Modern Quantum Mechanics

Lecture notes – FYS 4110

Jon Magne Leinaas
Department of Physics, University of Oslo

Preface

The course FYS 4110 gives an introduction to modern aspects of quantum physics, with subjects such as coherent quantum states, density operators and entanglement, elements from quantum information theory and the physics of photons and atoms. The focus is mainly on the elementary aspects of these topics, and the familiar quantum two-level system and the harmonic oscillator are used repeatedly and in different ways to illustrate the various aspects of quantum physics that are discussed.

The course extends the knowledge of quantum physics from the bachelor program, and the basics of quantum physics, in particular Dirac's bra-ket formulation, is thus supposed to be known from previous physics courses.

Jon Magne Leinaas
Department of Physics, University of Oslo,

Corrected, July 2019

Contents

Chapter 1

Quantum formalism

1.1 Summary of quantum states and observables

In this section we make a summary of the fundamental assumptions and postulates of quantum theory. We stress the correspondence with classical theory, but at the same time focus on the radically different way the quantum theory is interpreted. We summarize how an isolated quantum system is described in terms of abstract vectors and operators in a Hilbert space.

1.1.1 Classical and quantum states

The description of a classical system that is most closely related to the standard description of a quantum system is the *phase space description*. In this description the variables are the generalized coordinates $q = \{q_i \, ; i = 1, 2, ..., N\}$, each corresponding to a degree of freedom of the system, and the corresponding canonical momenta $p = \{p_i \, ; i = 1, 2, ..., N\}$. A complete specification of the state of the system is given by the full set of coordinates and momenta (q, p), which identifies a point in phase space.

There is a unique time evolution of the phase space coordinates $(q(t), p(t))$, with a given initial condition $(q_0, p_0) = (q(t_0), p(t_0))$ at time t_0. This is so, since the equations of motion, expressed in terms of the phase space coordinates are *first order* in time derivatives. For a Hamiltonian system the dynamics can be expressed in terms of the classical Hamiltonian, which is a function of the phase space variables, $H = H(q, p, t)$, and is normally identical to the energy function. The time evolution is expressed by Hamilton's equations as

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \ , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , i = 1, 2, ..., N \, , \tag{1.1}$$

where $\dot{q}_i$, in the usual way, means the time derivative of the coordinate q_i.

If a complete specification of the system cannot be given, a statistical description may often be used. The state of the system is then described in terms of a probability function $\rho(q, p, t)$ defined on the phase space. In statistical mechanics this function is a basic element of the description, and the time evolution is described through the time derivative of ρ,

$$\frac{d}{dt}\rho \ = \ \sum_i (\frac{\partial \rho}{\partial q_i}\dot{q}_i + \frac{\partial \rho}{\partial p_i}\dot{p}_i) + \frac{\partial}{\partial t}\rho$$

$$= \ \{\rho, H\}_{PB} + \frac{\partial}{\partial t}\rho \,, \tag{1.2}$$

where in this equation the *Poisson bracket*, defined by

$$\{A, B\}_{PB} \equiv \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right) \,, \tag{1.3}$$

has been introduced. One should note that in Eq.(1.2) $\frac{\partial}{\partial t}$ means the time derivative with fixed phase space coordinates, whereas $\frac{d}{dt}$ includes the time variation due to the motion in phase space.

The time evolution of ρ (and any other phase space variable), when written as in Eq. (1.2), shows a striking similarity with the *Heisenberg equation of motion* of the quantum system. The commutator between the variables then takes the place of the Poisson bracket. The quantum description of the system, like the classical description, involves the phase space variables (q, p). But these dynamical variables are in the quantum theory re-interpreted as operators that act on complex-valued wave functions $\psi(q, t)$ of the system. To specify the variables as operators they are often written $\hat{q}_i$ and $\hat{p}_i$, and we shall also in most places use this notation. In the standard way we refer to these as *observables*, and the fundamental relation between these observables is the (Heisenberg) commutation relation

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij} \,, \tag{1.4}$$

with $\hbar$ as (the "reduced") Planck's constant. A more general observable $\hat{A}$ may be viewed as a function of $\hat{q}_i$ and $\hat{p}_i$, and two observables $\hat{A}$ and $\hat{B}$ will in general not commute. We usually restrict observables to be Hermitian operators, which correspond to real-valued variables in the classical description.

The close relation between the classical and quantum description of a mechanical system is most clearly seen when the two descriptions are expressed in terms of the same phase space variables. In fact there exists a simple scheme for *quantizing* the classical system, referred to as *canonical quantization*, which defines a formal transition from the classical to the quantum description of the same physical system. In its simplest form this transition is viewed as a change from classical phase space variables to quantum observables

$$q_i \to \hat{q}_i, \ p_i \to \hat{p}_i \,, \tag{1.5}$$

where the quantum variables are assumed to satisfy the fundamental commutation relation (1.4). For general variables the transition can be expressed in the form of a substitution between Poisson brackets for the classical variables and commutators for the quantum variables

$$\{A, B\}_{PB} \to \frac{1}{i\hbar} \left[\hat{A}, \hat{B} \right] \,. \tag{1.6}$$

Clearly this simple substitution rule gives the right commutator between $\hat{q}_i$ and $\hat{p}_j$ when used on the Poisson brackets between q_i and p_j.[1]

---

[1]In general there will, however be an ambiguity in this substitution in the form of the so called *operator ordering problem*. Since classical observables commute, a composite variable $C = AB = BA$ can be written in several ways. The corresponding quantum observables may be different due to non-commutativity, $\hat{C} = \hat{A}\hat{B} \neq \hat{B}\hat{A} = \hat{C}'$. The *Weyl ordering* is one way to solve the ambiguity by replacing a product by its symmetrized version, $\hat{C} = \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A})$. However, a natural interpretation of the ambiguity is that the quantum description of a system is not fully determined by the classical description, without some additional specifications.

For the dynamical equations the quantization rules (1.5) and (1.6) lead from the classical Hamilton's equations to the Heisenberg's equation of motion for the quantum system. This correspondence between the classical and quantum dynamical equations is directly related to *Ehrenfest's theorem*, which states that the classical dynamical equations keep their validity also in the quantum theory, with the classical variables replaced by their corresponding quantum expectation values. Thus, the quantum expectation value $\langle q \rangle$ in many respects behaves like a classical variable q, and the time evolution of the expectation value follows a classical equation of motion. As long as the wave function is well localized (in the q-variable), the system behaves "almost classically". However when the wave functions spread out or divide into separated parts, then highly "non-classical effects" may arise.

The correspondence between the classical and quantum description of physical systems was in the early days of quantum theory used actively by Bohr and others in the form of the *correspondence principle*. Thus, even before quantum mechanics was fully developed the classical theory gave information about the quantum theory in the form of the "classical limit" of the theory, the limit where the effect of Planck's constant becomes negligible. In particular, for radiative transitions between atomic levels the correspondence principle would imply that the radiation formula of the quantum theory should reproduce the classical one for highly excited atoms, in the limit where the excitation energy approaches the ionization value. But at the formal level the correspondence between the classical and the fully developed quantum theory goes much further than simply to the requirement that the classical description should be recovered in the limit $\hbar \to 0$.

The close correspondence between the classical and quantum theory is in many respects rather surprising, since the physical interpretation of the two theories are radically different. The difference is linked to the statistical interpretation of the quantum theory, which is the subject of one of the later sections. Both classical and quantum descriptions of a system will often be of statistical nature, since the full information (especially for systems with a large number of degrees of freedom) may not be achievable. Often interactions with other systems (the surroundings) disturb the system in such a way that only a statistical description is meaningful. If such disturbances are negligible the system is referred to as an *isolated* or *closed system* and for an isolated classical system all the dynamical variables can in principle be ascribed sharp values.

For a quantum system that is not the case. The quantum state of an isolated system is described by the wave function $\psi(q)$, defined over the (classical) configuration space of the system. This is interpreted as a *probability amplitude*, which means that the absolute value $|\psi(q)|^2$ defines a probability distribution in configuration space. For a general observable $\hat{A}$ this leads to an uncertainty with respect to the measured value, usually expressed by the statistical *variance*

$$\Delta A^2 = \left\langle (\hat{A} - \left\langle \hat{A} \right\rangle)^2 \right\rangle . \tag{1.7}$$

Even if such a probability distribution, in principle, can be sharp in the set of variables q, it cannot at the same time be sharp in the conjugate variables p due to the fundamental commutation relation (1.4). This is quantified in *Heisenberg's uncertainty relation*

$$\Delta q_i \Delta p_j \geq \frac{\hbar}{2} \delta_{ij} . \tag{1.8}$$

The inherent probabilistic interpretation of the quantum theory in many respects makes it more closely related to a statistical description of the classical system than to a detailed non-statistical description. However, the standard description in terms of a wave function $\psi(q)$ defined on the configuration space seems rather different from a classical statistical description in terms of a phase space probability distribution $\rho(q, p)$. Quantum descriptions in terms of functions similar to $\rho(q, p)$ are possible and in some cases they are also useful. They are referred to as quasi-probability distributions, since they do not always satisfy the positivity condition of probabilities.[2] However, even if such a formulation brings the description closer to the classical, statistical description of the system, there is one important property of the wave functions that is hidden in such a reformulation. The *superposition principle* is a fundamental principle of quantum mechanics which implies that the theory is linear in the probability *amplitudes* $\psi(q)$. The quasi-probability distributions, derived from the standard quantum description, are quadratic in $\psi(q)$, and therefore this linearity is lost. In the classical statistical theory there is no counterpart to the quantum superposition principle.[3]

The description of quantum systems in terms of wave functions, defined as functions over the classical configuration space, is only one of many equivalent "representations" of the quantum theory. A more abstract formulation exists where the states are (abstract) vectors in a Hilbert space, and where different representations of the theory correspond to different choices of sets of basis vectors in this space. In the early days of quantum mechanics this representation theory of quantum physics was formulated and studied in a particular clear form by P.A.M. Dirac. His "bra-ket" formulation is still standard in quantum mechanics and will be applied also here, with a general (abstract) state vector denoted by $|\psi\rangle$ and the scalar product between two states as $\langle\phi|\psi\rangle$.

In the following a summary of this abstract (and formal) description will be given, in terms of what is often called the fundamental postulates of quantum theory.

1.1.2   The fundamental postulates

1. *A quantum state of an isolated physical system is described by a vector with unit norm in a Hilbert space. This is a complex vector space equipped with a scalar product.*[4]

---

[2]The Wigner function is a particular example of a quasi-probability distribution. For a particle in one dimension it is defined as $W(x, p) = (1/2\pi\hbar) \int dy\,\psi(x + y/2)^* \,\psi(x - y/2) \exp(iyp/\hbar)$, with $\psi(x)$ as the wave function of the particle. It shares the property with the classical probability distribution in phase space, that integrated over p it gives the probability distribution over x, $\int dp\, W(x, p) = |\psi(x)|^2$. Similarly, when integrated over x, it gives the probability distribution over p. However, $W(x, p)$ is not a true probability distribution, since it may become negative. These regions with negative W in phase space are often interpreted as signatures for the presence of quantum effects.

[3]Note, however that the *probabilities* of the classical theory also describe a linear system, but this linearity is different from that of the quantum theory, which is linear in the probability *amplitudes*. In a later section we shall discuss an extension of quantum theory from description in terms of wave functions to a description in terms of *density operators*. These operators are closely related to the quasi-probability distributions mentioned above. The extended theory is linear in these new operators, but the original superposition principle of quantum wave functions is no longer explicit in this extended formulation.

[4]A Hilbert space is more specifically a vector space with a scalar product (an inner product space) which is *complete* in the norm. This means that any (Cauchy) sequence of vectors $|n\rangle, n = 1, 2, ...$, where the norm of the relative vectors $|n, m\rangle = |n\rangle - |m\rangle$ goes to zero as $n, m \to \infty$, will have a limit (vector) belonging to the space. Usually the Hilbert space is assumed to be *separable*, which means that it is spanned by a *countable* orthonormal basis. These specifications are of importance when the vector space is infinite dimensional, and they imply that

In the Dirac notation a vector is represented by a "ket" $|\psi\rangle$, which can be expanded in any complete set of basis vectors $|i\rangle$,

$$|\psi\rangle = \sum_i c_i |i\rangle \,, \tag{1.9}$$

where the coefficients c_i are complex numbers. For an infinite dimensional Hilbert space the basis may be a discrete or continuous set of vectors. We refer to the vectors (kets) as *state vectors* and the vector space as the *state space*.

A "bra" $\langle\psi|$ is regarded as vector in the *dual vector space*, and is related to $|\psi\rangle$ by an *anti-linear* mapping (linear mapping + complex conjugation)

$$|\psi\rangle \rightarrow \langle\psi| = \sum_i c_i^* \langle i| \,. \tag{1.10}$$

The scalar product is a complex-valued composition of a bra and a ket, $\langle\phi|\psi\rangle$ which is a linear function of $|\psi\rangle$ and an antilinear function of $|\phi\rangle$. The quantum states are associated with the *normalized* vectors, so that $\langle\psi|\psi\rangle = 1$.

2. *Each physical observable of a system is associated with a hermitian operator acting on the Hilbert space. The eigenstates of each such operator define a complete, orthonormal set of vectors.*

With $\hat{A}$ as an observable, hermiticity means

$$\langle\phi|\hat{A}\psi\rangle = \langle\hat{A}\phi|\psi\rangle \equiv \langle\phi|\hat{A}|\psi\rangle \,. \tag{1.11}$$

If the observable has a discrete spectrum, the eigenstates are orthogonal and may be normalized as

$$\langle i|j\rangle = \delta_{ij} \,. \tag{1.12}$$

Completeness means

$$\sum_i |i\rangle\langle i| = \hat{\mathbb{1}} \,, \tag{1.13}$$

where $\hat{\mathbb{1}}$ is the unit operator. In general a hermitian operator will have partly a discrete and partly a continuous spectrum. For the continuous spectrum orthogonality is expressed in terms of Dirac's delta function.[5]

3. *The time evolution of the state vector, $|\psi\rangle = |\psi(t)\rangle$, is (in the Schrödinger picture) defined by the Schrödinger equation, of the form*

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \,. \tag{1.14}$$

---

many of the properties of finite dimensional vector spaces can be taken over almost directly.

[5]For an observable with a discrete spectrum the eigenstates are normalizable and belong to the Hilbert space. For a continuous spectrum the eigenstates are non-normalizable and therefore fall outside the Hilbert space. However, they can be included in an extension of the Hilbert space. Completeness holds within this extended space, but orthonormality of the vectors has to be expressed in terms of Dirac's delta function rather than the Kronecker delta.

The equation is first order in the time derivative, which means that the time evolution $|\psi\rangle = |\psi(t)\rangle$ is uniquely determined by the initial condition $|\psi_0\rangle = |\psi(t_0)\rangle$. $\hat{H}$ is the Hamiltonian of the system which is a *linear, hermitian* operator. It gives rise to a time evolution which is a *unitary*, time dependent mapping between quantum states.

4. *The measurable (physical) values associated with an observable $\hat{A}$ are defined by its eigenvalues a_n. With the physical system in the state $|\psi\rangle$ before a measurement of the observable, the probability p_n for finding a particular eigenvalue a_n in the measurement is*

$$p_n = |\langle n|\psi\rangle|^2 \,, \tag{1.15}$$

with $|n\rangle$ as the eigenvector corresponding to the eigenvalue a_n.

If the observable has a degeneracy, so that several (orthogonal) eigenvectors have the same eigenvalue, the probability is given as a sum over all eigenvectors with the same eigenvalue a_n.

The expectation value of an observable A in the state $|\psi\rangle$ is

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle \,. \tag{1.16}$$

It corresponds to the mean value obtained in an (infinite) series of identical measurements of the variable A, where the system before each measurement is prepared in the same state $|\psi\rangle$.

5. *An <u>ideal</u> measurement of observable A resulting in a value a_n projects the state vector from initial state $|\psi>$ to the final state*

$$|\psi> \rightarrow |\psi'> = P_n|\psi> \,, \tag{1.17}$$

where P_n is the projection on the eigenstate $|a_n\rangle$, or more generally on the subspace spanned by the vectors with eigenvalue a_n.[6]

Note that since the projected state in general will not be normalized to unity, the state should also be multiplied by a normalization factor in order to satisfy the standard normalization condition for physical states.

The effect of the measurement, that it projects the original state into the eigenstate which corresponds to the measured eigenvalue, is in a sense is a minimal disturbance of the system caused by the quantum measurement. The projection is often referred to as the "collapse of the wave function", and it corresponds to the "collapse" of a probability function of a classical system when additional information is introduced in the description without disturbing the system in any other way. But one should be aware of the far-reaching difference of this "collapse by adding new information" in the classical and quantum descriptions. In the classical case the ideal measurement corresponds to collecting new information without disturbing the system. In the quantum case the Heisenberg uncertainty principle implies that reducing the uncertainty of one observable by a measurement means increasing the uncertainty for other observables. In this sense an ideal measurement cannot be regarded as having no real influence on the quantum system.[7]

---

[6]Such idealized measurements are often referred to as *projective* measurements.

[7]There is an obvious question why elements of measurement theory are included in the *fundamental postulates*

1.1.3   Matrix representations and wave functions

The state vectors $|\psi\rangle$ and observables $\hat{A}$, as they appear in Dirac's bra-ket formalism, we often refer to as *abstract* vectors and operators, as opposed to concrete representations of these in the form of matrices, or wave functions and differential operators. A matrix representation is defined by the expansion coefficients of the vectors and observables in a *discrete* basis that spans the Hilbert space of the system. Usually this is a complete, orthogonal and normalized basis, often composed by the eigenstates of a set of commuting observables. With the expansion written as

$$|\psi\rangle = \sum_i \psi_i |i\rangle , \quad \psi_i = \langle i|\psi\rangle , \tag{1.18}$$

the matrix representation of a state vector is

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ . \\ . \end{pmatrix} . \tag{1.19}$$

The corresponding expansion of an observable is

$$\hat{A} = \sum_{i,j} A_{ij} |i\rangle\langle j| , \quad A_{ij} = \langle i|\hat{A}|j\rangle , \tag{1.20}$$

with the matrix representation

$$\mathbb{A} = \begin{pmatrix} A_{11} & A_{12} & . & . \\ A_{21} & A_{22} & . & . \\ . & . & . & . \end{pmatrix} . \tag{1.21}$$

In the matrix representations the actions of the observables as well as the scalar products between state vectors are reduced to matrix multiplications.

For an infinite dimensional Hilbertspace, the corresponding matrix dimensions will also be infinite. However, often truncation of the matrices to finite dimensional form can be done without loosing essential (relevant) information about the system.

If the state vectors and observables are expanded in a continuous rather than a discrete basis, this leads to a description of the quantum system in terms of wave functions and differential operators. We briefly discuss how this works. Let us then consider a *coordinate basis* defined by the continuous set of eigenvectors of the coordinate observables $\hat{q}_i$

$$\hat{q}_i |q\rangle = q_i |q\rangle , \tag{1.22}$$

---

of quantum theory. In classical theory that is usually not done, since the classical variables can in most cases be viewed as (in principle) measurable. Quantum theory is different since the basic elements (state vectors and observables) cannot be viewed (even in principle) as directly measurable. The postulates about (ideal) measurements are meant to express the fundamental *probabilistic* interpretation of quantum theory rather than describing realistic measurements. In a broader approach to quantum measurement theory other types of measurements than the idealized (projective) measurements will usually be introduced. But this does not imply any essential change in the (probabilistic) interpretation of the theory expressed by the above postulates.

where q denotes the set of coordinates $\{q_i\}$. For a Cartesian set of coordinates the standard normalization is

$$\langle q'|q\rangle = \delta(q' - q)\,, \tag{1.23}$$

where $\delta(q' - q)$ is the N-dimensional Dirac delta-function, with N as the dimension of the configuration space. The wave functions defined over the configuration space of the system are the components of the abstract state vector $|\psi\rangle$ on this basis,

$$\psi(q) = \langle q|\psi\rangle\,. \tag{1.24}$$

A general observable is in the coordinate representation specified by its matrix elements,

$$A(q',q) \equiv \langle q'|\hat{A}|q\rangle\,. \tag{1.25}$$

It acts on the wave function as an integral operator

$$\langle q|\hat{A}|\psi\rangle = \int dq'\, A(q,q')\,\psi(q')\,, \tag{1.26}$$

where dq' represents the N-dimensional volume element.

A potential function is an example of a *local* observable,

$$V(q',q) = V(q)\,\delta(q' - q)\,, \tag{1.27}$$

in which case the integral collapses to a simple multiplication

$$\langle q|\hat{V}|\psi\rangle = V(q)\,\psi(q)\,. \tag{1.28}$$

Similarly the momentum operator is *quasi-local* in the sense that it can be expressed as a derivative rather than an integral

$$\langle q|\hat{p}_i|\psi\rangle = -i\hbar\frac{\partial}{\partial q_i}\langle q|\psi\rangle\,. \tag{1.29}$$

Formally we can write the matrix elements of the momentum operator as the derivative of a delta function

$$\langle q|\hat{p}_i|q'\rangle = -i\hbar\frac{\partial}{\partial q_i}\delta(q - q')\,. \tag{1.30}$$

(Check this by use of the integration formula for observables in the coordinate representation.)

Since the interactions in a quantum system usually has a local character the Hamiltonian will be (quasi-)local in the above sense, and can therefore be expressed as a differential operator as in the standard Schrödinger's (wave) equation. However, occasionally we may have to deal with non-local operators, which have to be expressed as integrals rather than derivatives.

From the abstract formulation it is clear that the coordinate representation is only one of many equivalent representations of quantum states and observables. The *momentum representation* is defined analogous to the coordinate representation, but now with the momentum states $|p\rangle$ as basis vectors,

$$\psi(p) = \langle p|\psi\rangle\,. \tag{1.31}$$

The transition matrix elements between the two representations is (for Cartesian coordinates),

$$\langle q|p \rangle = (2\pi\hbar)^{-N/2} \exp(\frac{i}{\hbar} q \cdot p) \,, \tag{1.32}$$

with $q \cdot p = \sum_i q_i p_i$, which means that these two (conjugate) representations are related by a Fourier transformation.

Note that often a set of continuous (generalized) coordinates is not sufficient to describe the wave function. For example, the spin variable of a particle with spin has discrete eigenvalues and does not have a direct counterpart in terms of a continuous classical coordinate. With discrete variables present the wave function can be described as a multicomponent function

$$\psi_m(q) = \langle q, m|\psi \rangle \,, \tag{1.33}$$

where m represents the discrete variable, *e.g.* the spin component in the z-direction.

The coordinate representation and the momentum representation are only two specific examples of *unitarily equivalent representations* of the quantum system. In general the transition matrix elements between two representations, defined by orthonormal basis vectors $\{|a_n\rangle\}$ and $\{|b_m\rangle\}$,

$$U_{nm} = \langle a_n|b_m \rangle \,, \tag{1.34}$$

will satisfiy the condition

$$\sum_m U_{nm}(U_{mn'})^* = \sum_m \langle a_n|b_m\rangle\langle b_m|a'_n\rangle = \delta_{nn'} \,, \tag{1.35}$$

which means that U is a *unitary* matrix. In operator form this is expressed as

$$|b_n\rangle = \hat{U}\,|a_n\rangle \,, \quad \hat{U}\hat{U}^\dagger = \mathbb{1} \,, \tag{1.36}$$

and the corresponding representations are referred to as unitarily equivalent.

1.1.4 Spin-half system and the Stern Gerlach experiment

The postulates of quantum mechanics have far reaching implications. We have earlier stressed the close correspondence between the classical (phase space) theory and the quantum theory. Now we will study a special representation of the simplest quantum system, the *two-level system*, where some of the basic differences between the classical and quantum theory are apparent.

The electron spin gives an example of a *spin-half system*, and when the (orbital) motion of the electron is not taken into account the Hilbert space is reduced to a two-dimensional (complex) vector space. This two-dimensionality is directly related to the demonstration of Stern and Gerlach of the two spin states of silver atoms[8]. Their discovery is clearly incompatible with a classical model of spin as due to the rotation of a small body.

---

[8]When Stern and Gerlach performed the experiment in 1922, they did not realize that the measured spin could be identified as the intrinsic electron spin. However, a few years later the electron spin was discovered and a re-interpretation of the Stern-Gerlach experiment could be done.
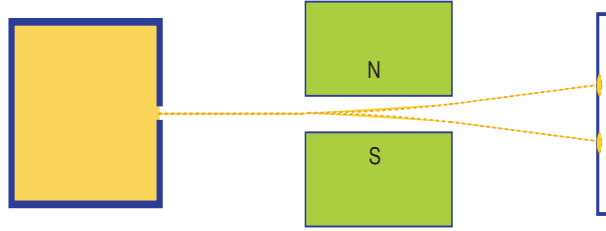
Figure 1.1: The Stern-Gerlach experiment. Atoms with spin 1/2 are sent as a beam from a hot source. When passing between two magnets, the atoms are deflected vertically, with an angle depending on the vertical spin component. Classically a smooth distribution is expected, since there is no preferred directions for the spin in the incoming beam. In reality only two directions are observed, consistent with the prediction of quantization of spin.

We focus on the Stern-Gerlach experiment as shown schematically in Figure 1. A beam of silver atoms is produced by a furnace with a small hole. Atoms with velocity sharply peaked around a given value are selected and sent through a strong magnetic field (in the z-direction). Due to a weak gradient in the magnetic field the particles in the beam are deflected, with a deflection angle depending on the component of the magnetic moment in the direction of the gradient. The degree of deflection is measured by registering the particles on a screen.

Let us first analyze the deflection from a classical point of view. We assume the atoms to have a magnetic moment $\boldsymbol{\mu} = (e/m_e)\mathbf{S}$, where $\mathbf{S}$ is the intrinsic electron spin, e is the charge and m_e is the electron mass. (The main contribution to the magnetic moment comes from the outermost electron.) Between the magnets the spin will rapidly precess around the magnetic field and the average value will be in the direction of the magnetic field. Furthermore, the gradient in the magnetic field will produce a force on the atom and change its momentum. Assuming the field vector to be dominated by its z component, we have

$$\dot{\boldsymbol{p}} = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \mathbf{B}) \approx \mu_z \frac{\partial B_z}{\partial z}\mathbf{k}\,, \tag{1.37}$$

which shows that the deflection angle is proportional to the component of the magnetic moment μ_z along the magnetic field. As a consequence we can regard the distribution of atoms on the screen to directly represent the distribution of the z-component of the magnetic moment (and spin) of the atoms in the incoming beam. Since we expect the spin direction of the emitted atoms to be randomly distributed in space, a classical reasoning will indicate that one should see a continuous distribution of atoms on the screen.

The experiment of Stern and Gerlach did not show such a continuous distribution. Instead the position of the atoms were rather strongly restricted to two spots, which according to the deflection formula would correspond to two possible measured values for the z-component of the magnetic moment,

$$\mu_z = \pm\mu\,. \tag{1.38}$$

This result cannot easily be explained within classical theory. To demonstrate this more directly, let us assume the y-component of the magnetic moment to be measured in a similar way by rotating the magnets. Since there is no preferred direction orthogonal to the beam, the

possible results of measuring the component of the magnetic moment the y- direction should be the same as for the z-direction,

$$\mu_y = \pm\mu \, . \tag{1.39}$$

Let us further consider the component of the magnetic moment of $\boldsymbol{\mu}$ in some rotated direction in the y, z-plane. For this component we have

$$\mu_\phi = \cos\phi\mu_y + \sin\phi\mu_z \, , \tag{1.40}$$

with ϕ as the rotation angle relative to the y-axis. Again we may argue that due to rotational symmetry, the possible measured values of μ_ϕ should be the same as for μ_y and μ_z,

$$\mu_\phi = \pm\mu \, . \tag{1.41}$$

This clearly leads to a contradiction. The condition of discrete values for the components (1.38), (1.39) and (1.41) is not consistent with the decomposition (1.40) for a continuous set of angles ϕ. Within the framework of classical theory the observation of the discreteness of the components of the magnetic moment thus leads to a paradoxical situation.

However, the results of the Stern-Gerlach experiment are consistent with the postulates of quantum mechanics. If we assume that the spin component in a given direction is an observable with only two eigenvalues

$$\hat{S}_x \left| \pm \right\rangle_x = \pm\frac{\hbar}{2} \left| \pm \right\rangle_x \, , \tag{1.42}$$

and the components satisfy the spin algebra

$$\left[\hat{S}_x, \hat{S}_y \right] = i\hbar\,\hat{S}_z \quad \text{(+ cycl. perm.)} \, , \tag{1.43}$$

then the component of the spin vector in *any* direction will have the two eigenvalues $\pm\hbar/2$. A similar conclusion is valid for components of the magnetic moment operator $\hat{\boldsymbol{\mu}} = (e/m_e)\hat{\mathbf{S}}$, so that Eqs.(1.38), (1.39) and (1.41) are valid if we interpret the equations as applying to the *eigenvalues* of these components.

Since the components of the magnetic moment operator do not commute, *i.e.*, they are *incompatible* observables, they cannot in general be ascribed sharp values at the same time. This incompatibility is directly related to the paradox discussed above when we in Eq.(1.40) ascribe sharp values to components in several different directions. The equation is valid also in the quantum description, but only if the components are interpreted as operators $\hat{\mu}_z$, $\hat{\mu}_y$ and $\hat{\mu}_\phi$. For the eigenvalues, which correspond to the measurable values of the components of the magnetic moment, the equation is not valid. This resolves the paradox.

In the Stern-Gerlach experiment we meet a situation where a vector, which can be *continuously* rotated, has components that nevertheless can take only *discrete* values. This cannot be explained within the framework of classical theory, but it can be explained in quantum theory.

1.2   Field quantization

The classical description of a physical system usually makes a clear distinction between particle and field variables. Quantum physics seems to blur this distinction due to what is known as *particle-wave duality*. For a particle like the electron this is apparent when the quantum dynamics is expressed in the form of a wave equation rather than a particle equation. However, also systems, which in the classical description appear as fields, will have a dual, particle nature. A well-known example is the photon description of the quantized electromagnetic field. At the formal level the quantum description of the two types of systems, particles and fields, is quite similar, with the physical variables expressed in the form of quantum states and observables. Also the transition from the classical to the quantum description, in the form of *canonical quantization*, can be formulated in much the same way. In this section we will discuss how to quantize by this method a simple, one-dimensional field theory, with the physical interpretation of a vibrating string. Even if this is a simple example of a field theory, it can be viewed as a prototype for more general theories. Later, in Chapt. 4, we will apply the method of canonical quantization to the electromagnetic field.

Let $\xi(x,t)$ denote the time dependent displacement of a pointlike element of the the string, with linear coordinate x. The displacement satisfies, for small deviations from equilibrium, the one-dimensional wave equation

$$\frac{\partial^2 \xi}{\partial t^2} - v^2 \frac{\partial^2 \xi}{\partial x^2} = 0\,, \tag{1.44}$$

with v as the wave velocity of the string. It is determined by the mass density μ and the string tension τ as $v = \sqrt{\tau/\mu}$. A general solution of the equation can be written as a combination of right- and left-going waves, the two types of motion defined by

$$\xi_\pm(x,t) = \xi_\pm(x \mp vt)\,. \tag{1.45}$$

Assuming the string to have fixed endpoints at $x = 0$ and $x = a$, the field $\xi(x,t)$ will satisfy the boundary conditions

$$\xi(0,t) = \xi(a,t) = 0\,, \tag{1.46}$$

with the general solution as a superposition of standing waves

$$\xi(x,t) = \sum_{k=1}^{\infty} \sin(k\pi\frac{x}{a})\xi_k(t)\,. \tag{1.47}$$

Thus k labels the independent vibrational modes, or normal modes, of the string. The field equation (1.44) implies that $\xi_k(t)$ satisfies

$$\frac{d^2 \xi_k}{dt^2} + \omega_k^2 \xi_k = 0\,, \qquad \omega_k \equiv k\frac{v\pi}{a}\,, \tag{1.48}$$

which clearly defines an infinite set of independent harmonic oscillator equations, for $k = 1, 2, ..., \infty$.

The variables ξ_k define a natural set of *generalized coordinates* for the system. To define the corresponding generalized momenta the classical Lagrangian L of the system is needed.

The field equation (1.44) can be identified with Lagrange's equation, with L defined in the standard way as $L = T - V$, where T is the kinetic and V is the potential energy of the vibrating string. Expressing L in terms of ξ_k and its time derivative, we find

$$L = \frac{1}{2} m \sum_k (\dot{\xi}_k^2 - \omega_k^2 \xi_k^2) \,, \tag{1.49}$$

with $m = a\mu/2$. The generalized momentum conjugate to ξ_k is then given by

$$\pi_k = \frac{\partial L}{\partial \dot{\xi}_k} = m \dot{\xi}_k \,, \tag{1.50}$$

and the classical Hamiltonian is

$$H = \sum_k \dot{\xi}_k \pi_k - L = \frac{1}{2m} \sum_k (\pi_k^2 + m^2 \omega_k^2 \xi_k^2) \,. \tag{1.51}$$

Quantization is now strait forward. The classical variables ξ_m and π_n are replaced by operators $\hat{\xi}_k$ and $\hat{\pi}_l$, which satisfy Heisenberg's commutation rule

$$\left[\hat{\xi}_k, \hat{\pi}_l \right] = i\hbar \delta_{kl} \,, \tag{1.52}$$

and by introducing the following linear combinations of the field components and their hermitian conjugate momenta

$$\hat{a}_k = \frac{1}{\sqrt{2m\hbar\omega_k}} (m\omega_k \hat{\xi}_k + i\hat{\pi}_k) \,, \quad \hat{a}_k^\dagger = \frac{1}{\sqrt{2m\hbar\omega_k}} (m\omega_k \hat{\xi}_k - i\hat{\pi}_k) \,, \tag{1.53}$$

the operators $\hat{a}_k$ and $\hat{a}_k^\dagger$ can be identified as ladder operators of the harmonic oscillators, with the standard commutation relations

$$\left[\hat{a}_k, \hat{a}_l^\dagger \right] = \delta_{kl} \,. \tag{1.54}$$

The quantum Hamiltonian, derived from (1.49), gets the standard form for a set of uncoupled harmonic oscillators

$$\hat{H} = \sum_k \left(\frac{1}{2m} \hat{\pi}_k^2 + \frac{1}{2} m\, \omega_k^2\, \hat{\xi}_k^2 \right) = \sum_k \hbar\omega_k (\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}) \,, \tag{1.55}$$

with the operators $\hat{a}_k^\dagger$ interpreted as creation operators for field quanta and $\hat{a}_k$ as annihilation operators for the same quanta.

The quantization condition for the variables of the normal modes, as discussed above, can be re-expressed more directly in terms of the field variable $\xi(x,t)$. To show this we first write the Lagrangian as $L = \int dx \mathcal{L}$, with the Lagrangian *density* given by

$$\mathcal{L} = \frac{1}{2} \mu \left[\dot{\xi}^2 - v^2 \left(\frac{\partial \xi}{\partial x} \right)^2 \right] \,, \tag{1.56}$$

where $\dot{\xi} = \frac{\partial \xi}{\partial t}$. The conjugate *field momentum density* is then defined as

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\xi}} = \mu \, \dot{\xi}(x) = \sum_{k=1}^{\infty} \sin(k\pi \frac{x}{a})\pi_k \,. \tag{1.57}$$

It is straight forward to verify that the commutator relation (1.52) then implies the following field commutator

$$\left[\hat{\xi}(x), \hat{\pi}(x')\right] = i\hbar \delta(x - x') \,. \tag{1.58}$$

Even if it is here derived from (1.52), it is quite standard to consider this rather as the fundamental commutation relation in the field description. Thus, field quantization implies first to establish the Lagrangian density of the field, and to derive the conjugate field momentum. Quantization is then introduced in the form of the fundamental commutator between the field variable and its conjugate field momentum. Due to the continuous character of the field variable the field commutator is expressed by a Dirac delta function rather than a Kronecker delta, which is the case when the variables are discrete.

The Hilbert space can be constructed in the same manner as for a single harmonic oscillator. This means that we first define the ground state $|0\rangle$ by

$$\hat{a}_k|0\rangle = 0 \,, \quad k = 1, 2, \dots \,. \tag{1.59}$$

Excited states are produced by acting on this state with the creation operators, and the a general energy eigenstate is thus characterized by a set of integers, which give the number of field quanta for each field mode

$$|\psi_E\rangle = |n_1, n_2, n_3, \dots\rangle = \mathcal{N}(\hat{a}_1^{\dagger})^{n_1}(\hat{a}_2^{\dagger})^{n_2} \, , (\hat{a}_3^{\dagger})^{n_3}\dots|0\rangle \,, \tag{1.60}$$

with $\mathcal{N}$ as a normalization factor.

This way of quantizing the field, as a system of many non-interacting harmonic oscillators, seems to work smoothly. However, the presence of an infinite number of field modes will in fact introduce some problems which have to be handled. A particular problem has to do with the ground state fluctuations of the fields. For a single harmonic oscillator these fluctuations give rise to the non-zero ground state energy $\hbar\omega/2$. For the field theory the corresponding ground state energy is

$$E_0 = \sum_{k=1}^{\infty} \frac{1}{2}\hbar\omega_k \,, \tag{1.61}$$

but the problem is that the sum does not converge to a finite value. However, a simple solution to this problem is to modify slightly the definition of the Hamiltonian, by subtracting the ground state energy. The new, well-defined expression for the Hamiltonian is

$$\hat{H} = \sum_{k} \hbar\omega_k \hat{a}_k^{\dagger}\hat{a}_k \,, \tag{1.62}$$

and since the subtracted term is a constant, this redefinition will not affect other observables of the system. All the observables of the theory can now be expressed in terms of the creation and

annihilation operators. This is in particular the case for the field operator of the string, which in the *Heisenberg picture* can be written as

$$\hat{\xi}(x,t) = \sum_{k=1}^{\infty} \sqrt{\frac{\hbar}{2m\omega_k}} [\sin(k\pi\frac{x}{a})e^{-i\omega_k t}\hat{a}_k + \sin(k\pi\frac{x}{a})e^{i\omega_k t}\hat{a}_k^{\dagger}] \,. \tag{1.63}$$

It has the same form as the classical field expanded in terms of normal modes, but with the expansion coefficients here replaced by the annihilation and creation operators.

The particle interpretation of the system can now be understood in the following way. The ground state $|0\rangle$ is identified as the vacuum state, *i.e.*, the state with no particle present. The single particle states are then the states created from the vacuum state by operators linear in the creation operators, $\hat{a}_k^{\dagger}$. Similarly the two-particle states are created by quadratic operators $\hat{a}_k^{\dagger}\hat{a}_{k'}^{\dagger}$ etc. Since the operator $\hat{a}_k^{\dagger}$ can be applied repeatedly to create many field quanta in the same state, this means that the corresponding particles should be classified as bosons.

The example discussed above shows a simple example of the general method used for (canonical) field quantization. Thus, quite generally the method consists in first identifying the free field part, which in the Lagangian include all terms quadratic in the fundamental fields. They define the independent field modes, which then are quantized like a collection of independent harmonic oscillators. The non-quadratic terms in the Lagrangian are identified as interaction terms. They do not affect the quantization procedure, but are re-expressed, after the quantization, in terms of creation and annihilation operators. Later in the course we shall show how to apply this method to electromagnetic fields interacting with electrons.

1.3 Quantum Dynamics

In this section we formulate the dynamical equation of a quantum system and discuss the relations between the unitarily equivalent descriptions known as the Schrödinger, Heisenberg and interaction pictures. We then examine the rather different Feynman's path integral formulation of quantum dynamics.

1.3.1 The different pictures of the time evolution

The Schrödinger picture.
The time evolution of an isolated quantum system is defined by the *Schrödinger equation*. Originally this was formulated as a wave equation, but it can be reformulated as a differential equation in the (abstract) Hilbert space of ket-vectors as

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \,. \tag{1.64}$$

With the state vector given for an initial time t_0, the equation will determine the state vector at later times t (and also at earlier times) as long as the system stays isolated. The information about the dynamics is contained in the Hamiltonian $\hat{H}$, which usually can be identified with the energy observable of the system. The original Schrödinger equation, described as a wave equation can be viewed as the coordinate representation of Eq.(1.64).

The dynamical evolution of the state vector can be expressed in terms of a *time evolution operator* $\hat{\mathcal{U}}(t, t_0)$, which is a unitary operator that relates the state vector of the system at time t with that of time t_0,

$$\hat{\mathcal{U}}(t, t_0)|\psi(t_0)\rangle = |\psi(t)\rangle\,. \tag{1.65}$$

The time evolution operator is determined by the Hamiltonian through the equation

$$i\hbar\frac{\partial}{\partial t}\hat{\mathcal{U}}(t, t_0) = \hat{H}\,\hat{\mathcal{U}}(t, t_0)\,, \tag{1.66}$$

which follows from the Schrödinger equation (1.64).

When $\hat{H}$ is a time independent, a closed form for the time evolution operator can be given[9]

$$\hat{\mathcal{U}}(t - t_0) = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}\,. \tag{1.68}$$

If however $\hat{H}$ is time dependent, so that the operator at different times do not commute, we may use a more general integral expression

$$\hat{\mathcal{U}}(t, t_0) = \sum_{n=0}^{\infty}\left(\frac{-i}{\hbar}\right)^n\int_{t_0}^t dt_1\int_{t_0}^{t_1}dt_2\cdots\int_{t_0}^{t_{n-1}}dt_n\,\hat{H}(t_1)\hat{H}(t_2)\cdots\hat{H}(t_n)\,. \tag{1.69}$$

The term corresponding to $n = 0$ in (1.69) is simply the unit operator $\hat{\mathbb{1}}$, and the full expression is generated from (1.66) by solving this equation iteratively as,

$$i\hbar\frac{\partial}{\partial t}\hat{\mathcal{U}}_{n+1}(t, t_0) = \hat{H}\,\hat{\mathcal{U}}_n(t, t_0)\,, \tag{1.70}$$

with $\hat{\mathcal{U}}_n(t, t_0)$ as the n'th order contribution to $\hat{\mathcal{U}}(t, t_0)$ in (1.69). Note that the product of the time dependent operators $\hat{H}(t_k)$ in each term of the expansion is a *time-ordered* product.

The Heisenberg picture.

The description of the quantum dynamics given above is referred to as the Schrödinger picture. From the discussion of different *representations* of the quantum system we know that a unitary transformation of states and observables leads to a different, but equivalent representation of the system. If we therefore denote the states of a system by $|\psi\rangle$ and the observables by $\hat{A}$ and make a unitary transformation $\hat{U}$ on *all* states and *all* observables,

$$|\psi\rangle \to |\psi'\rangle = \hat{U}|\psi\rangle\,, \quad \hat{A} \to \hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger\,, \quad \hat{U}^\dagger\hat{U} = \hat{\mathbb{1}}\,, \tag{1.71}$$

then all matrix elements are left unchanged,

$$\langle\phi'|\hat{A}'|\psi'\rangle = \langle\phi|\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\phi|\hat{A}|\psi\rangle\,, \tag{1.72}$$

---

[9]Note that a function of an observable $\hat{H}$, like $\exp(-\frac{i}{\hbar}\hat{H}(t - t_0))$ can be defined by its action on the eigenvectors $|E\rangle$ of $\hat{H}$,

$$e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}|E\rangle = e^{-\frac{i}{\hbar}E(t-t_0)}|E\rangle\,. \tag{1.67}$$

This follows since the eigenvectors form a complete set.

and since all measurable quantities can be expressed in terms of such matrix elements, the two descriptions related by a unitary transformation can be viewed as equivalent. This is true also when $\hat{U} = \hat{U}(t)$ is a time dependent transformation.

The transition to the *Heisenberg picture* is defined by a special time-dependent unitary transformation

$$\hat{U}(t) = \hat{\mathcal{U}}^\dagger(t, t_0) \,. \tag{1.73}$$

This is the inverse of the time-evolution operator, and when applied to the time-dependent state vector of the Schrödinger picture it will simply cancel the time dependence

$$|\psi\rangle_H = \hat{\mathcal{U}}^\dagger(t, t_0)|\psi(t)\rangle_S = |\psi(t_0)\rangle_S \,. \tag{1.74}$$

Here we have introduced a subscript S for the vector in the Schrödinger picture and H for the Heisenberg picture. (The initial time t_0 is arbitrary and is often chosen as $t_0 = 0$.) The time evolution is now carried by the observables, rather than the state vectors,

$$\hat{A}_H(t) = \hat{\mathcal{U}}^\dagger(t, t_0)\, \hat{A}_S\, \hat{\mathcal{U}}(t, t_0) \,, \tag{1.75}$$

and the Schrödinger equation is replaced by the Heisenberg equation of motion ,

$$\frac{d}{dt}\hat{A}_H = \frac{i}{\hbar}[\hat{H}, \hat{A}_H] + \frac{\partial}{\partial t}\hat{A}_H \,, \tag{1.76}$$

(with the Hamiltonian here assumed to be time independent). The partial derivative in this equation refers to a possible *explicit* time dependence of the observable in the Schrödinger picture,

$$\frac{\partial}{\partial t}\hat{A}_H = \hat{\mathcal{U}}^\dagger(t, t_0) \left(\frac{\partial}{\partial t}\hat{A}_S\right) \hat{\mathcal{U}}(t, t_0) \,. \tag{1.77}$$

This time dependence may be caused by some time varying external influence on the system, which in particular could also impose a time dependence on the Hamiltonian. The full time evolution of the observable $\hat{A}_H$ therefore may have two contributions, one is the dynamical contribution from the non-commutativity with the Hamiltonian and the other is the contribution from an explicit time dependence due to some external influence.

The interaction picture
A third representation of the unitary time evolution of a quantum system is the *interaction picture* which is particularly useful in the context of time-dependent perturbation theory. The Hamiltonian is of the form

$$\hat{H} = \hat{H}_0 + \hat{H}_1 \,, \tag{1.78}$$

where $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{H}_1$ is the (possibly time dependent) perturbation. We assume that the eigenvalue problem of $\hat{H}_0$ can be solved and that the corresponding time evolution operator is

$$\hat{\mathcal{U}}_0(t - t_0) = e^{-\frac{i}{\hbar}\hat{H}_0(t - t_0)} \,. \tag{1.79}$$

The transition from the Schrödinger picture to the interaction picture is defined by acting with the inverse of this on the state vectors

$$|\psi_I(t)\rangle = \hat{\mathcal{U}}_0^\dagger(t, t_0) \, |\psi_S(t)\rangle \,. \tag{1.80}$$

Note that the time variation of the state vector is only partly cancelled by this transformation, since the effect of the perturbation $\hat{H}_1$ is not included. The time evolution of the observables is given by

$$\hat{A}_I(t) = \hat{\mathcal{U}}_0^\dagger(t, t_0) \, \hat{A}_S \, \hat{\mathcal{U}}_0(t, t_0) \,. \tag{1.81}$$

This means that they satisfy the same Heisenberg equation of motion as for a system where the Hamiltonian is simply $\hat{H} = \hat{H}_0$. The remaining part of the dynamics is described by the interaction Hamiltonian

$$\hat{H}_I(t) = \hat{\mathcal{U}}_0^\dagger(t, t_0) \, \hat{H}_1 \, \hat{\mathcal{U}}_0(t, t_0) \,, \tag{1.82}$$

which acts on the state vectors through the (modified) Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi_I(t)\rangle = \hat{H}_I(t)|\psi_I(t)\rangle \,. \tag{1.83}$$

The corresponding time evolution operator has the same form as (1.69),

$$\hat{\mathcal{U}}_I(t, t_0) = \sum_{n=0}^{\infty} \left(\frac{-i}{\hbar} \right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \, \hat{H}_I(t_1)\hat{H}_I(t_2) \cdots \hat{H}_I(t_n) \,, \tag{1.84}$$

but it includes now only the interaction part of the Hamiltonian. This form of the time evolution operator gives a convenient starting point for a perturbative treatment of the effect of $\hat{H}_I$. We shall apply this method when studying the interaction between photons and atoms in a later chapter.

We summarize the difference between the three pictures by the following table

	States	Observables
Schrödinger	time dependent	time independent
Heisenberg	time independent	time dependent
Interaction	time dependent	time dependent

where we here have excluded the possibility of explicit time dependence of the observables.

1.3.2 Path integrals

Feynman's path integral method provides an approach to the dynamics of quantum systems that is rather different from the methods outlined above. Instead of applying the standard description of states as vectors in a Hilbert space, it focusses directly on transition matrix elements and describe these as integrals over classical trajectories of the system. The description has an intuitive appeal, since it is less abstract than the Hilbert space description. It describes the evolution of the system in terms of paths between the initial and final points of the evolution.

This makes the connection to the classical description rather close, but also shows the difference between the classical and quantum theories, since the system does not simply follow a single path from the initial to the final point. Instead the transition amplitude gets contribution from all paths with the given end points, as if the system during the time evolution tries out all possible trajectories.

