III(Ch1-6).tex — January 14, 2000

P.J.M. Bongaarts
Instituut Lorentz, University of Leiden
bongaart@lorentz.leidenuniv.nl

*Topics from $20^{th}$ century physics.*
*An introductory course for students in mathematics*

# III. QUANTUM THEORY: CHAPTERS 1 - 6

## – GENERAL PRINCIPLES

# 1. HISTORICAL BACKGROUND

## 1.1. *Introduction*

Quantum mechanics as we know it, essentially in the form in which it is now taught as a standard subject to physics undergraduates, came into being in the course of a few years, more or less in the period from 1925 to 1928. Its main creators were Schrödinger, Heisenberg, Max Born , Jordan, Pauli and Paul. It explained the structure of atoms, providing a theoretical basis for the model that had been suggested by the experiments of Rutherford, solving the problem of the discreteness of atomic spectra and predicting in fact the numerical values of such spectra with great accuracy. Because of this it was immediately accepted by the physics community even though it was long considered to be an extremely difficult and abstruse theory. This was due to its new and unusual physical features and to its use - often in an implicit way - of new mathematical methods.

## 1.2 *Atomic spectra*

Maxwell's theory of electromagnetic phenomena predicts *electromagnetic radiation*, i.e. electromagnetic waves propagating through empty space with a constant velocity $c = 2.99 \times 10^8$ m/s. An electromagnetic wave has a frequency $\nu$ and a wave length $\lambda$; one has of course the relation $c = \nu\lambda$. Radiation with frequencies $\nu$ roughly between the values $40 \times 10^{13}$ and $75 \times 10^{13}$ Hertz (number of oscillations per second) has been known for a long time as visible light. The particular mix of frequencies emitted by the sun in this frequency regime is seen by our eyes as 'white' light. Using refraction by a glass prisma Newton showed that it can be broken up into a band of coloured light, varying continuously from red (the lowest visible frequency) to violet (the highest visible frequency). Beyond these limits there is on one side first the ulta-violet regime with after this in the very high frequencies $X$-rays, and on the other side the infra-red region, first pure heat radiation and then in the lower frequencies the electromagnetic waves used for radio and television.

In certain circumstances, for instance by external stimulation, electric sparks passing through a gas, etc., atoms emit electromagnetic radiation. This radiation consists of a great number of sharply defined frequencies, some in the visible, some in the infra-red or ultra-violet region. Each element has its own very characteristic system of frequencies, its *spectrum*. The simplest is that of hydrogen. Its spectrum can be described by a simple general formula

$$\nu = cR\left(\frac{1}{n_a^2} - \frac{1}{n_b^2}\right),$$

with $c$ the velocity of light, $R$ the so-called Rydberg constant equal to $1.097 \times 10^7$ m$^{-1}$ and the $n_a$ and $n_b$ running through the positive integers, with the restriction $n_a < n_b$. For fixed $n_a = 1$ (and variable $n_b$) this gives what is called the Lyman series, lying in the ultra-violet; for $n_a = 2$ one has the Balmer series, in the visible region, and for $n_a = 3$ the Paschen series in the infra-red. Note that this formula was found in the late nineteenth century, *not* from theoretical arguments, but deduced from the results of spectroscopic measurents, which gave also the precise data of the spectra of many other types of atoms, much more complicated than the hydrogen spectrum.

A description of an atom in terms of the interaction of Newtonian mechanics and Maxwellian electromagnetism will indeed predict atomic radiation, but only radiation with a continuously varying frequency spectrum. The experimentally established systems of *discrete* atomic spectra, different and characteristic for different types of atoms, cannot be understood from classical physic, i.e. from a combination of classical mechanics and classical electromagnetism. This was a major fundamental problem in physics at the end of the nineteenth century physics, a problem quantum mechanics was able to solve.

1.3. *Earlier ideas*

The roots of quantum mechanics can be found in various earlier ideas:

a. In 1900 Planck gave a formula for the frequency distribution of radiation which correctly described the experimentally observed distribution. For the derivation of this formula he assumed that radiation with frequency $\nu$ could only be emitted or absorbed in discrete quantities, 'energy quanta' $h\nu$, with $h$ a universal constant. He made this discreteness assumption with great reluctance, because it did not fit in with classical notions of electromagnetic radiation. The only justification for it was that it led to a formula which was experimentally correct. The constant $h$, *Planck's constant* as it is now called, is in MKS (meter - kilogram - second) units equal to $6.626 \times 10^{-34}$ k.m$^2$/s, and in CGS (centimeter - gram - second) units $6.626 \times 10^{-27}$ g.cm$^2$/s, so it is extremely small. Because inconvenient factors $2\pi$ appear in many formulas, one often uses $\hbar = \frac{h}{2\pi}$ instead of $h$. This $\hbar$ is in MKS units equal to $1.054 \times 10^{-34}$ kg.m$^2$/s. The appearance of $\hbar$ or $h$ is typical for the formulas in quantum mechanics.

b. Certain metals, when irradiated with light, emit a stream of *electrons*, very light electrically charged particles, which had been discovered in 1897. Einstein gave in 1905 a theoretical explanation of this *photo-electric effect*, by further developing Planck's idea, assuming that light of frequency $\nu$ consisted in fact of a particular kind of particles, with energy $E = h\nu$, and moving with the velocity of light. These particles he called *photons*.

c. Arthur Compton, an American Physicist, found in 1924 that the results of experiments in which X-rays - electromagnetic radiation with much shorter wave length than visible light - were scattered by matter, could be explained by making the same assumption about the corpuscular nature of radiation. This particle character came out even stronger than in the photo-electric effect: the photons had not only energy but also momentum.

d. Rutherford, a physicist from New-Zealand working in England, concluded in 1911 from his scattering experiments, that an atom should be thought of as a very small planetary system, consisting of a heavy nucleus, surrounded by a cloud of electrons, compensating with their electrical charges the opposite charge of the nucleus. For this a theoretical model with very unusual and nonclassical features was postulated by his assistant Niels Bohr, a young Danish theoretical physicist. Bohr assumed that the electrons could only move in discrete stationary orbits, circles or ellipses, once in a while jumping suddenly between two such orbits and emitting in the process a quantum of radiation $E = h\nu$, with $E$ the difference in the energies of the two orbits. With this model he was able to obtain good numerical values for many atomic spectra.

