

Remarks on Quantum Mechanics

Norbert Dragon

Intended as completion, repetition and comment once the free particle, the potential well, the potential barrier, the hydrogen atom and scattering are understood. Among the numerous presentations of quantum mechanics I recommend as reliable, but also very demanding book [1].

The article does not have its final form, the most recent version can be found at <http://www.itp.uni-hannover.de/~dragon>. I am grateful for feedback concerning errors, including type slips, or inconceivable arguments.

I am indebted to Ulrich Theis for his help in translating this paper.

This article was typeset on 27th February 2006 using \LaTeX , more precisely $\text{\LaTeX} 2_{\epsilon}$ and the KOMA-Script-class scrbook.

Contents

1	Probabilities of results	1
2	Basic mathematical facts	3
2.1	Orthonormal basis	3
2.2	Bracket notation	4
2.3	Matrix algebra	5
2.4	Projectors, resolution of the identity	6
2.5	Finite norm	6
2.6	Rays in Hilbert space	7
3	Density matrix	9
3.1	Statistical sources	9
3.2	Mixing mixtures	10
4	Operators	11
4.1	Expectation value	11
4.2	Unbounded spectrum	12
4.3	Uncertainty	13
4.4	Commutator	15
4.5	Creation annihilation algebra	16
4.6	Angular momentum algebra	17
4.7	Measurement of a spin-1/2-mixture	20
4.8	Perturbation theory	23
5	Continuous spectrum	27
5.1	Wave function	27
5.2	Transformations of position	28
5.3	Translation and momentum	31
5.4	Rotation and angular momentum	32
5.5	Continuum basis	33
5.6	Multiparticle states	36
6	Time evolution	39
6.1	Schrödinger equation	39
6.2	Schrödinger picture, Heisenberg picture	42
6.3	Groundstate energy	44
6.4	Canonical quantization, normal order	45

6.5	Reduction of state	47
6.6	Time evolution of the two state system	49
6.7	Energy bands	50
7	Composite systems	55
7.1	Product space	55
7.2	Addition of angular momentum	56
7.3	Independent composite systems	57
7.4	Bell Inequalities	58
8	Basics of Thermodynamics	61
8.1	Entropy	61
8.2	Equilibrium	64
9	Decay of an unstable particle	67
9.1	Lorentz resonance	67
9.2	Deviation from exponential decay	68
9.3	Golden Rule	69
9.4	Decay into the continuum	71
9.5	General validity	76
9.6	Decay of moving particles	76
A	$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x+i\epsilon} = \text{PV} \frac{1}{x} - i\pi\delta(x)$	80
B	$\lim_{t \rightarrow \infty} \frac{\sin^2(tx)}{tx^2} = \pi\delta(x)$	81
C	Remark on Fourier transformation	82
D	Derivative of the determinant	83
	Bibliography	84

1 Probabilities of results

Physicists observe, measure and analyze properties of systems which are prepared in such a manner that they are simple enough.

Let us consider for definiteness that the system which is to be measured, the state, is a particle in a beam and that the instrument, which is used to measure the state splits the beam like a Stern-Gerlach device into different partial beams. The results $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ of the measurement are the discrete angles of deflection.

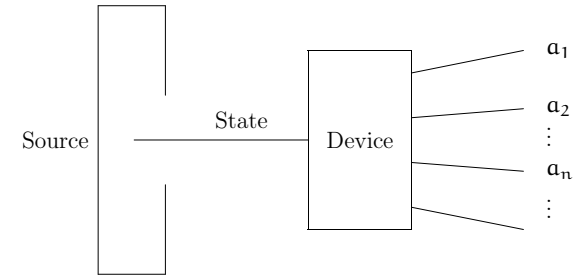


Figure 1.1: Setup of a Measurement

Quantum mechanics accounts for the following experimental observations

1. For all results α_i of an ideal measurement A one can prepare corresponding states Λ_i for which the result α_i is certain.
2. Even if the state Ψ which is to be measured has been prepared ideally then there are always measurements A with results $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ which cannot be predicted with certainty.

and assumes the following **Fundamental equation**:

If the state Ψ is measured with an instrument A then

$$w(i, A, \Psi) = |\langle \Lambda_i | \Psi \rangle|^2 . \tag{1.1}$$

is the probability that the i^{th} result α_i occurs.

For simplicity we suppose that the instrument A distinguishes precisely enough such that to each result α_i there corresponds only one state Λ_i . This state is called eigenstate of A corresponding to the eigenvalue α_i . If there is only one state corresponding to a given result of the measurement then this result is called non-degenerate.

2 Basic mathematical facts

2.1 Orthonormal basis

The formula (1.1) for probabilities has to be read in the following way: states like Λ_i and Ψ correspond to vectors in a Hilbert space \mathcal{H} . A Hilbert space is a vector space which means that for any two vectors Λ and Ψ their sum $\Lambda + \Psi$ and each complex multiple $c\Psi = \Psi c$, $c \in \mathbb{C}$ are vectors in Hilbert space. Moreover there is a scalar product $\langle \Lambda | \Psi \rangle \in \mathbb{C}$ for all pairs of vectors with the following properties

$$\langle \Lambda | \Psi \rangle^* = \langle \Psi | \Lambda \rangle, \quad (2.1)$$

$$\langle \Lambda | c_1 \Psi_1 + c_2 \Psi_2 \rangle = c_1 \langle \Lambda | \Psi_1 \rangle + c_2 \langle \Lambda | \Psi_2 \rangle \quad \forall c_1, c_2 \in \mathbb{C}. \quad (2.2)$$

The scalar product is linear in the second argument and because of (2.1) antilinear in the first argument

$$\langle c_1 \Psi_1 + c_2 \Psi_2 | \Lambda \rangle = c_1^* \langle \Psi_1 | \Lambda \rangle + c_2^* \langle \Psi_2 | \Lambda \rangle. \quad (2.3)$$

The scalar product of a vector with itself is positive definite and is used to define the length of vectors

$$0 \leq \langle \Psi | \Psi \rangle = \|\Psi\|^2 < \infty, \quad \|\Psi\| = 0 \Leftrightarrow \Psi = 0. \quad (2.4)$$

If a state Ψ is measured with an instrument \mathbf{A} , then, according to (1.1), the probability $w(i, \mathbf{A}, \Psi)$ for the i -th result \mathbf{a}_i is the square modulus of the scalar product $\langle \Lambda_i | \Psi \rangle$ of the state Ψ , which is being measured, with the eigenstate Λ_i corresponding to the i -th result. The scalar product $\langle \Lambda_i | \Psi \rangle$ is called the probability amplitude of the corresponding i -th result.

Equation (1.1) implies that the states Λ_i are normalized and mutually orthogonal.

$$\langle \Lambda_i | \Lambda_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (2.5)$$

For if an eigenstate Λ_j is measured then one is certain to obtain the result \mathbf{a}_j , all other results occur with probability 0, the probability distribution is $w(i, \mathbf{A}, \Lambda_j) = \delta_{ij}$. From these probabilities and from (2.4) one derives the probability amplitudes (2.5) of the states Λ_j .

In addition to (1.1) one assumes in quantum mechanics that the eigenstates Λ_i constitute a basis. Each state Ψ can be written as a linear combination of the states Λ_i .

$$\Psi = \sum_j \Lambda_j \psi_j \quad \psi_j \in \mathbb{C} \quad (2.6)$$

Using (2.5) the components ψ_i are obtained as scalar product with Λ_i

$$\psi_i = \langle \Lambda_i | \Psi \rangle . \quad (2.7)$$

The components of Ψ in the basis of the eigenstates of the measuring device are the probability amplitudes of the corresponding results.

If the state Ψ is still unknown, then the modulus of the components in the basis corresponding to the measurement can be taken from the probability distribution of the results. The phases of these components have to be taken from other measurements.

2.2 Bracket notation

If we insert the components into (2.6) we obtain

$$\Psi = \sum_j \Lambda_j \langle \Lambda_j | \Psi \rangle . \quad (2.8)$$

The scalar product with each vector Φ yields the following formula

$$\langle \Phi | \Psi \rangle = \langle \Phi | \left(\sum_j \Lambda_j \langle \Lambda_j | \Psi \rangle \right) \rangle = \sum_j \langle \Phi | \Lambda_j \rangle \langle \Lambda_j | \Psi \rangle . \quad (2.9)$$

Because this equation holds for all Φ , one skips the symbol “ $\langle \Phi$ ” and obtains the suggestive formula

$$|\Psi\rangle = \sum_j |\Lambda_j\rangle \langle \Lambda_j | \Psi \rangle = \sum_j |\Lambda_j\rangle \psi_j . \quad (2.10)$$

If one decomposes Ψ in the scalar product $\langle \Psi | \Phi \rangle$, and uses (2.1) one obtains analogously

$$\langle \Psi | = \sum_j \langle \Psi | \Lambda_j \rangle \langle \Lambda_j | = \sum_j \psi_j^* \langle \Lambda_j | . \quad (2.11)$$

Dirac has introduced the name ket vector for the part $|\Psi\rangle$ in the scalar product and the name bra vector for $\langle \Phi |$. The scalar product is a bracket $\langle \Phi | \Psi \rangle$, composed out of bra vector and ket vector. The admittedly trivial bijective map¹ of vectors to ket vectors $\Psi \rightarrow |\Psi\rangle$ is linear, the map to bra vectors $\Psi \rightarrow \langle \Psi |$ is antilinear:

$$\langle c\Psi | = c^* \langle \Psi | . \quad (2.12)$$

The map of bra vectors to ket vectors is a conjugation $|\Psi\rangle^* = \langle \Psi |$.

¹The map is similar to a military promotion where a chevron and a stripe are added.

2.3 Matrix algebra

Equation (2.9) expresses the scalar product in terms of components.

$$\langle \Phi | \Psi \rangle = \sum_j \Phi_j^* \psi_j \quad (2.13)$$

If one writes the components of a ket vector as a column and the components of a bra vector as row – according to (2.11) they are the complex conjugate components of the ket vector – then the scalar product is obtained by multiplication of the row with the column.

If one applies an operator A to a vector Ψ then one also obtains the components $(A\Psi)_n$ from matrix multiplication.

$$(A\Psi)_n = \langle \Lambda_n | A\Psi \rangle = \sum_m \langle \Lambda_n | A \Lambda_m \rangle \psi_m = \sum_m A_{nm} \psi_m \quad (2.14)$$

The column vector of components of $A\Psi$ is obtained by matrix multiplication of the matrix, which contains the matrix element $A_{nm} = \langle \Lambda_n | A \Lambda_m \rangle$ in the n^{th} row and the m^{th} column, with the column vector of components of Ψ .

The hermitean adjoint operator A^\dagger of a linear operator is defined by

$$\langle \Lambda | A\Psi \rangle = \langle A^\dagger \Lambda | \Psi \rangle \quad \forall \Lambda, \Psi \quad (2.15)$$

Hermitean conjugation reverses the order of factors of a product.

$$\langle \Lambda | AB\Psi \rangle = \langle A^\dagger \Lambda | B\Psi \rangle = \langle B^\dagger A^\dagger \Lambda | \Psi \rangle, \quad (AB)^\dagger = B^\dagger A^\dagger . \quad (2.16)$$

The hermitean adjoint of a complex number is the complex conjugate number. Hermitean adjoint matrices, i.e. the transposed and complex conjugated matrices, correspond to hermitean adjoint operators.

$$(A^\dagger)_{nm} = \langle \Lambda_n | A^\dagger \Lambda_m \rangle = \langle A \Lambda_n | \Lambda_m \rangle = \langle \Lambda_m | A \Lambda_n \rangle^* = A_{mn}^* \quad (2.17)$$

In bracket notation the operators A are written in terms of their matrix elements and the basis as

$$A = \sum_{nm} |\Lambda_n\rangle A_{nm} \langle \Lambda_m| , \quad (2.18)$$

or, more conventionally

$$A : \Psi \rightarrow \sum_{nm} \Lambda_n A_{nm} \langle \Lambda_m | \Psi \rangle . \quad (2.19)$$

2.4 Projectors, resolution of the identity

By (2.10) the operator $\sum_j |\Lambda_j\rangle\langle\Lambda_j|$ maps each vector $|\Psi\rangle$ to itself, the operator therefore is the identity $\mathbb{1}$.

$$\mathbb{1} = \sum_j |\Lambda_j\rangle\langle\Lambda_j|. \quad (2.20)$$

The single parts

$$P_j = |\Lambda_j\rangle\langle\Lambda_j| \quad (2.21)$$

are projectors

$$P_j^2 = P_j \quad (2.22)$$

which project to mutually orthogonal subspaces.

$$P_i P_j = 0 \text{ if } i \neq j. \quad (2.23)$$

The representation (2.20) of the $\mathbb{1}$ -operator as a sum of projection operators is called a resolution of the identity.

By help of a resolution of the identity and in bracket notation a change of basis is a simple algebraic task: If $|\Gamma_i\rangle$ and $|\Lambda_i\rangle$ are two orthonormal bases, then the components of a state Ψ in the two bases are related by an insertion of a resolution of the identity.

$$\langle\Gamma_i|\Psi\rangle = \sum_j \langle\Gamma_i|\Lambda_j\rangle\langle\Lambda_j|\Psi\rangle. \quad (2.24)$$

2.5 Finite norm

Depending on the states which are prepared, the Hilbert space is often not finite dimensional and the infinite sums which express vectors in terms of a basis have to be checked for convergence. In our discussion of quantum mechanics we neglect nearly all complications which are connected with questions of convergence. We only remark that vectors in Hilbert space have a finite scalar product

$$\langle\Psi|\Psi\rangle = \sum_j \psi_j^* \psi_j < \infty \quad (2.25)$$

and that their components ψ_j have to be square summable. Conversely, for given orthonormal basis, each square summable sequence ψ_n , $n = 1, 2, \dots$, defines a vector in Hilbert space.

More restrictively the square modulus of the components of physical states are the probabilities for the corresponding results of a measurement. Probabilities satisfy a sum rule: the sum of probabilities over a complete set of mutually exclusive events gives the overall probability 1.

$$1 = \sum_i w(i, \mathcal{A}, \Psi) = \sum_i |\psi_i|^2 = \langle\Psi|\Psi\rangle \quad (2.26)$$

Each vector in Hilbert space which corresponds to a physical state is normalized.

There is no correspondence which attributes to each vector in Hilbert space a physical state, though one can sometimes hear such statements. In particular the vector 0 in Hilbert space does not correspond to a physical state though it is tempting to mistake it for the ground state which is often denoted $|0\rangle$.

2.6 Rays in Hilbert space

By (2.5) and (2.26) states correspond to vectors on the unit sphere in Hilbert space. From the formula (1.1) for the probability of a result of a measurement one concludes that the unit vector Ψ and the vector which is multiplied by a phase $e^{i\alpha}\Psi$ correspond to the same physical state because for all measuring devices \mathcal{A} the probability distribution of both vectors agree

$$w(i, \mathcal{A}, \Psi) = w(i, \mathcal{A}, e^{i\alpha}\Psi) \quad \forall \alpha \in \mathbb{R}. \quad (2.27)$$

Therefore physical states correspond to equivalence classes of unit vectors with equivalence relation

$$\Psi \sim \Psi' \Leftrightarrow \exists \alpha \in \mathbb{R} : \Psi = e^{i\alpha}\Psi'. \quad (2.28)$$

A more elegant notion than “unit vector up to a phase” is the equivalent concept of a “ray in Hilbert space”. The ray corresponding to a vector $\Psi \neq 0$ is the set of all non-vanishing complex multiples of Ψ . If one attributes to physical states rays in Hilbert space then one has to adjust the formula (1.1) for the probabilities of results α_i such that it becomes independent of the normalization of the vectors.

$$w(i, \mathcal{A}, \Psi) = \frac{|\langle\Lambda_i|\Psi\rangle|^2}{\langle\Lambda_i|\Lambda_i\rangle\langle\Psi|\Psi\rangle} \quad (2.29)$$

For normalized vectors Λ_i and Ψ this formula agrees with (1.1). The probability does not depend on which vector one chooses as representative of its ray.

The linear structure of Hilbert space does not allow to add physical states! Physical states are rays in Hilbert space or, expressed less elegantly, unit vectors up to a phase. Vectors on the unit sphere do not form a linear space, one cannot combine states with a unique prescription which has the mathematical properties of addition and there is no physical correspondence to multiplication of a vector with a complex number.

These remarks are not just a mathematical subtlety. Quantum mechanics is not linear in all respects. The physical state behind a double slit is not the sum of two physical states though it can be written as sum of two parts. These parts are not physical states. The physical state passes both slits. The state which is composed of several parts depends sensitively on the phases of the parts, which in the example of a double slit one can control by the distance of the slits. In short: there is no natural addition of physical states.