The formal expression for the path integral is

$$\mathcal{G}(q_f t_f, q_i t_i) = \int \mathcal{D}[q(t)] \exp(\frac{i}{\hbar} S[q(t)]) \, . \tag{1.85}$$

In this expression q denotes a set of generalized coordinates of the system and $q(t)$ is a trajectory in the (classical) configuration space. $\mathcal{G}(q_f t_f, q_i t_i)$ is the *propagator* which defines the transition amplitude from an initial configuration q_i at time t_i to a final configuration q_f at time t_f. The integration, on the right hand side, is over all paths $q(t)$ that connect the initial and final configurations and $S[q(t)]$ is the action integral for a given path, so that

$$S[q(t)] = \int_{t_i}^{t_f} L(\dot{q}(t), q(t)) dt \, , \tag{1.86}$$

with $L(\dot{q}, q)$ as the Lagrangian of the system. The integral in Eq.(1.85) is referred to as a *functional integral*, since the integration variable is a function $q(t)$ rather than a set of discrete variables. The path integral is primarily a formal expression, since the conditions that the functions (the paths) have to satisfy are not specified in any strict way, and neither is the integration measure. In principle all curves included are equally important, since the weight factor of any curve is a phase factor of modulus 1. However, there seems to be an implicit suppression in importance of paths where the phase factor in (1.85) varies rapidly with changes in the path. For these paths the contribution to the integral is reduced due to destructive interference between contributions from nearby paths.
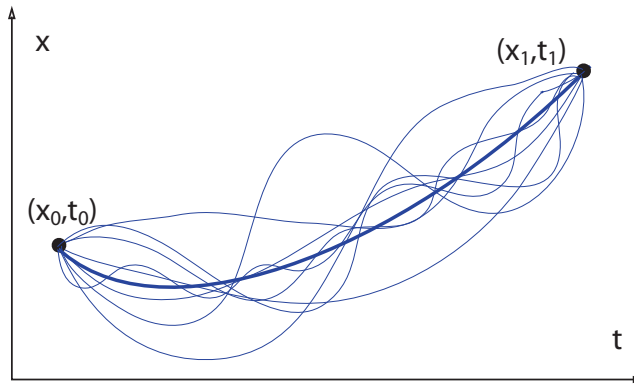

Figure 1.2: The path integral as a "sum over histories". All possible paths between the initial point (x_0, t_0) and the final point (x_1, t_1) contribute to the quantum transition amplitude between the points. The paths close to the classical path, here shown in dark blue, tend to be most important since their contributions interfere constructively.

Although intuitively attractive, it is well known that the path integral (1.85) is difficult to make mathematically precise. Only in the simplest cases it is possible to give a precise meaning to the the set of paths and to introduce a well defined integration measure on this set. Even so, the path integral method is an important method in physics and can often be used without a rigorous definition of the integral. It is often used together with semiclassical approximations and is particularly important in the study of non-perturbative effects. In quantum field theory it is an important tool, in the form of generating functionals for the correlation functions of the field, and it is also much used in the transformation between different sets of field variables for the physical system. In any case, the path integral method should be viewed as an important supplementary method, and not as a possible replacement of traditional quantum mechanical methods. In the computation of quantum effects, in particular when using perturbation theory, methods based on the Hilbert space formulation continue to be the most important ones.

Although the path integral method can be viewed as a fundamental approach to quantum theory, *i.e.*, a method that completely circumvents the standard description with state vectors and observables, it is often instead derived from the Hamiltonian formulation, and that is the approach we shall take also here. We will in this derivation meet some of the mathematical problems of the path integral approach, but will only comment on these and not go into any discussion of how to deal with these problems in a serious way.

Let us consider the time evolution of a quantum system as a wave function $\psi(q, t)$ in configuration space, where $q = \{q_1, q_2, ..., q_N\}$ is a set of continuous (generalized) coordinates. In the "bra-ket" notation we write it as

$$
\begin{aligned}
\psi(q, t) &= \langle q | \psi(t) \rangle \\
&= \langle q | \hat{\mathcal{U}}(t, t') | \psi(t') \rangle \\
&= \int dq' \langle q | \hat{\mathcal{U}}(t, t') | q' \rangle \langle q' | \psi(t') \rangle \\
&\equiv \int dq' \langle q\, t | q'\, t' \rangle \, \psi(t') \, ,
\end{aligned}
\tag{1.87}
$$

with dq' denoting the N-dimensional volume element. The information about the dynamics of the system is encoded in the matrix element of the time evolution operator, or transition matrix element,

$$
\langle q | \hat{\mathcal{U}}(t, t') | q' \rangle = \langle q\, t | q'\, t' \rangle \equiv \mathcal{G}(q\, t, q'\, t') \, ,
\tag{1.88}
$$

which we identify as the propagator previously expressed in the form of the path integral.

We will now see how a path integral representation of this propagator can be found in the simple case for a system with a one-dimensional configuration space. As a concrete realization we consider a particle moving on a line, with the set of coordinates q replaced by a single variable x.

The propagation between an initial time t_i and a final time t_f can be viewed as composed of the propagation between a series of intermediate times $t_k, k = 0, 1, ...n$ with $t_0 = t_i$ and $t_n = t_f$,

$$
\begin{aligned}
\mathcal{G}(x_f\, t_f, x_i\, t_i) = \int dx_{n-1}... \int dx_2 \int dx_1 \mathcal{G}(x_f\, t_f, x_{n-1}\, t_{n-1}) \\
\times \mathcal{G}(x_{n-1}\, t_{n-1}, x_{n-2}\, t_{n-2})...\mathcal{G}(x_2\, t_2, x_1\, t_1)\mathcal{G}(x_1\, t_1, x_i\, t_i) \, .
\end{aligned}
\tag{1.89}
$$

This follows from a repeated use of the composition rule satisfied by the time evolution operator

$$\hat{\mathcal{U}}(t_f, t_i) = \hat{\mathcal{U}}(t_f, t_m)\,\hat{\mathcal{U}}(t_m, t_i)\,, \tag{1.90}$$

where t_i, t_m and t_f are arbitrary chosen times. In the expression (1.89) for $\mathcal{G}(x_f\ t_f, x_i\ t_i)$ the intermediate times $t_1, t_2...$ may also be arbitrarily distributed between t_i and t_f, but for simplicity we think of them as having a fixed distance

$$t_{k+1} - t_k = \Delta t \equiv (t_f - t_i)/n\,. \tag{1.91}$$

The number n of time steps may be taken arbitrary large.

To proceed we assume a specific form for the Hamiltonian,

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x})\,, \tag{1.92}$$

which is that of a particle of mass m moving in a one-dimensional potential $V(x)$. The propagator for a small time interval Δt is

$$\begin{aligned}
\mathcal{G}(x\ t + \Delta t, x'\ t) &= \langle x|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|x'\rangle \\
&\approx \langle x|e^{-\frac{i}{\hbar}\frac{1}{2m}\hat{p}^2\Delta t}e^{-\frac{i}{\hbar}V(\hat{x})\Delta t}|x'\rangle \\
&= \langle x|e^{-\frac{i}{\hbar}\frac{1}{2m}\hat{p}^2\Delta t}|x'\rangle e^{-\frac{i}{\hbar}V(x')\Delta t}\,.
\end{aligned} \tag{1.93}$$

We have here made use of

$$e^{i(\hat{A}+\hat{B})\Delta t} = e^{i\hat{A}\Delta t}e^{i\hat{B}\Delta t} + \mathcal{O}(\Delta t^2)\,, \tag{1.94}$$

where $\hat{A}$ and $\hat{B}$ are two (non-commuting) operators and the $\mathcal{O}(\Delta t^2)$ term comes from the commutator between $\hat{A}$ and $\hat{B}$. In the present case $\hat{A} = \hat{p}^2/(2m\hbar)$, $\hat{B} = \hat{V}/\hbar$, and these clearly do not commute. However, we will take the limit $\Delta t \to 0$ ($n \to \infty$) and this allows us to neglect the correction term coming from the commutator, since this includes the factor Δt^2. The x-space matrix element of the kinetic term can be evaluated

$$\begin{aligned}
\langle x|e^{-\frac{i}{\hbar}\Delta t\frac{\hat{p}^2}{2m}}|x'\rangle &= \int dp\, \langle x|p\rangle e^{-\frac{i}{\hbar}\frac{p^2}{2m}\Delta t}\langle p|x'\rangle \\
&= \int \frac{dp}{2\pi\hbar}\, e^{\frac{i}{\hbar}p(x-x')}e^{-\frac{i}{\hbar}\frac{p^2}{2m}\Delta t} \\
&= \int \frac{dp}{2\pi\hbar}\, e^{-\frac{i}{\hbar}\frac{\Delta t}{2m}(p-m\frac{x-x'}{\Delta t})^2}e^{\frac{i}{\hbar}\Delta t\frac{m}{2}\left(\frac{x-x'}{\Delta t}\right)^2} \\
&= N_{\Delta t}\, e^{i\frac{m(x-x')^2}{2\hbar\Delta t}}\,,
\end{aligned} \tag{1.95}$$

where $N_{\Delta t}$ is an x-independent normalization constant,

$$\begin{aligned}
N_{\Delta t} &= \int \frac{dp}{2\pi\hbar}\, e^{-i\frac{\Delta t}{2m\hbar}(p-m\frac{x-x'}{\Delta t})^2} \\
&= \int \frac{dp}{2\pi\hbar}\, e^{-i\frac{\Delta t}{2m\hbar}p^2}\,.
\end{aligned} \tag{1.96}$$

This last expression may look somewhat mysterious, since the integral does not converge for large p. This is one of the places where we note that the path integral is not fully defined without some further specification. To make the expression well defined we focus on a related convergent integral, the Gaussian integral

$$\int_{-\infty}^{+\infty} dp\, e^{-\lambda\, p^2} = \sqrt{\frac{\pi}{\lambda}}\,, \tag{1.97}$$

where λ has a real, positive part. If we write the coefficient as

$$\lambda = i\frac{\Delta t}{2m\hbar} + \epsilon\,, \tag{1.98}$$

the integral is convergent for an arbitrarily small real part ϵ. This means that we can take the limit $\epsilon \to 0^+$, and obtain the expression

$$N_{\Delta t} = \sqrt{\frac{m}{2\pi i\hbar\Delta t}}\,. \tag{1.99}$$

For the matrix element of the time evolution operator we then get,

$$\langle x|e^{-\frac{i}{\hbar}\Delta t H}|x'\rangle = N_{\Delta t}\, e^{i\frac{m(x-x')^2}{2\hbar\Delta t}} e^{-\frac{i}{\hbar}V(x')\Delta t}\,, \tag{1.100}$$

and with $x' \to x_k$ and $x \to x_{k+1}$ this can be used for each term in the factorized expression (1.89) for the propagator. The result is

$$\mathcal{G}(x_f\, t_f, x_i\, t_i) = (N_{\Delta t})^n \int dx_{n-1}... \int dx_2 \int dx_1\, e^{\frac{i}{\hbar}\Delta t \sum_{k=0}^{n}\left[\frac{m}{2}(\frac{x_{k+1}-x_k}{\Delta t})^2 - V(x_k)\right]}\,. \tag{1.101}$$

The exponent can be further simplified in the limit $n \to \infty$,

$$\frac{i}{\hbar}\Delta t \sum_{k=0}^{n}\left[\frac{m}{2}(\frac{x_{k+1}-x_k}{\Delta t})^2 - V(x_k)\right] \to \frac{i}{\hbar}\int_{t_0}^{t} dt[\frac{1}{2}m(\frac{dx}{dt})^2 - V(x)]\,, \tag{1.102}$$

where we have now assumed that the sequence of intermediate positions x_k (which we integrate over) in the limit $n \to \infty$ defines a differentiable curve. The expression we arrive at can be identified as the (classical) action associated with the curve defined by the positions x_k as functions of time,

$$\mathcal{S}[x(t)] = \int_{t_0}^{t} L(x,\dot{x})dt = \int_{t_0}^{t}(\frac{1}{2}m\dot{x}^2 - V(x))dt\,. \tag{1.103}$$

In the continuum limit ($n \to \infty$) we therefore write the propagator as

$$\mathcal{G}(x_f\, t_f, x_i\, t_i) = \int \mathcal{D}[x(t)]e^{\frac{i}{\hbar}\mathcal{S}[x(t)]}\,, \tag{1.104}$$

which has the form (1.85) originally written for the path integral. We may now simply take the discretized form (1.101) as *defining* the path integral. This means that the formal expression (1.104) is interpreted as being identical to the multiple integral (1.101) in the limit $n \to \infty$. However, this is not completely satisfactorily since the independent integration over intermediate positions x_k is not really consistent with the picture of the integral as being an integration over *continuous* curves. So the derivation should rather be taken as suggestive for the idea that the path integral may be made well defined with some further specifications and that it can be related to the transition matrix element derived in the Schrödinger picture in the way outlined above.

1.3.3 Continuous paths for a free particle

The discretization of time is convenient when we examine the connection between the Hamiltonian formulation and the path integral formulation of quantum mechanics. However, as pointed out above, the discretization is a complication for the idea of regarding the path integral as a sum over contributions from *continuous* paths. We will here examine another formulation which respects more the idea of paths, and apply it to the example of a free particle.

We then consider a path as a continuous curve $x(t)$ which connects an initial point $x_i = x(t_i)$ with a final point $x_f = x(t_f)$, and denote the time difference as $T = t_f - t_i$. With the endpoints of the curve fixed, an arbitrary curve between these points can be written as

$$x(t) = x_{cl}(t) + \sum_{n=1}^{\infty} c_n \sin(n\pi \frac{t - t_i}{T}) \,, \tag{1.105}$$

where $x_{cl}(t)$ is a solution of the classical equation of motion with the given end points. The deviation from the classical curve is expanded in a Fourier series. We shall now interpret the path integral as an independent integration over each Fourier component c_n. Note that even if the variables c_n form a discrete set, the curve (1.105) may be continuous.

The action for a free particle is given by

$$
\begin{aligned}
S[x(t)] &= \int_{t_i}^{t_f} dt \frac{1}{2} m \dot{x}^2 \\
&= S[x_{cl}(t)] + \frac{1}{2} m \int_{t_i}^{t_f} dt \sum_{nn'} c_n c_{n'} \frac{nn'\pi^2}{T^2} \cos(n\pi \frac{t - t_i}{T}) \cos(n'\pi \frac{t - t_i}{T}) \\
&= S[x_{cl}(t)] + \frac{1}{2} m \int_{0}^{\pi} d\phi \sum_{nn'} c_n c_{n'} \frac{nn'\pi}{T} \cos(n\phi) \cos(n'\phi) \\
&= S[x_{cl}(t)] + \frac{m\pi^2}{4T} \sum_{n} n^2 c_n^2 \,, \tag{1.106}
\end{aligned}
$$

where the term linear in $\dot{x}_{cl}$ is absent since the action is stationary under first order variations in $x(t)$ about the classical path $x_{cl}(t)$. (See the discussion in the next subsection.) For the

propagator this gives

$$
\begin{aligned}
\mathcal{G}(x_f\ t_f, x_i\ t_i) &= \int \mathcal{D}[x(t)] e^{\frac{i}{\hbar}\mathcal{S}[x(t)]} \\
&= \mathcal{N} e^{\frac{i}{\hbar}\mathcal{S}[x_{cl}(t)]} \prod_n \int dc_n\, e^{i\frac{m\pi^2}{4T\hbar}n^2 c_n^2}\,,
\end{aligned}
\tag{1.107}
$$

with $\mathcal{N}$ as an unspecified normalization constant. The integrals have to be made well-defined by the same trick as before, by adding a small real part to the imaginary coefficient. This gives

$$
\begin{aligned}
\mathcal{G}(x_f\ t_f, x_i\ t_i) &= \mathcal{N} e^{\frac{i}{\hbar}\mathcal{S}[x_{cl}(t)]} \prod_n (\frac{2}{n}\sqrt{\frac{iT\hbar}{m\pi}}) \\
&= \mathcal{N}' e^{\frac{i}{\hbar}\mathcal{S}[x_{cl}(t)]} \\
&= \mathcal{N}' e^{\frac{i}{\hbar}\frac{1}{2}m\frac{(x_f-x_i)^2}{t_f-t_i}}\,.
\end{aligned}
\tag{1.108}
$$

Note that the product over n is not well-defined as a separate factor, but it has here been absorbed in $\mathcal{N}$ to form a new normalization factor $\mathcal{N}'$. This factor will depend on the precise definition of the path integral. The form of the product over n indicates that such a definition should include a prescription for regularizing the contributions for large n.

   In the simple case we consider here the propagator can be evaluated directly, and we use the expression to check our result from the path integral formulation and to determine $\mathcal{N}'$,

$$
\begin{aligned}
\mathcal{G}(x_f\ t_f, x_i\ t_i) &= \langle x_f|e^{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}T}|x_i\rangle \\
&= \int dp\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}T}\langle x_f|p\rangle\langle p|x_i\rangle \\
&= \int \frac{dp}{2\pi\hbar} e^{-\frac{i}{\hbar}[\frac{p^2}{2m}T - p(x_f-x_i)]} \\
&= \sqrt{\frac{m}{2\pi i\hbar T}} e^{i\frac{m(x_f-x_i)^2}{2\hbar(t_f-t_i)}}\,.
\end{aligned}
\tag{1.109}
$$

This agrees with the expression (1.108) and determines $\mathcal{N}'$. We note that the exponential factor is determined by the action of the classical path between the initial and final points while the path integral only determines the prefactor. A similar expression for the propagator, in terms of the action of the classical path, can be found when the action is quadratic in both q and $\dot{q}$.

1.3.4 The classical theory as a limit of the path integral

One of the advantages of the Feynman path integral is its close relation to the classical theory. This is clear already from the formulation in terms of the classical Lagrangian of the system. Let us write the path integral in the general form

$$
\begin{aligned}
\mathcal{G}(q_f\ t_f, q_i\ t_i) &= \int \mathcal{D}[q(t)] e^{\frac{i}{\hbar}\mathcal{S}[q(t)]} \\
&= \int \mathcal{D}[q(t)] \exp(\frac{i}{\hbar}\int_{t_i}^{t_f} L(q,\dot{q})dt)\,,
\end{aligned}
\tag{1.110}
$$

where the Lagrangian $L(q, \dot{q})$ depends on a set of generalized coordinates q and their derivatives $\dot{q}$. We note from this formulation that variations in the path $q(t)$ that give rise to rapid variations in the action $\mathcal{S}[q(t)]$ tend to give contributions to the path integral that add destructively. This is so because of the rapid change in the complex phase of the integrand.

The classical limit of a quantum theory is often thought of as a formal limit $\hbar \to 0$. From the expression for the path integral we note that smaller $\hbar$ means more rapid variation in the complex phase. This indicates that in the classical limit most of the paths will not contribute to the path integral, since variations in the action of the neighboring paths will tend to "wash out" the contributions due to destructive interference. The only paths which retain their importance are those where the action is stationary, *i.e.*, where the action does not change under small variations in the path.

The stationary paths are characterized by $\delta \mathcal{S} = 0$, with

$$
\begin{aligned}
\delta \mathcal{S} &= \sum_k \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q_k} \delta q_k + \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k \right) dt \\
&= \sum_k \int_{t_i}^{t_f} \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k dt \,.
\end{aligned}
\tag{1.111}
$$

In this expression δq_k denotes an (infinitesimal) variation in the path, and $\delta \dot{q}_k$ the corresponding variation in the time derivative. The last expression in (1.111) is found by a partial integration and applying the constraint on the variation that it vanishes in the end points, $\delta q_k(t_i) = \delta q_k(t_f) = 0$. This constraint follows from the fact that the end points of the paths are fixed by the coordinates of the propagator (1.110).

Thus, the important paths are those with stationary action, and these satisfy the Euler-Lagrange equations,

$$
\frac{\partial L}{\partial q_k} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) = 0 \,, \quad k = 1, 2, ..., N \,,
\tag{1.112}
$$

since $\delta \mathcal{S}$ should be 0 for *all* (infinitesimal) variations. In a Lagrangian formulation of the classical system, with dynamics determined by $L(q, \dot{q})$, these are exactly the classical equations of motion.

1.3.5 A semiclassical approximation

As discussed above, in the classical limit the relevant contributions to the path integral come from paths in the immediate neighborhood of the solutions to the classical equation of motion. We refer to these classical paths as $q_{cl}(t)$. This motivates a semiclassical approach, where we make a lowest order expansion of the action around the stationary paths. The coordinates of the paths in this neighborhood we write as

$$
q(t) = q_{cl}(t) + \eta(t) \,,
\tag{1.113}
$$

where $q(t)$ represents the full set of coordinates $\{q_k(t), k = 1, ..., N\}$. We assume the deviation $\eta(t)$ from the classical solution, $q_{cl}(t)$, which satisfies a given set of boundary conditions, to be

small. Since $q_{cl}(t)$ is supposed to satisfy the correct boundary conditions, $\eta(t)$ should vanish at the end points.

We introduce the approximation by assuming that the action can be expanded to second order in $\eta(t)$ and that higher orders can be neglected. Thus,

$$S[q(t)] = S[q_{cl}(t)] + \Delta S[q(t)] , \qquad (1.114)$$

with

$$\Delta S[q(t)] = \sum_{ij} \int dt \, \frac{1}{2} \left[\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \dot{\eta}_i \dot{\eta}_j + 2 \frac{\partial^2 L}{\partial q_i \partial \dot{q}_j} \eta_i \dot{\eta}_j + \frac{\partial^2 L}{\partial q_i \partial q_j} \eta_i \eta_j \right]$$

$$\equiv \sum_{ij} \int dt \, \frac{1}{2} \left[A_{ij} \dot{\eta}_i \dot{\eta}_j + 2 B_{ij} \eta_i \dot{\eta}_j + C_{ij} \eta_i \eta_j \right] . \qquad (1.115)$$

In this approximation, with $\eta(t)$ as a new set of variables, the Lagrangian is quadratic in the co-ordinates and velocities. The path integral for such a Lagrangian can (in principle) be evaluated and the general form is

$$\mathcal{G}(q_f t_f, q_i t_i) = N e^{\frac{i}{\hbar} S_{cl}(q_f t_f, q_i t_i)} , \qquad (1.116)$$

where N is the contribution from the integral over paths $\eta(t)$, and $S_{cl}(q_f t_f, q_i t_i)$ is the action of the classical path with the given end points. If the coefficients A_{ij}, B_{ij} and C_{ij} are time independent along the path, N will only depend on the length of the time interval. As a special case we have previously seen this in the evaluation of the propagator of a free particle.

In some cases there may be more than one classical path connecting the two points (q_i, t_i) and (q_f, t_f). In that case the path integral is given by a sum over the classical paths

$$\mathcal{G}(q_f t_f, q_i t_i) = \sum_{cl} N_{cl} e^{\frac{i}{\hbar} S_{cl}(q_f t_f, q_i t_i)} , \qquad (1.117)$$

and the strength of the transition amplitude depends on whether the contributions from different paths interfere constructively or destructively. In its simplest form one assumes that the normalization factors N_{cl}, which are determined by the integral over quadratic fluctuations around the classical paths, are all equal,

$$\mathcal{G}(q_f t_f, q_i t_i) = N \sum_{cl} e^{\frac{i}{\hbar} S_{cl}(q_f t_f, q_i t_i)} , \qquad (1.118)$$

and the propagator is then determined by the interference between the phase factors associated with each classical path.

The normalization factor N_{cl} in (1.117) can, in the semiclassical approximation, be determined to have the following form

$$N_{cl} = \sqrt{\frac{1}{2\pi i \hbar} \left| \frac{\partial^2 S_{cl}}{\partial x_f \partial x_i} \right|} (-i)^n , \qquad (1.119)$$

where n is an integer, called the Morse index. It is identical to the number of sign changes of $\frac{\partial^2 S}{\partial x_f \partial x_i}$ when x_f is continuously changed from the initial to the final position. As shown

by the expressions (1.117) and (1.119) the semiclassical propagator is fully determined by the classical action, that is, by the action integral restricted to classical paths.

We do not include a derivation of the expression (1.119) here, but we will check what it gives for the that semiclassical propagator of a free particle in one dimension, and compare this with the exact propagator which we have earlier discussed. For this system the classical action is

$$S_{cl}(x_f, t_f, x_i, t_i) = \frac{1}{2} m \frac{(x_f - x_i)^2}{t_f - t_i} \,. \tag{1.120}$$

This gives

$$\frac{\partial^2 S}{\partial x_f \partial x_i} = -\frac{m}{t_f - t_i} \,, \tag{1.121}$$

which is independent of the initial and final positions of the particle. As a consequence the morse index is $n = 0$, and the semiclassical expression for the propagator is

$$\mathcal{G}(x_f, t_f; x_i, t_i) = \sqrt{\frac{m}{2\pi i \hbar (t_f - t_i)}} \, e^{\frac{i}{2\hbar} m \frac{(x_f - x_i)^2}{t_f - t_i}} \,. \tag{1.122}$$

This expression coincides with the exact expression for the propagatator (1.109) of a free particle in one dimension. The lack of corrections to the semiclassical propagator is clearly related to the quadratic form of the free particle action, which means that there are no higher order corrections to the expansion of the action (1.114) around the classical solution.

1.3.6 The double slit experiment revisited

As a simple example let us apply the semiclassical approximation to the double slit experiment, schematically shown in Fig.1.3, and compare with conditions for constructive interference as derived in an elementary way from interference between partial waves.

We then consider a beam of electrons with sharply defined energy that impinge on a screen with two holes, as shown in the figure. Those of the electrons that pass the holes are registered on a second screen. When the experiment is running, an intensity distribution will gradually build up, and this will show an interference pattern in accordance with the quantum mechanical wave picture of the electrons.

When viewed as a propagating wave, the interference pattern is determined by the relative phase of the partial waves emerging from the two holes. Since the phase of each wave is determined by the distance from the hole measured in wave lengths, the condition for constructive interference at a given point is

$$\Delta L = n\lambda \,, \tag{1.123}$$

where ΔL is the difference in length of the two straight lines from the holes to the point considered. λ is the de Broglie wave length of the electrons and n is an integer.

In the path integral description the conditions for constructive interference also depends on length of paths, but now in a slightly different way. We consider paths beginning at the
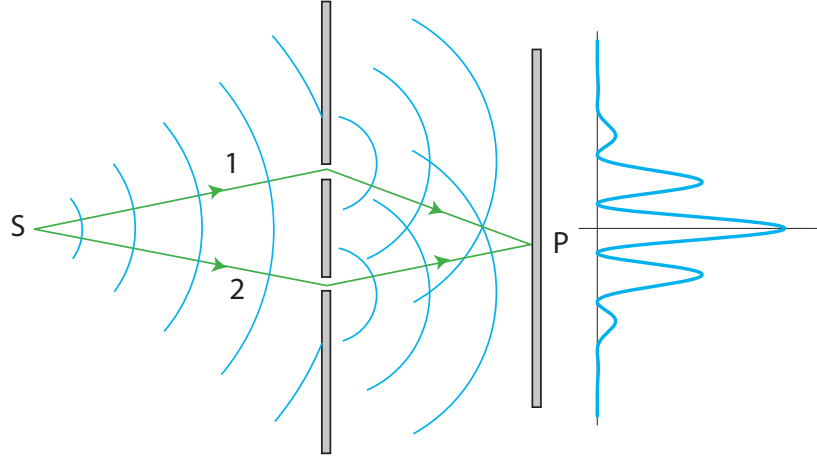
Figure 1.3: The double slit experiment. A beam of electrons is sent from the source S through a screen with two holes. The intensity of electrons is then registered on a second screen behind the first one. The intensity variations can be viewed as due to interference of the partial waves from the two slits. In the path integral approach the intensity at a given point P is determined by the difference in action of two (classical) paths from the source, passing through each of the openings, and and ending both in P.

electron source, with position x_S, at time $t = 0$ and ending at a point P on the second screen, with position x_P at a later time $t = T$. In principle all paths passing through one of the two holes contribute to the path integral, but in the semiclassical approximation we only include contributions from the classical paths and those nearby. In this case there are two classical paths to a given point. The first one describes a free particle moving to the upper hole where it is scattered and sent as a free particle towards the point P. On the second path the electrons is instead scattered in the lower hole and ends up in the same point P.

Note that all paths that contribute correspond to the same time T spent between the initial and final point. This means that if one of the paths is longer, the particle has to move more rapidly on that path. Therefore the situation is slightly different from when the particles have a sharply defined energy. But the result concerning the interference effect is essentially the same.

With L_1 as the length of the upper part and L_2 as the length of the lower one, the corresponding actions are

$$S_1 = \frac{1}{2}mv_1^2 T = \frac{1}{2}m\frac{L_1^2}{T} \ , \quad S_2 = \frac{1}{2}mv_2^2 T = \frac{1}{2}m\frac{L_2^2}{T} \ , \tag{1.124}$$

where we have used the fact that the velocity is constant along the classical paths. In the semiclassical approach the propagator between the the two points S and P can be expressed in terms of actions as

$$\mathcal{G}(x_P\,T, x_S\,0) = N(e^{\frac{i}{\hbar}S_1} + e^{\frac{i}{\hbar}S_2}) = Ne^{\frac{i}{\hbar}S_1}(1 + e^{\frac{i}{\hbar}\Delta S}) \ , \tag{1.125}$$

where $\Delta S = S_2 - S_1$ is the difference between the actions of the two paths. We introduce the average length and the difference in length of the two paths, $\bar{L} = (L_1 + L_2)/2$ and $\Delta L =$

$L_2 - L_1$. The difference in action is then

$$\Delta S = m\frac{\bar{L}}{T}\Delta L = m\bar{v}\Delta L\,, \tag{1.126}$$

where we have introduced the average velocity $\bar{v} = \bar{L}/T$. The condition for constructive interference is

$$\frac{1}{\hbar}\Delta S = 2\pi n \quad \Rightarrow \quad \frac{m\bar{v}}{\hbar}\Delta L = 2\pi n\,, \tag{1.127}$$

and if we assume the following form for the de Broglie relation, $m\bar{v} = 2\pi\hbar/\lambda$, the condition (1.127), derived from the path integral, agrees with (1.123), which is the condition for constructive interference between the partial waves that pass through the two slits.
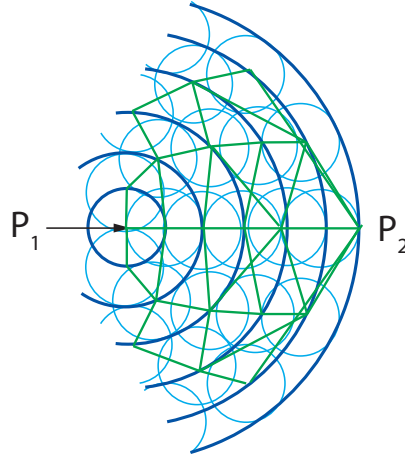
Figure 1.4: Huygens' principle and the sum over paths. The dark blue circles represent the wave fronts of an expanding wave. The smaller light blue circles represent the secondary waves that at each stage of the wave propagation recreates the wave front. The green lines between the centers of the secondary waves represent the different paths from the initial point P_1 to a final point P_2 on the largest wave front.

In this simple example we have demonstrated the close relation between the wave description of quantum mechanics and the path integral description. In fact, even if path integrals are usually associated with quantum physics, they can be seen more generally as a particular way to describe wave propagation. Viewed in this way the path integral formulation is like a geometrical optics approach which in a sense is complimentary to *Huygens' principle*. In Huygens' description wave propagation can be seen as continuous recreation of the wave front through interference between secondary waves that are created at each point of the front. In a ray representation each of these points is the source of rays that are sent in all directions, and if we follow a particular ray through these intermediate points, it traces out one of the paths between the initial and final points that enters into the path integral. This picture of wave propagation is sketched in Fig. 1.4, where a multitude of secondary waves are shown together with the rays between the points where these are emitted.

1.4   The two-level system and the harmonic oscillator

The (one-dimensional) harmonic oscillator is an important system to study, both in the context of classical and quantum physics. One reason is that in many respects it is the simplest dynamical system, and physicists always like to reduce more complicated problems to harmonic oscillators if possible. There are also many physical systems that are well described as harmonic oscillators. All periodic motions close to a stable equilibrium can be viewed as approximate harmonic oscillator motions, and both classically and quantum mechanically the modes of a free field can be viewed as a collection of independent harmonic oscillators.

Even if the harmonic oscillator, in many respects, can be regarded as the simplest system to study, there is a quantum mechanical system that in some respects is even simpler. That the two-level system, where the state space is two-dimensional rather than infinite dimensional as as in the case of the harmonic oscillator. A special realization of the two-level system is the spin-half system already discussed. There are also other physical systems which, to a good approximation, can be regarded as a two level system. Thus, for transitions between atomic levels sometimes only two of the levels will be active in the transitions, and a two-level model may be adequate. Atomic clocks are quantum systems of this type.

As opposed to the harmonic oscillator there is no classical analogue to the quantum two-level system. Classical spin does of course exist, but that corresponds to quantum systems with large spin (on the scale of $\hbar$) rather than to the simple spin-half system. In recent years the interest for two-level systems have increased with the interest for *quantum information*, since the fundamental *qubit* is described by a two-level system.

In this section we will study some of the properties of these two fundamental systems, which are of interest for the further discussion, first the two-level system.

1.4.1   The two-level system

The Hilbert space of this system is two-dimensional. Let us denote by $\{|k\rangle, \; k = 0, 1\}$ the basis vectors of an arbitrarily chosen orthonormal basis. Any state vector can be represented as a two-component, complex matrix Ψ,

$$|\psi\rangle = \sum_{k=1}^{2} \psi_k |k\rangle \;\; \Rightarrow \;\; \Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{1.128}$$

An operator $\hat{A}$ acting in this space has four independent components and can be represented as a 2×2 matrix A,

$$\hat{A}|k\rangle = \sum_{l} A_{lk} |l\rangle \;\; \Rightarrow \;\; A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{1.129}$$

It can be expressed in terms of the unit matrix and the Pauli spin matrices as,

$$A = a_0 \mathbb{1} + \sum_{m} a_m \sigma_m, \tag{1.130}$$

with

$$a_0 \;\; = \;\; \frac{1}{2}(A_{11} + A_{22}), \;\; a_1 = \frac{1}{2}(A_{12} + A_{21}),$$

$$a_2 = \frac{i}{2}(A_{12} - A_{21}), \quad a_3 = \frac{1}{2}(A_{11} - A_{22}). \tag{1.131}$$

When $\hat{A}$ is hermitian, all the coefficients a_k are real.

The standard representation of the Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.132}$$

They define the fundamental commutator relations of the observables of the two level system,

$$[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk} \sigma_k, \tag{1.133}$$

where ϵ_{ijk} is the Levi-Civita symbol, which is totally antisymmetric in the indices ijk and satisfies the identity $\epsilon_{123} = 1$. The Pauli matrices also satisfy the *anti-commutation* relations

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}. \tag{1.134}$$

Let us consider a general, rotated Pauli matrix

$$\sigma_{\mathbf{n}} = \mathbf{n} \cdot \boldsymbol{\sigma} = \begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}, \tag{1.135}$$

with $\mathbf{n}$ as a (three-component) unit vector and (θ, ϕ) as the corresponding polar angles. The two eigenstates of this operator are

$$\Psi_{\mathbf{n}} = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}, \quad \Psi_{-\mathbf{n}} = \begin{pmatrix} \sin\frac{\theta}{2} \\ -e^{i\phi}\cos\frac{\theta}{2} \end{pmatrix}, \tag{1.136}$$

with $\Psi_{\mathbf{n}}$ as the "spin up" state in the $\mathbf{n}$ direction and $\Psi_{-\mathbf{n}}$ as the "spin down" state.

It is interesting to note that an arbitrarily chosen, normalized vector in the state space,

$$\Psi = \begin{pmatrix} u \\ v \end{pmatrix}, \quad |u|^2 + |v|^2 = 1, \tag{1.137}$$

can be written in the form of the spin state $\Psi_{\mathbf{n}}$, with the polar angles of the unit vector determined by $\cos\frac{\theta}{2} = u$ and $e^{i\phi}\sin\frac{\theta}{2} = v$. This means, in particular, that the overall phase of the state vector Ψ is chosen to make u real. As follows from this, the *physically distinct* states of the two-level system can be identified uniquely by the three-dimensional unit vector $\mathbf{n}$, and the continuum of all physical states can therefore be represented by the two-dimensional surface of a sphere. The redundancy in the Hilbert space description is removed in this representation, since the rescaling of a Hilbert space vector by a complex factor, which represents no change in the physical state, does not affect the unit vector $\mathbf{n}$.

Note however that the points on the sphere do not correspond to independent configurations. There are only two *independent* states, for example $|1\rangle$ (the "south pole") and $|2\rangle$ (the "north pole"), corresponding to spin up and spin down in the z-direction. All other points on the sphere correspond to linear superpositions of these two states.

1.4.2   Spin dynamics and magnetic resonance

The spin realization of the two-level system has already been briefly discussed in the context of the Stern-Gerlach experiment. We here consider the spin dynamics in a constant magnetic field in more detail and proceed to show how to solve a time-dependent problem, where the spin is subject to a periodic field.[10]

The basic observables of the spin half-system are the three components of the spin vector

$$\hat{\mathbf{S}} = (\hbar/2)\boldsymbol{\sigma}\,, \tag{1.138}$$

where $\boldsymbol{\sigma}$ is a vector matrix with the three Pauli matrices as components. They correspond to the three space components of the spin vector.

The observable $\hat{\mathbf{S}}$ we will identify as the spin of an electron. The corresponding magnetic moment is given by

$$\boldsymbol{\mu} = \frac{e}{m_e}\hat{\mathbf{S}} \tag{1.139}$$

where e is the electron charge and m_e is the electron mass. The spin Hamiltonian is

$$\hat{H}_0 = -\frac{e}{m_e}\mathbf{B}_0 \cdot \hat{\mathbf{S}}. \tag{1.140}$$

with $\mathbf{B}_0$ as a constant external magnetic field. The Heisenberg equation of motion for the spin, which follows from (1.76) is

$$\frac{d}{dt}\hat{\mathbf{S}}_H = \omega_0\,\mathbf{n} \times \hat{\mathbf{S}}_H, \tag{1.141}$$

where

$$\omega_0 = -\frac{eB_0}{m_e}. \tag{1.142}$$

and $\mathbf{n}$ is a unit vector $\mathbf{B}_0 = B_0\,\mathbf{n}$.[11] Eq.(1.141) has exactly the same form as the classical spin precession equation with ω_0 as the precession frequency. A natural interpretation is then that the quantum spin precesses in the magnetic field in the same way as the classical spin. However, as our discussion of the Stern-Gerlach experiment has shown, the non-commutativity of the different components of $\hat{\mathbf{S}}$ makes the quantum spin variable qualitatively different from a classical spin.

The expression for the Hamiltonian (1.140) can be simplified by the choice of coordinate axes. If we choose the (1,2,3) components of the Pauli spin matrices to correspond to the (x,y,z) directions in space, and $e\mathbf{B}$ to point in the positive z-direction, the Hamiltonian gets the form

$$\begin{aligned}\hat{H}_0 &= -\frac{eB_0}{m_e}\,\hat{S}_z \\ &= \frac{1}{2}\hbar\omega_0\,\sigma_z.\end{aligned} \tag{1.143}$$

---

[10]The effect of a magnetic field on the electron spin, with its corresponding changes in the atomic energy levels is referred to as the *Zeeman effect*.

[11]ω_0 may be chosen to be positive by choosing eB_0 negative. For the electron, with $e < 0$ this means choosing $\mathbf{n}$ in the direction of $\mathbf{B}_0$, so that B_0 is positive. For a particle with positive charge e $\mathbf{n}$ is chosen in the *opposite* direction of $\mathbf{B}_0$, so that B_0 is negative.

The corresponding time evolution operator is

$$\hat{\mathcal{U}}_0(t) = e^{-\frac{i}{\hbar}\hat{H}_0 t} = e^{-\frac{i}{2}\omega_0 \sigma_z t}, \tag{1.144}$$

and it shows explicitly the spin rotation around the z-axis.

We now proceed to examine a two-state problem with a *time-dependent* Hamiltonian, which is directly relevant for the effect of so-called *spin magnetic resonance*. The system is the spin variable of a (bound) electron in a magnetic field $\mathbf{B}$ which now, in addition to a constant part has an oscillating part,

$$\mathbf{B} = B_0\,\mathbf{k} + B_1(\cos\omega t\,\mathbf{i} + \sin\omega t\mathbf{j})\,. \tag{1.145}$$

Both B_0 and B_1 are constants. The oscillating field is due to a circularly polarized electromagnetic field interacting with the electron.

The time variation of $\mathbf{B}$ gives rise to a time-dependent Hamiltonian. Usually a time-dependent problem like this can only be solved numerically or within some approximation scheme. But the present problem can be solved exactly. We show this by rewriting the Hamiltonian in the following form

$$
\begin{aligned}
\hat{H} &= -\frac{e}{m_e}[B_0\,\mathbf{k} + B_1(\cos\omega t\,\mathbf{i} + \sin\omega t\mathbf{j})]\cdot\hat{\mathbf{S}} \\
&= -\frac{e}{m_e}[B_0\hat{S}_z + B_1(\cos\omega t\,\hat{S}_x + \sin\omega t\,\hat{S}_y)] \\
&= -\frac{e}{m_e}\,e^{-\frac{i}{\hbar}\omega t\hat{S}_z}[B_0\hat{S}_z + B_1\hat{S}_x]e^{\frac{i}{\hbar}\omega t\hat{S}_z}\,,
\end{aligned} \tag{1.146}
$$

where the last expression follows from the commutation relations between the components of $\hat{\mathbf{S}}$. This expression for the Hamiltonian implies that it can be transformed to the time-independent form (1.140) by a unitary transformation.

In order to show this we perform a *time-dependent* transformation

$$|\psi(t)\rangle \to |\psi(t)\rangle_T = \hat{T}(t)\,|\psi(t)\rangle\,, \tag{1.147}$$

with $\hat{T}(t)$ as the unitary operator

$$\hat{T}(t) = e^{\frac{i}{\hbar}\omega t\hat{S}_z} \tag{1.148}$$

The transformed state vector satisfies the modified Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle_T = [\hat{T}(t)\hat{H}\hat{T}(t)^\dagger + i\hbar\frac{d\hat{T}}{dt}\hat{T}(t)^\dagger]|\psi(t)\rangle_T \tag{1.149}$$

where the right-hand side defines the Hamiltonian in the transformed picture

$$
\begin{aligned}
\hat{H}_T &= e^{\frac{i}{\hbar}\omega t\hat{S}_z}\hat{H}e^{-\frac{i}{\hbar}\omega t\hat{S}_z} - \omega\hat{S}_z \\
&= -\frac{e}{m_e}[B_0\hat{S}_z + B_1\hat{S}_x] - \omega\hat{S}_z \\
&= \frac{1}{2}\hbar[(\omega_0 - \omega)\sigma_z + \omega_1\sigma_x]\,,
\end{aligned} \tag{1.150}
$$

with $\omega_0 = -eB_0/m_e$ and $\omega_1 = -eB_1/m_e$.

It is of interest to note that the new Hamiltonian $\hat{H}_T$ is not simply identical to the unitary transformation $\hat{T}(t)\hat{H}\hat{T}(t)^\dagger$, but there is also a contribution from the time derivative of $\hat{T}(t)$. This is a consequence of the fact that the Hamiltonian is primarily defined as the operator that generates the time evolution of the state vector. For time-dependent transformations in general this means that the original and transformed Hamiltonians will not represent the same physical observable.

The transformation above can be interpreted as changing to a rotating reference frame, where the magnetic field looks time-independent, and where the Hamiltonian is therefore also independent of time. To re-write the Hamiltonian in a more standard form we introduce the parameters

$$\Omega = \sqrt{(\omega_0 - \omega)^2 + \omega_1^2} \tag{1.151}$$

and

$$\cos\theta = \frac{\omega_0 - \omega}{\sqrt{(\omega_0 - \omega)^2 + \omega_1^2}} \ , \quad \sin\theta = \frac{\omega_1}{\sqrt{(\omega_0 - \omega)^2 + \omega_1^2}} \ . \tag{1.152}$$

This gives for the Hamiltonian

$$\hat{H}_T = \frac{1}{2}\hbar\Omega\left(\cos\theta\sigma_z + \sin\theta\sigma_x\right), \tag{1.153}$$

which is the same expression as for a spin Hamiltonian in a constant magnetic field (see (1.143)), except that the magnetic field is now rotated by an angle θ relative to the z-axis. The time evolution in the transformed frame is then simply

$$\hat{\mathcal{U}}_T(t) = e^{-\frac{i}{2}\Omega t\,(\cos\theta\sigma_z + \sin\theta\sigma_x)}. \tag{1.154}$$

The time evolution operator, in the original frame, can now be found by applying the time dependent transformation in reverse,

$$\begin{aligned}\hat{\mathcal{U}}(t) &= \hat{T}(t)^\dagger\hat{\mathcal{U}}_T(t)\,\hat{T}(0) \\ &= e^{-\frac{i}{2}\omega t\sigma_z}e^{-\frac{i}{2}\Omega t(\cos\theta\sigma_z + \sin\theta\sigma_x)}, \end{aligned} \tag{1.155}$$

which in matrix form is

$$\begin{aligned}\hat{\mathcal{U}}(t) &= \begin{pmatrix} e^{-\frac{i}{2}\omega t} & 0 \\ 0 & e^{\frac{i}{2}\omega t} \end{pmatrix} \begin{pmatrix} \cos\frac{\Omega t}{2} - i\cos\theta\sin\frac{\Omega t}{2} & -i\sin\theta\sin\frac{\Omega t}{2} \\ -i\sin\theta\sin\frac{\Omega t}{2} & \cos\frac{\Omega t}{2} + i\cos\theta\sin\frac{\Omega t}{2} \end{pmatrix} \\ &= \begin{pmatrix} (\cos\frac{\Omega t}{2} - i\cos\theta\sin\frac{\Omega t}{2})e^{-\frac{i}{2}\omega t} & -i\sin\theta\sin\frac{\Omega t}{2}e^{-\frac{i}{2}\omega t} \\ -i\sin\theta\sin\frac{\Omega t}{2}e^{\frac{i}{2}\omega t} & (\cos\frac{\Omega t}{2} + i\cos\theta\sin\frac{\Omega t}{2})e^{\frac{i}{2}\omega t} \end{pmatrix}, \tag{1.156}\end{aligned}$$

The above result shows that the time-varying magnetic field $\mathbf{B}_1$ will induce oscillations in the spin between the two eigenstates $|0\rangle$ and $|1\rangle$ of the time-independent spin Hamiltonian $\hat{H}_0 = -(\hbar eB_0/2m_e)\sigma_z$,

$$|\psi(t)\rangle = c_0(t)|0\rangle + c_1(t)|1\rangle\,. \tag{1.157}$$

Let us choose as initial conditions, $c_0(0) = 1, c_1(0) = 0$, which means that the spin starts in the ground state of $\hat{H}_0$. This gives,

$$
\begin{aligned}
c_0(t) &= (\cos\frac{\Omega t}{2} + i\cos\theta\sin\frac{\Omega t}{2})e^{\frac{i}{2}\omega t} \\
c_1(t) &= -i\sin\theta\sin\frac{\Omega t}{2}e^{-\frac{i}{2}\omega t} .
\end{aligned}
\tag{1.158}
$$

The time-dependent occupation probability of the upper level is
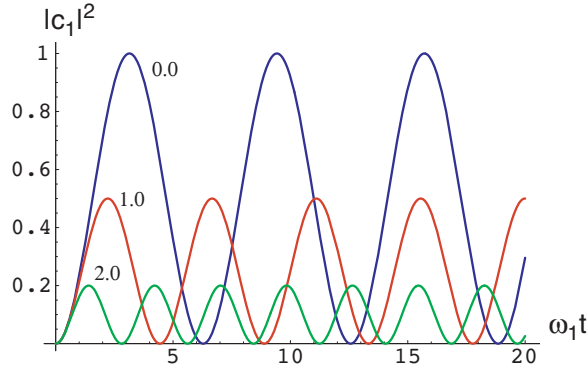
Figure 1.5: Rabi oscillations in a two-level system, caused by an oscillating field. The oscillation in the occupation of the upper level is shown as a function of time. The oscillation is shown for three different values of the detuning parameter $(\omega_0 - \omega)/\omega_1$.

$$
|c_1(t)|^2 = \sin^2\theta\sin^2\frac{\Omega t}{2} .
\tag{1.159}
$$

The amplitude of the oscillations is

$$
\sin^2\theta = \frac{\omega_1^2}{(\omega_0 - \omega)^2 + \omega_1^2} .
\tag{1.160}
$$

The expression shows a *resonance effect*, when the frequency of the oscillating field matches the energy difference between the two levels, $\omega = \omega_0$. For this frequency the maximum value of $|c_1(t)|^2$ is 1, which means that there is a complete transitions between the two levels $|0\rangle$ and $|1\rangle$ during the oscillations. The frequency of the oscillations is

$$
\Omega = \sqrt{(\omega_0 - \omega)^2 + \omega_1^2} ,
\tag{1.161}
$$

with a resonance value

$$
\Omega_{res} = |\omega_1| = \frac{|eB_1|}{m_e} .
\tag{1.162}
$$

The frequency thus depends on the amplitude of the oscillating field. In Fig.1.5 the oscillations are shown as a function of time for different values of the frequency ω.