### 1.4. *The emergence of quantum mechanics proper*

The great achievement of quantum mechanics, as it finally appeared in 1925 in the fundamental papers of Heisenberg, Born, Schrödinger and others, was that it gave a *fundamental* and *general* theoretical framework in which all these ideas found their natural place and in which in particular the ad-hoc atomic model of Rutherford and Bohr could be understood and derived from first principles.

### 1.5. *Quantum theory in $20^{th}$ century physics*

Quantum theory is one the two great revolutions that have changed the face of physics in the $20^{th}$ century, the other being Einstein's theory of relativity. Combining the two in a consistent way remains an open problem and is the most important challenge facing present-day theoretical physics.

## 2. THE BEGINNING OF QUANTUM MECHANICS

### 2.1. *Two different forms*

Quantum mechanics began in two seemingly different forms: Heisenberg's *matrix mechanics* and Schrödinger's *wave mechanics*.

In Heisenberg's approach the basic classical physical variables of position $x_j$ and momentum $p_j$ of a particle were replaced by somewhat mysterious algebraic quantities $\hat{x}_j$ and $\hat{p}_j$, which no longer commuted with each other, but satisfied instead the *commutation relation*

$$[\hat{p}_j, \hat{x}_k] = \hat{p}_j\hat{x}_k - \hat{x}_k\hat{p}_j = \frac{\hbar}{i}\delta_{jk},$$

with $j, k = 1, 2, 3$ and with $\hbar$ Planck's constant. Other physical quantities, of which the energy $E$ was the most important, were expressed in the basic variables, using as much as possible the classical expressions. It was soon realized that these algebraic *quantum variables* were in fact infinite matrices, and that the basic problem that had to be solved for each physical system in this approach was to find eigenvalues and eigenvectors of the matrix corresponding with the energy.

Schrödinger had been influenced by the ideas of the French physicist Louis de Broglie, who had suggested a few years earlier that a possible way to understand the discrete orbits in Bohr's model of an atom was to associate with a moving particle of momentum $p$ a wave with wave lenght $\lambda = h/p$, with $h$ Planck's (original) constant. Orbits were as a consequence restricted by the requirement that they had to consist of a certain number of wave lenghts fitted together. Schrödinger made this rather vague but intuitively appealing idea more precise by his proposal to describe the motion of an electron by a complex-valued function $\psi$ depending on the position variables $x_j$ and the time variable $t$. This 'wave function' had to satisfy a partial differential equation, which for an electron of mass $m$ moving in a potential $V(x_1, x_2, x_3)$ was

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi,$$

with $\Delta$ the Laplace operator $\frac{\partial^2}{\partial {x_1}^2} + \frac{\partial^2}{\partial {x_2}^2} + \frac{\partial^2}{\partial {x_3}^2}$. Schrödinger originally thought that the wave function described the electron 'smeared out' in space, but it was soon pointed out by others that it had a *statistical* interpretation, with the square of the absolute value of $\psi$ at each point in space describing the probability of finding the electron - still a point particle - there.

### 2.2. *Unification*

Schrödinger discovered that solving his differential equation amounted to solving the eigenvalue - eigenvector problem of the energy matrix in Heisenberg's approach; some time after this the full mathematical and physical equivalence of both approaches was established. The mathematician John von Neumann developed around 1930 a unified and rigorous framework for quantum mechanics in which this became completely clear. His framework is now generally accepted as the standard mathematical formulation of quantum mechanics. It will be the basis for the discussion of quantum mechanics in this course.

## 3. MODERN QUANTUM THEORY: GENERAL REMARKS

### 3.1. *A non-historic presentation*

In the preceding two chapters the historical background of quantum mechanics was sketched and a few remarks about its beginning were made. This historical path will not be followed further in this course. In this and subsequent chapters, quantum theory will be discussed as it is now. The presentation will be 'axiomatic'; a small number of principles will be postulated in a precise mathematical manner, on the basis of which the full theory, with various explicit examples, can be developed. The mathematics needed for this is separately reviewed in Appendices B (*Hilbert space*) and C (*Probability theory*).

### 3.2. *Von Neumann's formulation*

The mathematical basis of quantum theory, and in particular quantum mechanics, von Neumann's Hilbert space frame work, is well-understood. Hilbert space theory, like functional analysis in general, is linear algebra in infinite dimensional vector spaces, to which nontrivial topological and measure theoretical notions have been added to cope with limits, with infinite sums and integrals, needed to make it interesting and fruitful. The standard physics text books ignore this to a large extent; they can be said to follow in principle von Neumann's scheme - often without mentioning his name, or the term Hilbert space, reducing it to a heuristic level where problems of convergence and divergence in infinite sums are ignored and where it is considered to be sufficient to substitute integrals for discrete sums, when continuous eigenvalues make this necessary. It has to be admitted that this attitude can be justified on the practical ground that it does not cause great harm in most standard day-to-day applications of quantum theory to physical problems. The aim of this course is however *not* to teach quantum theory to future practitioners of the subject, but to present it as one of the great intellectual achievements of the natural sciences in the twentieth century. For this a satisfactory treatment of its mathematical foundations is required. This has the added advantage that many notions and arguments become clearer or seem less arbitrary when the underlying mathematical framework of quantum mechanics is kept in sight.