Rather than to use a vector Λ_i or Ψ as representative of a ray in Hilbert space one can represent rays by the projectors

$$P_{i,A} = \frac{|\Lambda_i\rangle\langle\Lambda_i|}{\langle\Lambda_i|\Lambda_i\rangle} \quad (2.30)$$

and

$$\rho = \frac{|\Psi\rangle\langle\Psi|}{\langle\Psi|\Psi\rangle} . \quad (2.31)$$

In this notation the probability for the i -th result \mathbf{a}_i is given by the trace of ρ times the projector $P_{i,A}$ to the corresponding subspace.

$$w(i, A, \rho) = \text{tr } \rho P_{i,A} . \quad (2.32)$$

We recall: the trace $\text{tr } A$ of an operator is defined as

$$\text{tr } A = \sum_j \langle\xi_j|A\xi_j\rangle , \quad (2.33)$$

where ξ_j constitute an orthonormal basis (e.g $\xi_j = \Lambda_j$). The trace of an operator does not depend on the basis and is cyclic $\text{tr } AB = \text{tr } BA$.

Equation (2.32) gives the probability distribution also in case of a degenerate result when several states $\Lambda_{a,k}$, $k = 1, 2, \dots$, which can be distinguished by finer instruments and are therefore mutually orthogonal, yield the same result \mathbf{a} . Then the projector $P_{a,A}$ has to be generalized to the operator which projects to the subspace of the states for which the result \mathbf{a} is certain.

$$P_{a,A} = \sum_k \frac{|\Lambda_{a,k}\rangle\langle\Lambda_{a,k}|}{\langle\Lambda_{a,k}|\Lambda_{a,k}\rangle} \quad (2.34)$$

If the state ρ is measured then the probability for the result \mathbf{a} is

$$w(a, A, \rho) = \text{tr } \rho P_{a,A} . \quad (2.35)$$

3 Density matrix

3.1 Statistical sources

The probability (1.1) of a result can be compared to the frequency with which it occurs in repeated measurements only if one prepares repeatedly the same state Ψ . For many sources, in particular if the source is an oven, this is not the case. If the source in figure (1.1) contains a random generator which prepares the state Ψ_1 with probability p_1 , the state Ψ_2 with probability p_2 and so on then the case that the state Ψ_1 is produced and result \mathbf{a}_i is obtained occurs with probability $p_1 w(i, A, \Psi_1)$, $p_2 w(i, A, \Psi_2)$ is the probability that Ψ_2 is produced and \mathbf{a}_i is measured and so on. Considering all possibilities for the occurrence of the i -th result one obtains the probability

$$w(i, A, \rho) = \sum_n p_n w(i, A, \Psi_n) = \sum_n p_n \langle\Lambda_i|\Psi_n\rangle\langle\Psi_n|\Lambda_i\rangle = \langle\Lambda_i|\rho\Lambda_i\rangle . \quad (3.1)$$

The density matrix ρ characterizes the mixture in all measurable properties.

$$\rho = \sum_n p_n |\Psi_n\rangle\langle\Psi_n| \quad (3.2)$$

The probability to obtain the i -th result of a measurement is the corresponding matrix element $\langle\Lambda_i|\rho\Lambda_i\rangle$ on the main diagonal of the density matrix in the basis of the eigenstates of the measuring device.

The probability (3.1) can be compared with the frequency of the result only if the production probabilities p_n are unchanged during the series of measurements and if the mixture can be prepared sufficiently often and if there exist instruments which are not part of the quantum mechanical system which is being measured.

It remains therefore unclear how to interpret a “wavefunction of the universe”. We are mercifully saved from this problem how to interpret the wave function of the universe because we do not know it.

The term mixture applies to the generic case that different states Ψ_n are prepared with probabilities p_n . If in a series of measurements one always prepares the same state Ψ then one calls the system which is to be measured a pure state. Pure states are special mixtures where one production probability is 1 and all the others vanish.

The states Ψ_n which constitute the mixture are normally not mutually orthogonal and normally do not constitute a basis. In the generic case one cannot reconstruct the individual parts $p_n |\Psi_n\rangle\langle\Psi_n|$ from the density matrix ρ similar to a sum which does not allow to tell its terms. One can, however, determine the eigenvalues ρ_n and orthonormalized eigenvectors Υ_n of ρ

$$\rho\Upsilon_n = \rho_n\Upsilon_n \quad \text{with } \langle\Upsilon_m|\Upsilon_n\rangle = \delta^m_n \quad (3.3)$$

and use them to write the density matrix in a confusingly similar form

$$\rho = \sum_{\mathbf{n}} \rho_{\mathbf{n}} |\Upsilon_{\mathbf{n}}\rangle \langle \Upsilon_{\mathbf{n}}| . \quad (3.4)$$

The eigenvalues $\rho_{\mathbf{n}}$ and the projectors to the corresponding eigenspaces are determined by ρ and the eigenvalue equation.

Each main diagonal element $\langle \Lambda | \rho | \Lambda \rangle$ of the density matrix is non-negative

$$\langle \Lambda | \rho | \Lambda \rangle = \sum_{\mathbf{n}} \langle \Lambda | \mathbf{p}_{\mathbf{n}} \Psi_{\mathbf{n}} \rangle \langle \Psi_{\mathbf{n}} | \Lambda \rangle = \sum_{\mathbf{n}} \mathbf{p}_{\mathbf{n}} |\langle \Lambda | \Psi_{\mathbf{n}} \rangle|^2 \geq 0 . \quad (3.5)$$

Therefore all eigenvalues $\rho_{\mathbf{n}}$ of a density matrix are non-negative. A main diagonal element $\langle \Lambda | \rho | \Lambda \rangle$ vanishes if and only if all products $\mathbf{p}_{\mathbf{n}} \langle \Psi_{\mathbf{n}} | \Lambda \rangle$ vanish which means that $\rho | \Lambda \rangle$ vanishes

$$\langle \Lambda | \rho | \Lambda \rangle = 0 \Leftrightarrow \rho | \Lambda \rangle = 0 . \quad (3.6)$$

The trace of the density matrix is fixed by the sum rule for probabilities.

$$\begin{aligned} \text{tr } \rho &= \sum_{\mathbf{i}} \langle \Lambda_{\mathbf{i}} | \rho | \Lambda_{\mathbf{i}} \rangle = \sum_{\mathbf{i}, \mathbf{n}} \langle \Lambda_{\mathbf{i}} | \mathbf{p}_{\mathbf{n}} \Psi_{\mathbf{n}} \rangle \langle \Psi_{\mathbf{n}} | \Lambda_{\mathbf{i}} \rangle = \sum_{\mathbf{n}} \langle \Psi_{\mathbf{n}} | \mathbf{p}_{\mathbf{n}} \Psi_{\mathbf{n}} \rangle = \sum_{\mathbf{n}} \mathbf{p}_{\mathbf{n}} \\ \text{tr } \rho &= 1 = \sum_{\mathbf{n}} \rho_{\mathbf{n}} \end{aligned} \quad (3.7)$$

3.2 Mixing mixtures

Let us imagine two sources in figure (1.1) which prepare mixtures $\hat{\rho}$ and $\tilde{\rho}$ and a device which combines the two beams in a random way such that with probability λ particles are taken from the first beam and in the remaining cases with probability $(1-\lambda)$ particles are taken from the second beam.

If the mixture, which is prepared by mixing two mixtures, is measured then the case that the first beam is chosen and $\mathbf{a}_{\mathbf{i}}$ is measured occurs with probability $\lambda \langle \Lambda_{\mathbf{i}} | \hat{\rho} | \Lambda_{\mathbf{i}} \rangle$, the case that the second beam is chosen and $\mathbf{a}_{\mathbf{i}}$ results has probability $(1-\lambda) \langle \Lambda_{\mathbf{i}} | \tilde{\rho} | \Lambda_{\mathbf{i}} \rangle$.

Altogether the result $\mathbf{a}_{\mathbf{i}}$ is measured with probability

$$\lambda \langle \Lambda_{\mathbf{i}} | \hat{\rho} | \Lambda_{\mathbf{i}} \rangle + (1-\lambda) \langle \Lambda_{\mathbf{i}} | \tilde{\rho} | \Lambda_{\mathbf{i}} \rangle = \langle \Lambda_{\mathbf{i}} | (\lambda \hat{\rho} + (1-\lambda) \tilde{\rho}) | \Lambda_{\mathbf{i}} \rangle . \quad (3.8)$$

Mixing two mixtures with mixing parameter λ yields for all measuring instruments the probability distributions of the density matrix

$$\rho(\lambda) = \lambda \hat{\rho} + (1-\lambda) \tilde{\rho} . \quad (3.9)$$

We will see that mixing increases the entropy, the lack of knowledge about the underlying states, and the uncertainty or standard deviation of each measurement.

4 Operators

4.1 Expectation value

Equation (2.32) specifies the probability distribution for all measurements and contains the complete information on the results of series of measurements. Often one is interested in less information, for example the mean value of the results of the measurement. For many probability distributions the most probable result is near the mean value and the mean value is then the result which one expects. Therefore physicists call the mean value expectation value. However, one should be warned that there are distributions with two or more humps, for example the fanned out beam after a Stern-Gerlach device, where results near the mean value are improbable and where the expectation value cannot be expected.

The mean value $\langle \mathbf{A} \rangle$ of the measured values of an instrument is the sum of the results weighted with their probabilities

$$\langle \mathbf{A} \rangle = \sum_{\mathbf{i}} \mathbf{a}_{\mathbf{i}} w(\mathbf{i}, \mathbf{A}, \rho) = \sum_{\mathbf{i}} \mathbf{a}_{\mathbf{i}} \langle \Lambda_{\mathbf{i}} | \rho | \Lambda_{\mathbf{i}} \rangle = \text{tr} \left(\sum_{\mathbf{i}} \mathbf{a}_{\mathbf{i}} | \Lambda_{\mathbf{i}} \rangle \langle \Lambda_{\mathbf{i}} | \right) \rho . \quad (4.1)$$

The mean value is therefore

$$\langle \mathbf{A} \rangle = \text{tr } \rho \mathbf{A} , \quad (4.2)$$

where \mathbf{A} does not only denote the instrument but also the operator

$$\mathbf{A} = \sum_{\mathbf{i}} \mathbf{a}_{\mathbf{i}} | \Lambda_{\mathbf{i}} \rangle \langle \Lambda_{\mathbf{i}} | . \quad (4.3)$$

The operator \mathbf{A} is characteristic for the instrument. The possible results and their probabilities for a given mixture ρ can be calculated from the operator. Succinctly, to each measuring instrument there corresponds an operator in Hilbert space.

It is, however, sobering, that producers of measuring instruments do not include the corresponding operator with their instructions of use.

Contrary to widespread hearsay the application of an operator to the physical state Ψ does not correspond to the act of measuring.

The notation $\langle \mathbf{A} \rangle$ for the mean value $\text{tr } \rho \mathbf{A}$ originates from the pure state. In that case, if one employs a normalized Ψ , one has more specifically

$$\langle \mathbf{A} \rangle = \langle \Psi | \mathbf{A} | \Psi \rangle . \quad (4.4)$$

Without changing anything one writes in bracket notation an additional “|” and stresses with the notation $\langle \mathbf{A} \rangle = \langle \Psi | \mathbf{A} | \Psi \rangle$ that it is irrelevant whether the operator \mathbf{A} acts on

the first or second argument of the scalar product. For \mathbf{A} is a linear, hermitean operator (2.15)

$$\mathbf{A} = \mathbf{A}^\dagger . \quad (4.5)$$

One easily confirms that the projection operators (2.21) are hermitean $\langle \Phi | \mathbf{P}_i | \Psi \rangle = \langle \Phi | \Lambda_i \rangle \langle \Lambda_i | \Psi \rangle = \langle \mathbf{P}_i \Phi | \Psi \rangle$ and that real linear combinations (4.3) of hermitean operators are hermitean.

By the same reason the density matrix ρ (3.2) is hermitean.

$$\rho = \rho^\dagger \quad (4.6)$$

From (2.5) one concludes immediately that the states Λ_i are eigenstates of the operator \mathbf{A} and that the eigenvalues are the possible results \mathbf{a}_i of the measurement.

$$\mathbf{A} \Lambda_i = \mathbf{a}_i \Lambda_i \quad (4.7)$$

This is how we have constructed the operator \mathbf{A} (4.3) from eigenstates and corresponding results.

Conversely, the results \mathbf{a}_i and the corresponding eigenvectors Λ_i up to a complex factor, i.e. the corresponding ray in Hilbert space, can be determined from the given operator \mathbf{A} as solutions to the eigenvalue equation.

The eigenvalues \mathbf{a} of an hermitean operator $\mathbf{A} = \mathbf{A}^\dagger$ are real. This can be seen from $\mathbf{A} \Lambda = \mathbf{a} \Lambda$ and $\langle \Lambda | \Lambda \rangle \neq 0$ and the following chain of arguments.

$$(\mathbf{a}^* - \mathbf{a}) \langle \Lambda | \Lambda \rangle = \langle \mathbf{a} \Lambda | \Lambda \rangle - \langle \Lambda | \mathbf{a} \Lambda \rangle = \langle \mathbf{A} \Lambda | \Lambda \rangle - \langle \Lambda | \mathbf{A} \Lambda \rangle = 0 \quad (4.8)$$

Eigenstates corresponding to different eigenvalues are mutually orthogonal.

$$(\mathbf{a}_i - \mathbf{a}_j) \langle \Lambda_i | \Lambda_j \rangle = \langle \mathbf{A} \Lambda_i | \Lambda_j \rangle - \langle \Lambda_i | \mathbf{A} \Lambda_j \rangle = 0, \quad \mathbf{a}_i \neq \mathbf{a}_j \Rightarrow \langle \Lambda_i | \Lambda_j \rangle = 0 \quad (4.9)$$

Unitary operators $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ leave invariant all scalar products.

$$\langle \mathbf{U} \Phi | \mathbf{U} \Psi \rangle = \langle \mathbf{U}^\dagger \mathbf{U} \Phi | \Psi \rangle = \langle \Phi | \Psi \rangle \quad (4.10)$$

Therefore Ψ and $\mathbf{U} \Psi$ have the same length and $\mathbf{U} \Psi = \lambda \Psi$, $\lambda \in \mathbb{C}$, is only possible with $|\lambda| = 1$. This is why eigenvalues of unitary operators have modulus 1.

$$\mathbf{U}^\dagger = \mathbf{U}^{-1} \text{ and } \mathbf{U} \Psi = \lambda \Psi \Rightarrow \lambda = e^{i\alpha}, \quad \alpha \in \mathbb{R} . \quad (4.11)$$

Each unitary operator \mathbf{U} can be written as $e^{i\mathbf{H}}$ with some hermitean operator $\mathbf{H} = \mathbf{H}^\dagger$.

4.2 Unbounded spectrum

The set of all eigenvalues of an operator – or to be more precise, the complement of the set of complex numbers $\lambda \in \mathbb{C}$ for which the resolvent $(\mathbf{A} - \lambda)^{-1}$ exists as operator in the whole Hilbert space – is called spectrum of \mathbf{A} . If the spectrum is not bounded, then

the operator is not defined on all vectors Ψ and a arbitrarily small change of a state, on which \mathbf{A} is defined, can give an arbitrary large change of the expectation value.

From every day life such a discontinuous behaviour of mean values of probability distributions of unbounded results is well known. One single student in his 40th term spoils the average. In statistics one masters this problem with additional arguments such as “A student in his 40th term is no longer a student” and just skips his data. Analyzing measured values one often proceeds similarly and drops runaway values from the determination of mean values.

In distinguished phrases this procedure is called regularization. If one intends to analyze sufficiently well behaved problems such as “Do students understand the subject faster after the introduction of the new curriculum?” then the old student is unimportant anyhow and the regularization is acceptable.

The mathematical difficulties with operators with unbounded spectrum show themselves for example with the expectation value of the energy of the harmonic oscillator. The energies are the eigenvalues of the Hamilton operator $\mathbf{H} = \hbar \omega \mathbf{a}^\dagger \mathbf{a}$. They are non-negative, integer multiples of $\hbar \omega$

$$\mathbf{H} | \Lambda_n \rangle = E_n | \Lambda_n \rangle, \quad E_n = n \hbar \omega, \quad n = 0, 1, 2, \dots . \quad (4.12)$$

We assume that the eigenvectors Λ_n are normalized. Then they constitute an orthonormal basis (2.5) and a general vector can be written as linear combination $|\Psi\rangle = \sum_n | \Lambda_n \rangle \psi_n$ with square summable components ψ_n .