The oscillations between the two spin levels caused by coupling to the electromagnetic field, are called *Rabi oscillations* and the frequency of the oscillations is referred to as the *Rabi frequency*. The effect studied here has important applications in the context of *electron spin resonance* (ESR) and *nuclear magnetic resonance* (NMR). In particular has NMR become important in the form of magnetic resonance imaging (MRI), which is now widely used as a scanning technique to obtain high quality images of the human body.

1.4.3   The Jaynes-Cummings model

We shall in this section study a model, which describes a two-level system interacting with a harmonic oscillator. It is called the Jaynes-Cummings model, and it has an important application in the field of *cavity QED*. The physical situation it describes is an atom interacting with the electromagnetic field within in a reflecting cavity. The assumption is that one of the atomic transitions is resonant with one of the modes of the electromagnetic field, and the simplification of the model is that only the two atomic levels involved in the transition is included in the description, and all the electromagnetic modes are neglected except the one which is strongly coupled to the resonant transition. Thus, the two-level system represents the atom and the harmonic oscillator represents the quantized field mode, in a similar way as discussed in Sect. 1.2. The model studied here is closely related to the one previously examined in Sect. 1.3.2, where a spin half system was coupled to a rotating magnetic field. The difference is that the electromagnetic field now is treated quantum mechanically. As we shall see, the *Rabi oscillations* observed in the previous case now is present in the form of energy oscillations between the atom and the cavity mode.

The Hamiltonian of the model has the following form

$$\hat{H} = \frac{1}{2}\hbar\omega_0\sigma_z + \hbar\omega a^\dagger a + i\hbar\lambda(a^\dagger\sigma_- - a\sigma_+) \equiv \hat{H}_0 + \hat{H}_1 \, , \qquad (1.163)$$

where the Pauli matrices act between the two atomic levels and the operators $\hat{a}$ and $\hat{a}^\dagger$ are photon annihilation and creation operators, $\hbar\omega_0$ is the energy difference between the two atomic levels, and $\hbar\omega$ is the photon energy. (The zero point of the energy has been adjusted to absorb the ground state energy of the harmonic oscillator and to place the energies of the two-level system symmetrically about $E = 0$.) The parameter λ gives the strength of the interaction and we use the notation $\sigma_\pm = (1/2)(\sigma_x \pm i\sigma_y)$ for the raising and lowering operators between the two atomic levels.

The Hamiltonian, in the above expression, has been separated in the non-interaction part $\hat{H}_0$ and the interaction $\hat{H}_1$, which can be interpreted as representing either a magnetic or electric dipole coupling between the atom and the electromagnetic field. This type of interaction we shall discuss further in Chapt. 4. The expression for the interaction has been simplified in a form referred to as the *rotating wave approximation*, by including only resonant terms. This approximation, and the other simplifications which are based on resonance between the two-level system and the harmonic oscillator, are valid when the two frequencies ω_0 and ω are sufficiently close in value.

The eigenvalue equation for $\hat{H}_0$ can be written in the following form

$$\hat{H}_0|\pm, n\rangle = \hbar(n\omega \pm \frac{1}{2}\omega_0)|\pm, n\rangle \, , \qquad (1.164)$$

with $\pm$ referring to the upper and lower atomic levels, and n referring to the number of photon being present in the cavity mode. Except for the ground state, all the eigenstates of the non-interacting Hamiltonian are doubly degenerate, with equal energy for $|-, n+1\rangle$ and $|+, n\rangle$, for $n \geq 0$. As follows from (1.163) interaction H_1 splits the degeneracy, but only couples states within such a pair. The ground state $|-, 0\rangle$ is unaffected by $\hat{H}_1$, and therefore remains the ground state of the full Hamiltonian.

In the $|\pm, n\rangle$ basis $\hat{H}$ thus has a simple block-diagonal form, where each block, except for the ground state, is a 2×2 matrix. This makes it easy to diagonalize the Hamiltonian, and we focus on the subspace spanned by $|-, n+1\rangle$ and $|+, n\rangle$, $n \geq 0$. The Hamiltonian restricted to this subspace we denote by $\hat{H}_n$. It has the matrix form

$$\hat{H}_n = \frac{1}{2}\hbar \begin{pmatrix} \Delta & -ig_n \\ ig_n & -\Delta \end{pmatrix} + \epsilon_n \mathbb{1}, \tag{1.165}$$

with $\Delta = \omega_0 - \omega$, $g_n = 2\lambda\sqrt{n}$ and $\epsilon_n = (n - \frac{1}{2})\hbar\omega$. The upper left matrix position has then been chosen to correspond to the atom being in the upper evergy level (of $\hat{H}_0$) and the lower right matrix position to correspond the atom being in the lower level. To simplify expressions it may be convenient to write the matrix elements in terms of new parameters Ω_n and θ_n defined by

$$\Omega = \sqrt{\Delta^2 + g_n^2}, \quad \cos\theta_n = \frac{\Delta}{\Omega_n}, \quad \sin\theta_n = \frac{g_n}{\Omega_n}. \tag{1.166}$$

This gives

$$\hat{H}_n = \frac{1}{2}\hbar\Omega_n \begin{pmatrix} \cos\theta_n & -i\sin\theta_n \\ i\sin\theta_n & -\cos\theta_n \end{pmatrix} + \epsilon_n \mathbb{1}. \tag{1.167}$$

We note that the matrix in the above expression has the form of a rotated Pauli matrix $\sigma_{\mathbf{n}}$, as written in (1.135), if we make the following identifications for the polar coordinates of $\mathbf{n}$, $\theta = \theta_n$ and $\phi = -\pi/2$. The matrix thus has eigenvalues ± 1 and eigenstates

$$\psi_n^+ = \begin{pmatrix} \cos\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} \end{pmatrix}, \quad \psi_n^- = \begin{pmatrix} -i\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}. \tag{1.168}$$

The corresponding energies are

$$\begin{aligned} E_n^\pm &= \epsilon_n \pm \frac{1}{2}\hbar\Omega_n \\ &= (n - \frac{1}{2})\hbar\omega \pm \sqrt{(\omega_0 - \omega)^2 + 4n\lambda^2}. \end{aligned} \tag{1.169}$$

A general time dependent state can be written as

$$\psi_n(t) = d_n^+(t)\,\psi_n^+ + d_n^-(t)\,\psi_n^- = \begin{pmatrix} c_{n1}(t) \\ c_{n2}(t) \end{pmatrix}, \tag{1.170}$$

with the coefficients related by

$$c_{n1} = i(d_n^+ \cos\frac{\theta_n}{2} + d_n^- \sin\frac{\theta_n}{2})$$

$$c_{n2} = -d_n^+ \sin\frac{\theta_n}{2} + d_n^- \cos\frac{\theta_n}{2}. \tag{1.171}$$

We assume, for simplicity, the following initial condition

$$c_{n1}(0) = 0, \quad c_{n2}(0) = 1, \tag{1.172}$$

which means that the atom starts in the lower level, with $n + 1$ photons present in the cavity mode. Expressed in terms of the d coefficients this is

$$d_n^+(0) = -\sin\frac{\theta_n}{2}, \quad d_n^-(0) = \cos\frac{\theta_n}{2}, \tag{1.173}$$

which for the time dependent coefficients gives

$$d_n^+(t) = e^{-\frac{i}{\hbar}E_+ t}d_n^+(0) = -\sin\frac{\theta_n}{2}e^{-\frac{i}{2}\Omega_n t}e^{-\frac{i}{\hbar}\epsilon_n t}$$

$$d_n^-(t) = e^{-\frac{i}{\hbar}E_- t}d_n^-(0) = \cos\frac{\theta_n}{2}e^{\frac{i}{2}\Omega_n t}e^{-\frac{i}{\hbar}\epsilon_n t}. \tag{1.174}$$

For the c coefficients this implies

$$c_{n1}(t) = -e^{-\frac{i}{\hbar}\epsilon_n t}\sin\theta_n\,\sin\frac{\Omega_n}{2}t$$

$$c_{n2}(t) = -e^{-\frac{i}{\hbar}\epsilon_n t}(\cos\frac{\Omega_n}{2}t + i\cos\theta_n\,\sin\frac{\Omega_n}{2}t). \tag{1.175}$$

The time dependent probabilities for the atom to be in the upper and lower levels thus are

$$|c_{n1}|^2 = \sin^2\theta_n\sin^2\frac{\Omega_n}{2}t, \quad |c_{n2}|^2 = 1 - |c_{n1}|^2. \tag{1.176}$$

The occupation probabilities oscillates between the upper and lower atomic levels, and the photon number at the same time oscillates between n and $n + 1$. The maximum oscillations happens when $\cos\theta_n = 0$, which corresponds to $\omega = \omega_0$. In this case the probabilities oscillate between 0 and 1.

   This situation should be compared with the oscillating occupation probabilities in Sect. 1.3.2, where a two-level system is interacting with a classical oscillating electromagnetic field. Thus, the expressions found here are very similar to the ones found in the previous case, and the correspondance between the photon number here and the strength of the magnetic feld in the other case is given by

$$g_n \leftrightarrow \omega_1 \Rightarrow 2\,\lambda\sqrt{n} \leftrightarrow -\frac{eB_1}{m_e c}. \tag{1.177}$$

The magnetic field, in this correspondance, is proportional to the square root of the photon number. We note that when n is large there is essentially no difference between n and $n + 1$, and the variation in the field strengt then is very small when the atom oscillates between the upper and lower level. The situation discussed in Sect. 1.3.2, with a classical oscillating field, can therefore be viewed as the limit of the situation considered here with the photon number being very large, $n >> 1$. In the opposite case, with $n \approx 1$, however, both the atom and the electromagnetic field have to be treated quantum mechanically, with the states of the two being strongly coupled at resonance, $\omega = \omega_0$.

1.4.4 Harmonic oscillator and coherent states

The Hamiltonian of a one-dimensional (quantum) harmonic oscillator we write in the standard way as

$$\hat{H} = \frac{1}{2m}(\hat{p}^2 + m^2\omega^2\hat{x}^2),$$ (1.178)

which means that it is realized as the energy observable of a particle with mass m in the oscillator potential $\hat{V} = (1/2)m\omega^2\hat{x}^2$.

The most elegant way to solve the energy eigenvalue problem is the algebraic method shown in all introductory text books on quantum mechanics. It is based on the closed commutator algebra formed by the operators $\hat{x}, \hat{p}$ and $\hat{H}$, [12]

$$[\hat{x}, \hat{p}] = i\hbar, \quad \left[\hat{x}, \hat{H}\right] = i\frac{\hbar}{m}\hat{p}, \quad \left[\hat{p}, \hat{H}\right] = -i\hbar m\omega^2\hat{x}.$$ (1.179)

This is conveniently reformulated in terms of the *raising* and *lowering* operators

$$\hat{a} = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega\hat{x} + i\hat{p}), \quad \hat{a}^\dagger = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega\hat{x} - i\hat{p}),$$ (1.180)

which gives

$$\hat{H} = \frac{1}{2}\hbar\omega(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})).$$ (1.181)

The commutator algebra is now reformulated as

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1, \quad \left[\hat{H}, \hat{a}\right] = -\hat{a}\,\hbar\omega, \quad \left[\hat{H}, \hat{a}^\dagger\right] = \hat{a}^\dagger\,\hbar\omega.$$ (1.182)

We briefly summarize the construction of energy eigenstates. Since $\hat{H}$ is a positive definite operator, there is a lowest energy state, which is annihilated by $\hat{a}$,

$$\hat{a}|0\rangle = 0.$$ (1.183)

The repeated application of $\hat{a}^\dagger$ on this state generates a series of states $|n\rangle, n = 0, 1, 2, ...,$ which according to the commutator with $\hat{H}$ all are energy eigenstates. These states form a complete set of states in the Hilbert space. The explicit action of the operators on these states follows from the algebraic relations and are here summarized as

$$\begin{aligned}
\hat{a}|n\rangle &= \sqrt{n}|n-1\rangle, \\
\hat{a}^\dagger|n-1\rangle &= \sqrt{n}|n\rangle, \\
\hat{H}|n\rangle &= \hbar\omega(n + \frac{1}{2})|n\rangle.
\end{aligned}$$ (1.184)

---

[12]The commutators of the observables of a quantum system defines a *Lie algebra*. When the Hamiltonian belongs to a finite-dimensional Lie algebra, the general methods for finding *representations* of Lie algebras can be used to construct the eigenstates and find the eigenvalues of the Hamiltonian. The algebraic method of solving the eigenvalue problem of a harmonic oscillator is a particular example of this more general algebraic approach.

Expansion of state vectors $|\psi\rangle$ in the orthonormal basis $\{|n\rangle\}$ defines a representation which we shall refer to as the energy (or n-) representation. The transition between this representation and the standard coordinate (or x-) representation is defined by the matrix elements

$$\langle x|n\rangle \equiv \psi_n(x)\,. \tag{1.185}$$

For given n this corresponds to the energy eigenfunction in the coordinate representation. We refer to standard treatments of the harmonic oscillator, where these eigenstates are expressed in terms of Hermite polynomials.

After this brief reminder on standard treatments of the harmonic oscillator, we turn to the main theme of this section, which is a discussion of the so-called *coherent states*. These are defined as the eigenstates of the annihilation operator $\hat{a}$,

$$\hat{a}|z\rangle = z|z\rangle\,. \tag{1.186}$$

What is unusual about this definition of states is that $\hat{a}$ is not a hermitian operator (and therefore not an observable in the usual sense). However, the states $|z\rangle$ defined in this way do form a complete set, in fact an over-complete set, and they define a new representation, the *coherent state representation* with many useful properties.

Note that, since $\hat{a}$ is non-hermitian, the eigenvalues z will in general be complex rather than real. Based on the relation between $\hat{a}$ and $\hat{x}$ and $\hat{p}$ it is useful to write z as,

$$z = \frac{1}{\sqrt{2m\hbar\omega}}(m\omega\, x_c + ip_c)\,. \tag{1.187}$$

This indicates that z can be interpreted as a complex *phase space* variable, with $\mathrm{Re}\,z$ proportional to x and $\mathrm{Im}\,z$ proportional to p. However, since the symbol x has already been used for the eigenvalues of $\hat{x}$ and p for the eigenvalues of $\hat{p}$, we have introduced x_c and p_c for the two phase-space components of z. Such a distinction is necessary, since $|z\rangle$ is not an eigenstate for $\hat{x}$ and $\hat{p}$, although the probability distributions over the eigenvalues will be strongly peaked around x_c and p_c, respectively. For the expectation values of $\hat{a}$ and $\hat{a}^\dagger$ we have have

$$\langle z|\hat{a}|z\rangle = z\,, \quad \langle z|\hat{a}^\dagger|z\rangle = z^*\,, \tag{1.188}$$

which for the expectation values of position and momentum gives

$$\langle x\rangle_z = x_c\,, \quad \langle p\rangle_z = p_c\,. \tag{1.189}$$

To study the coherent states further we first focus on the ground state of the harmonic oscillator. This is a particular coherent state, with $z = 0$, as follows from (1.183). For the ground state we have the following expectation values for x and p,

$$\begin{aligned} \langle x\rangle_0 &= \langle p\rangle_0 = 0\,, \\ \left\langle x^2\right\rangle_0 &= \frac{\hbar}{2m\omega}\langle 0|(\hat{a}+\hat{a}^\dagger)(\hat{a}+\hat{a}^\dagger)|0\rangle = \frac{\hbar}{2m\omega}\,, \\ \left\langle p^2\right\rangle_0 &= -\frac{m\omega\hbar}{2}\langle 0|(\hat{a}-\hat{a}^\dagger)(\hat{a}-\hat{a}^\dagger)|0\rangle = \frac{m\omega\hbar}{2}\,. \end{aligned} \tag{1.190}$$

From this follows that the uncertainties in x and p for the ground state satisfy

$$\Delta x_0^2 \, \Delta p_0^2 = \frac{\hbar}{2m\omega} \frac{m\omega\hbar}{2} = \frac{\hbar^2}{4} \, . \tag{1.191}$$

This is the minimum value for the product allowed by Heisenberg's uncertainty principle. Thus, the ground state is a *minimum uncertainty* state. A similar calculation for the excited energy states shows that they are not,

$$\Delta x_n^2 \Delta p_n^2 = \frac{\hbar^2}{4}(2n+1)^2 \, . \tag{1.192}$$

We shall proceed to show that all *coherent states are minimum uncertainty states*. To show this we introduce the unitary operator

$$
\begin{aligned}
\hat{\mathcal{D}}(z) &= e^{z\hat{a}^\dagger - z^*\hat{a}} \\
&= e^{\frac{i}{\hbar}(p_c\hat{x} - x_c\hat{p})} \, ,
\end{aligned}
\tag{1.193}
$$

where z is a complex number, related to x_c and p_c as in (1.187). It is the quantum version of a *displacement operator in phase space*. It transforms $\hat{a}$ and $\hat{a}^\dagger$ as

$$\hat{\mathcal{D}}(z)^\dagger \hat{a} \hat{\mathcal{D}}(z) = \hat{a} + z \, , \quad \hat{\mathcal{D}}(z)^\dagger \hat{a}^\dagger \hat{\mathcal{D}}(z) = \hat{a}^\dagger + z^* \, , \tag{1.194}$$

which is shown by use of the operator identity

$$e^{\hat{B}} \hat{A} e^{-\hat{B}} = A + \left[\hat{B}, \hat{A}\right] + \frac{1}{2}\left[\hat{B}, \left[\hat{B}, \hat{A}\right]\right] + \dots . \tag{1.195}$$

It acts on x and p in the following way

$$\hat{\mathcal{D}}(z)^\dagger \, \hat{x} \, \hat{\mathcal{D}}(z) = \hat{x} + x_c \, , \quad \hat{\mathcal{D}}(z)^\dagger \, \hat{p} \, \hat{\mathcal{D}}(z) = \hat{p} + p_c \, , \tag{1.196}$$

which explains the interpretation of $\hat{\mathcal{D}}$ as a displacement operator in phase space.

With the displacement operator $\hat{\mathcal{D}}$ acting on the ground state a continuum of new states can be generated,

$$
\begin{aligned}
|z\rangle &\equiv \hat{\mathcal{D}}(z)|0\rangle \\
&= e^{-z^*z} e^{z\hat{a}^\dagger}|0\rangle \, ,
\end{aligned}
\tag{1.197}
$$

and it follows directly that these new states are coherent states with eigenvalues z for $\hat{a}$,

$$
\begin{aligned}
\hat{a}\hat{\mathcal{D}}(z)|0\rangle &= \hat{\mathcal{D}}(z)\hat{\mathcal{D}}(z)^\dagger \hat{a}\hat{\mathcal{D}}(z)|0\rangle & (1.198) \\
&= \hat{\mathcal{D}}(z)(\hat{a} + z)|0\rangle & (1.199) \\
&= z\hat{\mathcal{D}}(z)|0\rangle \, . & (1.200)
\end{aligned}
$$

Since $\hat{\mathcal{D}}(z)$ simply adds a constant to observables $\hat{x}$ and $\hat{p}$, which means that $\hat{x} - \langle x \rangle$ and $\hat{p} - \langle p \rangle$ are unchanged by the displacements, the shifted state $|z\rangle$ has the same uncertainty in x and p as the ground state. Thus, all coherent states $|z\rangle$ are minimum uncertainty states.

The above discussion also shows that the parameter z can take any value in the complex plane, which means that there is a coherent state associated with each point in the (two-dimensional) phase space. Furthermore, since the coherent states are optimally focussed in x and p they can be viewed as the quantum states that are closest to the classical states, which are defined by sharply defined positions in phase space.

Coherent states in the coordinate representation
The coordinate representation of the coherent states are defined by

$$\psi_z(x) = \langle x|z\rangle\,,\tag{1.201}$$

where the *bra* corresponds to a position eigenstate and the *ket* to a coherent state. The $z = 0$ state is the ground state of the harmonic oscillator and is known to have the gaussian form

$$\psi_0(x) = (\frac{m\omega}{\pi\hbar})^{\frac{1}{4}}e^{-\frac{m\omega}{2\hbar}x^2}\,.\tag{1.202}$$

This expression can readily be generalized to arbitrary coherent states, since they satisfy a linear differential equation

$$\frac{1}{\sqrt{2m\hbar\omega}}(m\omega x + \hbar\frac{d}{dx})\psi_z(x) = z\psi_z(x)\,,\tag{1.203}$$

or,

$$\frac{d}{dx}\psi_z(x) = (-\frac{m\omega}{\hbar}x + \sqrt{\frac{2m\omega}{\hbar}}\,z)\,\psi_z(x)\,.\tag{1.204}$$

This differential equation follows directly from the definition (1.186) of the coherent states. The equation has the solution

$$\psi_z(x) = N_z e^{-(\frac{m\omega}{2\hbar}x^2 - \sqrt{\frac{2m\omega}{\hbar}}\,zx)}\,,\tag{1.205}$$

where N_z is a z-dependent normalization factor. We rewrite it in the form

$$\psi_z(x) = N_z' e^{-(\frac{m\omega}{2\hbar}(x-x_c)^2 - \frac{i}{\hbar}xp_c)}\,,\tag{1.206}$$

with N_z' as a new normalization factor, and with x_c and p_c as the phase space coordinates corresponding to the real and imaginary parts of z, as given by (1.187). Normalization of the wave function determines N_z' to be, up to a phase factor, the same as the prefactor of the ground state wave function (1.202), $|N_z'| = (m\omega/\pi\hbar)^{1/4}$. The complex phase of N_z' is undetermined by the normalization, and N_z' could simply be chosen to be real. However, implicitly the phase has already been fixed by the definition (1.197). To show this we shall find the expression for $\psi_z(x)$ in alternative, more direct way. We have

$$\begin{aligned}\psi_z(x) &= \langle x|\hat{\mathcal{D}}(z)|0\rangle\\ &= \langle x|e^{\frac{i}{\hbar}(p_c\hat{x}-x_c\hat{p})}|0\rangle\\ &= e^{-\frac{i}{2\hbar}x_cp_c}\langle x|e^{\frac{i}{\hbar}p_c\hat{x}}e^{-\frac{i}{\hbar}x_c\hat{p}}|0\rangle\\ &= e^{-\frac{i}{2\hbar}x_cp_c}e^{\frac{i}{\hbar}p_cx}e^{-x_c\frac{d}{dx}}\psi_0(x)\,,\end{aligned}\tag{1.207}$$

where we have used that in the coordinate representation we have the following relations, $\psi_0(x) = \langle x|0\rangle$ and $\exp(-\frac{i}{\hbar}x_c\hat{p}) = \exp(-x_c\frac{d}{dx})$. The latter is a translation operator in x-space, which gives, when applied to the ground state,

$$e^{-x_c\frac{d}{dx}}\psi_0(x) = \sum_{n=0}^{\infty}\frac{(-x_c)^n}{n!}\frac{d^n\psi_0}{dx^n}(x) = \psi_0(x - x_c)\,. \tag{1.208}$$

With the ground state wave function $\psi_0(x)$ given by (1.202), the full expression for the coherent state in the x-representation is then

$$\psi_z(x) = (\frac{m\omega}{\pi\hbar})^{\frac{1}{4}}e^{-(\frac{m\omega}{2\hbar}(x-x_c)^2 - \frac{i}{\hbar}xp_c + \frac{i}{2\hbar}x_cp_c)}\,. \tag{1.209}$$

This expression agrees with (1.206) and also gives the expression for the x-independent phase factor included in the normalization factor.

Time evolution of coherent states
In the Heisenberg picture the time evolution of the creation and annihilation operators are

$$\begin{aligned}\hat{a}^{\dagger}(t) &= \hat{\mathcal{U}}(t,0)^{\dagger}\,\hat{a}^{\dagger}\,\hat{\mathcal{U}}(t,0)\\ &= e^{i\omega t\,\hat{a}^{\dagger}\hat{a}}\hat{a}^{\dagger}e^{-i\omega t\,\hat{a}^{\dagger}\hat{a}}\\ &= e^{i\omega t}\hat{a}^{\dagger}\,,\end{aligned} \tag{1.210}$$

and

$$\begin{aligned}\hat{a}(t) &= \hat{\mathcal{U}}(t,0)^{\dagger}\,\hat{a}\,\hat{\mathcal{U}}(t,0)\\ &= e^{-i\omega t}\hat{a}\,.\end{aligned} \tag{1.211}$$

From the last one follows,

$$\hat{a}\,\hat{\mathcal{U}}(t,0)|z\rangle = e^{-i\omega t}\hat{\mathcal{U}}(t,0)\,\hat{a}|z\rangle = e^{-i\omega t}z\,\hat{\mathcal{U}}(t,0)|z\rangle\,, \tag{1.212}$$

which gives the time evolution

$$\hat{\mathcal{U}}(t,0)|z\rangle = e^{i\alpha(t)}|e^{-i\omega t}z\rangle\,, \tag{1.213}$$

where $\alpha(t)$ is an undetermined complex phase. The equation shows that a coherent state continues to be a coherent state during the time evolution. This means that it keeps its property of maximal localization in the phase space variables. The motion is given by

$$z(t) = e^{-i\omega t}\,z(0)\,, \tag{1.214}$$

which means for the phase space variables

$$\begin{aligned}x_c(t) &= \cos\omega t\,x_c(0) + \sin\omega t\,\frac{1}{m\omega}\,p_c(0)\,,\\ p_c(t) &= \cos\omega t\,p_c(0) - \sin\omega t\,m\omega\,x_c(0)\,.\end{aligned} \tag{1.215}$$

This shows that the coherent state moves in such a way that x_c and p_c change in exactly the same way as the phase space variables of a classical harmonic oscillator. This is consistent

with *Ehrenfest's theorem*, since x_c and p_c coincide with the expectation values $\langle x \rangle$ and $\langle p \rangle$. Since the coherent states keep their minimal uncertainty during the time evolution, they are as close as we can get to the classical, time dependent states of the harmonic oscillator within the quantum description.

The coherent state representation
The coherent states are expressed in the energy representation in the following way

$$
\begin{aligned}
\langle n|z\rangle &= \langle n|\hat{\mathcal{D}}(z)|0\rangle \\
&= \langle n|e^{z\hat{a}^\dagger - z^*\hat{a}}|0\rangle \\
&= \langle n|e^{-\frac{1}{2}|z|^2}e^{z\hat{a}^\dagger}e^{-z^*\hat{a}}|0\rangle \\
&= e^{-\frac{1}{2}|z|^2}\langle n|e^{z\hat{a}^\dagger}|0\rangle \\
&= e^{-\frac{1}{2}|z|^2}\langle n|\sum_{m=0}^{\infty}\frac{z^m}{m!}(\hat{a}^\dagger)^m|0\rangle \\
&= e^{-\frac{1}{2}|z|^2}\frac{z^n}{\sqrt{n!}}.
\end{aligned}
\tag{1.216}
$$

From this follows that the overlap between two coherent states is

$$
\begin{aligned}
\langle z|z'\rangle &= \sum_n \langle z|n\rangle\langle n|z'\rangle \\
&= e^{-\frac{1}{2}(|z|^2+|z'|^2)}\sum_n \frac{(z'z^*)^n}{n!} \\
&= e^{-\frac{1}{2}(|z|^2+|z'|^2)+z'z^*},
\end{aligned}
\tag{1.217}
$$

and for the absolute value this gives

$$
|\langle z|z'\rangle|^2 = e^{-|z-z'|^2}.
\tag{1.218}
$$

The coherent states corresponding to two different values of z are *not* orthogonal states, but the overlap falls off exponentially fast with the distance between the two points. This overlap gives a measure of the intrinsic uncertainty of the coherent state as a probability amplitude in phase space.

An interesting property of the coherent states is that, even if they are not orthogonal, they satisfy a completeness relation. To see this we calculate the following integral over the two-dimensional complex plane

$$
\begin{aligned}
\int d^2z|z\rangle\langle z| &= \int d^2z e^{-|z|^2}\sum_{n,m}\frac{z^n z^{*m}}{\sqrt{n!m!}}|n\rangle\langle m| \\
&= \int_0^{2\pi}d\theta\int_0^{\infty}dr\,r e^{-r^2}\sum_{n,m}\frac{r^{(n+m)}}{\sqrt{n!m!}}e^{i\theta(n-m)}|n\rangle\langle m| \\
&= 2\pi\int_0^{\infty}dr\,e^{-r^2}\sum_n \frac{r^{2n+1}}{n!}|n\rangle\langle n|
\end{aligned}
$$

$$= \pi \sum_n |n\rangle\langle n|$$

$$= \pi \, \hat{\mathbb{1}} \,, \tag{1.219}$$

where we have applied the identity

$$\int_0^\infty dr\, r^{2n+1} e^{-r^2} = \frac{1}{2} n! \,. \tag{1.220}$$

We rewrite the result (1.219) as the following completeness relation

$$\int \frac{d^2 z}{\pi} |z\rangle\langle z| = \hat{\mathbb{1}} \,. \tag{1.221}$$

With the help of the completeness relation the *coherent state representation* can be defined as an alternative to the coordinate representation and the momentum representation. The wave function, which is a function of the complex phase space variable z, is defined by the state vector $|\psi\rangle$ as

$$\psi(z) = \langle z|\psi\rangle \,, \tag{1.222}$$

and the inverse relation is

$$|\psi\rangle = \int \frac{d^2 z}{\pi} |z\rangle\langle z|\psi\rangle = \int \frac{d^2 z}{\pi} |z\rangle \psi(z) \,. \tag{1.223}$$

The two relations show that there is a 1-1 correspondence between abstract vectors $|\psi\rangle$ and wave functions $\psi(z)$, and the coherent state representation is therefore a well defined representation of the quantum theory.

One of the implications of the above relations is that the coherent states do not form a linearly independent set of states. They form instead an over-complete set. Thus,

$$|z\rangle = \int \frac{d^2 z'}{\pi} |z'\rangle\langle z'|z\rangle = \int \frac{d^2 z'}{\pi} |z'\rangle e^{-\frac{1}{2}(|z|^2 + |z'|^2) + z'^* z} \,, \tag{1.224}$$

and since the coherent state $|z\rangle$ can be expressed as a linear combination of other coherent states, this demonstrates the lack of linear independence. A consequence of this is that an expansion of a state vector $|\psi\rangle$ in terms of the set of coherent states is not uniquely defined. Nevertheless, the expansion given by (1.223) *is unique* because of constraints that implicitly are imposed on the wave functions $\psi(z)$. To see this we rewrite it in terms of the n-representation

$$\begin{aligned}
\psi(z) &= \sum_n \langle z|n\rangle\langle n|\psi\rangle \\
&= \sum_n \langle n|\psi\rangle e^{-\frac{1}{2}|z|^2} \frac{z^{*n}}{\sqrt{n!}} \\
&\equiv e^{-\frac{1}{2}|z|^2} f(z^*) \,. \tag{1.225}
\end{aligned}$$

The function

$$f(z^*) = \sum_n \langle n|\psi\rangle \frac{z^{*n}}{\sqrt{n!}} \,, \tag{1.226}$$

is an *analytic function* of z^* since it depends only on z^* and not on z. This is the constraint on the wave functions $\psi(z)$ that makes the coherent state representation well-defined, *the wave functions are up to a common factor $e^{-\frac{1}{2}|z|^2}$ restricted to be analytic functions*.

Thus, wave functions and observables of the originally one-dimensional problem can be rewritten in terms of analytic functions defined on the two-dimensional phase space. One should, however, be aware of the fact that several relations in this representation are unfamiliar, because of the non-orthogonality between the basis states $|z\rangle$.

The coherent states are important in many respects because of their close relation to classical states. They were introduced in the context of the quantum description of light, where they describe states of *classical light* within the quantum theory.

1.4.5 Fermionic and bosonic oscillators: an example of supersymmetry

There is a formal similarity between the two-level system and the harmonic oscillator which we shall examine in this section. To make the similarity explicit, we write the Hamiltonian of the two-level system as

$$\hat{H}_F = \frac{1}{2}\hbar\omega\sigma_z\,, \tag{1.227}$$

and introduce the raising and lowering operators

$$\hat{b}^\dagger = \sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y)\,, \quad \hat{b} = \sigma_- = \frac{1}{2}(\sigma_x - i\sigma_y)\,. \tag{1.228}$$

In matrix form the operators are

$$\hat{H}_F = \frac{1}{2}\hbar\omega\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad \hat{b}^\dagger = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\,, \quad \hat{b} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\,.$$

We now have the algebraic relations

$$\left\{\hat{b}, \hat{b}^\dagger\right\} = 1\,, \quad \hat{H}_F = \frac{1}{2}\hbar\omega\left[\hat{b}^\dagger, \hat{b}\right]\,, \tag{1.229}$$

where $\left\{\hat{b}, \hat{b}^\dagger\right\}$ is the *anticommutator* $\hat{b}\hat{b}^\dagger + \hat{b}^\dagger\hat{b}$. The corresponding relations for a harmonic oscillator are

$$\left[\hat{a}, \hat{a}^\dagger\right] = 1\,, \quad \hat{H}_B = \frac{1}{2}\hbar\omega\left\{\hat{a}^\dagger, \hat{a}\right\}\,. \tag{1.230}$$

We note that the (formal) transition between the two systems corresponds to interchanging commutators with anticommutators.

There are many physical realizations of these two systems. We will now focus on a simple *many-particle realization*. Let us assume that a single state is available for many identical particles. The state may be considered to be one of the field modes of a free field. The identical particles may be either fermions or bosons.

In the fermion case the state space will contain only two states. $|0\rangle$ is the *vacuum state* with no particle present and $|1\rangle$ is the excited state with one particle present. Due to the Pauli exclusion principle the single-particle state cannot be occupied by more than one particle. The

occupation energy of this state is $\hbar\omega$. With this interpretation of the two-level system the operator $\hat{b}^\dagger$ is a *creation operator* for a fermion and $\hat{b}$ is an annihilation operator. The Hamiltonian is defined so that the vacuum energy is $-\frac{1}{2}\hbar\omega$.

In the boson case there is an infinite number of states, since the single-particle state can be occupied by an arbitrary number of particles. The states $|n\rangle$ now are interpreted as states with n bosons present. The operator $\hat{a}^\dagger$ is a creation operator for bosons and $\hat{a}$ an annihilation operator. The boson vacuum state has a vacuum energy $+\frac{1}{2}\hbar\omega$ which is the ground state energy of the harmonic oscillator.

With the interpretation above in mind we may refer to the two-level system as a *fermionic oscillator* and the standard harmonic oscillator as a *bosonic oscillator*.

In recent years the idea has been extensively developed that nature has a (hidden) symmetry between fermions and bosons called *supersymmetry*. There is at this stage no physical evidence for the presence of this symmetry as a fundamental symmetry of nature. Nevertheless the idea has been pursued, with supersymmetry as an important input in *string theories* and *supergravity theories*.

We discuss here a simple realization of supersymmetry (or Fermi-Bose symmetry) as a symmetry for the two-level model and the quantum harmonic oscillator.

The symmetrized model has a Hamiltonian that can be written as a sum of the two Hamiltonians

$$\hat{H} = \hat{H}_F + \hat{H}_B = \frac{1}{2}\hbar\omega[\{\hat{a}^\dagger, \hat{a}\} + [\hat{b}^\dagger, \hat{b}]], \qquad (1.231)$$

where the operators of the two subsystems are assumed to commute. Since the level spacing of the two Hamiltonians have been chosen to be equal there is a double degeneracy of all the excited levels, while the ground state is non-degenerate, as shown in Fig.1.4.5.

The supersymmetry is made explicit in terms of a *supercharge*, defined as

$$\hat{Q} = \sqrt{\hbar\omega}\,\hat{a}^\dagger\hat{b}, \quad \hat{Q}^\dagger = \sqrt{\hbar\omega}\,\hat{a}\hat{b}^\dagger. \qquad (1.232)$$

Together with the Hamiltonian it defines a *supersymmetry algebra*

$$
\begin{aligned}
\{\hat{Q}, \hat{Q}^\dagger\} &= \hat{H}, \\
[\hat{Q}, \hat{H}] &= [\hat{Q}^\dagger, \hat{H}] = 0, \\
\{\hat{Q}, \hat{Q}\} &= 2\hat{Q}^2 = 0, \\
\{\hat{Q}^\dagger, \hat{Q}^\dagger\} &= 2\hat{Q}^{\dagger 2} = 0.
\end{aligned}
\qquad (1.233)
$$

This is not a commutator algebra (or Lie algebra), since it involves both commutators and anticommutators. It is referred to as a *graded (Lie) algebra*. $\hat{Q}$ and $\hat{Q}^\dagger$ are the fermionic (or odd) elements and $\hat{H}$ is the bosonic (or even) element of this graded algebra.

The supersymmetry gives, as a general feature, a ground state energy (vacuum energy) which is 0, due to cancellation of the contributions from the bosonic and fermionic variables. This type of cancellation is important in *supersymmetric quantum field theories*, where the divergent contributions to the vacuum energy are avoided.

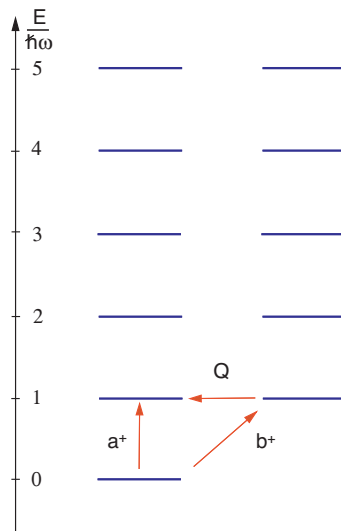Figure 1.6: The energy spectrum of a supersymmetric oscillator. The bosonic creation operator $\hat{a}^\dagger$ acts vertically, while the fermionic creation operator $\hat{b}^\dagger$ acts in the diagonal direction. The supercharge $\hat{Q}$ is a symmetry operator that maps between pairs of degenerate excited states. The ground state which is non-degenerate is annihilated by $\hat{Q}$ and $\hat{Q}^\dagger$.

.

Chapter 2

Quantum mechanics and probability

2.1 Classical and quantum probabilities

In this section we extend the quantum description to states that include "classical" uncertainties in addition to quantum probabilities. Such additional uncertainties may be due to lack of a full knowledge of the state vector of a system or due to a description of the system as member of a statistical ensemble. But it may also be due to *entanglement*, when the system is the part of a larger quantum system.

2.1.1 Pure and mixed states, the density operator

The state of a quantum system, described either as a wave function or an abstract vector in the state space, has a *probability interpretation*. Thus, the wave function is referred to as a *probability amplitude* and it predicts the result of a measurement performed on the system only in a statistical sense. The state vector therefore characterizes the state of the quantum system in a way that seems closer to the statistical description of a classical system than to a detailed, non-statistical description. However, in the standard interpretation, this uncertainty about the result of a measurement performed on the system is not ascribed to lack of information about the system. We refer to the quantum state described by a (single) state vector, as a *pure state* and consider this to contain *maximum available information* about the system. Thus, if we intend to acquire further information about the state of the system by performing measurements on the system, this will in general lead to a change of the state vector which we interpret as a *real modification* of the physical state due to the action of the measuring device. It cannot be interpreted simply as a (passive) collection of additional information.[1]

In reality we will often have less information about a quantum system than the maximally possible information contained in the state vector. Let us for example consider the spin state of the silver atoms emerging from the furnace of the Stern-Gerlach experiment. In principle

---

[1]An interesting question concerns the interpretation of a quantum state vector $|\psi\rangle$ as being the *objective* state associated with a *single* system. An alternative understanding, as stressed by Einstein, claims that the probability interpretation implies that the state vector can only be understood as an *ensemble variable*, it describes the state of an ensemble of identically prepared systems. From Einstein's point of view this means that additional information, beyond that included in the wave function, should *in principle* be possible to acquire when we consider single systems rather than ensembles.

each atom could be in a pure spin state $|\mathbf{n}\rangle$ with a quantized spin component in the $\mathbf{n}$-direction. However, the collection of atoms in the beam clearly have spins that are isotropically oriented in space and they cannot be described by a single spin vector. Instead they can be associated with a statistical ensemble of vectors. Since the spins are isotropically distributed, this indicates that the ensemble of spin vectors $|\mathbf{n}\rangle$ should have a uniform distribution over all directions $\mathbf{n}$. However, since the vectors of this ensemble are not linearly independent, the isotropic ensemble is equivalent to another ensemble with only two states, spin up and spin down in *any fixed direction* $\mathbf{n}$, with equal probability for the two states.

We therefore proceed to consider situations where a system is described not by a single state vector, but by an ensemble of state vectors, $\{|\psi\rangle_1, |\psi\rangle_2, ..., |\psi\rangle_n\}$ with a probability distribution $\{p_1, p_2, ..., p_n\}$ defined over the ensemble. We may consider this ensemble to contain both *quantum probabilities* carried by the state vectors $\{|\psi\rangle_k\}$ and *classical probabilities* carried by the distribution $\{p_k\}$. A system described by such an ensemble of states is said to be in a *mixed state*. There seems to be a clear distinction between the two types of probabilities, but as we shall see this is not fully correct. There are interesting examples of mixed states with no clear division between the quantum and classical probabilities.

The expectation value of a quantum observable in a state described by an ensemble of state vectors is

$$\langle A \rangle = \sum_{k=1}^{n} p_k \langle A \rangle_k = \sum_{k=1}^{n} p_k \langle \psi_k | \hat{A} | \psi_k \rangle . \tag{2.1}$$

This expression motivates the introduction of the *density operator* associated with the mixed state,

$$\hat{\rho} = \sum_{k=1}^{n} p_k \, |\psi_k\rangle\langle\psi_k| . \tag{2.2}$$

The corresponding matrix, defined by reference to an (orthogonal) basis $\{|\phi_i\rangle\}$, is called the *density matrix*,

$$\rho_{ij} = \sum_{k=1}^{n} p_k \, \langle\phi_i|\psi_k\rangle\langle\psi_k|\phi_j\rangle . \tag{2.3}$$

The important point to note is that *all measurable information about the mixed state is contained in the density operator of the state*, in the sense that the expectation value of *any* observable can be expressed in terms of $\hat{\rho}$,

$$
\begin{aligned}
\langle A \rangle &= \sum_{k=1}^{n} p_k \sum_{i} \langle\psi_k|\hat{A}|\phi_i\rangle\langle\phi_i|\psi_k\rangle \\
&= \sum_{i} \sum_{k=1}^{n} p_k \langle\phi_i|\psi_k\rangle\langle\psi_k|\hat{A}|\phi_i\rangle \\
&= \mathrm{Tr}(\hat{\rho}\hat{A}) .
\end{aligned}
\tag{2.4}
$$

There are certain general properties that any density operator has to satisfy,

$$a) \qquad \text{Hermiticity}: \quad p_k = p_k^* \;\; \Rightarrow \;\; \hat{\rho} = \hat{\rho}^\dagger ,$$

$$b) \qquad \text{Positivity}: \quad p_k \geq 0 \quad \Rightarrow \quad \langle \chi | \hat{\rho} | \chi \rangle \geq 0 \text{ for all } | \chi \rangle\,,$$

$$c) \qquad \text{Normalization}: \quad \sum_k p_k = 1 \quad \Rightarrow \quad \text{Tr } \hat{\rho} = 1\,. \qquad (2.5)$$

Thus, the eigenvalues are real and non-negative with a sum that equals one[2]. These conditions follow from (2.2) with the coefficients p_k interpreted as probabilities. We also note that

$$\text{Tr } \hat{\rho}^2 = \sum_k p_k^2 \quad \Rightarrow \quad 0 < \text{Tr } \hat{\rho}^2 \leq 1\,. \qquad (2.6)$$

This inequality follows from the fact that for all eigenvalues $p_k \leq 1$, which means that $\text{Tr } \hat{\rho}^2 \leq \text{Tr } \hat{\rho}$.

The pure states are the special case where one of the probabilities p_k is equal to 1 and the others are 0. In this case the density operator is the projection operator on a single state,

$$\hat{\rho} = |\psi\rangle\langle\psi| \quad \Rightarrow \quad \hat{\rho}^2 = \hat{\rho}\,. \qquad (2.7)$$

Therefore $Tr(\hat{\rho}^2) = 1$ for a pure state, while for all the (truly) mixed states $Tr(\hat{\rho}^2) < 1$.

A general density matrix can be written in the form

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k|\,, \qquad (2.8)$$

where the states $|\psi_k\rangle$ may be identified as members of a statistical ensemble of state vectors associated with the mixed state. Note, however, that *this expansion is not unique. There are many different ensembles that give rise to the same density matrix*. This means that even if all relevant information about the mixed state is contained in the density operator, in order for this to specify the expectation value of an arbitrary observable, there may in principle be additional information available that specifies the physical ensemble to which the system belongs.

An especially useful expansion of a density operator is the expansion in terms of its eigenstates. In this case the states $|\psi_k\rangle$ are orthogonal and the eigenvalues are the probabilities p_k associated with the eigenstates. This expansion is unique unless there are eigenvalues with degeneracies.

As opposed to the pure states, the mixed states are not providing the maximal possible information about the system. This is due to the classical probabilities contained in the mixed state, which to some degree makes it similar to a statistical state of a classical system. Thus, additional information may in principle be obtainable without interacting with the system. For example, two different observers may have different degrees of information about the system and therefore associate different density matrices to the system. By exchanging information they may increase their knowledge about the system without interacting with it. If, on the other hand, two observers have *maximal* information about the system, which means that they both describe it by a pure state, they have to associate the same state vector with the system in order to have a consistent description.

A mixed state described by the density operator (2.8) is sometimes referred to as an *incoherent* mixture of the states $|\psi_k\rangle$. A *coherent* mixture is instead a *superposition* of the states,

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle\,, \qquad (2.9)$$

---

[2]The positivity condition we often write in a simple form as $\hat{\rho} \geq 0$. It means that the operator is *positive semi-definite*, *i.e.*, it has only non-negative eigenvalues

which then represents a pure state. The corresponding density matrix can be written as

$$\hat{\rho} = \sum_k |c_k|^2 |\psi_k\rangle\langle\psi_k| + \sum_{k\neq l} c_k c_l^* |\psi_k\rangle\langle\psi_l|. \tag{2.10}$$

Comparing this with (2.8) we note that the first sum in (2.10) can be identified with the incoherent mixture of the states (with $p_k = |c_k|^2$). The second sum involves the *interference* terms of the superposition and these are essential for the coherence effect. Consequently, if the off-diagonal matrix elements of the density matrix (2.10) are erased, the pure state is reduced to a mixed state with the same probability p_k for the states $|\psi_k\rangle$.

The time evolution of the density operator for an isolated (closed) system is determined by the Schrödinger equation. As follows from the expression (2.8) the density operator satisfies the dynamical equation

$$i\hbar \frac{\partial}{\partial t}\hat{\rho} = \left[\hat{H}, \hat{\rho}\right]. \tag{2.11}$$

This looks similar to the Heisenberg equation of motion for an observable (except for a sign change), but one should note that Eq.(2.11) *is valid in the Schrödinger picture*. In the Heisenberg picture, the density operator, like the state vector is time independent.

One should note the close similarity between Eq.(2.11) and Liouville's equation for the classical probability density $\rho(q, p)$ in phase space

$$\frac{\partial}{\partial t}\rho = \{\rho, H\}_{PB}. \tag{2.12}$$

where $\{,\}_{PB}$ is the Poisson bracket.

2.1.2 Entropy

In the same way as one associates *entropy* with statistical states of a classical system, one associates entropy with mixed quantum states as a measure of the lack of (optimal) information about the state. The *von Neuman entropy* is defined as

$$S = -Tr(\hat{\rho}\log\hat{\rho}). \tag{2.13}$$

Rewritten in terms of the eigenvalues of $\hat{\rho}$ it has the form

$$S = -\sum_k p_k \log p_k, \tag{2.14}$$

which shows that it is closely related to the entropy defined in statistical mechanics and in information theory. [3]

The pure states are states with zero entropy. For mixed states the entropy measures how far the state is from being pure. The entropy increases when the probabilities get distributed

---

[3]In statistical mechanics the natural logarithm is usually used in the definition of entropy, but in information theory the base 2 logarithm is more common. There is no basic difference between these choices, since the difference is only a state independent, constant factor. To make a specific choice we may consider $\log$ in the following to mean the base 2 logarithm.

over many states. In particular we note that for a finite dimensional Hilbert space, a maximal entropy state exists where all states are equally probable. The corresponding density operator is

$$\hat{\rho}_{max} = \frac{1}{n} \sum_k |k\rangle\langle k| = \frac{1}{n} \hat{\mathbb{1}} \,, \tag{2.15}$$

where n is the dimension of the Hilbert space and $\{|k\rangle\}$ is an orthonormal set of basis vectors. Thus, the density operator is proportional to a projection operator that projects on the full Hilbert space. The corresponding maximal value of the entropy is

$$S_{max} = \log n \,. \tag{2.16}$$

A *thermal state* is a special case of a mixed state, with a (statistical) *Boltzmann distribution* over the energy levels. It is described by a *temperature dependent* density operator of the form

$$\hat{\rho} = N e^{-\beta \hat{H}} = N \sum_k e^{-\beta E_k} |\psi_k\rangle\langle\psi_k| \,, \tag{2.17}$$

with E_k as the energy eigenvalues, $|\psi_k\rangle$ as the energy eigenvectors, and $\beta = 1/(k_B T)$, with k_B as the Boltzmann constant. N is the normalization factor,

$$N^{-1} = \text{Tr}\, e^{-\beta \hat{H}} = \sum_k e^{-\beta E_k} \,, \tag{2.18}$$

which gives $\hat{\rho}$ the correct normalization (2.5), consistent with the probability interpretation. The close relation between the normalization factor and the *partition function* in (classical) statistical mechanics is apparent.