3.3. *Quantum theory and quantum mechanics*

A distinction will be made between the terms 'quantum mechanics' and 'quantum theory'.

a. *Quantum mechanics* is the theory invented by Heisenberg, Schrödinger and others for the description of atomic structure, as was sketched in the preceding chapters. It is called quantum *mechanics* because most of its observable quantities, in particular position, linear momentum and the expression of the energy, come from the *classical mechanics* of point particles. It is an important part of modern physics and is used is for a wide range of submicroscopic phenomena; it is particularly important for its applications in many areas of modern technology. Its mathematical basis - essentially von Neumann's Hilbert space formalism - is completely understood.

b. In the course of the past seventy-five years quantum mechanics has been generalized to cover other physical phenomena. We now have *quantum statistics* for collective phenomena such as in solids and fluids, *relativistic quantum field theory* for the description of subnuclear particles, various attempts at *quantum gravitation*, etc.. Not all of this is at present completely successful or well-understood.

By *quantum theory* I will mean the *general* theory which encompasses all these developments, and which is believed to describe in principle *all* known physical phenomena. It is mathematically formulated in a general manner in terms of *states*, as vectors in Hilbert space, *observables*, as selfadjoint operators, and *time evolution and symmetries*, as groups of unitary operators, even though much of this has not been made expicit in all cases. This means that there are challenging fundamental problems that remain to be solved. This is true in particular for *relativistic quantum field theory* and its derivatives.

In this course the formalism of quantum theory, in terms of Hilbert space vectors and operators, will be presented in its general form, together with its physical interpretation, but the explicit examples by which all this will be illustrated come from quantum mechanics. In this sense quantum theory will be restricted in this course to the special case of quantum mechanics.

### 3.4. *The main elements*

The main elements of the description of a physical system in quantum theory are *states*, *observables*, *time evolution* and *symmetries*. In Chapter 4 mathematical definitions of the notions of state and observable will be given, together with their physical interpretation. Chapter 5 will discuss time evolution; symmetries will be briefly touched upon in Chapter 6.

## 4. STATES AND OBSERVABLES

### 4.1. *Two basic principles*

In classical mechanics the *state* of a system of $N$ point particles, can be described completely at each moment in time by $6N$ numbers, the values of $3N$ coordinates and $3N$ momenta. *Observables* are functions of these variables. This is typical for all situations in classical physics. In quantum mechanics, and in quantum theory in general, the characterization of the state of a system and the description of observables is radically different; it includes a notion of indeterminacy described in probabilistic terms. We formulate the first two principles on which quantum theory is built:

I. *The state of a physical system, at each moment in time, is represented by a unit vector in a Hilbert space, the space of state vectors of the system.*

II. *Observable quantities are represented by selfadjoint operators in the Hilbert space of state vectors.*

*Remark*: Hilbert spaces in quantum theory are always over the complex numbers and separable. The selfadjoint operators that represent observables in quantum theory are usually unbounded.

*Example from quantum mechanics*: The Hilbert space $\mathcal{H}$ for a point particle in quantum mechanics is $L^2(R^3, d\vec{x})$, the space of square integrable complex-valued functions in the real variables $\vec{x} = (x_1, x_2, x_3)$. The functions that represent states have unit norm and are usually called *wave functions*. The selfadjoint operators representing the basic observables are the operators $Q_j$ for the position of the particle, acting on functions $\psi$ in $\mathcal{H}$ as

$$(Q_j\psi)(\vec{x}) = x_j\psi(\vec{x}),$$

and $P_j$ for the momentum, acting as

$$(P_j\psi)(\vec{x}) = \frac{\hbar}{i}\frac{\partial}{\partial x_j}\psi(\vec{x}),$$

for $j = 1, 2, 3$. In classical mechanics the energy for a free particle with mass $m$ is $E = \frac{P^2}{2m}$ and $E = \frac{P^2}{2m} + V$ if the particle moves under the influence of a potential $V(\vec{x})$. The energy *operator* in quantum mechanics is the classical expression in the position and momentum operators $P_j$ and $Q_j$, i.e. it acts as

$$(E\psi)(\vec{x}) = -\hbar^2 \sum_{j=1}^{3} \frac{\partial^2}{\partial {x_j}^2} \psi(\vec{x})$$

for a free particle and

$$(E\psi)(\vec{x}) = -\hbar^2 \sum_{j=1}^{3} \frac{\partial^2}{\partial {x_j}^2} \psi(\vec{x}) + V(\vec{x})\psi(\vec{x})$$

for a particle in a potential $V$. In this simple physical situation the quantum theory, in particular the system of quantum observables, is suggested by the elements of the classical description of the model in the picture of Hamiltonian mechanics. This is an example of a general procedure called *canonical quantization*, which will be discussed later in a more systematic manner.

The state space of a system of a quantum mechanical system of $N$ particles consists of square integrable functions in $3N$ variables, as will be discussed in some detail later. A classical *field*, such as an electromagnetic or gravitational field, is a system with an infinite number of 'degrees of freedom'. This is the reason why *quantum field theory* requires a mathematically much more sophisticated approach, in which, for instance, the Hilbert spaces of states of *quantum field theory* are no longer function spaces, but have to be constructed as fairly complicated algebraic objects.

