden variable theory that the measurement of the quantum spin ± 1 does not reveal all details of the hidden variable — in this case the direction of the classical spin $\vec{\lambda}$. In this sense, if the hidden variable theory were correct, there would be certain elements of reality that quantum mechanics does not “know” about. Let us first consider two special cases: \vec{a} and \vec{b} parallel or antiparallel. For $\vec{a} = \vec{b}$, we have

$$s_e(\vec{a}, \vec{\lambda}) = -s_p(\vec{a}, \vec{\lambda}) \Rightarrow P(\vec{a}, \vec{a}) = -1, \quad (8.3.3)$$

and for $\vec{a} = -\vec{b}$

$$s_e(\vec{a}, \vec{\lambda}) = s_p(-\vec{a}, \vec{\lambda}) \Rightarrow P(\vec{a}, -\vec{a}) = 1. \quad (8.3.4)$$

At least these special cases are consistent with quantum mechanics and thus with experiment. It is easy to show that, for general \vec{a} and \vec{b} , the model predicts

$$P(\vec{a}, \vec{b}) = \frac{2}{\pi} \arccos(\vec{a} \cdot \vec{b}) - 1. \quad (8.3.5)$$

Obviously, this contradicts quantum mechanics and is thus inconsistent with experiment. As we will see in the next section, there is a general proof (using Bell's inequality) which shows that, not only this model, but any local theory of hidden variables necessarily disagrees with quantum mechanics and thus with experiment.

8.4 Bell's Inequality

Let us assume that there are some hidden variables λ that pre-determine the spin of the electron and the positron in the moment they are created in the decay of the neutral pion. The result of the spin measurement of the electron, $s_e(\vec{a}, \lambda) = \pm 1$ (in units of $\frac{\hbar}{2}$), depends only on the quantization axis \vec{a} and on the value of the hidden variables λ . Similarly, the measured spin of the positron, $s_p(\vec{b}, \lambda) = \pm 1$, depends on the quantization axis \vec{b} and λ . If the quantization axes are chosen to be parallel ($\vec{a} = \vec{b}$), it is an experimental fact that the results of the two measurements are perfectly anti-correlated, i.e.

$$s_e(\vec{a}, \lambda) = -s_p(\vec{a}, \lambda). \quad (8.4.1)$$

We want to assume that the hidden variables are distributed according to some arbitrary probability distribution $\rho(\lambda) \geq 0$ which is normalized as

$$\int d\lambda \rho(\lambda) = 1, \quad (8.4.2)$$

so that

$$P(\vec{a}, \vec{b}) = \int d\lambda \rho(\lambda) s_e(\vec{a}, \lambda) s_p(\vec{b}, \lambda) = - \int d\lambda \rho(\lambda) s_e(\vec{a}, \lambda) s_e(\vec{b}, \lambda). \quad (8.4.3)$$

Next, we consider

$$\begin{aligned} P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c}) &= - \int d\lambda \rho(\lambda) [s_e(\vec{a}, \lambda) s_e(\vec{b}, \lambda) - s_e(\vec{a}, \lambda) s_e(\vec{c}, \lambda)] \\ &= \int d\lambda \rho(\lambda) [1 - s_e(\vec{b}, \lambda) s_e(\vec{c}, \lambda)] s_e(\vec{a}, \lambda) s_e(\vec{b}, \lambda). \end{aligned} \quad (8.4.4)$$

In the last step, we have used $s_e(\vec{b}, \lambda)^2 = 1$. Since

$$s_e(\vec{a}, \lambda)s_e(\vec{b}, \lambda) = \pm 1, \quad (8.4.5)$$

we have

$$\rho(\lambda)[1 - s_e(\vec{b}, \lambda)s_e(\vec{c}, \lambda)] \geq 0, \quad (8.4.6)$$

and hence

$$|P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c})| \leq \int d\lambda \rho(\lambda)[1 - s_e(\vec{b}, \lambda)s_e(\vec{c}, \lambda)] = 1 + P(\vec{b}, \vec{c}). \quad (8.4.7)$$

This is Bells' inequality.

It is remarkable that this result — derived for a general theory of local hidden variables — is actually inconsistent with the predictions of quantum mechanics. This is easy to see. Let us assume that the three quantization axes \vec{a} , \vec{b} , and \vec{c} are in the xy -plane with a 90 degrees angle between \vec{a} and \vec{b} , and with 45 degrees angles between \vec{a} and \vec{c} and between \vec{b} and \vec{c} . Then

$$P(\vec{a}, \vec{b}) = 0, \quad P(\vec{a}, \vec{c}) = P(\vec{b}, \vec{c}) = -\frac{1}{\sqrt{2}}, \quad (8.4.8)$$

so that

$$|P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c})| = \frac{1}{\sqrt{2}} > 1 - \frac{1}{\sqrt{2}} = 1 + P(\vec{b}, \vec{c}). \quad (8.4.9)$$

Obviously, this violates Bell's inequality. Consequently, the results of quantum mechanics (which are in excellent agreement with experiment) are inconsistent with any underlying theory of local hidden variables. Most physicists concluded from this results that hidden variables are simply not the correct idea. People who fell in love with hidden variables had to admit that they cannot be local. Thus, in any case, Bell's inequality tells us that Nature is fundamentally nonlocal at the quantum level.

Einstein was not amused by the “spooky action at a distance” that happens when the wave function collapses in the process of a measurement. Although no light signal could possibly communicate the result of the measurement of the electron spin, the measurement of the positron spin is always perfectly anticorrelated. Still, this apparent violation of causality is not as catastrophic as it may seem. In particular, the observer of the electron spin cannot use the “spooky” correlation to transmit information to the observer of the positron with velocities faster than light. This is because she has no way to influence the result of her spin measurement. After a long sequence of measurements, both observers simply get a long list of results $\pm\hbar/2$ which are perfectly random. Only when they compare their lists at the end of the day, they realize that both lists are identical, up to an overall minus sign.

8.5 Schrödinger's Cat

Let us also discuss the (in)famous experiment with Schrödinger's cat. The cute animal is put in a box together with a container of cyanide, a hammer, a Geiger counter, and a single unstable atomic nucleus. The life-time of the nucleus is such that during one hour it decays with a probability of 50 percent. When this happens, the Geiger counter registers the event, activates the hammer, which hits the container, releases the cyanide, and kills the cat. On the other hand, with the other 50 percent probability, nothing happens. Before we open the box, we do not know if the cat is dead or alive. In fact, quantum mechanically, the cat is in a quantum state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\text{dead}\rangle + |\text{alive}\rangle). \quad (8.5.1)$$

Here we have again used Dirac's bracket notation which will be explained in the next section. Of course, Dirac can also not help to make sense of $|\Psi\rangle$. A macroscopic object like a cat can hardly be in a linear superposition of dead and alive.

In any case, Schrödinger then puts us in the following situation. When we open the box and look at the cat, we make the measurement that forces the cat to "make up its mind" if it wants to be dead or alive. In some sense, we kill the cat (with 50 percent probability) by opening the box. The delicacy of quantum measurements becomes obvious when situations are driven to such extremes. Is it really us who decides the fate of the cat? Probably not. We can blame everything on the Geiger counter. Once it registers the radioactive decay, a classical event has happened and the fate of the cat is determined. When we open the box after one hour, the thing has long happened. However, this thought experiment should still bother us. There may be something very deep about the quantum mechanical measurement process that we presently do not comprehend. This cannot stop us from doing quantum mechanics by using what we have learned to describe the physics at microscopic scales. However, I invite you to remain curious and open about new ideas, even if they may sound crazy at first.

Chapter 9

Abstract Form of Quantum Mechanics

9.1 Dirac's Bracket Notation

In chapter 4 we have investigated the formal structure of quantum mechanics. In particular, we found that wave functions can be viewed as unit-vectors in an abstract infinite-dimensional Hilbert space, and that physical observables are described by Hermitean operators acting on those vectors. The value of a wave function $\Psi(x)$ at the point x can be viewed as a component of a vector with infinitely many components — one for each space point. From ordinary vector algebra, we know that it is useful to describe vectors in a basis-free notation \vec{a} instead of using the Cartesian coordinate representation (a_1, a_2, a_3) . For wave functions in the Hilbert space we have thus far used the position space representation $\Psi(x)$ or the momentum space representation $\Psi(k)$, and we will now proceed to the basis-free abstract notation $|\Psi\rangle$ introduced by Dirac. In chapter 4 we have already written the scalar product of two wave functions $\Psi(x)$ and $\Phi(x)$ as $\langle\Phi|\Psi\rangle$. Dirac dissects this bracket expression into a “bra”-vector $\langle\Phi|$ and a “ket”-vector $|\Psi\rangle$. As in ordinary vector algebra, Dirac's basis-free notation is quite useful in many calculations in quantum mechanics. It should be pointed out that we are not really learning anything new in this section. We are just reformulating what we already know in a more compact notation.

Let us investigate the analogy with ordinary vector algebra in some detail. The scalar product of two ordinary vectors $\vec{a} = (a_1, a_2, a_3)$ and $\vec{b} = (b_1, b_2, b_3)$

takes the form

$$\langle a|b\rangle = \vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3, \quad (9.1.1)$$

and the norm of a vector can be written as

$$\|a\|^2 = \langle a|a\rangle = \vec{a} \cdot \vec{a} = a_1^2 + a_2^2 + a_3^2. \quad (9.1.2)$$

Similarly, the scalar product of two wave functions takes the form

$$\langle \Phi|\Psi\rangle = \int dx \Phi(x)^* \Psi(x), \quad (9.1.3)$$

and the norm of a wave function is given by

$$\|\Psi\|^2 = \langle \Psi|\Psi\rangle = \int dx \Psi(x)^* \Psi(x) = \int dx |\Psi(x)|^2. \quad (9.1.4)$$

Introducing the ordinary unit-vectors \vec{i} in the Cartesian i -direction one can write

$$a_i = \langle i|a\rangle = \vec{i} \cdot \vec{a}, \quad (9.1.5)$$

i.e. the component a_i of the basis-free abstract vector \vec{a} in the Cartesian basis of unit-vector \vec{i} is simply the projection onto those basis vectors. Similarly, one can introduce abstract vectors $|x\rangle$ as analogs of \vec{i} in the quantum mechanical Hilbert space. The ket-vectors $|x\rangle$ are so-called position eigenstates. They describe the wave function of a particle completely localized at the point x . The usual wave function $\Psi(x)$ is the analog of the Cartesian component a_i , and the abstract ket-vector $|\Psi\rangle$ is the analog of \vec{a} . In analogy to eq.(9.1.5) one can now write

$$\Psi(x) = \langle x|\Psi\rangle, \quad (9.1.6)$$

i.e. the ordinary wave function $\Psi(x)$ is a component of the abstract state vector $|\Psi\rangle$ in the basis of position eigenstates $|x\rangle$ and is again given by the projection $\langle x|\Psi\rangle$.

Obviously, the j -component of the Cartesian unit vectors \vec{i} (in Cartesian coordinates) is given by

$$i_j = \vec{j} \cdot \vec{i} = \langle j|i\rangle = \delta_{ij}. \quad (9.1.7)$$

Similarly, the position eigenstates obey

$$\langle x'|x\rangle = \delta(x - x'). \quad (9.1.8)$$

Since the position x is a continuous variable, the Kronecker δ for the discrete indices i and j is replaced by a Dirac δ -function. Since the unit-vectors \vec{i} form

a complete basis of orthogonal vectors, one can write the 3×3 unit-matrix as a sum of tensor products

$$\sum_i i_j i_k = \sum_i \delta_{ij} \delta_{ik} = \delta_{jk} = \mathbb{1}_{jk}. \quad (9.1.9)$$

In a basis-free notation one can write this equation as

$$\sum_i (\vec{i} \otimes \vec{i}) = \mathbb{1}. \quad (9.1.10)$$

Here the direct tensor product \otimes turns two 3-component vectors into a 3×3 matrix. In basis-free notation the corresponding equation for position eigenstates in the Hilbert space takes the form

$$\int dx |x\rangle\langle x| = \mathbb{1}, \quad (9.1.11)$$

where the sum over the discrete index i is replaced by an integral over the continuous coordinate x . In the basis of position eigenstates the same equation reads

$$\int dx \langle x'|x\rangle\langle x|x''\rangle = \int dx \delta(x-x')\delta(x-x'') = \delta(x'-x'') = \langle x'|\mathbb{1}|x''\rangle. \quad (9.1.12)$$

Using the completeness relation eq.(9.1.11) we can write

$$|\Psi\rangle = \mathbb{1}|\Psi\rangle = \int dx |x\rangle\langle x|\Psi\rangle = \int dx \Psi(x)|x\rangle, \quad (9.1.13)$$

i.e. the wave function $\Psi(x) = \langle x|\Psi\rangle$ is the component of an abstract state vector $|\Psi\rangle$ in the complete orthonormal basis of position eigenstates $|x\rangle$. Multiplying this equation with $\langle x'|$ one consistently obtains

$$\langle x'|\Psi\rangle = \int dx \Psi(x)\langle x'|x\rangle = \int dx \Psi(x)\delta(x-x') = \Psi(x'). \quad (9.1.14)$$

The two equations for ordinary vectors corresponding to eqs.(9.1.13,9.1.14) take the form

$$\vec{a} = \mathbb{1}\vec{a} = \sum_i (\vec{i} \otimes \vec{i})\vec{a} = \sum_i (\vec{i} \cdot \vec{a})\vec{i} = \sum_i a_i \vec{i}, \quad (9.1.15)$$

and

$$\vec{j} \cdot \vec{a} = \sum_i a_i \vec{j} \cdot \vec{i} = \sum_i a_i \delta_{ij} = a_j. \quad (9.1.16)$$

Just as the Cartesian basis of unit-vectors \vec{i} is not the only basis in real space, the basis of position eigenstates $|x\rangle$ is not the only basis in Hilbert space. In

particular, the momentum eigenstates $|k\rangle$ (with $p = \hbar k$) also form an orthonormal complete basis. In this case we have

$$\frac{1}{2\pi} \int dk |k\rangle\langle k| = \mathbb{1}. \quad (9.1.17)$$

Hence, we can also write

$$|\Psi\rangle = \mathbb{1}|\Psi\rangle = \frac{1}{2\pi} \int dk |k\rangle\langle k|\Psi\rangle. \quad (9.1.18)$$

This raises the question how the object $\langle k|\Psi\rangle$ defined in the basis of momentum eigenstates $|k\rangle$ is related to the familiar wave function $\Psi(x) = \langle x|\Psi\rangle$ in the basis of position eigenstates $|x\rangle$. Expressed in the basis of position eigenstates, the momentum eigenstates take the form

$$\langle x|k\rangle = \exp(ikx) = \exp\left(\frac{i}{\hbar}px\right), \quad (9.1.19)$$

i.e. they are the familiar wave functions of free particles with momentum $p = \hbar k$. Hence, one obtains consistently

$$\langle x|\Psi\rangle = \langle x|\mathbb{1}|\Psi\rangle = \frac{1}{2\pi} \int dk \langle x|k\rangle\langle k|\Psi\rangle = \frac{1}{2\pi} \int dk \Psi(k) \exp(ikx) = \Psi(x). \quad (9.1.20)$$

We identify the object $\langle k|\Psi\rangle$ as the momentum space wave function $\Psi(k)$ that we encountered earlier as the Fourier transform of the position space wave function $\Psi(x)$. Similarly, one obtains

$$\langle k|x\rangle = \langle x|k\rangle^* = \exp(-ikx). \quad (9.1.21)$$

and thus

$$\langle k|\Psi\rangle = \langle k|\mathbb{1}|\Psi\rangle = \int dx \langle k|x\rangle\langle x|\Psi\rangle = \int dx \Psi(x) \exp(-ikx) = \Psi(k). \quad (9.1.22)$$

We are already familiar with Hermitean operators A representing physical observables. We are also familiar with the object

$$\langle \Phi|A\Psi\rangle = \int dx \Phi(x)^* A\Psi(x). \quad (9.1.23)$$

Now we consider the operator A in a basis-free notation and write

$$\langle \Phi|A|\Psi\rangle = \langle \Phi|\mathbb{1}A|\Psi\rangle = \int dx \langle \Phi|x\rangle\langle x|A|\Psi\rangle = \int dx \Phi(x)^*\langle x|A|\Psi\rangle. \quad (9.1.24)$$

Hence, we easily identify $\langle x|A|\Psi\rangle = A\Psi(x)$.

As we have seen earlier, the eigenfunctions $\chi_n(x)$ that solve the eigenvalue problem $A\chi_n(x) = a_n\chi_n(x)$ form a complete orthonormal basis in Hilbert space. In basis-free notation this is expressed as

$$A|\chi_n\rangle = a_n|\chi_n\rangle, \quad (9.1.25)$$

with

$$\langle\chi_m|\chi_n\rangle = \delta_{mn}, \quad (9.1.26)$$

and

$$\sum_n |\chi_n\rangle\langle\chi_n| = \mathbb{1}. \quad (9.1.27)$$

An arbitrary state $|\Psi\rangle$ can be decomposed as

$$|\Psi\rangle = \sum_n c_n|\chi_n\rangle. \quad (9.1.28)$$

Multiplying this equation with $\langle\chi_m|$ we obtain

$$\langle\chi_m|\Psi\rangle = \sum_n c_n\langle\chi_m|\chi_n\rangle = \sum_n c_n\delta_{mn} = c_m. \quad (9.1.29)$$

We now consider the eigenvalue problem of the Hamilton operator

$$\mathbf{H}|\chi_n\rangle = E_n|\chi_n\rangle, \quad (9.1.30)$$

which is nothing but the time-independent Schrödinger equation. For example, the Hamilton operator for a single particle moving in a potential $V(x)$

$$\mathbf{H} = \mathbf{T} + \mathbf{V}, \quad (9.1.31)$$

separates into the kinetic energy operator \mathbf{T} and the potential energy operator \mathbf{V} . In a basis of coordinate eigenstates one has

$$\langle x|\mathbf{T}|\chi_n\rangle = -\frac{\hbar^2}{2m} \frac{d^2\chi_n(x)}{dx^2}, \quad (9.1.32)$$

as well as

$$\langle x|\mathbf{V}|x'\rangle = V(x)\delta \Rightarrow \langle x|\mathbf{V}|\chi_n\rangle = \int dx \langle x|\mathbf{V}|x'\rangle\langle x|\chi_n\rangle = V(x)\chi_n(x), \quad (9.1.33)$$

and hence

$$\langle x|\mathbf{H}|\chi_n\rangle = \langle x|\mathbf{T} + \mathbf{V}|\chi_n\rangle = -\frac{\hbar^2}{2m} \frac{d^2\chi_n(x)}{dx^2} + V(x)\chi_n(x) = E_n\chi_n(x). \quad (9.1.34)$$

In basis-free Dirac notation the time-dependent Schrödinger equation (with a time-independent Hamiltonian) takes the form

$$i\hbar\partial_t|\Psi(t)\rangle = \mathbf{H}|\Psi(t)\rangle. \quad (9.1.35)$$

Expanding

$$|\Psi(t)\rangle = \sum_n c_n(t)|\chi_n\rangle, \quad (9.1.36)$$

in the basis of stationary energy eigenstates $|\chi_n\rangle$ one obtains

$$c_n(t) = \langle\chi_n|\Psi(t)\rangle. \quad (9.1.37)$$

Thus, multiplying the time-dependent Schrödinger equation with $\langle\chi_n|$ one finds

$$\begin{aligned} i\hbar\partial_t c_n(t) &= \langle\chi_n|\mathbf{H}|\Psi(t)\rangle = \sum_m \langle\chi_n|\mathbf{H}|\chi_m\rangle \langle\chi_m|\Psi(t)\rangle \\ &= \sum_m E_m \delta_{mn} \langle\chi_m|\Psi(t)\rangle = E_n c_n(t), \end{aligned} \quad (9.1.38)$$

which implies

$$c_n(t) = c_n(0) \exp\left(-\frac{i}{\hbar} E_n t\right), \quad (9.1.39)$$

and hence

$$|\Psi(t)\rangle = \sum_n \langle\chi_n|\Psi(0)\rangle \exp\left(-\frac{i}{\hbar} E_n t\right) |\chi_n\rangle. \quad (9.1.40)$$

It should be pointed out again that all results of this section have been derived in earlier chapters. We have just presented them in the more compact basis-free Dirac notation.