The Hamilton operator maps Ψ to $\mathbf{H} \Psi$ with components

$$\langle \Lambda_n | \mathbf{H} \Psi \rangle = \hbar \omega n \langle \Lambda_n | \Psi \rangle = \hbar \omega n \psi_n . \quad (4.13)$$

One can easily specify square summable sequences, e. g. $\psi_n = 1/(n+1)$, such that the sequence $n \psi_n$ is not square summable. The operator \mathbf{H} is not defined on the corresponding vectors. If one takes from this sequence only terms until some large \mathbf{N} and adds it with a small coefficient to a physical state, one recognizes that in each neighbourhood of each state there exists another state whose energy expectation value surpasses each given bound. This is a mathematical nuisance but unimportant for physics: the large expectation value results from very rare results with high energy.

Better behaved than operators with unbounded spectrum are projectors (2.21) to the eigenspaces corresponding to the results of the measurement. If one measures a mixture ρ then only these projectors are needed to calculate the probability distribution of the results.

4.3 Uncertainty

Next in importance to the mean value is the deviation $\Delta_\rho \mathbf{A}$, which characterizes the spread of the results if a mixture ρ (3.2) is measured with an instrument \mathbf{A} . More precisely the square of the deviation is

$$(\Delta_\rho \mathbf{A})^2 = \langle (\mathbf{A} - \langle \mathbf{A} \rangle)^2 \rangle = \sum_n p_n \langle \Psi_n | (\mathbf{A} - \langle \mathbf{A} \rangle)^2 | \Psi_n \rangle = \langle \mathbf{A}^2 \rangle - \langle \mathbf{A} \rangle^2 \quad (4.14)$$

$\Delta_\rho \mathbf{A}$ is called deviation or uncertainty of \mathbf{A} in the mixture ρ . The uncertainty depends on both the hermitean operator \mathbf{A} and the mixture ρ .

The quantity $(\Delta_\rho \mathbf{A})^2$ is non-negative, for (4.14) is a sum of squares $\|(\mathbf{A} - \langle \mathbf{A} \rangle) \Psi_n\|^2$ weighted with non-negative probabilities.

$$\langle \Psi | (\mathbf{A} - \langle \mathbf{A} \rangle)^2 \Psi \rangle = \langle (\mathbf{A} - \langle \mathbf{A} \rangle) \Psi | (\mathbf{A} - \langle \mathbf{A} \rangle) \Psi \rangle = \|(\mathbf{A} - \langle \mathbf{A} \rangle) \Psi\|^2 \quad (4.15)$$

The uncertainty vanishes if and only if the mixture is mixed from eigenstates Λ_n with the same eigenvalue $\mathbf{a} = \langle \mathbf{A} \rangle$

$$0 = \sum_n p_n \|(\mathbf{A} - \langle \mathbf{A} \rangle) \Lambda_n\|^2 \Leftrightarrow (\mathbf{A} - \mathbf{a}) \Lambda_n = 0 \text{ or } p_n = 0. \quad (4.16)$$

The sum $\sum_n p_n \|(\mathbf{c}_A (\mathbf{A} - \langle \mathbf{A} \rangle) + i \mathbf{c}_B (\mathbf{B} - \langle \mathbf{B} \rangle)) \Psi_n\|^2$ is non-negative. If one inspects hermitean operators \mathbf{A} and \mathbf{B} and real numbers \mathbf{c}_A and \mathbf{c}_B then this observation yields a general, lower bound for the product $\Delta_\rho \mathbf{A} \Delta_\rho \mathbf{B}$ of the uncertainties of \mathbf{A} and \mathbf{B} in the mixture ρ . With the notation

$$[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} \quad (4.17)$$

for the commutator of \mathbf{A} and \mathbf{B} one has

$$\begin{aligned} 0 &\leq \sum_n p_n \|(\mathbf{c}_A (\mathbf{A} - \langle \mathbf{A} \rangle) + i \mathbf{c}_B (\mathbf{B} - \langle \mathbf{B} \rangle)) \Psi_n\|^2 \\ &= \sum_n p_n \langle \Psi_n | (\mathbf{c}_A^2 (\mathbf{A} - \langle \mathbf{A} \rangle)^2 + \mathbf{c}_B^2 (\mathbf{B} - \langle \mathbf{B} \rangle)^2 + i \mathbf{c}_A \mathbf{c}_B [\mathbf{A}, \mathbf{B}]) \Psi_n \rangle \\ &= (\mathbf{c}_A \Delta_\rho \mathbf{A} + \mathbf{c}_B \Delta_\rho \mathbf{B})^2 - \mathbf{c}_A \mathbf{c}_B (2 \Delta_\rho \mathbf{A} \Delta_\rho \mathbf{B} - i \sum_n p_n \langle \Psi_n | [\mathbf{A}, \mathbf{B}] \Psi_n \rangle). \end{aligned} \quad (4.18)$$

We exploit this inequality for $\mathbf{c}_A = -\Delta_\rho \mathbf{B}$ and $\mathbf{c}_B = \Delta_\rho \mathbf{A}$. Then the first term vanishes. If neither $\Delta \mathbf{A}$ nor $\Delta \mathbf{B}$ vanish then $-2 \mathbf{c}_A \mathbf{c}_B > 0$ and we obtain the general uncertainty relation

$$\Delta_\rho \mathbf{A} \Delta_\rho \mathbf{B} \geq \frac{1}{2} |[\mathbf{A}, \mathbf{B}]|. \quad (4.19)$$

This relation holds for the modulus of $\langle [\mathbf{A}, \mathbf{B}] \rangle$ because we can repeat our considerations with \mathbf{B} and \mathbf{A} exchanged. Thereby the left side is unchanged and the commutator $[\mathbf{A}, \mathbf{B}]$ changes sign. The inequality also holds for vanishing uncertainty $\Delta_\rho \mathbf{A} = 0$ (or $\Delta_\rho \mathbf{B} = 0$) for then the mixture ρ consists of eigenstates to \mathbf{A} (or \mathbf{B}) with eigenvalue \mathbf{a} . The expectation value of a commutator $[\mathbf{A}, \mathbf{B}]$ vanishes in each eigenstate of \mathbf{A} or \mathbf{B} .

$$\langle [\mathbf{A}, \mathbf{B}] \rangle = \sum_n p_n \langle \Psi_n | [\mathbf{A}, \mathbf{B}] \Psi_n \rangle = \sum_n p_n \langle \Psi_n | (\mathbf{a}\mathbf{B} - \mathbf{B}\mathbf{a}) \Psi_n \rangle = 0 \quad (4.20)$$

Mixing does not decrease the deviation. The square uncertainty of a mixture of mixtures

$$(\Delta_{\rho(\lambda)} \mathbf{A})^2 = \text{tr } \rho(\lambda) \mathbf{A}^2 - (\text{tr } \rho(\lambda) \mathbf{A})^2 \quad \text{with } \rho(\lambda) = \lambda \hat{\rho} + (1 - \lambda) \bar{\rho} \quad (4.21)$$

is a polynomial in λ with non-positive second derivative $-2(\text{tr } \hat{\rho} \mathbf{A} - \text{tr } \bar{\rho} \mathbf{A})^2$ and therefore a convex function of the mixing parameter

$$(\Delta_{\rho(\lambda)} \mathbf{A})^2 \geq \lambda (\Delta_{\hat{\rho}} \mathbf{A})^2 + (1 - \lambda) (\Delta_{\bar{\rho}} \mathbf{A})^2. \quad (4.22)$$

The square uncertainty of a mixture of mixtures is at least the proportionate sum of the square uncertainties and coincides with the proportionate sum only if the mean values $\text{tr } \hat{\rho} \mathbf{A}$ and $\text{tr } \bar{\rho} \mathbf{A}$ agree.

4.4 Commutator

In spite of the mathematical complication which originate from unbounded operators one prefers to specify properties of quantum mechanical systems in terms of operators.

The simplest algebraic relation is that one can exchange the order of two operators $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$. ‘‘The operators commute’’ or, in other words, the commutator

$$[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} \quad (4.23)$$

vanishes. If \mathbf{A} and \mathbf{B} commute and if they are diagonalizable, e. g. because they are hermitean or unitary, then the eigenvectors of \mathbf{A} can be chosen to be also eigenvectors of \mathbf{B} and vice versa. For \mathbf{B} maps the eigenspace \mathcal{H}_i of \mathbf{A} with eigenvalue \mathbf{a}_i to \mathcal{H}_i

$$[\mathbf{A}, \mathbf{B}] = 0 \wedge (\mathbf{A} - \mathbf{a}_i) \Lambda_i = 0 \Rightarrow (\mathbf{A} - \mathbf{a}_i) (\mathbf{B} \Lambda_i) = \mathbf{B} (\mathbf{A} - \mathbf{a}_i) \Lambda_i = 0 \quad (4.24)$$

and can be diagonalized in this subspace. If the dimension \mathbf{d}_i of \mathcal{H}_i is larger than 1, then \mathbf{a}_i is degenerate and there are linearly independent eigenvectors Λ_{ij} for \mathbf{A} and \mathbf{B}

$$\mathbf{A} \Lambda_{ij} = \mathbf{a}_i \Lambda_{ij} \quad \mathbf{B} \Lambda_{ij} = \mathbf{b}_{ij} \Lambda_{ij} \quad j = 1, \dots, \mathbf{d}_i. \quad (4.25)$$

One can then construct a finer instrument which measures in one measurement \mathbf{A} and \mathbf{B} and decomposes the beams \mathbf{a}_i in figure (1.1) into finer beams \mathbf{b}_{ij} .

If \mathbf{B} is degenerate in the same subspaces as \mathbf{A} then $\mathbf{B} = f(\mathbf{A})$. In this case \mathbf{B} is not essentially different from \mathbf{A} . It only uses a different scale similar to a Volt- and Ampère-meter.

With respect to a set of measurements, which correspond to mutually commuting operators quantum mechanical systems behave like classical statistical systems. All states Ψ are characterized with respect to these measurements completely by classical probability distributions given by the square modulus of the scalar products $\langle \Lambda | \Psi \rangle$ with the eigenstates Λ of the commuting operators. Only measurements corresponding to non-commuting operators become sensitive to the phases of the complex components of Ψ .

The main algebraic properties of the commutator are antisymmetry, linearity and the product rule

$$[\mathbf{A}, \mathbf{B}] = -[\mathbf{B}, \mathbf{A}], \quad (4.26)$$

$$[\mathbf{A}, \lambda_1 \mathbf{B} + \lambda_2 \mathbf{C}] = \lambda_1 [\mathbf{A}, \mathbf{B}] + \lambda_2 [\mathbf{A}, \mathbf{C}] \quad \forall \lambda_1, \lambda_2 \in \mathbb{C}, \quad (4.27)$$

$$[\mathbf{A}, \mathbf{B}\mathbf{C}] = [\mathbf{A}, \mathbf{B}]\mathbf{C} + \mathbf{B}[\mathbf{A}, \mathbf{C}]. \quad (4.28)$$

Taking the hermitean adjoint reverses the order (2.16) and therefore and because the commutator is antisymmetric the commutator of hermitean operators is antihermitean.

Because of the product rule and because of linearity the operation “taking the commutator with an operator” behaves like a derivative. This derivative does not change the order of factors.

The product rule implies the Jacobi identity

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] = [[\mathbf{A}, \mathbf{B}], \mathbf{C}] + [\mathbf{B}, [\mathbf{A}, \mathbf{C}]] \quad (4.29)$$

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = 0. \quad (4.30)$$

4.5 Creation annihilation algebra

Algebraic relations generate structure in the Hilbert space of states and restrict this space. For example the Heisenberg commutation relation

$$[\mathbf{X}, \mathbf{P}] = i\hbar \quad (4.31)$$

of a hermitean position operator \mathbf{X} with the corresponding hermitean momentum operator \mathbf{P} cannot hold in a Hilbert space \mathcal{H} with finite dimension n . In a finite dimensional space one would calculate $\text{tr}(\mathbf{X}\mathbf{P} - \mathbf{P}\mathbf{X}) = 0$ in contradiction to $\text{tr}(i\hbar) = ni\hbar$. In an infinite dimensional space the trace cannot be defined on all operators.

If for a real number $x_0 \in \mathbb{R}$, $x_0 \neq 0$ there exist the complex linear combinations

$$\mathbf{a} = \frac{1}{\sqrt{2}}\left(\frac{\mathbf{X}}{x_0} + \frac{i}{\hbar}x_0\mathbf{P}\right), \quad \mathbf{a}^\dagger = \frac{1}{\sqrt{2}}\left(\frac{\mathbf{X}}{x_0} - \frac{i}{\hbar}x_0\mathbf{P}\right) \quad (4.32)$$

of the position operator and the momentum operator, then these combinations satisfy the creation annihilation algebra

$$[\mathbf{a}, \mathbf{a}] = 0, \quad [\mathbf{a}^\dagger, \mathbf{a}^\dagger] = 0, \quad [\mathbf{a}, \mathbf{a}^\dagger] = 1, \quad (4.33)$$

and there exists an orthonormal basis

$$\Lambda_{n,\tau} = \frac{(\mathbf{a}^\dagger)^n}{\sqrt{n!}}\Lambda_{0,\tau}, \quad n = 0, 1, 2, \dots, \quad (4.34)$$

on which the operator \mathbf{a} acts as annihilation operator and \mathbf{a}^\dagger as creation operator

$$\mathbf{a}\Lambda_{n,\tau} = \sqrt{n}\Lambda_{n-1,\tau}, \quad \mathbf{a}^\dagger\Lambda_{n,\tau} = \sqrt{n+1}\Lambda_{n+1,\tau}. \quad (4.35)$$

This results from the following analysis of the hermitean operator $\mathbf{a}^\dagger\mathbf{a}$. In anticipation of later results we call $\mathbf{a}^\dagger\mathbf{a}$ the number operator and denote its eigenvalues with n .

$$\mathbf{a}^\dagger\mathbf{a}\Lambda_n = n\Lambda_n. \quad (4.36)$$

From the creation annihilation algebra (4.33) one concludes that the commutator with the number operator $\mathbf{a}^\dagger\mathbf{a}$ maps the operators \mathbf{a} and \mathbf{a}^\dagger to multiples of themselves.

$$[\mathbf{a}^\dagger\mathbf{a}, \mathbf{a}] = -\mathbf{a}, \quad [\mathbf{a}^\dagger\mathbf{a}, \mathbf{a}^\dagger] = \mathbf{a}^\dagger. \quad (4.37)$$

Therefore the states $\mathbf{a}\Lambda_n$ and $\mathbf{a}^\dagger\Lambda_n$ either vanish or are eigenstates of the number operator with eigenvalues $n-1$ and $n+1$ respectively.

$$\mathbf{a}^\dagger\mathbf{a}\mathbf{a}\Lambda_n = ([\mathbf{a}^\dagger\mathbf{a}, \mathbf{a}] + \mathbf{a}\mathbf{a}^\dagger\mathbf{a})\Lambda_n = (-1+n)\mathbf{a}\Lambda_n \quad (4.38)$$

$$\mathbf{a}^\dagger\mathbf{a}\mathbf{a}^\dagger\Lambda_n = ([\mathbf{a}^\dagger\mathbf{a}, \mathbf{a}^\dagger] + \mathbf{a}^\dagger\mathbf{a}^\dagger\mathbf{a})\Lambda_n = (1+n)\mathbf{a}^\dagger\Lambda_n \quad (4.39)$$

Because the operator \mathbf{a} lowers the eigenvalue of the number operator it is called annihilation operator. Analogously \mathbf{a}^\dagger is the creation operator. For a normalized state Λ_n one can calculate the norm of $\mathbf{a}\Lambda_n$ and $\mathbf{a}^\dagger\Lambda_n$ from the algebra and the eigenvalue equation

$$\langle \mathbf{a}\Lambda_n | \mathbf{a}\Lambda_n \rangle = \langle \Lambda_n | \mathbf{a}^\dagger\mathbf{a}\Lambda_n \rangle = n \quad (4.40)$$

$$\langle \mathbf{a}^\dagger\Lambda_n | \mathbf{a}^\dagger\Lambda_n \rangle = \langle \Lambda_n | \mathbf{a}\mathbf{a}^\dagger\Lambda_n \rangle = \langle \Lambda_n | ([\mathbf{a}, \mathbf{a}^\dagger] + \mathbf{a}^\dagger\mathbf{a})\Lambda_n \rangle = n+1. \quad (4.41)$$

All these norms are non-negative (2.4). Therefore n is non-negative. However, repeated application of the annihilation operator \mathbf{a} decreases the eigenvalue of $\mathbf{a}^\dagger\mathbf{a}$ in integer steps and, before n becomes negative, has to lead to a state $\Lambda_0 \neq 0$ which by further application of \mathbf{a} is mapped to 0.

$$\mathbf{a}\Lambda_0 = 0. \quad (4.42)$$

Such a state is called ground state. According to (4.40) it has number eigenvalue $n=0$. Therefore each eigenvalue n is integer and non-negative. The spectrum of the number operator $\mathbf{a}^\dagger\mathbf{a}$ consists of the integer and non-negative numbers.

$$\mathbf{a}^\dagger\mathbf{a}\Lambda_n = n\Lambda_n, \quad n = 0, 1, 2, \dots. \quad (4.43)$$

In the space of all ground states one chooses an orthonormal basis $\Lambda_{0,\tau}$ and considers the vectors (4.34), which are generated by n -fold application of the creation operator \mathbf{a}^\dagger from the ground state $\Lambda_{0,\tau}$. Up to a factor these states coincide with the states from which the ground states had been generated by n -fold application of \mathbf{a}

$$\begin{aligned} (\mathbf{a}^\dagger)^n\mathbf{a}^n\Lambda_n &= (\mathbf{a}^\dagger)^{n-1}(\mathbf{a}^\dagger\mathbf{a})\mathbf{a}^{n-1}\Lambda_n = (\mathbf{a}^\dagger)^{n-1}(n-(n-1))\mathbf{a}^{n-1}\Lambda_n = \\ &= 1 \cdot (\mathbf{a}^\dagger)^{n-2}(\mathbf{a}^\dagger\mathbf{a})\mathbf{a}^{n-2}\Lambda_n = 1 \cdot 2 \cdot (\mathbf{a}^\dagger)^{n-3}(\mathbf{a}^\dagger\mathbf{a})\mathbf{a}^{n-3}\Lambda_n = \dots = n!\Lambda_n. \end{aligned} \quad (4.44)$$

Therefore the ground states are degenerate in exactly the same way as all other eigenstates of the number operator.