4.2. *The third principle: interpretation of I and II*

The interpretation of I and II is based on the *spectral theorem* for selfadjoint operators in Hilbert space; we formulate it as a third basic principle, first in a provisional and more elementary form, sufficient for the case of discrete spectrum:

$\mathrm{III}_0$. *The possible outcomes of measurements of an observable $A$ in a state $\psi$ are given by a probability distribution concentrated on the spectrum of $A$, such that the expectation is given by*

$$\overline{A} = (\psi, A\psi),$$

*and more generally the $n^{th}$ moments by*

$$\overline{A^n} = (\psi, A^n\psi).$$

Our definitive, more sophisticated version applies to general spectra:

III. *A pair $(\psi, A)$, consisting of a unit vector $\psi$ and a selfadjoint operator $A$, representing respectively a state of the system and an observable, determine a* *stochastic variable in the sense of standard probability theory. The distribution function of this variable is given by $F_\psi(\alpha) = (\psi, E_\alpha \psi)$, with the projection operators $E_\alpha$ belonging to the spectral resolution of $A$. This implies that it is concentrated on the spectrum of $A$ and that it is such that its mean or expected value - or expectation value as it is called in physics - is equal to*

$$\overline{A} = (\psi, A\psi) = \int_{-\infty}^{+\infty} \alpha \, d(\psi, E_\alpha \psi),$$

*with more generally the $n^{th}$ moments equal to*

$$\overline{A^n} = (\psi, A^n \psi) = \int_{-\infty}^{+\infty} \alpha^n \, (\psi, E_\alpha \psi).$$

*Remark*: The expectation, or more generally the $n^{th}$ moment, of a probability distribution may not be finite. Here $\overline{A^n}$ is finite whenever $\psi$ is in the domain of the operator $A^n$.

4.3. *The case of discrete spectrum*

Let us consider first the technically simple case of a selfadjoint operator $A$ with a purely discrete spectrum, consisting of eigenvalues $\alpha_1, \alpha_2, \ldots$ which for the sake of further simplicity are supposed to be nondegenerate. The spectral theorem in its simplest form tells us that there is an orthogonal basis of eigenvectors $\phi_1, \phi_2, \ldots$, uniquely determined up to phase factors. The vector $\psi$ can be expanded as

$$\psi = \sum_{j=1}^{\infty} c_j \phi_j,$$

with the complex coefficients $c_j$ satisfying the normalization condition

$$\sum_{j=1}^{\infty} |c_j|^2 = (\psi, \psi) = 1.$$

This means that the real numbers $\rho_j = |c_j|^2$ form a discrete probability distribution. The expression for the expectation becomes

$$(\psi, A\psi) = \sum_{j,k=1}^{\infty} \overline{c_j} c_k \, \alpha_k (\phi_j, \phi_k) = \sum_{j=1}^{\infty} \alpha_j |c_j|^2 = \sum_{j=1}^{\infty} \alpha_j \rho_j.$$

and for the $n^{th}$ moment

$$(\psi, A^n \psi) = \sum_{j=1}^{\infty} \alpha_j^n \rho_j.$$

An example of this simple situation, which will be discussed later, is the energy of a one-dimensional harmonic oscillator. Its energy operator has nondegenerate eigenvalues $E_n = \hbar\omega(n+\frac{1}{2})$, for $n = 0, 1, \ldots$, and with $\omega$ the basic oscillator frequency . The generalization to the situation in which (some of ) the eigenvalues $\alpha_j$ are degenerate is simple and can be left to the reader.

If $\psi$ happens to be one of the eigenvectors of the operator $A$, then the probablity distibution will be concentrated on the corresponding eigenvalue, i.e. a measurement will give with certainty this eigenvalue.

4.4. *The case of continuous spectrum*

For the case of an operator with continuous spectrum we need the more sophisticated version of III, based on the general version of the spectral theorem. Eigenvectors no longer exists, at least not as proper normalized vectors in $\mathcal{H}$; instead of a sequence of projection operators on eigenspaces one has to use a continuous system of spectral projections, the spectral resolution of the general form of the spectral theorem, involving an operator valued Stieltjes integral, as discussed in Appendix B.

*Example*: Let us consider for the sake of simplicity a particle in 1-dimensional space. Its quantum mechanical state space is the Hilbert space $\mathcal{H} = L^2(R^1, dx)$. There is a single position operator $Q$. It has the spectral projections operators $E_\lambda$ defined as

$$(E_\lambda\psi)(x) = \psi(x),$$

for $x \leq \lambda$ and

$$(E_\lambda\psi)(x) = 0,$$

for $x > 0$. For a particular choice of state vector $\psi$ one has as distribution function $F_\psi$ in this case

$$F_\psi(\lambda) = (\psi, E_\lambda\psi) = \int_{-\infty}^{\lambda} |\psi(x)|^2 dx$$

which means that the probability distribution determined by the position operator $Q$ and the normalized state vector $\psi$ is absolutely continuous with respect to the Lebèsgue measure on $R^1$, i.e. can be given by the probablity density $\rho(x) = |\psi(x)|^2$. The expectation value for $Q$ in the state $\psi$ is

$$(\psi, Q\psi) = \int_{-\infty}^{+\infty} x\rho(x)dx = \int_{-\infty}^{+\infty} x|\psi(x)|^2 dx.$$

Note that this is consistent with the fact that a function $\psi(x)$ is in the domain of the operator $Q$ if and only if $\psi(x)$ *and* $x\psi(x)$ are square integrable.

We have found in this way the well-known physical interpretation of the 'wave function' in quantum mechanics:

*The square of the absolute value of the wave function of a quantum mechanical particle is the probability density for the measurement of the position of that particle.*

In general a selfadjoint operator may have both discrete and continuous spectrum. An example is the operator that represents the energy of a hydrogen atom. It has an infinite system of discrete values, starting at a lowest nondegenerate 'ground state energy' and then converging to a point where a continous spectrum begins which goes all the way to $+\infty$. A hydrogen atom is thought of as a system consisting of an electron moving in the electrostatic field of a nucleus. The discrete spectrum represents the *bound states* of this system, the continuous spectrum corresponds to the situation where the atom is *ionized*, i.e. with the electron separated from the nucleus. All this will be discussed later.