9.2 Unitary Time-Evolution Operator

The time-dependent Schrödinger equation can be solved formally by writing

$$|\Psi(t)\rangle = \exp\left[-\frac{i}{\hbar} \mathbf{H}(t - t_0)\right] |\Psi(t_0)\rangle, \quad (9.2.1)$$

because then indeed

$$i\hbar\partial_t|\Psi(t)\rangle = \mathbf{H} \exp\left[-\frac{i}{\hbar} \mathbf{H}(t - t_0)\right] |\Psi(t_0)\rangle = \mathbf{H}|\Psi(t)\rangle. \quad (9.2.2)$$

The operator

$$\mathbf{U}(t, t_0) = \exp\left[-\frac{i}{\hbar} \mathbf{H}(t - t_0)\right], \quad (9.2.3)$$

is known as the time-evolution operator because it determines the time-evolution

$$|\Psi(t)\rangle = \mathbf{U}(t, t_0)|\Psi(t_0)\rangle, \quad (9.2.4)$$

of the physical state $|\Psi(t)\rangle$ starting from an initial state $|\Psi(t_0)\rangle$. The time-evolution operator is unitary, i.e.

$$\mathbf{U}(t, t_0)^\dagger \mathbf{U}(t, t_0) = \mathbf{U}(t, t_0) \mathbf{U}(t, t_0)^\dagger = \mathbb{1}. \quad (9.2.5)$$

This property follows from the Hermiticity of the Hamiltonian because $\mathbf{H}^\dagger = \mathbf{H}$ implies

$$\mathbf{U}(t, t_0)^\dagger = \exp\left[\frac{i}{\hbar} \mathbf{H}^\dagger(t - t_0)\right] = \exp\left[\frac{i}{\hbar} \mathbf{H}(t - t_0)\right] = \mathbf{U}(t, t_0)^{-1}. \quad (9.2.6)$$

Unitarity of the time-evolution operator guarantees conservation of probability because

$$\langle \Psi(t) | \Psi(t) \rangle = \langle \Psi(t_0) | \mathbf{U}(t, t_0)^\dagger \mathbf{U}(t, t_0) | \Psi(t_0) \rangle = \langle \Psi(t_0) | \Psi(t_0) \rangle = 1. \quad (9.2.7)$$

9.3 Schrödinger versus Heisenberg Picture

Until now we have formulated quantum mechanics in the way it was first done by Schrödinger. In his approach the Hermitean operators that describe physical observables are time-independent and the time-dependence of the wave function is obtained by solving the time-dependent Schrödinger equation. At the end, the results of measurements of an observable quantity A are given as expectation values $\langle \Psi(t) | A | \Psi(t) \rangle$ which inherit their time-dependence from the wave function while the operator A itself is time-independent.

Heisenberg had a different (but equivalent) picture of quantum mechanics. He preferred to think of the wave function as being time-independent, i.e. $|\Psi(t)\rangle = |\Psi(0)\rangle$ and attribute the time-dependence to the operators. In particular, a time-dependent Hermitean operator $A(t)$ in the Heisenberg picture is related to the time-independent operator A in the Schrödinger picture by a unitary transformation with the time-evolution operator

$$A(t) = \mathbf{U}(t, t_0)^\dagger A \mathbf{U}(t, t_0). \quad (9.3.1)$$

In the Heisenberg picture, the result of a measurement is given by

$$\langle \Psi(t_0) | A(t) | \Psi(t_0) \rangle = \langle \Psi(t_0) | \mathbf{U}(t, t_0)^\dagger A \mathbf{U}(t, t_0) | \Psi(t_0) \rangle = \langle \Psi(t) | A | \Psi(t) \rangle, \quad (9.3.2)$$

which indeed agrees with Schrödinger's result. Since Heisenberg worked with time-dependent operators he needed an analog of the Schrödinger equation to describe the operator dynamics. One finds

$$\begin{aligned}
\partial_t \mathbf{A}(t) &= \partial_t [\mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0)] \\
&= \partial_t \mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0) + \mathbf{U}(t, t_0)^\dagger \mathbf{A} \partial_t \mathbf{U}(t, t_0) \\
&= \frac{i}{\hbar} \mathbf{H} \mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0) - \mathbf{U}(t, t_0)^\dagger \mathbf{A} \frac{i}{\hbar} \mathbf{H} \mathbf{U}(t, t_0) \\
&= \frac{i}{\hbar} [\mathbf{H}, \mathbf{A}(t)].
\end{aligned} \tag{9.3.3}$$

This is the so-called Heisenberg equation of motion. It is physically completely equivalent to the Schrödinger equation. It should be noted that the Hamilton operator itself remains time-independent in the Heisenberg picture, i.e.

$$\mathbf{H}(t) = \mathbf{U}(t, t_0)^\dagger \mathbf{H} \mathbf{U}(t, t_0) = \mathbf{H}. \tag{9.3.4}$$

This is a consequence of energy conservation.

It should be noted that the time-dependent operator $\mathbf{A}(t)$ in the Heisenberg picture is Hermitean at all times if the time-independent operator in the Schrödinger picture \mathbf{A} is Hermitean, i.e.

$$\mathbf{A}(t)^\dagger = [\mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0)]^\dagger = \mathbf{U}(t, t_0)^\dagger \mathbf{A}^\dagger \mathbf{U}(t, t_0) = \mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0) = \mathbf{A}(t). \tag{9.3.5}$$

Furthermore, the commutation relations of two operators $\mathbf{A}(t)$ and $\mathbf{B}(t)$ are covariant under time-evolution, i.e.

$$\begin{aligned}
[\mathbf{A}(t), \mathbf{B}(t)] &= [\mathbf{U}(t, t_0)^\dagger \mathbf{A} \mathbf{U}(t, t_0), \mathbf{U}(t, t_0)^\dagger \mathbf{B} \mathbf{U}(t, t_0)] \\
&= \mathbf{U}(t, t_0)^\dagger [\mathbf{A}, \mathbf{B}] \mathbf{U}(t, t_0) = [\mathbf{A}, \mathbf{B}](t).
\end{aligned} \tag{9.3.6}$$

9.4 Time-dependent Hamilton Operators

Until now, we have only considered closed systems with time-independent Hamilton operators. Interestingly, as a consequence of energy conservation, the Hamilton operator remained time-independent even in the Heisenberg picture. For open quantum systems that are influenced by external forces, on the other hand, time-translation invariance is broken and, consequently, energy is in general no longer conserved. In that case one has a time-dependent Hamilton operator already in the Schrödinger picture. The Schrödinger equation then takes the form

$$i\hbar \partial_t |\Psi(t)\rangle = \mathbf{H}(t) |\Psi(t)\rangle. \tag{9.4.1}$$

Again, we introduce a time-evolution operator $U(t, t_0)$ that evolves the quantum state

$$|\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle, \quad (9.4.2)$$

from an initial time t_0 to a later time t . Inserting eq.(9.4.2) into the Schrödinger equation one obtains

$$i\hbar\partial_t U(t, t_0) = H(t)U(t, t_0). \quad (9.4.3)$$

A formal integration of this equation yields

$$U(t, t_0) = \mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 H(t_1)U(t_1, t_0). \quad (9.4.4)$$

Iterating this equation one obtains

$$U(t, t_0) = \mathbb{1} + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n H(t_1)H(t_2)\dots H(t_n). \quad (9.4.5)$$

In particular, for a time-independent Hamilton operator H one again obtains

$$\begin{aligned} U(t, t_0) &= \mathbb{1} + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n H^n \\ &= \mathbb{1} + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n H^n \\ &= \exp\left(-\frac{i}{\hbar}H(t-t_0)\right). \end{aligned} \quad (9.4.6)$$

Even for general time-dependent (Hermitian) Hamilton operators the time-evolution operator is unitary. Taking the Hermitian conjugate of eq.(9.4.3) one finds

$$-i\hbar\partial_t U(t, t_0)^\dagger = U(t, t_0)^\dagger H(t). \quad (9.4.7)$$

Thus, we have

$$i\hbar\partial_t [U(t, t_0)^\dagger U(t, t_0)] = -U(t, t_0)^\dagger H(t)U(t, t_0) + U(t, t_0)^\dagger H(t)U(t, t_0) = 0, \quad (9.4.8)$$

and, consequently,

$$U(t, t_0)^\dagger U(t, t_0) = U(t_0, t_0)^\dagger U(t_0, t_0) = \mathbb{1}. \quad (9.4.9)$$

9.5 Dirac's Interaction picture

In the Schrödinger picture of quantum mechanics the time-dependence resides in the wave function and possibly in an explicitly time-dependent Hamiltonian. In the Heisenberg picture, on the other hand, the time-dependence resides in the operators. Dirac introduced the so-called interaction picture which is a hybrid of the Schrödinger and Heisenberg pictures and has time-dependence in both the wave function and the operators. The Dirac picture is useful when a closed system with time-independent Hamilton operator H_0 is influenced by time-dependent external forces described by the interaction Hamiltonian $V(t)$ such that the total Hamilton operator

$$H(t) = H_0 + V(t), \quad (9.5.1)$$

is indeed time-dependent. In this case, one may perform a unitary transformation with the time-evolution operator

$$U_0(t, t_0) = \exp\left[-\frac{i}{\hbar}H_0(t - t_0)\right], \quad (9.5.2)$$

of the unperturbed closed system. This transformation turns the state $|\Psi(t)\rangle$ of the Schrödinger picture into the state

$$|\Psi_D(t)\rangle = U_0(t, t_0)^\dagger |\Psi(t)\rangle, \quad (9.5.3)$$

in Dirac's interaction picture. Similarly, an operator A in the Schrödinger picture turns into the operator

$$A_D(t) = U_0(t, t_0)^\dagger A U_0(t, t_0), \quad (9.5.4)$$

in the Dirac picture. For example, one finds

$$H_{0D}(t) = U_0(t, t_0)^\dagger H_0 U_0(t, t_0) = H_0, \quad V_D(t) = U_0(t, t_0)^\dagger V(t) U_0(t, t_0). \quad (9.5.5)$$

In the interaction picture the Schrödinger equation takes the form

$$\begin{aligned} i\hbar\partial_t |\Psi_D(t)\rangle &= i\hbar\partial_t [U_0(t, t_0)^\dagger |\Psi(t)\rangle] \\ &= -U_0(t, t_0)^\dagger H_0 |\Psi(t)\rangle + U_0(t, t_0)^\dagger i\hbar\partial_t |\Psi(t)\rangle \\ &= U_0(t, t_0)^\dagger (H(t) - H_0) |\Psi(t)\rangle = V_D(t) |\Psi_D(t)\rangle, \end{aligned} \quad (9.5.6)$$

i.e. the equation of motion for the wave function is driven by the interaction $V_D(t)$. Similarly, in the Dirac picture the analog of the Heisenberg equation for the time-dependence of operators takes the form

$$\begin{aligned} \partial_t A_D(t) &= \partial_t [U_0(t, t_0)^\dagger A U_0(t, t_0)] \\ &= U_0(t, t_0)^\dagger \frac{i}{\hbar} H_0 A U_0(t, t_0) - U_0(t, t_0)^\dagger A \frac{i}{\hbar} H_0 U_0(t, t_0) \\ &= \frac{i}{\hbar} U_0(t, t_0)^\dagger [H_0, A] U_0(t, t_0) = \frac{i}{\hbar} [H_{0D}, A_D(t)], \end{aligned} \quad (9.5.7)$$

i.e. the equation of motion for operators is driven by H_{0D} . In the Dirac picture the time-evolution operator is defined by

$$|\Psi_D(t)\rangle = U_D(t, t_0)|\Psi_D(t_0)\rangle. \quad (9.5.8)$$

It takes the form

$$U_D(t, t_0) = \mathbb{1} + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n V_D(t_1)V_D(t_2)\dots V_D(t_n). \quad (9.5.9)$$

In the absence of time-dependent external forces, i.e. for $V(t) = 0$, the Dirac picture reduces to the Heisenberg picture.

Chapter 10

Approximation Methods

10.1 The Variational Approach

In many cases of physical interest it is impossible to find an analytic solution of the Schrödinger equation. Numerical methods are a powerful tool to find accurate approximate solutions. A thorough discussion of numerical methods would deserve a whole course and will not be attempted here. Instead, we will concentrate on approximate analytic methods. An interesting method is the variational approach. It is based on the observation that the time-independent Schrödinger equation

$$H\chi_n(x) = E_n\chi_n(x), \quad (10.1.1)$$

can be derived from a variational principle. The Schrödinger equation results when one varies the wave function $\chi_n(x)$ such that

$$\langle \chi_n | H | \chi_n \rangle - E_n \langle \chi_n | \chi_n \rangle = \int dx \chi_n(x)^* (H - E_n) \chi_n(x), \quad (10.1.2)$$

is minimized. Here the energy E_n appears as a Lagrange multiplier that enforces the constraint

$$\langle \chi_n | \chi_n \rangle = \int dx \chi_n(x)^* \chi_n(x) = 1. \quad (10.1.3)$$

It is not difficult to derive the Schrödinger equation from this variational principle. For example, using

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x), \quad (10.1.4)$$

one can write

$$\begin{aligned} \langle \chi_n | (\mathbf{H} - E_n) | \chi_n \rangle &= \\ \int dx \left[\frac{\hbar^2}{2m} \frac{d\chi_n(x)^*}{dx} \frac{d\chi_n(x)}{dx} + \chi_n(x)^* (V(x) - E_n) \chi_n(x) \right] &= \\ \int dx \mathcal{L}(\chi_n(x), \frac{d\chi_n(x)}{dx}). \end{aligned} \quad (10.1.5)$$

Here we have introduced the Lagrange density

$$\mathcal{L}(\chi_n, \frac{d\chi_n}{dx}) = \frac{\hbar^2}{2m} \frac{d\chi_n(x)^*}{dx} \frac{d\chi_n(x)}{dx} + \chi_n(x)^* (V(x) - E_n) \chi_n(x). \quad (10.1.6)$$

The Euler-Lagrange equation that results from the variational principle then takes the form

$$\frac{d}{dx} \frac{\delta \mathcal{L}}{\delta (d\chi_n^*/dx)} - \frac{\delta \mathcal{L}}{\delta \chi_n^*} = \frac{\hbar^2}{2m} \frac{d^2 \chi_n(x)}{dx^2} - (V(x) - E_n) \chi_n(x) = 0, \quad (10.1.7)$$

which is indeed nothing but the Schrödinger equation. Varying the Lagrange density with respect to $\chi_n(x)$ yields the same equation in complex conjugated form.

The variational approximation now proceeds as follows. When one cannot find the analytic solution $\chi_0(x)$ for a ground state wave function, one makes an ansatz for a trial function $\tilde{\chi}_0(x)$ that depends on a number of variational parameters α_i . Then one computes $\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle$ as well as $\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle$ and one puts

$$\frac{d}{d\alpha_i} \frac{\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle}{\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle} = 0. \quad (10.1.8)$$

The resulting equations for the parameters α_i determine the best approximate solution $\tilde{\chi}_0(x)$ given the chosen ansatz. The approximate ground state energy results as

$$\tilde{E}_0 = \frac{\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle}{\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle}. \quad (10.1.9)$$

The true ground state energy E_0 is the absolute minimum of the right-hand side under arbitrary variations of $\tilde{\chi}_0(x)$. Hence, $\tilde{E}_0 \geq E_0$, i.e. the variational result for the ground state energy is an upper limit on the true value.

Let us illustrate the variational method with an example — a modified 1-dimensional harmonic oscillator problem with the potential

$$V(x) = \frac{1}{2} m \omega^2 x^2 + \lambda x^4. \quad (10.1.10)$$

In the presence of the quartic term λx^4 it is impossible to solve the Schrödinger equation analytically. We now make the variational ansatz

$$\tilde{\chi}_0(x) = A \exp(-\frac{1}{2}\alpha^2 x^2), \quad (10.1.11)$$

for the ground state wave function. This particular ansatz depends on only one variational parameter α . In this case, one finds

$$\begin{aligned} \langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle &= \frac{|A|^2 \sqrt{\pi}}{\alpha} \left[\frac{\hbar^2 \alpha^2}{4m} + \frac{m\omega^2}{4\alpha^2} + \frac{3\lambda}{4\alpha^4} \right], \\ \langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle &= \frac{|A|^2 \sqrt{\pi}}{\alpha}. \end{aligned} \quad (10.1.12)$$

Hence, the variation with respect to α takes the form

$$\frac{d}{d\alpha} \frac{\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle}{\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle} = \frac{\hbar^2 \alpha}{2m} - \frac{m\omega^2}{2\alpha^3} - \frac{3\lambda}{\alpha^5} = 0. \quad (10.1.13)$$

For $\lambda = 0$ this equation implies $\alpha^2 = m\omega/\hbar$ in agreement with our earlier results for the harmonic oscillator. Using this optimized value for the variational parameter α one obtains

$$\tilde{E}_0 = \frac{\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle}{\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle} = \frac{\hbar^2 \alpha^2}{4m} + \frac{m\omega^2}{4\alpha^2} = \frac{\hbar\omega}{2}, \quad (10.1.14)$$

which is indeed the exact ground state energy of a harmonic oscillator. In general, one cannot expect to obtain the exact result from a variational calculation. However, for $\lambda = 0$ our family of ansatz wave functions $\tilde{\chi}_0(x)$ indeed contains the exact wave function. Hence, in this case the variational calculation gives the exact result.

For $\lambda \neq 0$ the equation for α can be written as

$$\alpha^6 - \frac{m^2 \omega^2 \alpha^2}{\hbar^2} - \frac{6\lambda m}{\hbar^2} = 0. \quad (10.1.15)$$

This cubic equation for α^2 can still be solved in closed form, but the result is not very transparent. Hence, we specialize on small values of λ and we write

$$\alpha^2 = \frac{m\omega}{\hbar} + c\lambda + \mathcal{O}(\lambda^2), \quad (10.1.16)$$

which implies $c = 3/m\omega^2$. The variational result for the ground state energy then becomes

$$\tilde{E}_0 = \frac{\langle \tilde{\chi}_0 | \mathbf{H} | \tilde{\chi}_0 \rangle}{\langle \tilde{\chi}_0 | \tilde{\chi}_0 \rangle} = \frac{\hbar^2 \alpha^2}{4m} + \frac{m\omega^2}{4\alpha^2} + \frac{3\lambda}{4\alpha^4} = \frac{\hbar\omega}{2} + \frac{3\lambda \hbar^2}{4m^2 \omega^2}. \quad (10.1.17)$$

10.2 Time-independent Nondegenerate Perturbation Theory to First and Second Order

Perturbation theory is another powerful approximation method. It is applicable when a quantum system for which an analytic solution of the Schrödinger equation exists is influenced by a small additional perturbation. First, we consider time-independent perturbations V and we write the total Hamilton operator as

$$H = H_0 + \lambda V, \quad (10.2.1)$$

where λ is a small parameter that controls the strength of the perturbation. The Hamiltonian H_0 is the one of the unperturbed system for which an analytic solution of the time-independent Schrödinger equation

$$H_0 |\chi_n^{(0)}\rangle = E_n^{(0)} |\chi_n^{(0)}\rangle, \quad (10.2.2)$$

is known. Here $|\chi_n^{(0)}\rangle$ and $E_n^{(0)}$ are the eigenstates and energy eigenvalues of the unperturbed system. The full time-independent Schrödinger equation (including the perturbation) takes the form

$$H |\chi_n\rangle = (H_0 + \lambda V) |\chi_n\rangle = E_n |\chi_n\rangle. \quad (10.2.3)$$

Now we expand the full solution as a superposition of unperturbed states

$$|\chi_n\rangle = \sum_m c_{mn}(\lambda) |\chi_m^{(0)}\rangle. \quad (10.2.4)$$

The expansion coefficients are given by

$$c_{mn}(\lambda) = \langle \chi_m^{(0)} | \chi_n \rangle. \quad (10.2.5)$$

In perturbation theory these coefficients are expanded in a power series in the strength parameter λ , i.e.

$$c_{mn}(\lambda) = \sum_{k=0}^{\infty} c_{mn}^{(k)} \lambda^k. \quad (10.2.6)$$

Similarly, the energy eigenvalues of the perturbed problem are written as

$$E_n(\lambda) = \sum_{k=0}^{\infty} E_n^{(k)} \lambda^k. \quad (10.2.7)$$

First, we consider nondegenerate perturbation theory, i.e. we assume that the unperturbed states $|\chi_n^{(0)}\rangle$ that we consider have distinct energy eigenvalues $E_n^{(0)}$.