Quantum statistical mechanics is based on definitions of density matrices associated with different statistical ensembles. Thus, the density matrix (2.17) is associated with a *canonical ensemble* of quantum states. Furthermore the *thermodynamic entropy* is, in the quantum statistical mechanics, identical to the von Neuman entropy (2.13) apart from a factor proportional to the Boltzmann constant k_B. We will in the following make some further study of the entropy, but focus mainly on its *information* content rather than thermodynamic relevance. Whereas the thermodynamic entropy is most relevant for systems with a large number of degrees of freedom, the von Neuman entropy (2.13) is also highly relevant for small systems in the context of quantum information.

2.1.3 Mixed states for a two-level system

For the two-level system we can give an explicit (geometrical) representation of the density operators of mixed (and pure) states. Since the density operators are hermitian, with trace 1, they may be written as 2x2 matrices in the form

$$\hat{\rho} = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma}) \,, \tag{2.19}$$

with $\mathbf{r}$ as a real, three-component vector, $\mathbb{1}$ is the unit matrix and $\boldsymbol{\sigma}$ as the vector form of the Pauli matrices. The condition for the density operator to represent a pure state is

$$
\begin{aligned}
\hat{\rho}^2 &= \hat{\rho} \\
\Rightarrow \quad \frac{1}{4}(\mathbb{1} + \mathbf{r}^2 + 2\mathbf{r} \cdot \boldsymbol{\sigma}) &= \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma}) \\
\Rightarrow \quad \mathbf{r}^2 &= 1 \,.
\end{aligned}
\tag{2.20}
$$

Thus, a pure state can be written as

$$
\hat{\rho} = \frac{1}{2}(\mathbb{1} + \mathbf{n} \cdot \boldsymbol{\sigma}) \,,
\tag{2.21}
$$

with $\mathbf{n}$ as a unit vector. This matrix represents the projection operator that projects on the one-dimensional space spanned by the *spin up* eigenvector of $\mathbf{n} \cdot \boldsymbol{\sigma}$. We note that this is consistent with the discussion of section 1.3.1, where it was shown that *any state* in the two-dimensional space would be the spin up eigenstate of the operator $\mathbf{n} \cdot \boldsymbol{\sigma}$ for some unit vector $\mathbf{n}$.

For a general, mixed state we can write the density matrix as

$$
\begin{aligned}
\rho &= \frac{1}{2}(\mathbb{1} + r\mathbf{n} \cdot \boldsymbol{\sigma}) \\
&= \frac{1}{2}(1 + r)\frac{\mathbb{1} + \mathbf{n} \cdot \boldsymbol{\sigma}}{2} + \frac{1}{2}(1 - r)\frac{\mathbb{1} - \mathbf{n} \cdot \boldsymbol{\sigma}}{2} \\
&= \frac{1}{2}(1 + r)P_+(\mathbf{n}) + \frac{1}{2}(1 - r)P_-(\mathbf{n}) \,,
\end{aligned}
\tag{2.22}
$$

with $\mathbf{r} = r\mathbf{n}$ and $P_\pm(\mathbf{n})$ as the matrices that project on spin up (+) or spin down (-). This means that the eigenvectors of ρ, also in this case are the eigenvectors of $\mathbf{n} \cdot \boldsymbol{\sigma}$, but the eigenvalue p_+ of the spin up state and the eigenvalue p_- of the spin down state are given by

$$
p_\pm = \frac{1}{2}(1 \pm r) \,.
\tag{2.23}
$$

The interpretation of $p_\pm$ as probabilities means that they have to be positive (and less than one). This is satisfied if $0 \leq r \leq 1$. This result shows that all the physical states, both pure and mixed can be represented by points in a three dimensional sphere of radius $r = 1$. The points on the surface ($r = 1$) correspond to the pure states. As r decreases the states gets less pure and for $r = 0$ we find the state of maximal entropy. The entropy, defined by (2.14), is a monotonic function of r, with maximum ($\log 2$) for $r = 0$ and minimum (0) for $r = 1$. The explicit expression for the entropy is

$$
\begin{aligned}
S &= -(\frac{1}{2}(1 + r)\log\frac{1}{2}(1 + r) + \frac{1}{2}(1 - r)\log\frac{1}{2}(1 - r)) \\
&= -\frac{1}{2}\log[(1 + r)^{1+r}(1 - r)^{1-r}] + \log 2 \,.
\end{aligned}
\tag{2.24}
$$

The *sphere of states* for the two-level system is called the *Bloch sphere*[4].

---

[4]The Bloch sphere is a special case of a *convex set*, since all matrices on a line between two density matrices is
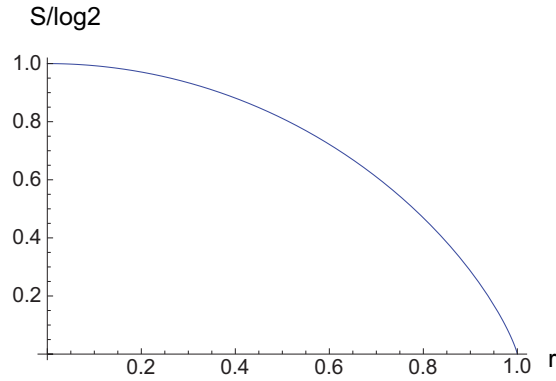
Figure 2.1: The entropy of mixed states for a two level system. The entropy S is shown as a function of the radial variable r for the *Bloch sphere representation* of the states.

2.2   Entanglement

In this section we consider *composite systems*, which can be considered as consisting of two or more *subsystems*. These subsystems will in general be *correlated*, often caused by (previous) interactions between the two parts. In quantum systems the correlations may in some sense be stronger than what is possible in a classical system. We refer to this as due to quantum *entanglement* between the subsystems. Entanglement is considered as one of the clearest marks of the difference between classical and quantum physics.

2.2.1   Composite systems

Up to this point we have mainly considered *isolated quantum systems*, with the systems restricted to pure states. Only the variables of the isolated system then enters in the quantum description, in the form of state vectors and observables, while the dynamical variables of other systems are irrelevant. As long as the system stays isolated, *i.e.*, with interactions with other systems being negligible, the system will continue to be in a pure state, with the time evolution described by the Schrödinger equation. For a mixed state, however, the degree of "non-purity" measured by the entropy will stay constant as long as the system is isolated. This follows from the fact that the time evolution is unitary and the eigenvalues of the density operator therefore do not change with time.

Clearly an isolated system is an abstraction, since interactions with other systems (generally referred to as the *environment*) can never be totally absent. However, in some cases this idealization works perfectly well. If interactions with the environments cannot be neglected, also variables associated with the environment have to be taken into account. If these act *ran-*

---

also a density matrix. This property is generally valid, as follows from the original definition (2.2), which expresses a general density operator as a convex combination of pure states. The pure states define the *extreme points* of the convex set of density matrices, which generate all other points (matrices) of the convex set through convex combinations, but which cannot themselves be expressed as convex combinations of other points in the set. For a two-level system the pure states define the boundary of the set of density operators. This does not happen for quantum systems of higher dimensions, where the boundary includes density matrices that are not pure.

domly on the system they tend to introduce *decoherence* in the system. This means that the state develops in the direction of being less pure, *i.e.*, the entropy increases. On the other hand, a *systematic* manipulation of the system may change the state in the direction of being more pure. Thus, a measurement performed on the system will usually be of this kind.

Even if a system $\mathcal{A}$ cannot be considered as isolated, sometimes it will be a part of a *larger* isolated system $\mathcal{S}$. We will consider this situation and assume that the total system $\mathcal{S}$ can be described in terms of a set of variables for system $\mathcal{A}$ and a set of variables for the rest of the system, which we denote $\mathcal{B}$. We assume that these two sets of variables can be regarded as independent (they are associated with independent degrees of freedom for the two subsystems $\mathcal{A}$ and $\mathcal{B}$), and that the interactions between $\mathcal{A}$ and $\mathcal{B}$ do not destroy this relative independence. We further assume that the totality of dynamical variables for $\mathcal{A}$ and $\mathcal{B}$ provides a complete set of variables for the full system $\mathcal{S}$.

Correlations is a typical feature of interacting systems, both at the classical and quantum level. Thus, the interactions between the two parts of a composite system will in general correlate their behavior. Such correlations will persist after the two subsystems have ceased to interact, and when that is the case they bring information about interactions in the past. For a classical system the correlations are often expressed in terms of *statistical correlation functions*, which can be expressed in terms of mean values of products of variables belonging to the two systems. In quantum mechanics the corresponding correlation functions are expectation values of products of observables for the subsystems. In the following we focus on such correlations with the aim of studying the *difference* between quantum and classical correlations. We first give a brief resumé of classical statistical correlations.

2.2.2   Classical statistical correlations

We consider a physical system $\mathcal{S}$ with two clearly distinguishable subsystems $\mathcal{A}$ and $\mathcal{B}$. Let us assume A to be a variable (observable) of system $\mathcal{A}$, and for simplicity assume this to have a discrete set of possible values $a_1, a_2, ...$ Similarly B is a variable of subsystem $\mathcal{B}$, with possible values $b_1, b_2, ...$. In a classical statistical description the variables A and B may each be associated with a probability distribution, so that p_k^A is the probability for the variable A to take the value a_k and p_l^B is the probability for the variable B to take the value b_l. To study correlations between the subsystems we introduce *joint probabilities*, where p_{kl} measures the probability for the variable A to take the value a_k *and* the variable B (at the same time) to take the value b_l. Clearly we have the relations

$$p_k^A = \sum_l p_{kl} \ , \quad p_l^B = \sum_k p_{kl} \ . \tag{2.25}$$

The statistical mean values for each subsystem then are given by

$$\langle A \rangle = \sum_k p_k^A a_k \ , \quad \langle B \rangle = \sum_l p_l^B b_l \ . \tag{2.26}$$

And for the composite system the mean value of the product variable AB is

$$\langle AB \rangle = \sum_{kl} p_{kl} a_k b_l \ . \tag{2.27}$$

The subsystems of the composite system are uncorrelated when all the information about the statistical mean values are contained in the probability distributions $\{p_k^A\}$ and $\{p_l^B\}$. This means that the joint probability has the product form

$$p_{kl} = p_k^A p_l^B \,, \tag{2.28}$$

and the mean value of the product AB therefore also factorizes,

$$\langle AB \rangle = \langle A \rangle \langle B \rangle \,. \tag{2.29}$$

If the probability distribution, and thereby the mean values, do not factorize in this way, that expresses the presence of statistical correlations between the subsystems. The correlations can then be studied in the form of correlation functions defined by

$$C(A, B) = \langle AB \rangle - \langle A \rangle \langle B \rangle \,. \tag{2.30}$$

We note that if we, in the classical setting, have full information about the composite system $\mathcal{S}$, and thereby full information about the subsystems $\mathcal{A}$ and $\mathcal{B}$, then the value of the product AB will trivially factorize. In this sense there are no correlations between the subsystems. Thus, for correlations between the subsystems to be a meaningful concept, the state of the full system, as well as of its parts, have to be described by statistical ensembles. For a quantum system this is different. Even if we have maximal information about the composite system, in the sense that it is in a pure state, we do not necessarily have maximal information about its parts. In this case there are correlations between the subsystems, and since the full system is in a pure state, these correlations are of non-classical nature. The systems $\mathcal{A}$ and $\mathcal{B}$ are entanglement.

2.2.3 States of a composite quantum system

In mathematical terms we describe the Hilbert space $\mathcal{H}$ of a composite quantum system as a *tensor product* of two Hilbert spaces $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ associated with the two subsystems, and write this as

$$\mathcal{H} = \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B} \,. \tag{2.31}$$

With $\{|\alpha\rangle\}$ as a basis for $\mathcal{H}_\mathcal{A}$ and $\{|\beta\rangle\}$ as a basis for $\mathcal{H}_\mathcal{B}$, a combination of vectors from the two sets, which we refer to as the tensor product of the vectors, forms a basis for $\mathcal{H}$,

$$|\alpha\beta\rangle = |\alpha\rangle \otimes |\beta\rangle \,, \tag{2.32}$$

and a general vector in $\mathcal{H}$ is then a linear combination of these tensor product vectors,

$$|\chi\rangle = \sum_{\alpha,\beta} c_{\alpha\beta} |\alpha\rangle \otimes |\beta\rangle \,. \tag{2.33}$$

Note that a vector in $\mathcal{H}$ is generally not a tensor product of vectors from $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$. For a tensor product vector the coefficients of the superposition factorizes, $c_{\alpha\beta} = c_\alpha^\mathcal{A} c_\beta^\mathcal{B}$, so that,

$$|\chi\rangle = (\sum_\alpha c_\alpha^\mathcal{A} |\alpha\rangle) \otimes (\sum_\beta c_\beta^\mathcal{B} |\beta\rangle) \equiv |\chi_\mathcal{A}\rangle \otimes |\chi_\mathcal{B}\rangle \,. \tag{2.34}$$

With $n_\mathcal{A}$ as the dimension of $\mathcal{H}_\mathcal{A}$ and $n_\mathcal{B}$ as the dimension of $\mathcal{H}_\mathcal{B}$ the dimension of the full Hilbert space $\mathcal{H}$ is clearly $n = n_\mathcal{A}n_\mathcal{B}$.

The observables of subsystems $\mathcal{A}$ and $\mathcal{B}$ are built into the Hilbert space description of the full system in the form of *commuting observables*. Thus the action of these observables are defined by the action on the product basis vectors as

$$\hat{A}|\alpha\beta\rangle = \sum_{\alpha'} A_{\alpha'\alpha}|\alpha'\beta\rangle \,, \quad \hat{B}|\alpha\beta\rangle = \sum_{\beta'} B_{\beta'\beta}|\alpha\beta'\rangle \,, \tag{2.35}$$

and the commutativity then readily follows

$$\hat{A}\hat{B}|\alpha\beta\rangle = \sum_{\alpha'\beta'} A_{\alpha'\alpha}B_{\beta'\beta}|\alpha'\beta'\rangle = \hat{B}\hat{A}|\alpha\beta\rangle \,. \tag{2.36}$$

We note that in the product basis, matrices have a double set of indices, with reference to basis vectors of both subsystem $\mathcal{A}$ and $\mathcal{B}$. This is in particular the case for the density matrix of the full system,

$$\rho_{\alpha'\beta',\alpha\beta} = \langle\alpha'\beta'|\,\hat{\rho}\,|\alpha\beta\rangle \,. \tag{2.37}$$

The expectation values of observables acting only on subsystem $\mathcal{A}$ are determined by the *reduced density matrix* of system $\mathcal{A}$. This matrix is obtained by taking the *partial trace* with respect to the coordinates of system $\mathcal{B}$,

$$\hat{\rho}^\mathcal{A} = Tr_\mathcal{B}(\hat{\rho}) \quad \Leftrightarrow \quad \rho_{\alpha\alpha'}^\mathcal{A} = \sum_{\beta} \rho_{\alpha\beta,\alpha'\beta} \,. \tag{2.38}$$

In the same way we define the reduced density matrix for system $\mathcal{B}$

$$\hat{\rho}^\mathcal{B} = Tr_\mathcal{A}(\hat{\rho}) \quad \Leftrightarrow \quad \rho_{\beta\beta'}^\mathcal{B} = \sum_{\alpha} \rho_{\alpha\beta,\alpha\beta'} \,. \tag{2.39}$$

In general the reduced density matrices will correspond to mixed states for the subsystems $\mathcal{A}$ and $\mathcal{B}$, even if the full system is in a pure state.

For a composite system there exist certain general relations between the entropy of the full system and the entropies of the subsystems. These relations are not identical for a (classical) statistical description and a quantum description of the composite system. Thus, for a bipartite system, consisting of the two parts $\mathcal{A}$ and $\mathcal{B}$ as discussed above, the following inequalities are satisfied in the classical description

$$a)\ S \leq S_\mathcal{A} + S_\mathcal{B} \,, \quad b)\ S \geq \max\{S_\mathcal{A}, S_\mathcal{B}\} \quad \text{(classical)} \,, \tag{2.40}$$

whereas the corresponding constraints in the quantum description are

$$a)\ S \leq S_\mathcal{A} + S_\mathcal{B} \,, \quad b)\ S \geq |S_\mathcal{A} - S_\mathcal{B}| \quad \text{(quantum)} \,. \tag{2.41}$$

The inequality a) is the same in the two cases. Equality here means that the two subsytems are uncorrelated, and if the entropy S of the full system is smaller than the sum $S_\mathcal{A} + S_\mathcal{B}$ that indicates the presence of correlations between the two subsystems. More interesting is inequality b). For the classical system it states that the entropy of the full system cannot be smaller than

the entropy of any of its parts. This seems quite natural in the sense that it can be interpreted as "the full system cannot be less ordered than any of its parts". However, in the quantum case the entropy of the full system can be lower than the entropy of its parts. A corresponding situation for a classical system would be rather paradoxical, but for a quantum system that can happen. The extreme situation appears when the full system is in a pure quantum state, with $S = 0$, while both entropies S_A and S_B of the subsystems are non-vanishing and positive. We interpret this as telling us that quantum correlations are different from, and can be stronger than correlations allowed in a classical statistical description.

2.2.4 Correlations and entanglement

In a quantum system we distinguish between different levels of correlations. We begin with product states

$$\hat{\rho} = \hat{\rho}^A \otimes \hat{\rho}^B \,, \tag{2.42}$$

which are states with *no correlations* between the subsystems A and B. For any observable $\hat{A}$ acting on A and $\hat{B}$ acting on B the expectation value of the product operator $\hat{A}\hat{B}$, for such a state is simply the product of the expectation values

$$\langle AB \rangle = \langle A \rangle_A \langle B \rangle_B \,. \tag{2.43}$$

Written in terms of the density operators,

$$Tr(\hat{\rho}\hat{A}\hat{B}) = Tr_A(\hat{\rho}^A \hat{A}) \, Tr_B(\hat{\rho}^B \hat{B}) \,. \tag{2.44}$$

Conversely, if the product relation (2.43) is correct for *all* observables $\hat{A}$ acting on A and *all* observables $\hat{B}$ acting on B, the two subsystems are uncorrelated and the density operator can be written in the product form (2.42).

We next consider mixed states of the form

$$\hat{\rho} = \sum_{kl} p_{kl} \, \hat{\rho}_k^A \otimes \hat{\rho}_l^B \,. \tag{2.45}$$

With $\{p_{kl}\}$ as a probability distribution, this density matrix can be viewed as describing a *statistical ensemble* of product states, *i.e.*, of states which, when regarded separately, do not contain correlations between the two subsytems. Such a state we refer to as *separable*. The correlations contained in a density operator of this form we may refer to as *classical* correlations, since these are of the same form that we find in a classical, statistical description of a composite system (see Eq. (2.27)). The expectation value of a product of observables for A and B in this case is

$$\begin{aligned} \langle AB \rangle &= \sum_{kl} p_{kl} \langle A \rangle_k \langle B \rangle_l \\ &= \sum_{kl} p_{kl} \, \text{Tr}_A(\hat{\rho}_k^A \hat{A}) \, \text{Tr}_B(\hat{\rho}_l^B \hat{B}) \,, \end{aligned} \tag{2.46}$$

and with $p_k^A = \sum_l p_{kl}$ and $p_l^B = \sum_k p_{kl}$ the product of expectation values is

$$
\begin{aligned}
\langle A \rangle \langle B \rangle &= \sum_k p_k^A \langle A \rangle_k \sum_l p_l^B \langle B \rangle_l \\
&= \sum_{kl} p_k^A \, p_l^B \, Tr_{\mathcal{A}}(\hat{\rho}_k^A \hat{A}) \, Tr_{\mathcal{B}}(\hat{\rho}_l^B \hat{B}) \, .
\end{aligned}
\tag{2.47}
$$

In general $\langle AB \rangle \neq \langle A \rangle \langle B \rangle$, and only when the probability distribution factorizes, $p_{kl} = p_k^A p_l^B$, the correlations between the two subsystems disappear. The situation is completely analogous to the situation with classical statistical correlations between the subsystem.

However, for a quantum system there is a next level of correlations. Such correlations are present for states described by density operators that are not of the product form (2.42) or of the separable form (2.45). Correlations of this kind is what we refer to as entanglement. In recent years there has been much effort devoted to give a concise *quantitative* meaning to entanglement in composite systems. However, for bipartite systems in mixed states and for multipartite systems in general, it is not obvious how to quantify deviations from classical correlations. But for a bipartite system in a pure state a unique definition of entanglement can be given, in terms of the entropy of the reduced density matrices of the subsystems. We focus on this case.

We consider then a general *pure* state of the total system $\mathcal{S} = \mathcal{A} + \mathcal{B}$. The entropy of such a state clearly vanishes. The state has the general form

$$
|\chi\rangle = \sum_{\alpha,\beta} c_{\alpha\beta} |\alpha\rangle_{\mathcal{A}} \otimes |\beta\rangle_{\mathcal{B}}
\tag{2.48}
$$

but can also be written in the diagonal form

$$
|\chi\rangle = \sum_n d_n |n\rangle_{\mathcal{A}} \otimes |n\rangle_{\mathcal{B}} \, ,
\tag{2.49}
$$

where $\{|n\rangle_{\mathcal{A}}\}$ is a set of orthonormal states for the $\mathcal{A}$ system and $\{|n\rangle_{\mathcal{B}}\}$ is a set of orthonormal states for the $\mathcal{B}$ system. This form of the state vector as a simple sum over product vectors, rather than a double sum, is referred to as the *Schmidt decomposition* of $|\chi\rangle$. It will be shown below that such an expansion is generally valid for a composite bipartite system.

The form of the state vector $|\chi\rangle$ is similar to that of density operator (2.45) for the correlated state. In the same way as for the density operator the system is uncorrelated if the sum includes only one term, *i.e.*, if the state has a product form. However, the correlations implied by the general form of the state vector (2.49) are different from those of the (classically) correlated state (2.45). From the Schmidt decomposition follows that the density matrix of the full system corresponding to the pure state $|\chi\rangle$ is

$$
\hat{\rho} = \sum_{nm} d_n d_m^* \, |n\rangle_{\mathcal{A}} \langle m|_{\mathcal{A}} \otimes |n\rangle_{B} \langle m|_{\mathcal{B}} \, .
\tag{2.50}
$$

This is not of the form (2.45). The state (2.49) is a linear combination (superposition) of products of *state vectors* of the two subsystems. But the corresponding density operator is not separable, *i.e.*, it cannot be written as a linear combination of products of *density matrices*.

From (2.50) also follows that the reduced density operators of the two subsystems have the form

$$\hat{\rho}^{\mathcal{A}} = \sum_n |d_n|^2 |n\rangle_{\mathcal{A}}\langle n|_{\mathcal{A}} , \quad \hat{\rho}^{\mathcal{B}} = \sum_n |d_n|^2 |n\rangle_{\mathcal{B}}\langle n|_{\mathcal{B}} . \tag{2.51}$$

The expressions show that all the eigenvalues of the two reduced density matrices are the same. This implies that the von Neuman entropy of the two subsystems is the same,

$$S_{\mathcal{A}} = S_{\mathcal{B}} = -\sum_n |d_n|^2 \log |d_n|^2 . \tag{2.52}$$

In this case, with the full system $\mathcal{A}+\mathcal{B}$ in a pure state, the entropy of one of the subsystems ($\mathcal{A}$ or $\mathcal{B}$) is taken as a quantitative measure of the entanglement between the subsystems. It is referred to as the *entanglement entropy* of the composite system.

The Schmidt decomposition
We show here that a general state of the composite system can be written in the form (2.49). The starting point is the general expression (2.48). We introduce a unitary transformation U for the basis of system $\mathcal{A}$,

$$|\alpha\rangle_{\mathcal{A}} = \sum_n U_{n\alpha} |n\rangle_{\mathcal{A}} , \tag{2.53}$$

and rewrite the state vector as

$$
\begin{aligned}
|\chi\rangle &= \sum_{\alpha,\beta} \sum_n c_{\alpha\beta} U_{n\alpha} |n\rangle_{\mathcal{A}} \otimes |\beta\rangle_{\mathcal{B}} \\
&= \sum_n |n\rangle_{\mathcal{A}} \otimes (\sum_{\alpha,\beta} c_{\alpha\beta} U_{n\alpha} |\beta\rangle_{\mathcal{B}}) \\
&\equiv \sum_n |n\rangle_{\mathcal{A}} \otimes |\tilde{n}\rangle_{\mathcal{B}} .
\end{aligned}
\tag{2.54}
$$

If the unitary transformation can be chosen to make the set of states $\{|\tilde{n}\rangle\}$ an *orthogonal* set of states, then (2.49) follows. The scalar product between two of the states is

$$
\begin{aligned}
\langle\tilde{n}|\tilde{m}\rangle_{\mathcal{B}} &= \sum_{\alpha,\alpha'} \sum_{\beta,\beta'} c_{\alpha\beta}^* U_{n\alpha}^* c_{\alpha'\beta'} U_{m\alpha'} \langle\beta|\beta'\rangle_{\mathcal{B}} \\
&= \sum_{\alpha,\alpha'} \sum_{\beta} c_{\alpha\beta}^* U_{n\alpha}^* c_{\alpha'\beta} U_{m\alpha'} \\
&= (UCC^{\dagger}U^{\dagger})_{mn} .
\end{aligned}
\tag{2.55}
$$

In the last expression the coefficients $c_{\alpha\beta}$ have been treated as the matrix elements of the matrix C. We note that the matrix $M = CC^{\dagger}$ is hermitian and positive definite, and the unitary matrix U can therefore be chosen to diagonalize M. With the (non-negative) eigenvalues written as $|d_n|^2$ we have

$$\langle\tilde{n}|\tilde{m}\rangle_{\mathcal{B}} = |d_n|^2 \delta_{mn} , \tag{2.56}$$

and if the normalized vectors $|m\rangle$ are introduced by

$$|\tilde{m}\rangle = d_m|m\rangle\,. \tag{2.57}$$

then the Schmidt form (2.49) of the vector $|\chi\rangle$ is reproduced with the basis sets of both systems $\mathcal{A}$ and $\mathcal{B}$ as orthonormal vectors.[5]

2.2.5 Entanglement in a two-spin system

As an example, we consider entanglement in the simplest possible composite system, which is a system that consists of two two-level subsystems (for example two spin-half systems). We shall consider several sets of vectors with different degrees of entanglement.

An orthogonal basis of product states is given by the four vectors

$$
\begin{aligned}
|\uparrow\uparrow\rangle &= |\uparrow\rangle_\mathcal{A} \otimes |\uparrow\rangle_\mathcal{B}\,,\\
|\uparrow\downarrow\rangle &= |\uparrow\rangle_\mathcal{A} \otimes |\downarrow\rangle_\mathcal{B}\,,\\
|\downarrow\uparrow\rangle &= |\downarrow\rangle_\mathcal{A} \otimes |\uparrow\rangle_\mathcal{B}\,,\\
|\downarrow\downarrow\rangle &= |\downarrow\rangle_\mathcal{A} \otimes |\downarrow\rangle_\mathcal{B}\,,
\end{aligned}
\tag{2.58}
$$

where $\{|\uparrow\rangle, |\downarrow\rangle\}$ is an orthonormal basis set for the spin-half system.

Another basis is given by states with well-defined total spin. As is well-known by the rules of addition of angular momentum, two spin half systems will have states of total spin 0 or 1. The spin 0 state is the antisymmetric (spin singlet) state

$$|0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)\,, \tag{2.59}$$

while spin 1 is described by the symmetric (spin triplet) states

$$|1,1\rangle = |\uparrow\uparrow\rangle\,, \quad |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\,, \quad |1,-1\rangle = |\downarrow\downarrow\rangle\,. \tag{2.60}$$

Clearly the two states $|1,1\rangle = |\uparrow\uparrow\rangle$ and $|1,-1\rangle = |\downarrow\downarrow\rangle$ are product states with no correlation between the two spin systems. However, the two remaining states $|0\rangle$ and $|1,0\rangle$ are entangled. These states may be included as two of the states of a third basis of orthonormal states, called the *Bell states*. They are defined by

$$
\begin{aligned}
|a,\pm\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle)\,,\\
|c,\pm\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle)\,,
\end{aligned}
\tag{2.61}
$$

and are all states of *maximal entanglement* between the two subsystems. We note that the two spins are strictly *anti-correlated* for the first two states ($|a,\pm\rangle$) and strictly *correlated* for the two other states ($|c,\pm\rangle$).

---

[5]In mathematical terminology the Schmidt decomposition is usually referred to as *singular value decomposition*.

Let us focus on the spin singlet state $|a, -\rangle$. It is a pure state (of system $\mathcal{A}+\mathcal{B}$) with density matrix

$$
\begin{aligned}
\hat{\rho}(a, -) &= |a, -\rangle\langle a, -| \\
&= \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow|) .
\end{aligned} \tag{2.62}
$$

The corresponding reduced density matrices of both systems A and B are of the same form

$$
\hat{\rho}^{\mathcal{A}} = \hat{\rho}^{\mathcal{B}} = \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) . \tag{2.63}
$$

The symmetric state $\hat{\rho}(a, +)$ has the same reduced density matrices as $\hat{\rho}(a, -)$. In fact this is true for *all* the four Bell states. Thus, the information contained in the density matrices of the subsystems do not distinguish between these four states of the total system.

The loss of information when we consider the reduced density matrices is demonstrated explicitly in taking the partial trace of (2.62) with respect to subsystem $\mathcal{A}$ or subsystem $\mathcal{B}$. The two last terms in (2.62) simply do not contribute. If we leave out the two last terms of (2.62) we have the following density matrix, which also have the same reduced density matrices,

$$
\hat{\rho}(a) = \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|) . \tag{2.64}
$$

This is still strictly anti-correlated in the spin of the two particles, but the correlation is now *classical* in the sense that the full density matrix is written in the form (2.45). Thus, the terms that are important for the quantum entanglement between the two subsystems are the ones that are left out when we take the partial trace. These terms are the off-diagonal *interference terms* of the full density matrix, and this means that we can view the entanglement as a special type of interference effect associated with the composite system.

When the reduced density operators are written as 2×2 matrices they have the form

$$
\hat{\rho}^{\mathcal{A}} = \hat{\rho}^{\mathcal{B}} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{2.65}
$$

Since they are proportional to the 2×2 unit operator they correspond to states with maximal entropy $S_A = S_B = \log 2$. With the entropy of the subsystems taken as a measure of entanglement this means that the spin singlet state is a state of *maximal entanglement*. This is true for all the Bell states, as already has been mentioned.

Finally, note that the reduced density matrix (2.63) is *rotationally invariant*. This is however not the case for the "classical" density matrix (2.64) of the full system. It describes a state where the spins of the two particles *along one particular direction* is strictly anti-correlated. It is interesting to note that we cannot define a density matrix of this form which predicts strict anti-correlation for the spin in *any* direction. Also note that neither is the pure state $|a, +\rangle$ rotationally invariant. In fact the state $|a, -\rangle$ is the only Bell state that is rotationally invariant, since it corresponds to *spin 0*. This makes the spin singlet state particularly interesting, since it describes a situation where the spins are strictly anti-correlated *along any direction in space*.

2.3   Quantum states and physical reality

In 1935 Einstein, Podolsky and Rosen published a paper where they focussed on a central question in quantum mechanics. This question has to do with the relation between the *formalism of quantum mechanics*, with its probability interpretation, and the underlying *physical reality* that quantum mechanics describes. They pointed out, by way of a simple thought experiment, that one of the implications of quantum physics is that it blurs the distinction between the *objective reality of nature* and the *subjective description* used by the physicist. Their thought experiment has been referred to as the EPR paradox, and it has challenged physicists in their understanding of the relation between physics and the reality of natural phenomena up to this day. As conclusion drawn from the paradox Einstein, Podolsky and Rosen suggested that quantum mechanics is an *incomplete* theory of nature. Later, in 1964 John Bell showed that the conflict between quantum mechanics and intuitive notions of reality goes deeper. Quantum theory allows types of correlations that cannot be found in classical theories that obey the basic assumptions of *locality* and *reality*.

In this section we examine the EPR paradox in a form introduced by David Bohm and proceed to examine how the limitations of classical theory is broken in the form of Bell inequalities. As first emphasized by Erwin Schrödinger the basic element of quantum physics involved in these considerations is that of *entanglement*.

2.3.1   EPR-paradox

We consider a thought experiment, where two spin half particles (particle A and particle B) are produced in a spin singlet state (with total spin zero). The two particles move apart, but since they are considered to be well separated from any disturbance, they keep their correlations so that the spin state is left unchanged. The full wave function of the two-particle system can be viewed as the product of a spin state and a position state, but for our purpose only the spin state is of interest. Concerning the position state it is sufficient to know that after some time the particles are separated by a large distance.

At a point in time when the two particles are well separated, a spin measurement is performed on particle A. This will modify the spin state of this particle, but since particle B is far away, *the measurement cannot affect particle B in any real sense*. However, and this is the paradox, *the change in the spin state caused by the measurement will also influence the possible outcomes of spin measurements performed on particle B*.

The experiment is schematically shown in Fig.(2.3.1). From the time when the particles are emitted and until the time when the spin experiment is performed the two particles are in the entangled state

$$|a, -\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).\tag{2.66}$$

In this state the spin of the two particles are strictly *anti-correlated*. This means that if particle A is measured to be in a spin up state, particle B is necessarily in a spin down state. But each particle, when viewed separately, is with equal probability found with spin up and spin down.

The reference to *spin up* ($\uparrow$) and *spin down* ($\downarrow$) in the state vector (2.66) seems to indicate that we have given preference to some particular direction in space. However, the singlet state
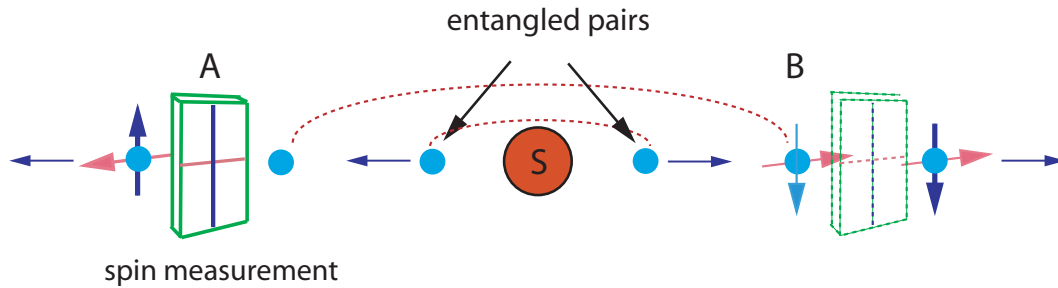
Figure 2.2: The set-up of an EPR experiment. Pairs of entangled particles are prepared in a singlet spin state and sent in opposite directions from a source S. (In the figure the entangled pairs are shown connected with dashed red curves.) The spin direction of each particle is, in this initial state, totally undetermined. At some distance from the source, one particle (A) from each pair passes a measuring device (shown by the green window) which measures the spin component in the z- direction. After passing the measuring apparatus the particle appears with quantized spin in the z-direction, either spin up as shown in the figure or with spin down. Due to the strict anti-correlation between the spin orientation of particle A and particle B in a pair, at the same time as particle A appears with spin up in the z-direction particle B will appear with spin down along the z-axis. In this way the measuring of the spin of particle A effectively acts as a measurement of the spin of particle B (indicated by the dashed green window) even if no measuring device acts on this particle. Following EPR we make the following deduction. Since particle B has been subject to no physical action, spin down along the z-axis for particle B after the measurement on A implies that B had spin down along the z-action also before the measurement (even if that was not known prior to the measurement). But before the measurement was performed we could in principle have decided to measure the spin in the x-direction. If that was done the spin of particle B in the x-direction (red arrow) would have been quantized. In this way, by choosing how to perform the measurement on A, we can decide whether to quantize the spin of B in the z- or the x-direction. Again, since no physical interaction with B has taken place, this implies that the spin components both in the z- and the x-directions must have had quantized, sharp (but unknown) values before the measurement. This conclusion is not consistent with quantum mechanics, since the two components of the spin are *incompatible* observables, hence the paradox.

is *rotationally invariant*. Therefore the state is left unchanged if we redefine this direction in space. Thus, we do not have to specify whether the z-axis, the x-axis or any other direction has been chosen, they all give rise to the same (spin 0) state.

We now consider, in this hypothetical experiment, that the spin of particle A is measured along the z-axis. If the result is *spin up*, we know with certainty that the spin of particle B in the z-direction is *spin down*. Likewise, if the result of measurement on A is *spin down* the spin of particle B in the z-direction is *spin up*. In both cases, a measurement of the z spin component of particle A will with necessity project particle B into a state with well defined (quantized) z-component of the spin. Since no real change can have been introduced in the state of particle B (it is far away) it seems natural to conclude that the spin component along the z-axis must have had a sharp (although unknown value) also *before* the measurement was actually performed on particle A.

This conclusion is on the other hand in conflict with the standard interpretation of quantum mechanics. Thus, if the z-component of the spin of particle B has a sharp value before the measurement, it will have a sharp value even if the measurement along the z-axis is not per-

formed at all, and even if we choose to measure the spin along the x-axis instead. But now the argument can be repeated for the x-component of the spin. This will in the same way lead to the conclusion that the x-component of the spin of particle B has a sharp value before the measurement. Consequently, both the z-component and the x-component of the spin must have a sharp (but unknown) values before the measurement is performed on particle A. But this is not in accordance with the standard interpretation of quantum mechanics. The z-component and the x-component of the spin operator are *incompatible observables*, since they do not commute as operators. Therefore they cannot simultaneously be assigned sharp values.

The conclusion Einstein, Podolsky and Rosen drew from the thought experiment is that both observables will in reality have sharp values, but since quantum mechanics is not able to predict these values with certainty (only with probabilities) quantum theory cannot be a complete theory of nature. From their point of view there seems to be some missing variables in the theory. If such variables are introduced in the quantum description, we refer to the theory as a "hidden variable" theory.

Note the two elements that enter into the above argument:
Locality — Since particle B is far away from where the measurement takes place we conclude that no real change can take place concerning the state of particle B.
Reality — When the measurement on particle A makes it possible to predict *with certainty* the outcome of a measurement performed on particle B, we draw the conclusion that this represents a property of B which is there even without actually performing a measurement on B.

From later studies we know that the idea of Einstein, Podolsky and Rosen that quantum theory is an incomplete theory does not really resolve the EPR paradox. Hidden variable theories can in principle be introduced, but to be consistent with the predictions of quantum mechanics, they cannot satisfy the conditions of locality and reality used by EPR. Thus, there is a real *conflict* between quantum mechanics and basic principles of classical theories of nature.

From the discussion above it is clear that *entanglement* between the two particles is the source of the (apparent) problem. It seems natural to draw the conclusion that if a measurement is performed on one of the partners of an entangled system there will be (immediately) a change in the state also of the other partner. But if we phrase the effect in this way, one should be aware of the fact that the effect of entanglement is always hidden in correlations between measurements performed on the two parts. The measurement performed on particle A does not lead to any local change at particle B that can be seen without consulting the result of the measurement performed on A. Thus, no change that can be interpreted as a signal sent from A to B has taken place. This means that the immediate change that is affecting the total system A+B due to the measurement on A does not introduce results that are in conflict with causality.

2.3.2   Bell's inequality

Correlation functions

The EPR paradox shows that quantum mechanics is radically different from classical statistical theories, when entanglement between systems is involved. In the original presentation of Einstein, Podolsky and Rosen it was left open as a possibility that a more complete description of nature could replace quantum mechanics. However, by studying correlations between spin systems, John Bell concluded in 1964 that the predictions of quantum mechanics is in direct conflict with what can be deduced from a "more complete" theory with hidden variables, when

this satisfy what is known as *Einstein locality*. This was expressed in terms of a certain inequality (Bell inequality), that had to be satisfied by a modified spin theory, if this should have the form of a local and realistic theory. Quantum theory was shown not to satisfy, in general, this inequality, and the Bell inequality therefore set a borderline between classical and quantum correlations in this system.

To examine this difference between classical and quantum correlations, we follow John Bell and reconsider the EPR experiment, but focus on correlations between measurements of spin components performed on both particles A and B. Let us assume that the particles move in the y-direction. The spin component in a rotated direction in the x, z-plane will have the form

$$\hat{S}_\theta = \cos\theta \hat{S}_z + \sin\theta \hat{S}_x \,. \tag{2.67}$$

We consider now the following correlation function between spin measurements on the two particles,

$$E(\theta_A, \theta_B) \equiv \frac{4}{\hbar^2} \langle S_{\theta_A} S_{\theta_B} \rangle \,. \tag{2.68}$$

Thus, we are interested in correlations between spin directions that are rotated by *arbitrary* angles θ_A and θ_B in the x, z-plane. The factor $\hbar^2/4$ is included to make $E(\theta_A, \theta_B)$ a dimensionless function, normalized to take values in the interval $(-1, +1)$.

For the spin singlet state it is easy to calculate this correlation function. It is

$$E(\theta_A, \theta_B) = -\cos(\theta_A - \theta_B) \,. \tag{2.69}$$

With the two spin measurements oriented in the same direction, $\theta_A = \theta_B \equiv \theta$, we have

$$E(\theta, \theta) = -1 \,, \tag{2.70}$$

which shows the anti-correlation between the two spin vectors.

Let us now leave the predictions of quantum mechanics and look at the limitations of a *local, realistic hidden variable theory*. We write the measured spin values as $S_A = \frac{\hbar}{2}\sigma_A$ and $S_B = \frac{\hbar}{2}\sigma_B$ so that σ_A and σ_B are normalized to unity in absolute value. We make the following assumptions, which build in the requirements of a local and realistic hidden variable theory together with the facts that are known about spin half systems[6],

1. *Hidden variables.* The measured spin values S_A of particle A and S_B of particle B are determined (before the measurement) by one or more unknown variables, denoted λ. Thus,

$\sigma_A = \sigma_A(\lambda)$ and $\sigma_B = \sigma_B(\lambda)$.

2. *Einstein locality.* The spin value S_A will depend on the orientation θ_A, but not on the orientation θ_B, when the two measuring devices are separated by a large distance. Similarly S_B will be independent of θ_A. We write this as,

---

[6]Bell's analysis focusses on the conflict between the possible correlations in a quantum and a classical system. But to assume sharp values for the components of the particle spin in different directions creates a problem even without making reference to correlations. As discussed in the context of the Stein-Gerlach experiment, the fact that only *spin up* and *spin down* can be measured for a spin component seems incompatible with rotational invariance when spin is treated as a classical variable. This means that the classical spin theory considered by Bell is unconventional, in the sense that the measured spin values in different directions cannot be regarded as the components of a spin vector that can be continuously rotated in space.

$$\sigma_A = \sigma_A(\theta_A, \lambda) \quad \sigma_B = \sigma_B(\theta_B, \lambda) \,.$$

3. *Spin quantization.* We assume the possible outcome of measurements are consistent with spin quantization,

$$\sigma_A = \pm 1 \quad \sigma_B = \pm 1 \,.$$

4. *Anti-correlation.* We assume that the measured values are consistent with the known anti-correlation of the singlet state

$$\sigma_A(\theta, \lambda) = -\sigma_B(\theta, \lambda) \equiv \sigma(\theta, \lambda) \,.$$

Since the variable λ is unknown we assume that particles are emitted from the source S with some probability distribution $\rho(\lambda)$ over λ. Thus, the expectation value (2.68) can be written as

$$
\begin{aligned}
E(\theta_A, \theta_B) &= \int d\lambda \, \rho(\lambda) \, \sigma_A(\theta_A, \lambda) \, \sigma_B(\theta_B, \lambda) \\
&= -\int d\lambda \, \rho(\lambda) \, \sigma(\theta_A, \lambda) \, \sigma(\theta_B, \lambda) \,.
\end{aligned}
\tag{2.71}
$$

We now consider the difference between correlation function E for two different sets of angles

$$
\begin{aligned}
E(\theta_1, \theta_2) - E(\theta_1, \theta_3) &= -\int d\lambda \, \rho(\lambda)(\sigma(\theta_1, \lambda)\sigma(\theta_2, \lambda) - \sigma(\theta_1, \lambda)\sigma(\theta_3, \lambda)) \\
&= \int d\lambda \, \rho(\lambda)\sigma(\theta_1, \lambda)\sigma(\theta_2, \lambda) \, (\sigma(\theta_2, \lambda)\sigma(\theta_3, \lambda) - 1) \,.
\end{aligned}
\tag{2.72}
$$

In the last expression we have used $\sigma(\theta, \lambda)^2 = 1$, which follows from assumption 3. From this we deduce the following inequality

$$
\begin{aligned}
|E(\theta_1, \theta_2) - E(\theta_1, \theta_3)| &\leq \int d\lambda \, \rho(\lambda) \, (1 - \sigma(\theta_2, \lambda)\sigma(\theta_3, \lambda)) \\
&= 1 + E(\theta_2, \theta_3) \,,
\end{aligned}
\tag{2.73}
$$

which is one of the forms of the *Bell inequality*.

Let us make the further assumption that the correlation function only depends on the relative angle (rotational invariance),

$$E(\theta_1, \theta_2) = E(\theta_1 - \theta_2) \,. \tag{2.74}$$

Let us also assume that $E(\theta)$ increases monotonically with θ in the interval $(0, \pi)$. (Both these assumptions are true for the quantum mechanical correlation function (2.69).) We introduce the probability function

$$P(\theta) = \frac{1}{2}(E(\theta) + 1) \,. \tag{2.75}$$

This function gives the probability for measuring either *spin up* or *spin down* on *both spin measurements* (see the discussion below). It satisfies the simplified inequality

$$P(2\theta) \leq 2P(\theta) \,, \quad 0 \leq \theta \leq \frac{\pi}{2} \,. \tag{2.76}$$

Due to strict anti-correlation between the spin measurements for $\theta = 0$, which means strict correlation for $\theta = \pi$, the function satisfies the boundary conditions

$$P(0) = 0 , \quad P(\pi) = 1 . \tag{2.77}$$

We further note that the inequality (2.76) is satisfied if $P(\theta)$ is a *concave function* of θ (func-
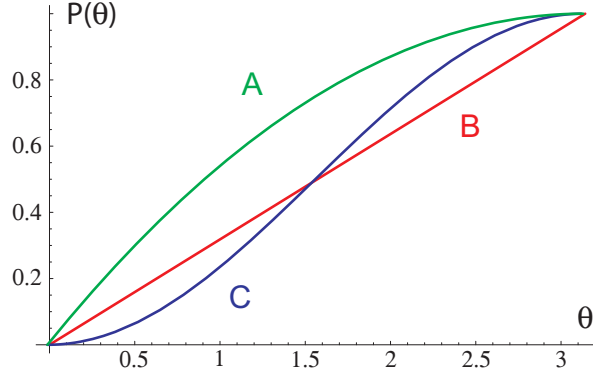

Figure 2.3: The probability $P(\theta)$ for measuring two *spin up*s or two *spin down*s as a function of the relative angle θ between the spin directions. Curve A satisfies the Bell inequality, Curve B is the limiting case where the Bell inequality is replaced by equality and Curve C is the probability predicted by quantum mechanics. Curve C is not consistent with the Bell inequality, but it is the curve confirmed by experiments.

tion with a negative second derivative), with the given boundary values. This is illustrated in Fig.(2.3.2), where $P(\theta) = \theta/\pi$ is a limiting function which satisfies (2.76) *with equality* rather than inequality.

We now turn to the predictions of quantum theory. As follows from (2.69) the function $P(\theta)$ is given by

$$P(\theta) = \sin^2 \frac{\theta}{2} . \tag{2.78}$$

This function does not satisfy the Bell inequality. This is clear from taking a special value $\theta = \pi/3$. We have

$$
\begin{aligned}
P(2\pi/3) &= \sin^2 \frac{\pi}{3} = \frac{3}{4} , \\
2P(\pi/3) &= 2\sin^2 \frac{\pi}{6} = \frac{1}{2} .
\end{aligned}
\tag{2.79}
$$

Thus, $P(2\pi/3) > 2P(\pi/3)$ in clear contradiction to Bell's inequality. This breaking of the inequality is seen also in Fig.(2.3.2).

Measured frequencies
To get a better understanding of the physical content of the Bell inequality we will discuss its

meaning in the context of spin measurements performed on the two particles, where a series of results, either *spin up* or *spin down*, is registered for each particle. We are particularly interested in correlations between the results of pairs of entangled particles.