4.5. *Systems of observables*

So far we have only discussed single observables. In the case of *systems* of observables, there are new and interesting phenomena, that follow from the basic assumptions and which are typical for quantum theory. Let us consider two observables represented by selfadjoint operators $A$ and $B$. Two different cases should be distinguished:

a. The operators $A$ and $B$ *commute*. Note that the correct definition for this is that the spectral projections of $A$ commute with those of $B$; the relation $[A, B] = 0$ would in general involve domain problems.

Suppose first that $A$ and $B$ have only discrete eigenvalues, say $\alpha_1 < \alpha_2 < \ldots$ and $\beta_1 < \beta_2 < \ldots$. A generalization of the spectral theorem to *systems* of commuting operators would mean for this case that there exist an orthonormal basis of common eigenvectors of $A$ and $B$. For general spectrum there is a 2-parameter spectral resolution $\{E_{\alpha\beta}\}$. An arbitrary unit vector $\psi$ in $\mathcal{H}$ gives then a joint distribution function $F_\psi$, in the sense of probability theory, according to $F_\psi(\alpha, \beta) = (\psi, E_{\alpha\beta}\psi)$. The obvious physical interpretation of this is that it is the joint probability for finding the values $\alpha$ and $\beta$ in a *simultaneous* measurement of the two observables in the state represented by $\psi$.

For a system of $n$ observables, represented by $n$ commuting selfadjoint operators $A_1, \ldots, A_n$ we formulate a generalization of III:

IV. *A unit vector $\psi$ together with n commuting selfadjoint operators $A_1, \ldots, A_n$, representing a state and n observable quantities of a quantum system, determine a system of stochastic variables in the sense of ordinary probability theory. The joint distribution function of these variables is given by*

$$F_\psi(\alpha_1, \ldots, \alpha_n) = (\psi, E_{\alpha_1 \ldots \alpha_n}\psi),$$

*with $\{E_{\alpha_1 \ldots \alpha_n}\}$ the spectral resolution associated with $A_1, \ldots, A_n$, according to the general spectral theorem for systems of selfadjoint operators.*

DEFINITION: A system of observables represented by commuting selfadjoint operators will be called a system of *commensurable* observables.

Principle IV means that the physical interpretation of a system of commensurable observables is part of standard probability theory. An illustration of this is the expression for the covariance of two commensurable observables $A$ and $B$ in a state $\psi$, which is equal to

$$(\overline{AB})_\psi = (\psi, AB\psi) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \alpha\beta \, dF_\psi(\alpha,\beta),$$

and which in case of absolute continuity can be written as an integral over a probability density $\rho_\psi$ as

$$(\overline{AB})_\psi = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \alpha\beta\, \rho_\psi(\alpha,\beta)\, d\alpha d\beta.$$

*Example*: Consider a point particle in 3-dimensional space. Its quantum state space $\mathcal{H}$ is $L^2(R^3, d\vec{x})$. The three components of the position are represented by the commuting selfadjoint operators $Q_j$, defined as

$$(Q_j\psi)(\vec{x}) = x_j\psi(\vec{x}),$$

for $j = 1,2,3$. The spectral resolution of this system consists of projection operators $E_{\lambda_1\lambda_2\lambda_3}$ given by

$$(E_{\lambda_1\lambda_2\lambda_3}\psi)(x_1,x_2,x_3) = \psi(x_1,x_2,x_3),$$

for $\lambda_j \leq x_j$, $j = 1,2,3$ and

$$(E_{\lambda_1\lambda_2\lambda_3}\psi)(x_1,x_2,x_3) = 0,$$

for the other values of $x_1, x_2, x_2$. For a state vector $\psi$ with $||\psi|| = 1$ this gives a joint distibution function

$$F_\psi(\lambda_1,\lambda_2,\lambda_3) = (\psi, E_{\lambda_1\lambda_2\lambda_3}\psi) = \int_{-\infty}^{\lambda_1}\int_{-\infty}^{\lambda_2}\int_{-\infty}^{\lambda_3} ||\psi(x_1,x_2,x_3)||^2\, dx_1dx_2dx_3$$

Because of absolute continuity everything can be written in terms of the joint probablity density $\rho(x_1,x_2,x_3) = ||\psi(x_1,x_2,x_3)||^2$. All this means that *as far as the position of the particle is concerned*, the quantum mechanical description remains conceptually within the framework of standard probability theory.

b. The operators $A$ and $B$ do *not* commute. The occurrence of such systems of *incommensurable observables* is a typical quantum phenomenon. After a choice of a state vector $\psi$ the $A$ and $B$ give *separate* stochastistic variables; together these do *not* form a *system* of stochastic variables in the sense of standard probability theory: i.e. there does not exist a common probability measure on which these variables are measurable functions and there is no joint distibution function or probability density. There are nevertheless relations between the results of measurements of the two observables, relations which again are typical for quantum theory.

*Example*: The operators for position and momentum of a 1-dimensional particle, $P$ and $Q$, do not commute, are incommensurable. The same is true for $Q_j$ and

$P_j$ for a 3-dimensional particle. There are no common probability densities in these cases. There is however an interesting and important relation between the measurement of (the corresponding components of) of position and moment. It is called the *Heisenberg uncertainty relation* and will be discussed later.