10.2. TIME-INDEPENDENT NONDEGENERATE PERTURBATION THEORY TO FIRST AND SECOND ORDER

In that case, the perturbed state $|\chi_n\rangle$ reduces to the unperturbed state $|\chi_n^{(0)}\rangle$ as the perturbation is switched off, and hence

$$c_{mn}(0) = c_{mn}^{(0)} = \delta_{mn}. \quad (10.2.8)$$

Let us consider nondegenerate perturbation theory up to second order in λ , such that

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots, \quad (10.2.9)$$

and

$$|\chi_n\rangle = |\chi_n^{(0)}\rangle + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \lambda^2 \sum_m c_{mn}^{(2)} |\chi_m^{(0)}\rangle + \dots \quad (10.2.10)$$

Inserting this into the Schrödinger equation one obtains

$$\begin{aligned} & (\mathbf{H}_0 + \lambda \mathbf{V}) [|\chi_n^{(0)}\rangle + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \lambda^2 \sum_m c_{mn}^{(2)} |\chi_m^{(0)}\rangle + \dots] = \\ & (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) \times \\ & [|\chi_n^{(0)}\rangle + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \lambda^2 \sum_m c_{mn}^{(2)} |\chi_m^{(0)}\rangle + \dots]. \end{aligned} \quad (10.2.11)$$

Let us consider this equation order by order in λ . To zeroth order we simply obtain the unperturbed Schrödinger equation

$$\mathbf{H}_0 |\chi_n^{(0)}\rangle = E_n^{(0)} |\chi_n^{(0)}\rangle, \quad (10.2.12)$$

while to first order in λ one obtains

$$\mathbf{H}_0 \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \mathbf{V} |\chi_n^{(0)}\rangle = E_n^{(0)} \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + E_n^{(1)} |\chi_n^{(0)}\rangle, \quad (10.2.13)$$

which implies

$$\sum_m c_{mn}^{(1)} (E_m^{(0)} - E_n^{(0)}) |\chi_m^{(0)}\rangle + \mathbf{V} |\chi_n^{(0)}\rangle = E_n^{(1)} |\chi_n^{(0)}\rangle. \quad (10.2.14)$$

Projecting this equation on the state $\langle \chi_l^{(0)} |$ one obtains

$$c_{ln}^{(1)} (E_l^{(0)} - E_n^{(0)}) + \langle \chi_l^{(0)} | \mathbf{V} | \chi_n^{(0)} \rangle = E_n^{(1)} \delta_{ln}. \quad (10.2.15)$$

Hence, putting $l = n$ one obtains

$$E_n^{(1)} = \langle \chi_n^{(0)} | \mathbf{V} | \chi_n^{(0)} \rangle, \quad (10.2.16)$$

i.e. the leading order correction to the energy is given by the expectation value of the perturbation V in the unperturbed state $|\chi_n^{(0)}\rangle$. Putting $l \neq n$ and using the fact that the energies are nondegenerate, i.e. $E_l^{(0)} \neq E_n^{(0)}$, one also finds

$$c_{ln}^{(1)} = \frac{\langle \chi_l^{(0)} | V | \chi_n^{(0)} \rangle}{E_n^{(0)} - E_l^{(0)}}. \quad (10.2.17)$$

It should be noted that the coefficient $c_{nn}^{(1)}$ is not determined by this equation. In fact, it is to some extent arbitrary. Through a choice of phase for the state $|\chi_n\rangle$ one can always make $c_{nn}^{(1)}$ real. Then it can be determined from the normalization condition

$$\langle \chi_n | \chi_n \rangle = \sum_m |c_{mn}(\lambda)|^2 = 1 + 2\lambda c_{nn} + \mathcal{O}(\lambda^2) = 1 \Rightarrow c_{nn}^{(1)} = 0. \quad (10.2.18)$$

As an example, we consider again an anharmonic perturbation $V(x) = \lambda x^4$ to a 1-dimensional harmonic oscillator. The unperturbed ground state wave function is then given by

$$\langle x | \chi_0^{(0)} \rangle = A \exp(-\frac{1}{2}\alpha^2 x^2), \quad (10.2.19)$$

with $\alpha^2 = m\omega/\hbar$ and $|A|^2 = \alpha/\sqrt{\pi}$. We thus obtain

$$E_0^{(1)} = \langle \chi_0^{(0)} | V | \chi_0^{(0)} \rangle = \int dx \lambda x^4 |A|^2 \exp(-\alpha^2 x^2) = \frac{3\lambda\hbar^2}{4m^2\omega^2}. \quad (10.2.20)$$

This result happens to agree with the one of the previous variational calculation expanded to first order in λ . In general, one would not necessarily expect agreement between a variational and a perturbative calculation. The perturbative calculation is based on a systematic power series expansion in the strength λ of the perturbation. Hence, it gives correct results order by order. The precision of a variational calculation, on the other hand, depends on the choice of variational ansatz for the wave function. In our example, the variational ansatz contained the exact wave function of the unperturbed problem. Since that wave function alone determines the first order correction to the energy, in this particular case the variational calculation gives the correct result to first order in λ .

Next, we consider second order perturbation theory. The terms of second order in λ in eq.(10.2.11) imply

$$\begin{aligned} H_0 \sum_m c_{mn}^{(2)} |\chi_m^{(0)}\rangle + V \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle = \\ E_n^{(0)} \sum_m c_{mn}^{(2)} |\chi_m^{(0)}\rangle + E_n^{(1)} \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + E_n^{(2)} |\chi_n^{(0)}\rangle. \end{aligned} \quad (10.2.21)$$

Projecting this equation on the state $\langle \chi_n^{(0)} |$ one finds

$$\begin{aligned} E_n^{(2)} &= \sum_m c_{mn}^{(1)} \langle \chi_n^{(0)} | \mathbf{V} | \chi_m^{(0)} \rangle - E_n^{(1)} c_{nn}^{(1)} \\ &= \sum_{m \neq n} c_{mn}^{(1)} \langle \chi_n^{(0)} | \mathbf{V} | \chi_m^{(0)} \rangle = \sum_{m \neq n} \frac{|\langle \chi_n^{(0)} | \mathbf{V} | \chi_m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \end{aligned} \quad (10.2.22)$$

It should be noted that $E_0^{(2)}$ for the ground state is always negative.

10.3 Time-independent Degenerate Perturbation Theory to First Order

Until now we have assumed that the unperturbed level is not degenerate. Now we consider the degenerate case, i.e. we assume that there are N states $|\chi_t^{(0)}\rangle$ ($t \in \{1, 2, \dots, N\}$) with the same energy $E_n^{(0)}$. This degeneracy may be lifted partly or completely by the perturbation \mathbf{V} . As in the nondegenerate case we write

$$|\chi_n\rangle = \sum_m c_{mn}(\lambda) |\chi_m^{(0)}\rangle. \quad (10.3.1)$$

Since the states $|\chi_t^{(0)}\rangle$ are now degenerate, any linear combination of them is also an energy eigenstate. Consequently, in contrast to the nondegenerate case, one can no longer assume that $c_{mn}(0) = c_{mn}^{(0)} = \delta_{mn}$. Instead, in the degenerate case one has

$$c_{tn}^{(0)} = U_{tn}, \quad (10.3.2)$$

for $t, n \in \{1, 2, \dots, N\}$. Here U is a unitary transformation that turns the arbitrarily chosen basis of unperturbed degenerate states into the basis of (not necessarily degenerate) perturbed states in the limit $\lambda \rightarrow 0$. Hence, to first order in λ we have

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \dots, \quad (10.3.3)$$

and

$$|\chi_n\rangle = \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn} + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \dots \quad (10.3.4)$$

Again, inserting into the Schrödinger equation one obtains

$$\begin{aligned} & (\mathbf{H}_0 + \lambda \mathbf{V}) \left[\sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn} + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \dots \right] = \\ & (E_n^{(0)} + \lambda E_n^{(1)} + \dots) \left[\sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn} + \lambda \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \dots \right]. \end{aligned} \quad (10.3.5)$$

To first order in λ one now obtains

$$\mathbf{H}_0 \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + \mathbf{V} \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn} = E_n^{(0)} \sum_m c_{mn}^{(1)} |\chi_m^{(0)}\rangle + E_n^{(1)} \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn}, \quad (10.3.6)$$

which implies

$$\sum_m c_{mn}^{(1)} (E_m^{(0)} - E_n^{(0)}) |\chi_m^{(0)}\rangle + \mathbf{V} \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn} = E_n^{(1)} \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn}. \quad (10.3.7)$$

Projecting on $\langle \chi_s^{(0)} |$ with $s \in \{1, 2, \dots, N\}$ one obtains

$$\sum_{t=1}^N \langle \chi_s^{(0)} | \mathbf{V} | \chi_t^{(0)} \rangle U_{tn} = E_n^{(1)} U_{sn}. \quad (10.3.8)$$

Multiplying from the left by U_{ls}^\dagger , summing over s , and using $U_{ls}^\dagger U_{sn} = \delta_{ln}$, one obtains

$$E_n^{(1)} \delta_{ln} = \sum_{s,t=1}^N U_{ls}^\dagger \langle \chi_s^{(0)} | \mathbf{V} | \chi_t^{(0)} \rangle U_{tn}, \quad (10.3.9)$$

i.e. to first order in λ , the perturbed energies $E_n^{(1)}$ result from diagonalizing the perturbation \mathbf{V} in the space of degenerate unperturbed states by the unitary transformation U .

10.4 The Hydrogen Atom in a Weak Electric Field

Let us apply perturbation theory to the problem of the hydrogen atom in a weak external electric field. We consider a homogeneous electric field E_z in the z -direction. The corresponding potential energy for an electron at the position \vec{r} is then given by

$$V(\vec{r}) = eE_z z. \quad (10.4.1)$$

The unperturbed Hamilton operator is the one of the hydrogen atom without the external field, i.e.

$$H_0 = -\frac{\hbar^2}{2\mu} - \frac{Ze^2}{|\vec{r}|}. \quad (10.4.2)$$

The unperturbed energy spectrum of bound states is

$$E_n^{(0)} = -\frac{Z^2 e^4 \mu}{2\hbar^2 n^2}. \quad (10.4.3)$$

The corresponding wave functions are

$$\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi), \quad (10.4.4)$$

where $Y_{lm}(\theta, \varphi)$ are the spherical harmonics. In particular, the ground state is a nondegenerate s-state with a spherically symmetric wave function

$$\psi_{100}(\vec{r}) = R_{10}(r)Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}R_{10}(r). \quad (10.4.5)$$

To first order in the perturbation, the shift of the ground state energy,

$$E_n^{(1)} = \langle \psi_{100} | V | \psi_{100} \rangle = \frac{eE_z}{4\pi} \int d^3r R_{10}(r)^2 z = \frac{eE_z}{4\pi} \int d^3r R_{10}(r)^2 r \cos \theta = 0, \quad (10.4.6)$$

vanishes. The first excited state of the unperturbed hydrogen atom is 4-fold degenerate. There are one s-state

$$\psi_{200}(\vec{r}) = R_{20}(r)Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}R_{20}(r). \quad (10.4.7)$$

and three p-states

$$\begin{aligned} \psi_{210}(\vec{r}) &= R_{20}(r)Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}}R_{20}(r)\cos\theta, \\ \psi_{21\pm 1}(\vec{r}) &= R_{20}(r)Y_{1\pm 1}(\theta, \varphi) = \mp\sqrt{\frac{3}{8\pi}}R_{20}(r)\sin\theta\exp(\pm i\varphi). \end{aligned} \quad (10.4.8)$$

We must now evaluate the perturbation in the space of degenerate unperturbed states. The perturbation $V(\vec{r}) = eE_z z$ is invariant against rotations around the z -axis. Consequently, the perturbed Hamiltonian still commutes with L_z (but no longer with L_x and L_y). Thus, the perturbation cannot mix states with different m quantum numbers. The four degenerate states have $m = \pm 1$ for two of the p-states and $m = 0$ for the other p-state and the s-state. Hence, only the $m = 0$ states can mix. The $m = \pm 1$ states are not even shifted in their energies because

$$\langle \psi_{21\pm 1} | V | \psi_{21\pm 1} \rangle = 0. \quad (10.4.9)$$

The energy shifts of the two $m = 0$ states are the eigenvalues of the matrix

$$V = \begin{pmatrix} \langle \psi_{200} | \mathbf{V} | \psi_{200} \rangle & \langle \psi_{200} | \mathbf{V} | \psi_{210} \rangle \\ \langle \psi_{210} | \mathbf{V} | \psi_{200} \rangle & \langle \psi_{210} | \mathbf{V} | \psi_{210} \rangle \end{pmatrix}. \quad (10.4.10)$$

Parity symmetry implies that the diagonal elements vanish

$$\langle \psi_{200} | \mathbf{V} | \psi_{200} \rangle = \langle \psi_{210} | \mathbf{V} | \psi_{210} \rangle = 0. \quad (10.4.11)$$

In addition, for the off-diagonal matrix elements one obtains

$$\langle \psi_{200} | \mathbf{V} | \psi_{210} \rangle = \langle \psi_{210} | \mathbf{V} | \psi_{200} \rangle = -3 \frac{\hbar^2 E_z}{Zem}. \quad (10.4.12)$$

Hence, the energy shifts of the two $m = 0$ states are

$$E^{(1)} = \pm 3 \frac{\hbar^2 E_z}{Zem}. \quad (10.4.13)$$

10.5 Time-independent Nondegenerate Perturbation Theory to All Orders

In the next step we investigate nondegenerate perturbation theory to all orders. In this calculation we put $\lambda = 1$ and we thus write $\mathbf{H} = \mathbf{H}_0 + \mathbf{V}$. Let us consider the so-called resolvent

$$\mathbf{R}(z) = \frac{\mathbb{1}}{z - \mathbf{H}}, \quad (10.5.1)$$

as well as its expectation value in the unperturbed state $|\chi_n^{(0)}\rangle$, i.e. the function

$$R_n(z) = \langle \chi_n^{(0)} | \frac{\mathbb{1}}{z - \mathbf{H}} | \chi_n^{(0)} \rangle. \quad (10.5.2)$$

One pole of $R_n(z)$ is located at the energy eigenvalue E_n of the perturbed Hamiltonian \mathbf{H} . We write

$$\frac{\mathbb{1}}{z - \mathbf{H}} = \frac{\mathbb{1}}{z - \mathbf{H}_0} (z - \mathbf{H} + \mathbf{V}) \frac{\mathbb{1}}{z - \mathbf{H}} = \frac{\mathbb{1}}{z - \mathbf{H}_0} + \frac{\mathbb{1}}{z - \mathbf{H}_0} \mathbf{V} \frac{\mathbb{1}}{z - \mathbf{H}}. \quad (10.5.3)$$

Iterating this equation one obtains

$$\frac{\mathbb{1}}{z - \mathbf{H}} = \frac{\mathbb{1}}{z - \mathbf{H}_0} \sum_{k=0}^{\infty} \left[\mathbf{V} \frac{\mathbb{1}}{z - \mathbf{H}_0} \right]^k, \quad (10.5.4)$$

which implies

$$\begin{aligned}
R_n(z) &= \frac{1}{z - E_n^{(0)}} + \frac{1}{(z - E_n^{(0)})^2} \langle \chi_n^{(0)} | \mathbb{V} \sum_{k=0}^{\infty} \left[\frac{\mathbb{1}}{z - \mathbf{H}_0} \mathbb{V} \right]^k | \chi_n^{(0)} \rangle \\
&= \frac{1}{z - E_n^{(0)}} + \frac{1}{(z - E_n^{(0)})^2} \langle \chi_n^{(0)} | \mathbb{V} \sum_{k=0}^{\infty} \left[\left(\frac{\mathbf{P}_0}{z - E_n^{(0)}} + \frac{\mathbf{Q}_0}{z - \mathbf{H}_0} \right) \mathbb{V} \right]^k | \chi_n^{(0)} \rangle \\
&= \frac{1}{z - E_n^{(0)}} + \frac{1}{(z - E_n^{(0)})^2} \langle \chi_n^{(0)} | \mathbf{A} \sum_{k=0}^{\infty} \left[\frac{\mathbf{P}_0}{z - E_n^{(0)}} \mathbf{A} \right]^k | \chi_n^{(0)} \rangle. \tag{10.5.5}
\end{aligned}$$

Here we have introduced the projection operators

$$\mathbf{P}_0 = |\chi_n^{(0)}\rangle\langle\chi_n^{(0)}|, \quad \mathbf{Q}_0 = \mathbb{1} - \mathbf{P}_0, \tag{10.5.6}$$

on the unperturbed state $|\chi_n^{(0)}\rangle$ and on the rest of the Hilbert space of unperturbed states. We have also defined the operator

$$\mathbf{A} = \mathbb{V} \sum_{k=0}^{\infty} \left[\frac{\mathbf{Q}_0}{z - \mathbf{H}_0} \mathbb{V} \right]^k, \tag{10.5.7}$$

which satisfies

$$\mathbf{A} = \mathbb{V} + \mathbb{V} \frac{\mathbf{Q}_0}{z - \mathbf{H}_0} \mathbf{A}. \tag{10.5.8}$$

Similarly, the operator

$$\mathbf{B} = \langle \chi_n^{(0)} | \mathbb{V} \sum_{k=0}^{\infty} \left[\left(\frac{\mathbf{P}_0}{z - E_n^{(0)}} + \frac{\mathbf{Q}_0}{z - \mathbf{H}_0} \right) \mathbb{V} \right]^k, \tag{10.5.9}$$

obeys the equation

$$\mathbf{B} = \mathbb{V} + \mathbb{V} \left(\frac{\mathbf{P}_0}{z - E_n^{(0)}} + \frac{\mathbf{Q}_0}{z - \mathbf{H}_0} \right) \mathbf{B}. \tag{10.5.10}$$

The last equality in eq.(10.5.5), namely

$$\mathbf{B} = \mathbf{A} + \sum_{k=0}^{\infty} \left[\frac{\mathbf{P}_0}{z - \mathbf{H}_0} \mathbf{A} \right]^k, \tag{10.5.11}$$

is equivalent to

$$\mathbf{B} = \mathbf{A} + \mathbf{A} \frac{\mathbf{P}_0}{z - \mathbf{H}_0} \mathbf{B}. \tag{10.5.12}$$

This is indeed consistent because

$$\begin{aligned}
B &= V + V\left(\frac{P_0}{z - E_n^{(0)}} + \frac{Q_0}{z - H_0}\right)B \\
&= A + V\frac{Q_0}{z - H_0}(B - A) + A\frac{P_0}{z - H_0}B + (V - A)\frac{P_0}{z - H_0}B \\
&= B + V\frac{Q_0}{z - H_0}(B - A) + (V - A)\frac{P_0}{z - H_0}B \\
&= B + V\frac{Q_0}{z - H_0}A\frac{P_0}{z - H_0}B - V\frac{Q_0}{z - H_0}A\frac{P_0}{z - H_0}B = B. \quad (10.5.13)
\end{aligned}$$

Next we introduce the function

$$A_n(z) = \langle \chi_n^{(0)} | A | \chi_n^{(0)} \rangle, \quad (10.5.14)$$

and we evaluate

$$\begin{aligned}
\langle \chi_n^{(0)} | A \sum_{k=0}^{\infty} \left[\frac{P_0}{z - E_n^{(0)}} A \right]^k | \chi_n^{(0)} \rangle &= \langle \chi_n^{(0)} | A \sum_{k=0}^{\infty} \left[\frac{|\chi_n^{(0)}\rangle\langle\chi_n^{(0)}|}{z - E_n^{(0)}} A \right]^k | \chi_n^{(0)} \rangle = \\
A_n(z) \sum_{k=0}^{\infty} \left(\frac{A_n(z)}{z - E_n^{(0)}} \right)^k &= (z - E_n^{(0)}) \frac{A_n(z)}{z - E_n^{(0)} - A_n(z)}. \quad (10.5.15)
\end{aligned}$$

Inserting this in eq.(10.5.5) we finally obtain

$$R_n(z) = \frac{1}{z - E_n^{(0)}} + \frac{1}{z - E_n^{(0)}} \frac{A_n(z)}{z - E_n^{(0)} - A_n(z)} = \frac{1}{z - E_n^{(0)} - A_n(z)}. \quad (10.5.16)$$

Hence, there is a pole in $R_n(z)$ if $z = E_n^{(0)} - A_n(z)$. The pole position determines the perturbed energy $E_n = z$ and hence

$$E_n - E_n^{(0)} = A_n(E_n) = \langle \chi_n^{(0)} | V \sum_{k=0}^{\infty} \left[\frac{Q_0}{E_n - H_0} V \right]^k | \chi_n^{(0)} \rangle. \quad (10.5.17)$$

This is an implicit equation for the perturbed energy E_n valid to all orders in the perturbation V . Expanding this equation to first order we obtain

$$E_n - E_n^{(0)} = E_n^{(1)} = \langle \chi_n^{(0)} | V | \chi_n^{(0)} \rangle, \quad (10.5.18)$$

in complete agreement with our previous result. To second order we find

$$E_n - E_n^{(0)} = E_n^{(1)} + E_n^{(2)} = \langle \chi_n^{(0)} | V | \chi_n^{(0)} \rangle + \langle \chi_n^{(0)} | V \frac{Q_0}{E_n - H_0} V | \chi_n^{(0)} \rangle. \quad (10.5.19)$$

10.6. TIME-INDEPENDENT DEGENERATE OR QUASI-DEGENERATE PERTURBATION THEORY

To order V^2 in the term $Q_0/(E_n - H_0)$ the energy E_n must be replaced by $E_n^{(0)}$ and hence

$$\begin{aligned} E_n^{(2)} &= \langle \chi_n^{(0)} | V \frac{Q_0}{E_n^{(0)} - H_0} V | \chi_n^{(0)} \rangle = \langle \chi_n^{(0)} | V \sum_{m \neq n} \frac{|\chi_m^{(0)}\rangle\langle \chi_m^{(0)}|}{E_n^{(0)} - E_m^{(0)}} V | \chi_n^{(0)} \rangle \\ &= \sum_{m \neq n} \frac{|\langle \chi_n^{(0)} | V | \chi_m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}, \end{aligned} \quad (10.5.20)$$

again in agreement with our previous result.