Let us therefore consider a spin measurement experiment where the orientations of the two measurement devices are set to fixed directions. We assume N pairs of entangled particles are used for the experiment. The results are registered for each pair, and we denote by n_+ the number of times where the two particles are registered both with *spin up* or both with *spin down*. The number of times where one of the particles are registered with *spin up* and the other with *spin down* we denote by n_-. Clearly $n_+ + n_- = N$.

Spin measurements

Series I: $\Theta_A = \Theta_B = 0$, $n_+ = 0$

S_A + − + + − + − − + + − +

S_B − + − − + − + + − − + −

Series II: $\Theta_A = 0$, $\Theta_B = \pi/3$, $n_+ = 3$

S_A + − + + − + − − + + − +

S_B − + ⊕ − + − ⊖ + ⊕ − + −

Series III: $\Theta_A = -\pi/3$, $\Theta_B = 0$, $n_+ = 3$

S_A ⊖ − + ⊖ − + − − ⊖ + − +

S_B − + − − + − + + − − + −

Series IV: $\Theta_A = -\pi/3$, $\Theta_B = \pi/3$, $n_+ = 4$

S_A ⊖ − + ⊖ − + − − ⊖ − − +

S_B − + ⊕ − + − ⊖ + ⊕ − + −

Figure 2.4: Four series with results of spin measurement on pairs of entangled particles (S_A and S_B). *Spin up* is represented by + and *spin down* by - . *Series I* is considered as the result of a real experiment with alignment of the directions of the measuring devices M_A and M_B. A strict anti-correlation between the result is observed for each pair. *Series II* gives the hypothetical results that would have been obtained in the same experiment if the direction of M_B were tilted relative to M_A. About $1/4$ of the spins S_B would be flipped. These are indicated by red circles. This would lead to four pairs with correlated spins, indicated by green lines. *Series III* is a similar hypothetical situation with M_A rotated. Finally *Series IV* gives the results if both M_A and M_B were rotated. In one of the pairs both spins are flipped relative to *Series I*. This reduces the number of correlated spins relative to the sum of correlated pairs for *Series II* and *Series III*.

The functions $E(\theta)$ and $P(\theta)$ are for large N approximated by the frequencies

$$
\begin{aligned}
E(\theta)_{exp} &= \frac{n_+ - n_-}{N}, \\
P(\theta)_{exp} &= \frac{1}{2}(E(\theta)_{exp} + 1) = \frac{n_+}{N}.
\end{aligned}
\tag{2.80}
$$

Thus $P(\theta)$ represents the probability for having the same result (both *spin up* or both *spin down*) for an entangled pair, as already mentioned.

Let us first focus at the list of results called *Series I* in Fig.(2.3.2). We consider this list as the outcome of a series of measurements for a setup where the directions of the two measuring devices are aligned ($\theta_A = \theta_B = 0$). Each pair of entangled particles is separated in a *particle A* whose spin is measured in M_A and a *particle B* whose spin is measured in M_B.

The list indicates, as we should expect, that the results of measurements at M_A, when considered separately, correspond to *spin up* and *spin down* with equal probability. The same is the case for the measurements at M_B. The list also reveals the strict *anti-correlation* between measurements on each entangled pair. We write the result as

$$
P(I)_{exp} = \frac{n_+(I)}{N} = 0.
\tag{2.81}
$$

Instead of performing a new series of measurements we next make some theoretical considerations based on the assumption of *locality*. We ask the question: *What would have happened if in the experiment performed the spin detector M_B had been rotated to another direction* $\theta_B = \pi/3$. Locality is now interpreted as meaning that if M_B were rotated, that could in no way have influenced the results at M_A. The series of measurements S_A would therefore have been unchanged. The results S_B would, however, have to change since the results for $\theta \neq 0$ are not strictly anti-correlated. Thus, some of the results would be different compared to those listed in Series I. Quantum mechanics in this case predicts $P(\pi/3) = 1/4$. In the list of Series II we have indicated a possible outcome consistent with this, where

$$
P(II)_{exp} = \frac{n_+(II)}{N} = \frac{3}{12}.
\tag{2.82}
$$

Since we have a symmetric situation between M_A and M_B we clearly could have used the same argument if M_A were rotated instead of M_B. This situation is shown in Series III, where the series of measurements S_B is left unchanged, but some of the results of S_A are changed. In this case we have chosen $\theta_A = -\pi/3$. We also in this list have

$$
P(III)_{exp} = \frac{n_+(III)}{N} = \frac{3}{12},
\tag{2.83}
$$

consistent with the predictions of quantum mechanics.

Finally, we combine these two results to a two-step argument for what would have happened if both measurement devices had been rotated, to the directions $\theta_A = -\pi/3$ and $\theta_B = \pi/3$. We can start with either Series II or Series III. Rotation of the second measuring device would then lead to a change of about $1/4$ of the results in the series that was not changed in the first step. The total number of correlated pairs would now be

$$
n(IV) = n(II) + n(III) - \Delta.
\tag{2.84}
$$

In this expression $n(II)$ is the number of flips in the result of S_B and $n(III)$ the number of flips in the results of S_A. Such flips would change the result from *anti-correlated* to *correlated*. However that would be true only if *only one* of the spin results of an entangled pair were flipped. If both spins were flipped we would be back to the situation with *anti-correlations*. Thus the number of correlated pairs would be equal to the sum of the number of flips in each series minus the number where two flips is applied to the same pair. The last number is represented by Δ in the formula (2.84). In the result of Series IV this is represented by the result

$$
\begin{aligned}
P(IV)_{exp} \quad &= \quad \frac{n_+(IV)}{N} \quad = \quad \frac{4}{12} \\
< \quad P(II)_{exp} \quad &+ P(III)_{exp} = \frac{6}{12}\,.
\end{aligned}
\tag{2.85}
$$

The subtraction of Δ in (2.84) explains the Bell inequality, which here takes the form

$$
P(2\pi/3) \leq 2P(\pi/3)\,.
\tag{2.86}
$$

The results extracted from the series of spin measurements listed in Fig.(2.3.2) gives (2.85) in accordance with the Bell inequality (2.86).

The arguments given above, which reproduce Bell's inequality, may seem to implement the assumption of *locality* in a very straight forward way: *Changes in the set up of the measurements at M_A cannot influence the result of measurements performed at M_B* when these devices are separated by a large distance. The predictions of quantum mechanics, on the other hand, are not consistent with the results of these hypothetical experiments, and *real experiments* have confirmed quantum mechanics rather than the Bell inequality.

So is there anything about the arguments given for the hypothetical experiments that indicates that they may be wrong? We note at least one disturbing fact. Only one of the four series of results in Fig.(2.3.2) can be associated with a *real* experiment. The others must be based on assumptions of what would have happened if *the same series of measurements* would have been performed under somewhat different conditions. This is essential for the result. If the four series of measurements should instead correspond to *four different (real) experiments* the situation would be different. We then would have no reason to believe that the series of results S_A would be preserved when going from I to II or S_B when going from I to II. Each series would in this case rather give a new (random) distribution of results for particle A (or particle B).

In any case the conclusion seems inevitable, that quantum mechanics are in conflict with *Einstein locality*, *i.e.*, with the basic assumptions that lead to the Bell inequality. Does this mean that *some kind of influence* is transmitted from A to B when measuring on particle A, even when the two particles are very far apart?

Chapter 3

Quantum physics and information

In recent years there has been an increasing interest in questions concerning the relation between quantum physics and information theory. The present understanding is that the characteristic features of quantum physics that distinguishes it from classical physics, namely quantum interference in general and quantum entanglement in particular, creates the physical foundation for an approach to communication and to processing of information that is qualitatively different from the traditional one. At present there is only a partial understanding of this new approach, but the belief of many physicists is that a new type of *quantum information theory* should be developed as an alternative to classical information theory. This belief is supported by the discovery of algorithms that could speed up the computation of certain types of mathematical problems in a *quantum computer* and by the development of methods for secure and efficient communication by use of *entangled qubits*.

The development of this new approach to information and communication poses important challenges to the manipulation of quantum systems. This is so since *quantum coherence* is important for the methods to work, and in a system with many degrees of freedom *decoherence* will under normal conditions rapidly destroy the important quantum correlations. The very difficult challenge is to create a quantum system where, on one side, the quantum states are effectively protected from outside disturbances and, on the other side, the variables can rapidly be addressed and manipulated in a controlled way in order for the system to perform the task in question.

In this chapter we focus on some basic *theoretical* elements in this new approach to physics and information, while for the discussion of the present status of *implementations* of the ideas on physical systems we refer to several recent books on the subject. We first focus on an example of how quantum physics admits the possibility of acquiring information in a radically new way through an *interaction-free measurement*. We then proceed to study how *qubits* can replace *bits* as the fundamental unit of information.

3.1 An interaction-free measurement

The usual picture of measurements performed on a quantum system is that they involve a non-negligible, minimal disturbance quantified by Planck's constant. If photons are used to examine a physical object, the minimal disturbance corresponds to letting a single photon interact with

the system. The energy of the photon may be made small, but since this means making the corresponding wave length long, one will thereby loose resolution. Thus if a certain resolution is required, a minimal energy has to be carried by the photon and this gives rise to a finite perturbation of the object. The picture of measurements in classical theory is different. There the energy that is carried by light of a given frequency can be made arbitrarily small by reducing the amplitude, and therefore there is no lower limit to how much an (idealized) measurement will have to disturb the object studied.

However, this is not the complete picture. Quantum mechanics opens up the possibility for other types of more "intelligent" interactions than the direct "mechanical" interaction between the object and the measuring apparatus. With the use of quantum superposition (or interference) certain types of measurements can be performed which involve *no mechanical interaction* with the object. A particular example is discussed here.[1]

Let us assume that a measurement should be performed in order to examine whether or not an object is present within a small transparent box. If the object is not present the box is transparent to light, while if it is present the box is not transparent, since the object will absorb or scatter the photon. Let us further assume that measurements are performed with single photons.

A direct measurement would be to send a photon through the box, and to register whether or not the photon is transmitted through the box. This would clearly give the information required. If the object is present, the information about this situation is achieved by a direct (mechanical) interaction between the photon and the object. Apparently this is the least interaction with the object that can be made in order to detect its presence. However, this is not the correct conclusion to draw, as is outlined in Fig.3.1. The figure shows a Mach-Zehnder interferometer where an incoming photon can follow two different paths and eventually be registered in of the two detectors. We first consider the case where both paths are open. The photon will first meet a beam splitter, that with equal probability will direct the photon horizontally into the lower path or vertically into the upper part. On both paths the photon will meet a mirror that redirects it towards a second (50/50) beam splitter. Here the two components of the photon wave functions will meet and form a superposition that can either direct it in the horizontal direction towards a detector A or vertically towards a detector B.

If we neglect the coherence effect and assume an incoherent scattering of the photon by the beam splitters, with equal probability in the two directions, then we expect that the probability for detecting the photon by detector A to be 50%. With the same probability the photon will be detected by B. However, the quantum description of the transmission of the photon through the apparatus implies that the photon at intermediate times is not located (with a certain probability) on one of the paths, but is rather in a *superposition* of being on each of the two paths. This means that the two signals following the upper and lower paths will interfere when they meet at the second beam splitter. In the following we will assume the experimental setup to be adjusted to a situation where the interference acts constructively for a photon directed towards detector A and destructively for a photon directed towards detector B. Thus, the probability for detecting the photon by A is 1 and the probability for detecting the photon by B is 0. This will be the case as long as both paths are open. Clearly, if one of the paths is closed, the photons

---

[1]This example is taken from A.C. Elitzur and L. Vaidman, Found. Phys. **23** (1993) 987.
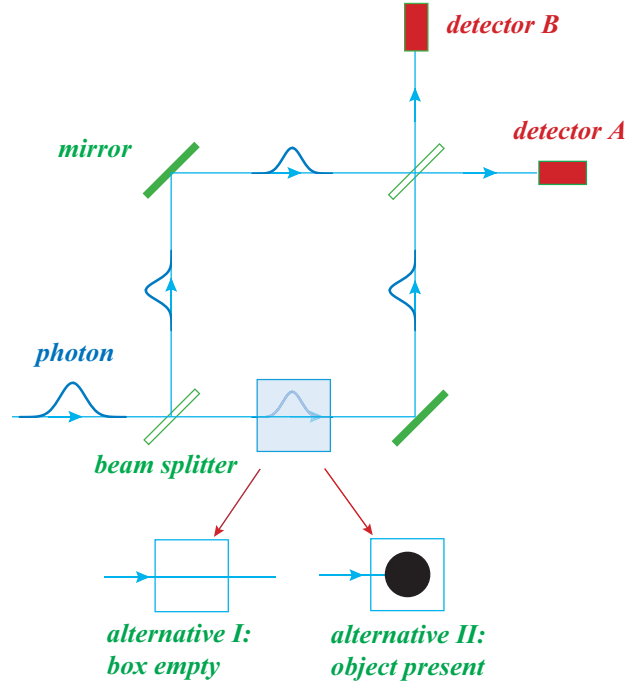
Figure 3.1: The set up of a single-photon measurement to detect the presence of an object in a transparent box by use of a Mach-Zehnder interferometer. A photon is sent through a beam splitter that directs it, in a superposition, either in the horizontal or vertical direction. On the lower path the box is placed. Therefore, if the object is present, the lower path is blocked, while if it is not present both paths are open. The two paths meet again at a second beam splitter and then the photon is directed towards one out of two possible detectors. The interferometer is arranged so that if both paths are open, destructive interference prevents the photon from reaching detector B. Thus, if the photon is registered in B this provides the information that the lower path is blocked and that the object is present in the box.

that get through are instead detected with equal probability by A and B.

To describe the situation more formally let us denote the state of a photon moving in the horizontal direction by $|0\rangle$ and a photon moving in the vertical direction by $|1\rangle$. (With this notation we do not make any distinction between where in the apparatus the photon is.) The action of the beam splitters on a photon is described by the mapping

$$
\begin{aligned}
|0\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\,, \\
|1\rangle &\rightarrow \frac{1}{\sqrt{2}}(|1\rangle + i|0\rangle)\,,
\end{aligned}
\tag{3.1}
$$

while the action of the mirrors is given by

$$
\begin{aligned}
|0\rangle &\rightarrow i|1\rangle \ \ (\text{lower mirror})\,, \\
|1\rangle &\rightarrow i|0\rangle \ \ (\text{upper mirror})\,.
\end{aligned}
\tag{3.2}
$$

The lengths of the paths are assumed to be the same, so that the difference in phase acquired by the photon on the two paths is only due to the phase shifts produced at the mirrors and beam

splitters. The photon is subject to a series of transformations of the form (3.1) and (3.2) in the interferometer.

Let us consider the mapping from the incoming state to the outgoing state (before detection) when only the upper path is open,

$$|0\rangle \rightarrow i|1\rangle \rightarrow -|0\rangle \rightarrow -\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle),$$ (3.3)

where the changes in the state take place at the mirror and at beam splitters on the path. Since we are interested in the state of a photon that exits from the last beam splitter, the intermediate states have been normalized to 1, thus neglecting the probability that the photon is absorbed on the lower path. The interesting point to note is that the final state is a superposition with equal probability for exit in the horizontal and vertical direction.

If both paths are open the mapping from the initial to the final state is instead

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \rightarrow \frac{1}{\sqrt{2}}(-|0\rangle + i|1\rangle) \rightarrow -|0\rangle,$$ (3.4)

and we note that only one component survives. The photon will exit (with probability 1) in the horizontal direction.

We will now turn to the original problem and consider the the situation where the box, which either is empty or not empty, is placed on one of the paths of the photon. The intention is to send one photon through the interferometer in order to investigate whether the box is empty or not. We note that if the box is empty we have the situation where both photon paths are open. If the box is not empty only one of the paths will be open.

Let us consider the possible outcomes of the experiment where a photon is sent through the interferometer:

1. There is no photon registered in any of the two detectors.
   We conclude that the object is present, the photon has interacted with the object and has not got through to the detectors.

2. The photon is registered by detector A.
   The result is inconclusive. Whether the object is there or not there is always a chance for the photon to be detected by A. The experiment has to be repeated.

3. The photon is registered by detector B.
   We conclude that the object is present, since the probability of detecting the photon by B with both paths open is 0.

Of the possible outcomes we focus on 3, which is the interesting one. In this case the presence of the object has been detected without any interaction with it. This is so since the *detection* of the photon implies that no interaction has taken place. A natural explanation seems to be that the photon has followed the upper (open) path, but that the detection of the photon provides information about the lower path (that it is closed). This result depends crucially on the possibility of quantum superposition. In a classical theory this would not happen. Note however, the curious fact that when the object blocks the path, in reality no superposition takes place. The result of the measurement will be the same as in a classical theory with a certain

probability for the photon to follow the upper path. It is *our knowledge* of the possible outcomes when both paths are open that allows us to draw the conclusion that one path is closed.

In conclusion this thought experiment shows that a direct interaction in a measurement is not always needed. But there has to be a *possibility* for the interaction to take place. A superposition between two states, where one of them interacts (when the object is there) and the other does not, is an important ingredient in the set up.

The example discussed here shows that alternative ways to collect information with quantum mechanical methods is a possibility. Quantum coherence or superposition is important for such methods to work. Also the importance of using "intelligent ways" to address the measuring problem, instead of a (naive) direct measurement is emphasized.

3.2 The No-Cloning Theorem

In classical physics there is no theoretical lower limit to how much the extraction of information about the system will disturb the system. Thus, the information can be aquired with measurements that in principle introduce only arbitrarily small changes in the state of the system. As a consequence of this, the state of a classical system can, in principle, be copied with arbitrarily good precission to another classical system. We may refer to this as cloning of the classical state.

However, in quantum physics such a faithful copying is not possible, and this impossibility is expressed in what is known as the *No-Cloning Theorem.* We shall now look at why this is impossible, but before doing so, let us first note that in some special situations quantum states can in fact be duplicated. Thus, in the following two cases we can, in principle, produce or aquire a multitude of copies of the same quantum state.

a) When we know how the original state has been prepared and can repeat this operation on other quantum systems.

b) When nature offers a number of identical systems, such as for example the hydrogen atom, where the energy eigenstates of the different systems are the the same.

These cases do not contradict the No-Cloning Theorem, which we may phrase in the following way

Assume a physical system A is in an unknown quantum state ψ. It is impossible to reproduce the same state ψ in a second system B, while preserving in system A the state ψ.

To make the situation somewhat more specific, let us assume that we have two identical systems A and B available, which for simplicity may be two two-level systems. Let us assume that system A has been prepared in an unknown linear superposition of two basis states, $|\psi\rangle = \alpha|1\rangle + \beta|2\rangle$. For simplicity we also assume the the basis states to be degenerate in energy, so there is no time evolution of the state. The challenge is now to prepare the second system B in the same state $|\psi\rangle$, without knowing the coefficients α and β, while leaving the state of A unchanged. This could possibly be done by letting the two systems interact directly, or by letting them interact indirectly through an additional third system. However, the theorem states that this is impossible to do.

To show this is in fact rather simple, the result follows from the linearity of quantum theory. Thus let us consider the two systems initially to be disconnected and to be in a tensor product

state $|\psi\rangle \otimes |\phi\rangle$. To change the state of system B we assume that there follows an intermedi-
ate, interacting phase where the two systems, possibly together with an auxillary third system,
forms an interacting composite system. After this intermediate interacting period the two sys-
tems A and B should end up in a new product state $|\psi\rangle \otimes |\psi\rangle$ if the cloning has been successful.
From general quantum theory we know that there will be a linear, unitary transformation $\hat{U}$ that
relates the final state to the initial state,

$$|\psi\rangle \otimes |\psi\rangle = \hat{U} |\psi\rangle \otimes |\phi\rangle . \tag{3.5}$$

The form of this equation is however non-linear, as a direct consequence of the state $|\psi\rangle$ in the
final state of the composite system. To show this explicitly, let us duplicate two different states
$|\psi_1\rangle$ and $|\psi_2\rangle$,

$$|\psi_1\rangle \otimes |\psi_1\rangle = \hat{U} |\psi_1\rangle \otimes |\phi\rangle ,$$
$$|\psi_2\rangle \otimes |\psi_2\rangle = \hat{U} |\psi_2\rangle \otimes |\phi\rangle . \tag{3.6}$$

Linearity of the operator $\hat{U}$ can now applied in two different ways, first simply by adding the
two equations, which gives

$$\hat{U} \left(|\psi_1\rangle \otimes |\phi\rangle + |\psi_2\rangle \otimes |\phi\rangle \right) = |\psi_1\rangle \otimes |\psi_1\rangle + |\psi_2\rangle \otimes |\psi_2\rangle , \tag{3.7}$$

and then by applying it to the superposition

$$
\begin{aligned}
\hat{U} \left(|\psi_1\rangle + |\psi_2\rangle \right) \otimes |\phi\rangle) &= \left(|\psi_1\rangle + |\psi_2\rangle \right) \otimes \left(|\psi_1\rangle + |\psi_2\rangle \right) \\
&= |\psi_1\rangle \otimes |\psi_1\rangle + |\psi_2\rangle \otimes |\psi_2\rangle |\psi_2\rangle \otimes |\psi_2\rangle \\
&\quad + |\psi_1\rangle \otimes |\psi_2\rangle + |\psi_2\rangle \otimes |\psi_1\rangle .
\end{aligned}
\tag{3.8}
$$

Since the left-hand sides in the two equations above are identical, the right-hand sides should
also be identical, but they are not. The reason for this contradiction is that the the original
equation (3.5) is intrinsically non-linear.

   We therefore have to conclude that the No-Cloning Theorem follows from one of the basic
principles of quantum theory.

3.3   Quantum teleportation

Is it possible to transfer, in a faithful way, an unknown quantum state from one physical system
to another? The answer to this question is affirmative, but to avoid conflict with the No-Cloning
theorem, this can only happen provided that the quantum state of the first system is changed.
We will show that this can be done, not in the form of a general proof, but by way of a simple
example.

   Consider the following situation. A person A is in the possession of two two-level systems,
which we label 1 and 2, while a second person B is in possession of a third two-level system,
denoted 3. System 1 is in an unknown state

$$|\chi\rangle_1 = a|+\rangle_1 + b|-\rangle_1 , \tag{3.9}$$

where the state has been expanded in the two basis state $|+\rangle$ and $|-\rangle$. Person A now wants to transfer this state from system 1 to system 3. A and B have in advance prepared the two systems 2 and 3 in a maximally entangled Bell state,

$$|\psi^-\rangle_{23} = \frac{1}{\sqrt{2}}(|+-\rangle_{23} - |-+\rangle_{23}), \quad |+-\rangle \equiv |+\rangle \otimes |-\rangle. \tag{3.10}$$

This state, of course, does not contain any information about the unknown state $|\chi\rangle$. The state of the full system, consisting of systems 1, 2 and 3, then is

$$|\psi\rangle_{123} = |\chi\rangle_1 \otimes |\psi^-\rangle_{23}. \tag{3.11}$$

This state may be re-expressed in terms of the Bell states of the composite system 12. These states are

$$|\psi^\pm\rangle_{12} = \frac{1}{\sqrt{2}}(|+-\rangle_{12} \pm |-+\rangle_{12}),$$

$$|\phi^\pm\rangle_{12} = \frac{1}{\sqrt{2}}(|++\rangle_{12} \pm |--\rangle_{12}). \tag{3.12}$$

The inverted expressions are

$$|\pm\mp\rangle_{12} = \frac{1}{\sqrt{2}}(|\psi^+\rangle_{12} \pm |\psi^-\rangle_{12}),$$

$$|\pm\pm\rangle_{12} = \frac{1}{\sqrt{2}}(|\phi^+\rangle_{12} \pm |\phi^-\rangle_{12}). \tag{3.13}$$

Inserted in the expression (3.11), we find

$$
\begin{aligned}
|\psi\rangle_{123} &= \frac{1}{\sqrt{2}}(a|+\rangle_1 + b|-\rangle_1) \otimes (|+-\rangle_{23} - |-+\rangle_{23}) \\
&= \frac{1}{\sqrt{2}}(a|++-\rangle_{123} - b|--+\rangle_{123} - a|+-+\rangle_{123} + b|-+-\rangle_{123}) \\
&= \frac{1}{2}(a(|\phi^+\rangle_{12} + |\phi^-\rangle_{12}) \otimes |-\rangle_3 - b(|\phi^+\rangle_{12} - |\phi^-\rangle_{12}) \otimes |+\rangle_3 \\
&\quad -a(|\psi^+\rangle_{12} + |\psi^-\rangle_{12}) \otimes |+\rangle_3 + b(|\psi^+\rangle_{12} - |\psi^-\rangle_{12}) \otimes |-\rangle_3 \\
&= \frac{1}{2}(|\phi^+\rangle_{12} \otimes (a|-\rangle_3 - b|+\rangle_3) + |\phi^-\rangle_{12} \otimes (a|-\rangle_3 + b|+\rangle_3)) \\
&\quad -|\psi^+\rangle_{12} \otimes (a|+\rangle_3 - b|-\rangle_3) - |\psi^-\rangle_{12} \otimes (a|+\rangle_3 + b|-\rangle_3)). \tag{3.14}
\end{aligned}
$$

Let us now define operators that act on a two-level system in the following way

$$
\begin{aligned}
V_1|\pm\rangle &= -|\pm\rangle, \\
V_2|\pm\rangle &= \mp|\pm\rangle, \\
V_3|\pm\rangle &= \mp|\mp\rangle, \\
V_4|\pm\rangle &= |\mp\rangle. \tag{3.15}
\end{aligned}
$$

With the help of these the state of the full system can be written

$$
\begin{aligned}
|\psi\rangle_{123} \;\; = \;\; & \frac{1}{2}(|\phi^+\rangle_{12} \otimes (\hat{V}_3|\chi\rangle_3) + |\phi^-\rangle_{12} \otimes (\hat{V}_4|\chi\rangle_3)) \\
& + |\psi^+\rangle_{12} \otimes (\hat{V}_2|\chi\rangle_3) + |\psi^-\rangle_{12} \otimes (\hat{V}_1|\chi\rangle_3)) \,,
\end{aligned}
\tag{3.16}
$$

which gives the state vector expressed in the Bell basis of the subsystem 12. The four Bell states are orthogonal and may be distinguished by measurement of an observable with these as eigenvectors and with different eigenvalues for the four states. Assume A is performing such a measurement on subsystem 12. The state of system 12 is then projected to one of the Bell states, and the measurement result gives A the information about which state. The entanglement with system 3 implies that in the measurement this subsystem is projected to one of the states $\hat{V}_k|\chi\rangle_3$ where the value of k $(= 1, , 2, 3, 4)$ is linked to the surviving Bell state.

A knows the value of k from the result of the measurement and conveys this information to B, who performs the transformation $\hat{V}_k$ on system 3. Since $\hat{V}_k^2 = \mathbb{1}$ for any k, this means that the state of system 3 is changed to $|\chi\rangle_3$. The final state of the full system 123 therefore is

$$
|\psi^{fin}\rangle_{123} = |\psi^{Bell}\rangle_{12} \otimes |\chi\rangle_3 \,,
\tag{3.17}
$$

where $|\psi^{Bell}\rangle_{12}$ is one of the four Bell states. This shows that the unknown state $|\chi\rangle$ which was initially in system 1 has now been transferred to system 3. At the same time the entanglement between systems 2 and 3 has been replaced by entanglement between systems 1 and 3.

It is interesting to note that only classical information has been sent from A to B, but this information allows B to reconstruct in system 3 the quantum state that was originally in system 1.

3.4   From bits to qubits

In the classical information theory, developed by C. Shannon, information is quantified in terms of discrete *units of information*. Thus, the elementary information unit is a *bit*, and this is viewed as a function which can take two possible values, normally the numerical values 0 and 1. The information lies in specifying which of the two values to assign to the bit. The idea is that a general message, to an arbitrary good precision, can be expressed in terms of a finite sequence of bits.

The idea of *discretizing information* creates the basis for the general theory of information. But, as we all know, it also creates the basis for a practical approach to information and communication in the form of digital signals. Whereas communication (by telephone, radio or TV) used to be in the form of *analog signals*, presently the use of digital signals are preferred because this admits a more precise determination (and correction) of the information content of the signal.

Information theory can be viewed (and is normally so) as a mathematical discipline. However, from a physics point of view, it is natural to focus on the implementation of the theoretical ideas in terms of physical signals. Thus, the information will normally be coded into signals that are created and manipulated in physical (electronic) devices. They are transmitted by physical mediators (electromagnetic waves or electric signals) and are again manipulated and decoded in (electronic) receivers.

A message consisting of a certain number of bits can be viewed as a *state* of a physical system. With N bits there are 2^N states, which represent all the different messages that can be encoded in N bits. In this picture the factorization of the message into single bits corresponds to a separation of the physical system into N two-state subsystems. Thus, the information unit *bit* corresponds to the *two-state system* as a physical unit. Such a system can be realized in many ways, as a physical system that can easily be switched between two stable states, or an electric signal with two values ("on" or "off"), etc.

The important point to note is that the two-state system considered in this way is a *classical* system. And the interesting question, which has been addressed in recent years, is whether quantum physics should introduce a new picture of the (physical) unit of information. The classical two-state system has its counterpart in the quantum two-state or two-level system, and for the quantum system a new feature is that *coherent superpositions* between states are possible. In the same way as the classical two-state system is associated with a bit of information, the quantum two-level system is now associated with a new information unit, a *qubit*. While the possible values of a bit is restricted to 0 and 1, the qubit takes values in a two-dimensional Hilbert space spanned by two vectors $|0\rangle$ and $|1\rangle$. Thus a general qubit state is

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle \,, \tag{3.18}$$

with α and β as complex coefficients.

Let us therefore assume that a message in this new picture is encoded not in a classical state of a system, but in a quantum state of a finite-dimensional Hilbert space. Such a Hilbert space is unitarily equivalent to a tensor product of N two-level systems, provided we restrict the dimension of the Hilbert space to $M = 2^N$. In this sense we can view the qubit as the elementary building block of information. Note however that the general state is not a product state of the qubits, since also superpositions should be included. This means that the general state involves *entanglement* between the qubits.

Apparently there is much more than one bit of information contained in each qubit, since the qubit states form a continuum that interpolates between the "classical" states $|0\rangle$ and $|1\rangle$. However, one should be aware of the fact that even if more information is contained in the *specification* of the qubit state, all this information cannot be read out by making a measurement on the qubit. This is due to the probabilistic interpretation of the state. One may compare this to a situation with a (classical) probability distribution over the two states 0 and 1. Since the probability distribution depends on a continuous parameter, much more than one bit of information is needed to specify the value of this parameter. However, since each classical two-state system can only be in the states 0 or 1, an ensemble of these systems is needed to provide the information about the probability distribution.

In the case of qubits the situation is somewhat similar, but not completely so. Unlike a classical probability distribution, the superposition of states is a resource that can be exploited for certain types of information processing. This has been demonstrated by specific examples.

3.5 Communication with qubits

New possibilities open up in communication when we can exploit quantum interference and quantum entanglement. We show here a simple example of *dense coding* of information with
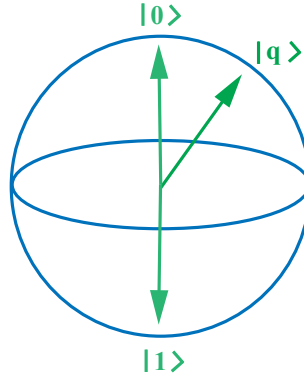
Figure 3.2: The physical states of a qubit can be viewed as points on a sphere. The poles of the sphere correspond to the two classical one-bit states 0 and 1

the help of entangled qubits.

Let us assume that a sender A (often referred to as *Alice*) wants to send a *two-bit* message to a receiver B (referred to as *Bob*). The question that is posed is whether this can be done by transmitting a single qubit, since the claim is that a qubit carries more information than a bit. The apparent answer is no: If Alice prepares the qubit in a pure state and sends it to Bob, he can read out the information by measuring the state in a given basis (corresponding to measuring the spin component in some direction). The result is 0 or 1, where the probability for getting these two results is determined by the decomposition of the prepared state on the two basis states $|0\rangle$ and $|1\rangle$. It seems that the best they can do in order to send the message is to agree on what basis to use. Then Alice can choose between two possible states $|0\rangle$ and $|1\rangle$ and Bob can determine which of the states is chosen by making a measurement in the same basis as used by Alice. But in this way a qubit can communicate only one bit of information.

However, more intelligent ways to do it exist. Let us assume that Alice and Bob in advance have shared a pair of qubits with maximum entanglement. They may for example be in the state

$$|c, +\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) . \tag{3.19}$$

We assume the qubits are kept in a safe way so that the entanglement is kept unchanged until the qubits are used for communicating the message.

The are four different two-bit messages, which written in the standard digital form are denoted 00, 01, 10 and 11. Let us assume that these four messages are associated with four orthogonal states of the two-qubit system in the following way,

$$
\begin{aligned}
00 &\rightarrow & \mathbb{1} \otimes \mathbb{1} \, |c, +\rangle &= \; |c, +\rangle , \\
01 &\rightarrow & i\sigma_z \otimes \mathbb{1} \, |c, +\rangle &= i \, |c, -\rangle , \\
10 &\rightarrow & i\sigma_x \otimes \mathbb{1} \, |c, +\rangle &= i \, |a, +\rangle , \\
11 &\rightarrow & i\sigma_y \otimes \mathbb{1} \, |c, +\rangle &= \; |a, -\rangle .
\end{aligned}
\tag{3.20}
$$

The important point is that all these four states can be created from the original state (3.19) by

transforming only on one of the qubits, as indicated in (3.20). We assume this to be Alice's qubit.

We consider then the following situation: Alice wants to submit one of the four messages 00, 01, 10 or 11. She then performs one of the unitary transformations on her qubit which brings the two-qubit system into the corresponding state, shown in the list (3.20). In the first case (message 00) no change is made to her qubit, in the second case (01) a rotation of π around the z-axis is performed, in the subsequent cases ((10) and (11)) rotations around the y-axis and z-axis are performed. (We here envisage the qubit states as spin states.) We note that in all cases no change is done to the B-qubit (which is not available for Alice), and in all cases the maximal entanglement is kept by transforming the original state into another Bell state.

Alice now transmits her qubit to Bob, who is free to make measurements on *both* entangled qubits. At this point one of the four different (two-bit) messages has been encoded in the two-qubit system, in the form of one of the four orthogonal Bell states. Bob knows that the system is in one of these states, and since he now has both qubits at his disposal, he can perform an measurement that distinguishes between the four states. We do not specify how this should be done, but in principle he can identify an observable with different eigenvalues for the four states, and measure which of the eigenvalues that is realized in the two-qubit system. From the measurement he identifies which of the messages 00, 01, 10 or 11 that Alice is transmitting, and in this way he has obtained the full two-bit message from the measurement.

In this way Alice has managed to transfer the two-bit message to Bob, by encoding the message into a single qubit which, is afterwards transmitted to Bob. There is also a second qubit involved, the qubit belonging to Bob, but this is not affected by the manipulations performed by Alice. Through the entanglement with Alice's qubit it plays an important role, as a resource for transmitting the message, but it is passive in this process until the final stage when Bob makes the measurement to identify which of the four Bell state that is realized in the transmission.

The entanglement is essential for being able to transmit the full message by a single qubit. In fact, if we consider the two qubits separately, they are at all time in the same, maximally mixed state, described by the density matrix

$$\hat{\rho}_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \hat{\rho}_B \,, \tag{3.21}$$

which contains no information about the message that is transferred. This means that *all the information is contained in the entanglement between the two qubits*, since no information lies in the reduced states of the two qubits. From this we also note the important point: The message can only be read by the receiver who has the second qubit of the entangled pair. In this way the message is protected from others than Bob. This principle of protecting information by encoding the information into entangled pairs of qubits is the basis for *quantum cryptography*, which is a field of research that has been rapidly developed in recent years.

3.6 Principles for a quantum computer

The probably most interesting suggestion for applications of quantum physics to information technology is in the form of *quantum computers*. The idea is that a computer built on quantum principles will manipulate information in a qualitatively different way than a classical computer, and thereby will be able to solve certain types of problems much more efficiently.

A specific type of computational problem that is interesting for physicists is the *simulation of quantum systems*. Today numerical solutions of physics problems are important for research in almost any field of physics, but the capacity of present day computers gives a clear limitation to the size of the problem that can be handled. This has to do with how the complexity of the computation scales with the size of the system. For a quantum system with N degrees of freedom, the Hilbert space dimension is m^N, if each degree of freedom is described by an m-dimensional space. This means that m^N complex parameters are needed to specify a Hilbert space vector, and the number of computational variables therefore grows exponentially with N. This gives a strong limitation for the simulations of quantum systems on a classical computer.

The idea is that a quantum computer should *work as a quantum mechanical system*, with the computation performed by unitary transformations on superpositions of qubit states. For the simulation of quantum systems there is an obvious gain, since the number of qubits needed to represent the wave function grows linearly with the number of degrees of freedom N rather than exponentially. In addition to the simulation of quantum systems there are also certain other types of mathematical problems that can be solved more efficiently with the use of quantum superposition. Two algorithms that have gained much interest are the *Shor algorithm* for factorizing large numbers and the *Grover algorithm* for making efficient search through data bases.

The typical feature of a quantum computer is to work with superpositions of (qubit) states. From a computational point of view this can be seen as new type of *(quantum) parallel computing*. In the picture of path integrals we may view a classical computation as a (classical) path, where each logical operation corresponds to making a (new) direction for the path. Parallel processing in this picture corresponds to working simultaneously with several paths in the "space of logical operations". In the quantum computer many paths are, in a natural way, involved at the same time in the form of a superposition of states, and the final result is obtained by quantum mechanical interference between contributions from all the (classical) paths. Clearly, if a problem should be solved much more efficiently on a quantum computer than on a classical computer, superposition of states has to be used extensively in the computation. This means that the qubits will be highly entangled during the computation. The serious challenge for constructing a quantum computer is therefore to be able to preserve and operate on such highly entangled states.

3.6.1   A universal quantum computer

The idea of a universal quantum computer is similar to that of a universal classical computer. Thus, a universal quantum computer is designed to solve general types of problems by reducing the computation to elementary qubit operations. This means that the input wave function is encoded in a set of (input) qubits, and the computational program acts on these by performing *logical operations* in the form of unitary transformations on the qubits. A standard set of unitary *one-qubit* and *two-qubit* operations is used, where each operation is performed at a logical gate. Together the logical gates form a *computational network* of gates.

In Fig. 3.6.1 a schematic picture of a universal quantum computer is shown. The input data are encoded by preparation of an input quantum state. The computer program acts on
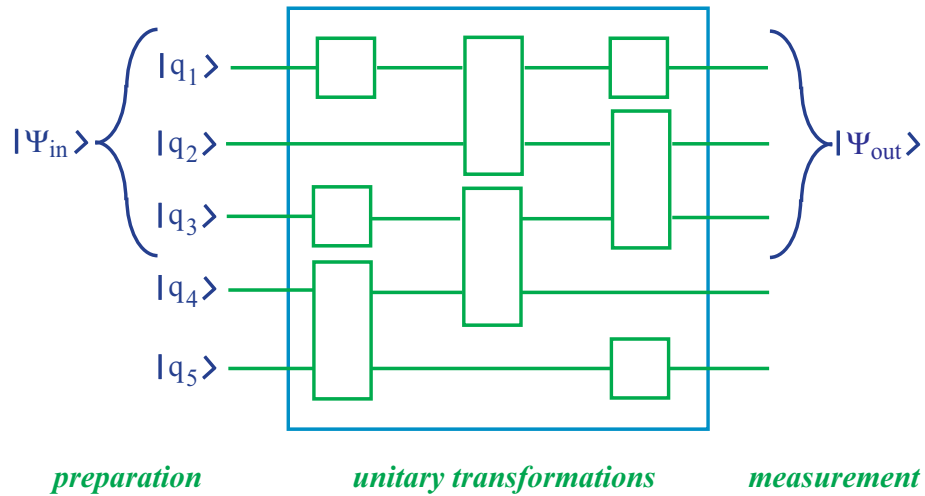
Figure 3.3: A schematic picture of a universal quantum computer. An input state is prepared as a quantum state of a set of qubits. A network of logical gates, that perform one-qubit and two-qubit transformations, operates on the input qubits (and a set of additional work qubits) to produce an output state. The result of the computation is read out by measuring the state of each output qubit. In the diagram the horizontal lines represent the qubits and the boxes represent the gates or logical operations performed on the qubits.

the input state and by a unitary transformation produces an output state where the result of the computation is read out in a quantum measurement. The computational task is specified partly by the unitary transformation performed on the input state and partly on how the measurement is performed on the output state.

This picture of a quantum computer is quite analogous to that of a classical computer, where bits of information are processed at logical gates that together form a logical network. The main difference is that in the quantum computer the information is processed as quantum superposition between states. And the computation is *reversible* since the unitary transformations are all invertible. This is different from a (standard) classical computer where some of the standard logical operations are irreversible in the sense that the mapping between the input state and the output state is not one-to-one.

A universal set of logical gates

The idea of a universal quantum computer is based on the possibility of reducing any unitary transformation that acts on the quantum states of an N qubit system to a repeated operation of a small number of standard one-qubit and two-qubit transformations. We will here only outline how such a reduction can be performed. It involves the following steps,

1. A unitary transformation acting on a finite-dimensional Hilbert space can be factorized in terms of *two-level* unitary transformations. These transformations act on two-level subsystems spanned by two orthonormalized vectors of a chosen basis in the full Hilbert

space,

$$\hat{U} = \prod_n \hat{U}(i_n, j_n), \tag{3.22}$$

where i_n and j_n denotes the basis states affected by the nth transformation

2. A two-level unitary transformation acting between basis states of the N-qubit Hilbert space can be factorized in terms of a set of one-qubit and two-qubit operations. With the number of terms in the factorization restricted to a finite number, a general unitary transformation can only be represented in an approximate form. Thus, the continuous parameters of the unitary transformation are replaced by a set of discrete parameter values.
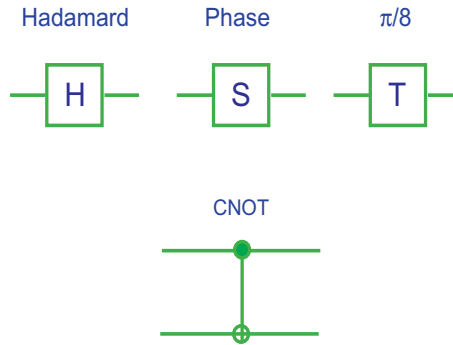

Figure 3.4: Symbolic representation of logical gates. In the representation of the CNOT gate the upper line corresponds to the control qubit and the lower line to the target qubit

The following one-qubit and two-qubit transformations give an example of a universal set of qubit operations,

1. *The Hadamard transformation.*
   This is a single-qubit transformation defined by the operations on the qubit states in the following way,

$$\hat{H}\,|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$
$$\hat{H}\,|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{3.23}$$

In matrix form this is,

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{3.24}$$

with $|0\rangle$ corresponding to the to the upper place and $|1\rangle$ to the lower place of the matrix.

2. *The Phase transformation.*
   This is also a single-qubit transformation, defined by

$$\begin{aligned} \hat{S}\,|0\rangle &= |0\rangle\,, \\ \hat{S}\,|1\rangle &= i\,|1\rangle\,, \end{aligned} \tag{3.25}$$

which in matrix form is

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{3.26}$$

3. *The $\pi/8$ transformation.*
   This is the third single-qubit transformation. It is defined by

$$\begin{aligned} \hat{T}\,|0\rangle &= |0\rangle\,, \\ \hat{T}\,|1\rangle &= e^{i\pi/4}|1\rangle\,, \end{aligned} \tag{3.27}$$

with the matrix form

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}. \tag{3.28}$$

4. *The $CNOT$ transformation.*
   This is a two-qubit transformation, which is a quantum version of a *controlled not* operation. It is defined by

$$\begin{aligned} \hat{C}_{NOT}\,|0\rangle \otimes |0\rangle &= |0\rangle \otimes |0\rangle\,, \\ \hat{C}_{NOT}\,|0\rangle \otimes |1\rangle &= |0\rangle \otimes |1\rangle\,, \\ \hat{C}_{NOT}\,|1\rangle \otimes |0\rangle &= |1\rangle \otimes |1\rangle\,, \\ \hat{C}_{NOT}\,|1\rangle \otimes |1\rangle &= |1\rangle \otimes |0\rangle\,. \end{aligned} \tag{3.29}$$

We note that the state of the first qubit is left unchanged. It acts as a *control qubit* on the second qubit: If the first qubit is in the state $|0\rangle$ the second qubit is left unchanged. If the first qubit is in the state $|1\rangle$ the second qubit switches states, $|0\rangle \leftrightarrow |1\rangle$. With the basis vectors of the product space written in matrix form,

$$|0\rangle \otimes |0\rangle \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\,, \quad |0\rangle \otimes |1\rangle \to \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\,,$$

$$|1\rangle \otimes |0\rangle \to \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\,, \quad |1\rangle \otimes |1\rangle \to \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\,, \tag{3.30}$$

the $CNOT$ operation corresponds to the following 4×4 matrix,

$$C_{NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.31}$$

The definitions above give the action of qubit operations on a set of basis vectors for the single-qubit and-two qubit spaces. With specification of the qubits involved, the action of these operators on a complete set of basis vectors for the full N- qubit space is determined, and thereby the action of the operators on any state vector, by the principle of linear superposition.

3.6.2   A simple algorithm for a quantum computation

As already mentioned, certain algorithms have been designed for solving mathematical problems more efficiently on a quantum computer than can be done on a classical computer. A famous example is *Shor's algorithm*, that addresses the question of how to factorize large numbers. This is a hard problem on a classical computer, since the computational time used to factorize large numbers will in general increase exponentially with the number of digits. (This fact is the basis for making use of factorization of large numbers as keys in cryptographic schemes.) The demonstration by P. Shor that the factorization can be done more efficiently on a quantum computer is one of the reasons for the boost of interest for quantum computing over the last decade.

In this section we will focus on a simpler algorithm introduced by D. Deutsch some time ago. The intention is to use this algorithm as a simple demonstration of how superposition makes it possible to address certain types of problems more efficiently.

The problem addressed is to study *one-bit functions*. Such a function gives a mapping

$$f(x): \ \{0,1\} \to \{0,1\} \,. \tag{3.32}$$

We will need to add these functions, and note that addition between one-bit numbers can be deduced from ordinary addition if the result is defined *modulo 2*. Thus, the explicit addition rule is

$$0+0 = 0 \,, \ \ 0+1 = 1+0 = 1 \,, \ \ 1+1 = 0 \,. \tag{3.33}$$

We also note that there exist four different one-bit functions (3.32),

$$
\begin{aligned}
f_a &: \{0,1\} \to \{0,1\} \,, \quad f_b : \{0,1\} \to \{1,0\} \,, \\
f_c &: \{0,1\} \to \{0,0\} \,, \quad f_d : \{0,1\} \to \{1,1\} \,,
\end{aligned}
\tag{3.34}
$$

where the sets of input values and output values are here considered as *ordered sets*. We distinguish between two types of functions, the set $A = \{f_a, f_b\}$ are the *invertible* functions, while the set $B = \{f_c, f_d\}$ are *constant* functions.

Let us assume that a function $f(x)$ is known only *operationally*, *i.e.*, in a classical computation one assigns to the input variable x one of the two possible values 0 and 1, from which one of the two possible output values $f(0)$ and $f(1)$ are produced. (We may think of the function as a simple *computational program* that produces the output result from a given input.) Initially the function $f(x)$ is not known, and as a concrete problem we want to determine whether the function belongs to set A or set B. Thus, we do not seek a detailed knowledge of the function. However, by classical computations, as explained above, we clearly have to use two input values $x = 0$ and $x = 1$ in order to determine which set the function belongs to, which means that a complete determination of the function has to be done.

The intention is to show that a quantum computation, which makes use of superpositions, can distinguish between set A and set B in one operation. We first have to discuss how such a quantum computation can be implemented. Let us consider the possibility of simple extension of the classical operation to a linear transformation, that acts on a qubit in the following way,

$$T_f : |q\rangle = \alpha|0\rangle + \beta|1\rangle \to \alpha|f(0)\rangle + \beta|(f(1)\rangle\,. \tag{3.35}$$

This is considered as a single (qubit) operation even if both results $f(0)$ and $f(1)$ are present in the output state. However, for the non-invertible functions (f_c and f_d) the operation T_f is non-invertible and therefore non-unitary. In order to represent f as a unitary transformation we therefore modify the one-bit operation T_f to a unitary two-qubit operation U_f in the following way

$$U_f : |x\rangle \otimes |y\rangle \to |x\rangle \otimes |y + f(x)\rangle\,, \tag{3.36}$$

where $|x\rangle$ and $|y\rangle$ denote standard qubit basis states ($|0\rangle$ and $|1\rangle$). We note that the first qubit ($|x\rangle$) is left unchanged by the transformation; it acts as a control qubit on the second qubit. Thus, if $f(x) = 0$ the state of the second qubit is left unchanged, if $f(x) = 1$ the state of the second qubit is flipped ($0 \leftrightarrow 1$). It is straight forward to check that (3.36) defines a unitary transformation.
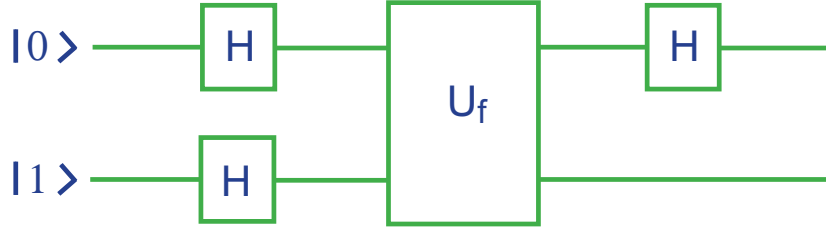

Figure 3.5: Schematic representation of qubit transformations for Deutsch's algorithm. The first qubit is prepared in the $|0\rangle$ state and the second qubit in the $|1\rangle$ state. They are both transformed by a *Hadamard* operation before a two-qubit operation is performed. This operation, U_f, is determined by the (unknown) one-bit function $f(x)$. Finally a second *Hadamard* transformation is performed on the first qubit before a measurement is performed.