4.6 Angular momentum algebra

Another example for algebraic relations is the algebra of angular momentum

$$[\mathbf{L}_i, \mathbf{L}_j] = i\hbar\epsilon_{ijk}\mathbf{L}_k, \quad i, j, k \in \{1, 2, 3\}. \quad (4.45)$$

A vector space endowed with a bilinear product $[\mathbf{A}, \mathbf{B}]$, which is antisymmetric and satisfies the Jacobi identity (4.30), is called a Lie algebra. The angular momentum operators therefore are a basis of a Lie algebra, more precisely of the Lie algebra corresponding to the group $\text{SO}(3)$ of rotations in three dimensions.

Provided the hermitean angular momentum operators L_i exist, the Hilbert space \mathcal{H} has an orthonormal basis $\Lambda_{l,m,\tau}$. τ is an index of degeneracy. $2l$ is non-negative and integral, thus l can take the values $0, 1/2, 1, 3/2, \dots$. The algebra does not imply which one of the allowed values of l occurs and how they are degenerated. For fixed l , m takes the values $-l, -l+1, \dots, +l$. On the orthonormal basis $\Lambda_{l,m,\tau}$ the angular momentum operators can be given explicitly.

$$L_3 \Lambda_{l,m,\tau} = \hbar m \Lambda_{l,m,\tau} \quad (4.46)$$

$$(L_1 + iL_2) \Lambda_{l,m,\tau} = \hbar \sqrt{l(l+1) - m(m+1)} \Lambda_{l,m+1,\tau} \quad (4.47)$$

$$(L_1 - iL_2) \Lambda_{l,m,\tau} = \hbar \sqrt{l(l+1) - m(m-1)} \Lambda_{l,m-1,\tau} \quad (4.48)$$

This can be deduced from the angular momentum algebra in the following way. One verifies that the total angular momentum $L^2 = (L_1^2 + L_2^2 + L_3^2)$ commutes with each of the angular momentum operators L_1 , L_2 and L_3 , for angular momentum operators generate rotations and leave invariant the square of the length of vectors such as $x^2 + y^2 + z^2$ or L^2 .

$$[L_i, L^2] = 0 \quad (4.49)$$

The angular momentum operators therefore (4.24) map angular momentum multiplets, i.e. subspaces \mathcal{H}_l of eigenstates of L^2 with eigenvalue $\hbar^2 l(l+1)$, to themselves. Hence, one can find common eigenstates Λ_{lm} of L^2 and L_3

$$L^2 \Lambda_{lm} = \hbar^2 l(l+1) \Lambda_{lm}, \quad L_3 \Lambda_{lm} = \hbar m \Lambda_{lm}. \quad (4.50)$$

The denomination of the eigenvalue of L^2 is chosen in anticipation of later results.

The angular momentum algebra (4.45) implies that the commutator with L_3 maps the complex linear combinations L_+ and $L_- = (L_+)^{\dagger}$

$$L_+ = L_1 + iL_2, \quad L_- = L_1 - iL_2 \quad (4.51)$$

to a multiple of L_+ and L_- , respectively, and that their commutator gives $2\hbar L_3$

$$[L_3, L_+] = +\hbar L_+, \quad [L_3, L_-] = -\hbar L_-, \quad [L_+, L_-] = 2\hbar L_3. \quad (4.52)$$

Therefore $L_+ \Lambda_{lm}$ and $L_- \Lambda_{lm}$ are either zero or eigenstates of L_3 with eigenvalue $\hbar(m+1)$ and $\hbar(m-1)$, respectively

$$L_3 L_+ \Lambda_{lm} = ([L_3, L_+] + L_+ L_3) \Lambda_{lm} = \hbar(1+m) L_+ \Lambda_{lm} \quad (4.53)$$

$$L_3 L_- \Lambda_{lm} = ([L_3, L_-] + L_- L_3) \Lambda_{lm} = \hbar(-1+m) L_- \Lambda_{lm}. \quad (4.54)$$

Since L_+ and L_- raise and lower the eigenvalues of L_3 by a constant amount, they are also called ladder operators or raising and lowering operators. If Λ_{lm} has unit length, the norm of $L_+ \Lambda_{lm}$ and $L_- \Lambda_{lm}$ follows from

$$L^2 = L_+ L_- + L_3^2 - \hbar L_3 = L_- L_+ + L_3^2 + \hbar L_3. \quad (4.55)$$

$$\begin{aligned} \langle L_+ \Lambda_{lm} | L_+ \Lambda_{lm} \rangle &= \langle \Lambda_{lm} | L_- L_+ \Lambda_{lm} \rangle = \langle \Lambda_{lm} | (L^2 - L_3^2 - \hbar L_3) \Lambda_{lm} \rangle \\ &= \langle \Lambda_{lm} | \hbar^2 (l(l+1) - m(m+1)) \Lambda_{lm} \rangle = \hbar^2 (l(l+1) - m(m+1)) \end{aligned} \quad (4.56)$$

$$\langle L_- \Lambda_{lm} | L_- \Lambda_{lm} \rangle = \hbar^2 (l(l+1) - m(m-1)). \quad (4.57)$$

These norms are non-negative (2.4), so for given l the quantum number m is bounded from above and below.

Applied to the eigenstate $\Lambda_{l,m_{\max}}$ with maximal eigenvalue of L_3 , L_+ must vanish, and thus $l(l+1) - m_{\max}(m_{\max}+1) = 0$. Similarly, $l(l+1) - m_{\min}(m_{\min}-1)$ must be zero. Hence, the quadratic equations

$$(m_{\max} + \frac{1}{2})^2 = (m_{\min} - \frac{1}{2})^2 = l(l+1) + \frac{1}{4} = (l + \frac{1}{2})^2. \quad (4.58)$$

hold. We choose $l + \frac{1}{2}$ to be positive. Because of $m_{\max} + \frac{1}{2} > m_{\min} - \frac{1}{2}$ the unique solution is

$$m_{\max} = l, \quad m_{\min} = -l. \quad (4.59)$$

Since repeated application of L_+ to the state with minimal eigenvalue of L_3 increases the quantum number m in integral steps until one obtains the state with $m_{\max} = l$, the difference $m_{\max} - m_{\min} = 2l$ must be integral and non-negative. Thus $l \in \{0, \frac{1}{2}, 1, \dots\}$ is integral or half-integral. The angular momentum operators act in a $2l+1$ -dimensional space, the angular momentum multiplet with total angular momentum l , which is spanned by the basis states Λ_{lm} with $m = -l, -l+1, \dots, +l$.

For $l = 1/2$ the angular momentum operators, the spin-1/2 operators S_1, S_2 , and S_3 , act in a two-dimensional space on the spinors of the spin-1/2 multiplet with basis states $\Lambda_{l,m}$ with $l = 1/2$ and $m = \pm 1/2$. Because of (4.46), (4.47) and (4.48), the spin operators are given in this basis by $\hbar/2$ times the Pauli matrices σ_1, σ_2 and σ_3

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (4.60)$$

$$S_i = \frac{\hbar}{2} \sigma_i, \quad i \in \{1, 2, 3\}. \quad (4.61)$$

The Pauli matrices satisfy the algebraic relations

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i \epsilon_{ijk} \sigma_k. \quad (4.62)$$

Angular momentum operators generate rotations of states. If the axis of rotation points into the direction of the unity vector \vec{e} , then to a rotation around an angle α there corresponds the unitary operator

$$U(\vec{e}, \alpha) = \exp\left(-\frac{i\alpha}{\hbar} \vec{L} \cdot \vec{e}\right). \quad (4.63)$$

For spin-1/2 the matrix corresponding to (4.63) can be easily given, since due to $(i\vec{\sigma} \cdot \vec{e})^2 = -\mathbb{1}$ the exponential series can be summed similarly to the Euler formula $e^{-i\alpha} = \cos \alpha - i \sin \alpha$.

$$U_{1/2}(\vec{e}, \alpha) = \exp\left(-\frac{i\alpha}{2} \vec{\sigma} \cdot \vec{e}\right) = \mathbb{1} \cos \alpha/2 - i \vec{e} \cdot \vec{\sigma} \sin \alpha/2 \quad (4.64)$$

A rotation of a spin-1/2 spinor around $\alpha = 2\pi$ results in its negative $U_{1/2}(\vec{e}, 2\pi) = -1$.

4.7 Measurement of a spin-1/2-mixture

For a system with two basis states Λ_1, Λ_2 all measurements of all states can be characterized by a few parameters. A prominent example for systems with two states are spin-1/2 particles, which are examined with Stern-Gerlach devices, another example are atoms if due to the experimental setup only two of the energy states need to be considered.

In a Stern-Gerlach device a beam of spin-1/2 particles is split into two separate beams ($n = 2$) as shown in picture (1.1). The intensity of the upper beam divided by the intensity of the incoming beam is the probability with which the spin points up into the direction into which the Stern-Gerlach device splits the beams. For given beam ρ we would like to know how the intensity depends on the direction of the device.

The density matrix ρ , which characterizes the mixture in the two state system, is given by a hermitean matrix, once an orthonormal basis has been chosen, with trace 1. Hermitean $n \times n$ matrices constitute a real n^2 -dimensional vector space. This means in our case that one can write each hermitean 2×2 -matrix as real linear combination of 4 basis matrices. We choose as basis the $\mathbb{1}$ -matrix and the traceless, hermitean Pauli-Matrices (4.60). Because of the normalization condition $\text{tr } \rho = 1$ the coefficient of the $\mathbb{1}$ -matrix is $1/2$ and the most general density matrix of a two state system has the form

$$\rho = \frac{1}{2} \mathbb{1} + \mathbf{a}\sigma_1 + \mathbf{b}\sigma_2 + \mathbf{c}\sigma_3 = \begin{pmatrix} 1/2 + c & \mathbf{a} - i\mathbf{b} \\ \mathbf{a} + i\mathbf{b} & 1/2 - c \end{pmatrix} \quad \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R} . \quad (4.65)$$

By choice of the basis in the two state system we can simplify the density matrix. As long as there is no other choice of a basis we just choose the eigenvectors of ρ as basis. Then the matrix corresponding to ρ is diagonal and simplifies

$$\rho = \begin{pmatrix} 1/2 + c & 0 \\ 0 & 1/2 - c \end{pmatrix} . \quad (4.66)$$

More precisely we choose as first eigenvector the one which corresponds to the larger eigenvalue of ρ , then c is non-negative. Moreover the second diagonal element $1/2 - c$ is non-negative (3.5). So the density matrix is characterized by the basis and the eigenvalue $\rho_1 = 1/2 + c$, $0 \leq c \leq 1/2$.

The values of the results \mathbf{a}_1 and \mathbf{a}_2 are not particularly important for the measuring device: the instrument does not change in an essential way if we introduce new scales. Important is the probability with which the first result occurs. According to (3.1) we need the first normalized eigenvector Λ of the device Λ to calculate it. We write the components of Λ as modulus times phase. The square modulus have to add up to one because of $\langle \Lambda | \Lambda \rangle = 1$. The modulus are therefore sine and cosine of some angle $\theta/2$. A common phase of the components is irrelevant¹, the relative phase of the two components

¹There is no choice of the phases of the components, such that the components are continuous, 2π -periodic in φ and independent of φ for $\theta = 0$ and $\theta = \pi$. So the components cannot be chosen as continuous functions on the unit sphere. The problem comes from the phases of the vector. The corresponding rays in Hilbert space vary continuously on the sphere.

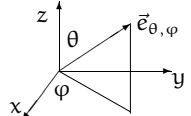
is split and attributed to both components in opposite parts.

$$\Lambda = \begin{pmatrix} \cos(\theta/2)e^{-i\varphi/2} \\ \sin(\theta/2)e^{+i\varphi/2} \end{pmatrix} \quad (4.67)$$

The angles θ and φ have geometric meaning. The state Λ (4.67) is eigenstate corresponding to the result $\hbar/2$ of the spin-1/2-operator

$$S_{\theta, \varphi} = \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\varphi} \\ \sin \theta e^{i\varphi} & -\cos \theta \end{pmatrix} = \frac{\hbar}{2} (\sigma_x e_x + \sigma_y e_y + \sigma_z e_z) , \quad (4.68)$$

which measures spin in the direction of $\vec{e}(\theta, \varphi)$.

$$\vec{e}(\theta, \varphi) = \begin{pmatrix} e_x \\ e_y \\ e_z \end{pmatrix} = \begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix} \quad \begin{array}{c} z \\ \theta \\ \vec{e}_{\theta, \varphi} \\ \varphi \\ x \end{array} \quad (4.69)$$


The vector $\vec{e}(\theta, \varphi)$ forms an angle θ with the z -axis and its projection into the x - y -plane forms the angle φ with the x -axis.

A spin measurement in direction $\vec{e}(\theta, \varphi)$ yields the result spin up with probability $w(\uparrow_{\theta, \varphi})$ which, according to (3.1), is given by the matrix element $\langle \Lambda | \rho | \Lambda \rangle$ Using (4.66) and (4.67) one calculates

$$w(\uparrow_{\theta, \varphi}) = 1/2 + c \cos \theta . \quad (4.70)$$

This probability distribution does not depend on the angle φ . It is invariant under rotations around the z -axis.

Figure (4.1) shows for the values $c = 0$, $c = 1/2$ and $c = 1/4$ the probability to find spin up as a function of the angle θ which the direction, into which the beam is split by the Stern-Gerlach device, forms with the z -axis.

The choice to employ the eigenvectors of ρ as basis for the spin states turns out to be a choice of the z -axis. The z -axis is the direction in which most particles in the mixture ρ are measured to have spin up.

The representation of the probability in the range $0 \leq \theta \leq 2\pi$ is redundant. The angle θ to the z -axis only ranges in the interval $0 \leq \theta \leq \pi$ and denotes for values $\pi < \theta \leq 2\pi$ angles $\theta' = 2\pi - \theta$.

In figure (4.1) the range $0 \leq \theta \leq 2\pi$ has been chosen to show that results of measurements of a spin-1/2-system and the probabilities of the results are unaltered by a rotation around an angle 2π .

Spin-1/2-spinors (4.67) are mapped by a rotation around 2π to their negative (4.64). Take for example the eigenstate $\Lambda_{\theta, \varphi} = \Lambda_{\frac{\pi}{2}, 0}$ with spin up in x -direction and rotate it around the z -axis around an angle 2π , then the angle φ increases from 0 to 2π and $\Lambda_{\frac{\pi}{2}, 0}$ is mapped to its negative.

$$\Lambda_{\frac{\pi}{2}, 0} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{2\pi} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -\Lambda_{\frac{\pi}{2}, 0} \quad (4.71)$$

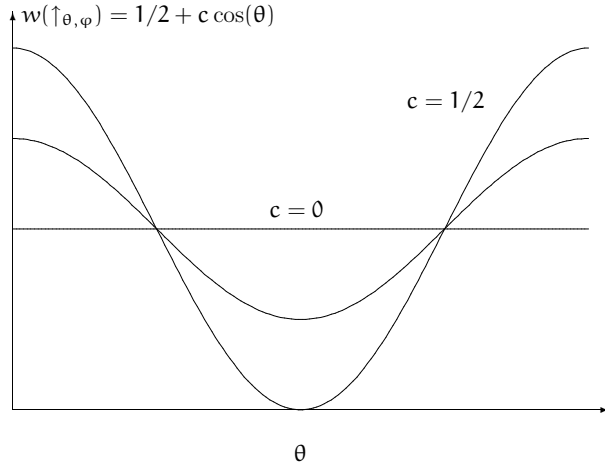


Figure 4.1: Probability of spin up of spin 1/2 particles

This does not mean that after the rotation the spin points down in x -direction $-\Lambda_{\frac{\pi}{2},0} \neq \Lambda_{\frac{\pi}{2},\pi}$. The negative sign is an unmeasurable phase. It is the ray in Hilbert space, that is the vector Ψ up to a non-vanishing number, which corresponds to a physical state. This ray is mapped to itself by a rotation around 2π . The negative phase can only be measured if one splits the spin-1/2-state, e.g. in a double slit, rotates one part around 2π and detects the change of phase relative to the second part in an interference pattern.