10.6 Time-independent Degenerate or Quasi-Degenerate Perturbation Theory to All Orders

Let us now consider degenerate perturbation theory to all orders. Again, we assume that there are N states $|\chi_s^{(0)}\rangle$ ($s \in \{1, 2, \dots, N\}$) with the same energy $E_n^{(0)}$. We introduce the projection operators

$$P_0 = \sum_{s=1}^N |\chi_s^{(0)}\rangle\langle \chi_s^{(0)}|, \quad Q_0 = \mathbb{1} - P_0, \quad (10.6.1)$$

on the space of degenerate unperturbed states and its complement in the Hilbert space. We now consider the $N \times N$ matrix-function

$$A_{st}(z) = \langle \chi_s^{(0)} | A | \chi_t^{(0)} \rangle = \langle \chi_s^{(0)} | V \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H_0} V \right]^k | \chi_t^{(0)} \rangle, \quad (10.6.2)$$

which we diagonalize by a unitary transformation

$$|\phi_n^{(0)}\rangle = \sum_{t=1}^N |\chi_t^{(0)}\rangle U_{tn}, \quad \langle \phi_l^{(0)}| = \sum_{s=1}^N U_{ls}^\dagger \langle \chi_s^{(0)}|, \quad (10.6.3)$$

such that

$$\langle \phi_l^{(0)} | V \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H_0} V \right]^k | \phi_n^{(0)} \rangle = \sum_{s,t=1}^N U_{ls}^\dagger A_{st}(z) U_{tn} = A_n(z) \delta_{ln}. \quad (10.6.4)$$

It should be noted that the projector on the set of unitarily transformed unperturbed states $|\phi_n^{(0)}\rangle$ is still given by

$$\begin{aligned} \sum_{n=1}^N |\phi_n^{(0)}\rangle\langle\phi_n^{(0)}| &= \sum_{n,s,t=1}^N |\chi_t^{(0)}\rangle U_{tn} U_{ns}^\dagger \langle\chi_s^{(0)}| \\ &= \sum_{s,t=1}^N |\chi_t^{(0)}\rangle \delta_{ts} \langle\chi_s^{(0)}| = \sum_{s=1}^N |\chi_s^{(0)}\rangle\langle\chi_s^{(0)}| = P_0. \end{aligned} \quad (10.6.5)$$

At this stage we have reduced the degenerate case to the previously discussed nondegenerate case. In particular, the perturbed energy eigenvalues are solutions of the implicit equation

$$E_n - E_n^{(0)} = A_n(E_n) = \langle\phi_n^{(0)}|\mathbf{V} \sum_{k=0}^{\infty} \left[\frac{\mathbf{Q}_0}{E_n - \mathbf{H}_0} \mathbf{V} \right]^k |\phi_n^{(0)}\rangle. \quad (10.6.6)$$

To first order in the perturbation \mathbf{V} we obtain

$$E_n - E_n^{(0)} = E_n^{(1)} = \langle\phi_n^{(0)}|\mathbf{V}|\phi_n^{(0)}\rangle. \quad (10.6.7)$$

To that order the perturbed wave functions are simply the states $|\phi_n^{(0)}\rangle$ ($n \in \{1, 2, \dots, N\}$) in which the perturbation is diagonalized. It should be noted that the original degeneracy is often at least partially lifted by the perturbation. In particular, if the perturbation \mathbf{V} has fewer symmetries than the unperturbed Hamiltonian \mathbf{H}_0 some degeneracies may disappear in the presence of the perturbation.

Sometimes the states of an unperturbed Hamiltonian are almost — but not quite — degenerate. In this case of quasi-degeneracy one might think that non-degenerate perturbation theory would be appropriate. However, if the perturbation gives rise to energy shifts that are comparable with or even larger than the splittings between the almost degenerate levels, one must use quasi-degenerate perturbation theory. Let us consider a group of quasi-degenerate states $|\chi_s^{(0)}\rangle$ ($s \in \{1, 2, \dots, N\}$) with energies $E_s^{(0)}$ that are all close to an energy $E^{(0)}$. In that case, it is useful to consider a modified unperturbed Hamiltonian

$$\mathbf{H}'_0 = \mathbf{H}_0 + \sum_{t=1}^N (E^{(0)} - E_t^{(0)}) |\chi_t^{(0)}\rangle\langle\chi_t^{(0)}|, \quad (10.6.8)$$

and a modified perturbation

$$\mathbf{V}' = \mathbf{V} - \sum_{t=1}^N (E^{(0)} - E_t^{(0)}) |\chi_t^{(0)}\rangle\langle\chi_t^{(0)}|, \quad (10.6.9)$$

10.6. TIME-INDEPENDENT DEGENERATE OR QUASI-DEGENERATE PERTURBATION THEORY

such that the whole problem

$$H'_0 + V' = H_0 + V, \quad (10.6.10)$$

remains unchanged. By construction, the unperturbed quasi-degenerate eigenstates $|\chi_t^{(0)}\rangle$ of the original Hamiltonian are exactly degenerate eigenstates of modified unperturbed Hamiltonian, i.e.

$$\begin{aligned} H'_0 |\chi_s^{(0)}\rangle &= H_0 |\chi_s^{(0)}\rangle + \sum_{t=1}^N (E^{(0)} - E_t^{(0)}) |\chi_t^{(0)}\rangle \langle \chi_t^{(0)} | \chi_s^{(0)}\rangle \\ &= E_s^{(0)} |\chi_s^{(0)}\rangle + (E^{(0)} - E_s^{(0)}) |\chi_s^{(0)}\rangle = E^{(0)} |\chi_s^{(0)}\rangle. \end{aligned} \quad (10.6.11)$$

At this stage one can use the formalism of degenerate perturbation theory. In particular, we consider the function

$$\begin{aligned} A'_{st}(z) &= \langle \chi_s^{(0)} | V' \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H'_0} V' \right]^k | \chi_t^{(0)} \rangle \\ &= \langle \chi_s^{(0)} | (V - \sum_{u=1}^N (E^{(0)} - E_u^{(0)}) |\chi_u^{(0)}\rangle \langle \chi_u^{(0)} |) \\ &\quad \times \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H'_0} (V - \sum_{v=1}^N (E^{(0)} - E_v^{(0)}) |\chi_v^{(0)}\rangle \langle \chi_v^{(0)} |) \right]^k | \chi_t^{(0)} \rangle \\ &= \langle \chi_s^{(0)} | (V - \sum_{u=1}^N (E^{(0)} - E_u^{(0)}) |\chi_u^{(0)}\rangle \langle \chi_u^{(0)} |) \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H'_0} V \right]^k | \chi_t^{(0)} \rangle \\ &= \langle \chi_s^{(0)} | V \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H'_0} V \right]^k | \chi_t^{(0)} \rangle \\ &\quad + (E_s^{(0)} - E^{(0)}) \langle \chi_s^{(0)} | \sum_{k=0}^{\infty} \left[\frac{Q_0}{z - H'_0} V \right]^k | \chi_t^{(0)} \rangle \\ &= A_{st}(z) + (E_s^{(0)} - E^{(0)}) \delta_{st}, \end{aligned} \quad (10.6.12)$$

which we again diagonalize by a unitary transformation such that

$$\sum_{s,t=1}^N U_{ls}^\dagger A'_{st}(z) U_{tn} = A'_n(z) \delta_{ln}. \quad (10.6.13)$$

Hence, the perturbed energy eigenvalues are now solutions of the implicit equation

$$E_n - E^{(0)} = A'_n(E_n). \quad (10.6.14)$$

Chapter 11

Particle in an Electromagnetic Fields

11.1 The Classical Electromagnetic Field

In this chapter we will study the quantum mechanics of a charged particle (e.g. an electron) in a general classical external electromagnetic field. In principle, the electromagnetic field itself should also be treated quantum mechanically. This is indeed possible and naturally leads to quantum electrodynamics (QED). QED is a relativistic quantum field theory — a subject beyond the scope of this course. Here we will limit ourselves to classical electrodynamics. Hence, we will only treat the charged particle moving in the external field (but not the field itself) quantum mechanically.

In previous chapters we have investigated the quantum mechanics of a point particle with coordinate \vec{x} moving in an external potential $V(\vec{x})$. In particular, we have studied the motion of a nonrelativistic electron with negative charge $-e$ in the Coulomb field of an atomic nucleus of positive charge Ze . In that case the potential energy is given by

$$V(\vec{x}) = -e\Phi(\vec{x}) = -\frac{Ze^2}{r}, \quad (11.1.1)$$

where $r = |\vec{x}|$. The electrostatic potential

$$\Phi(\vec{x}) = \frac{Ze}{r}, \quad (11.1.2)$$

is related to the static electric field of the nucleus by

$$\vec{E}(\vec{x}) = -\vec{\nabla}\Phi(\vec{x}) = \frac{Ze}{r^2}\vec{e}_r, \quad (11.1.3)$$

where $\vec{e}_r = \vec{x}/|\vec{x}|$ is the unit-vector in the direction of \vec{x} . The electric field of the atomic nucleus (and hence $\Phi(\vec{x})$) can be obtained as a solution of the first Maxwell equation

$$\vec{\nabla} \cdot \vec{E}(\vec{x}, t) = 4\pi\rho(\vec{x}, t), \quad (11.1.4)$$

with the point-like static charge density of the nucleus given by

$$\rho(\vec{x}, t) = Ze\delta(\vec{x}). \quad (11.1.5)$$

Here we want to investigate the quantum mechanical motion of a charged particle in a general classical external electromagnetic field. For this purpose, we remind ourselves of Maxwell's equations

$$\begin{aligned} \vec{\nabla} \cdot \vec{E}(\vec{x}, t) &= 4\pi\rho(\vec{x}, t), \\ \vec{\nabla} \times \vec{E}(\vec{x}, t) + \frac{1}{c}\partial_t\vec{B}(\vec{x}, t) &= 0, \\ \vec{\nabla} \cdot \vec{B}(\vec{x}, t) &= 0, \\ \vec{\nabla} \times \vec{B}(\vec{x}, t) - \frac{1}{c}\partial_t\vec{E}(\vec{x}, t) &= \frac{4\pi}{c}\vec{j}(\vec{x}, t). \end{aligned} \quad (11.1.6)$$

Adding the time-derivative of the first and c times the divergence of the last equation one obtains the continuity equation

$$\partial_t\rho(\vec{x}, t) + \vec{\nabla} \cdot \vec{j}(\vec{x}, t) = 0, \quad (11.1.7)$$

which guarantees charge conservation.

The electromagnetic fields $\vec{E}(\vec{x}, t)$ and $\vec{B}(\vec{x}, t)$ can be expressed in terms of scalar and vector potentials $\Phi(\vec{x}, t)$ and $\vec{A}(\vec{x}, t)$ as

$$\begin{aligned} \vec{E}(\vec{x}, t) &= -\vec{\nabla}\Phi(\vec{x}, t) - \frac{1}{c}\partial_t\vec{A}(\vec{x}, t), \\ \vec{B}(\vec{x}, t) &= \vec{\nabla} \times \vec{A}(\vec{x}, t). \end{aligned} \quad (11.1.8)$$

Then the homogeneous Maxwell equations

$$\begin{aligned} &\vec{\nabla} \times \vec{E}(\vec{x}, t) + \frac{1}{c}\partial_t\vec{B}(\vec{x}, t) \\ &= -\vec{\nabla} \times \vec{\nabla} \cdot \Phi(\vec{x}, t) - \frac{1}{c}\vec{\nabla} \times \partial_t\vec{A}(\vec{x}, t) + \frac{1}{c}\partial_t\vec{\nabla} \times \vec{A}(\vec{x}, t) = 0, \\ &\vec{\nabla} \cdot \vec{B}(\vec{x}, t) = \vec{\nabla} \cdot \vec{\nabla} \times \vec{A}(\vec{x}, t) = 0, \end{aligned} \quad (11.1.9)$$

11.2. CLASSICAL MOTION OF A CHARGED PARTICLE IN AN EXTERNAL ELECTROMAGNETIC FIELD

are automatically satisfied. The inhomogeneous equations can be viewed as four equations for the four unknown functions $\Phi(\vec{x}, t)$ and $\vec{A}(\vec{x}, t)$.

All fundamental forces in Nature are described by gauge theories. This includes the electromagnetic, weak, and strong forces and even gravity. Gauge theories have a high degree of symmetry. In particular, their classical equations of motion (such as the Maxwell equations in the case of electrodynamics) are invariant against local space-time dependent gauge transformations. In electrodynamics a gauge transformation takes the form

$$\begin{aligned}\Phi(\vec{x}, t)' &= \Phi(\vec{x}, t) + \frac{1}{c}\partial_t\varphi(\vec{x}, t), \\ \vec{A}(\vec{x}, t)' &= \vec{A}(\vec{x}, t) - \vec{\nabla}\varphi(\vec{x}, t).\end{aligned}\tag{11.1.10}$$

Under this transformation the electromagnetic fields

$$\begin{aligned}\vec{E}(\vec{x}, t)' &= -\vec{\nabla}\Phi(\vec{x}, t)' - \frac{1}{c}\partial_t\vec{A}(\vec{x}, t)' = -\vec{\nabla}\Phi(\vec{x}, t) - \frac{1}{c}\partial_t\vec{A}(\vec{x}, t) \\ &\quad - \frac{1}{c}\vec{\nabla}\partial_t\varphi(\vec{x}, t) + \frac{1}{c}\partial_t\vec{\nabla}\varphi(\vec{x}, t) = \vec{E}(\vec{x}, t), \\ \vec{B}(\vec{x}, t)' &= \vec{\nabla} \times \vec{A}(\vec{x}, t)' = \vec{\nabla} \times \vec{A}(\vec{x}, t) - \vec{\nabla} \times \vec{\nabla}\varphi(\vec{x}, t) = \vec{B}(\vec{x}, t),\end{aligned}\tag{11.1.11}$$

remain unchanged — they are gauge invariant. As a consequence, Maxwell's equations themselves are gauge invariant as well. In fact, in a gauge theory only gauge invariant quantities have a physical meaning. The scalar and vector potentials $\Phi(\vec{x}, t)$ and $\vec{A}(\vec{x}, t)$ vary under gauge transformations and are not physically observable. Instead they are mathematical objects with an inherent unphysical gauge ambiguity. Instead, as gauge invariant quantities, the electromagnetic fields $\vec{E}(\vec{x}, t)$ and $\vec{B}(\vec{x}, t)$ are physically observable.

11.2 Classical Motion of a Charged Particle in an External Electromagnetic Field

The motion of a point particle is governed by Newton's equation

$$m\vec{a}(t) = \vec{F}(t).\tag{11.2.1}$$

For a particle with charge $-e$ moving in an external electromagnetic field the force is given by

$$\vec{F}(t) = -e[\vec{E}(\vec{x}(t), t) + \frac{\vec{v}(t)}{c} \times \vec{B}(\vec{x}(t), t)].\tag{11.2.2}$$

Newton's equation can be derived from the action

$$S[\vec{x}(t)] = \int dt \frac{m}{2} \vec{v}(t)^2 - \int dt d^3y [\rho(\vec{y}, t)\Phi(\vec{y}, t) - \vec{j}(\vec{y}, t) \cdot \frac{1}{c} \vec{A}(\vec{y}, t)], \quad (11.2.3)$$

where

$$\begin{aligned} \rho(\vec{y}, t) &= -e\delta(\vec{y} - \vec{x}(t)), \\ \vec{j}(\vec{y}, t) &= -e\vec{v}(t)\delta(\vec{y} - \vec{x}(t)), \end{aligned} \quad (11.2.4)$$

are the charge and current densities of the charged particle at position $\vec{x}(t)$. It is easy to show that charge is conserved, i.e.

$$\partial_t \rho(\vec{y}, t) + \vec{\nabla} \cdot \vec{j}(\vec{y}, t) = 0. \quad (11.2.5)$$

Inserting eq.(11.2.4) into eq.(11.2.3), for the action one obtains

$$S[\vec{x}(t)] = \int dt \left[\frac{m}{2} \vec{v}(t)^2 + e\Phi(\vec{x}(t), t) - e\frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t) \right]. \quad (11.2.6)$$

This action is indeed invariant under gauge transformations because

$$\begin{aligned} &\int dt [\Phi(\vec{x}(t), t)' - \frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t)'] = \\ &\int dt [\Phi(\vec{x}(t), t) + \frac{1}{c} \partial_t \varphi(\vec{x}(t), t) - \frac{\vec{v}(t)}{c} \cdot (\vec{A}(\vec{x}(t), t) - \vec{\nabla} \varphi(\vec{x}(t), t))] = \\ &\int dt [\Phi(\vec{x}(t), t) - \frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t) + \frac{1}{c} \frac{d}{dt} \varphi(\vec{x}(t), t)], \end{aligned} \quad (11.2.7)$$

and because the total derivative

$$\frac{d}{dt} \varphi(\vec{x}(t), t) = \partial_t \varphi(\vec{x}(t), t) + \vec{v} \cdot \vec{\nabla} \varphi(\vec{x}(t), t), \quad (11.2.8)$$

integrates to zero as long as $\varphi(\vec{x}(t), t)$ vanishes in the infinite past and future. Identifying the Lagrange function

$$L = \frac{m}{2} \vec{v}(t)^2 + e\Phi(\vec{x}(t), t) - e\frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t), \quad (11.2.9)$$

it is straightforward to derive Newton's equation as the Euler-Lagrange equation

$$\frac{d}{dt} \frac{\delta L}{\delta v_i(t)} - \frac{\delta L}{\delta x_i} = 0. \quad (11.2.10)$$

The theory can also be formulated in terms of a classical Hamilton function

$$H = \vec{p}(t) \cdot \vec{v}(t) - L, \quad (11.2.11)$$

where \vec{p} is the momentum canonically conjugate to the coordinate \vec{x} . One finds

$$m\vec{v}(t) = \vec{p}(t) + \frac{e}{c}\vec{A}(\vec{x}(t), t), \quad (11.2.12)$$

and thus one obtains

$$H = \frac{1}{2m}[\vec{p}(t) + \frac{e}{c}\vec{A}(\vec{x}(t), t)]^2 - e\Phi(\vec{x}(t), t). \quad (11.2.13)$$

This is indeed consistent because

$$v_i(t) = \frac{dx_i(t)}{dt} = \frac{\partial H}{\partial p_i(t)} = \frac{1}{m}[p_i(t) + \frac{e}{c}A_i(\vec{x}(t), t)]. \quad (11.2.14)$$

The other equation of motion is

$$\frac{dp_i(t)}{dt} = -\frac{\partial H}{\partial x_i(t)} = -\frac{e}{mc}[p_j(t) + \frac{e}{c}A_j(\vec{x}(t), t)]\partial_i A_j(\vec{x}(t), t) + e\partial_i\Phi(\vec{x}(t), t). \quad (11.2.15)$$

It is straightforward to show that these equations of motion are again equivalent to Newton's equation.

11.3 Gauge Invariant Form of the Schrödinger Equation

Remarkably, the gauge invariance of electrodynamics is intimately related to the phase ambiguity of the quantum mechanical wave function. As we have seen earlier, the Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{x}, t) = -\frac{\hbar^2}{2m}\Delta\Psi(\vec{x}, t) + V(\vec{x})\Psi(\vec{x}, t), \quad (11.3.1)$$

determines the wave function only up to a global phase ambiguity

$$\Psi(\vec{x}, t)' = \Psi(\vec{x}, t)\exp(i\phi). \quad (11.3.2)$$

Here ϕ is a constant, independent of space and time.

We now apply the gauge principle to the Schrödinger equation, i.e. we demand that the physics is invariant even under local transformations

$$\Psi(\vec{x}, t)' = \Psi(\vec{x}, t)\exp(i\phi(\vec{x}, t)), \quad (11.3.3)$$

with a space-time dependent phase $\phi(\vec{x}, t)$. Of course, if the wave function $\Psi(\vec{x}, t)$ solves the original Schrödinger equation (11.3.1), the wave function $\Psi(\vec{x}, t)'$ of eq.(11.3.3) in general does not. This is easy to see because

$$\partial_t \Psi(\vec{x}, t)' = [\partial_t \Psi(\vec{x}, t) + i\Psi(\vec{x}, t)\partial_t \phi(\vec{x}, t)] \exp(i\phi(\vec{x}, t)), \quad (11.3.4)$$

contains the second term on the right hand side that was not present in the original Schrödinger equation. However, if the potential energy $V(\vec{x})$ is replaced by a scalar potential $-e\Phi(\vec{x}, t)$, the Schrödinger equation takes the form

$$i\hbar D_t \Psi(\vec{x}, t) = -\frac{\hbar^2}{2m} \Delta \Psi(\vec{x}, t), \quad (11.3.5)$$

with the covariant derivative

$$D_t \Psi(\vec{x}, t) = \partial_t \Psi(\vec{x}, t) - i\frac{e}{\hbar} \Phi(\vec{x}, t) \Psi(\vec{x}, t). \quad (11.3.6)$$

Using the gauge transformation property

$$\Phi(\vec{x}, t)' = \Phi(\vec{x}, t) + \frac{1}{c} \partial_t \varphi(\vec{x}, t), \quad (11.3.7)$$

of the electromagnetic scalar potential, one obtains

$$\begin{aligned} D_t \Psi(\vec{x}, t)' &= \partial_t \Psi(\vec{x}, t)' - i\frac{e}{\hbar} \Phi(\vec{x}, t)' \Psi(\vec{x}, t)' \\ &= [\partial_t \Psi(\vec{x}, t) + i\Psi(\vec{x}, t)\partial_t \phi(\vec{x}, t)] \exp(i\phi(\vec{x}, t)) \\ &\quad - i\frac{e}{\hbar} [\Phi(\vec{x}, t) + \frac{1}{c} \partial_t \varphi(\vec{x}, t)] \Psi(\vec{x}, t) \exp(i\phi(\vec{x}, t)) \\ &= D_t \Psi(\vec{x}, t) \exp(i\phi(\vec{x}, t)), \end{aligned} \quad (11.3.8)$$

provided that we identify

$$\phi(\vec{x}, t) = \frac{e}{\hbar c} \varphi(\vec{x}, t). \quad (11.3.9)$$

We also introduce a space-like covariant derivative

$$\vec{D} \Psi(\vec{x}, t) = \vec{\nabla} \Psi(\vec{x}, t) + i\frac{e}{\hbar c} \vec{A}(\vec{x}, t) \Psi(\vec{x}, t), \quad (11.3.10)$$

which also transforms as

$$\begin{aligned} \vec{D} \Psi(\vec{x}, t)' &= \vec{\nabla} \Psi(\vec{x}, t)' + i\frac{e}{\hbar c} \vec{A}(\vec{x}, t)' \Psi(\vec{x}, t)' \\ &= [\vec{\nabla} \Psi(\vec{x}, t) + i\Psi(\vec{x}, t)\vec{\nabla} \phi(\vec{x}, t)] \exp(i\phi(\vec{x}, t)) \\ &\quad + i\frac{e}{\hbar c} [\vec{A}(\vec{x}, t) - \vec{\nabla} \varphi(\vec{x}, t)] \Psi(\vec{x}, t) \exp(i\phi(\vec{x}, t)) \\ &= \vec{D} \Psi(\vec{x}, t) \exp(i\phi(\vec{x}, t)), \end{aligned} \quad (11.3.11)$$

under a gauge transformation. Using $\vec{p} = (\hbar/i)\vec{\nabla}$ one obtains

$$\frac{\hbar}{i}\vec{D} = \vec{p} + \frac{e}{c}\vec{A}(\vec{x}, t), \quad (11.3.12)$$

which is the quantum version of $m\vec{v}$ from eq.(11.2.12) that we encountered in the classical theory.