### 4.6. *The measurement process*

The prediction of results of measurements in quantum theory in terms of probability theory, as described in the preceding sections, is uncontroversial and generally accepted. This is less so for the description of the actual measurement process. According to the orthodox 'Copenhague' interpretation of quantum mechanics, of which Bohr was the main advocate, a distinction must be made between the (microscopic) system that is the subject of the measuring operation, and the (macroscopic) measurement apparatus which is seen as classical. As a consequence of the measurement the wave function which describes the system is supposed to ondergo an instantaneous discontinous change. If, for instance, the measurement of an observable represented by the operator $A$ gives as result the value $a$, neccessarily an eigenvalue of $A$, then the wave function is projected onto the corresponding eigenvector. (A more precise formulation is necessary in the case of continuous spectrum). This idea of the 'reduction of the wave packet' has led to much discussion of a philosophical kind. Questions whether objective and subject aspects of reality can be truly distinguished, whether an object exists when it is not observed, etc., etc.. In recent years the Copenhague interpretation has been criticized and alternatives have been put forward. These matters cannot be discussed here.

## 5. TIME EVOLUTION

### 5.1. *Time evolution as a* 1*-parameter group of operators*

The evolution of a quantum system from an initial time to a later time is described by an *evolution operator* acting on the state vectors in the Hilbert space. It should be invertible and should map unit vectors to unit vectors; this together means that we require it to be a *unitary* operator. We assume furthermore that the systems that we consider are *autonomous* i.e. that their dynamics is not *explicitly* time-dependent. This means that the unitary evolution operator depends only on the *difference* of initial and later times. We think therefore of a system of unitary operators $\{U(t)\}$, carrying the system from a state at an initial time $t_0$ to a later time $t_0 + t$, satisfying the following natural conditions:

a. $U(0) = 1$,

b. $U(t_1)U(t_2) = U(t_1 + t_2)$, for all real $t_1$ and $t_2$.

Note that these two conditions together imply a third property, n.l.

c. $U(t)^{-1} = U(-t)$, for all $t$ in $R$.

The operators $U(t)$ obviously form an abelian group. We assume that the $U(t)$ depend continuously on $t$, in the sense of strong operator convergence. The system $\{U(t)\}$ is what is called a 1*-parameter group of unitary operators*. See

Appendix B, Chapter 7. It is a general property of such a 1-parameter group that it has a selfadjoint generator by which it is uniquely determined.

All this together gives us the general principle describing time evolution in quantum theory:

V. *Time evolution in quantum theory is described by a continuous* 1*-parameter group of unitary operators* $\{U(t)\}$*, acting in the Hilbert space of states* $\mathcal{H}$*. It can be written as* $U(t) = e^{-\frac{i}{\hbar}tH}$*, with* $H$ *a selfadjoint operator which is called the Hamitonian of the system.*

5.2. *The Schrödinger equation*

To describe the dynamics of a concrete physical situation it is the Hamiltonian operator that is explicitly given, not the 1-parameter group of evolution operators. Note that it is called 'Hamiltonian' because – at least in quantum mechanics – it is usually obtained by taking the Hamiltonian function which determines the time evolution of the corresponding classical system and substituting operators for the classical variables of position and momentum. Note also that the Hamiltonian – both in the classical and the quantum case – determines the time evolution *and* represents usually also the *energy* as observable. State vectors evolving in time can be generally written as $\psi(t) = e^{-\frac{i}{\hbar}tH}\psi(0)$. Starting from the Hamiltonian one obtains such time $\psi(t)$ as solutions of the first order equation

$$\frac{d}{dt}\psi(t) = H\psi(t).$$

This equation, which is valid for all $\psi(t)$ in the domain of $H$, may be called the *'abstract' or 'vector-valued' Schrödinger equation*.

*Example*: For a 1-dimensional point particle with mass $m$, moving in a potential $V(q)$, the classical Hamiltonian is $H_{class}(p,q) = \frac{p^2}{2m} + V(q)$. The Hamiltonian of the corresponding quantum system is therefore $H = \frac{P^2}{2m} + V(Q)$, an operator acting on the state vectors, i.e. 'wave functions' $\psi(x)$, as

$$(H\psi)(x) = -\hbar^2 \frac{d^2}{dx^2}\psi(x) + V(x)\psi(x).$$

A time-dependent wave function $\psi(x,t)$ therefore has to satisfy the equation

$$\frac{\partial}{\partial t}\psi(x,t) = -\hbar^2 \frac{d^2}{dx^2}\psi(x) + V(x)\psi(x).$$

This is, for this situation, the 'concrete' *Schrödinger equation*, the central partial differential equation of standard quantum mechanics.

Solving the Schrödinger equation for all state vectors $\psi(0)$ is in principle equivalent to solving the generalized eigenvalue - eigenvector problem for the Hamiltonian operator $H$, i.e. finding its spectral resolution. To understand the idea behind this statement, consider the simple case of an operator $H$ with a purely discreet nondegenerate spectrum. Suppose that its eigenvalues $\alpha_0, \alpha_1, \ldots$ are known, together with an orthonormal system of eigenvectors $\phi_0, \phi_1, \ldots$. An arbitrary vector $\psi$ can be expanded as $\psi = \sum_{n=0}^{\infty} c_n \phi_n$, with the coefficients $c_n$ determined by the inverse formula $c_n = (\psi, \phi_n)$, for $j = 0, 1, \ldots$. For a time-dependent $\psi(t)$ this expansion is $\psi(t) = \sum_{n=0}^{\infty} c_n(t) \phi_n$, with time-dependent coefficients $c_n(t) = (\psi(t), \phi_n)$. Application of the 'vector-valued' Schrödinger equation gives, for each $n$ separately, $\frac{d}{dt} c_n(t) = -\frac{i}{\hbar} E_n c_n(t)$, with solution $c_n(t) = e^{-\frac{i}{\hbar} t E_n} c_n(0)$. This allows us to express $\psi(t)$ in the initial state $\psi(0)$ as $\psi(t) = \sum_{n=0}^{\infty} e^{-\frac{i}{\hbar} t E_n} (\psi(0), \phi_n)$. For the case of a general Hamiltonian the result is essentially the same, although more sophistication is needed to formulate it properly. It explains in any case the relation between Schrödinger's wave mechanics, in which a wave equation had to be solved, and Heisenberg's matrix quantum mechanics, which involved solving the eigenvalue - eigenvector problem of an infinite dimensional matrix, the two versions of early quantum mechanics, which looked quite different at first sight.