We now consider the computation is performed that is shown in diagrammatic form in Fig. 3.5. It corresponds to the following sequence of unitary transformations

$$
\begin{aligned}
|0\rangle \otimes |1\rangle \quad &\to \quad \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
&\to \quad \frac{1}{\sqrt{2}}((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
&\to \quad \frac{1}{2}\Big[((-1)^{f(0)} + (-1)^{f(1)})|0\rangle \\
&\qquad + ((-1)^{f(0)} - (-1)^{f(1)})|1\rangle\Big] \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
&\equiv \quad |q_1\rangle \otimes |q_2\rangle\,. \tag{3.37}
\end{aligned}
$$

We note that the for the first qubit there are two possible final states depending on whether $f(0)$ and $f(1)$ are equal or different. Thus,

$$\begin{aligned}
f(0) + f(1) &= 0 \Rightarrow |q_1\rangle = (-1)^{f(0)}|1\rangle\,, \\
f(0) + f(1) &= 1 \Rightarrow |q_1\rangle = (-1)^{f(0)}|0\rangle\,.
\end{aligned} \tag{3.38}$$

This implies that by making a measurement on this qubit, which projects it either to the state $|0\rangle$ or $|1\rangle$, we can decide the value of $f(0)+f(1)$. This does not fully determine the function $f(x)$, but it determines whether it belongs to set A (with $f(0)+f(1) = 1$) or set B (with $f(0)+f(1) = 0$). This demonstrates the point that the use of quantum superpositions makes it possible to perform in one operation what would normally correspond to two classical computations.

This use of superposition is what we have referred to as *quantum parallel processing*. There is a close relation between the evaluation discussed here, where the quantum state contains information about both functional values $f(0)$ and $f(1)$, and the interaction-free measurement discussed earlier, where the state vector contains information about both the paths that the photon may follow.

The example given by Deutsch's algorithm is too simple to convincingly justify the claim that the quantum computation is more efficient than the classical computation. Obviously the unitary transformation U_f corresponding to the function $f(x)$ has to be supplemented by other qubit operations and this does not make it completely clear that there is a net gain. The important point to make is that only one evaluation which involves $f(x)$ has been done rather than two. For the other algorithms mentioned one can demonstrate more explicitly the gain by showing that the number of qubit operations scales in a different way than the number of operations in a classical computation. This makes it clear that quantum parallelism may indeed speed up certain types of calculations.

3.6.3   Can a quantum computer be constructed?

The considerations on how a quantum computer may more efficiently solve certain types of problems makes it a very interesting idea. But is it feasible that this idea can be implemented in the form of a real physical computer? The difficulties to overcome are extremely demanding. At present qubit operations with a small number of qubits can be performed, but the idea that thousands and thousands of qubits work coherently together at the quantum level is at this stage an attractive dream. Some hysicists are rather pessimistic that the necessary control of the quantum states can in reality be made. In particular the problem of *decoherence* is extremely demanding, although algorithms for correcting quantum states that are modified due to decoherence have been suggested. However, other physicists, who work in this field, remain optimistic, and at the level of making controlled quantum operations on a few qubits there has been an impressive progress.

There is in fact a competition between different types of realizations of physical qubits. In the context of electronic systems interesting developments are based on the use of electronic spin as the two-level variable. In the context of quantum optics the use of trapped two-level atoms or ions has been extensively studied. One particularly interesting application is in the form of *optical lattices*, where a collection of laser beams are used to trap atoms in a periodic potential. A pair of atomic levels may define a qubit, and one-qubit operations on this atom

may be be performed in a controlled way by manipulating the atom with a laser beam. The two-qubit operations are similarly obtained by manipulating pairs of atoms in such a way that interactions between the atoms will induce the relevant qubit operations.

One should note that even if the construction of a *universal* quantum computer at this stage may seem to be far into the future, there may be partial goals that can be more readily achieved. Ideas of quantum cryptography have already been implemented, and in the field of computation a *quantum simulator* may be a much closer goal than a universal computer. Recent developments suggests that the use of optical lattices may give a realistic approach towards this goal, and for example the simulation in this way of *quantum spin lattices* (which are important systems to study in condensed matter theory) may be possible in a not too distant future.

Chapter 4

Photons and atoms

In this chapter we study the quantum theory of light and its interaction with matter. Historically the understanding of how light is created and absorbed by atoms was central for development of quantum theory, starting with Planck's revolutionary idea of *energy quanta* in the description of black body radiation. Today a more complete description exists, with the theory of *quantum electrodynamics* (QED) as a part of a more complete relativistic quantum field theory, which describes the electro-weak physics of the elementary particles. Even so the non-relativistic theory of photons and atoms has continued to be important, and has been developed further, in directions that are referred to as *quantum optics*. We will study here the basics of the non-relativistic description of interactions between photons and atoms, in particular with respect to the processes of spontaneous and stimulated emission. As a particular application we study a simple model of a laser as a source of *coherent light*.

4.1   Classical electromagnetism

Maxwell's theory of electromagnetism is the basis for the classical as well as the quantum description of radiation. With some modifications the quantum theory can be derived from the classical theory by a similar quantization method as previously used to quantize the simple field theory of the vibrating string. In the present case the natural choice of generalized coordinates are the *field amplitudes* of the electromagnetic potentials, but in order to secure that these correspond to *independent* degrees of freedom, we have to sort out some problems related to the *gauge invariance* of the electromagnetic theory. We also have to take into account the additional degrees of freedom related to polarization.

We first, in this section, make a brief summary of the classical theory and show how a Lagrangian and Hamiltonian formulation of electromagnetic fields interacting with point charges can be given. At the next step this forms the basis for the quantization of the free field and for the formulation of the complete quantum theory of interacting fields and charges. We make use of the *non-covariant* form of Maxwell's equations, which is convenient when only coupling to non-relativistic electrons are considered. When relativistic electrons interact with the electromagnetic field the covariant formulation is often better, even if the treatment of the gauge invariance then is more cumbersome.

4.1.1 Maxwell's equations

Maxwell's equations in non-covariant form are

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0}, \\
\boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \mathbf{j}, \\
\boldsymbol{\nabla} \cdot \mathbf{B} &= 0, \\
\boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0,
\end{aligned}
\tag{4.1}
$$

with $\mathbf{E}$ and $\mathbf{B}$ as the electric and magnetic fields and ρ and $\mathbf{j}$ as the charge and current densities. The constant ϵ_0 is the *vacuum permittivity* and μ_0 is the *vacuum permeability*. They are related to the speed of light through the equation $\epsilon_0\mu_0 = 1/c^2$, with c, as usual, representing the speed of light.

The electric and magnetic fields, when expressed in terms of the scalar potential ϕ and the vector potential $\mathbf{A}$, take the form

$$
\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial}{\partial t}\mathbf{A}, \quad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}.
\tag{4.2}
$$

Expressed in this way, the two homogeneous equations in (4.1) are satisfied as identities, while the remaining two equations take the form

$$
\begin{aligned}
\boldsymbol{\nabla}^2\phi + \frac{\partial}{\partial t}\boldsymbol{\nabla} \cdot \mathbf{A} &= -\frac{\rho}{\epsilon_0}, \\
\boldsymbol{\nabla}^2\mathbf{A} - \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{A}) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} - \frac{1}{c^2}\frac{\partial}{\partial t}\boldsymbol{\nabla}\phi &= -\mu_0\mathbf{j}.
\end{aligned}
\tag{4.3}
$$

The potentials are however not uniquely determined by the electromagnetic fields. A *gauge transformation* of the electromagnetic potentials has the following form,

$$
\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} - \boldsymbol{\nabla}\chi, \quad \phi \rightarrow \phi' = \phi + \frac{\partial}{\partial t}\chi,
\tag{4.4}
$$

with χ as a (scalar) function of space and time. It is straight forward to show that such a transformation leaves both the electric and the magnetic field, $\mathbf{E}$ and $\mathbf{B}$, invariant. The usual way to view the invariance of the fields under this transformation is that it reflects the presence of *non-physical degrees of freedom* in the potentials. Thus, the potentials define an overcomplete set of variables for the electromagnetic field.

For the Hamiltonian formulation to work it is necessary to identify the independent, physical degrees of freedom. This can be done by introducing a constraint, or gauge conditions, which removes the unphysical degrees of freedom associated with gauge invariance. In the non-covariant formulation of the theory this is usually done in the form of the Coulomb (or radiation) gauge condition,

$$
\boldsymbol{\nabla} \cdot \mathbf{A} = 0.
\tag{4.5}
$$

This condition will not constrain the physical fields $\mathbf{E}$ and $\mathbf{B}$, but it simplifies the two remaining field equations (4.2)

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0} \,, \tag{4.6}$$

$$(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2)\mathbf{A} = \mu_0 \mathbf{j}_T \,, \tag{4.7}$$

where

$$\mathbf{j}_T \equiv \mathbf{j} - \epsilon_0 \frac{\partial}{\partial t}\nabla\phi \tag{4.8}$$

is the *transverse* component of the current density.[1] It satisfies the transversality condition as a consequence of the continuity equation for charge,

$$
\begin{aligned}
\nabla \cdot \mathbf{j}_T &= \nabla \cdot \mathbf{j} - \epsilon_0 \frac{\partial}{\partial t}\nabla^2 \phi \\
&= \nabla \cdot \mathbf{j} + \frac{\partial}{\partial t}\rho \\
&= 0 \,.
\end{aligned}
\tag{4.9}
$$

We note that the equation (4.6) for the scalar field ϕ contains no time derivatives, and can be solved in terms of the charge distribution,

$$\phi(\mathbf{r}, t) = \int d^3 r' \frac{\rho(\mathbf{r}', t)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} \,. \tag{4.10}$$

This is the same as the electrostatic (Coulomb) potential of a *stationary* charge distribution which coincides with the true charge distribution $\rho(\mathbf{r}, t)$ at time t. Since the ϕ-field, in this formulation, is fully determined by the position of the charges, the dynamics of the Maxwell field is determined solely by the vector potential $\mathbf{A}$. In the quantum description this means that the photons are associated only with the vector potential $\mathbf{A}$ and not with ϕ.

4.1.2 Lagrange formulation

The field equation (4.7) of the vector potential can be derived from the following Lagrangian density,

$$\mathcal{L} = \frac{\epsilon_0}{2}\dot{\mathbf{A}}^2 - \frac{1}{2\mu_0}(\nabla \times \mathbf{A})^2 + \mathbf{j} \cdot \mathbf{A} \,, \tag{4.11}$$

where we have used the freedom to replace, in the Lagrangian, the transverse current $\mathbf{j}_T$ with the full current $\mathbf{j}$. This can be done since the difference gives rise to an irrelevant derivative term.[2]

---

[1]Any vector field $\mathbf{j}(\mathbf{r})$ can be written as $\mathbf{j} = \mathbf{j}_T + \mathbf{j}_L$, with $\nabla \cdot \mathbf{j}_T = 0$ and $\nabla \times \mathbf{j}_L = 0$. $\mathbf{j}_T$ is referred to as the *transverse* or *solenoidal* part of the field and $\mathbf{j_L}$ as the *longitudinal* or *irrotational* part of the field. In the present case, with $\mathbf{j}_L = -\nabla\dot{\phi}$ the irrotational form of $\mathbf{j}_L$ follows directly, while the transversality of $\mathbf{j}_T$ follows from charge conservation.

[2]Note, however, that when the full current is used, the transversality condition $\nabla \cdot \mathbf{A}$ has to be imposed as a constraint to derive the correct field equation from the Lagrangian. When the $\mathbf{A}$-field is coupled only to the transverse part, that is not needed.

So far we have not worried about the degrees of freedom associated with the charges. We will now include them in the description by assuming the charges to be carried by point particles. This means that we express the charge density and current as

$$
\begin{aligned}
\rho(\mathbf{r}, t) &= \sum_i e_i \delta(\mathbf{r} - \mathbf{r}_i(t)) \,, \\
\mathbf{j}(\mathbf{r}, t) &= \sum_i e_i \mathbf{v}_i(t) \delta(\mathbf{r} - \mathbf{r}_i(t)) \,,
\end{aligned}
\tag{4.12}
$$

where the sum is over all the particles with $\mathbf{r}_i(t)$ and $\mathbf{v}_i(t)$ as the position and velocity of particle i as functions of time. The full Lagrangian of the interacting system we write as

$$
L = L_{field} + L_{int} + L_{part} \,,
\tag{4.13}
$$

where L_{field} is the part of L which depends only on the field variable $\mathbf{A}$, L_{int} is the part which depends both on the field and particle variables, and L_{part} depends only on the particle variables. The Lagrangian is defined as the space integral of the Lagrangian *density*. Thus, the two first terms of (4.13) is derived from the integral of the Lagrangian density (4.11), while the last term includes the contributions from the charged particles, in the form of kinetic and potential (Coulomb) energy. This gives the following expression for the full Lagrangian L,

$$
\begin{aligned}
L &= \int d^3 r \left[\frac{\epsilon_0}{2} \dot{\mathbf{A}}^2 - \frac{1}{2\mu_0} (\boldsymbol{\nabla} \times \mathbf{A}) \right] + \sum_i e_i \mathbf{v}_i \cdot \mathbf{A}(\mathbf{r}_i) \\
&\quad + \sum_i \frac{1}{2} m_i \mathbf{v}_i^2 - \frac{1}{2} \sum_{i \neq j} \frac{e_i e_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \,.
\end{aligned}
\tag{4.14}
$$

To find the corresponding Hamiltonian we make use of the standard relation between the Lagrangian and Hamiltonian of a system,

$$
H = \int d^3 r \, \boldsymbol{\Pi} \cdot \dot{\mathbf{A}} + \sum_i \mathbf{p}_i \cdot \mathbf{v}_i - L \,,
\tag{4.15}
$$

where

$$
\boldsymbol{\Pi} = \epsilon_0 \dot{\mathbf{A}} \equiv -\epsilon_0 \mathbf{E}_T \,,
\tag{4.16}
$$

is the canonical conjugate field momentum density and

$$
\mathbf{p}_i = m \mathbf{v}_i + e_i \mathbf{A}(\mathbf{r}_i) \,,
\tag{4.17}
$$

is the canonical conjugate particle momentum. The full expression for the Hamiltonian then is

$$
H = \int d^3 r \, \frac{1}{2} \left(\epsilon_0 \mathbf{E}_T^2 + \frac{1}{\mu_0} \mathbf{B}^2 \right) + \sum_i \frac{1}{2 m_i} (\mathbf{p}_i - e_i \mathbf{A}(\mathbf{r}_i))^2 + \sum_{i<j} \frac{e_i e_j}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \,.
\tag{4.18}
$$

This expression for the Hamiltonian is derived for the classical system of interacting fields and particles, but it has the same form as the Hamiltonian *operator* that is used to describe the quantum system of non-relativistic electrons interacting with the electromagnetic field. However, for spinning particles there is an additional term to include, which is coupling term of the

magnetic dipole moment of the charged particles and the magnetic field. The standard form of this magnetic dipole term is

$$H_{spin} = -\sum_i \frac{g_i e_i}{2m_i} \mathbf{S}_i \cdot \mathbf{B}(\mathbf{r}_i) \,, \tag{4.19}$$

where g_i is the g-factor of particle i, which is close to 2 for electrons. Often the spin term is small corresponding to the other interaction terms and can be neglected.

4.2 Photons – the quanta of light

We consider now the quantum description of the system of particles and fields. With the degrees of freedom of the electromagnetic field and of the electrons disentangled, we may assume the space of states of the full quantum system to be the product space of a *field*-state space and a *particle*-state space,

$$\mathcal{H} = \mathcal{H}_{field} \otimes \mathcal{H}_{particle} \,. \tag{4.20}$$

The quantization of the electromagnetic field is then independent of the quantization of the particle degrees of freedom. Since the interaction is not important for the the field quantization, we may simply focus on the quantum description of the *free* electromagnetic field, which defines the space $\mathcal{H}_{field}$ and the operators (observables) acting there. The energy eigenstates of the free field are the *photon states*. When we as a next step include interactions, these will introduce processes where photons are emitted and absorbed. The method we use to make the transition from the classical to the quantum description of the free electromagnetic field is the essentially the same as earlier used in Sect. 1.2 to quantize the vibrating string, however here taking into account the effects of the Coulomb gauge condition and the presence of the polarization degrees of freedom.

4.2.1 The quantized field

To deal with the transversality condition $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ it is convenient to expand $\mathbf{A}(\mathbf{r})$ in a Fourier series, with the wave vector $\mathbf{k}$ as the Fourier variable conjugate to $\mathbf{r}$. It is also convenient to confine the fields to a finite, but large box with linear dimensions L, and to assume periodic conditions for the fields. The components of $\mathbf{k}$ then take discrete values, $k_i = 2\pi n_i/L$, with n_i as a set of integers, and the field amplitude can be written as a discrete Fourier sum

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \sum_{a=1}^{2} A_{\mathbf{k}a}(t) \boldsymbol{\epsilon}_{\mathbf{k}a} e^{i\mathbf{k} \cdot \mathbf{r}} \,, \tag{4.21}$$

where the $V = L^3$ is the normalization volume, and the vectors $\boldsymbol{\epsilon}_{\mathbf{k}a}$ are unit vectors which satisfy $\mathbf{k} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a} = 0$ as a consequence of the transversality condition $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$. Since there are no constraints on the amplitudes $A_{\mathbf{k}a}$, these variables forms a convenient set of generalized coordinates for the independent degrees of freedom of the field. Note, however that the amplitudes for opposite wave vectors $\mathbf{k}$ and $-\mathbf{k}$ will represent the same degrees of freedom, as we shall discuss below.

The *normal modes*, which are defined by the amplitudes

$$A_{\mathbf{k}a}(t) = A_{\mathbf{k}a}^0 e^{\pm i\omega_k t} \;,\;\; \omega_k = ck \;\;\; (k = |\mathbf{k}|)\,, \tag{4.22}$$

represent plane waves of *monochromatic, polarized light*. The two vectors $\boldsymbol{\epsilon}_{\mathbf{k}a}$, with $a = 1, 2$, are the *polarization vectors*, which are unit vectors in the plane perpendicular to $\mathbf{k}$. We will in the following chose these vectors to be real, with the orthonormality condition

$$\boldsymbol{\epsilon}_{\mathbf{k}a} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a'} = \delta_{aa'}\,. \tag{4.23}$$

Real polarization vectors corresponds to linear polarization of the field modes. Also complex vectors can be chosen, corresponding to circular or more generally to elliptic polarization. These different sets of polarization vectors are equivalent in the sense that the unit vectors of one set can be written as linear combinations of the unit vectors of any other set.

The fact that the field $\mathbf{A}(\mathbf{r}, t)$ is a real field gives the following relation between the field amplitudes at wave vectors $\mathbf{k}$ and $-\mathbf{k}$,

$$\sum_a A_{-\mathbf{k}a}\boldsymbol{\epsilon}_{-\mathbf{k}a} = \sum_a A_{\mathbf{k}a}^*\boldsymbol{\epsilon}_{\mathbf{k}a}\,. \tag{4.24}$$

The polarization vectors can be chosen so that

$$\boldsymbol{\epsilon}_{-\mathbf{k}a} = \boldsymbol{\epsilon}_{\mathbf{k}\bar{a}}\,, \tag{4.25}$$

where $\bar{a}$ is determined from a by interchanging the two polarization directions 1 and 2. For the field amplitude this give the relation

$$A_{\mathbf{k}a}^* = A_{-\mathbf{k}\bar{a}}\,. \tag{4.26}$$

The Lagrangian of the free electromagnetic field expressed in terms of the Fourier amplitudes is

$$L = \frac{1}{2}\epsilon_0 \sum_{\mathbf{k}a} \left[\dot{A}_{\mathbf{k}a}^* \dot{A}_{\mathbf{k}a} - \omega_k^2 A_{\mathbf{k}a}^* A_{\mathbf{k}a}\right]\,. \tag{4.27}$$

We note that there is no coupling between the different Fourier components, and for each component the Lagrangian has the same form as for a harmonic oscillator of frequency $\omega = ck$. This is similar to the situation previously discussed in the quantization of the field theory of the vibrating string. However, in the present case the field variables are complex rather than real, since reality of the field is represented by the relation (4.26) with each Fourier component being complex. A rewriting of the Lagrangian in terms of real variables is straight forward, but it is more convenient to continue to work with the complex variables.

We also note two other differences when comparing with the string field thory. The first one is the presence of additional degrees of freedom associated with polarization, and the other is the use of periodic boundary conditions which allows propagating waves rather than standing waves as the fundamental modes. However, the important common feature is the decomposition of the systems into non-interacting harmonic oscillators, one for each field mode. This reflects the free field properties of the systems, which means that the fields satisfy linear field equations.

To introduce conjugate momenta for the variables $A_{\mathbf{k}a}$ is now unproblematic,

$$\Pi_{\mathbf{k}a} = \epsilon_0 \dot{A}^*_{\mathbf{k}a} = -\epsilon_0 E^*_{\mathbf{k}a}, \tag{4.28}$$

where $E_{\mathbf{k}a}$ is the Fourier component of the electric field. From this the form of the free field Hamiltonian is found

$$
\begin{aligned}
H &= \sum_{\mathbf{k}a} \Pi_{\mathbf{k}a} \dot{A}_{\mathbf{k}a} - L \\
&= \sum_{\mathbf{k}a} \frac{1}{2} \epsilon_0 (E^*_{\mathbf{k}a} E_{\mathbf{k}a} + \omega_k^2 A^*_{\mathbf{k}a} A_{\mathbf{k}a}),
\end{aligned} \tag{4.29}
$$

which is consistent with the earlier expression found for the Hamiltonian of the electromagnetic field. It is represented by the first term of (4.15), here written in Fourier transformed field components.

Quantization of the theory means that the classical field amplitude $\mathbf{A}$ and field strength $\mathbf{E}$ are now replaced by operators $\hat{\mathbf{A}}$ and $\hat{\mathbf{E}}$, while the complex conjugate fields are replaced by hermitian conjugate operators. The conjugate field variables satisfy the canonical commutation relations in the form

$$\left[\hat{E}^\dagger_{\mathbf{k}a}, \hat{A}_{\mathbf{k}'b} \right] = -\frac{1}{c} \left[\dot{\hat{A}}^\dagger_{\mathbf{k}a}, \hat{A}_{\mathbf{k}'b} \right] = i \frac{\hbar}{\epsilon_0} \delta_{\mathbf{k}\mathbf{k}'} \delta_{ab}, \tag{4.30}$$

while the different components of the $\mathbf{A}$ commute

$$\left[\hat{A}_{\mathbf{k}a}, \hat{A}_{\mathbf{k}'b} \right] = 0. \tag{4.31}$$

It is convenient to change to new variables,

$$
\begin{aligned}
\hat{A}_{\mathbf{k}a} &= \sqrt{\frac{\hbar}{2\omega_k \epsilon_0}} (\hat{a}_{\mathbf{k}a} + \hat{a}^\dagger_{-\mathbf{k}\bar{a}}), \\
\hat{E}_{\mathbf{k}a} &= i \sqrt{\frac{\hbar \omega_k}{2\epsilon_0}} (\hat{a}_{\mathbf{k}a} - \hat{a}^\dagger_{-\mathbf{k}\bar{a}}),
\end{aligned} \tag{4.32}
$$

where the reality condition (4.24) (here meaning hermiticity of $\hat{\mathbf{A}}(\mathbf{r}, t)$ and $\hat{\mathbf{E}}(\mathbf{r}, t)$) has been made explicit. In terms of the new variables the Hamiltonian, with the vacuum energy subtracted, takes the form

$$\hat{H} = \sum_{\mathbf{k}a} \hbar \omega_k \hat{a}^\dagger_{\mathbf{k}a} \hat{a}_{\mathbf{k}a}, \tag{4.33}$$

and the canonical commutation relations are

$$
\begin{aligned}
\left[\hat{a}_{\mathbf{k}a}, \hat{a}^\dagger_{\mathbf{k}'b} \right] &= \delta_{\mathbf{k}\mathbf{k}'} \delta_{ab}, \\
\left[\hat{a}_{\mathbf{k}a}, \hat{a}_{\mathbf{k}'b} \right] &= 0.
\end{aligned} \tag{4.34}
$$

Expressed in this way the Hamiltonian has exactly the form of a collection of independent quantum oscillators, one for each field mode, with $\hat{a}_{\mathbf{k}a}$ as lowering operator and $\hat{a}^\dagger_{\mathbf{k}a}$ as raising operator in the energy spectrum of the oscillator labeled by $(\mathbf{k}a)$.

It is convenient to work, in the following, in the Heisenberg picture, where the observables are time dependent. In this picture the vector potential is described by the *field operators*,

$$\hat{\mathbf{A}}(\mathbf{r},t) = \sum_{\mathbf{k}a} \sqrt{\frac{\hbar}{2V\omega_k\epsilon_0}} \left[\hat{a}_{\mathbf{k}a}\boldsymbol{\epsilon}_{\mathbf{k}a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + \hat{a}^\dagger_{\mathbf{k}a}\boldsymbol{\epsilon}_{\mathbf{k}a}e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \right], \qquad (4.35)$$

and the electric and mangnetic fields are field operators derived from this by differensiation with respect to t and $\mathbf{r}$,

$$\hat{\mathbf{E}}(\mathbf{r},t) = -\frac{\partial}{\partial t}\hat{\mathbf{A}}(\mathbf{r},t), \quad \hat{\mathbf{B}}(\mathbf{r},t) = \boldsymbol{\nabla} \times \hat{\mathbf{A}}(\mathbf{r},t). \qquad (4.36)$$

4.2.2   Constructing Fock space

The state space of the free electromagnetic field then can be viewed as a product space of harmonic oscillator spaces, one for each normal mode of the field. The vacuum state is defined as the ground state of the Hamiltonian (4.33), which means that it is the state where all oscillators are unexcited,

$$\hat{a}_{\mathbf{k}a}|0\rangle = 0 \quad \text{for all } (\mathbf{k},a). \qquad (4.37)$$

The operator $\hat{a}^\dagger_{\mathbf{k}a}$, which excites one of the oscillators, is interpreted as a *creation operator* which creates one *photon* from the vacuum,

$$\hat{a}^\dagger_{\mathbf{k}a}|0\rangle = |1_{\mathbf{k}a}\rangle. \qquad (4.38)$$

By repeatedly applying the creation operator an arbitrary number of photons can be created in the same state $(\mathbf{k}a)$,

$$(\hat{a}^\dagger_{\mathbf{k}a})^{n_{\mathbf{k}a}}|0\rangle = \sqrt{n_{\mathbf{k}a}!}\,|n_{\mathbf{k}a}\rangle. \qquad (4.39)$$

The operator $\hat{a}_{\mathbf{k}a}$ is an *annihilation operator* which reduces the number of photons in the mode $(\mathbf{k}a)$,

$$\hat{a}_{\mathbf{k}a}|n_{\mathbf{k}a}\rangle = \sqrt{n_{\mathbf{k}a}}\,|n_{\mathbf{k}a}-1\rangle. \qquad (4.40)$$

The photon number operator is

$$\hat{N}_{\mathbf{k}a} = \hat{a}^\dagger_{\mathbf{k}a}\hat{a}_{\mathbf{k}a}. \qquad (4.41)$$

It counts the number of photons present in the state $(\mathbf{k}a)$,

$$\hat{N}_{\mathbf{k}a}|n_{\mathbf{k}a}\rangle = n_{\mathbf{k}a}\,|n_{\mathbf{k}a}\rangle. \qquad (4.42)$$

The general photon state, often referred to as a *Fock state*, is a product state with a well-defined number of photons for each set of quantum numbers $(\mathbf{k}a)$. It is specified by a set of occupation numbers $\{n_{\mathbf{k}a}\}$ for the single-photon states. The set of Fock states $|\{n_{\mathbf{k}a}\}\rangle$ form a basis of orthonormal states that span the state space, the *Fock space*, of the quantized field.

Thus a general quantum state of the free electromagnetic field is a linear superposition of Fock states,

$$|\psi\rangle = \sum_{\{n_{\mathbf{k}a}\}} c(\{n_{\mathbf{k}a}\})|\{n_{\mathbf{k}a}\}\rangle \,, \tag{4.43}$$

with $c(\{n_{\mathbf{k}a}\})$ as the expansion coefficients in this basis.

The observables of the theory are assumed to have the same form as the corresponding classical variables when expressed in terms of the electric and magnetic fields. Thus, the energy operator has the same form as the classical field energy

$$
\begin{aligned}
\hat{H} &= \int d^3r \frac{1}{2}(\epsilon_0 \hat{E}^2 + \frac{1}{\mu_0}\hat{B}^2) \\
&= \sum_{\mathbf{k}a} \frac{1}{2}\hbar\omega_k(\hat{a}^\dagger_{\mathbf{k}a}\hat{a}_{\mathbf{k}a} + \hat{a}_{\mathbf{k}a}\hat{a}^\dagger_{\mathbf{k}a}) \\
&= \sum_{\mathbf{k}a} \hbar\omega_k(\hat{N}_{\mathbf{k}a} + \frac{1}{2}) \,,
\end{aligned} \tag{4.44}
$$

and is identical to the free field part of the Hamiltonian previously discussed, except that the vacuum energy is here included. In the following this contribution will usually be ommited.

In a similar way the classical field momentum, defined by Poyntings vector, is replaced by an operator of the same form,

$$
\begin{aligned}
\hat{\mathcal{P}} &= \int d^3r \epsilon_0(\hat{\mathbf{E}} \times \hat{\mathbf{B}}) \\
&= \sum_{\mathbf{k}a} \frac{1}{2}\hbar\mathbf{k}(\hat{a}^\dagger_{\mathbf{k}a}\hat{a}_{\mathbf{k}a} + \hat{a}_{\mathbf{k}a}\hat{a}^\dagger_{\mathbf{k}a}) \\
&= \sum_{\mathbf{k}a} \hbar\mathbf{k}\hat{N}_{\mathbf{k}a} \,.
\end{aligned} \tag{4.45}
$$

In this case the vacuum contribution disappears when summing over $\mathbf{k}$. Both $\hat{H}$ and $\hat{\mathcal{P}}$ are diagonal in the Fock basis.

The expressions for $\hat{H}$ and $\hat{\mathcal{P}}$ show that a single photon in the mode $(\mathbf{k}a)$ has energy $E = \hbar\omega_k$ and momentum $\mathbf{p} = \hbar\mathbf{k}$, in accordance with the de Broglie relations. Furthermore, the relation $\omega_k = c\mathbf{k}$, generally valid for electromagnetic waves, means that the energy-momentum relation is $E = cp$, which is the correct one for a *massless* relativistic particle. This is consistent with the idea of the photon as a massless *particle*, not simply as an energy quantum of an electromagnetic mode. In addition, the degree of freedom associated with the two possible polarizations is readily re-interpreted as due to the intrinsic spin of the photon. Thus, the photon carries one unit of spin, and the existence of only two spin states, rather than three, is related to the fact that the photon is massless. A natural choice for two orthogonal spin states are the *helicity* states, corresponding to spin either in the direction of the momentum or in the direction opposite to the momentum. In the language of polarization these two possibilities correspond to circular polarization, which may either be right-handed or left-handed. The photons are bosons, since an arbitrary number of photons can be created in the same state.

4.2.3   Coherent and incoherent photon states

In a similar way as for the quantum description of particles, we may view the expectation values of the electromagnetic field operators to represent classical variables within the quantum theory. For example, the expectation value of the electric field operator takes the usual form of a classical electric field. Expanded in plane wave components it is

$$\left\langle \hat{\mathbf{E}}(\mathbf{r}, t) \right\rangle = i \sum_{\mathbf{k}a} \sqrt{\frac{\hbar \omega_k}{2V\epsilon_0}} \left[\alpha_{\mathbf{k}a} \boldsymbol{\epsilon}_{\mathbf{k}a} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} - \alpha_{\mathbf{k}a}^* \boldsymbol{\epsilon}_{\mathbf{k}a} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \right], \quad (4.46)$$

where the complex expansion coefficients $\alpha_{\mathbf{k}a}$ and $\alpha_{\mathbf{k}a}^*$ are the expectation values of the annihilation and creation operators,

$$\alpha_{\mathbf{k}a} = \langle \hat{a}_{\mathbf{k}a} \rangle \;,\; \alpha_{\mathbf{k}a}^* = \left\langle \hat{a}_{\mathbf{k}a}^\dagger \right\rangle . \quad (4.47)$$

The quantum field in addition has (quantum) fluctuations around the classical configuration.

We note, from the above expression, the curious fact that for Fock states, with a sharply defined set of photon numbers, the expectation value of the electric (as well as the magnetic) field vanishes. Such states, which may be highly excited in energy while the expectation values for the fields vanish, we may regard as being highly *non-classical states*. In these states the fields, in a sense, are dominated by the quantum fluctuations. Classical fields, on the other hand, may be represented by states where the expectation values dominate the quantum fluctuations. For a single harmonic oscillator we have already discussed such "classical" states in the form of *coherent states*, states with minimum uncertainty in their phase space positions $\hat{x}$ and $\hat{p}$. For the electromagnetic field the corresponding variables are $\hat{\mathbf{A}}(\mathbf{r}, t)$ and $\hat{\mathbf{E}}(\mathbf{r}, t)$, and since each field mode can be regarded as a harmonic oscillator, the definition of oscillator coherent states can be directly applied to define coherent states of the electromagnetic field. Such states correspond to a particular form of superpositions of Fock states.

For a single electromagnetic field mode $(\mathbf{k}a)$, the coherent state has the same form as discussed in Sect. 1.3.3 for the harmonic oscillator,

$$|\alpha_{\mathbf{k}a}\rangle = \sum_{n_{\mathbf{k}a}} e^{-\frac{1}{2}|\alpha_{\mathbf{k}a}|^2} \frac{(\alpha_{\mathbf{k}a})^{n_{\mathbf{k}a}}}{\sqrt{n_{\mathbf{k}a}!}} |n_{\mathbf{k}a}\rangle , \quad (4.48)$$

where $\alpha_{\mathbf{k}a}$ is related to the expectation values of $\hat{a}_{\mathbf{k}a}$ and $\hat{a}_{\mathbf{k}a}^\dagger$ as in (4.47). The full coherent state of the electromagnetic field is then a product state of the form

$$|\psi\rangle = \prod_{\mathbf{k}a} |\alpha_{\mathbf{k}a}\rangle , \quad (4.49)$$

and the corresponding expectation value of the electric field is given by (4.46).

Vacuum fluctuations

The vacuum state is a special case of a coherent state, and in the same way as for a single harmonic oscillator, the quantum fluctuations of this state are the same as for any other coherent state. This follows since the coherent state can be considered as obtained from the vacuum state

by shifting the field operators by a classical field contribution[3]. The quantum fluctuations of the fields in a given, sharply defined point in space, are ill defined. The *two-point correlation function*, however, gives finite expressions for the fluctuations. We examine here the vacuum correlation function

$$
\begin{aligned}
C_{ij}(\mathbf{r} - \mathbf{r}') &\equiv \langle E_i(\mathbf{r}) E_j(\mathbf{r}') \rangle - \langle E_i(\mathbf{r}) \rangle \langle E_j(\mathbf{r}') \rangle_{vac} \\
&= \sum_{\mathbf{k}} \frac{\hbar \omega_k}{2V\epsilon_0} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} (\delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2}) \\
&\rightarrow \frac{c\hbar}{2(2\pi)^3\epsilon_0} \int d^3k\, k\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} (\delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2}),
\end{aligned}
\tag{4.50}
$$

where in the last step we have taken the infinite volume limit. For large k this integral has an undamped oscillatory behavior, but the integral can be made well defined by introducing a damping factor $e^{-\epsilon k}$, taking ϵ to 0 after the integration. We note that the correlation function only depends on the relative coordinates of the two points and therefore simply put $\mathbf{r}' = 0$, so that $\mathbf{r}$ represents the relative coordinate of the two points. The correlation function can be re-expressed in the following way

$$
C_{ij}(\mathbf{r}) = \frac{c\hbar}{2(2\pi)^3\epsilon_0} \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} - \delta_{ij} \sum_k \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_k} \right) \int d^3k\, \frac{1}{k} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-\epsilon k},
\tag{4.51}
$$

with the integral evaluated as

$$
\begin{aligned}
\int d^3k\, \frac{1}{k} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-\epsilon k} &= 2\pi \int_0^\infty dk\, k \int_0^\pi d\theta \sin\theta\, e^{ikr\cos\theta - \epsilon k} \\
&= 2\pi \int_0^\infty dk\, k\, e^{-\epsilon k} \int_{-1}^1 du\, e^{ikru} \\
&= \frac{4\pi}{r} \int_0^\infty dk \sin kr\, e^{-\epsilon k} \\
&= \frac{4\pi}{r^2},
\end{aligned}
\tag{4.52}
$$

where in the last step we have taken the limit $\epsilon \rightarrow 0^+$. This gives

$$
\begin{aligned}
C_{ij}(\mathbf{r}) &= \frac{c\hbar}{2(2\pi)^3\epsilon_0} \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} - \delta_{ij} \sum_k \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_k} \right) \frac{4\pi}{r^2} \\
&= -\frac{c\hbar}{2\pi^2\epsilon_0 r^4} (\delta_{ij} - 4\frac{x_i x_j}{r^2}).
\end{aligned}
\tag{4.53}
$$

We note the fall-off of the correlation with increased separation between the two points and also note that the function diverges as $r \rightarrow 0$. This divergence is typical for field theories, and reflects the fact that the *physical* fields cannot be defined with infinite resolution. A way to avoid such infinities is to define the physical fields by averaging over a small volume. This introduces effectively a momentum cut-off in the Fourier transformation of the field.

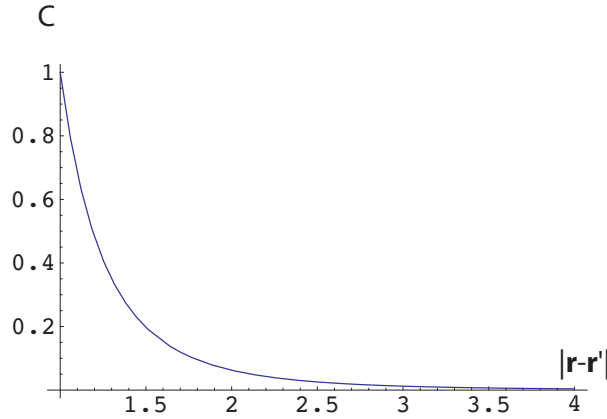[3]*i.e.*, by a contribution proportional to the identity operator

Figure 4.1: The correlation function for the electric field in a coherent state, $C_{ij}(\mathbf{r} - \mathbf{r}') \equiv \langle E_i(\mathbf{r})E_j(\mathbf{r}')\rangle - \langle E_i(\mathbf{r})\rangle \langle E_j(\mathbf{r}')\rangle$. The form of the non-vanishing elements, $i = j$, is shown as the function of distance $\mathbf{r} - \mathbf{r}'$ between two points in space.

A similar calculation for the magnetic field shows that the correlation function is the same as for the electric field. This is typical for the coherent state which is symmetric in its dependence of the field variable and its conjugate momentum. Since $\mathbf{A}$ and $\mathbf{E}$ are conjugate variables, the electric and magnetic fields do not commute as operator fields. Formally the commutator is, for equal times,

$$\left[B_i(\mathbf{r}), E_j(\mathbf{r}')\right] = i\frac{\hbar}{\epsilon_0}\epsilon_{ijk}\,\frac{\partial}{\partial x_k}\delta(\mathbf{r} - \mathbf{r}')\,, \tag{4.54}$$

where one should consider the Fourier transform of the above expression, in order to give the derivative of the delta-function a precise meaning. There are states where the fluctuations in the $\mathbf{E}$-field are suppressed relative to that of the vacuum state, but due to the non-vanishing of the commutator with $\mathbf{B}$, the fluctuations in the $\mathbf{B}$-field will then be larger than that of the vacuum. States with reduced fluctuations in one of the fields, and which still satisfy the condition of minimum (Heisenberg) uncertainty for conjugate variables, are often referred to as *squeezed states*.

Radio waves, produced by oscillating currents in an antenna, can clearly be regarded as classical electromagnetic waves and are well described as coherent states of the electromagnetic field within the quantum theory of radiation. The mean value of the fields at the receiver antenna induces the secondary current that creates the electric signal to the receiver. Also for shorter wavelengths in the microwave and the optical regimes coherent states of the electromagnetic can be created, but not by oscillating macroscopic currents. In masers and lasers the intrinsic tendency of atoms to correlate their behavior in a strong electromagnetic field is used to create a monochromatic beam with a high degree of coherence. Ordinary light, on the other hand, as emitted by a hot source is highly incoherent, since the emission from different atoms only to a low degree is correlated. The expectation value for the field values at any point in space vanishes, and in this sense the light from a hot source is non-classical. However, this light does not corresponding to a *pure* quantum state, like the Fock state, but instead to a *mixed* state, with a probability distribution over Fock states.

The thermal photon gas

We examine now the quantum state of radiation from a hot source, which we assume to have the form of *black body radiation*, with a thermal *Boltzmann distribution* over energy eigenstates. The density operator of this state is

$$\hat{\rho} = \mathcal{N} e^{-\beta \hat{H}} \,, \quad \mathcal{N} = (\mathrm{Tr}\, e^{-\beta \hat{H}})^{-1} \,, \quad \beta = (k_B T)^{-1} \,, \tag{4.55}$$

where $\hat{H} = \sum_{\mathbf{k}a} \hbar \omega_k \hat{a}_{\mathbf{k}a}^\dagger \hat{a}_{\mathbf{k}a}$ is the Hamiltonian of the free electromagnetic field, k_B is Boltzmann's constant and T is the temperature of the radiation. We first calculate the expectation value of the photon numbers,

$$
\begin{aligned}
n_{\mathbf{k}a} &\equiv \left\langle \hat{N}_{\mathbf{k}a} \right\rangle \\
&= \mathcal{N} \,\mathrm{Tr}\, (e^{-\beta \hat{H}} \, \hat{a}_{\mathbf{k}a}^\dagger \hat{a}_{\mathbf{k}a}) \\
&= \mathcal{N} \,\mathrm{Tr}(e^{-\beta \hat{H}} \, \hat{a}_{\mathbf{k}a}^\dagger e^{\beta \hat{H}} \, e^{-\beta \hat{H}} \hat{a}_{\mathbf{k}a}) \\
&= \mathcal{N} \,\mathrm{Tr}(e^{-\beta \hbar \omega_k} \, \hat{a}_{\mathbf{k}a}^\dagger \, e^{-\beta \hat{H}} \hat{a}_{\mathbf{k}a}) \\
&= e^{-\beta \hbar \omega_k} \mathcal{N} \,\mathrm{Tr}(\, e^{-\beta \hat{H}} \hat{a}_{\mathbf{k}a} \hat{a}_{\mathbf{k}a}^\dagger) \\
&= e^{-\beta \hbar \omega_k} \mathcal{N} \,\mathrm{Tr}(e^{-\beta \hat{H}} (\hat{a}_{\mathbf{k}a}^\dagger \hat{a}_{\mathbf{k}a} + 1)) \\
&= e^{-\beta \hbar \omega_k} (n_{\mathbf{k}a} + 1) \,, \tag{4.56}
\end{aligned}
$$

where we have made use of the relation

$$e^{-\beta \hat{H}} \, \hat{a}_{\mathbf{k}a}^\dagger e^{\beta \hat{H}} = e^{-\beta \hbar \omega_k} \, \hat{a}_{\mathbf{k}a}^\dagger \,. \tag{4.57}$$

From (4.56) then follows

$$n_{\mathbf{k}a} = \frac{1}{e^{\beta \hbar \omega_k} - 1} \,, \tag{4.58}$$

with $\omega_k = ck$. This is the well-known *Bose-Einstein distribution* for photons in thermal equilibrium with a heat bath. From this distribution the *Planck spectrum* of the black body radiation can be determined. The total radiation energy is

$$
\begin{aligned}
\mathcal{E} &= \sum_{\mathbf{k}a} \hbar \omega_k n_{\mathbf{k}a} \\
&\to 2 \frac{V}{(2\pi)^3} \int d^3k \, \frac{\hbar \omega_k}{e^{\beta \hbar \omega_k} - 1} \\
&= \frac{V \hbar}{\pi^2 c^3} \int d\omega \, \frac{\omega^3}{e^{\beta \hbar \omega} - 1} \,. \tag{4.59}
\end{aligned}
$$

This corresponds to the following energy density per frequency unit to be

$$u(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta \hbar \omega} - 1} \,, \tag{4.60}$$
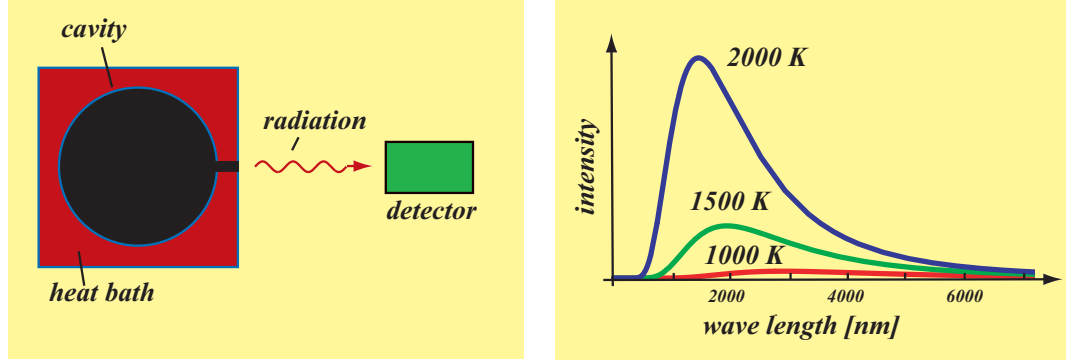
which displays the form of the Planck spectrum.

Figure 4.2: The Planck spectrum. A schematic experimental set up is shown where thermal light is created in a cavity surrounded by a heat bath (an oven). The radiation escapes through a narrow hole and can be analyzed in a detector. The right part of the figure shows the intensity (energy density) of the radiation as a function of wave length for three different temperatures.

We next consider the expectation values of the photon annihilation and creation operators. A similar calculation as for the number operator gives

$$
\begin{aligned}
\left\langle \hat{a}_{\mathbf{k}a}^{\dagger} \right\rangle &= \mathcal{N} \, \mathrm{Tr}(e^{-\beta H} \, \hat{a}_{\mathbf{k}a}^{\dagger}) \\
&= \mathcal{N} \, \mathrm{Tr}(e^{-\beta H} \, \hat{a}_{\mathbf{k}a}^{\dagger} e^{\beta H} \, e^{-\beta H}) \\
&= \mathcal{N} \, \mathrm{Tr}(e^{-\beta \hbar \omega_k} \, \hat{a}_{\mathbf{k}a}^{\dagger} \, e^{-\beta H}) \\
&= e^{-\beta \hbar \omega_k} \mathcal{N} \, \mathrm{Tr}(\, e^{-\beta H} \hat{a}_{\mathbf{k}a}^{\dagger}) \\
&= e^{-\beta \hbar \omega_k} \left\langle \hat{a}_{\mathbf{k}a}^{\dagger} \right\rangle .
\end{aligned}
\tag{4.61}
$$

This shows that the expectation value of the creation operator vanishes. A similar reasoning applies to the annihilation operator, so that

$$
\left\langle \hat{a}_{\mathbf{k}a}^{\dagger} \right\rangle = \left\langle \hat{a}_{\mathbf{k}a} \right\rangle = 0 \,.
\tag{4.62}
$$

As a consequence the expectation value of the electric and magnetic field vanish identically,

$$
\left\langle \mathbf{E}(\mathbf{r}, t) \right\rangle = \left\langle \mathbf{B}(\mathbf{r}, t) \right\rangle = 0 \,.
\tag{4.63}
$$

There are several ways to interpret this result, that the expectation values of the electromagnetic fields vanishes. We have represented the quantum state of the black body radiation as a statistical distribution over Fock states, and this indicates that the light from a hot source should be viewed as non-classical in the same sense as the pure Fock states. However, the vanishing of the expectation values of the fields only reflects rotational invariance, or the fact that black body radiation is *unpolarized*. This can be obtained also with statistical distributions over classical field configurations. In fact, the mixed quantum state of black body radiation can be expressed as statistical distributions over two completely different ensembles of pure states.