Finally, in the Schrödinger equation we replace $\Delta\Psi(\vec{x}, t) = \vec{\nabla} \cdot \vec{\nabla}\Psi(\vec{x}, t)$ with $\vec{D} \cdot \vec{D}\Psi(\vec{x}, t)$ and we obtain

$$i\hbar D_t\Psi(\vec{x}, t) = -\frac{\hbar^2}{2m}\vec{D} \cdot \vec{D}\Psi(\vec{x}, t). \quad (11.3.13)$$

Inserting the explicit form of the covariant derivatives, the Schrödinger equation for a charged particle in an arbitrary external electromagnetic field takes the form

$$i\hbar\left[\partial_t - i\frac{e}{\hbar}\Phi(\vec{x}, t)\right]\Psi(\vec{x}, t) = -\frac{\hbar^2}{2m}\left[\vec{\nabla} + i\frac{e}{\hbar c}\vec{A}(\vec{x}, t)\right] \cdot \left[\vec{\nabla} + i\frac{e}{\hbar c}\vec{A}(\vec{x}, t)\right]\Psi(\vec{x}, t). \quad (11.3.14)$$

This equation is invariant under gauge transformations of the form

$$\begin{aligned} \Phi(\vec{x}, t)' &= \Phi(\vec{x}, t) + \frac{1}{c}\partial_t\varphi(\vec{x}, t), \\ \vec{A}(\vec{x}, t)' &= \vec{A}(\vec{x}, t) - \vec{\nabla}\varphi(\vec{x}, t), \\ \Psi(\vec{x}, t)' &= \Psi(\vec{x}, t)\exp\left(i\frac{e}{\hbar c}\varphi(\vec{x}, t)\right). \end{aligned} \quad (11.3.15)$$

Under this transformation, both sides of the Schrödinger equation change by a factor $\exp(i(e/\hbar c)\varphi(\vec{x}, t))$. Canceling this factor out, the equation remains invariant.

As usual, the wave function $\Psi(\vec{x}, t)$ that solves the gauged Schrödinger equation (11.3.14) can be interpreted as the probability amplitude for finding the particle at position \vec{x} at time t . In particular, the probability density

$$\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2, \quad (11.3.16)$$

is gauge invariant and hence physically meaningful. Again, probability conservation follows from a continuity equation

$$\partial_t\rho(\vec{x}, t) + \vec{\nabla} \cdot \vec{j}(\vec{x}, t) = 0. \quad (11.3.17)$$

However, in the presence of electromagnetic fields the usual probability current must be modified by replacing ordinary with covariant derivatives such that now

$$\vec{j}(\vec{x}, t) = \frac{\hbar}{2mi}[\Psi(\vec{x}, t)^*\vec{D}\Psi(\vec{x}, t) - (\vec{D}\Psi(\vec{x}, t))^*\Psi(\vec{x}, t)]. \quad (11.3.18)$$

11.4 Magnetic Flux Tubes and the Aharonov-Bohm Effect

Let us consider a charged particle moving in the electromagnetic field of an idealized solenoid. Such a solenoid generates magnetic flux inside it but has no outside magnetic or electric field. Hence a classical particle moving outside the solenoid is not affected by it at all. We will see that this is not the case quantum mechanically. We consider an infinitely thin and infinitely long solenoid oriented along the z -axis. The corresponding vector potential is given by

$$A_x(\vec{x}, t) = -\frac{\Phi}{2\pi} \frac{y}{x^2 + y^2}, \quad A_y(\vec{x}, t) = \frac{\Phi}{2\pi} \frac{x}{x^2 + y^2}, \quad A_z(\vec{x}, t) = 0, \quad (11.4.1)$$

while the scalar potential $\Phi(\vec{x}, t) = 0$ vanishes. The corresponding electric and magnetic fields vanish. However, there is an infinitely thin tube of magnetic flux along the z -direction. The quantity Φ (which has nothing to do with the scalar potential $\Phi(\vec{x}, t)$) is the magnetic flux. This follows from

$$\int_S d\vec{f} \cdot \vec{B}(\vec{x}, t) = \int_S d\vec{f} \cdot \vec{\nabla} \times \vec{A}(\vec{x}, t) = \int_{\partial S} d\vec{l} \cdot \vec{A}(\vec{x}, t) = \Phi. \quad (11.4.2)$$

Here S is any 2-d surface pinched by the z -axis and its boundary ∂S is a closed curve winding around this axis. Hence, the magnetic field takes the form

$$B_x(\vec{x}, t) = B_y(\vec{x}, t) = 0, \quad B_z(\vec{x}, t) = \Phi \delta(x) \delta(y), \quad (11.4.3)$$

i.e. it vanishes everywhere except along the z -axis. Since both the electric and the magnetic field vanish outside the solenoid, a classical particle moving outside the solenoid is totally unaffected by it and travels in a straight line with constant velocity. We will see that this is not the case quantum mechanically.

To see this, we first consider a modified double slit experiment, similar to the one discussed in chapter 1, section 1.7. Again, we consider a planar screen oriented in the y - z -plane with two slits located at $x = z = 0$, the first at $y_1 = -d/2$ and the second at $y_2 = d/2$. We place a source of electrons at the point $x = -l, y = z = 0$ a distance l before the screen, and we detect an interference pattern at a parallel detection screen a distance l behind the screen with the two slits. An electron detected at position y at time T may have passed through the slit at $y_1 = -d/2$ or through the one at $y_2 = d/2$, but we don't know through which slit it went. In the first case it has traveled a total distance squared $|\vec{x}(T)|^2 = 2l^2 + d^2/4 + (y + d/2)^2$, while in the second case we have $|\vec{x}(T)|^2 = 2l^2 + d^2/4 + (y - d/2)^2$. Using Feynman's path integral method, the

total probability amplitude is a superposition of contributions from two paths

$$\begin{aligned}
 \Psi(y) &= A \exp\left(\frac{im(2l^2 + d^2/4 + (y - d/2)^2)}{2\hbar T}\right) \\
 &+ A \exp\left(\frac{im(2l^2 + d^2/4 + (y + d/2)^2)}{2\hbar T}\right) \\
 &= A \exp\left(\frac{im(2l^2 + y^2 + d^2/2)}{2\hbar T}\right) \left[\exp\left(\frac{imy d}{2\hbar T}\right) + \exp\left(-\frac{imy d}{2\hbar T}\right) \right] \\
 &= A \exp\left(\frac{im(2l^2 + y^2 + d^2/2)}{2\hbar T}\right) 2 \cos\left(\frac{my d}{2\hbar T}\right). \tag{11.4.4}
 \end{aligned}$$

To obtain the probability density we take the absolute value squared

$$\rho(y) = |\Psi(y)|^2 = 4|A|^2 \cos^2\left(\frac{my d}{2\hbar T}\right) = 4\rho_0 \cos^2\left(\frac{my d}{2\hbar T}\right). \tag{11.4.5}$$

This is the typical interference pattern of a double slit experiment.

Now we modify the situation by adding the solenoid discussed before. It is oriented along the z -axis within the screen, right between the two slits. As a result, electrons passing through the different slits pass the solenoid on different sides. Classically, this would make no difference because there is no electromagnetic field outside the solenoid. Quantum mechanically, however, there is an effect, as first suggested by Aharonov and Bohm. The set-up with the two slits and the solenoid is known as an Aharonov-Bohm experiment and its result is the Aharonov-Bohm effect. To investigate the Aharonov-Bohm effect, we must include the electromagnetic contribution

$$\begin{aligned}
 S_{em}[\vec{x}(t)] &= \int dt \left[e\Phi(\vec{x}(t), t) - e\frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t) \right] \\
 &= \int dt e\frac{\vec{v}(t)}{c} \cdot \vec{A}(\vec{x}(t), t) = \frac{e}{c} \int_{\mathcal{C}} \vec{dl} \cdot \vec{A}(\vec{x}, t) \tag{11.4.6}
 \end{aligned}$$

to the action in the Feynman path integral. Here \mathcal{C} is the path along which the particle travels. The relative phase between the two paths \mathcal{C}_1 and \mathcal{C}_2 contributing in the double slit experiment, that is generated by this term, is given by

$$\begin{aligned}
 \frac{1}{\hbar} \{ S_{em}[\vec{x}_1(t)] - S_{em}[\vec{x}_2(t)] \} &= \frac{e}{\hbar c} \left\{ \int_{\mathcal{C}_1} \vec{dl} \cdot \vec{A}(\vec{x}, t) - \int_{\mathcal{C}_2} \vec{dl} \cdot \vec{A}(\vec{x}, t) \right\} = \\
 \frac{e}{\hbar c} \int_{\mathcal{C}} \vec{dl} \cdot \vec{A}(\vec{x}, t) &= \frac{e\Phi}{\hbar c}. \tag{11.4.7}
 \end{aligned}$$

Here $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ is the closed curve obtained by combining the two paths \mathcal{C}_1 and \mathcal{C}_2 . By construction, this curve winds around the solenoid and hence the

corresponding integral yields the magnetic flux φ . The resulting interference pattern of the modified double slit experiment is now given by

$$\rho(y)' = |\Psi(y)|^2 = 4\rho_0 \cos^2\left(\frac{myd}{2\hbar T} + \frac{e\Phi}{2\hbar c}\right). \quad (11.4.8)$$

In other words, the presence of the magnetic flux Φ causes an observable effect in the interference pattern. Interestingly, when the magnetic flux is quantized in integer units n , i.e. when

$$\Phi = 2\pi n \frac{\hbar c}{e}, \quad (11.4.9)$$

the solenoid becomes invisible to the electron and thus behaves classically.

11.5 Flux Quantization for Monopoles and Superconductors

In 1931 Dirac investigated the quantum mechanics of magnetically charged particles — so-called magnetic monopoles. There is no experimental evidence for such particles, and Maxwell's equation $\vec{\nabla} \cdot \vec{B}(\vec{x}, t) = 0$ indeed says that magnetic monopoles do not exist. Still, there are some theories — the so-called grand unified extensions of the standard model of particle physics — that naturally contain magnetically charged particles. The magnetic field of a point-like magnetic charge (a so-called Dirac monopole) is determined from

$$\vec{\nabla} \cdot \vec{B}(\vec{x}, t) = 4\pi g \delta(\vec{x}), \quad (11.5.1)$$

where g is the magnetic charge of the monopole. The resulting magnetic field is given by

$$\vec{B}(\vec{x}, t) = \frac{g}{r^2} \vec{e}_r, \quad (11.5.2)$$

and the corresponding magnetic flux through any surface S surrounding the monopole is given by

$$\Phi = \int_S d\vec{f} \cdot \vec{B}(\vec{x}, t) = 4\pi g. \quad (11.5.3)$$

The magnetic field of a monopole can not be obtained from a vector potential as $\vec{B}(\vec{x}, t) = \vec{\nabla} \times \vec{A}(\vec{x}, t)$. In order to still be able to work with a vector potential Dirac introduced a mathematical trick — the so-called Dirac string. A Dirac string is an infinitely thin solenoid that emanates from the monopole and carries its magnetic flux g to infinity. Of course, the Dirac string (being just a mathematical trick)

must be physically invisible. This is the case if the magnetic charge g and thus the magnetic flux of the solenoid obeys the Dirac quantization condition

$$g = \frac{\Phi}{4\pi} = \frac{\hbar c}{2e} \Rightarrow eg = \frac{\hbar c}{2}. \quad (11.5.4)$$

The Dirac quantization condition implies that the units e and g of electric and magnetic charge are related. In particular, the existence of magnetic monopoles would immediately explain why electric charge is quantized. This is one reason why many physicists find grand unified theories an attractive idea for going beyond the standard model of particle physics. However, despite numerous experimental efforts, nobody has ever observed a single magnetic charge. Grand unified theories predict that magnetic monopoles must have existed immediately after the big bang. Then, why don't we see any of them today? This question has been answered by Alan Guth from MIT. In his scenario of the inflationary universe with an exponential expansion early on magnetic monopoles get so much diluted that it would indeed be impossible to find any in the universe today.

In the presence of magnetic monopoles flux quantization would simply follow the Dirac quantization condition. Interestingly, magnetic flux is also quantized inside superconductors, despite the fact that no magnetic monopoles seem to exist in the universe. Superconductivity is a very interesting effect that occurs in metals at very low temperatures. Below a critical temperature, electric currents flow in a superconductor without any resistance. This effect is due to a pair formation of electrons. The so-called Cooper pairs (consisting of two electrons) are bosons and can thus undergo Bose-Einstein condensation. The electrons in the condensate of Cooper pairs are in a coherent quantum state that leads to superconductivity. In order to form a Cooper pair, the electrons must overcome their Coulomb repulsion. This is possible inside a crystal lattice because the quantized lattice oscillations — the phonons — mediate an attractive force between electrons. There are other interesting materials — the more recently discovered ceramic high-temperature superconductors — for which the mechanism of Cooper pair formation is not yet understood. Both the ordinary low-temperature superconductors and the more exotic high-temperature superconductors show the so-called Meissner effect: magnetic fields are expelled from the superconductor. When a ring of non-superconducting metal is placed in an external magnetic field, the magnetic flux permeates the material. When the ring is cooled down below the critical temperature it becomes superconducting and the magnetic flux is expelled. Some of the flux is expelled to the outside, some to the inside of the ring. It turns out that the magnetic flux trapped inside the ring is quantized in units of

$$\Phi = 2\pi n \frac{\hbar c}{2e}. \quad (11.5.5)$$

Interestingly, the quantization is such that the trapped flux does not give rise to an observable Aharonov-Bohm phase for a Cooper pair of charge $2e$ moving around the ring. The wave function of an electron (of charge e), on the other hand, would pick up a minus sign when it moves around the ring. The flux quantization in superconductors is a dynamical effect. If the trapped magnetic flux would not be quantized, the Aharonov-Bohm effect for Cooper pairs would destroy their coherence and hence the superconductivity. It is energetically favorable to maintain the superconducting state and have an invisible quantized trapped flux. It is irrelevant that the trapped flux is visible to unpaired electrons because those do not superconduct.

11.6 Charged Particle in a Constant Magnetic Field

Let us now consider the case of a constant homogeneous external magnetic field $\vec{B} = B\vec{e}_z$ in the z -direction which can be obtained from the potentials

$$\Phi(\vec{x}, t) = 0, \quad A_x(\vec{x}, t) = 0, \quad A_y(\vec{x}, t) = Bx, \quad A_z(\vec{x}, t) = 0. \quad (11.6.1)$$

It should be noted that this choice of $\Phi(\vec{x}, t)$ and $\vec{A}(\vec{x}, t)$ is to some extent ambiguous. Other gauge equivalent choices yield the same physical results.

First, we consider a classical particle moving in the x - y -plane perpendicular to the magnetic field. The particle experiences the Lorentz force

$$\vec{F}(t) = -e\frac{\vec{v}(t)}{c} \times B\vec{e}_z, \quad (11.6.2)$$

which forces the particle on a circular orbit of some radius r . It moves along the circle with an angular velocity ω , which implies the linear velocity $v = \omega r$ and the acceleration $a = \omega^2 r$. Hence, Newton's equation takes the form

$$m\omega^2 r = e\frac{\omega r}{c} B \Rightarrow \omega = \frac{eB}{mc}. \quad (11.6.3)$$

The so-called cyclotron frequency ω is independent of the radius r .

Next, we consider the same problem semiclassically, i.e. using Bohr-Sommerfeld quantization. For this purpose, we first compute the action of a classical periodic cyclotron orbit $\vec{x}_c(t)$. Using eq.(11.2.9) one obtains

$$S[\vec{x}_c(t)] = \int_0^T dt L = \frac{1}{2}m\omega^2 r^2 T + \frac{eB\pi r^2}{c}. \quad (11.6.4)$$

Similarly, using eq.(11.2.13) one finds for the energy

$$E = H = \frac{1}{2}m\omega^2 r^2. \quad (11.6.5)$$

Hence, the Bohr-Sommerfeld quantization condition takes the form

$$S[\vec{x}_c(t)] - ET = 2\pi\hbar n \Rightarrow r^2 = \frac{2\hbar n}{m\omega}. \quad (11.6.6)$$

Consequently, in quantum mechanics the allowed radii of cyclotron orbits are now quantized. Inserting the above expression for r^2 into the energy one obtains

$$E = \hbar\omega n. \quad (11.6.7)$$

Interestingly, up to a constant $\hbar\omega/2$ the semiclassical quantized energy values are those of a harmonic oscillator with the cyclotron frequency ω .

Finally, we consider the problem fully quantum mechanically. The Schrödinger equation (11.3.14) then takes the form

$$\begin{aligned} i\hbar\partial_t\Psi(\vec{x}, t) &= -\frac{\hbar^2}{2m}[\partial_x^2 + (\partial_y + i\frac{eBx}{\hbar c})(\partial_y + i\frac{eBx}{\hbar c}) + \partial_z^2]\Psi(\vec{x}, t) \\ &= -\frac{\hbar^2}{2m}[\partial_x^2 + \partial_y^2 + 2i\frac{eBx}{\hbar c}\partial_y - (\frac{eBx}{\hbar c})^2 + \partial_z^2]\Psi(\vec{x}, t). \end{aligned} \quad (11.6.8)$$

The corresponding time-independent Schrödinger equation takes the form

$$-\frac{\hbar^2}{2m}[\partial_x^2 + \partial_y^2 + 2i\frac{eBx}{\hbar c}\partial_y - (\frac{eBx}{\hbar c})^2 + \partial_z^2]\Psi(\vec{x}) = E\Psi(\vec{x}). \quad (11.6.9)$$

Again, we want to consider motion in the x - y -plane. We make the factorized ansatz $\Psi(\vec{x}) = \psi(x) \exp(ip_y y/\hbar)$ (which implies $p_z = 0$) and we obtain

$$[-\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m\omega^2(x + \frac{p_y}{m\omega})^2]\psi(x) = E\psi(x). \quad (11.6.10)$$

Indeed, this is the equation of motion of a (shifted) harmonic oscillator. Hence, the quantum mechanical energy spectrum takes the form

$$E = \hbar\omega(n + \frac{1}{2}). \quad (11.6.11)$$

Interestingly, the energy of the charged particle is completely independent of the transverse momentum p_y . As a result, the energy levels — which are known as Landau levels — have an infinite degeneracy.

11.7 The Quantum Hall Effect

The quantum Hall effect is one of the most remarkable effects in condensed matter physics in which quantum mechanics manifests itself on macroscopic scales. One distinguishes three effects: the classical Hall effect discovered by the student Hall in 1879, the integer quantum Hall effect discovered by von Klitzing in 1980 (Nobel prize in 1985), and the fractional quantum Hall effect discovered by Tsui, Störmer and Gossard in 1982 and theoretically explained by Laughlin (Nobel prize in 1998).