For $c = 1/2$ the density matrix ρ describes a mixture in which spin up is found with certainty if one measure spin in z -direction ($\theta = 0$). The probability $w(\uparrow_{\theta=0,\varphi})$ for $c = 1/2$ is 1 and the mixture is a pure state, the eigenstate spin up in z -direction.

$$\rho_{|c=1/2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}. \quad (4.72)$$

It is not always so simple to see whether a density matrix $\rho = \sum_j p_j |\Psi_j\rangle \langle \Psi_j|$ has rank 1 and can be written with one term. In such a case one probability p_j is 1, the other terms vanish and $\rho = \rho_{\text{pure}}$ is a projector $\rho_{\text{pure}}^2 = \rho_{\text{pure}}$. Because of $\text{tr } \rho = 1$ we then also have

$$\text{tr } \rho_{\text{pure}}^2 = 1. \quad (4.73)$$

If one evaluates the trace in an eigenbasis of ρ one can see that this equation is also sufficient for ρ to be a pure state. The trace is the sum of the eigenvalues of ρ . These eigenvalues lie between 0 and 1 and sum to 1. Therefore their squares sum to 1 if and only if one eigenvalue is 1 and if all other eigenvalues are 0.

The difference between the maximal probability w_{max} and the minimal probability w_{min} in relation to the maximal probability is the polarization P of the beam of spin-1/2-particles.

$$P = \frac{w_{\text{max}} - w_{\text{min}}}{w_{\text{max}}} \quad (4.74)$$

For a pure state ($c = 1/2$) the polarization is 100%. For $c = 0$ the beam is completely unpolarized and each spin measurement irrespective of its direction splits the beam into two equally intense beams.

The difference between a pure state and a mixture is seen experimentally only after measuring with different, in the case at hand rotated, devices. The corresponding probability distributions depend strongly on the apparatus, in our example on θ , if a pure state is measured. This contrast becomes smaller if mixtures are measured. Another example is the interference pattern of light behind a double slit: the interference pattern fades if the light, as in the case of light from the sun, is a mixture of different colours.

4.8 Perturbation theory

We investigate discrete eigenvalues and normalizable eigenstates of a differentiable set $H(\lambda)$ of hermitean operators. If one knows the spectrum for $\lambda = 0$, for instance, one may try to approximate the spectrum and the eigenstates for neighbouring values of λ by a Taylor series.

$$(H(\lambda) - E_n(\lambda)) \Psi_n(\lambda) = 0 \quad (4.75)$$

We assume that the operator $H(\lambda)$, its eigenvalues $E_n(\lambda)$ and its eigenstates $\Psi_n(\lambda)$ depend on λ in a differentiable way. All results of stationary perturbation theory follow from (4.75) by repeated differentiation with respect to λ , because the expansion coefficients of a Taylor series around $\lambda = 0$ are given there by repeated differentiation.

The eigenvalue equation does not fix the corresponding eigenvector $\Psi_n(\lambda) \neq 0$ completely. All complex multiples of $\Psi_n(\lambda)$ also satisfy the equation. To fix the normalization and the phase of $\Psi_n(\lambda)$ we require

$$\langle \Psi_m(\lambda) | \Psi_n(\lambda) \rangle = \delta^m_n \quad (4.76)$$

$$\langle \Psi_m(\lambda) | \frac{d}{d\lambda} \Psi_n(\lambda) \rangle_{|m=n} = 0. \quad (4.77)$$

The condition (4.76) is satisfied for $E_n \neq E_m$, since eigenvectors of a hermitean operator with different eigenvalues are orthogonal (4.9). In every subspace where an eigenvalue is degenerate, an orthonormal basis can be chosen and (4.76) can be satisfied.

Differentiation of (4.76) for $n = m$ shows that $\langle \Psi_n(\lambda) | \frac{d}{d\lambda} \Psi_n(\lambda) \rangle = i f(\lambda)$ is imaginary. By choosing the phase

$$\tilde{\Psi}_n(\lambda) = e^{i\alpha(\lambda)} \Psi_n(\lambda) \quad \alpha(\lambda) = - \int_0^\lambda d\lambda' f(\lambda') \quad (4.78)$$

equation (4.77) can be satisfied with $\tilde{\Psi}_n$. We assume that the equations (4.76) and (4.77) hold already without a redefinition of the phases.

Differentiating (4.75) with respect to λ results in

$$\left(\frac{d}{d\lambda}H - \frac{d}{d\lambda}E_n\right)\Psi_n + (H - E_n)\frac{d}{d\lambda}\Psi_n = 0. \quad (4.79)$$

Taking the scalar product with Ψ_n leads to $\langle\Psi_n|(\frac{d}{d\lambda}H - \frac{d}{d\lambda}E_n)\Psi_n\rangle = 0$, hence

$$\frac{d}{d\lambda}E_n = \langle\Psi_n|(\frac{d}{d\lambda}H)\Psi_n\rangle. \quad (4.80)$$

The scalar product with Ψ_m , $m \neq n$, gives

$$\langle\Psi_m|(\frac{d}{d\lambda}H)\Psi_n\rangle + (E_m - E_n)\langle\Psi_m|\frac{d}{d\lambda}\Psi_n\rangle = 0. \quad (4.81)$$

If an eigenvalue E is degenerate, i.e. if $E_{m_1} = E_{m_2} = \dots = E_{m_k} = E$ holds for a value of λ for some orthonormal states Ψ_{m_i} which span a k -dimensional subspace, then these states can depend on the perturbation parameter in a differentiable way only if the perturbation operator $(\frac{d}{d\lambda}H)$ does not lead to transitions between these states, i.e. if

$$\langle\Psi_{m_i}|(\frac{d}{d\lambda}H)\Psi_{m_j}\rangle = 0 \text{ for } E_{m_i} = E_{m_j} \text{ and } m_i \neq m_j. \quad (4.82)$$

In the subspace where an eigenvalue is degenerate, the orthonormal basis is to be chosen such that the perturbation operator $\frac{d}{d\lambda}H$ restricted to this subspace is diagonal.

Equations (4.81), (4.82) and (4.77) fix the scalar products of $\frac{d}{d\lambda}\Psi_n$ with all basis vectors Ψ_m . Therefore

$$\frac{d}{d\lambda}\Psi_n = - \sum_{m: E_m \neq E_n} \Psi_m \frac{\langle\Psi_m|(\frac{d}{d\lambda}H)\Psi_n\rangle}{E_m - E_n}. \quad (4.83)$$

The coefficients of $\frac{d}{d\lambda}\Psi_n$ are square summable if the vector Ψ_n depends on λ in a differentiable way,

$$\sum_{m: E_m \neq E_n} \left| \frac{\langle\Psi_m|(\frac{d}{d\lambda}H)\Psi_n\rangle}{E_m - E_n} \right|^2 < \infty \quad (4.84)$$

The equations (4.80) and (4.83) are a coupled system of differential equations for E_n and Ψ_n , from which by means of repeated differentiation all higher derivatives and thus the series expansion in λ can be determined algebraically.

If the Hamiltonian $H(\lambda)$ depends linearly on λ , the second derivative of the ground state energy $E_0(\lambda)$ is negative and it is decreased in second order. The ground state energy therefore is a convex function of the perturbation parameter.

$$\frac{d^2E_0}{d\lambda^2} = -2 \sum_{m: E_m > E_0} \frac{|\langle\Psi_m|(\frac{d}{d\lambda}H)\Psi_0\rangle|^2}{E_m - E_0} \leq 0 \quad (4.85)$$

In relativistic theories one wants a Poincaré invariant ground state with vanishing energy for every value of the coupling constant. The equation $H(\lambda)\Psi_0 = 0$ should hold

identically in the coupling λ . Then H cannot simply depend linearly on λ , otherwise the ground state energy would be a convex function of λ .

If we consider the spin operator $S_{\theta,\varphi}$ for $\theta = \frac{\pi}{2}$ as a function of φ and vary φ on a full circle, then the operator is mapped to itself

$$S_{\frac{\pi}{2},0} = S_{\frac{\pi}{2},2\pi}. \quad (4.86)$$

The corresponding eigenstate with the spin pointing upwards, whose phase and normalization are fixed by (4.77) and (4.76), changes into itself only up to a phase, or more generally in the case of an operator with degenerate states up to a unitary transformation,

$$\Psi_{\frac{\pi}{2},0} = e^{i\pi} \Psi_{\frac{\pi}{2},2\pi}. \quad (4.87)$$

5 Continuous spectrum

5.1 Wave function

Many measurement devices, in particular the measurement of position or momentum, have a continuum of possible results, which may be measured together with discrete values, called spin in the following. In the basis of eigenstates of the commuting operators that correspond to the measurement, Ψ is given by the probability amplitude $\psi_i(\mathbf{x})$ for continuous real measured values \mathbf{x} and for the i -th discrete measured values \mathbf{a}_i , $i \in I$, counted by an index set I . The state Ψ is given by a map from the set of jointly measurable real values $\mathbb{D} \subset (I \times \mathbb{R}^n)$ into the complex numbers \mathbb{C} .

$$\Psi : (i, \mathbf{x}) \rightarrow \psi_i(\mathbf{x}) \quad (5.1)$$

If \mathbf{x} belongs to the position measurement, the functions $\psi_i(\mathbf{x})$ are called position wave functions.

The square modulus $|\psi_i(\mathbf{x})|^2$ is a probability density, i.e. the probability to measure the position within a domain Δ and to measure the i -th spin result \mathbf{a}_i is

$$w(i, \Delta, \Psi) = \int_{\Delta} d^n \mathbf{x} |\psi_i(\mathbf{x})|^2 . \quad (5.2)$$

For a small domain Δ which is so small that the probability density $|\psi_i(\mathbf{x})|^2$ is nearly constant there, the integral can be approximated. Denoting the size of the domain with $d^n \mathbf{x}$, we obtain

$$w(i, \Delta, \Psi) \approx |\psi_i(\mathbf{x})|^2 d^n \mathbf{x} . \quad (5.3)$$

The probability to measure the particle around \mathbf{x} within a small domain and that the spin has the i -th value \mathbf{a}_i is the square modulus of the wave function $|\psi_i(\mathbf{x})|^2$ multiplied with the size $d^n \mathbf{x}$ of the domain.

Since probabilities are dimensionless, wave functions carry dimension

$$\dim(\psi_i(\mathbf{x})) = (\dim(d^n \mathbf{x}))^{-1/2} . \quad (5.4)$$

If the domain Δ comprises the set of all possible continuous measurement values and if one sums over all possible spin values, the sum rule for probabilities implies that Ψ is normalized.

$$\sum_i \int d^n \mathbf{x} |\psi_i(\mathbf{x})|^2 = 1 \quad (5.5)$$

Here one can read off the scalar product.

$$\langle \Phi | \Psi \rangle = \sum_i \int d^n x \phi_i^*(x) \psi_i(x) \quad (5.6)$$

Strictly speaking, one integrates only over all possible measurement values $(i, x) \in \mathbb{D} \subset \mathbb{I} \times \mathbb{R}^n$. We can easily take into consideration this restriction by confining ourselves to the Hilbert space of square integrable functions that vanish outside of \mathbb{D} .

Applied to wave functions the operators X^l , $l \in \{1, 2, \dots, n\}$ corresponding to the continuous measurement values give the probability amplitude multiplied with the measured value

$$X^l : \Psi \rightarrow X^l \Psi \quad X^l \Psi : (i, x) \rightarrow x^l \psi_i(x) . \quad (5.7)$$

Functions $f(X)$ of the operators X^l , for example the potential $V(X)$ or a plane wave $e^{ik \cdot X}$, act by multiplication with $f(x)$

$$f(X) : \Psi \rightarrow f(X) \Psi \quad f(X) \Psi : (i, x) \rightarrow f(x) \psi_i(x) . \quad (5.8)$$

The operators X^l are defined only on states Ψ whose corresponding wave function $\psi_i(x)$ remains square integrable after multiplication with x^l . The operators $e^{ik \cdot X}$ are defined for all $k \in \mathbb{R}^n$ in the entire Hilbert space.

5.2 Transformations of position

The notion of a position wave function easily transfers to manifolds. Equation (5.3) gives the probability of finding a particle with i -th spin quantum number α_i in the range of points that belong to the coordinate interval Δ . The equation holds in all coordinate systems if under general coordinate transformations $x'(x)$ the wave function transforms as a density of weight $1/2$

$$\psi'_i(x') = \left| \det \frac{\partial x}{\partial x'} \right|^{\frac{1}{2}} \psi_i(x(x')) . \quad (5.9)$$

This defines unitary transformations $U(T)$ of states corresponding to invertible maps of the manifold onto itself.

$$T : x \rightarrow x' = T(x) \quad (5.10)$$

If dT denotes the Jacobi matrix of partial derivatives

$$(dT)^k_l = \frac{\partial x'^k}{\partial x^l} \quad (5.11)$$

the unitary transformation is

$$U(T) \Psi = |\det dT|^{-\frac{1}{2}} \Psi \circ T^{-1} \quad (5.12)$$

The operators $U(T)$ are linear and unitary. Linearity in Ψ is obvious. Unitarity means invariance of scalar products. This invariance follows from the definition of $U(T)$ and the integral substitution theorem

$$\begin{aligned} \langle U\Phi | U\Psi \rangle &= \sum_i \int d^n x' (U\Phi)_i^*(x') (U\Psi)_i(x') \\ &= \sum_i \int d^n x' \left| \det \frac{\partial x}{\partial x'} \right| \phi_i^*(x(x')) \psi_i(x(x')) = \sum_i \int d^n x \phi_i^*(x) \psi_i(x) = \langle \Phi | \Psi \rangle . \end{aligned} \quad (5.13)$$

Invertible maps T of the manifold onto itself form a group with successive application of transformations as the group multiplication and the identity map as the unit element. The unitary transformations (5.12) are a representation of this group in the Hilbert space, i.e. they are linear transformations of the Hilbert space and satisfy the multiplication law

$$U(T_2)U(T_1) = U(T_2 \circ T_1) , \quad (5.14)$$

which connects successive transformations.