One is the ensemble of Fock states, that we have already discussed. The other is an ensemble of coherent states. Thus, if we write the density operator (4.55) as a product over modes,

$$\hat{\rho} = \prod_{\mathbf{k}a} \hat{\rho}_{\mathbf{k}a} \,, \tag{4.64}$$

the density operator of each mode has a thermal form with the two equivalent expansions

$$\hat{\rho}_{\mathbf{k}a} = \mathcal{N} \sum_n \exp(-\beta\hbar\omega n)|n\rangle\langle n| = \mathcal{N}' \int d^2\alpha \exp(-(e^{\beta\hbar\omega} - 1)|\alpha|^2)|\alpha\rangle\langle\alpha| \,. \tag{4.65}$$

In this expression $\omega = ck$ is the frequency of the mode, α is a coherent-state variable and $\mathcal{N}$ and $\mathcal{N}'$ are normalization factors. The identity between the two expansions is readily verified by using the expansion of coherent states in energy eigenstates.

This situation, that the density operator of a mixed state can be viewed as statistical distributions over completely different ensembles of pure quantum states, is something we have noticed already in the introductory discussion of pure and mixed states in Sect. 2.1.1. In the present case this means that it is our subjective choice to see the quantum state of the radiation either as a probability distribution over states with well defined photon numbers (the particle representation) or as a probability distribution over coherent states (the wave representation). Thus, the distinction between "classical light" and "non-classical light" is not so obvious for mixed quantum states as it is for pure states of the electromagnetic field.

Let us finally consider the question of how these different views of light are related to the classical demonstrations of the wave nature of light, in particular as shown in Young's interference experiments. These demonstrations may seem to favor the wave representation rather than the particle representation of light. However, even if this type of experiments demonstrate the wave nature of light, one should remember that interference is a direct consequence of *quantum superpositions* and is *not* depending of coherent behavior of many photons. We may compare this with the double slit experiment for electrons where interference can be seen as a single particle effect. Many electrons are needed to build up the interference pattern, but no coherent effect between the states of *different* electrons is needed. The interference can be seen as a single-electron effect, but the geometry of the experiment introduces a correlation in space between the pattern associated with each particle so that a macroscopic pattern can be built. In the same way we may interpret interference in incoherent light to be a single-photon effect. Each photon interferes with itself, without any coherent behavior with other photons. But they all see the same geometrical structure of the slits and this creates a correlated interference pattern.

4.2.4 Photon emission and photon absorption

We shall in the following consider processes where only a single electron is involved. To be more specific we may consider transitions in a hydrogen atom, or a hydrogen-like atom, where a single atomic electron can either absorb or emit a photon. The full Hamiltonian has the form

$$\hat{H} = \hat{H}_0^{field} + \hat{H}_0^{atom} + \hat{H}_{int} \,, \tag{4.66}$$

where

$$\hat{H}_0^{atom} = \frac{\hat{p}^2}{2m} + V(\hat{\mathbf{r}}) \,, \tag{4.67}$$

is the unperturbed Hamiltonian of the electron, which moves in the electrostatic potential V set up by the charges of the rest of the atom (*i.e.*, the nucleus and possibly other, more tightly bound electrons). The transitions are induced by the interaction part of the Hamiltonian of the electron and electromagnetic field,

$$\hat{H}_{int} = -\frac{e}{m}\hat{\mathbf{A}}(\hat{\mathbf{r}}) \cdot \hat{\mathbf{p}} + \frac{e^2}{2m}\hat{\mathbf{A}}(\hat{\mathbf{r}})^2 - \frac{e}{m}\hat{\mathbf{S}} \cdot \hat{\mathbf{B}}(\hat{\mathbf{r}}) \,, \tag{4.68}$$

where $\hat{\mathbf{r}}$ is the electron coordinate and $\hat{\mathbf{p}}$ is the (conjugate) momentum operator. The g-factor of the electron has here been set to 2.

The two first terms in H_{int} are *charge interaction terms* which describe interactions between the charge of the electron and the electromagnetic field. The third term is the *spin interaction term*, which describes interactions between the magnetic dipole moment of the electron and the magnetic field. We note that to lowest order in perturbation expansion, the first and third term of the interaction Hamiltonian (4.68) describe processes where a single photon is either absorbed or emitted. The second term describe scattering processes for a single photon and *two-photon* emission and absorption processes. This term is generally smaller than the first term and in a perturbative treatment it is natural to collect first order contributions from the second term with second order contributions from the first term. This means that we treat the perturbation series as an expansion in powers or the charge e (or rather the dimensionless fine-structure constant) instead of as an expansion in the interaction H_{int}.

The spin interaction term is also generally smaller than the first (charge interaction) term. However there are different selection rules for the transitions induced by these two terms, and when the direct contribution from the first term is forbidden the spin term may give an important contribution to the transition. Here we shall in the following restrict the discussion to transitions where the contribution from the first term is dominant and we therefore can neglect the two other terms. For simplicity we use the same notation H_{int} when only the first term is included.

The interaction Hamiltonian we may now separated in a creation (emission) part and an annihilation (absorption) part

$$\hat{H}_{int} = \hat{H}_{emis} + \hat{H}_{abs} \,. \tag{4.69}$$

Separately they are are non-hermitian with $\hat{H}^{\dagger}_{emis} = \hat{H}_{abs}$. Expressed in terms of photon creation and annihilation operators they are

$$\hat{H}_{emis} = -\frac{e}{m}\sum_{\mathbf{k}a}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}\,\hat{\mathbf{p}} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}\hat{a}^{\dagger}_{\mathbf{k}a}e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \,,$$

$$\hat{H}_{abs} = -\frac{e}{m}\sum_{\mathbf{k}a}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}\,\hat{\mathbf{p}} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}\hat{a}_{\mathbf{k}a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \,. \tag{4.70}$$

Note that when written in this way the operators are expressed in the *interaction picture* where the time evolution is determined by the free (non-interacting) theory. The time evolution of the state vectors are in this picture determined by the interaction Hamiltonian only, not by the free (unperturbed) Hamiltonian. This picture is most conveniently used in perturbative expansions, as previously discussed in Sect. 1.2.1.

From the above expressions we can find the interaction matrix elements corresponding to emission and absorption of a single photon. We note that to first order the photon number only

of one mode $(\mathbf{k}, a)$ is changed and we write the photon number only of this mode explicitly. We write the initial and final states as

$$
\begin{aligned}
|i\rangle &= |A, n_{\mathbf{k}a}\rangle, \\
|f\rangle &= |B, n_{\mathbf{k}a} \pm 1\rangle,
\end{aligned}
\tag{4.71}
$$

where $|A\rangle$ is the (unspecified) initial state of the atom (*i.e.*, the electron state), $|B\rangle$ is the final state of the atom, and $n_{\mathbf{k}a}$ is the photon number of the initial state. This means that we consider transitions between Fock states of the electromagnetic field.

For absorption the matrix element is

$$
\begin{aligned}
\langle B, n_{\mathbf{k}a} - 1|\hat{H}_{abs}|A, n_{\mathbf{k}a}\rangle &= -\frac{e}{m}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}\, \langle B, n_{\mathbf{k}a} - 1|\hat{\mathbf{p}} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}\, \hat{a}_{\mathbf{k}a} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}|A, n_{\mathbf{k}a}\rangle \\
&= -\frac{e}{m}\sqrt{\frac{\hbar n_{\mathbf{k}a}}{2V\epsilon_0\omega_k}}\, \boldsymbol{\epsilon}_{\mathbf{k}a} \cdot \langle B|\hat{\mathbf{p}}e^{i\mathbf{k}\cdot\mathbf{r}}|A\rangle e^{-i\omega_k t},
\end{aligned}
\tag{4.72}
$$

and the corresponding expression for photon emission is

$$
\begin{aligned}
\langle B, n_{\mathbf{k}a} + 1|\hat{H}_{emis}|A, n_{\mathbf{k}a}\rangle &= -\frac{e}{m}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_k}}\, \langle B, n_{\mathbf{k}a} + 1|\hat{\mathbf{p}} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}\, \hat{a}_{\mathbf{k}a}^{\dagger} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}|A, n_{\mathbf{k}a}\rangle \\
&= -\frac{e}{m}\sqrt{\frac{\hbar(n_{\mathbf{k}a} + 1)}{2V\epsilon_0\omega_k}}\, \boldsymbol{\epsilon}_{\mathbf{k}a} \cdot \langle B|\hat{\mathbf{p}}e^{-i\mathbf{k}\cdot\mathbf{r}}|A\rangle e^{i\omega_k t}.
\end{aligned}
\tag{4.73}
$$

In the final expressions of both (4.72) and (4.73) note that only the matrix elements for the electron operator between the initial and final states A and B remain, while the effect of the photon operators is absorbed in the prefactor, which now depends on the photon number of the initial state.

It is of interest to note that the electron matrix elements found for the interactions with a *quantized* electromagnetic field is quite analogous to those found for interaction with a classical *time-dependent* electromagnetic field of the form

$$
\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_0\, e^{i(\mathbf{k}\cdot r - \omega_k t)} + \mathbf{A}_0^*\, e^{-i(\mathbf{k}\cdot r - \omega_k t)},
\tag{4.74}
$$

where the positive frequency part of the field (*i.e.*, the term proportional to $e^{-i\omega_k t}$) corresponds to the absorption part of the matrix element and the negative frequency part (proportional to $e^{i\omega_k t}$) corresponds to the emission part. The use of this expression for the $\mathbf{A}$-field gives a *semi-classical* approach to radiation theory, which in many cases is completely satisfactory. It works well for states with a large number $n_{\mathbf{k}a}$ of photons, where the strength of the matrix elements for absorption and emission are essentially equal. However, for small photon numbers there is a non-negligible difference in the strength of the two matrix elements, a difference which is not reflected in the classical amplitude (4.74). In the case of *spontaneous emission*, with $n_{\mathbf{k}a} = 0$ in the initial state, the quantum theory correctly describes the transition of the electron from an excited state by emission of a photon, whereas a semiclassical description does not, since electronic transitions in this approach depends on the presence of an oscillating electromagnetic field.

4.2.5   Dipole approximation and selection rules

For radiation from an atom, the wave length of the radiation field is typically much larger than the dimension of the atom. This difference is exemplified by the wave length of blue light and the Bohr radius of the hydrogen atom

$$\lambda_{blue} \approx 400\text{nm} , \quad a_0 = 4\pi\epsilon_0 \frac{\hbar^2}{me^2} \approx 0.05\text{nm} . \tag{4.75}$$

This means that the effect of spacial variations in the electromagnetic field over the dimensions of an atom are small, and therefore the *time-variations* rather than the space variations of the field are important. In the expression for the transition matrix elements (4.72) and (4.73) this justifies an expansion of the phase factors in powers of $\mathbf{k} \cdot \mathbf{r}$,

$$e^{\pm i\mathbf{k}\cdot\mathbf{r}} = 1 \pm i\mathbf{k}\cdot\mathbf{r} - \frac{1}{2}(\mathbf{k}\cdot\mathbf{r})^2 + ... , \tag{4.76}$$

where the first term is dominant. The approximation where only this term is kept is referred to as the *dipole approximation*. The other terms give rise to higher *multipole* contributions. These may give important contributions to atomic transitions only when the contribution from the first term vanishes due to a selection rule. However the transitions dominated by higher multipole terms are normally much slower than the ones dominated by the dipole contribution.

When the dipole approximation is valid, the matrix elements of the interaction Hamiltonian simplify to

$$\langle B, n_{\mathbf{k}a} - 1 | \hat{H}_{abs} | A, n_{\mathbf{k}a} \rangle = -\frac{e}{m}\sqrt{\frac{\hbar n_{\mathbf{k}a}}{2V\epsilon_0\omega_k}}\ \boldsymbol{\epsilon}_{\mathbf{k}a} \cdot \mathbf{p}_{BA}\ e^{-i\omega_k t} ,$$

$$\langle B, n_{\mathbf{k}a} + 1 | \hat{H}_{emis} | A, n_{\mathbf{k}a} \rangle = -\frac{e}{m}\sqrt{\frac{\hbar(n_{\mathbf{k}a}+1)}{2V\epsilon_0\omega_k}}\ \boldsymbol{\epsilon}_{\mathbf{k}a} \cdot \mathbf{p}_{BA}\ e^{i\omega_k t} , \tag{4.77}$$

where $\mathbf{p}_{BA}$ is the matrix element of the operator $\hat{\mathbf{p}}$ between the states $|A\rangle$ and $|B\rangle$. It is convenient to re-express it in terms of the matrix elements of the position operator $\hat{\mathbf{r}}$, which can be done by use of the form of the (unperturbed) electron Hamiltonian,

$$\hat{H}_0^{atom} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}) , \tag{4.78}$$

where $V(\hat{\mathbf{r}})$ is the (Coulomb) potential felt by the electron. This gives

$$\left[\hat{H}_0^{atom}, \hat{\mathbf{r}}\right] = -i\frac{\hbar}{m}\hat{\mathbf{p}} . \tag{4.79}$$

With the initial state $|A\rangle$ and the final state $|B\rangle$ as eigenstates of $\hat{H}_0^{atom}$, with eigenvalues E_A and E_B, we find for the matrix element of the momentum operator between the two states

$$\mathbf{p}_{BA} = i\frac{m}{\hbar}(E_B - E_A)\mathbf{r}_{BA} = im\omega_{BA}\ \mathbf{r}_{BA} , \tag{4.80}$$

where we have introduced $\omega_{AB} = (E_A - E_B)/\hbar$. Note that in the expression for the interaction, the change from $\mathbf{p}_{BA}$ to $\mathbf{r}_{BA}$ corresponds to the transformation

$$-\frac{e}{m}\mathbf{A}\cdot\mathbf{p} \rightarrow e\mathbf{r}\cdot\dot{\mathbf{A}} = -e\mathbf{r}\cdot\mathbf{E} , \tag{4.81}$$

where the last term is identified as the electric dipole energy of the electron. This form justifies the name *electric dipole transitions* for transitions induced by this operator.

By use of the identity (4.80) the interaction matrix elements are finally written as

$$
\begin{aligned}
\langle B, n_{\mathbf{k}a} - 1|\hat{H}_{abs}|A, n_{\mathbf{k}a}\rangle &= -ie\sqrt{\frac{\hbar n_{\mathbf{k}a}\omega_{BA}^2}{2V\epsilon_0\omega_k}}\ \boldsymbol{\epsilon}_{\mathbf{k}a}\cdot\mathbf{r}_{BA}\ e^{-i\omega_k t}\,, \\
\langle B, n_{\mathbf{k}a} + 1|\hat{H}_{emis}|A, n_{\mathbf{k}a}\rangle &= ie\sqrt{\frac{\hbar (n_{\mathbf{k}a}+1)\omega_{AB}^2}{2V\epsilon_0\omega_k}}\ \boldsymbol{\epsilon}_{\mathbf{k}a}\cdot\mathbf{r}_{BA}\ e^{i\omega_k t}\,.
\end{aligned}
$$
(4.82)

As we shall later see these expressions may in most cases be simplified by use of *energy conservation*, in the form $\omega_k = \omega_{BA}$ for photon absorption and $\omega_k = \omega_{AB}$ for photon emission.

Selection rules
The matrix elements

$$
\mathbf{r}_{BA} = \langle B|\hat{\mathbf{r}}|A\rangle\,,
$$
(4.83)

are subject to certain *selection rules* which follow from conservation of spin and parity. Thus, the operator $\hat{\mathbf{r}}$ transforms as a *vector* under rotation and changes sign under space inversion. Since a vector is a *spin 1* quantity, the operator can change the spin of the state $|A\rangle$ by maximally one unit of spin. Physically we interpret this as due to the spin carried by the photon. The change in sign under space inversion corresponds to the parity of the photon being -1. This change implies that the parity of the final state is opposite that of the initial state.

Let us specifically consider (orbital) angular momentum states of the atomic electron, characterized by the quantum numbers l and m (total angular momentum and angular momentum in the z-direction). The initial atomic state A is then characterized by quantum numbers l_A and m_A, while the parity is $P_A = (-1)^{l_A}$. The corresponding quantum numbers for the state B are l_B, m_A and $P_B = (-1)^{l_B}$. We note that the change in parity forces the angular moment to change by one unit ($l_A \neq l_B$) and the formula for spin addition gives $l_B = l_A \pm 1$. The selection rules for the electric dipole transitions (referred to as $E1$ transitions) are therefore

$$
\begin{aligned}
\Delta l &= \pm 1\ (l_A \neq 0)\,, \quad \Delta l \equiv l_B - l_A\,, \\
\Delta l &= +1\ (l_A = 0)\,, \\
\Delta m &= 0, \pm 1\,, \qquad \Delta m \equiv m_B - m_A\,.
\end{aligned}
$$
(4.84)

The transitions that do not satisfy these rules are "forbidden" in the sense that the interaction matrix element vanishes in the dipole approximation. Nevertheless such transitions may take place, but as a much slower rate than the $E1$ transitions. They may be induced by higher multipole terms in the expansion (4.76), or by higher order terms in $\mathbf{A}$ which give rise to multi-photon processes. The multipole terms include higher powers in the components of the position operator, and these terms transform differently under rotation and space inversion than the vector $\mathbf{r}$. As a consequence they are restricted by other selection rules. Physically we may consider the higher multipole transitions as corresponding to non-central photon emission and absorption, where the total spin transferred is not only due to the intrinsic photon spin but also orbital angular momentum of the photon.

4.3   Photon emission from excited atom

In this section we examine the photon emission process for an excited atom in some more detail. We make the assumption that the atom at a given instant $t = t_i$ is in an excited state $|A\rangle$ and consider the amplitude for making a transition to a state $|B\rangle$ at a later time $t = t_f$ by emission of a single photon. The amplitude is determined to first order in the interaction, and the transition probability per unit time and and unit solid angle for emission of the photon in a certain direction is evaluated and expressed in terms of the dipole matrix element. We subsequently discuss the effect of decay of the initial atomic state and its relation to the formation of a *line width* for the photon emission line.

4.3.1   First order transition and Fermi's golden rule

The perturbation expansion of the time evolution operator in the interaction picture is (see Sect. 1.3.1)

$$\hat{U}_{int}(t_f, t_i) = \mathbf{1} - \frac{i}{\hbar} \int_{t_i}^{t_f} dt \hat{H}_{int}(t) + \frac{1}{2}(-\frac{i}{\hbar})^2 \int_{t_i}^{t_f} dt \int_{t_i}^{t} dt' \hat{H}_{int}(t) \hat{H}_{int}(t') + ... , \quad (4.85)$$

where the time evolution of the interaction Hamiltonian is

$$\hat{H}_{int}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t} \hat{H}_{int} e^{-\frac{i}{\hbar}\hat{H}_0 t} , \quad (4.86)$$

with $\hat{H}_0$ as the unperturbed Hamiltonian. The transition matrix element between an initial state $|i\rangle$ at time t_i and final state $|f\rangle$ at time t_f is

$$\langle f|\hat{U}_{int}(t_f, t_i)|i\rangle = \langle f|i\rangle - \frac{i}{\hbar}\langle f|\hat{H}_{int}|i\rangle \int_{t_i}^{t_f} dt e^{\frac{i}{\hbar}(E_f - E_i)t}$$

$$+\frac{1}{2}(-\frac{i}{\hbar})^2 \sum_m \langle f|\hat{H}_{int}|m\rangle\langle m|\hat{H}_{int}|i\rangle \int_{t_i}^{t_f} dt \int_{t_i}^{t} dt' e^{\frac{i}{\hbar}(E_f - E_m)t} e^{\frac{i}{\hbar}(E_m - E_i)t'} + ... (4.87)$$

In this expressions we have assumed that the initial state $|i\rangle$ and the final state $|f\rangle$ as well as a complete set of intermediate states $|m\rangle$ are eigenstates of the unperturbed Hamiltonian $\hat{H}_0$. The corresponding eigenvalues are E_i, E_f and E_m. We perform the time integrals, and in order to simplify expressions we introduce the notation $\omega_{fi} = (E_f - E_i)/\hbar$, $T = t_f - t_i$ and $\bar{t} = (t_i + t_f)/2$. To second order in the interaction the transition matrix element is

$$\langle f|\hat{U}_{int}(t_f, t_i)|i\rangle = \langle f|i\rangle$$

$$-2i\frac{\sin[\frac{1}{2}\omega_{fi}T]}{\hbar\omega_{fi}}e^{i\omega_{fi}\bar{t}}\Big[\langle f|\hat{H}_{int}|i\rangle - \sum_m \frac{\langle f|\hat{H}_{int}|m\rangle\langle m|\hat{H}_{int}|i\rangle}{\hbar\omega_{mi}} + ...\Big]$$

$$-2ie^{i\omega_{fi}\bar{t}}\sum_m \sin[\frac{1}{2}\omega_{fm}T]e^{i\omega_{mi}T}\frac{\langle f|\hat{H}_{int}|m\rangle\langle m|\hat{H}_{int}|i\rangle}{\hbar^2\omega_{fm}\,\omega_{mi}} + ... , \quad (4.88)$$

where we assume the *diagonal* matrix elements of $\hat{H}_{int}$ to vanish in order to avoid ill-defined terms in the expansion. The factor depending on $\bar{t}$ is unimportant and can be absorbed in a redefinition of the time coordinate so, that $\bar{t} = 0$. (The interesting time dependence lies in the relative coordinate $T = t_f - t_i$.) The last term in (4.88) does not contribute (at average) to low order due to rapid oscillations. Without this term the result simplifies to

$$\langle f|\hat{U}_{int}(t_f, t_i)|i\rangle = \langle f|i\rangle - 2i\frac{\sin[\frac{1}{2}\omega_{fi}\, T]}{\hbar\omega_{fi}}e^{i\omega_{fi}\bar{t}}\,\mathcal{T}_{fi}\,, \tag{4.89}$$

where $\mathcal{T}_{fi}$ is the *T-matrix* element

$$\mathcal{T}_{fi} = \left[\langle f|\hat{H}_{int}|i\rangle - \sum_m \frac{\langle f|\hat{H}_{int}|m\rangle\langle m|\hat{H}_{int}|i\rangle}{\hbar\omega_{mi}} + ...\right]. \tag{4.90}$$

The transition probability for $f \neq i$ is

$$W_{fi} = \left(\frac{2\sin[\frac{1}{2}\omega_{fi}\, T]}{\hbar\omega_{fi}}\right)^2 |\mathcal{T}_{fi}|^2\,. \tag{4.91}$$

Also this is an oscillating function, but for large T it gives a contribution proportional to T. To see this we consider the function

$$f(x) = \left(\frac{\sin x}{x}\right)^2\,. \tag{4.92}$$

The function is shown in fig.4.3. It is localized around $x = 0$ with oscillations that are damped
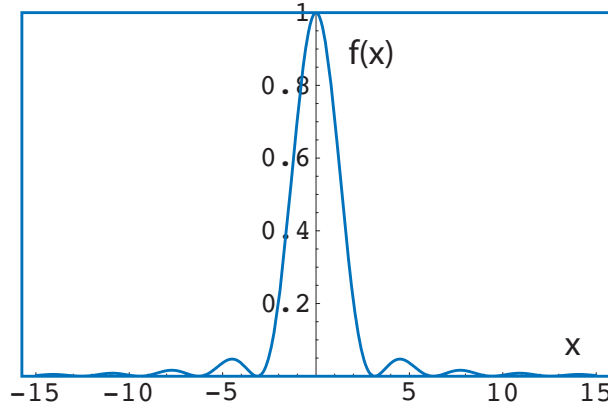

Figure 4.3: Frequency dependence of the transition probability. For finite transition time T the function has a non-vanishing width. In the limit $T \to \infty$ the function tends to a delta-function.

like $1/x^2$ for large x. The integral of the function is

$$\int_{-\infty}^{\infty} f(x)dx = \pi\,. \tag{4.93}$$

The prefactor of (4.91) can be expressed in terms of the function $f(x)$ as

$$\left(\frac{2\sin[\frac{1}{2}\omega_{fi}\,T]}{\hbar\omega_{fi}}\right)^2 = \frac{T^2}{\hbar^2}f(\frac{1}{2}\omega_{fi}\,T)\,. \tag{4.94}$$

When regarded as a function of ω_{fi} this function gets increasingly localized around $\omega_{fi} = 0$ as T increases, and in the limit $T \to \infty$ it approaches a delta function, with the strength of the delta function being determined by the (normalization) integral (4.93),

$$\left(\frac{2\sin[\frac{1}{2}\omega_{fi}\,T]}{\hbar\omega_{fi}}\right)^2 \to \frac{2\pi}{\hbar}T\delta(\hbar\omega_{fi})\,. \tag{4.95}$$

This gives a constant transition rate between the initial and final states,

$$w_{fi} = \frac{W_{fi}}{T} = \frac{2\pi}{\hbar}|\mathcal{T}_{fi}|^2\delta(E_f - E_i)\,, \tag{4.96}$$

where $\mathcal{T}_{fi}$ to lowest order in the interaction is simply the interaction matrix element. When applied to transitions in atoms, the expression (4.96) for the constant transition rate is often referred to as *Fermi's golden rule*. The delta function expresses energy conservation in the process. Note, however, that only in the limit $T \to \infty$ the energy dependent function is really a delta function. For finite time intervals there is a certain width of the function which means that E_f can deviate slightly from E_i. This apparent breaking of energy conservation for finite times may happen since E_f and E_i are eigenvalues of the *unperturbed* Hamiltonian rather than the *full* Hamiltonian.

In reality the time T cannot be taken to infinity for an atomic emission process, since any excited state will have a *finite life time*. The width of the energy-function then has a physical interpretation in terms of a *line width* for the emission line. At this point we neglect this effect of *level broadening*, but we shall return to a description of this effect a coming section.

4.3.2   Emission rate

We consider now the case where initially the atom is in an excited state A and finally in a state B with one photon being emitted. The initial and final states of the full quantum system are

$$|i\rangle = |A, 0\rangle\,, \quad |f\rangle = |B, 1_{\mathbf{k}a}\rangle\,, \tag{4.97}$$

where 0 in the initial state indicates the photon vacuum, while $1_{\mathbf{k}a}$ indicates one photon with quantum numbers $\mathbf{k}a$. We will be interested in finding an expression for the *differential* transition rate, *i.e.*, the transition probability per unit time and unit solid angle, as well as the *total* transition rate.

We write the transition rate summed over all final states of the photon as

$$\begin{aligned}
w_{BA} &= \sum_{\mathbf{k}a}\frac{2\pi}{\hbar}|\langle B, 1_{\mathbf{k}a}|\hat{H}_{emis}|A, 0\rangle|^2\delta(E_A - E_B - \hbar\omega_k) \\
&\to \frac{V}{(2\pi)^3}\int d^3k\sum_{a}\frac{2\pi}{\hbar}|\langle B, 1_{\mathbf{k}a}|\hat{H}_{emis}|A, 0\rangle|^2\delta(E_A - E_B - \hbar\omega_k)
\end{aligned}$$

$$= \frac{V}{(2\pi)^2 \hbar} \int d\Omega \int_0^\infty dk k^2 \sum_a |\langle B, 1_{\mathbf{k}a}|\hat{H}_{emis}|A, 0\rangle|^2 \delta(E_A - E_B - \hbar\omega_k)$$

$$= \frac{V\omega_{BA}^2}{(2\pi)^2 c^3 \hbar^2} \int d\Omega \sum_a |\langle B, 1_{\mathbf{k}a}|\hat{H}_{emis}|A, 0\rangle|^2 , \tag{4.98}$$

where the infinite volume limit has been taken in the change from a discrete momentum sum to a momentum integral. The subsequent integration of $k = \omega_k/c$ over the delta function then fixes the frequency of the emitted photon to match the atomic frequency, $\omega_k = \omega_{AB} = (E_A - E_B)/\hbar$. By use of the expression for the dipole transmission matrix element (4.82), we find the following expression for differential emission rate,

$$\frac{dw_{BA}}{d\Omega} = \frac{e^2 \omega_{AB}^3}{8\pi^2 \epsilon_0 \hbar c^3} |\mathbf{r}_{BA} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}^*|^2 , \tag{4.99}$$

where $\boldsymbol{\epsilon}_{\mathbf{k}a}$ is the polarization vector of the emitted photon.

Summed over photon states we have

$$\sum_a |\mathbf{r}_{BA} \cdot \boldsymbol{\epsilon}_{\mathbf{k}a}^*|^2 = |\mathbf{r}_{BA}|^2 - \frac{|\mathbf{r}_{BA} \cdot \mathbf{k}|^2}{\mathbf{k}^2} . \tag{4.100}$$

From this we find the total transition rate into all final states of the photon,

$$\begin{aligned}
w_{BA} &= \frac{e^2 \omega_{AB}^3}{8\pi^2 \epsilon_0 \hbar c^3} \int d\Omega \left[|\mathbf{r}_{BA}|^2 - \frac{|\mathbf{r}_{BA} \cdot \mathbf{k}|^2}{\mathbf{k}^2} \right] \\
&= \frac{e^2 \omega_{AB}^3}{4\pi \epsilon_0 \hbar c^3} |\mathbf{r}_{BA}|^2 \int_0^\pi d\theta \sin\theta (1 - \cos^2\theta) \\
&= \frac{e^2 \omega_{AB}^3}{4\pi \epsilon_0 \hbar c^3} |\mathbf{r}_{BA}|^2 \int_{-1}^{+1} du (1 - u^2) \\
&= \frac{e^2 \omega_{AB}^3}{3\pi \epsilon_0 \hbar c^3} |\mathbf{r}_{BA}|^2 \\
&= \frac{4\alpha}{3c^2} \omega_{AB}^3 |\mathbf{r}_{BA}|^2 , \tag{4.101}
\end{aligned}$$

where α is the fine structure constant. This expression shows how the transition rate depends on the dipole matrix element $|\mathbf{r}_{BA}|^2$ and on the energy $\hbar\omega_{AB}$ released in the transition.[4]

The formalism employed here for the case of spontaneous emission can also be used to describe photon *absorption* processes and *scattering* of photons on atoms. In the latter case the amplitude should be calculated to second order in the electric charge, and that would involve the $\mathbf{A} \cdot \mathbf{A}$-term as well as the $\mathbf{A} \cdot \mathbf{p}$-term of the interaction. We note from the expression (4.101) that the emission rate increases strongly with the energy of the emitted photon. To a part that can be seen as a *density of states* effect, since for high energy photons many more states are available (within a fixed energy interval) than for low energy photons. This is demonstrated

---

[4]Note that in the evaluation we have treated $\mathbf{r}_{BA}$ as a real vector. In reality $\mathbf{r}_{BA}$ may be complex, but this does not change the result, one only has to repeat the evaluation for the real and imaginary parts separately.

explicitly by (4.98) which shows that a factor ω_{BA}^2 can be ascribed to the frequency dependence of the density of states. A similar effect takes place when light scattering is considered. This gives a qualitative explanation for why blue light is more readily scattered than red light, and thereby why the sky is blue and the sunset is red.

4.3.3   Life time and line width

Fermi's golden rule, which gives a constant transition rate from the excited to the lower energy level of the atom, can be correct only in an approximate sense. Thus, the corresponding expression for the integrated transition probability,

$$W_{BA}(t) = t w_{BA} \,, \tag{4.102}$$

which shows that W_{BA} increases linearly with time, can obviously be correct only for a sufficiently short time interval $t < w_{BA}^{-1}$. One way to view this problem is that the expression (4.91) does not take into account the fact that the probability of state A to be occupied is reduced during the transition. This motivates the following modification of the expression for the transition rate,

$$\frac{d}{dt} W_{BA} = w_{BA} P_A(t) \,, \tag{4.103}$$

where w_{BA} is the time independent rate determined from the golden rule and $P_A(t)$ is the probability for level A to be occupied after time t. This probability is on the other hand related to the sum of transition rates over all final states B,

$$P_A(t) = 1 - \sum_B W_{BA}(t) \,, \tag{4.104}$$

and taking the time derivative and making use of (4.103), we find

$$
\begin{aligned}
\frac{d}{dt} P_A &= -\sum_B \frac{d}{dt} W_{BA} \\
&= -\sum_B w_{BA} P_A \\
&\equiv -P_A/\tau_A \,, \tag{4.105}
\end{aligned}
$$

where $\tau_A = (\sum\limits_B w_{BA})^{-1}$. This can be integrated to give

$$P_A(t) = e^{-t/\tau_A} \,, \tag{4.106}$$

which shows an exponential decay of the excited state A, with τ_A as the *life time* of the state. From this follows that the corrected transition rate is

$$\frac{d}{dt} W_{BA} = w_{BA} e^{-t/\tau_A} \,, \tag{4.107}$$

which integrates to

$$W_{BA} = w_{BA}\, \tau_A \left(1 - e^{-t/\tau_A}\right) , \tag{4.108}$$

The original expression (4.91) for the transition probability corresponds to (4.108) expanded to first order in t/τ_A, and is clearly a good approximation to the full expression only when $t << \tau_A$.

The decay of the initial state can be built into the expressions for the transitions to lower energy states in a simple way by including a damping factor $exp(-t/2\tau_A)$ as a normalization factor in the amplitude. For the transition amplitude this gives a modified expression

$$
\begin{aligned}
\langle f|\hat{U}_{int}(t_f,t_i)|i\rangle &= -\frac{i}{\hbar}\langle f|\hat{H}_{int}|i\rangle \int_{t_i}^{t_f} dt e^{\frac{i}{\hbar}(E_f-E_i)t} e^{-\frac{(t-t_i)}{2\tau_A}} + ... \\
&= -\frac{i}{\hbar}\langle f|\hat{H}_{int}|i\rangle e^{\Gamma_A t_i/2\hbar} \int_{t_i}^{t_f} dt e^{\frac{i}{\hbar}(E_f-E_i+\frac{i}{2}\Gamma_A)t} + ... \\
&= -\frac{\langle f|\hat{H}_{int}|i\rangle}{E_f-E_i+\frac{i}{2}\Gamma_A}(e^{\frac{i}{\hbar}(E_f-E_i)t_f}e^{-\frac{1}{2\hbar}\Gamma_A(t_f-t_i)} - e^{\frac{i}{\hbar}(E_f-E_i)t_i}) + ... \\
&\to \frac{\langle f|\hat{H}_{int}|i\rangle}{E_f-E_i+\frac{i}{2}\Gamma_A}e^{\frac{i}{\hbar}(E_f-E_i)t_i} + ... ,
\end{aligned}
\tag{4.109}
$$

where we have introduced $\Gamma_A = \hbar/\tau_A$ and at the last step assumed $t_f - t_i >> \tau_A$. With the initial energy $E_i = E_A$ and the final energy $E_f = E_B + \hbar\omega_k$, the expression shows that the correction due to a finite life time of the atomic level A can be viewed as represented by an imaginary contribution to the energy of this level, $E_A \to E_A - (i/2)\Gamma_A$.

The transition probability is then, for $t_f - t_i >> \tau_A$, given by

$$
W_{fi} = \frac{|\langle f|\hat{H}_{int}|i\rangle|^2}{(E_A - E_B - \hbar\omega_k)^2 + \frac{1}{4}\Gamma_A^2} \cdot + ... \tag{4.110}
$$

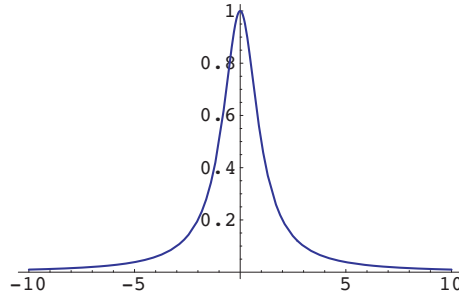Viewed as a function of ω_k, the transition probability is no longer proportional to a delta

Figure 4.4: Level broadening. The emission probability given by $W_{fi} = |\langle f|\hat{H}_{int}|i\rangle|^2/[(\hbar\omega_{fi})^2 + \frac{1}{4}\Gamma_A^2]$ is shown as a function of ω_{fi}.

function in the difference $\omega_k - \omega_{AB}$. The delta function is instead replaced by a *Lorentzian* (Fig.4.4), which is strongly peaked around $\omega_{fi} = 0$, but which has a finite width proportional

to Γ_A. For the emitted photon this is translated to a finite *line width* for the emission line corresponding to the transition $A \to B$.

If we compare the two figures 4.3 and 4.4 we note that a finite cut-off in the time integral (*i.e.*, a finite value of $T = t_f - t_i$) gives essentially the same frequency dependence for the emitted photon as the one obtained by an exponential damping due to the finite life time of the initial atomic state. We finally note that since Γ_A depends on the transition probabilities, to include it in the lowest order transition amplitude means that we effectively have included higher order contributions. This means that the perturbative expansion is no longer an order by order expansion in the interaction Hamiltonian.

4.4   Stimulated photon emission and the principle of lasers

As we have already noticed, the rate for *stimulated* emission of a photon by an atomic transition is larger than the corresponding *spontaneous* emission. Thus, to lowest order in the interaction, the transition probability by photon emission into a given mode is enhanced by a factor n equal to the number of photons already present in the mode. This means that if an atom in an excited state emits a photon with equal probability in all directions, the same atom, when placed in a strong field which resonates with the atomic transition, will preferably emit the photon into the mode that is already occupied. However, in free space, since a continuum of modes is available, the probability for emission into the preferred will may still be small.
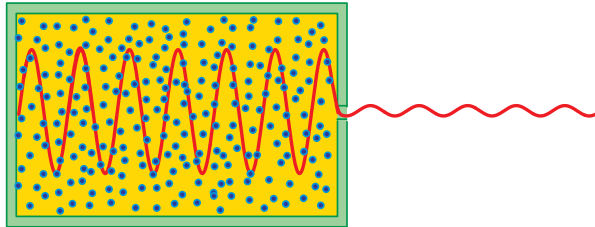

Figure 4.5: Schematic representation of a laser. A gas of atoms is trapped inside a *Fabry-Perot* cavity with reflecting walls. The atoms are *pumped* to an excited state that subsequently make transitions to a lower energy mode. The emitted photons resonate with a longitudinal mode of the cavity and build up a strong field in this mode. Some of the light from this mode escapes through a semi-reflective mirror in one end of the cavity to form a monochromatic laser beam.

A laser is based on the principle of stimulated emission, but the probability for emission into the preferred mode is enhanced by use of a reflecting cavity. The boundary conditions imposed by the reflecting mirrors of the cavity reduce the number of available electromagnetic modes to a discrete set, and the trapping of the photons makes it possible to build a large population of photons in one of the modes.

A schematic picture of a laser is shown in fig.4.5 where a elongated cavity is filled with gas of (Rubidium) atoms. These are continuously excited (pumped) from the ground state to a higher energy state. The excited atoms subsequently make transitions to a lower energy state and emit photons with energy that matches one of the modes selected by the distance between

the mirrors in the longitudinal direction. The emission tends to increase the excitation of the preferred mode to a high level. Some of the light that builds up inside the cavity escapes through a small hole in one of the mirrors in the form of a laser beam.

The laser light has the same characteristics as that of a *classical, momochromatic* beam of light. This means that it has a sharply defined frequency, with a long temporal phase coherence. It is therefore usually considered to be well described as a coherent state of photons. The coherence of the laser light follows from how the stimulated emission builds up a strong field in a single mode of the laser cavity. The coupling to the atoms involved in the stimulated emission, and to the electromagnetic field outside the cavity, will however introduce some degree of decoherence in the laser light.

4.4.1 Three-level model of a laser

We consider a simple model of a laser where a single mode is populated by stimulated emission. A three-level model is used for the atom, where most atoms are in the ground state $|0\rangle$, but where there is a continuous rate of "pumping" of atoms to a higher energy level $|2\rangle$, by a light source or some other way of excitation. The atoms in the upper level next make transitions to an intermediate level $|1\rangle$ by emission of a photon. This may be by stimulated emission of a photon to the field mode that is already strongly populated, or by spontaneous emission to one of the other photon states in the emission line that are allowed by the boundary conditions of the cavity. From the state $|1\rangle$ there is a fast transition back to the ground state due to a strong coupling between the levels $|1\rangle$ and $|0\rangle$.

Due to the assumed strength of the transitions we have

$$N_0 >> N_2 >> N_1 \,. \tag{4.111}$$

This means that the number of atoms in the ground state N_0 is essentially identical to the total number of atoms N. When we consider the two levels 1 and 2, which are relevant for populating the laser mode, the inequality means that there is *population inversion*, since the upper level is more strongly populated than the lower mode. Population inversion is a condition for the atoms to be able to "feed" the preferred photon mode, since the probability of reabsorbing a photon from the preferred mode by the inverse transition $|1\rangle \rightarrow |2\rangle$ is much smaller than emission of a photon by the process $|2\rangle \rightarrow |1\rangle$.

When the atoms are pumped to a higher energy level, this creates initially a situation where the photon modes within the line width of the transition $2 \rightarrow 1$ begins to get populated, but where no single mode is preferred. If the pumping rate R is sufficiently large, this is an unstable situation, due to the effect of stimulated emission to preferably populate a mode which already is excited. As a result one of the modes will spontaneously tend to grow at the expense of the others. We will consider the situation after one of the modes has been preferred in this way.

The transition from state 2 to 1 may now go in two ways, either by spontaneous emission to one of the modes that have not been populated or by stimulated emission to the preferred (laser) mode. We denote the transition rate by spontaneous emission Γ_{sp} and by stimulated emission $\Gamma_{st}\, n$, where n is the photon number of the laser mode. The difference between Γ_{sp} and Γ_{st} is due to the large number available for spontaneous emission compared to the single
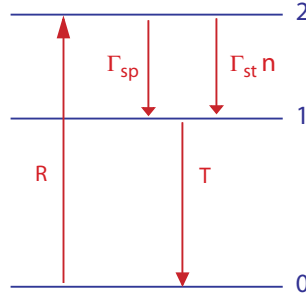
Figure 4.6: Transitions between three atomic levels in a laser. R is the pumping rate from the ground state 0 to an excited state 2. The transition from level 2 to an intermediate level 1 can go either by spontaneous emission (Γ_{sp}), or by stimulated emission ($\Gamma_{st}\ n$) to the laser mode. n is the photon number of the laser mode. A fast transition T brings the atoms back to the ground state. The direct transition rate $2 \to 0$ is assumed to be negligible.

mode available for stimulated emission. Typical values are

$$\Gamma_{sp} \approx 10^7 s^{-1} , \ \ \Gamma_{st} \approx 1 s^{-1} . \tag{4.112}$$

The smallness of Γ_{st} is compensated by the factor n and when a stationary situation is reached the photon number n in the laser mode will be so large so that the probabilities of the two types of transitions are comparable. The ratio

$$n_s = \frac{\Gamma_{sp}}{\Gamma_{st}} , \tag{4.113}$$

is referred to as the *saturation photon number*.

The quantum state of the laser mode should be described by a (mixed state) density matrix $\rho_{nn'}$ rather than a (pure state) wave function. This is so since we cannot regard the electromagnetic field as a closed (isolated) system. It is a part of a larger system consisting of both atoms and field, but even that is not a closed system due to coupling of the atoms to the pumping field and due to the leak from the laser mode to the escaping laser beam and to the surroundings.

We will not approach the general problem of describing the time evolution of the density matrix $\rho_{nn'}$, but rather show that the steady state form of the photon probability distribution $p_n = \rho_{nn}$ can be found with some simplifying assumptions[5]. We first note that for the steady state there is a balance between transitions to and from atomic level 2,

$$N_0 R = N_2(\Gamma_{sp} + \Gamma_{st} n) , \tag{4.114}$$

where R is the transition rate, per atom, from the ground state to the excited state 2. We assume at this point that the photon number n in this equation is identical to the expectation value $\langle n \rangle$. This means that we neglect the (quantum and statistical) fluctuations in the photon number, which we regard as small compared to its mean value.

---

[5]The discussion here mainly follows the approach of Rodney Loudon, *The Quantum Theory of Light*, Oxford Science Publications, 2000.

We next consider the photon probability distribution p_n. If this distribution is stationary this means that for any photon number n there is balance between the processes that tend to increase the photon number and to decrease the photon number. We write this as a balance equation

$$N_2\Gamma_{st}(np_{(n-1)} - (n+1)p_n) - \Gamma_{cav}(np_n - (n+1)p_{(n+1)}) = 0\,, \qquad (4.115)$$

where we have introduced the *cavity loss rate per photon*, Γ_{cav}. With $|\mathcal{T}|^2$ as the transmission probability to the outside for each reflection and L as the length of the cavity, it is given by $\Gamma_{cav} = c|\mathcal{T}|^2/L$. By multiplying the equation with n and averaging over the probability distribution, we find the following expression for the photon expectation value[6],

$$N_2\Gamma_{st}(1 + \langle n \rangle) = \Gamma_{cav}\langle n \rangle\,. \qquad (4.116)$$

This equation relates the number N_2 of atoms in level 2 to the average photon number in the laser mode. We insert this in the steady state equation (4.114), with n replaced by $\langle n \rangle$, and we also replace N_0 by the total number of atoms N,

$$\Gamma_{st}\Gamma_{cav}\langle n \rangle^2 - (NR\Gamma_{st} - \Gamma_{sp}\Gamma_{cav})\langle n \rangle - NR\Gamma_{st} = 0\,. \qquad (4.117)$$

By introducing the coefficient

$$C = \frac{NR\Gamma_{st}}{\Gamma_{sp}\Gamma_{cav}}\,, \qquad (4.118)$$

the equation simplifies to

$$\langle n \rangle^2 - (C-1)n_s\langle n \rangle - Cn_s = 0\,, \qquad (4.119)$$

with solution

$$\langle n \rangle = \frac{1}{2}\left[(C-1)n_s + [(C-1)^2n_s^2 + 4Cn_s]^{\frac{1}{2}}\right]\,. \qquad (4.120)$$

The expectation value for the photon number is in Fig. 4.7 shown as a function of the parameter C for $n_s = 10^7$. As a characteristic feature one notes that around $C = 1$ the photon number rapidly increases from a small number to a number of the order of n_s. After this rapid increase there is a continued less dramatic increase where $\langle n \rangle$ changes linearly with C.

The curve demonstrates the presence of a threshold for the laser around $C = 1$. For smaller values the effects of spontaneous emission and emission from the cavity prevents the build up of the laser mode, while for larger values of C there is a net input of photons into the mode which allows the photon number to grow to a large number.

---

[6]Correlations between quantum fluctuations in N_2 and n are then assumed to be unimportant
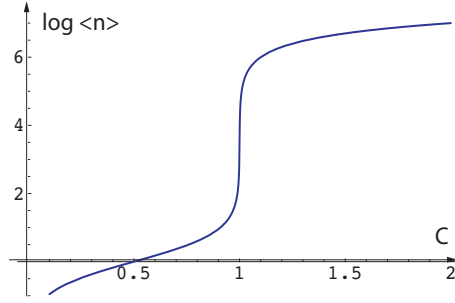
Figure 4.7: The photon number of the laser mode as a function of the parameter $C = NR/(n_s\Gamma_{cav})$. The curve corresponds to a saturation photon number $n_s = 10^7$.

4.4.2   Laser light and coherent photon states

So far we have considered the *expectation value* of the photon number, which can increase to a large value for $C > 1$. This implies that the mode is strongly excited, but does not necessarily mean that it is in a coherent state that can be described by a (classical) monochromatic wave. For this to be the case *fluctuations* of the field variables have to be restricted. The fluctuations in the photon number can be determined from the evolution equation of the probability distribution (4.115). This should be satisfied for all n, including $n = 0$, and with $p_{-1} = 0$ we conclude that following simpler equation has to be satisfied

$$N_2\Gamma_{st}p_{(n-1)} = \Gamma_{cav}p_n \,, \tag{4.121}$$

With the occupation number N_2 determined by (4.114), we get

$$
\begin{aligned}
p_n &= \frac{NR\Gamma_{st}}{\Gamma_{cav}(\Gamma_{sp} + \Gamma_{st}n)}p_{(n-1)} \\
&= \frac{Cn_s}{n_s + n}p_{(n-1)} \cdot
\end{aligned}
\tag{4.122}
$$

By repeated use of the equation we find

$$p_n = \frac{(Cn_s)^n n_s!}{(n_s + n)!}p_0 \,, \tag{4.123}$$

where p_0 is determined by the normalization of the probability distribution.