The basic experimental set-up is the same in the three cases. Some flat (quasi 2-dimensional) rectangular sample of size $L_x \times L_y$ (oriented in the x - y -plane) of some conducting or semi-conducting material is placed in a strong magnetic field $\vec{B} = B\vec{e}_z$ perpendicular to the sample. A current with density

$$\vec{j} = -nev\vec{e}_y, \quad (11.7.1)$$

is flowing through the sample in the y -direction, i.e. $j_y = -nev$. Here n is the number of electrons per unit area which move with the velocity $\vec{v} = v\vec{e}_y$. The electrons moving in the magnetic field experience the Lorentz force

$$\vec{F} = -e\frac{\vec{v}}{c} \times \vec{B}. \quad (11.7.2)$$

The resulting sideways motion of the electrons leads to an excess of negative charge on one edge of the sample which generates an electric field transverse to the current. One then observes a voltage drop along the transverse direction, the x -direction in this case. More electrons move to the edge of the sample until the resulting electric field

$$\vec{E} = -\frac{\vec{v}}{c} \times \vec{B}, \quad (11.7.3)$$

exactly compensates the Lorentz force. Hence, in equilibrium we have

$$\vec{E} = \frac{\vec{j}}{nec} \times \vec{B}. \quad (11.7.4)$$

The resistivity tensor ρ is defined by

$$\vec{E} = \rho\vec{j}, \quad (11.7.5)$$

such that we can identify

$$\rho = \frac{B}{nec} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (11.7.6)$$

The conductivity tensor σ is the matrix inverse of the resistivity tensor ρ and it is defined by

$$\vec{j} = \sigma \vec{E}, \quad (11.7.7)$$

such that

$$\sigma = \frac{ne c}{B} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (11.7.8)$$

Since both the resistivity and the conductivity tensor are purely off-diagonal the material behaves paradoxically. On the one hand, it looks insulating because $\sigma_{xx} = 0$, i.e. there is no current in the direction of the electric field. On the other hand, it looks like a perfect conductor because $\rho_{xx} = 0$. The Hall resistivity is given by

$$\rho_{xy} = \frac{B}{ne c}. \quad (11.7.9)$$

This is indeed the result of the classical Hall effect. The resistivity is linear in the magnetic field B and depends on the electron density n . The classical Hall effect can thus be used to determine the electron density by measuring the Hall resistivity as well as the applied magnetic field.

Remarkably, at sufficiently low temperatures eq.(11.7.9) is not the correct result in real samples. In fact, one observes deviations from the linear behavior in the form of quantum Hall plateaux, i.e. the Hall resistivity becomes independent of the magnetic field in regions that correspond to the quantized values

$$\rho_{xy} = \frac{h}{\nu e^2}. \quad (11.7.10)$$

Here h is Planck's quantum, e is the electric charge, and ν is an integer. This so-called integer quantum Hall effect was first observed by von Klitzing and his collaborators. Later, quantum Hall plateaux were also observed at fractional values (e.g. at $\nu = 1/3, 2/5, 3/7$) by Tsui, Störmer and Gossard. This is now known as the fractional quantum Hall effect. Very surprisingly, the value of the resistivity is independent of any microscopic details of the material, the purity of the sample, the precise value of the magnetic field, etc. In fact, the quantum Hall effect is now used to maintain the standard of resistance. At each of the quantum Hall plateaux the dissipative resistivity ρ_{xx} practically drops to zero (by about 13 orders of magnitude).

In addition to the previous discussion of Landau levels we should now add the transverse electric field $\vec{E} = E_x \vec{e}_x$ generated by the sideways moving electrons.

The Schrödinger equation then takes the form

$$\begin{aligned} & \left[-\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m\omega^2\left(x + \frac{p_y}{m\omega}\right)^2 + eE_x x\right]\psi(x) = \\ & \left[-\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m\omega^2\left(x + \frac{p_y}{m\omega} + \frac{eE_x}{m\omega^2}\right)^2 - \frac{p_y eE_x}{m\omega} - \frac{e^2 E_x^2}{2m\omega^2}\right]\psi(x) = E\psi(x). \end{aligned} \quad (11.7.11)$$

Again, this is the equation of motion of a shifted harmonic oscillator, now centered around

$$x_0 = -\frac{p_y}{m\omega} - \frac{eE_x}{m\omega^2}. \quad (11.7.12)$$

It is straightforward to construct the current density and to verify explicitly that eq.(11.7.3) is indeed satisfied for the quantum mechanical solution.

Let us count the number of states in each of the Landau levels for a Hall sample of size $L_x \times L_y$. For simplicity, we choose periodic boundary conditions in the y -direction, such that

$$p_y = \frac{2\pi\hbar}{L_y}n_y, \quad (11.7.13)$$

with $n_y \in \mathbb{Z}$. Since the wave function must be located inside the Hall sample one must have $x_0 \in [0, L_x]$. The number N of n_y values for which this is the case, and hence the degeneracy of a given Landau level, is

$$N = \frac{m\omega L_x L_y}{2\pi\hbar} = \frac{eBL_x L_y}{2\pi\hbar c}. \quad (11.7.14)$$

When ν Landau levels are completely occupied by $N\nu$ electrons, the number density (per unit area) of electrons is given by

$$n = \frac{N\nu}{L_x L_y} = \frac{eB\nu}{2\pi\hbar c}. \quad (11.7.15)$$

Using eq.(11.7.9) one then finds

$$\rho_{xy} = \frac{B}{ne c} = \frac{h}{\nu e^2}, \quad (11.7.16)$$

i.e. the quantum Hall plateaux correspond to completely filled Landau levels. It is clear that complete filling implies a very special situation for the electron system in the Hall sample. In particular, in that case there are no available low-lying energy levels into which interacting electrons could scatter, and thus dissipation is extremely suppressed. A full understanding of the integer quantum Hall effect requires the incorporation of impurities in the sample. In particular,

it is nontrivial to understand the enormous precision with which eq.(11.7.9) is realized despite numerous imperfections in actual materials. To understand this goes beyond the scope of the present discussion.

In contrast to the integer quantum Hall effect, the fractional quantum Hall effect relies heavily on electron-electron correlations. It is a complicated collective phenomenon in which electrons break up into quasiparticles with fractional charge and with fractional spin — so-called anyons. Again, the explanation of the fractional quantum Hall effect is beyond the scope of this course.

11.8 The “Normal” Zeeman Effect

Let us briefly discuss the so-called “normal” Zeeman effect, which describes the shift of energy levels of hydrogen-like atoms in a weak homogeneous external magnetic field $\vec{B} = B\vec{e}_z$ ignoring the effects of spin. When spin is included one speaks of the “anomalous” Zeeman effect. This is clearly a case of bad nomenclature because normally spin plays an important role and the “anomalous” Zeeman effect is the more common phenomenon. In this case, it is natural to choose the so-called symmetric gauge

$$A_x(\vec{x}, t) = -\frac{By}{2}, \quad A_y(\vec{x}, t) = \frac{Bx}{2}, \quad A_z(\vec{x}, t) = 0. \quad (11.8.1)$$

Then to leading order in the magnetic field we have

$$(p_x + \frac{e}{c}A_x)^2 + (p_y + \frac{e}{c}A_y)^2 = p_x^2 + p_y^2 + \frac{eB}{c}(xp_y - yp_x) + \mathcal{O}(B^2). \quad (11.8.2)$$

Hence, the “normal” Zeeman effect results from the Hamilton operator

$$H = H_0 + \frac{eBL_z}{2\mu c}, \quad (11.8.3)$$

where H_0 is the unperturbed problem (at $B = 0$) and μ is the reduced mass of the electron and the atomic nucleus. Due to the rotation invariance of the unperturbed problem $[H_0, \vec{L}] = 0$. The external magnetic field explicitly breaks rotation invariance. Only the rotations around the z -axis remain a symmetry of the perturbed problem and hence $[H, L_z] = 0$. Since the unperturbed states of hydrogen-like atoms can be chosen as eigenstates of L_z with eigenvalue $\hbar m$, they remain eigenstates when the magnetic field is switched on. However, the energy eigenvalue is now shifted to

$$E = E_0 + \frac{eB\hbar m}{2\mu c}. \quad (11.8.4)$$

As a result, the $(2l + 1)$ -fold degeneracy of states with angular momentum quantum number l is lifted.

Chapter 12

Coupling of Quantum Mechanical Angular Momenta

12.1 Quantum Mechanical Angular Momentum

Angular momentum is a fundamental quantity that is conserved in any known physical process. Angular momentum conservation is a consequence of the isotropy of space — the laws of Nature are invariant against spatial rotations. Of course, rotation invariance may be broken explicitly under certain conditions, for example, in the presence of external electric or magnetic fields. In that case, the subsystem without the fields is not rotation invariant. Still, the total system behaves in the same way when everything including the fields is spatially rotated. Angular momentum is a vector. In quantum mechanics its components cannot be measured simultaneously because the corresponding operators do not commute. This has interesting consequences for the physical behavior of quantum mechanical particles under spatial rotations.

Besides the orbital angular momentum $\vec{L} = \vec{r} \times \vec{p}$ familiar from classical mechanics, quantum mechanical particles can carry an internal angular momentum known as spin. While orbital angular momentum is quantized in integer units, spin may be quantized in integer or half-integer units. Interestingly, there is an intimate connection between spin and statistics: particles with half-integer spin obey Fermi-Dirac statistics and are thus fermions, while particles with integer spin obey Bose-Einstein statistics and are hence bosons. This connection between spin and statistics can be understood in the framework of relativistic quantum field theory — but not from quantum mechanics alone.

When particles carry both orbital angular momentum \vec{L} and spin \vec{S} , in general only their total angular momentum $\vec{J} = \vec{L} + \vec{S}$ is conserved. In that case, one is confronted with the problem of coupling two angular momenta together. The same problem arises when several particles add their angular momenta together to the conserved angular momentum of the total system. Performing the corresponding angular momentum “gymnastics” is an important tool of the quantum mechanic. Although the subject is a bit formal, its understanding is vital in atomic, molecular and particle physics, as well as in other branches of our field.

Let us consider the commutation relations of an arbitrary angular momentum \vec{J} in quantum mechanics. Here \vec{J} may be an orbital angular momentum, a spin, or any combination of these. For an orbital angular momentum \vec{L} we have already derived the commutation relation

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k, \quad (12.1.1)$$

from the definition $\vec{L} = \vec{r} \times \vec{p}$ and from the fundamental commutation relation $[x_i, p_j] = i\hbar\delta_{ij}$. Now we postulate

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k, \quad (12.1.2)$$

for any angular momentum in quantum mechanics. In particular, different components of the angular momentum vector do not commute with one another. However, all components commute with the magnitude $\vec{J}^2 = J_x^2 + J_y^2 + J_z^2$, i.e. $[J_i, \vec{J}^2] = 0$. As a consequence, one can construct simultaneous eigenstates of the operator \vec{J}^2 and one component J_i . Usually one chooses J_z , i.e. the arbitrary quantization direction is then the z -direction. In general, this choice does not imply a physical violation of rotation invariance. One could have chosen any other quantization axis without any effect on the physics. When rotation invariance is already broken, for example, by the direction of an external electric or magnetic field, it is very convenient to choose the quantization axis along the same direction. As usual, we choose the z -direction as our quantization axis and we thus construct simultaneous eigenstates $|j, m\rangle$ of both \vec{J}^2 and J_z ,

$$\vec{J}^2|j, m\rangle = \hbar^2 j(j+1)|j, m\rangle, \quad J_z|j, m\rangle = \hbar m|j, m\rangle. \quad (12.1.3)$$

It will turn out that both j and m are either integer or half-integer.

For convenience, we introduce the angular momentum raising and lowering operators

$$J_{\pm} = J_x \pm iJ_y. \quad (12.1.4)$$

They obey the commutation relations

$$[J_+, J_-] = 2\hbar J_z, \quad [\vec{J}^2, J_{\pm}] = 0, \quad [J_z, J_{\pm}] = \pm\hbar J_{\pm}. \quad (12.1.5)$$

We have

$$J_z J_{\pm} |j, m\rangle = (J_{\pm} J_z + [J_z, J_{\pm}]) |j, m\rangle = \hbar(m \pm 1) J_{\pm} |j, m\rangle, \quad (12.1.6)$$

i.e. the state $J_{\pm} |j, m\rangle$ is also an eigenstate of J_z and has the quantum number $m \pm 1$. Similarly

$$\vec{J}^2 J_{\pm} |j, m\rangle = J_{\pm} \vec{J}^2 |j, m\rangle = \hbar^2 j(j+1) J_{\pm} |j, m\rangle, \quad (12.1.7)$$

i.e. $J_{\pm} |j, m\rangle$ is still also an eigenstate of \vec{J}^2 with the unchanged quantum number j . Since j determines the magnitude of the angular momentum vector, one expects that for fixed j the m quantum number that measures the z -component of the angular momentum vector should be bounded from above and from below. On the other hand, for any state $|j, m\rangle$ with quantum number m one can construct the states

$$J_{\pm} |j, m\rangle = C_{j,m} |j, m \pm 1\rangle, \quad (12.1.8)$$

with quantum number $m \pm 1$. The apparent contradiction is resolved only if the constant $C_{j,m}$ vanishes for a given $m = m_{max}$ or $m = m_{min}$. Let us compute $C_{j,m}$ from the normalization condition $\langle j, m \pm 1 | j, m \pm 1 \rangle = 1$. Using the relation

$$J_{\pm}^{\dagger} J_{\pm} = J_{\mp} J_{\pm} = \vec{J}^2 - J_z^2 \mp \hbar J_z, \quad (12.1.9)$$

we obtain

$$\begin{aligned} |C_{j,m}|^2 &= |C_{j,m}|^2 \langle j, m \pm 1 | j, m \pm 1 \rangle = \langle j, m | J_{\pm}^{\dagger} J_{\pm} | j, m \rangle \\ &= \langle j, m | \vec{J}^2 - J_z^2 \mp \hbar J_z | j, m \rangle \\ &= \hbar^2 [j(j+1) - m(m \pm 1)]. \end{aligned} \quad (12.1.10)$$

For fixed j , the maximal value m_{max} of the quantum number m is determined by $C_{j,m_{max}} = 0$, which implies $m_{max} = j$. Similarly, the minimal value is determined by $C_{j,m_{min}} = 0$, which implies $m_{min} = -j$. Hence, for fixed j there are $2j + 1$ possible m values

$$m \in \{m_{min}, m_{min} + 1, \dots, m_{max} - 1, m_{max}\} = \{-j, -j + 1, \dots, j - 1, j\}. \quad (12.1.11)$$

Since the difference $m_{max} - m_{min} = 2j$ is an integer, j can be an integer or a half-integer. Both possibilities are realized in Nature.

Let us now couple two angular momenta \vec{J}_1 and \vec{J}_2 together to a total angular momentum

$$\vec{J} = \vec{J}_1 + \vec{J}_2. \quad (12.1.12)$$

The angular momenta \vec{J}_1 and \vec{J}_2 could, for example, be orbital angular momentum and spin of the same particle, or angular momenta of two different particles.

In any case, since they act in different Hilbert spaces the two angular momentum operators commute with one another, i.e.

$$[J_{1i}, J_{2j}] = 0. \quad (12.1.13)$$

As a consequence, the total angular momentum operator $\vec{J} = \vec{J}_1 + \vec{J}_2$ indeed obeys the usual commutation relations

$$[J_i, J_j] = [J_{1i}, J_{1j}] + [J_{2i}, J_{2j}] = i\hbar\varepsilon_{ijk}(J_{1k} + J_{2k}) = i\hbar\varepsilon_{ijk}J_k. \quad (12.1.14)$$

Let us assume that the states of subsystem 1 have a fixed quantum number j_1 and are given by $|j_1, m_1\rangle$. Similarly, the states of subsystem 2 have quantum number j_2 and are given by $|j_2, m_2\rangle$. Hence, the combined system has $(2j_1 + 1)(2j_2 + 1)$ product states $|j_1, m_1\rangle|j_2, m_2\rangle$ which span the Hilbert space of the total system. How does this space decompose into sectors of definite total angular momentum? It will turn out that the possible values for j are restricted by

$$j \in \{|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2\}. \quad (12.1.15)$$

Indeed, the total number of states then is

$$(2|j_1 - j_2| + 1) + (2|j_1 - j_2| + 3) + \dots + (2(j_1 + j_2) + 1) = (2j_1 + 1)(2j_2 + 1). \quad (12.1.16)$$

Also the question arises how one can construct states

$$|(j_1, j_2)j, m\rangle = \sum_{m_1, m_2} C_{m_1, m_2} |j_1, m_1\rangle |j_2, m_2\rangle, \quad (12.1.17)$$

as linear combinations of the product states? This is the “gymnastics” problem of coupling together two angular momenta. The factors C_{m_1, m_2} are known as Clebsch-Gordan coefficients.

12.2 Coupling of Spins

Let us consider the spins of two spin 1/2 particles, for example, a proton and a neutron forming the atomic nucleus of heavy hydrogen (deuterium). The corresponding bound state of proton and neutron is known as a deuteron. What are the possible total spins of the coupled system? In this case $j_1 = j_2 = 1/2$ and thus

$$j \in \{|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2\} = \{0, 1\}, \quad (12.2.1)$$

i.e. the total spin j is either 0 (a singlet) or 1 (a triplet). Altogether, there are four states. This is consistent because there are also four product states $|j_1, m_1\rangle|j_2, m_2\rangle$ with $m_1 = \pm 1/2$ and $m_2 = \pm 1/2$. For these four states we introduce the short-hand notation

$$\begin{aligned} |\frac{1}{2}, \frac{1}{2}\rangle|\frac{1}{2}, \frac{1}{2}\rangle &= |\uparrow\uparrow\rangle, & |\frac{1}{2}, \frac{1}{2}\rangle|\frac{1}{2}, -\frac{1}{2}\rangle &= |\uparrow\downarrow\rangle, \\ |\frac{1}{2}, -\frac{1}{2}\rangle|\frac{1}{2}, \frac{1}{2}\rangle &= |\downarrow\uparrow\rangle, & |\frac{1}{2}, -\frac{1}{2}\rangle|\frac{1}{2}, -\frac{1}{2}\rangle &= |\downarrow\downarrow\rangle. \end{aligned} \quad (12.2.2)$$

These product states are eigenstates of $J_z = J_{1z} + J_{2z}$,

$$\begin{aligned} J_z|\uparrow\uparrow\rangle &= \hbar(\frac{1}{2} + \frac{1}{2})|\uparrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle, \\ J_z|\uparrow\downarrow\rangle &= \hbar(\frac{1}{2} - \frac{1}{2})|\uparrow\downarrow\rangle = 0, & J_z|\downarrow\uparrow\rangle &= \hbar(-\frac{1}{2} + \frac{1}{2})|\downarrow\uparrow\rangle = 0, \\ J_z|\downarrow\downarrow\rangle &= \hbar(-\frac{1}{2} - \frac{1}{2})|\downarrow\downarrow\rangle = -\hbar|\downarrow\downarrow\rangle. \end{aligned} \quad (12.2.3)$$

The first and the last of the four states must belong to $j = 1$ because they have $m = \pm 1$, i.e.

$$|(\frac{1}{2}, \frac{1}{2})1, 1\rangle = |\uparrow\uparrow\rangle, \quad |(\frac{1}{2}, \frac{1}{2})1, -1\rangle = |\downarrow\downarrow\rangle. \quad (12.2.4)$$

The two remaining states have $m = 0$. One linear combination of them is the $m = 0$ state of the triplet, and the orthogonal combination is the $m = 0$ state with $j = 0$. In order to identify the $m = 0$ state with $j = 1$ we act with the lowering operator

$$|(\frac{1}{2}, \frac{1}{2})1, 0\rangle = \frac{1}{\sqrt{2}}J_-|(\frac{1}{2}, \frac{1}{2})1, 1\rangle = \frac{1}{\sqrt{2}}(J_{1-} + J_{2-})|\uparrow\uparrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle). \quad (12.2.5)$$

The orthogonal combination

$$|(\frac{1}{2}, \frac{1}{2})0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (12.2.6)$$

should hence be the state with $j = 0$. This can be checked explicitly, for example, by acting with the operator $\vec{J}_+ = J_{1+} + J_{2+}$, i.e.

$$\vec{J}_+|(\frac{1}{2}, \frac{1}{2})0, 0\rangle = (J_{1+} + J_{2+})\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle - |\uparrow\uparrow\rangle) = 0. \quad (12.2.7)$$

Hence, according to the previous discussion the spin of the deuteron could be 0 or 1. The proton and neutron that form the deuteron nucleus attract each other through the so-called strong interaction. This interaction is spin-dependent.

There is a term proportional to $-\vec{J}_1 \cdot \vec{J}_2$ (where \vec{J}_1 and \vec{J}_2 are the spin operators of proton and neutron) in the proton-neutron potential. One can write

$$\vec{J}^2 = (\vec{J}_1 + \vec{J}_2)^2 = \vec{J}_1^2 + \vec{J}_2^2 + 2\vec{J}_1 \cdot \vec{J}_2, \quad (12.2.8)$$

and hence

$$-\vec{J}_1 \cdot \vec{J}_2 = \frac{1}{2}(\vec{J}_1^2 + \vec{J}_2^2 - \vec{J}^2). \quad (12.2.9)$$

Since both the proton and the neutron have spin 1/2 we have

$$\vec{J}_1^2 = \vec{J}_2^2 = \hbar^2 \frac{1}{2} \left(\frac{1}{2} + 1 \right) = \frac{3}{4} \hbar^2. \quad (12.2.10)$$

In the spin singlet state we have $\vec{J}^2 = 0$ and hence

$$-\vec{J}_1 \cdot \vec{J}_2 |(\frac{1}{2}, \frac{1}{2})0, 0\rangle = \frac{3}{4} \hbar^2 |(\frac{1}{2}, \frac{1}{2})0, 0\rangle. \quad (12.2.11)$$

In the spin triplet state, on the other hand, $\vec{J}^2 = 2\hbar^2$ such that

$$-\vec{J}_1 \cdot \vec{J}_2 |(\frac{1}{2}, \frac{1}{2})1, m\rangle = -\frac{1}{4} \hbar^2 |(\frac{1}{2}, \frac{1}{2})1, m\rangle. \quad (12.2.12)$$

In the triplet channel the proton-neutron interaction is attractive while in the singlet channel it is repulsive. As a consequence, the deuteron has spin 1, while in the spin 0 channel there is no bound state.