If we use the chain rule

$$(dT_2 \cdot dT_1)^k_l = \frac{\partial x''^k}{\partial x'^m} \frac{\partial x'^m}{\partial x^l} = \frac{\partial x''^k}{\partial x^l} = (dT_2 \circ T_1)^k_l \quad (5.15)$$

the representation property follows from

$$\begin{aligned} U(T_2)U(T_1)\Psi &= |\det dT_2|^{-\frac{1}{2}} (U(T_1)\Psi) \circ T_2^{-1} \\ &= |\det dT_2|^{-\frac{1}{2}} |\det dT_1|^{-\frac{1}{2}} \Psi \circ T_1^{-1} \circ T_2^{-1} \\ &= |\det dT_2 \cdot dT_1|^{-\frac{1}{2}} \Psi \circ (T_2 \circ T_1)^{-1} \\ &= |\det d(T_2 \circ T_1)|^{-\frac{1}{2}} \Psi \circ (T_2 \circ T_1)^{-1} = U(T_2 \circ T_1)\Psi . \end{aligned} \quad (5.16)$$

We consider a one-parameter continuous group T_α of transformations, for example rotations or translations, which are parameterized such that $T_{\alpha+\beta} = T_\alpha T_\beta$. Then $\alpha = 0$ corresponds to the identity mapping $T_0 = \text{id}$ and one has $(T_\alpha)^{-1} = T_{-\alpha}$. If α varies, then $T_\alpha x = x'(\alpha, x)$ as a function of α for each fixed x traces out a curve, the orbit, with tangent vectors

$$\frac{d(T_\alpha x)^m}{d\alpha} = \xi^m(T_\alpha(x)) . \quad (5.17)$$

The tangent vectors to these curves define a vector field, which due to $T_{\alpha+\epsilon}(x) - T_\alpha(x) = T_\alpha \circ (T_\epsilon(x) - T_0(x)) = (T_\epsilon - T_0) \circ T_\alpha(x)$ depends on α and x only via $T_\alpha(x)$. The vector field is obtained by differentiation at $\alpha = 0$.

$$\xi^m(x) = \left. \frac{d(T_\alpha x)^m}{d\alpha} \right|_{\alpha=0} \quad (5.18)$$

The vector field $\xi^m(x)$ is called an infinitesimal transformation. The solution $x(\alpha)$ to the corresponding system of differential equations

$$\frac{dx^m}{d\alpha} = \xi^m(x(\alpha)) \quad (5.19)$$

defines T_α as a map of the initial values $x(0)$ to $x(\alpha)$.

$$T_\alpha(x(0)) = x(\alpha) \quad (5.20)$$

If we differentiate the transformation law (5.12) for a one-parameter continuous group T_α at $\alpha = 0$, or if we expand $x^m = x^m + \alpha \xi^m$ and $U(T_\alpha) = e^{-\frac{i}{\hbar} \alpha N}$ with respect to α , we obtain the infinitesimal form

$$-\frac{i}{\hbar} (N\Psi)_i(x) = -\frac{1}{2} (\partial_m \xi^m) \psi_i(x) - \xi^m \partial_m \psi_i(x). \quad (5.21)$$

Here $N = i\hbar U^{-1} \partial_\alpha U$ denotes the Hermitean operator which generates the unitary transformation $U(T_\alpha)$

$$U(T_\alpha) = e^{-\frac{i}{\hbar} \alpha N}. \quad (5.22)$$

It is Hermitean, as follows from the unitarity condition $U^\dagger = U^{-1}$. The derivative of $|\det dT_\alpha|^{-\frac{1}{2}}$ contributes the term $-\frac{1}{2} (\partial_{x^m} \xi^m)$ in (5.21), because the determinant $\det dT_\alpha$ has the expansion (D.5)

$$\det \frac{\partial x'^m}{\partial x^n} = 1 + \alpha \partial_m \xi^m + O(\alpha^2). \quad (5.23)$$

On manifolds the components X^k of the position operator lose their significance since coordinates x serve only as labels of the positions, their value is irrelevant. On the circle, for instance, there exists no Hermitean position operator: spinless states on a circle with circumference l are rays in the Hilbert space of the l -periodic position wave functions $\psi(x) = \psi(x + l)$, which are square integrable in the interval $0 \leq x \leq l$. But $x\psi(x)$ is not periodic. X is not an operator in the Hilbert space of the wave functions on a circle.

The fact that there is no operator X on the circle is the solution to the puzzle of why for a normalized momentum eigenstate on the circle $\psi_n(x) = \frac{1}{\sqrt{l}} e^{i\frac{2\pi n}{l} x}$ with momentum $p = \frac{2\pi\hbar}{l} n$ the expectation value of $[X, P] \stackrel{?}{=} i\hbar$ gives, depending on the calculation, $i\hbar$ on the one hand and 0 on the other.

$$i\hbar \langle \Psi | \Psi \rangle \stackrel{?}{=} \langle \Psi | [X, P] \Psi \rangle = \langle \Psi | (XP - PX) \Psi \rangle = \langle \Psi | (Xp - pX) \Psi \rangle = 0$$

If one examines the same calculational steps on the real axis instead of the circle, then the Hermitean operators X and P do exist, but there is no normalized eigenstate of P or X .

The position measurement on the circle measures an angle and corresponds to a unitary operator

$$U : \Psi \rightarrow U\Psi, \quad U\Psi : x \rightarrow e^{i\frac{2\pi}{l} x} \psi(x). \quad (5.24)$$

From its eigenvalues $e^{i\lambda}$ one can read off the position $x = \frac{\lambda l}{2\pi}$ up to multiples of l .

To a periodic potential $V(x + l) = V(x)$ there corresponds the operator $V\Psi(x) = V(x)\psi(x)$. The potential can be written as a Fourier series $V(x) = \sum_n c_n e^{in\frac{2\pi}{l} x}$ and the operator V therefore as a series in U

$$V = \sum_n c_n U^n. \quad (5.25)$$

5.3 Translation and momentum

The requirement that translations can be defined and that no translation, apart from $T_0 = \text{id}$, keeps a point fixed $T_\alpha : x \rightarrow T_\alpha x = x + \mathbf{a}$ determines the possible values of position measurements: $\mathbb{D} = I \times \mathbb{R}^n$, where n is the dimension of the space.

According to (5.12) translations map states Ψ unitarily to shifted states $U(T_\alpha)\Psi$ in a natural way. One has $\det(dT_\alpha) = 1$ and the transformed wave functions have at the point $T_\alpha x = x + \mathbf{a}$ the same value as ψ_i at the inverse image x .

$$(U(T_\alpha)\Psi)_i(x) = \psi_i(x - \mathbf{a}) \quad (5.26)$$

We obtain the infinitesimal form (5.21) of this transformation if we differentiate the one-parameter transformations $T_{\alpha, \mathbf{a}}$ at $\alpha = 0$. The generating vector field $\xi^k = \mathbf{a}^k$ is x -independent and thus divergence-free $\partial_k \xi^k = 0$. Hence, the right-hand side of (5.21) is simply $-\mathbf{a}^k \partial_k \psi_i(x)$. Therefore the operator N , which generates the unitary transformation $U(T_\alpha)$, is linear in \mathbf{a}^k : $N = P_k \mathbf{a}^k$. Here the generating operators P_k , which belong to translations in the coordinate direction x^k , are by definition the momenta P_k belonging to the coordinates. A comparison of coefficients at the parameters \mathbf{a}^k in (5.21) implies that the momentum operator differentiates the position wave function.

$$(P_k \Psi)_i(x) = -i\hbar \partial_k \psi_i(x) \quad (5.27)$$

The operators P_k generate the unitary transformation $U(T_\alpha)$ (5.26), which corresponds to finite translations.

$$U(T_\alpha) = e^{-\frac{i}{\hbar} P \cdot \alpha} \quad (5.28)$$

The momentum operators are defined on vectors in the Hilbert space which correspond to differentiable wave functions with a square integrable derivative. The operators $U(T_\alpha) = e^{-\frac{i}{\hbar} P \cdot \alpha}$ are defined for all $\alpha \in \mathbb{R}^n$ in the entire Hilbert space, if \mathbb{D} admits translations.

On vectors which allow multiple applications of position and momentum operators, the components of the position operator as well as the components of the momentum operator commute due to $x^k x^l = x^l x^k$ and $\partial_k \partial_l = \partial_l \partial_k$. Because of

$$((X^k P_l - P_l X^k) \Psi)_i(x) = -i\hbar x^k \partial_l \psi_i(x) + i\hbar \partial_l (x^k \psi_i(x)) = (i\hbar \delta_l^k \Psi)_i(x)$$

position and momentum operator satisfy the Heisenberg commutation relations

$$[X^k, X^l] = 0, \quad [P_k, P_l] = 0, \quad [X^k, P_l] = i\hbar \delta_l^k. \quad (5.29)$$

Thus the position uncertainty ΔX^k and the momentum uncertainty ΔP_k in the same direction cannot be made small simultaneously by preparation of the state, because the Heisenberg uncertainty relation

$$\Delta X^k \Delta P_l \geq \frac{\hbar}{2} \delta_l^k \quad (5.30)$$

follows from the general uncertainty relation (4.19) and the Heisenberg commutation relation. It is certainly possible to focus the position in a plane by means of an aperture and to prepare a definite momentum perpendicular to this plane in the third direction. In this way one prepares particle beams. If one narrows the aperture, the unfocused momentum in these two directions is noticeable as diffraction at the aperture.

5.4 Rotation and angular momentum

Rotations are linear transformations of the position $D : \mathbf{x} \rightarrow D\mathbf{x}$, which leave invariant all lengths squared.

$$\sum_k (D^k_l x^l)^2 = D^k_l D^k_m x^l x^m = x^k x^k \quad \forall \mathbf{x} \Leftrightarrow D^k_l D^k_m = \delta_{lm} \quad (5.31)$$

The corresponding matrices D thus satisfy the orthogonality relation

$$D^T = D^{-1} . \quad (5.32)$$

They form the group $O(n)$ of the orthogonal transformations of \mathbb{R}^n . Here and in the following we allow ourselves to employ the convenient and, among physicists, customary convention and do not distinguish between the transformations and the corresponding matrices.

Equation (5.32) implies that $\det D = (\det D)^{-1}$, so $\det D = \pm 1$. Orthogonal transformations whose determinant has the special value 1 form the subgroup $SO(n)$ of the special orthogonal transformations.

Every one-parameter subgroup of rotations is a set of matrices $D_\alpha = e^{\alpha\omega}$ with generating matrix ω , which because of $D_\alpha^{-1} = e^{-\alpha\omega} = D_\alpha^T = e^{\alpha\omega^T}$ (5.32) is antisymmetric.

$$(\omega)^k_l = -(\omega)^l_k \quad (5.33)$$

In $n=3$ space dimensions the matrix ω is therefore a linear combination of three antisymmetric basis matrices τ_m , whose matrix elements we write using the ϵ -tensor

$$\omega^k_l = \varphi^m \epsilon_{kml} , \quad (\tau_m)^k_l = \epsilon_{kml} . \quad (5.34)$$

If $\vec{\varphi} = \vec{e}$ is a unity vector, then α is the rotation angle, for $D_\alpha = e^{\alpha\omega}$ has the following properties: ω acts on every vector \vec{v} as a vector product $\omega\vec{v} = \vec{e} \times \vec{v}$. Thus $\omega\vec{e}$ vanishes and \vec{e} marks the rotation axis $D_\alpha\vec{e} = \vec{e}$. A unit vector \vec{n}_1 orthogonal to \vec{e} is mapped by ω to the unit vector \vec{n}_2 which is orthogonal to both.

$$\omega\vec{n}_1 = \vec{n}_2 , \quad \omega\vec{n}_2 = -\vec{n}_1 . \quad (5.35)$$

If one applies the series $e^{\alpha\omega}$ to \vec{n}_1 and \vec{n}_2 and separates the even and odd powers of ω , one obtains the series of cosine and sine

$$e^{\alpha\omega}\vec{n}_1 = \vec{n}_1 \cos \alpha + \vec{n}_2 \sin \alpha , \quad e^{\alpha\omega}\vec{n}_2 = -\vec{n}_1 \sin \alpha + \vec{n}_2 \cos \alpha . \quad (5.36)$$

In particular a rotation about 2π is the identity $D_{2\pi} = \mathbb{1}$.

The vector field $\xi(\mathbf{x}) = \partial_\alpha D_\alpha \mathbf{x}|_{\alpha=0}$ corresponding to the transformation $\mathbf{x}' = D_\alpha \mathbf{x}$ is $\xi^k = \omega^k_l x^l = (\vec{\varphi} \times \vec{x})^k$. The vector field has vanishing divergence $\partial_k \xi^k = \delta_k^l \omega^k_l = 0$ and the infinitesimal transformation (5.21) of the wave function is

$$-\frac{i}{\hbar} (N\Psi)_i(\mathbf{x}) = -\omega^k_l x^l \partial_k \psi(\mathbf{x}) = -\varphi^m \epsilon_{kml} x^l \partial_k \psi(\mathbf{x}) . \quad (5.37)$$

The right-hand side is linear in φ^m , thus the operator N is linear in φ^m and of the form $N = L_m \varphi^m$. By definition the operators L_m that occur here are the components of the orbital angular momentum: they generate rotations around the coordinate axes, $\vec{L} \cdot \vec{e}$ generates rotations around \vec{e} . A comparison of coefficients of φ^m gives

$$(L_m \Psi)_i(\mathbf{x}) = -i\hbar \epsilon_{mkl} x^k \partial_{x^l} \psi_i(\mathbf{x}) , \quad L_m = \epsilon_{mkl} X^k P_l . \quad (5.38)$$

With the Heisenberg algebra (5.29) it follows that the components of the orbital angular momentum satisfy the angular momentum algebra (4.45).

$$\begin{aligned} [L_i, L_j] &= \epsilon_{ikl} \epsilon_{jmn} [X^k P_l, X^m P_n] \\ &= \epsilon_{ikl} \epsilon_{jmn} ([X^k, X^m P_n] P_l + X^k [P_l, X^m P_n]) \\ &= i\hbar \epsilon_{ikl} \epsilon_{jmn} (X^m \delta_n^k P_l - X^k \delta_l^m P_n) \\ &= i\hbar (\epsilon_{ikl} \epsilon_{jmk} - \epsilon_{imk} \epsilon_{jkl}) X^m P_l \\ &= i\hbar (\delta_{im} \delta_{lj} - \delta_{ij} \delta_{lm} + \delta_{ij} \delta_{ml} - \delta_{il} \delta_{mj}) X^m P_l \\ &= i\hbar \epsilon_{ijk} \epsilon_{kmn} X^m P_n = i\hbar \epsilon_{ijk} L_k . \end{aligned} \quad (5.39)$$

The finite unitary transformation belonging to a rotation around the axis \vec{e} with an angle α is

$$(\mathbf{U}(\vec{e}, \alpha)\Psi)_i(\mathbf{x}) = (\exp(-\frac{i\alpha}{\hbar} \vec{L} \cdot \vec{e})\Psi)_i(\mathbf{x}) = \psi_i(D_{\vec{e}, \alpha}^{-1}(\mathbf{x})) . \quad (5.40)$$

Rotations around $\alpha = 2\pi$ map positions to themselves $D_{\vec{e}, 2\pi}^{-1} \mathbf{x} = \mathbf{x}$. For orbital angular momenta (5.38) one therefore has the restriction $\mathbf{U}(\vec{e}, 2\pi) = \exp(-\frac{2\pi i}{\hbar} \vec{L} \cdot \vec{e}) = \mathbb{1}$. Applied to L_3 -eigenstates this means for rotations around the z -axis $\exp(-2\pi i m) = 1$. Thus the m -quantum numbers of the orbital angular momentum and hence also the l -quantum numbers can take integer values only.

5.5 Continuum basis

If we introduce suitable generalized basis elements $\Lambda_{i,\mathbf{x}}$, we can write states Ψ with the position wave functions as expansion coefficients.

$$\Psi = \sum_i \int d^n \mathbf{x} \Lambda_{i,\mathbf{x}} \psi_i(\mathbf{x}) \quad (5.41)$$

The scalar product (5.61) with a similarly expanded vector Φ fixes the scalar products of the basis elements.

$$\begin{aligned} \sum_{i,j} \int d^n \mathbf{x} d^n \mathbf{x}' \phi_i^*(\mathbf{x}) \langle \Lambda_{i,\mathbf{x}} | \Lambda_{j,\mathbf{x}'} \rangle \psi_j(\mathbf{x}') &= \sum_i \int d^n \mathbf{x} \phi_i^*(\mathbf{x}) \psi_i(\mathbf{x}) \quad \forall \Phi, \Psi \\ \Leftrightarrow \langle \Lambda_{i,\mathbf{x}} | \Lambda_{j,\mathbf{x}'} \rangle &= \delta^n(\mathbf{x} - \mathbf{x}') \delta_j^i \end{aligned} \quad (5.42)$$

We see that $\Lambda_{i,\mathbf{x}}$ has no finite length and is not a vector in the Hilbert space, but that $\Lambda_{i,\mathbf{x}}$ is a distribution. Only the integral (5.41) with the square integrable wave functions $\psi_i(\mathbf{x})$ yields a vector in the Hilbert space. A generalized basis with scalar products given by δ -functions as in (5.42) is called continuum normalized.