Well above the laser threshold, $C >> 1$, we have

$$\langle n \rangle = (C - 1)n_s \,, \tag{4.124}$$

as follows from Eq. (4.120). By use of this expression for C, the distribution can be re-written as

$$p_n = \frac{(n_s + \langle n \rangle)^n}{(n + n_s)!}\, n_s!p_0 \,. \tag{4.125}$$

We note that the above distribution has a form very similar to that of a coherent state,

$$p_n^{cs} = \frac{\langle n \rangle^n}{(n)!} \, e^{-\langle n \rangle} \,, \tag{4.126}$$

which has the form of a *Poisson distribution* over photon numbers n. The n-dependence is essentially the same in (4.125), except for the shift $n \to n + n_s$ in the case of the laser. For the coherent state the width of the distribution has the variance

$$(\Delta n)_{cs}^2 = \langle n \rangle \,, \tag{4.127}$$

whereas, due to the shift, the width of the probability distribution (4.125) is

$$(\Delta n)^2 = \langle n \rangle + n_s \,. \tag{4.128}$$

For large photon numbers, $\langle n \rangle \gg n_s$ the last term can be neglected and the variance is
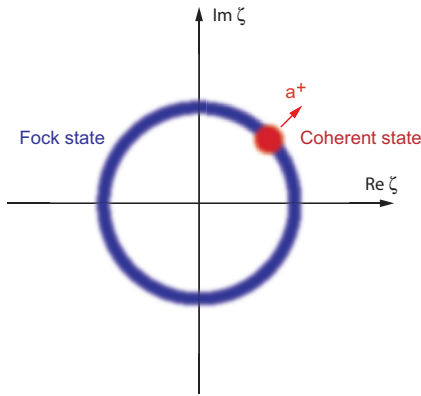

Figure 4.8: The complex phase plane of a single electromagnetic mode. A photon state is represented by the overlap with a *coherent state* $|z\rangle$. The radial coordinate correspond to the direction of increasing energy (or photon number), while the angular variable correspond to the phase of the electromagnetic wave. A Fock state (blue), with sharp photon number, is completely delocalized in the angular direction. A coherent state is well localized in both directions.

the same as expected for a coherent state. However, for smaller values n_s introduces a non-negligible contribution to the fluctuation in the photon number. This can be interpreted as the influence of spontaneous emission: It affects the fluctuations in the occupation number N_2 which in turn influences the fluctuations in the photon number of the laser mode through stimulated emission.

The expression we have found for the photon distribution of the laser gives an indication that laser light is to a good approximation described as coherent light. This is indeed the standard picture of light produced by a laser, and historically the introduction of *coherent states* of light was introduced in this context. However, the photon distribution in reality does not give information of the *phase* relations between the different photon components of the quantum state, which are essential in order to specify this as a coherent state. In Fig. 4.8 the

phase of the coherent state corresponds to the angular distribution of the state, while the photon number distribution corresponds to the radial distribution.

There has in fact quite recently been a discussion of this question, to what extent laser light is well described as a coherent state. This question is raised since the complex phase of the quantum state is not available for detection in any measurement, due to the rapid rotation of the phase angle. (This phase rotation defined by the angular frequency ω of the light.) However, the *relative* phase angle for different points along the laser beam is measurable, and the corresponding phase stability which characterize laser light is well established experimentally. In fact, this phase stability makes it possible to create interference between the light of two different lasers, as has been shown experimentally. The phase stability can be understood as a (phase) coherence of the quantum state, but with *coherence* in a more general sense than in the definition of a *coherent state*. We shall not go into further details about this question here.

4.5   Open quantum systems

An open quantum system is a system which is coupled to another, larger system, often referred to as the surroundings, with degrees of freedom that are not included in the description in a detailed way. Thus, the full set of degrees of freedom are separated in two groups, the set of internal degrees of freedom of the smaller system, which are treated explicitly, and the set of external degrees of freedom that are treated in an average way. Due to the coupling between the two systems the natural setting for the quantum description of the smaller system is to use the reduced density operator of this system. The time evolution of the density operator is assumed to depend only on a small set of coupling parameters to the external system, while the internal degrees of freedom are given a complete quantum treatment.

A general time evolution equation of the smaller system exists in the form of the *Lindblad equation*, which we will discuss below. The validity of the average treatment of the external degrees of freedom depends on their interaction with the internal degrees of freedom in a sense is acting only one way. Thus, there is essentially no delayed interaction between the two subsystems. In statistical physics, where the surroundings, in this way, has no memory of the earlier interactions with the smaller system, is called a *Markovian system*. In a previous section we have in fact met such a system in the form of a single atom in open space (the smaller system), which acts by spontaneous emission on the radiation field (the surroundings). In this case the system has the Markovian property, since the photons emitted by the atom will simply be absorbed by the radiation field, with no delayed reaction back on the atom. (For an atom in a small reflecting cavity, however, this may not be a good approximation, since an emitted photon there can be reflected back to react again with the atom.) Also the laser can be regarded as an open quantum system, with the internal degrees of freedom being those of the atoms in the cavity together with the laser mode. The external degrees of freedom will then include all other modes of the electromagnetic field and the degrees of freedom of the system that feeds energy to the laser.

The idea now is to start from the description of spontaneous emission in Sect. 4.3.3 where an imaginary energy contribution was given to the decaying atomic level to account for the effects of the emission process. The discussion was based on arguments of probability, and we shall see that by completing these arguments we will be able to reproduce the Lindblad

equation.

To include the imaginary contributions to the energy we define an effective atomic Hamiltonion of the form

$$\hat{H}_{eff} = \hat{H}_0 - \frac{i}{2} \sum_i \Gamma_i |i\rangle\langle i| = \sum_i (E_i - \frac{i}{2}\Gamma_i)|i\rangle\langle i|\,, \tag{4.129}$$

where $\hat{H}_0$ is the atomic Hamiltonian without decay, $|i\rangle$ are the corresponding eigenvectors, and E_i the eigenvalues. Note that the operator $\hat{H}_{eff}$ acts only on the atomic states, while the effect of the coupling to the radiation field is only through the decay parameters Γ_i. Theses parameters are nonnegative, Γ_i, and in principle all different from zero, exept for the ground state.

The corresponding Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi\rangle = \hat{H}_{eff}|\psi\rangle\,, \tag{4.130}$$

gives decay of the occupation probabilities for the states $|i\rangle$ when $\Gamma_i \neq 0$. However since $\hat{H}_{eff}$ is non-hermitian, the norm of $|\psi\rangle$ is not preserved, and this we take a sign that something is missing in the description of the time evolution of the atomic state.

A qualitative understanding of what is missing is not so difficult to get. Assume the photon emission is linked to an atomic transition $|i\rangle \rightarrow |j\rangle$. The Hamiltonian (4.129) takes correctly into account the reduction of the occupation probability of state $|i\rangle$ in this transition, but it does not include the corresponding increase in the occupation of state $|j\rangle$. To compensate for this we cannot simply add new terms to the Hamiltonian, we have to modify the time evolution equation of the density operator. Let us therefore first rewrite Eq.(4.130) as an equation for the atomic density operator $\hat{\rho} = |\psi\rangle\langle\psi|$,

$$\begin{aligned} i\hbar\frac{d}{dt}\hat{\rho} &=& \hat{H}_{eff}\hat{\rho} - \hat{\rho}\hat{H}_{eff}^\dagger \\ &=& \left[\hat{H}_0, \hat{\rho}\right] - \frac{i}{2}\sum_i \Gamma_i(|i\rangle\langle i|\hat{\rho} + \hat{\rho}|i\rangle\langle i|)\,. \end{aligned} \tag{4.131}$$

This gives

$$\frac{d}{dt}\operatorname{Tr}\hat{\rho} = -\frac{1}{\hbar}\sum \Gamma_i\langle i|\hat{\rho}|i\rangle < 0\,, \tag{4.132}$$

which shows the decrease in the norm of the state during the time evolution. The idea is now to add a term in the equation for the time evolution of $\hat{\rho}$ that compensates for this and makes $\operatorname{Tr}\hat{\rho}$ constant (equal to 1).

To proceed we first introduce what is often referred as a "jump operator", which maps the initial state $|i\rangle$ of the radiative transition to the final state $|j\rangle$. The operator and its hermitian conjugate are

$$\hat{\alpha}_{ji} = |j\rangle\langle i|\,, \quad \hat{\alpha}_{ji}^\dagger = |i\rangle\langle j|\,. \tag{4.133}$$

We furter write Γ_i as a sum over contributions from all transitions with $|i\rangle$ as initial state

$$\Gamma_i = \sum_j \gamma_{ji}\,, \quad \gamma_{ji} = \hbar w_{ji}\,, \tag{4.134}$$

where w_{ji} is the transition rate for the transition $|i\rangle \to |j\rangle$. This relation between Γ_i and w_{ji} is consistent with what has been established in Sect. 4.3.3. From this follows

$$
\begin{aligned}
\sum_i \Gamma_i |i\rangle\langle i| &= \sum_{ij} \gamma_{ji} |i\rangle\langle i| \\
&= \sum_{ij} \gamma_{ji} |i\rangle\langle j|j\rangle\langle i| \\
&= \sum_{ij} \gamma_{ji}\, \hat{\alpha}_{ji}^\dagger \hat{\alpha}_{ji}\,.
\end{aligned}
\tag{4.135}
$$

Note that we may in these expressions have unrestricted summation over i and j, provided we set $\gamma_{ji} = 0$ for all cases where $|i\rangle \to |j\rangle$ does not correspond to a radiative transition. In particular $E_i > E_j$ is a necessary condition for $\gamma_{ji} \neq 0$.

The two equations (4.131) and (4.132) can then be written as

$$
\begin{aligned}
i\hbar \frac{d}{dt}\hat{\rho} &= \hat{H}_{eff}\hat{\rho} - \hat{\rho}\hat{H}_{eff}^\dagger \\
&= \left[\hat{H}_0, \hat{\rho}\right] - \frac{i\hbar}{2}\sum_{ij}\gamma_{ji}(\,\hat{\alpha}_{ji}^\dagger\hat{\alpha}_{ji}\hat{\rho} + \hat{\rho}\alpha_{ji}^\dagger\hat{\alpha}_{ji})\,,
\end{aligned}
\tag{4.136}
$$

and

$$
i\hbar \frac{d}{dt}\operatorname{Tr}\hat{\rho} = -i\sum_{ij}\gamma_{ji}\operatorname{Tr}(\hat{\alpha}_{ji}\,\hat{\rho}\,\hat{\alpha}_{ji}^\dagger)\,.
\tag{4.137}
$$

The last equation suggests a way to prevent decay of the norm. Since we have the identity

$$
Tr(\hat{\alpha}_{ji}^\dagger\hat{\alpha}_{ji}\hat{\rho} + \hat{\rho}\alpha_{ji}^\dagger\hat{\alpha}_{ji} - 2\hat{\alpha}_{ji}\hat{\rho}\hat{\alpha}_{ji}^\dagger) = 0\,,
\tag{4.138}
$$

we may correct the expression for the time derivative of the density operator in the following way,

$$
i\hbar\frac{d}{dt}\hat{\rho} = \left[\hat{H}_0, \hat{\rho}\right] - \frac{i}{2}\sum_{ij}\gamma_{ji}(\,\hat{\alpha}_{ji}^\dagger\hat{\alpha}_{ji}\hat{\rho} + \hat{\rho}\alpha_{ji}^\dagger\hat{\alpha}_{ji} - 2\hat{\alpha}_{ji}\hat{\rho}\hat{\alpha}_{ji}^\dagger)\,.
\tag{4.139}
$$

The laddition of the last term, which we may write as

$$
\hat{\alpha}_{ji}\hat{\rho}\hat{\alpha}_{ji}^\dagger = |j\rangle\langle i|\hat{\rho}|i\rangle\langle j|\,,
\tag{4.140}
$$

increases the occupation probability of state $|j\rangle$, and compensates presicely for the lost occupation probability of state $|i\rangle$, in the transition $|i\rangle \to |j\rangle$.

Eq. (4.139) is what we have referred to as the Lindblad equation. It has been reached in a stepwise way, by use of probability arguments, and has been applied to a concrete situation with decay of exitation probabilities for an atom interacting with the radiation field. However, the equation has much wider range of validity. It can be applied to open quantum systems more generally, under the condition of Markovian behavior of the system. The equation does not only describe the decay of separate excited atomic levels, as previously discussed, but also accounts for the effects of decay for more general atomic states, corresponding to superposition between the energy eigenstates.

4.5.1 Application to a two-level atom

Consider a two-level model of an atom with ground state $|0\rangle$ and excited state $|1\rangle$. The Hamiltonian, when the coupling to the radiation field is neglected, we write as

$$\hat{H}_0 = \frac{1}{2}\Delta\sigma_z\,, \tag{4.141}$$

where σ_z is diagonal in the energy eigenstate basis, with eigenvalues -1 for the ground state and $+1$ for the excited state. Coupling to the radiation field introduces a transition from the excited state to the ground state, with a decay parameter we denote by γ. The time evolution of the density operator is governed by the Lindblad equation, here on the form

$$i\hbar\frac{d}{dt}\hat{\rho} = \left[\hat{H}_0, \hat{\rho}\right] - \frac{i\hbar}{2}\gamma(\hat{\alpha}^\dagger\hat{\alpha}\hat{\rho} + \hat{\rho}\hat{\alpha}^\dagger\hat{\alpha} - 2\hat{\alpha}\hat{\rho}\hat{\alpha}^\dagger)\,, \tag{4.142}$$

where the jump operator is $\alpha = |0\rangle\langle 1|$.

Let us write the density matrix, in the energy eigenstate basis, as

$$\rho = \begin{pmatrix} p_1 & d \\ d^* & p_0 \end{pmatrix}\,, \tag{4.143}$$

with p_1 as the occupation probability of the upper level, p_0 of the ground state, and d as the matrix element between the upper and lower level. We note that the term in (4.142) which contains $\hat{H}_0$ only gives contributions to the off-diagonal terms, whereas the last term, proportional tot $\alpha\hat{\rho}\alpha^\dagger$ only gives contributions to the diagonal terms of the density matrix. We find, for the different matrix elements

$$\begin{aligned}
\frac{d}{dt}p_1 &= -\frac{\gamma}{\hbar}p_1 \\
\frac{d}{dt}p_0 &= \frac{\gamma}{\hbar}p_1 \\
\frac{d}{dt}d &= -(\frac{\gamma}{2\hbar} + \frac{i}{\hbar}\Delta)\,d\,.
\end{aligned} \tag{4.144}$$

The equations give rise, as expected, to exponential decay of the occupation probability p_1 of the excited level, and a corresponding increase in the occupation probability p_0 of the ground state. This is consistent with the total probability being conserved, $p_0 + p_1 = 1$. The off-diagonal matrix element d also decays, but with a time rate that is half of that for p_1. There is in addition a phase rotation of d caused by the energy difference between the two levels. Expressed in terms of initial values $p_1(0)$ and $d(0)$, for $t = 0$, Eqs. (4.144) give the following expressions for the time dependent parameters,

$$\begin{aligned}
p_1(t) &= p_1(0)\,e^{-\frac{\gamma}{\hbar}t}\,, \\
p_0(t) &= 1 - p_1(0)e^{-\frac{\gamma}{\hbar}t}\,, \\
d(t) &= d(0)e^{-(\frac{\gamma}{2\hbar}-i\frac{\Delta}{\hbar})t}\,.
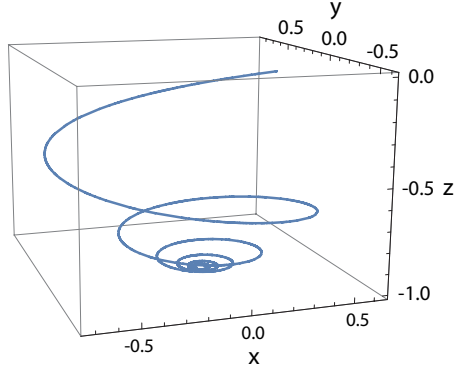\end{aligned} \tag{4.145}$$

The density matrix may also be parametrized as

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\cdot\boldsymbol{\sigma}) = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}\,. \tag{4.146}$$

Expressed in these variables the general time evolution is

$$
\begin{aligned}
x(t) &= \left(x(0)\cos(\frac{\Delta}{\hbar}t) - y(0)\sin(\frac{\Delta}{\hbar}t) \right) e^{-\frac{\gamma}{2\hbar}t}, \\
y(t) &= \left(x(0)\sin(\frac{\Delta}{\hbar}t) + y(0)\cos(\frac{\Delta}{\hbar}t) \right) e^{-\frac{\gamma}{2\hbar}t}, \\
z(t) &= (z(0)+1)e^{-\frac{\gamma}{\hbar}t} - 1.
\end{aligned}
\tag{4.147}
$$

If the initial state is a *statistical* mixture of the exited state and the ground state, we have $x = y = 0$ under the time evolution, and z approaches exponentially the ground state value -1. If $x \neq 0$ and/or $y \neq 0$ this means that the state includes superpositions between the two state. Also in this case the state will approach exponentially the ground state, but with a different decay rate for the horizontal coordinates x, y and the vertical coorinate z.

In the figure the initial values are $x = 1, y = z = 0$, which gives

$$
\begin{aligned}
x(t) &= \cos(\frac{\Delta}{\hbar}t)e^{-\frac{\gamma}{2\hbar}t}, \\
y(t) &= \sin(\frac{\Delta}{\hbar}t)e^{-\frac{\gamma}{2\hbar}t}, \\
z(t) &= e^{-\frac{\gamma}{\hbar}t} - 1.
\end{aligned}
\tag{4.148}
$$

The motion is a combination of rotation in the x, y-plane and exponential decay of the variables towards the ground state value.

Chapter 5

Quantum mechanics and geometry

Physics and geometry are closely linked in many different ways. In its original form Euclidean geometry describes the mathematical structure of physical space and the shapes of physical bodies. In the theory of relativity this geometry of space is extended to a geometry of space-time. Although general relativity represents in many ways the clearest realization of geometry in physics, there are also many other examples, both in classical and quantum physics, of how a geometrical formulation is both possible and important in order to understand the structure of the theory.

In quantum theory geometry is usually not emphasized to the same degree as it is in the theory of relativity. But there are some parts of quantum theory where the underlying geometry is quite clear. One example has to do with gauge invariance of electromagnetism and the generalization of electromagnetic theory to gauge theories. Another example has to do with geometrical phases that appear when a quantum state is evolving in a periodic fashion. However, beyond these particular cases a geometrical view of quantum theory at a more fundamental level is possible, and an interesting question to consider is whether the geometrical view of quantum physics may be of importance beyond the applications studied today.

In this chapter the intention is not to make an extensive study of the geometry of quantum mechanics, but rather to examine some basic elements in such a geometrical formulation and to discuss some simple examples where the geometry is clearly revealed. In particular I will focus attention on the geometrical phases, but I will also discuss the general aspects of quantum geometry that have to do with the fact that the natural geometry of quantum states is complex rather than real.

5.1   Geometry of quantum states

The state space of a quantum system is a vector space, and in the same way as for three-dimensional space, or four-dimensional space-time, the natural geometry of this vector space is defined by the scalar product. In particular the there is a notion of length, or *metric*, defined by

$$\Delta s^2 \equiv \langle \Delta \psi | \Delta \psi \rangle \,, \tag{5.1}$$

where $\Delta\psi$ is the relative vector of two "points" in the Hilbert space, defined by two vectors $|\psi\rangle$ and $|\psi + \Delta\psi\rangle$. Clearly there are some important differences between the geometry of physical (Euclidean) space and the Hilbert space of a quantum system. One difference is that the Hilbert space is complex rather than real, and therefore the geometry defined by the scalar product is also complex. Another difference has to do with the interpretation of *distance*. In three-dimensional space distance is a basic variable. It defines the separation and relative motion of physical bodies. In quantum theory the scalar product instead is linked to the probability interpretation of the theory, since the squared scalar product of two vectors $|\langle\phi|\psi\rangle|^2$ can be viewed as the probability of finding the system in state ϕ when it is prepared in state ψ (or the other way around). This implies that the distance defined by (5.1) does not have the dimension of *physical length*, but is instead a dimensionless quantity.

However, the metric defined by the scalar product of a Hilbert space is not precisely the metric we will identify as the metric of the *quantum states*. That is because these states are not in a one-to-one correspondence to Hilbert space vectors. Thus, if a vector is is multiplied by a complex number z,

$$|\psi\rangle \rightarrow |\psi'\rangle = z|\psi\rangle \,, \tag{5.2}$$

this does not correspond to a change of the quantum state[1]. The ambiguity in representing quantum states in terms of state vectors is reduced when normalization is introduced, $\langle\psi|\psi\rangle = 1$, but still the possibility of multiplying the vector by a complex phase factor remains. Since such a transformation does not represent a physical change, it should not affect the distance between two states. The definition of distance (5.1) does not satisfy this requirement, but it can be corrected by a simple modification. For an infinitesimal change we write the (corrected) metric, for normalized vectors, as

$$ds^2 \equiv \langle d\psi|d\psi\rangle - \langle d\psi|\psi\rangle\langle\psi|d\psi\rangle \,, \tag{5.3}$$

where the correction term can be interpreted as a subtraction of the contribution from changing the phase of the vector. Invariance of the corrected metric under multiplication with a (state dependent) phase factor can readily be demonstrated.

The reason that the distance in (5.3) is defined for an infinitesimal, rather than for a finite change of the vector, is due to the fact that the corrected metric is not a flat (Euclidean) metric. In that respect it is different from the original Hilbert space metric (5.1). We shall return to this point.

To express the metric in coordinate form, we introduce a complete set of Hilbert space vectors and expand a general vector in this basis,

$$|\psi\rangle = \sum_i \psi^i |\chi_i\rangle \,. \tag{5.4}$$

The metric then can be written as

$$ds^2 = \sum_{ij} \gamma_{ij} \, d\psi^{*i} d\psi^j \,, \tag{5.5}$$

---

[1]A quantum state may be viewed as corresponding to a *ray* of Hilbert space vectors. A ray is defined as an *equivalence class* of vectors, where vectors that differ by a complex number belong to the same class, $|\psi\rangle \simeq z|\psi\rangle$. The space of rays is not a vector space, but is referred to as a *projective space*. Since the rays can be associated with *directions* in the Hilbert space, the topology of this space is similar to that of a sphere.

with

$$\gamma_{ij} = \langle \chi_i | \chi_j \rangle - \langle \chi_i | \psi \rangle \langle \psi | \chi_j \rangle \, , \tag{5.6}$$

where $d\psi^i$ represent the coordinates of the change $|d\psi\rangle$ of the Hilbert space vector. The matrix γ_{ij} is the *metric tensor*, which here is complex and hermitian. The expression (5.5) should be compared with the more conventional expression for the metric of a real, curved space,

$$ds^2 = \sum_{ij} g_{ij}(x) dx^i dx^j \, , \tag{5.7}$$

where the metric tensor g_{ij} is a real and symmetric matrix. The tensor γ_{ij} depends on the vector $|\psi\rangle$, and this dependence, which is similar to the x-dependence of g_{ij}, reflects the fact that it is the metric of a curved space.

So far the discussion has been restricted to *pure* quantum states. However, an extension to the full set of pure and mixed states can readily be done. In order to do so, we first re-write the metric (5.5) of pure states in terms of density operators. The density operator of a pure state has the form $\hat{\rho} = |\psi\rangle\langle\psi|$ (assuming a normalized vector), and for an infinitesimal change that gives

$$d\hat{\rho} = |d\psi\rangle\langle\psi| + |\psi\rangle\langle d\psi| \, . \tag{5.8}$$

By use of this expression it is straight forward to check that the metric (5.3) can be written as

$$ds^2 = \frac{1}{2} \operatorname{Tr}(d\hat{\rho}^2) \, , \tag{5.9}$$

and this expression, which is well-defined also for mixed states, extends the notion of distance, in a natural way, to the set of all quantum states, both pure and mixed.

It is interesting to note that right-hand-side of (5.9) is proportional to the norm of $d\hat{\rho}$ with respect to the natural scalar product of hermitian matrices

$$\langle A | B \rangle = \operatorname{Tr}(AB) \, . \tag{5.10}$$

These matrices define a real vector space, which is a flat, Euclidean space with respect to the vector norm. In this vector space the density operators form a continuous, flat subset, restricted by the positivity condition $\hat{\rho} \geq 0$ and the norm condition $\operatorname{Tr} \hat{\rho} = 1$. The situation is illustrated in a schematic way in Fig. 5.1, which shows the subset of density matrices in the vector space of hermitian matrices. The positive matrices are shown as the interior of the *positive cone* of matrices, and the density matrices is the subset of these in the (flat) hyperplane[2] selected by the trace condition. If the Hilbert space dimension of the quantum system is n, the dimension of the corresponding vector space of hermitian matrices is n^2 (since a hermitian $n \times n$ matrix is specified by n^2 real parameters), and the dimension of the set of density matrices is $n^2 - 1$ due to the constraint $\operatorname{Tr} \hat{\rho} = 1$. The pure states form a continuous set of dimension $2n - 2$, which are boundary points of the set of density matrices. (A general complex vector is specified by $2n$ real parameters, and a physical pure state satisfy two constraints, the norm condition and a condition that specifies the complex phase of the vector. This reduces the number of independent parameters to $2n - 2$.) As a specific example we shall next examine the geometry of quantum states for a two-level system, in which case the geometry can be illustrated in a simple way.

---

[2]A *hyperplane* refers to a flat subspace not restricted to two dimensions.
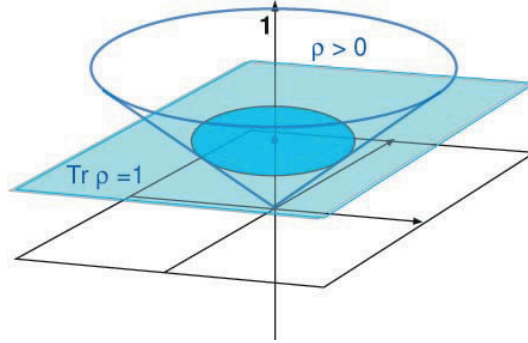
Figure 5.1: A schematic representation of the set of density matrices in the vector space of hermitian matrices. The positive matrices form a cone about the axis defined by the unit matrix, and the normalization condition restricts the density matrices to a convex subset, here represented by the shaded circle.

5.1.1 Example: Geometry of the two-level system

As previously discussed, a general hermitian 2x2-matrix has the following representation in terms of Pauli matrices

$$A = \frac{1}{2}(a\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma})\,, \tag{5.11}$$

where hermiticity means that a as well as the three components of $\mathbf{r}$ are real. Thus, the 2x2 hermitian matrices can be viewed as vectors in a four-dimensional real vector space, spanned by the four matrices $\mathbb{1}$, σ_x, σ_y and σ_z. These matrices satisfy the orthogonality relations

$$\frac{1}{2}\,\mathrm{Tr}(\sigma_\mu\sigma_\nu) = \delta_{\mu\nu}\,, \quad \mu, \nu = 1, 2, 3, 4\,, \tag{5.12}$$

where we for simplicity have defined $\sigma_4 = \mathbb{1}$. The eigenvalues of the matrix A are $(a \pm r)/2$, and positivity is therefore expressed by the following condition,

$$0 \le r = \sqrt{x^2 + y^2 + z^2} \le a\,. \tag{5.13}$$

Geometrically this condition means that the positive matrices are represented by vectors that are located within a cone, which is centered around the axis defined by the identity matrix $\mathbb{1}$ (see Fig. 5.1 for a schematic illustration).

The trace condition, satisfied by the density matrices, fixes one of the coordinates, $a = 1$ and reduces thereby four-dimensional space to a three-dimensional hyperplane in this space. The coordinates of this space are now the three components of $\mathbf{r}$, and the set of positive matrices defines the Bloch sphere, with $r \le 1$, as we have previously discussed. The center of the sphere corresponds to the point $(0, 0, 0, \frac{1}{2})$ in the four dimensional representation, and it represents the maximally mixed state $\hat{\rho} = \frac{1}{2}\mathbb{1}$.

To examine the metric in the space of density matrices, we write a general density matrix in the standard way as

$$\hat{\rho} = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma})\,, \tag{5.14}$$

and consider an infinitesimal displacement

$$d\hat{\rho} = \frac{1}{2} d\mathbf{r} \cdot \boldsymbol{\sigma} \, . \tag{5.15}$$

The corresponding infinitesimal distance is given by,

$$
\begin{aligned}
ds^2 &= \frac{1}{2} \operatorname{Tr}(d\hat{\rho}^2) \\
&= \frac{1}{8} \sum_{ij} \operatorname{Tr}(\sigma_i \sigma_j) \\
&= \frac{1}{4} d\mathbf{r}^2 \, . \tag{5.16}
\end{aligned}
$$

This shows that the metric of the quantum states that we have previously defined, in the three-dimensional space of density matrices of the two-level system, is proportional to the standard Euclidean metric of three-dimensional space.

Written in polar coordinates the expression for the distance is

$$ds^2 = \frac{1}{4}(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2) \, , \tag{5.17}$$

and therefore, if we restrict the density matrices to pure quantum states, the metric is

$$ds^2 = \frac{1}{4}(d\theta^2 + \sin^2\theta \, d\phi^2) \, . \tag{5.18}$$

This is, up to a constant factor, identical to the standard metric of the surface of a unit sphere. This also shows what has been stated as a general situation, that the pure states form a continuous, curved surface at the boundary of the set of density matrices, while the full set, with the mixid states in the interior included, is a (compact) subset of a flat, Euclidean space.

5.1.2 Geometrical structures in parameter space

In the following we again restrict the discussion to the geometry of *pure* quantum states. For the two level system they can be identified with the surface of the sphere, and a specific state can be identified by the values of the two real parameters θ and ϕ. We will now consider a similar description for more general sets of quantum states.

Let us therefore consider a set of n real coordinates $x = \{x^1, x^2, ..., x^n\}$, which label a continuous set of quantum states. In the two-level case there are two parameters θ and ϕ and for a three-level system four parameters are needed to label the full set of quantum states. However, in the following we do not specify precisely this set, it may be the complete set of all states of a quantum system, or a subset of quantum states that are relevant for the dynamics of the system. We shall later consider a situation where the relevant states are low-energy states of the system.

Each set of coordinates x we consider as specifying a point in *parameter space* and to each such point we associate a unique quantum state described by a state vector $|\chi_x\rangle$. As already discussed there is not a unique mapping between physical states and state vectors, but we assume that, for all x, the vector $|\chi_x\rangle$ is normalized and that the complex phase is specified

by making a choice between equivalent vectors. In this way we have created a one-to-one mapping between the continuous set of quantum states and the parameter space. We consider first how the metric structure on the set of (pure) quantum states can be described as a structure on the parameter space.

Since a variation of the state vector $|\chi_x\rangle$ can be associated with a change in the parameters x, we write this change as

$$|d\chi_x\rangle = \sum_{k=1}^{n} |\partial_k\chi_x\rangle dx^k \,, \qquad |\partial_k\chi_x\rangle \equiv \frac{\partial}{\partial x^k}|\chi_x\rangle \,. \tag{5.19}$$

The metric (5.3) can then be written in the form

$$
\begin{aligned}
ds^2 &= \sum_{kl} [\langle\partial_k\chi_x|\partial_l\chi_x\rangle - \langle\partial_k\chi_x|\chi_x\rangle\langle\chi_x|\partial_l\chi_x\rangle]\, dx^k dx^l \\
&= \sum_{kl} g_{kl}(x)\, dx^k dx^l \,.
\end{aligned}
\tag{5.20}
$$

Expressed in this way the metric tensor g_{kl} is real and can be assumed to be symmetric. It is given by

$$
\begin{aligned}
g_{kl} &= \frac{1}{2}[\langle\partial_k\chi_x|\partial_l\chi_x\rangle - \langle\partial_k\chi_x|\chi_x\rangle\langle\chi_x|\partial_l\chi_x\rangle + \langle\partial_l\chi_x|\partial_k\chi_x\rangle - \langle\partial_l\chi_x|\chi_x\rangle\langle\chi_x|\partial_k\chi_x\rangle] \\
&= \mathrm{Re}[\langle\partial_k\chi_x|\partial_l\chi_x\rangle - \langle\partial_k\chi_x|\chi_x\rangle\langle\chi_x|\partial_l\chi_x\rangle] \,.
\end{aligned}
\tag{5.21}
$$

We now introduce the *projected derivative* in the following way

$$|D_k\chi_x\rangle \equiv |\partial_k\chi_x\rangle - |\chi_x\rangle\langle\chi_x|\partial_k\chi_x\rangle \,. \tag{5.22}$$

The second term is introduced, like in the definition of the complex metric tensor γ_{ij}, to remove the contribution that is "unphysical" in the sense that it only changes the phase of the vector. For the scalar product between two such state vectors we find

$$\langle D_k\chi_x|D_l\chi_x\rangle = \langle\partial_k\chi_x|\partial_l\chi_x\rangle - \langle\partial_k\chi_x|\chi_x\rangle\langle\chi_x|\partial_l\chi_x\rangle \,, \tag{5.23}$$

and the metric tensor can therefore be expressed as

$$g_{kl} = \mathrm{Re}[\langle D_k\chi_x|D_l\chi_x\rangle] \,. \tag{5.24}$$

To examine further the meaning of the projected derivative we introduce the following vector field

$$a_k(x) = -i\langle\chi_x|\partial_k\chi_x\rangle \,, \tag{5.25}$$

which is a real-valued field, due to the normalization condition,

$$\langle\chi_x|\chi_x\rangle \quad \Rightarrow \quad \langle\partial_k\chi_x|\chi_x\rangle = -\langle\chi_x|\partial_k\chi_x\rangle \,. \tag{5.26}$$

The projected derivative then can be re-written as

$$|D_k\chi_x\rangle = |\partial_k\chi_x\rangle - ia_k|\chi_x\rangle \,, \tag{5.27}$$

or simply

$$D_k = \partial_k - ia_k \,. \tag{5.28}$$

This form for the derivative gives a clear resemblance to *minimal coupling* in electromagnetism, where the interaction of an electron with the electromagnetic field can be introduced in the Schrödinger equation by a simple substitution $\partial_k \to D_k \equiv \partial_k - i\frac{e\hbar}{c}A_k$, with A_k as the electromagnetic potential. The differentiation D_k is there called a *covariant derivative*; it transforms in a covariant way under a combined gauge transformation of the electromagnetic potential and of the quantum mechanical wave function. This correspondence to electromagnetism is no accident, and we shall refer to the field a_k as the *geometrical potential*. Also here D_k can be interpreted as a covariant derivative, and to show this we examine what gauge invariance means in the present context.

As already discussed there is an intrinsic ambiguity in associating a unique state vector $|\chi_x\rangle$ to each quantum state, labeled by the set of parameters x. Thus, a choice has to be made for the complex phase of each state vector, and a change in this choice, corresponding to multiplication with an x-dependent phase factor, gives rise to an equivalent set of state vectors,

$$|\chi_x\rangle \to |\chi'_x\rangle = e^{i\phi(x)}|\chi_x\rangle \,. \tag{5.29}$$

This transformation introduces a change of the geometric potential

$$a_k \to a'_k = -i\langle\chi'_x|\partial_k\chi'_x\rangle = a_k + \partial_k\phi \,, \tag{5.30}$$

which is like the gauge transformation of the electromagnetic potential. The corresponding transformation of the projected derivative is

$$D_k \to D'_k \equiv \partial_k - ia'_k = e^{i\phi(x)}D_k e^{-i\phi(x)} \,, \tag{5.31}$$

and from this follows that the derivative of a state vector transforms in the same way as the state vector itself,

$$|D_k\chi_x\rangle \to |D'_k\chi'_x\rangle = e^{i\phi(x)}|D_k\chi_x\rangle \,. \tag{5.32}$$

This justifies the name of a *covariant* derivative.

It is clear from the expression (5.24) that the metric tensor is gauge invariant, in the sense that it does not change under the phase transformation (5.29). However, this expression for g_{kl} also indicates that there exists another related, gauge-invariant field, which we write as

$$f_{kl} = 2\,\mathrm{Im}\langle D_k\chi_x|D_l\chi_x\rangle \,, \tag{5.33}$$

(the numerical factor 2 is simply introduced for convenience). This field has also a geometric interpretation, and we shall soon discuss that further. Note that f_{kl} is antisymmetric in the indices k and l, whereas g_{kl} is symmetric. Also note that when expressed in terms of ordinary derivatives, g_{kl} has an additional term which depends on the potential a_k, whereas f_{kl} has no such term

$$\begin{aligned} g_{kl} &= \mathrm{Re}\langle\partial_k\chi_x|\partial_l\chi_x\rangle - a_k a_l \,, \\ f_{kl} &= 2\,\mathrm{Im}\langle\partial_k\chi_x|\partial_l\chi_x\rangle \,. \end{aligned} \tag{5.34}$$

However, as follows directly from the definition of a_k, another useful expression for f_{kl} exists,

$$f_{ij} = \partial_i a_j - \partial_j a_i \,, \tag{5.35}$$

and this shows that f_{kl} is analogous to the electromagnetic field tensor. In the following we shall show that this field gives rise to a *geometrical phase factor* in the time evolution of the state vector of the quantum system.

5.2   Periodic motion and the geometric phase

We consider the situation of periodic motion, where the quantum state after a certain time T returns to the inital state. Expressed in terms of the time dependent state vector $|\psi(t)\rangle$, the vectors at time $t = 0$ and time $t = T$ will be equal up to a phase factor, $|\psi(T)\rangle = \exp(-i\alpha(T))|\psi(0)\rangle$. As we shall show the phase $\alpha(T)$ consists of two contributions, a dynamical term which depends on the expectation value of the energy along the path, and another one which we identify as the geometric phase associated with the closed path in the space of quantum states.

   We assume that the relevant state vectors are parametrized in the way previously discussed, so that x represents a set of real parameters which identifies a quantum state, with $|\chi_x\rangle$ as the corresponding, normalized state vector. Expressed in terms of these vectors, the time dependent state vector takes the form

$$|\psi(t)\rangle = e^{-i\alpha(t)}|\chi_{x(t)}\rangle \,. \tag{5.36}$$

This gives

$$\langle\psi(t)|\frac{\partial}{\partial t}|\psi(t)\rangle = -i\dot{\alpha} + \sum_k \dot{x}_k\langle\chi_{x(t)}|\partial_k\chi_{x(t)}\rangle \,. \tag{5.37}$$

By application of Schrödinger's equation, we find the following expression for the expectation value of the energy

$$\begin{aligned} E(t) &\equiv \langle\psi(t)|\hat{H}|\psi(t)\rangle \\ &= \hbar(\dot{\alpha} + i\dot{x}_k\langle\chi_{x(t)}|\partial_k\chi_{x(t)}\rangle) \,. \end{aligned} \tag{5.38}$$

Integrating this over a closed path C in x space we get

$$\alpha(T) = \frac{1}{\hbar}\int_0^T dt\, E(t) - i\oint_C dx_k\langle\chi_x|\partial_k\chi_x\rangle \,. \tag{5.39}$$

Note that this result is valid whether the Hamiltonian is time dependent or not. The main point is that the quantum state performs periodic motion, and returns to the original state at time T. The first contribution, which is the dynamical one, depends on the integrated energy $E(t)$. If the Hamiltonian is time independent, and the expectation value of the energy also is so, the dynamical phase factor gets the recognizable form $\exp(-\frac{i}{\hbar}ET)$. The second term is the geometric phase. It only depends on the path C described by the state vector and not on the energy of the state or the time spent on traversing the closed path. The geometric phase,

often referred to as the Berry phase, can be expressed in terms of the geometric fields a_i and f_{ij} earlier introduced as

$$\alpha_B = \oint_C dx_k a_k(x) = \frac{1}{2} \int_S f_{ij} d\sigma_{ij} , \qquad (5.40)$$

where S is a surface in parameter space with C as boundary, and $d\sigma_{ij}$ is the infinitesimal surface element on S. We note that for periodic motion the phase (5.40) is gauge independent, in the sense that it does not depend on the phases chosen for the x-dependent basis vectors $|\chi_x^0\rangle$. However, for an open path that is not the case, since the geometrical phase will depend on the relative phases of the two basis vectors associated with the end points of the path.

5.2.1 Example: Spin motion in a magnetic field

We consider first the geometric phases for closed paths in space of quantum states of a spin half system. As previously discussed, the full set of states can be viewed as the surface of a unit sphere, with the unit vector $\mathbf{n}$ as variable. Let us parametrize the corresponding state vectors by the polar angles (ϕ, θ) of the unit vector $\mathbf{n}$. These angles thus correspond to the parameters x_k in the previous section. From earlier discussion of the spin states we have the following complete set of vectors parametrized by the angles angles (ϕ, θ),

$$\chi_{\mathbf{n}} = \chi_{\phi, \theta} = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix} . \qquad (5.41)$$

From this we derive the two components of the geometric potential

$$a_\theta = -i \langle \chi_{\mathbf{n}} | \partial_\theta \chi_{\mathbf{n}} \rangle = 0 , \qquad a_\phi = -i \langle \chi_{\mathbf{n}} | \partial_\phi \chi_{\mathbf{n}} \rangle = \frac{1}{2}(1 - \cos \theta) . \qquad (5.42)$$

The corresponding antisymmetric field f_{kl} has one (independent) non-vanishing component,

$$f_{\theta\phi} = -f_{\phi\theta} = \frac{1}{2} \sin \theta . \qquad (5.43)$$

For the geometric phase associated with a closed path C on the unit sphere this gives

$$\alpha_B = \int_C a_\phi d\phi = \int_S f_{\theta\phi} d\theta d\phi = \frac{1}{2} \int_S \sin \theta d\theta d\phi , \qquad (5.44)$$

where S is a surface with C as boundary. Thus, the geometric phase is identical to half the solid angle defined by the closed loop on the sphere. One should note the ambiguity in the definition of S, as the loop defines *two* surfaces on the sphere, with solid angles adding to 4π. However, this only means that the geometric phase is defined modulo 2π.

We next introduce dynamics by considering the time evolution of the spin in a constant magnetic field, which points in the z-direction. The spin Hamiltonian is

$$H = -\frac{e\hbar}{2m} \mathbf{B} \cdot \boldsymbol{\sigma} = \frac{1}{2} \hbar \omega_B \, \sigma_z . \qquad (5.45)$$

and the the time evolution, for a general initial state, then is

$$\psi(t) = e^{-\frac{i}{2} \omega_B t \sigma_z} \psi(0) . \qquad (5.46)$$

The motion is periodic, with the same period $T = 2\pi/\omega_B$ for any initial state. With $\psi(T) = -\psi(0)$ the phase change during one period is $\alpha(T) = \pi \pmod{2\pi}$.

We will now check that this simple result is consistent with the general result given in Eqs. (5.39) and (5.40). To this end we specify the initial state as

$$\psi(0) = \begin{pmatrix} \cos\frac{\theta_0}{2} \\ \sin\frac{\theta_0}{2} \end{pmatrix} , \tag{5.47}$$

which means $\phi = 0$ and θ_0 unspecified. The expectation value of the energy then is

$$E = \frac{1}{2}\hbar\omega_B(\cos^2\frac{\theta_0}{2} - \sin^2\frac{\theta_0}{2}) = \frac{1}{2}\hbar\omega_B \cos\theta_0 , \tag{5.48}$$

which gives the time integral

$$\frac{i}{\hbar}\int_0^T E dt = \pi\cos\theta_0 . \tag{5.49}$$

The geometric phase associated with the closed path is

$$\alpha_B = \frac{1}{2}\int_0^{2\pi} d\phi \int_0^{\theta_0} \sin\theta = \pi(1 - \cos\theta_0) . \tag{5.50}$$

The sum is thus

$$\alpha(T) = \frac{i}{\hbar}\int_0^T E dt + \alpha_B = \pi , \tag{5.51}$$

which is consistent with the result derived above in a more direct way.

We consider next the case where the magnetic field contains a rotating component

$$\mathbf{B} = B_0\,\mathbf{k} + B_1(\cos\omega t\,\mathbf{i} + \sin\omega t\mathbf{j}) , \tag{5.52}$$

which gives the time dependent hamiltonian

$$\hat{H}(t) = \frac{1}{2}\hbar[\omega_0\,\sigma_z + \omega_1(\cos\omega t\,\sigma_x + \sin\omega t\,\sigma_y)] , \tag{5.53}$$

where $\omega_0 = -eB_0/m$ and $\omega_1 = -eB_1/m$. The spin motion in this field was found in Sect.1.4.2, by transforming to a rotating frame where the Hamiltonian was time independent. In this frame the spin motion is a precession around the effective magnetic field, and consequently, in the non-rotating frame the motion is a combination of this precession with the rotation of the frame. The situation is illustrated in Fig. 5.2.

The Hamiltonian in the rotating frame has the form

$$\hat{H}_R = \frac{1}{2}\hbar[(\omega_0 - \omega)\sigma_z + \omega_1\sigma_x] , \tag{5.54}$$

where the effect of the rotation is seen to reduce the contribution from the magnetic field in the z-direction. We focus on one of the eigenstates of $\hat{H}_R$, with spin vector in the direction of the effective magnetic field, with unit vector $\mathbf{n}$ defined by polar angles with $\phi = 0$ and
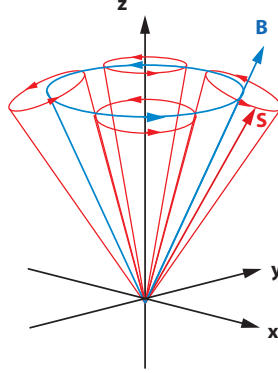
Figure 5.2: Spin precession in a slowly varying magnetic field. The direction of the magnetic field **B** is slowly rotating around the z-axis (indicated by the blue cone), while the spin vector $\mathbf{S} = (\hbar/2)\mathbf{n}$ is rapidly precessing around the **B** field (indicated by the red cones).

$\tan\theta = \omega_1/(\omega_0 - \omega)$. Since the Hamiltonian H_R is time independent, this vector is fixed in the rotating frame, but rotates with angular velocity ω around the z-axis in the non-rotating frame. The time dependent state vector is therefore

$$\psi(t) = e^{-i\alpha(t)}\chi_{\mathbf{n}(t)} , \tag{5.55}$$

with $\mathbf{n}(t)$ as the unit vector of the rotating spin. The spin motion is periodic, with period $T = 2\pi/\omega$.

We will also here examine the phase factor associated with one period of motion of the spin vector. Since the Hamiltonian in the non-rotating frame is time dependent the phase factor of the two-component state vector cannot be determined quite so easily as in the time independent case. We can find it by use of the time evolution operator, which we have earlier found, but rather than doing this we will use Eq. (5.39).

We first find the energy expectation value of the rotating spin state

$$
\begin{aligned}
E &= \langle \psi(t)|\hat{H}(t)|\psi(t)\rangle \\
&= \langle \psi(0)|\hat{H}(0)|\psi(0)\rangle \\
&= \frac{1}{2}\hbar\langle\chi_{\mathbf{n}}|\omega_0\,\sigma_z + \omega_1\sigma_x|\chi_{\mathbf{n}}\rangle \\
&= \frac{1}{2}\hbar(\omega_0\cos\theta + \omega_1\sin\theta) \\
&= \frac{1}{2}\hbar\frac{\omega_0(\omega_0 - \omega) + \omega_1^2}{\Omega} .
\end{aligned}
\tag{5.56}
$$

The result is time independent since the state vector rotates with the same angular frequency as the time dependent part of the magnetic field. The time integral of the expectation value then is

$$\frac{i}{\hbar}\int_0^T E\,dt = \pi\frac{\omega_0(\omega_0 - \omega) + \omega_1^2}{\omega\,\Omega} . \tag{5.57}$$

The geometric phase of the closed path is, as given in (5.50),

$$\alpha_B = \pi(1 - \cos\theta) = \pi(1 - \frac{\omega_0 - \omega}{\Omega}),\tag{5.58}$$

and together they give the phase

$$\alpha(T) = \pi(1 + \frac{\Omega}{\omega}).\tag{5.59}$$

Let us consider what happens in the *adiabatic limit*, that is when the motion of the magnetic field is sufficiently slow so that the state vector $\chi_{\mathbf{n}(t)}$, during the motion will be an eigenvector of the time dependent Hamiltonian $\hat{H}(t)$. For the angular frequency this means $\omega << \omega_0$ and $\omega << \omega_1$. Expanding the phase angle in powers of ω, we find

$$\begin{aligned}
\alpha(T) &= \pi\left(1 + \frac{1}{\omega}\sqrt{(\omega_0 - \omega)^2 + \omega_1^2}\right) \\
&= \pi\left(1 + \frac{1}{\omega}\sqrt{\omega_0^2 + \omega_1^2} - \frac{\omega_0}{\sqrt{\omega_0^2 + \omega_1^2}}\right) + \mathcal{O}(\omega) \\
&= \frac{1}{2}T\Omega_0 + \pi(1 - \frac{\omega_0}{\Omega_0}) + \mathcal{O}(1/T),
\end{aligned}\tag{5.60}$$

with $\Omega_0 = \sqrt{\omega_0^2 + \omega_1^2}$. The first term, which is proportional to the period T of the motion, is the dynamic part, with $\hbar\Omega_0$ as the energy of the spin state. The second term, which is independent of T, is the geometric phase, where we identify $\cos\theta_0 = \omega_0/\Omega_0$ with θ_0 as the angle between the spin vector and the z-axis in the adiabatic limit.

This result is a special case of the situation with adiabatic evolution in a system with a time dependent Hamiltonian. If such a system initially is in one of the eigenstates of the Hamiltonian, it will continue to be so during the slow evolution, even when the energy slowly changes.The phase factor picked up in a full period, will consist of two parts. The dynamical phase is proportional to the time integrated energy, and the non-dynamical part is identical to the geometric phase of the curve in parameter space, as in (5.59). This is in turn consistent with the more general expression for the phase factor, represented in Eqs.(5.39) and (5.40).