### 5.3. *The Heisenberg picture*

All observable effects can be expressed in probabilities of measuring certain observables $A$ in certain states $\psi$. Time evolution is no exception; we do not observe the time-dependence of states directly, but only its consequences for the time-dependence of probabilities. One has for instance for the expectation value of the observable $A$ in of a state $\psi$ the formula

$$(\overline{A})_{\psi}(t) = (\psi(t), A\psi(t)) = (e^{-\frac{i}{\hbar} tH} \psi(0), A e^{-\frac{i}{\hbar} tH} \psi(0)).$$

This can be written as

$$\overline{A})_{\psi}(t) = (\psi(0), e^{\frac{i}{\hbar} tH} A\, e^{-\frac{i}{\hbar} tH} \psi(0)) = (\psi(0), A(t)\psi(0)).$$

Similar formulas hold for the time-dependence of general moments and the distribution function it self. This means that instead of putting the time time-dependence in the state vectors as

$$\psi(t) = e^{-\frac{i}{\hbar} tH} \psi(0),$$

with the observables remaing constant, one may as well let the 1-parameter group $\{U(t) = e^{-\frac{i}{\hbar} tH}$ act on the operators as

$$A(t) = e^{\frac{i}{\hbar} tH} A(0)\, e^{-\frac{i}{\hbar} tH},$$

with the state vectors kept constant. The first procedure, which is the usual one in most of elementary quantum mechanics, is called the *Schrödinger picture*. The second way of looking at time evolution, with constant states and time-dependent observables, is known as the *Heisenberg picture*; it is used in some of the more advanced parts of quantum theory. Both pictures are physically completely equivalent; the choice between them is purely a matter of convenience. In the Heisenberg picture the basic dynamical equation is a differential equation for operators. By differentiating $e^{\frac{i}{\hbar}tH}A(0)\,e^{-\frac{i}{\hbar}tH}$ with respect to time one obtains

$$\frac{d}{dt}A(t) = \frac{i}{\hbar}[H, A(t)],$$

the *Heisenberg equation*, which in the Heisenberg picture plays the role of the Schrödinger equation in the Schrödinger picture.

5.4. *Stationary states and constants of the motion*

DEFINITION: A *stationary state* of a quantum system is a state vector $\psi$ for which the probabilities of measurements of *all observables* are constant in time.

THEOREM: A unit vector in $\mathcal{H}$ is a stationary state if and only if it is an eigenvector of the Hamiltonian.

*Proof*: The action of $U(t) = e^{-\frac{i}{\hbar}tH}$ on an eigenvector $\phi_n$ multiplies it by a phase factor $e^{-\frac{i}{\hbar}tE_n}$. Such a phase factor does not show up in any of the formulas used to calculate probabilities. From this it follows easily that $\phi_n$ is a stationary state. Suppose on the other hand that $\phi$ is a stationary state. Denote $e^{-\frac{i}{\hbar}tH}\phi$ as $\phi(t)$. Then one has for an arbitrary selfadjoint operator $A$ with spectral resolution $\{E_\alpha\}$ that the distribution function $F^A_{\phi,t}(\alpha) = (\phi(t), E_\alpha\phi(t))$ is independent of time, for all $\alpha$ in $R$. Because $A$ is arbitrary, we can choose as projection $E_\alpha$ the 1-dimensional projection $E_\phi$ on $\phi$. This gives that $(\phi(t),\phi)(\phi,\phi(t)) = |(\phi(t),\phi)|^2$ is independent of $t$ which implies $|(\phi(t),\phi)|^2 = |(\phi,\phi)|^2$, or because $||\phi(t)|| = ||\phi||$, $|(\phi(t),\phi)| = ||(\phi(t)||\,||\phi)||$. The equality part of the statement of the inequality of Schwarz then gives that $\phi(t)$ is a scalar multiple of $\phi$, for every real $t$. See Appendix B, Section 2.4. Some further simple arguments give finally that $\phi$ is an eigenvector of $H$. This was to be proved.

*Remark*: A freely moving point particle has a Hamiltonian with purely continuous spectrum, so this system has no stationary states. The Hamiltonian of a harmonic oscillator, has a a purely discrete spectrum, and has therefor an orthonormal basis consisting of stationary states. The Hamiltonian of a hydrogen atom has both discrete and continuous spectrum, corresponding with stationary states, the 'Bohr orbits', and nonstationary 'ionized states' respectively.

DEFINITION: A *constant of the motion* or *conserved quantity* is an observable for which the probabilities of measuring it in *all states* are constant in time.

THEOREM: An observable is a constant of the motion if and only if it commutes with the Hamiltonian.