Because of (5.42) the position wave function $\psi_i(\mathbf{x})$ is given by the scalar products of Ψ with the position basis

$$\psi_i(\mathbf{x}) = \langle \Lambda_{i,\mathbf{x}} | \Psi \rangle. \quad (5.43)$$

In particular the position wave functions of the basis elements $\Lambda_{j,\mathbf{x}'}$ are delta-functions $\delta^n(\mathbf{x} - \mathbf{x}') \delta_j^i$. The basis elements are generalized eigenvectors of the position operator

$$\mathbf{X}^k \Lambda_{j,\mathbf{x}'} = \mathbf{x}'^k \Lambda_{j,\mathbf{x}'} . \quad (5.44)$$

If we insert into (5.41), we obtain in bracket notation

$$|\Psi\rangle = \sum_i \int d^n \mathbf{x} |\Lambda_{i,\mathbf{x}}\rangle \langle \Lambda_{i,\mathbf{x}} | \Psi \rangle . \quad (5.45)$$

Analogously to (2.20) the identity can be resolved with the continuum basis $\Lambda_{i,\mathbf{x}}$.

$$\mathbb{1} = \sum_i \int d^n \mathbf{x} |\Lambda_{i,\mathbf{x}}\rangle \langle \Lambda_{i,\mathbf{x}}| \quad (5.46)$$

The generalized eigenstates $\Gamma_{j,\mathbf{p}}$ of the momentum operator

$$\mathbf{P}_k \Gamma_{j,\mathbf{p}} = \mathbf{p}_k \Gamma_{j,\mathbf{p}} \quad (5.47)$$

with the eigenvalues \mathbf{p}_k , $\mathbf{p} \in \mathbb{R}^n$ form a continuum basis similar to the position states $\Lambda_{i,\mathbf{k}}$. Their position wave functions $(\Gamma_{j,\mathbf{p}})_i(\mathbf{x}) = \langle \Lambda_{i,\mathbf{x}} | \Gamma_{j,\mathbf{p}} \rangle$ are solutions to the eigenvalue equation

$$-i\hbar \partial_{\mathbf{x}^k} (\Gamma_{j,\mathbf{p}})_i(\mathbf{x}) = \mathbf{p}_k (\Gamma_{j,\mathbf{p}})_i(\mathbf{x}) \quad (5.48)$$

and, after a suitable choice of normalization factors \mathbf{c}_j^i , are given by

$$\langle \Lambda_{i,\mathbf{x}} | \Gamma_{j,\mathbf{p}} \rangle = (\Gamma_{j,\mathbf{p}})_i(\mathbf{x}) = \frac{1}{\sqrt{(2\pi\hbar)^n}} \delta_j^i e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \quad (5.49)$$

As one sees with

$$\int \frac{d^n \mathbf{x}}{(2\pi)^n} e^{-i\mathbf{x} \cdot (\mathbf{y} - \mathbf{y}')} = \delta^n(\mathbf{y} - \mathbf{y}') \quad (5.50)$$

the basis $\Gamma_{j,\mathbf{p}}$ is continuum normalized

$$\langle \Gamma_{i,\mathbf{p}} | \Gamma_{j,\mathbf{p}'} \rangle = \delta^n(\mathbf{p} - \mathbf{p}') \delta_j^i \quad (5.51)$$

and yields a resolution of the identity.

$$\mathbb{1} = \sum_i \int d^n \mathbf{p} |\Gamma_{i,\mathbf{p}}\rangle \langle \Gamma_{i,\mathbf{p}}| \quad (5.52)$$

Analogously to the position wave function one defines the momentum wave functions¹ of a state Ψ as the scalar product with the continuum basis of momentum eigenstates

$$\tilde{\psi}_j(\mathbf{p}) = \langle \Gamma_{j,\mathbf{p}} | \Psi \rangle = \sum_i \int d^n \mathbf{x} \langle \Gamma_{\mathbf{p},i} | \Lambda_{i,\mathbf{x}} \rangle \langle \Lambda_{i,\mathbf{x}} | \Psi \rangle . \quad (5.53)$$

So the momentum wave function is the Fourier transform of the position wave function and, up to a sign, vice versa

$$\tilde{\psi}_i(\mathbf{p}) = \int \frac{d^n \mathbf{x}}{\sqrt{(2\pi\hbar)^n}} e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \psi_i(\mathbf{x}), \quad \psi_i(\mathbf{x}) = \int \frac{d^n \mathbf{p}}{\sqrt{(2\pi\hbar)^n}} e^{\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \tilde{\psi}_i(\mathbf{p}) . \quad (5.54)$$

We may represent a state Ψ by the momentum wave function $\tilde{\psi}_i(\mathbf{p})$ or the position wave function $\psi_i(\mathbf{x})$. One can construct one from the other.

Due to $\langle \Gamma_{j,\mathbf{p}} | \mathbf{P}_k \Psi \rangle = \mathbf{p}_k \langle \Gamma_{j,\mathbf{p}} | \Psi \rangle$ the momentum wave function that belongs to $\mathbf{P}_k \Psi$ is $\mathbf{p}_k \tilde{\psi}_j(\mathbf{p})$. The momentum wave function that belongs to $\mathbf{X}^k \Psi$ is

$$\begin{aligned} \langle \Gamma_{j,\mathbf{p}} | \mathbf{X}^k \Psi \rangle &= \sum_i \int d^n \mathbf{x} \langle \Gamma_{j,\mathbf{p}} | \Lambda_{i,\mathbf{x}} \rangle \langle \Lambda_{i,\mathbf{x}} | \mathbf{X}^k \Psi \rangle \\ &= \int \frac{d^n \mathbf{x}}{\sqrt{(2\pi\hbar)^n}} e^{-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}} \mathbf{x}^k \psi_j(\mathbf{x}) = i\hbar \partial_{\mathbf{p}^k} \tilde{\psi}_j(\mathbf{p}), \\ (\mathbf{P}_k \Psi)^{\sim}_i(\mathbf{p}) &= \mathbf{p}_k \tilde{\psi}_i(\mathbf{p}), \quad (\mathbf{X}^k \Psi)^{\sim}_i(\mathbf{p}) = i\hbar \partial_{\mathbf{p}^k} \tilde{\psi}_i(\mathbf{p}). \end{aligned} \quad (5.55)$$

The square modulus of the momentum wave function is the probability density for momentum measurements. The probability to find the momentum in the interval $\Delta_{\mathbf{p}}$ and to obtain the i -th discrete result \mathbf{a}_i is

$$\mathbf{w}(\mathbf{i}, \Delta_{\mathbf{p}}, \Psi) = \int_{\Delta_{\mathbf{p}}} d^n \mathbf{p} |\tilde{\psi}_i(\mathbf{p})|^2 . \quad (5.56)$$

The orbital angular momentum $\vec{\mathbf{L}} = \vec{\mathbf{X}} \times \vec{\mathbf{P}}$ rotates the arguments of the momentum wave function $\tilde{\psi}_i(\mathbf{p})$ in the same way as the position arguments of the position wave function (5.40).

$$(\mathbf{L}_m \Psi)^{\sim}_i(\mathbf{p}) = -i\hbar \epsilon_{mkl} \mathbf{p}^k \frac{\partial}{\partial \mathbf{p}^l} \tilde{\psi}_i(\mathbf{p}) \quad (5.57)$$

$$(\mathbf{U}(\vec{\boldsymbol{\alpha}}, \alpha) \Psi)^{\sim}_i(\mathbf{p}) = (\exp(-\frac{i\alpha}{\hbar} \vec{\mathbf{L}} \cdot \vec{\boldsymbol{\epsilon}}) \Psi)^{\sim}_i(\mathbf{p}) = \tilde{\psi}_i(\mathbf{D}_{\vec{\boldsymbol{\alpha}}, \alpha}^{-1}(\mathbf{p})) . \quad (5.58)$$

¹To distinguish the position wave functions from the momentum wave functions we mark the latter with “ \sim ”.

5.6 Multiparticle states

The consideration of measurement results in a higher dimensional continuum, such as the six position coordinates of a two particle system, and the consideration of additional discrete measurement values, like the spins of the two particles for instance, is obvious. Such a two particle state Ψ assigns a probability amplitude for results in a six dimensional continuum with two discrete quantum numbers

$$\Psi : (i, \vec{x}, j, \vec{y}) \rightarrow \psi_{ij}(\vec{x}, \vec{y}) \quad (5.59)$$

and is given by wave functions $\psi_{ij}(\vec{x}, \vec{y})$. Here

$$w(i, \vec{x}, d^3x, j, \vec{y}, d^3y, \Psi) \approx |\psi_{ij}(\vec{x}, \vec{y})|^2 d^3x d^3y \quad (5.60)$$

is the probability of measuring the first particle with spin i at \vec{x} in the range d^3x and the second particle with spin j at \vec{y} in the range d^3y . From the probability formula one reads off the scalar product

$$\langle \Phi | \Psi \rangle = \sum_{ij} \int d^3x d^3y \phi_{ij}^*(\vec{x}, \vec{y}) \psi_{ij}(\vec{x}, \vec{y}) . \quad (5.61)$$

One has identical particles, if for all two-particle states the probability of measuring the first particle at \vec{x} with spin i and the second particle at \vec{y} with spin j coincides with the one of measuring the first particle at \vec{y} with spin j and the second particle at \vec{x} with spin i , i.e. if for all states of the two identical particles the wave function $\psi_{ji}(\vec{y}, \vec{x})$ coincides with $\psi_{ij}(\vec{x}, \vec{y})$ up to a phase. If the phase is 1 the particles are called bosons, if it is -1 they are called fermions. More precisely the wave functions for n -particle states of identical bosons are invariant under every permutation $\pi : (1, \dots, n) \rightarrow (\pi(1), \dots, \pi(n))$.

$$(\Psi_{\text{Boson}})_{i_1, \dots, i_n}(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\Psi_{\text{Boson}})_{i_{\pi(1)}, \dots, i_{\pi(n)}}(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(n)}) \quad (5.62)$$

Upon odd permutations $\text{sign}(\pi) = -1$, n -particle wave functions of identical fermions are mapped to their negative.

$$(\Psi_{\text{Fermion}})_{i_1, \dots, i_n}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \text{sign}(\pi) (\Psi_{\text{Fermion}})_{i_{\pi(1)}, \dots, i_{\pi(n)}}(\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(n)}) \quad (5.63)$$

As a consequence fermions are subject to the Pauli exclusion principle, which prohibits that multi-fermion states contain a product of the same one-particle states or, colloquially, that two identical fermions occupy the same state. But it is possible for the ground state wave function of the two electrons in the helium atom, for instance, to be a product of the same position wave function χ , since it is antisymmetric in the spin quantum number

$$\psi_{ij}(\vec{x}, \vec{y}) = \epsilon_{ij} \chi(\vec{x}) \chi(\vec{y}), \quad \epsilon_{ij} = -\epsilon_{ji}, \quad \epsilon_{\uparrow\downarrow} = 1, \quad i, j \in \{\uparrow, \downarrow\}. \quad (5.64)$$

Slater determinants are totally antisymmetric n -particle states. They arise from a product (7.1) of orthonormal one-particle states χ_i which is antisymmetrized.

$$\begin{aligned} \Psi_n &= \frac{1}{\sqrt{n!}} \sum_{\pi} \text{sign}(\pi) \chi_{\pi(1)} \otimes \chi_{\pi(2)} \otimes \cdots \otimes \chi_{\pi(n)} \\ &= \frac{1}{\sqrt{n!}} \epsilon_{i_1 i_2 \dots i_n} \chi_{i_1} \otimes \chi_{i_2} \otimes \cdots \otimes \chi_{i_n} \end{aligned} \quad (5.65)$$

The Pauli exclusion principle and the fact that electrons carry spin $1/2$ explain the basic structure of the periodic system of the elements, if one has understood the hydrogen atom, and are crucial for chemistry. Similarly condensed matter physics is ruled by the Pauli exclusion principle, for instance with the implication that in the ground state electrons occupy all one-particle states up to the Fermi energy.

Fermions carry half-integer spin, bosons have integer spin. This is first of all an experimental result. As the spin-statistics theorem these properties follow from the basic assumptions of relativistic quantum mechanics.

For states of identical particles there exist no operators which measure individual quantum numbers, say the momentum \mathbf{p}_1 of the first particle, because the eigenstates Λ of the measurement operators are also subject to the Bose- or Fermi-symmetry and the operators preserve the Bose- or Fermi-symmetry of the states. The individual quantum numbers \mathbf{p}_1 and \mathbf{p}_2 can be reconstructed only up to particle permutation from the eigenvalues of the symmetric operators $P_1 + P_2$ and $P_1^2 + P_2^2$.

6 Time evolution

6.1 Schrödinger equation

In figure (1.1) we can vary the distance between the source and the measuring device and thereby vary the time of flight. We want to discuss how the probability to obtain a result a_i depends on this time and consider for simplicity pure states. The input of the measuring device is a state $\Psi(t)$ which depends on the time of flight t .

The time evolution maps rays $\Psi(0)$ in Hilbert space to rays $\Psi(t)$. If we represent rays by normalized vectors then one has to have

$$\langle \Psi(t) | \Psi(t) \rangle = 1 \quad (6.1)$$

for all times and a changed phase of $\Psi(0)$ must change $\Psi(t)$ by a phase at most. With these restrictions the time evolution maps the unit sphere in Hilbert space to itself.

If we differentiate with respect to time we obtain

$$\langle \partial_t \Psi(t) | \Psi(t) \rangle + \langle \Psi(t) | \partial_t \Psi(t) \rangle = 0 . \quad (6.2)$$

The infinitesimal time evolution should be given by a differential equation of first order, otherwise $\Psi(0)$ would not characterize the system completely at time $t = 0$ and additional data such as $\partial_t \Psi$ could be prepared and would become measurable in the course of time.

The assumption that the time evolution is linear in Ψ is the true content of the superposition principle of quantum mechanics. It postulates the Schrödinger equation.

$$i\hbar \partial_t \Psi = H\Psi \quad H = H^\dagger \quad (6.3)$$

For if $\partial_t \Psi = O\Psi$ holds for some linear operator O then this operator is antihermitean by (6.2). The Schrödinger equation expresses this fact after factorization of i and Planck's constant \hbar . The Hamilton operator generates the time evolution. It has the dimension of an energy.

A mixture $\rho = \sum_j p_j |\Psi_j\rangle \langle \Psi_j|$ changes with time because the states Ψ_j change. The probabilities p_j are the probabilities with which the states $\Psi_j(0)$ are produced in the source and with which $\Psi_j(t)$ occurs at time t . These probabilities remain unchanged by time evolution which is generated by the Hamiltonian.

$$\rho(t) = \sum_j p_j |\Psi_j(t)\rangle \langle \Psi_j(t)| \quad (6.4)$$

If one differentiates with respect to time and observes that bra vectors are antilinear (2.12) and therefore satisfy

$$-i\hbar\partial_t\langle\Psi(t)| = \langle i\hbar\partial_t\Psi(t)| = \langle H\Psi(t)| \quad (6.5)$$

then one obtains from the Schrödinger equation (6.3) the von-Neumann equation for the time evolution of the density matrix

$$i\hbar\partial_t\rho = H\rho - \rho H = [H, \rho] . \quad (6.6)$$

It is compatible with the basic structures of quantum mechanics that $\Psi(t)$ depends non-linearly on $\Psi(0)$ and is given by a map $\Psi(t) = \Phi_t(\Psi(0))$ of $\mathcal{H} - \{0\}$ to itself, which maps rays in Hilbert space to rays

$$\Phi_t(\lambda\Psi) = f_t(\lambda, \Psi)\Phi_t(\Psi), \quad \forall\Psi \neq 0, \quad \forall\lambda \neq 0, \quad (6.7)$$

where the complex factor $f_t(\lambda, \Psi) \in \mathbb{C}$ does not vanish. The corresponding differential equations for the time evolution are called “non-linear Schrödinger equation”. Non-linear time evolution is studied in ongoing research, its physical relevance is unclear. We follow the widespread terminology and restrict the term quantum mechanical system to physical systems with a linear time evolution.

It is one of the peculiarities of quantum mechanics that time is not measured as a property of the physical state in figure (1.1) but that time and the measuring device is part of the outer world. From measuring the state one cannot tell the time though one can compare the evolution of the system and a clock and tell from the clock how much time has passed.

In quantum mechanics there is no operator which corresponds to the measurement of time. This has the welcome consequence that there are no eigenstates of time. Such states would be at one and no other time, for these states time would not pass.

From the general uncertainty relation (4.19) and the Schrödinger equation one can deduce a lower bound of the product of the uncertainty of time and the uncertainty of energy. One defines the uncertainty of time Δt to be given by the time which passes in a state until the expectation value $\langle A \rangle$ has changed by the uncertainty of A .

$$\Delta t = \frac{\Delta A}{\left| \frac{d}{dt}\langle A \rangle \right|} \quad (6.8)$$

Then (4.19) and (6.3) imply

$$\Delta A \Delta H \geq \frac{1}{2} | \langle [A, H] \rangle | = \frac{\hbar}{2} \left| \frac{d}{dt}\langle A \rangle \right|, \quad (6.9)$$

and from the definition of Δt one concludes

$$\Delta t \Delta E \geq \frac{\hbar}{2} \quad (6.10)$$

irrespective of which device A is used for the measurement.

Even in relativistic quantum mechanics the variable t of time dependent wave functions is basically different from the position variables x . Multiparticle states have wave functions with several position variables but with one and only one time. In the scalar product \mathbf{t} is no integration variable, $w(\Delta, \Psi(t))$ is not the probability to find the result \mathbf{t} in an interval dt . The time t rather parameterizes the states $\Psi(t)$ between source and measuring device. Where exactly the source ends and the device starts is irrelevant. The setup in figure (1.1) is not changed essentially if we count part of the evolution of the state to the preparation in the source or to the measurement in the device. The time t between preparation and measurement is positive.