The strong interaction is mediated by gluons, just as electromagnetic interactions are mediated by photons. The strong interaction analogs of electrons and positrons are quarks and anti-quarks — the basic building blocks of protons and neutrons. There are two u-quarks and one d-quark in each proton. A u-quark has electric charge 2/3 and a d-quark has $-1/3$. Hence, the charge of a proton is indeed $2 \times 2/3 - 1/3 = 1$. Similarly, a neutron contains one u-quark and two d-quarks and thus has charge $2/3 - 2 \times 1/3 = 0$. Like electrons, quarks are fermions with spin 1/2. What is the possible total spin of a bound system of three quarks? As we learned before, two spin 1/2 particles can couple to a total spin $j = 0$ or $j = 1$. When a third spin 1/2 particle is added to the $j = 0$ state, the total spin is 1/2. If it is added to the $j = 1$ state the total spin can be either $j - 1/2 = 1/2$ or $j + 1/2 = 3/2$. Again, the strong interactions between quarks are spin-dependent and they favor the total spin 1/2 states corresponding to proton and neutron. The spin 3/2 state also exists but is unstable. It is known in particle physics as the Δ -isobar.

12.3 Coupling of Orbital Angular Momentum and Spin

Just as there are spin-dependent strong interactions between protons and neutrons or between quarks, there are also spin-dependent electromagnetic interactions between electrons and protons. As we already discussed in the context of the relativistic Dirac equation, there are spin-orbit coupling terms proportional to

$$\vec{L} \cdot \vec{S} = \frac{1}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2) = \frac{\hbar^2}{2}(j(j+1) - l(l+1) - \frac{3}{4}). \quad (12.3.1)$$

Let us consider the coupling of an orbital angular momentum l and a spin $s = 1/2$ of an electron in more detail. In this case, there are $2(2l+1)$ product states $|l, m_l\rangle|s, m_s\rangle$. The possible values of the total angular momentum are

$$j \in \{|l-s|, l+s\} = \{l - \frac{1}{2}, l + \frac{1}{2}\}. \quad (12.3.2)$$

Indeed, there are again

$$2(l - \frac{1}{2}) + 1 + 2(l + \frac{1}{2}) + 1 = 2(2l+1) \quad (12.3.3)$$

states. The direct product state

$$|(l, \frac{1}{2}) l + \frac{1}{2}, l + \frac{1}{2}\rangle = |l, l\rangle|\uparrow\rangle \quad (12.3.4)$$

has $m = l + 1/2$ and must thus have $j = l + 1/2$. We can construct the other states with $j = l + 1/2$ and with lower m -values by acting with $J_- = L_- + S_-$. For example,

$$\begin{aligned} |(l, \frac{1}{2}) l + \frac{1}{2}, l - \frac{1}{2}\rangle &= \frac{1}{\sqrt{2l+1}} J_- |l, l\rangle|\uparrow\rangle \\ &= \frac{1}{\sqrt{2l+1}} (L_- + S_-) |l, l\rangle|\uparrow\rangle \\ &= \frac{1}{\sqrt{2l+1}} (|l, l\rangle|\downarrow\rangle + \sqrt{2l}|l, l-1\rangle|\uparrow\rangle). \end{aligned} \quad (12.3.5)$$

The orthogonal combination

$$|(l, \frac{1}{2}) l - \frac{1}{2}, l - \frac{1}{2}\rangle = \frac{1}{\sqrt{2l+1}} (\sqrt{2l}|l, l\rangle|\downarrow\rangle - |l, l-1\rangle|\uparrow\rangle), \quad (12.3.6)$$

has $j = l - 1/2$ and $m = l - 1/2$. Again, by acting with J_- one can generate all other states with $j = l - 1/2$ and smaller m -values.

Chapter 13

Systems of Identical Particles

13.1 Bosons and Fermions

Identical elementary particles (like two photons or two electrons) are indistinguishable from one another — they are like perfect twins. Ultimately, this is a consequence of quantum field theory. Elementary particles are quantized excitations of fields that fill all of space. Since the properties of the field are the same everywhere in space, a particle created here has exactly the same properties as a particle created somewhere else. In classical physics one usually thinks about a particle in a different way. Starting from some initial positions, one can use Newton's equation to predict the particles' paths and one can always keep track of where a given particle came from. Classically, one can always distinguish two particles by referring to their initial positions. One can say: "This is the same particle that originally came from the position \vec{x}_1 ". It is classically distinguishable from another particle that started from the position \vec{x}_2 . However, photons or electrons are not classical particles. As elementary particles they must be treated quantum mechanically, and at the quantum level the concept of a classical path breaks down. In quantum mechanics it is impossible to keep track of a particle by constantly watching it as it travels along a path. Instead a quantum system of N particles is described by a multi-particle wave function $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$, where $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$ are the positions of the particles. The absolute value $|\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2$ determines the probability density for finding the particles at the positions $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$. Consequently, the normalization condition takes the form

$$\int d^3x_1 d^3x_2 \dots d^3x_N |\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2 = 1. \quad (13.1.1)$$

Similarly, the probability density for finding one of the particles at the position \vec{x}_1 while the other particles are in arbitrary positions is given by

$$\rho(\vec{x}_1) = \int d^3x_2 \dots d^3x_N |\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2 = 1. \quad (13.1.2)$$

As we said above, identical elementary particles are absolutely indistinguishable. Hence, when two identical particles are interchanged, this has no physically observable effects. Let us denote the pair permutation operator of two identical particles by P_{ij} . It acts on a wave function by interchanging the corresponding particle coordinates. For example,

$$P_{12}\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \Psi(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N). \quad (13.1.3)$$

For identical particles the states $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ and its permutation partner $\Psi(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N)$ are physically equivalent. However, this does not imply that the two wave functions are the same. In fact, they may still differ by a complex phase which drops out from physical observables.

The Hamilton operator of a system of nonrelativistic identical particles can be written as

$$H = -\frac{\hbar^2}{2m} \sum_{i=1}^N \Delta_i + V(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N), \quad (13.1.4)$$

Here m is the particle mass and Δ_i is the Laplace operator that takes second derivatives with respect to the coordinates of \vec{x}_i . It should be noted that the masses of all the particles are the same — otherwise they would not be indistinguishable. The potential energy $V(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ depends on the coordinates of all particles. For indistinguishable particles the potential energy must be invariant against particle permutations P_{ij} , for example,

$$P_{12}V(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = V(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N) = V(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N). \quad (13.1.5)$$

One obtains

$$\begin{aligned} P_{12}H\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) &= \left[-\frac{\hbar^2}{2m} \sum_{i=1}^N \Delta_i + V(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N)\right]\Psi(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N) \\ &= HP_{12}\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N), \end{aligned} \quad (13.1.6)$$

and thus $P_{12}H = HP_{12}$. The same is true for all permutations and hence

$$[H, P_{ij}] = 0. \quad (13.1.7)$$

This means that the eigenstates of H can be chosen such that they are simultaneously eigenstates of P_{ij} . For example, an eigenstate of P_{12} obeys

$$P_{12}\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \lambda\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N), \quad (13.1.8)$$

where λ is the corresponding eigenvalue. When we apply the permutation operator twice we obtain

$$\begin{aligned} P_{12}^2\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) &= P_{12}\Psi(\vec{x}_2, \vec{x}_1, \dots, \vec{x}_N) = \Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \\ &= \lambda^2\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \Rightarrow \lambda = \pm 1. \end{aligned} \quad (13.1.9)$$

It should be noted that, in general, different pair permutation operators do not commute. For example, for three particles the operators P_{12} and P_{23} do not commute. The combinations $P_{312} = P_{12}P_{23}$ and $P_{231} = P_{23}P_{12}$ correspond to two different cyclic permutations. As a consequence, in general one cannot find simultaneous eigenstates to all permutation operators P_{ij} .

The permutations P_{ij} form a finite group — the permutation group S_N . Since different permutations, in general, do not commute, the group S_N is non-Abelian. In this respect it is like the group of spatial rotations in three dimensions, $SO(3)$. However, unlike the permutation group S_N , the rotation group $SO(3)$ is continuous: there is an infinite (continuous) number of rotations, but there is only a finite (discrete) number of permutations. Mathematically speaking, the angular momentum multiplets characterized by j form $(2j + 1)$ -dimensional irreducible representations of the group $SO(3)$. Similarly, the irreducible representations of the group S_N have different dimensions. There are two different irreducible representations of the permutation group of two particles S_2 : the symmetric and the antisymmetric one, which are both 1-dimensional. In fact, all permutation groups S_N have two 1-dimensional representations: the totally symmetric and the totally antisymmetric one. However, for $N \geq 3$ the permutation group S_N also has irreducible representations of larger dimension. For example, besides the 1-dimensional symmetric and antisymmetric representations the group S_3 has a 2-dimensional irreducible representation of mixed symmetry.

The indistinguishability of elementary particles has profound consequences for their wave function. For example, in a quantum state in which two particles are located at the positions \vec{x}_1 and \vec{x}_2 , it makes no sense to ask which particle is in which position. Hence, the distinct states $|\vec{x}_1\rangle|\vec{x}_2\rangle$ and $|\vec{x}_2\rangle|\vec{x}_1\rangle$ cannot be correct wave functions for such a quantum state. Only if the two particles were distinguishable it would matter if the particle of type 1 is at position \vec{x}_1 and the particle of type 2 is at position \vec{x}_2 ($|\vec{x}_1\rangle|\vec{x}_2\rangle$) or vice versa ($|\vec{x}_2\rangle|\vec{x}_1\rangle$). For indistinguishable particles there is only one state describing the two identical

particles at the positions \vec{x}_1 and \vec{x}_2 . Since the Hamiltonian commutes with the permutation operator one can choose this state as a simultaneous eigenstate of H and P_{12} . Since P_{12} has eigenvalues ± 1 , one can construct a symmetric state

$$|s\rangle = \frac{1}{\sqrt{2}}(|\vec{x}_1\rangle|\vec{x}_2\rangle + |\vec{x}_2\rangle|\vec{x}_1\rangle), \quad (13.1.10)$$

as well as an antisymmetric state

$$|a\rangle = \frac{1}{\sqrt{2}}(|\vec{x}_1\rangle|\vec{x}_2\rangle - |\vec{x}_2\rangle|\vec{x}_1\rangle). \quad (13.1.11)$$

By construction, we have

$$P_{12}|s\rangle = |s\rangle, \quad P_{12}|a\rangle = -|a\rangle. \quad (13.1.12)$$

For three identical particles one can again construct a totally symmetric state

$$\begin{aligned} |s\rangle &= \frac{1}{\sqrt{6}}(|\vec{x}_1\rangle|\vec{x}_2\rangle|\vec{x}_3\rangle + |\vec{x}_2\rangle|\vec{x}_1\rangle|\vec{x}_3\rangle + |\vec{x}_2\rangle|\vec{x}_3\rangle|\vec{x}_1\rangle \\ &+ |\vec{x}_3\rangle|\vec{x}_2\rangle|\vec{x}_1\rangle + |\vec{x}_3\rangle|\vec{x}_1\rangle|\vec{x}_2\rangle + |\vec{x}_1\rangle|\vec{x}_3\rangle|\vec{x}_2\rangle), \end{aligned} \quad (13.1.13)$$

as well as an antisymmetric state

$$\begin{aligned} |a\rangle &= \frac{1}{\sqrt{6}}(|\vec{x}_1\rangle|\vec{x}_2\rangle|\vec{x}_3\rangle - |\vec{x}_2\rangle|\vec{x}_1\rangle|\vec{x}_3\rangle + |\vec{x}_2\rangle|\vec{x}_3\rangle|\vec{x}_1\rangle \\ &- |\vec{x}_3\rangle|\vec{x}_2\rangle|\vec{x}_1\rangle + |\vec{x}_3\rangle|\vec{x}_1\rangle|\vec{x}_2\rangle - |\vec{x}_1\rangle|\vec{x}_3\rangle|\vec{x}_2\rangle). \end{aligned} \quad (13.1.14)$$

For any pair permutation operator P_{ij} we then have

$$P_{ij}|s\rangle = |s\rangle, \quad P_{ij}|a\rangle = -|a\rangle. \quad (13.1.15)$$

Besides the totally symmetric and antisymmetric states, for three particles there are also states of mixed symmetry.

13.2 The Pauli Principle

Remarkably, there is an intimate relation between the statistics of particles (i.e. their properties under permutations) and their spin. Bosons obeying Bose-Einstein statistics have a totally symmetric wave function and have integer spins. Fermions obeying Fermi-Dirac statistics, on the other hand, have a totally antisymmetric wave function and have half-integer spin. This connection between spin and statistics is known as the Pauli principle. It cannot be derived within the

theoretical framework of quantum mechanics alone. However, it can be derived from quantum field theory in the form of the so-called spin-statistics theorem.

First, we consider bosons with spin 0 whose spin wave function is trivial. The Pauli principle implies that their orbital wave function $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ must be totally symmetric, i.e. for any pair permutation

$$P_{ij}\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N). \quad (13.2.1)$$

For example, two (or three) spinless bosons located at the positions \vec{x}_1, \vec{x}_2 (and \vec{x}_3) are described by the states $|s\rangle$ constructed above.

For particles with spin the situation is a bit more complicated. Let us now consider fermions with spin 1/2, whose spin wave function is nontrivial. The Pauli principle dictates that the total wave function (orbital and spin) is antisymmetric under particle exchange. This does not imply that the orbital or spin wave functions alone must be antisymmetric. Two spin 1/2 particles can couple their spins to a total spin $j = 0$ or $j = 1$. The singlet state with spin 0 given by

$$|(\frac{1}{2}, \frac{1}{2})0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (13.2.2)$$

is antisymmetric under the exchange of the two spins. In that case, in order to obtain a totally antisymmetric wave function, the orbital wave function $\Psi(\vec{x}_1, \vec{x}_2)$ must be symmetric, i.e. $P_{12}\Psi(\vec{x}_1, \vec{x}_2) = \Psi(\vec{x}_1, \vec{x}_2)$. On the other hand, the triplet states with spin $j = 1$ given by

$$\begin{aligned} |(\frac{1}{2}, \frac{1}{2})1, 1\rangle &= |\uparrow\uparrow\rangle, \\ |(\frac{1}{2}, \frac{1}{2})1, 0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \\ |(\frac{1}{2}, \frac{1}{2})1, -1\rangle &= |\downarrow\downarrow\rangle, \end{aligned} \quad (13.2.3)$$

are symmetric under spin exchange. Hence, in that case the orbital wave function must be antisymmetric, i.e. $P_{12}\Psi(\vec{x}_1, \vec{x}_2) = -\Psi(\vec{x}_1, \vec{x}_2)$.

Let us now consider three identical spin 1/2 fermions, for example, three quarks inside a proton or neutron. As we have discussed before, the spins of three spin 1/2 particles can couple to a total spin 1/2 or 3/2. For three quarks the coupling to a total spin 1/2 is energetically favored by the spin-dependence of the strong interactions. In that case, the spin wave function is of mixed symmetry. Here we consider the simpler case of coupling to total spin 3/2 which leads to the unstable Δ -isobar. For example, if we combine three u-quarks we obtain the

Δ^{++} particle with electric charge $3 \times 2/3 = 2$. In this state the quark spins are coupled to a total spin $j = 3/2$. The corresponding $2j + 1 = 4$ states

$$\begin{aligned} |(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\frac{3}{2}, \frac{3}{2}\rangle &= |\uparrow\uparrow\uparrow\rangle, \\ |(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\frac{3}{2}, \frac{1}{2}\rangle &= \frac{1}{\sqrt{3}}(|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle), \\ |(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\frac{3}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}}(|\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\rangle), \\ |(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\frac{3}{2}, -\frac{3}{2}\rangle &= |\downarrow\downarrow\downarrow\rangle, \end{aligned} \quad (13.2.4)$$

are totally symmetric under spin exchange. It turns out that the orbital wave function of the Δ -isobar is also totally symmetric. This seems to contradict the Pauli principle. The quarks are spin 1/2 fermions and still they seem to have a totally symmetric wave function. In order to resolve this puzzle a new quantum number for quarks was postulated by Gell-Mann. Besides orbital quantum numbers and spin quarks also carry “color” — the generalized charge of the strong interactions. There are three different color states of a quark, arbitrarily called red, green and blue. Indeed, the color wave function of the three quarks inside a Δ -isobar (and also of a proton or neutron) is totally antisymmetric

$$|a\rangle = \frac{1}{\sqrt{6}}(|rgb\rangle - |rbg\rangle + |gbr\rangle - |grb\rangle + |brg\rangle - |bgr\rangle). \quad (13.2.5)$$

This state forms a 1-dimensional representation of the color gauge group $SU(3)$ (a color singlet). In particular, the three quarks all have different colors. The strong interactions are not only spin-dependent but, most important, color-dependent. In particular, quarks are always in a color-singlet state. This leads to one of the most important properties of the strong interactions — the so-called quark confinement. Despite numerous experimental efforts, no single quark has ever been observed. They are always confined together in color-singlet groups of three, forming protons, neutrons, Δ -isobars or other strongly interacting elementary particles.

13.3 The Helium Atom

The helium atom consists of three particles: the atomic nucleus and two electrons, interacting through their mutual Coulomb potentials. The Helium nucleus contains $Z = 2$ protons and two neutrons and thus has electric charge $Ze = 2e$. The two negatively charged electrons make the whole atom electrically neutral.

The helium nucleus is about 8000 times heavier than an electron and can thus, to a very good approximation, be treated as fixed in space. The Hamiltonian then takes the form

$$H = H_1 + H_2 + V_{12}. \quad (13.3.1)$$

Here

$$H_i = -\frac{\hbar^2}{2M}\Delta_i - \frac{Ze^2}{|\vec{r}_i|} \quad (13.3.2)$$

is the Hamilton operator for electron $i \in \{1, 2\}$ at position \vec{r}_i in the Coulomb field of the nucleus located at the origin. The Laplace operator Δ_i entering the kinetic energy of electron i takes second derivatives with respect to the coordinates \vec{r}_i . The term

$$V_{12} = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \quad (13.3.3)$$

represents the Coulomb repulsion between the two electrons. The Hamilton operator H does not contain any spin-dependent forces because we have neglected tiny relativistic effects like the spin-orbit interaction. Still, as a consequence of the Pauli principle, we will find that the spectrum depends on the spin.

The spins $s = 1/2$ of the two electrons can couple to a total spin $S = 0$ or $S = 1$. The corresponding spin states are antisymmetric for $S = 0$, i.e.

$$|(\frac{1}{2} \frac{1}{2})0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (13.3.4)$$

and symmetric for $S = 1$, i.e.

$$\begin{aligned} |(\frac{1}{2} \frac{1}{2})1, 1\rangle &= |\uparrow\uparrow\rangle, \\ |(\frac{1}{2} \frac{1}{2})1, 0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \\ |(\frac{1}{2} \frac{1}{2})1, -1\rangle &= |\downarrow\downarrow\rangle. \end{aligned} \quad (13.3.5)$$

The $S = 0$ states are referred to as para-, and the $S = 1$ states are referred to as ortho-helium. The total wave function of the two electrons is given by

$$|\Psi\rangle = \Psi(\vec{r}_1, \vec{r}_2)|(\frac{1}{2} \frac{1}{2})S, S_z\rangle. \quad (13.3.6)$$

Here the orbital wave function $\Psi(\vec{r}_1, \vec{r}_2)$ solves the time-independent Schrödinger equation

$$H\Psi(\vec{r}_1, \vec{r}_2) = E\Psi(\vec{r}_1, \vec{r}_2). \quad (13.3.7)$$

In addition, according to the Pauli principle, $\Psi(\vec{r}_1, \vec{r}_2)$ must be symmetric under the exchange of the coordinates \vec{r}_1 and \vec{r}_2 for para-helium ($S = 0$), and antisymmetric for ortho-helium ($S = 1$). Symmetric and antisymmetric wave functions will in general have different energies, such that a spin-dependent spectrum emerges although the Hamiltonian does not contain explicitly spin-dependent forces.