*Proof*: Let $A$ be an observable, i.e. a selfadjoint operator, with spectral resolution $\{E_\alpha\}$. Define for each $t$ the selfadjoint operator $A(t) = e^{\frac{i}{\hbar}tH}A\,e^{-\frac{i}{\hbar}tH}$, with a spectral resolution consisting of the projections $E_\alpha(t) = e^{\frac{i}{\hbar}tH}E_\alpha\,e^{-\frac{i}{\hbar}tH}$. Let $\psi$ be an arbitrary state vector, with $\psi(t) = e^{-\frac{i}{\hbar}tH}\psi$. We have distribution functions $F^A_{\psi,t}(\alpha) = (\psi(t), E_\alpha\psi(t)) = (\psi, E_\alpha(t)\psi)$. The operator $A$ is a constant of the motion if and only if $(\psi, E_\alpha(t)\psi) = (\psi, E_\alpha\psi)$, for all states vectors $\psi$ and all real $t$ and $\alpha$. This is equivalent to the operator identity $E_\alpha(t) = E_\alpha$, for all $\alpha$ and $t$, or $E_\alpha\,e^{-\frac{i}{\hbar}tH} = e^{-\frac{i}{\hbar}tH}E_\alpha$, for all $\alpha$ and $t$, which is in turn equivalent to the statement that all spectral projections of $A$ commute with all spectral projections of $H$. This is the definition of $A$ commuting with $H$, which proves the theorem.

An example of a constant of the motion is the momentum of a free particle. Another important example is formed by the three components of a 3-dimensional particle moving in a rotationally invariant potential. Constants of the motion usually are connected with *symmetries*.

## 6. SYMMETRIES

### 6.1. *Groups of symmetries*

Symmetry is a an important notion in physics. It simplifies the solution of concrete problems and is often a guide in finding new models for physical phenomena.

The symmetries of a mathematical object are its *automorphisms*, i.e. the invertible maps from the object onto itself, which leave its characteristic structure invariant. For a Hilbert space these are the unitary operators. A quantum theory, as a mathematical object a pair $(\mathcal{H}, \{U(t)\}_{t\in R})$, consisting of a Hilbert space of states $\mathcal{H}$ and a 1-parameter group $\{U(t)\}_{t\in R}$ of unitary time-evolution operators, has as its automorphisms the unitary operators in $\mathcal{H}$ that commute with all the unitary operators $U(t)$ of the time-evolution group, or equivalently with the selfadjoint generator $H$ of this group. We state therefore as basic principle:

VI. *Symmetries in quantum theory are unitary operators which commute with the* 1*-parameter group of time involution, or equivalently, with its generator, the Hamiltonian.*

Examples of symmetries in quantum mechanics are *space reflection* in the case of a 1-dimensional point particle, when the particle is moving in a potential $V$ with $V(x) = V(-x)$, *space translation* for a free particle, and *spatial rotation* for a 3-dimensional particle moving in a rotation invariant potential.

Of particular importance are *groups* of symmetries:

DEFINITION: The group $\mathcal{G}$ is a symmetry group of a quantum system if there is a unitary representation $\pi$ of $\mathcal{G}$ in $\mathcal{H}$ with the property that, for all $g$ in $\mathcal{G}$, the operators $\pi(g)$ commute with the time evolution operators $U(t)$, for all $t$ in $R$, or equivalently, with the Hamiltonian $H$.

### 6.2. *Infinitesimal symmetries*

Many of the symmetry groups in physics are 'continuous' groups, i.e. Lie groups. A Lie group $\mathcal{G}$ has a Lie algebra $\mathcal{L}(\mathcal{G})$. Elements of $\mathcal{L}(\mathcal{G})$ can be obtained – roughly – by differentiating 1-parameter groups in $\mathcal{G}$ at the identity element. (See Appendix D for a brief review of the properties of Lie groups, Lie algebras, etc.) A representation $\pi$ of $\mathcal{G}$ gives a (linear) representation $\hat{\pi}$ of $\mathcal{L}(\mathcal{G})$. If

$\mathcal{G}$ is simply connected then $\pi$ can – in principle – be recovered from $\hat{\pi}$ by exponentiation

$$\pi(e^{-i\tau h}) = e^{-\tau\hat{\pi}(h)},$$

for all $h$ in $\mathcal{L}(\mathcal{G})$ and all real $\tau$. If $\mathcal{G}$, together with the representation $\pi$, is a symmetry, all operators $\hat{\pi}(h)$ will commute with time evolution. This suggests the following definition:

DEFINITION: A Lie algebra $\mathcal{L}$ is an *infinitesimal symmetry* of a quantum system if there is a representation $\hat{\pi}$ of $\mathcal{L}$ in $\mathcal{H}$ such that, for all $h$ in $\mathcal{L}$, the operators $\hat{\pi}(h)$ commute with the time evolution operators $U(t)$, for all $t$ in $R$, or equivalently, with the Hamiltonian $H$.

Lie groups are manifolds. The great advantage of Lie algebras over Lie groups is that Lie algebras are linear spaces. An $n$-dimensional Lie group $\mathcal{G}$ has an $n$-dimensional Lie algebra $\mathcal{L}(\mathcal{G})$, in which a basis can be chosen, say $e_1, \ldots e_n$. This simplifies the condition for infinitesimal symmetry, and for 'noninfinitesimal' symmetry in the case of a simply connected group Lie group, to a finite set of relations, the equations

$$[\hat{\pi}(e_j), H] = 0,$$

for $j = 1, \ldots, n$. The $\hat{\pi}(e_j)$ are usually called the *generators* of the symmetry. Working with Lie algebra generators instead of with the full group is very popular in concrete physical applications of symmetry in quantum theory. Note that the generators are usually unbounded operators, and that therefore some of the above statements should be made more precise by taking into account domain questions. We will not worry about this, as it does not do much harm in practice, at least not in elementary quantum mechanics.

It should finally be remarked that symmetry generators are *constants of the motion*. This in another reason why symmetries are important. Some of the best-known constants of motion are connected with symmetry in this way: conservation of linear momentum and angular momentum is a consequence of symmetry under spatial translations and spatial rotatations, respectively. This will be discussed later.