Solving the Schrödinger equation for the helium atom is much more difficult than for hydrogen and can, in fact, not be done in closed form. Instead, one can use numerical methods, perturbation theory, or a variational approach. Here we perform an approximate calculation treating V_{12} as a small perturbation. As a first step, we completely neglect the Coulomb repulsion between the electrons and we put $V_{12} = 0$. Then the Hamiltonian $H = H_1 + H_2$ separates into two hydrogen-like problems. The corresponding single-particle Schrödinger equation

$$H_i \Psi_{n_i, l_i, m_i}(\vec{r}_i) = E_{n_i} \Psi_{n_i, l_i, m_i}(\vec{r}_i) \quad (13.3.8)$$

has been solved earlier. The energy eigenvalues for bound states are given by

$$E_n = -\frac{Z^2 e^4 M}{2\hbar^2 n^2}. \quad (13.3.9)$$

The correctly normalized and (anti-)symmetrized two-particle orbital wave function then takes the form

$$\Psi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} (\Psi_{n_1, l_1, m_1}(\vec{r}_1) \Psi_{n_2, l_2, m_2}(\vec{r}_2) \pm \Psi_{n_1, l_1, m_1}(\vec{r}_2) \Psi_{n_2, l_2, m_2}(\vec{r}_1)). \quad (13.3.10)$$

The plus-sign corresponds to para-helium (spin state antisymmetric, orbital wave function symmetric) and the minus-sign corresponds to ortho-helium (spin state symmetric, orbital wave function antisymmetric). Neglecting V_{12} , the ground state has $n_1 = n_2 = 1$, $l_1 = l_2 = 0$ and thus $m_1 = m_2 = 0$. In that case, the antisymmetric orbital wave function

$$\Psi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} (\Psi_{1,0,0}(\vec{r}_1) \Psi_{1,0,0}(\vec{r}_2) - \Psi_{1,0,0}(\vec{r}_2) \Psi_{1,0,0}(\vec{r}_1)), \quad (13.3.11)$$

simply vanishes. Hence, the ground state is para-helium (a non-degenerate spin singlet) with the symmetric orbital wave function

$$\begin{aligned} \Psi(\vec{r}_1, \vec{r}_2) &= \frac{1}{\sqrt{2}} (\Psi_{1,0,0}(\vec{r}_1) \Psi_{1,0,0}(\vec{r}_2) + \Psi_{1,0,0}(\vec{r}_2) \Psi_{1,0,0}(\vec{r}_1)) \\ &= \sqrt{2} \Psi_{1,0,0}(\vec{r}_1) \Psi_{1,0,0}(\vec{r}_2). \end{aligned} \quad (13.3.12)$$

13.4. PERTURBATION THEORY FOR GROUND AND EXCITED STATES OF THE HELIUM ATOM

This wave function seems to be incorrectly normalized to 2. However, the correct normalization condition for two identical electrons takes the form

$$\frac{1}{2} \int d^3r_1 d^3r_2 |\Psi(\vec{r}_1, \vec{r}_2)|^2 = 1. \quad (13.3.13)$$

The prefactor $1/2$ arises due to the indistinguishability of the electrons. The value of the ground state energy is given by $E = 2E_1$, which is a factor $2Z^2 = 8$ larger than the ground state energy of the hydrogen atom. Of course, we should keep in mind that we have not yet included the electron repulsion term V_{12} . The first excited states still have $n_1 = 1$, $l_1 = m_1 = 0$ (a 1s state), but now $n_2 = 2$ and hence the total energy is $E = E_1 + E_2$. The possible values for the angular momentum are $l_2 = m_2 = 0$ (a 2s state) and $l_2 = 1$, $m_2 = 0, \pm 1$ (one of three 2p states). In this case, both para- and ortho-helium states exist with the orbital wave functions

$$\Psi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} (\Psi_{1,0,0}(\vec{r}_1) \Psi_{2,l,m}(\vec{r}_2) \pm \Psi_{1,0,0}(\vec{r}_2) \Psi_{2,l,m}(\vec{r}_1)). \quad (13.3.14)$$

Altogether there are 16 degenerate excited states: the four spin states combined with the four possible choices for l_2, m_2 .

13.4 Perturbation Theory for Ground and Excited States of the Helium Atom

In the next step we will include the electron-electron Coulomb repulsion V_{12} using perturbation theory. First, we consider the correction ΔE to the ground state energy $E = 2E_1$, which is given by

$$\Delta E = \langle \Psi | V_{12} | \Psi \rangle = \int d^3r_1 d^3r_2 \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} |\Psi_{1,0,0}(\vec{r}_1)|^2 |\Psi_{1,0,0}(\vec{r}_2)|^2. \quad (13.4.1)$$

From our study of the hydrogen atom, we know the ground state wave function

$$\Psi_{1,0,0}(\vec{r}) = \frac{1}{\sqrt{\pi a^3}} \exp(-|\vec{r}|/a), \quad (13.4.2)$$

where $a = \hbar^2 / (Ze^2 M)$ is the Bohr radius. In the next step we write

$$|\vec{r}_1 - \vec{r}_2| = \sqrt{|\vec{r}_1|^2 + |\vec{r}_2|^2 - 2\vec{r}_1 \cdot \vec{r}_2} = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}. \quad (13.4.3)$$

Inserting this and performing all angular integrations except the one over the angle θ between the vectors \vec{r}_1 and \vec{r}_2 one obtains

$$\Delta E = \frac{8\pi^2 e^2}{\pi^2 a^6} \int dr_1 dr_2 d\theta r_1^2 r_2^2 \sin \theta \frac{\exp(-2r_1/a) \exp(-2r_2/a)}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}}. \quad (13.4.4)$$

Substituting $x = \cos \theta$, $dx = -\sin \theta d\theta$ one finds

$$\begin{aligned} \int_0^\pi d\theta \sin \theta \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta}} &= \int_{-1}^1 dx \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 x}} = \\ &= -\frac{1}{r_1 r_2} (\sqrt{r_1^2 + r_2^2 - 2r_1 r_2} - \sqrt{r_1^2 + r_2^2 + 2r_1 r_2}) = \\ \frac{r_1 + r_2 - |r_1 - r_2|}{r_1 r_2} &= \begin{cases} 2/r_1 & \text{for } r_1 > r_2 \\ 2/r_2 & \text{for } r_2 > r_1. \end{cases} \end{aligned} \quad (13.4.5)$$

Hence, we now obtain

$$\begin{aligned} \Delta E &= \frac{16e^2}{a^6} \int_0^\infty dr_1 r_1 \exp(-2r_1/a) \\ &\times \left[\int_0^{r_1} dr_2 r_2^2 \exp(-2r_2/a) + r_1 \int_{r_1}^\infty dr_2 r_2 \exp(-2r_2/a) \right] \\ &= \frac{5e^2}{8a} = \frac{5Ze^4 M}{8\hbar^2}. \end{aligned} \quad (13.4.6)$$

Chapter 14

Quantum Mechanics of Elastic Scattering

In order to investigate the structure of molecules, atoms, atomic nuclei, or individual protons, one must interact with these systems from outside. To study these fundamental objects, one can probe them with other elementary particles such as electrons, photons, or neutrinos. For this purpose, one performs scattering experiments. For example, by bombarding atoms with electrons Rutherford and his collaborators were able to demonstrate that the atomic nucleus is tiny compared to the entire atom. Similarly, electron scattering experiments have revealed that protons are not truly elementary but consist of quarks and gluons. In Rutherford's experiments the atomic nucleus remained intact and in its ground state, i.e. the electrons were scattered elastically. The experiments that revealed the quark structure of the proton, on the other hand, were deeply inelastic. In such a scattering event the proton is broken up into numerous fragments. In this chapter we consider the quantum mechanics of elastic scattering.

14.1 Differential Cross Section and Scattering Amplitude

Let us consider projectile particles (for example, electrons) moving along the z -direction with constant momentum $\vec{k} = k\vec{e}_z$. They hit a heavy target (for example, an atom) that is at rest. The target exerts a potential $V(\vec{x})$ (in our example the Coulomb potential) on the projectile. Here we assume that the

target is static and unaffected by the projectile (elastic scattering), and can be described entirely in terms of the scattering potential $V(\vec{x})$. In reality the target, of course, also suffers some recoil which we neglect here. It would be easy to incorporate this by describing the collision in the center of mass frame.

The incoming projectile particles are described by an incident probability current density j which measures the number of incident projectile particles per beam area and time. The projectile particles are then affected by the scattering potential and in general change their direction of motion until they leave the interaction region and disappear to infinity. The scattering angles (θ, φ) describe the direction of the momentum of the projectile after the collision. When we place a detector covering an angle $d\Omega$ in that direction, we measure the current dJ which counts the number of projectile particles per unit time. The differential cross section is defined as

$$\frac{d\sigma(\theta, \varphi)}{d\Omega} = \frac{dJ}{d\Omega} \frac{1}{Nj}. \quad (14.1.1)$$

Here N is the number of scattering centers (in this case atoms) in the target. The differential cross section thus counts the number of projectiles scattered into the direction (θ, φ) , normalized to the incident current and the number of scattering centers. Here we assume a dilute target. Then each projectile particle scatters only on a single target particle and the current dJ is proportional to N . For multiple scattering events this would not be the case. The theory allows us to derive the differential cross section $d\sigma(\theta, \varphi)/d\Omega$ from the scattering potential $V(\vec{x})$. Experiments, on the other hand, directly measure $d\sigma(\theta, \varphi)/d\Omega$ and thus indirectly yield information on $V(\vec{x})$ and thus on the structure of the target. One can also define the total cross section

$$\sigma = \int d\Omega \frac{d\sigma(\theta, \varphi)}{d\Omega}. \quad (14.1.2)$$

We limit the discussion to scattering potentials that vanish at infinity. The energy of the incident projectile particles is then given by

$$E = \frac{\hbar^2 k^2}{2m}. \quad (14.1.3)$$

We are looking for a stationary scattering wave function $|\Psi(\vec{k})\rangle$ that solves the time-independent Schrödinger equation

$$H|\Psi(\vec{k})\rangle = E|\Psi(\vec{k})\rangle. \quad (14.1.4)$$

At asymptotic distances the wave function can be decomposed into the incident plane wave $\exp(ikz)$ and a scattered wave, i.e.

$$\langle \vec{x} | \Psi(\vec{k}) \rangle \sim A [\exp(ikz) + f(\theta, \varphi) \frac{\exp(ikr)}{r}]. \quad (14.1.5)$$

Here $r = |\vec{x}|$ is the distance from the target. The factor $\exp(ikr)/r$ describes a radial wave with an angular-dependent amplitude $f(\theta, \varphi)$ — the so-called scattering amplitude.

Let us now derive a relation between the scattering amplitude $f(\theta, \varphi)$ and the differential cross section $d\sigma(\theta, \varphi)/d\Omega$. The total probability current density is given by

$$\vec{j}(\vec{x}) = \frac{\hbar}{2mi} [\langle \Psi(\vec{k}) | \vec{x} \rangle \vec{\nabla} \langle \vec{x} | \Psi(\vec{k}) \rangle - (\vec{\nabla} \langle \Psi(\vec{k}) | \vec{x} \rangle) \vec{\nabla} \langle \vec{x} | \Psi(\vec{k}) \rangle]. \quad (14.1.6)$$

For the incident plane wave $A \exp(ikz)$ one obtains

$$\vec{j}_{in}(\vec{x}) = \frac{\hbar k}{m} |A|^2 \vec{e}_z. \quad (14.1.7)$$

The probability current density of the scattered wave, on the other hand, is given by

$$\begin{aligned} \vec{j}_{out}(\vec{x}) &= \frac{\hbar}{2mi} |A|^2 [f(\theta, \varphi)^* \frac{\exp(-ikr)}{r} (\vec{\nabla} f(\theta, \varphi) \frac{\exp(ikr)}{r}) \\ &\quad - (\vec{\nabla} f(\theta, \varphi)^* \frac{\exp(-ikr)}{r}) f(\theta, \varphi) \frac{\exp(ikr)}{r}] \\ &= \frac{\hbar k}{m} |A|^2 |f(\theta, \varphi)|^2 \frac{1}{r^2} + \mathcal{O}(\frac{1}{r^3}). \end{aligned} \quad (14.1.8)$$

Hence, we can identify the differential cross section as

$$\frac{d\sigma(\theta, \varphi)}{d\Omega} = \frac{dJ}{d\Omega} \frac{1}{jN} = \frac{|\vec{j}_{out}| r^2}{|\vec{j}_{in}|} = |f(\theta, \varphi)|^2. \quad (14.1.9)$$

This shows that the asymptotic form of the wave function determines the differential cross section.

14.2 Green Function for the Schrödinger Equation

Let us consider the Schrödinger equation

$$[-\frac{\hbar^2}{2m} \Delta + V(\vec{x})] \langle \vec{x} | \Psi(\vec{k}) \rangle = E \langle \vec{x} | \Psi(\vec{k}) \rangle \quad (14.2.1)$$

for scattering states of positive energy $E = \hbar^2 k^2 / 2m$. First, we replace the potential by a δ -function

$$V(\vec{x}) = \frac{\hbar^2}{8m\pi} \delta(\vec{x}), \quad (14.2.2)$$

and we consider the following equation for the Green function $G_0(\vec{x})$

$$[\Delta + k^2]G_0(\vec{x}) = -4\pi\delta(\vec{x}). \quad (14.2.3)$$

In momentum space this equation takes the form

$$\tilde{G}_0(\vec{k}') = \frac{4\pi}{k'^2 - k^2}. \quad (14.2.4)$$

Here $\tilde{G}_0(\vec{k}')$ is the Fourier transform of $G_0(\vec{x})$, which is obtained by an inverse Fourier transform

$$\begin{aligned} G_0(\vec{x}) &= \frac{1}{(2\pi)^3} \int d^3k' \tilde{G}(\vec{k}') \exp(i\vec{k}' \cdot \vec{x}) \\ &= \frac{1}{\pi} \int_0^\infty dk' \frac{k'^2}{k'^2 - k^2} \int_{-1}^1 d\cos\theta \exp(ik'|\vec{x}|\cos\theta) \\ &= \frac{1}{\pi i|\vec{x}|} \int_{-\infty}^\infty dk' \frac{k' \exp(ik'|\vec{x}|)}{(k' - k)(k' + k)} = \frac{\exp(ik|\vec{x}|)}{|\vec{x}|}. \end{aligned} \quad (14.2.5)$$

In the last step we have closed the integration contour in the complex plane and we have used the residue theorem.

14.3 The Lippmann-Schwinger Equation

The Schrödinger equation is equivalent to the Lippmann-Schwinger equation

$$\langle \vec{x} | \Psi(\vec{k}) \rangle = \langle \vec{x} | \vec{k} \rangle - \frac{1}{4\pi} \frac{2m}{\hbar^2} \int d^3x' \frac{\exp(ik|\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|} V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle. \quad (14.3.1)$$

Here $\langle \vec{x} | \vec{k} \rangle = A \exp(i\vec{k} \cdot \vec{x})$ is the incident plane wave. Acting with the Laplacian $[\Delta + k^2]$ and using the defining property of the Green function we indeed obtain

$$\begin{aligned} \frac{\hbar^2}{2m} [\Delta + k^2] \langle \vec{x} | \Psi(\vec{k}) \rangle &= \\ -\frac{1}{4\pi} \int d^3x' [\Delta + k^2] \frac{\exp(ik|\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|} V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle &= \\ \int d^3x' \delta(\vec{x} - \vec{x}') V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle &= V(\vec{x}) \langle \vec{x} | \Psi(\vec{k}) \rangle, \end{aligned} \quad (14.3.2)$$

which is nothing but the Schrödinger equation.

Let us consider the scattering wave function $\langle \vec{x} | \Psi(\vec{k}) \rangle$ at asymptotic distances \vec{x} . The wave vector pointing in the corresponding direction is then given by $\vec{k}' = k\vec{x}/|\vec{x}|$ and we can write

$$\begin{aligned} k|\vec{x} - \vec{x}'| &= k\sqrt{x^2 + x'^2 - 2\vec{x} \cdot \vec{x}'} \approx k|\vec{x}|\sqrt{1 - 2\vec{x} \cdot \vec{x}'/|\vec{x}|^2} \\ &\approx k|\vec{x}| - k\vec{x} \cdot \vec{x}'/|\vec{x}| = k|\vec{x}| - \vec{k}' \cdot \vec{x}'. \end{aligned} \quad (14.3.3)$$

At asymptotic distances the Lippmann-Schwinger equation thus takes the form

$$\langle \vec{x} | \Psi(\vec{k}) \rangle \approx \langle \vec{x} | \vec{k} \rangle - \frac{1}{4\pi} \frac{2m}{\hbar^2} \frac{\exp(ik|\vec{x}|)}{|\vec{x}|} \int d^3x' \exp(i\vec{k}' \cdot \vec{x}') V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle. \quad (14.3.4)$$

Hence, we can identify the scattering amplitude as

$$f(\vec{k}') = -2\pi^2 \frac{2m}{\hbar^2} \int d^3x' \exp(i\vec{k}' \cdot \vec{x}') V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle = -2\pi^2 \frac{2m}{\hbar^2} \langle \vec{k}' | V | \Psi(\vec{k}) \rangle. \quad (14.3.5)$$

14.4 Abstract Form of the Lippmann-Schwinger Equation

Let us now define a free propagator as

$$G_0(E) = \frac{1}{E - p^2/2m}, \quad (14.4.1)$$

where p is the momentum operator. The corresponding matrix element

$$\begin{aligned} \langle \vec{x} | G_0(E) | \vec{x}' \rangle &= \langle \vec{x} | \frac{1}{E - p^2/2m} | \vec{x}' \rangle = \int d^3k' \langle \vec{x} | \vec{k}' \rangle \frac{2m}{\hbar^2} \frac{1}{k^2 - k'^2} \langle \vec{k}' | \vec{x}' \rangle \\ &= \frac{1}{(2\pi)^3} \frac{2m}{\hbar^2} \int d^3k' \frac{\exp(i\vec{k}' \cdot (\vec{x} - \vec{x}'))}{k^2 - k'^2} = G_0(\vec{x} - \vec{x}'), \end{aligned} \quad (14.4.2)$$

is just the Green function defined above. This allows us to write the Lippmann-Schwinger equation in the more abstract form

$$|\Psi(\vec{k})\rangle = |\vec{k}\rangle + G_0(E)V|\Psi(\vec{k})\rangle. \quad (14.4.3)$$

Indeed, by left-multiplication with $\langle \vec{x} |$ one obtains

$$\begin{aligned} \langle \vec{x} | \Psi(\vec{k}) \rangle &= \langle \vec{x} | \vec{k} \rangle + \langle \vec{x} | G_0(E) V | \Psi(\vec{k}) \rangle \\ &= \langle \vec{x} | \vec{k} \rangle + \int d^3x' \langle \vec{x} | G_0(E) | \vec{x}' \rangle \langle \vec{x}' | V | \Psi(\vec{k}) \rangle \\ &= \langle \vec{x} | \vec{k} \rangle + \int d^3x' G_0(\vec{x} - \vec{x}') V(\vec{x}') \langle \vec{x}' | \Psi(\vec{k}) \rangle. \end{aligned} \quad (14.4.4)$$

In the next step we define a full interacting propagator

$$G(E) = \frac{1}{E - H} = \frac{1}{E - p^2/2m - V}, \quad (14.4.5)$$

and we write

$$\begin{aligned} G(E) - G_0(E) &= \frac{1}{E - p^2/2m - V} - \frac{1}{E - p^2/2m} \\ &= \frac{1}{E - p^2/2m - V} (E - p^2/2m) \frac{1}{E - p^2/2m} \\ &\quad - \frac{1}{E - p^2/2m - V} (E - p^2/2m - V) \frac{1}{E - p^2/2m} \\ &= \frac{1}{E - p^2/2m - V} V \frac{1}{E - p^2/2m} \\ &= G(E) V G_0(E). \end{aligned} \quad (14.4.6)$$

Then the abstract form of the Lippmann-Schwinger equation takes the form

$$\begin{aligned} |\Psi(\vec{k})\rangle &= |\vec{k}\rangle + G_0(E) V |\Psi(\vec{k})\rangle \\ &= |\vec{k}\rangle + G(E) V |\Psi(\vec{k})\rangle - G(E) V G_0(E) V |\Psi(\vec{k})\rangle \\ &= |\vec{k}\rangle + G(E) V |\Psi(\vec{k})\rangle - G(E) V (|\Psi(\vec{k})\rangle - |\vec{k}\rangle) \\ &= |\vec{k}\rangle + G(E) V |\vec{k}\rangle. \end{aligned} \quad (14.4.7)$$

Indeed, multiplying this equation with $E - H$, we obtain

$$(E - H) \Psi(\vec{k}) = (E - H) |\vec{k}\rangle + V |\vec{k}\rangle = (E - p^2/2m) |\vec{k}\rangle = 0, \quad (14.4.8)$$

which is nothing but the Schrödinger equation.

Iterating the abstract form of the Lippmann-Schwinger equation one obtains

$$\begin{aligned} |\Psi(\vec{k})\rangle &= |\vec{k}\rangle + G_0(E) V |\Psi(\vec{k})\rangle \\ &= |\vec{k}\rangle + G_0(E) V |\vec{k}\rangle + G_0(E) V G_0(E) V |\Psi(\vec{k})\rangle \\ &= \sum_{n=0}^{\infty} (G_0(E) V)^n |\vec{k}\rangle. \end{aligned} \quad (14.4.9)$$