If the Hamilton operator H does not change with time and if the initial state $\Psi(t=0)$ is an eigenstate of H then the probabilities of all results do not depend on time, because

$$i\hbar\partial_t\Psi = H\Psi = E\Psi \quad (6.11)$$

implies $\Psi(t) = e^{-\frac{i}{\hbar}Et}\Psi(0)$ and the ray in Hilbert space corresponding to $\Psi(t)$ does not change with time. Energy eigenstates are therefore termed stationary states.

If the Hamilton operator is time independent and if it commutes with a hermitean operator A

$$[H, A] = 0, \quad (6.12)$$

which corresponds to a time independent measuring device then the probability $w(i, A, \Psi(t))$ (1.1) to obtain the i -th result a_i does not change with time. This holds because the corresponding normalized eigenstate Λ_i can be chosen as eigenstate of H (4.24) $H\Lambda_i = E_i\Lambda_i$ and the probability amplitude $\langle \Lambda_i | \Psi(t) \rangle$ changes by a phase only

$$\begin{aligned} i\hbar\partial_t\langle \Lambda_i | \Psi(t) \rangle &= \langle \Lambda_i | H\Psi(t) \rangle = \langle H\Lambda_i | \Psi(t) \rangle = E_i\langle \Lambda_i | \Psi(t) \rangle \\ \langle \Lambda_i | \Psi(t) \rangle &= e^{-\frac{i}{\hbar}E_it}\langle \Lambda_i | \Psi(0) \rangle. \end{aligned} \quad (6.13)$$

In particular if $\Psi(0)$ is eigenstate of A corresponding to the result \mathbf{a} then it stays eigenstate and the quantum number \mathbf{a} is a conserved quantity. This is the reason for the outstanding role of energy, momentum and angular momentum. Outside of interaction regions, in particular before and after scattering, energy, momentum and angular momentum are conserved.

If the Hamilton operator is time independent and if one knows its eigenvalues $\{E_1, E_2, \dots, E_i, \dots\}$ together with their corresponding, orthonormalized eigenstates $\Lambda_1, \Lambda_2, \dots, \Lambda_i, \dots$, then the problem to determine the time evolution of an arbitrary state Ψ is simplified to the task to determine the components of the initial state in the eigenbasis of the Hamilton operator and to compose the state at later time with phase shifted components

$$\Psi(t) = \sum_i e^{-\frac{i}{\hbar}E_it}\Lambda_i\psi_i, \quad \psi_i = \langle \Lambda_i | \Psi(t=0) \rangle. \quad (6.14)$$

Consequently there corresponds to each given Hamilton operator the standard task to determine its spectrum and its eigenstates.

The eigenvalue equation of the Hamilton operator

$$(\mathbf{H} - E_i)\Lambda_i = 0 \quad (6.15)$$

is the time independent Schrödinger equation. Contrary to widespread opinion this equation does not allow to calculate the state of the quantum mechanical system. Unstable particles or wave packets of free particles or more generally all non-stationary states are no energy eigenstates. However, parts of a state with different energy separate automatically, if one just waits sufficiently long, if these parts are separated by a finite energy difference and if their velocity is different. In this way energy eigenstates are often prepared automatically.

If the quantum mechanical state is a particle moving in one dimension without spin, then the most general state (5.41) is a linear combination of position eigenstates Λ_x with expansion coefficients given by the wave function $\psi(x)$

$$\Psi = \int dx \Lambda_x \psi(x) . \quad (6.16)$$

The Hamilton operator which corresponds to motion in a potential consists of kinetic energy $\frac{p^2}{2m}$ and potential energy $V(X)$. Applied to position wave functions the momentum operator acts as derivative $(P\Psi)(x) = -i\hbar \frac{d\Psi(x)}{dx}$ (5.27) and the potential multiplies the position wave function $(V\Psi)(x) = V(x)\psi(x)$ (5.8). So the stationary Schrödinger equation $(\mathbf{H} - E)\Psi = 0$ reads

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) - E\right)\psi(x) = 0 . \quad (6.17)$$

The solutions have to be normalizable, if they are to correspond to vectors in the Hilbert space of square integrable wave functions. In addition, all solutions are of interest, which do not grow with $x \rightarrow \pm\infty$, because from these solutions one can construct wave packets Ψ which approximate the energy eigenvalue equation $\|(\mathbf{H} - E)\Psi\|^2 < \epsilon$ to each given precision $\epsilon > 0$. For example, if the potential vanishes then the generalized momentum eigenstates $\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar} p x}$ are generalized energy eigenstates with $E = \frac{p^2}{2m}$. They belong to the continuous spectrum of the kinetic energy which consists of the real, non-negative numbers $E \geq 0$.

To keep the mathematical problems simple one prefers to investigate the time independent Schrödinger equation (6.17) for simplified potentials like the potential well or the potential barrier.

6.2 Schrödinger picture, Heisenberg picture

In quantum mechanics one cannot decide whether the probability distributions of results change because the state Ψ changes in the course of time or because the measuring device changes and the state remains unchanged.

In the Schrödinger picture, which we use in our discussion, one attributes the time evolution to the states and uses time independent operators for the measuring devices.

If one is given a set of unitary operators which depend in a differentiable way on the time t

$$\mathbf{U}^\dagger(t) = \mathbf{U}^{-1}(t) \quad (6.18)$$

and if one uses $\Psi'(t)$ as state and $\Lambda'_i(t)$ as eigenstate of the operator $\mathcal{A}'(t)$

$$\Psi'(t) = \mathbf{U}(t)\Psi(t), \quad \Lambda'_i(t) = \mathbf{U}(t)\Lambda_i, \quad \mathcal{A}'(t) = \mathbf{U}(t)\mathcal{A}\mathbf{U}^{-1}(t), \quad (6.19)$$

rather than $\Psi(t)$ and Λ_i and \mathcal{A} , then one obtains for all times and all measuring devices and all physical states unchanged probability amplitudes and unchanged eigenvalues of the operators corresponding to the measuring devices

$$\langle \Lambda'_i(t) | \Psi'(t) \rangle = \langle \mathbf{U}(t)\Lambda_i | \mathbf{U}(t)\Psi(t) \rangle = \langle \Lambda_i | \mathbf{U}^\dagger(t)\mathbf{U}(t)\Psi(t) \rangle = \langle \Lambda_i | \Psi(t) \rangle .$$

$$\mathcal{A}\Lambda_i = \alpha_i \Lambda_i \Leftrightarrow \mathbf{U}(t)\mathcal{A}\mathbf{U}^{-1}(t) \mathbf{U}(t)\Lambda_i = \alpha_i \mathbf{U}(t)\Lambda_i .$$

Differentiating $\partial_t(\mathbf{U}(t)\mathbf{U}^{-1}(t)) = 0$ one obtains from the product rule

$$\partial_t \mathbf{U}^{-1} = -\mathbf{U}^{-1}(\partial_t \mathbf{U})\mathbf{U}^{-1} . \quad (6.20)$$

Therefore the product $\mathbf{U}^{-1}(t)\partial_t \mathbf{U}(t)$ is antihermitean.

$$(\mathbf{U}^\dagger \partial_t \mathbf{U})^\dagger = (\partial_t \mathbf{U}^{-1})\mathbf{U} = -\mathbf{U}^{-1} \partial_t \mathbf{U} . \quad (6.21)$$

If we write

$$\mathbf{U}^{-1}(t)i\hbar\partial_t \mathbf{U}(t) = -\mathbf{H}_0(t) \quad \mathbf{H}_0 = \mathbf{H}_0^\dagger \quad (6.22)$$

then the Schrödinger equation for $\Psi(t)$ implies the time evolution

$$i\hbar\partial_t \Psi' = \mathbf{H}'(t)\Psi' \quad \text{where } \mathbf{H}' = \mathbf{U}(\mathbf{H} - \mathbf{H}_0)\mathbf{U}^{-1} . \quad (6.23)$$

Operators which correspond to measuring devices satisfy

$$i\hbar\partial_t \mathcal{A}'(t) = -[\tilde{\mathbf{H}}, \mathcal{A}'(t)] \quad \text{where } \tilde{\mathbf{H}} = \mathbf{U}\mathbf{H}_0\mathbf{U}^{-1} . \quad (6.24)$$

In particular, if one chooses $\mathbf{H}_0 = \mathbf{H}$ and determines $\mathbf{U}(t)$ as solution of $i\hbar\partial_t \mathbf{U} = -\mathbf{U}\mathbf{H}$ with $\mathbf{U}(0) = 1$, then Ψ' is time independent and $\tilde{\mathbf{H}} = \mathbf{H}$. The primed quantities are the states and measuring operators in the Heisenberg picture. They satisfy

$$i\hbar\partial_t \Psi_H = 0, \quad i\hbar\partial_t \rho_H = 0, \quad i\hbar\partial_t \mathcal{A}_H = -[\mathbf{H}, \mathcal{A}_H] . \quad (6.25)$$

At $t = 0$ the states and measuring operators in the Heisenberg picture and in the Schrödinger picture coincide.

If states traverse an interaction region then for early and late time the interaction \mathbf{H}_{int} vanishes. In the interaction picture one chooses \mathbf{H}_0 in such a manner that $\mathbf{H}_{\text{int}} = \mathbf{U}(\mathbf{H} - \mathbf{H}_0)\mathbf{U}^{-1}$. Then the states become time independent for early and late times

and the limit $\lim_{t \rightarrow \pm\infty} \Psi_W(t)$ exists. This picture is advantageous for the discussion of scattering.

Even if all pictures are mathematically equivalent they are nevertheless differently intuitive. In the Schrödinger picture one can intuitively understand that a free wave packet of a massive particle spreads because the wave packet consists of parts with different momentum and, because the mass is not zero, with different velocity. The spread of the wave packet is completely analogous to a race where the competitors start at the same line and finish one after the other. In the equivalent Heisenberg picture the state does not change but the measuring instrument as if not the competitors but the referees at the finishing line separated from each other.

6.3 Groundstate energy

Strictly speaking the eigenvalues of the Hamilton operator which generates the time evolution, the energy, is not measurable. Only differences of energies can be determined. In particular the energy of the ground state cannot be inferred from the time evolution.

For one cannot distinguish the states $\Psi(t)$ from $\Psi'(t) = e^{i\alpha t/\hbar} \Psi(t)$ for arbitrary, real α . Ψ' and Ψ yield at all times and for all measuring instruments the same probability distribution of results. If Ψ satisfies the Schrödinger equation (6.3) with some Hamilton operator H then $\Psi'(t)$ satisfies the Schrödinger equation with $H' = H - \alpha$. Therefore one cannot distinguish between H and $H - \alpha$. Discussions about the magnitude of the ground state energy are therefore similar to medieval considerations, how many angels fit on the tip of a needle. At least we know that we cannot determine the ground state energy from any measurement of the time evolution.

In the Heisenberg picture it is even simpler to see that the time evolution of the measuring operators is unchanged, if one adds a number α to the Hamilton operator, because a number α commutes with each operator.

Not the time evolution but other effects can fix the absolute value of the energy. Density of energy or more precisely the energy momentum tensor contributes in general relativity to gravity and curvature of spacetime. From the observations one concludes that contribution of the energy density of the ground state to the curvature of spacetime is small and compatible with a vanishing ground state energy. However, there is no theory which successfully unites quantum mechanics and general relativity.

To a free, non-relativistic particle with momentum \vec{p} one attributes the energy $E = \vec{p}^2/(2m) \geq 0$ and has thereby chosen the groundstate energy such that the energy at rest vanishes.

The hydrogen atom has a continuous non-negative energy of motion of the center of mass. In bound states the relative motion of proton and electron has a discrete, negative spectrum with energies $E_{n,l,m} = -\frac{Ry}{n^2}$ and eigenstates $\Lambda_{n,l,m}$. In this formula $Ry = \frac{ue^4}{2\hbar^2}$ is the Rydberg constant. For each natural number $n = 1, 2, \dots$ and for each $l = 0, 1, \dots, n-1$, which ranges in integer steps between 0 and an upper bound $n-1$ for each fixed energy, there are $2l+1$ states corresponding to $m = -l, -l+1, \dots, l$. If more precisely one takes into account the spin of the electron, then the number of states

double and if one considers the spin of the proton, they double again. The spin of the electron is crucial for the periodic system of the elements, the spin of the proton makes itself felt only in the hyperfine structure of the energies and is often not even mentioned in text books. Above the discrete spectrum there is the continuous spectrum of positive energies of the ionized electron proton pair. It is natural to attribute the energy 0 to the lowest ionized state.

In relativistic theories in flat space the energy of the groundstate is fixed. There the Hamilton operator $H = cP^0$ is a component of the four vector P^m , $m = 0, 1, 2, 3$, which satisfies together with the generators of Lorentz transformations $M^{mn} = -M^{nm}$ the following commutation relations

$$[M^{mn}, P^l] = -i(\eta^{ml}P^n - \eta^{nl}P^m) \quad m, n, l \in \{0, 1, 2, 3\}. \quad (6.26)$$

These relations do not allow to add numbers to P^m . In relativistic theories the groundstate, the vacuum, must have energy 0 if a normalizable groundstate exists at all. Moreover, a free particle with momentum \vec{p} and mass m must have the energy $E = \sqrt{m^2c^4 + \vec{p}^2c^2}$.

Traditionally the groundstate of the harmonic oscillator is attributed the energy $\hbar\omega/2$. In his derivation of the black body energy density, with which quantum mechanics was founded in 1900, Planck knew this better: he attributed to states with n photons the energy $n\hbar\omega$ rather than $(n+1/2)\hbar\omega$, which moreover would be incompatible with a relativistic covariant description of photons.

We have already argued that the energy of the groundstate is not measurable, therefore one should choose it such that calculations become simple and finite. A finite volume allows for infinitely many frequencies ω_i of photons. If one attributes to each frequency a contribution $\hbar\omega_i/2$ to the groundstate energy then already the state without photons has infinite energy $\sum_i \hbar\omega_i/2 = \infty$.

The misconception that the groundstate energy is well defined starts already in classical physics. The choice of the Hamilton function

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \quad (6.27)$$

disposes of the energy in such a way that the state with lowest energy, the point $(x = 0, p = 0)$ in phase space has energy 0. This choice simplifies the algebraic expression for the potential energy $V(x)$. However, one could have chosen $V(x) = 1/2 m\omega^2x^2 - \hbar\omega/2$ equally well.

6.4 Canonical quantization, normal order

The misconception that the groundstate energy is well defined is continued with canonical quantization. Canonical quantization is the prescription to read in the algebraic expression for the Hamilton function $H(p, x)$ the symbols p and x as hermitean operators P and X which satisfy the Heisenberg commutation relations (5.29). From the

Hamilton function of the one dimensional harmonic oscillator one obtains for example the Hamilton operator

$$H_{\text{Oszillator}} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (6.28)$$

This operator has eigenvalues $(n + 1/2)\hbar\omega$, $n = 0, 1, 2, \dots$.

Simple as canonical quantization seems to be, it is not even defined: canonical quantization is not a map of Hamilton functions to Hamilton operators. The result of canonical quantization does not only depend on the function in phase space $H(p, x)$ but on its algebraic form.

We can write the Hamilton function of the harmonic oscillator as square modulus of complex phase space coordinates using $x_0 = \sqrt{\frac{\hbar}{m\omega}} \neq 0$.

$$a = \frac{1}{\sqrt{2}}\left(\frac{x}{x_0} + \frac{i}{\hbar}x_0 p\right), \quad a^\dagger = \frac{1}{\sqrt{2}}\left(\frac{x}{x_0} - \frac{i}{\hbar}x_0 p\right) \quad (6.29)$$

$$H = \hbar\omega a^\dagger a \quad (6.30)$$

If we quantize the Hamilton function in this algebraic form then a and a^\dagger become annihilation and creation operators (4.33) and the Hamilton operator corresponding to (6.30) has eigenvalues $n\hbar\omega$ (section 4.5).

If we write the Hamilton function as $H = \hbar\omega((1 - \lambda)a^\dagger a + \lambda a a^\dagger)$ with some $\lambda \in \mathbb{R}$, then we obtain an arbitrary ground state energy $\lambda\hbar\omega$ upon canonical quantization.

In any case the groundstate satisfies the equation $a\Psi_0 = 0$, which determines the position wave function

$$\left(\frac{x}{x_0} + x_0 \partial_x\right)\psi_0(x) = 0. \quad (6.31)$$

The groundstate wave function of the harmonic oscillator is a Gauß function

$$\psi_0(x) = (\pi^{-\frac{1}{4}} x_0^{-\frac{1}{2}}) e^{-\frac{x^2}{2x_0^2}}. \quad (6.32)$$

The result of canonical quantization and in particular the groundstate energy depend on the algebraic form of the classical Hamilton function. Canonical quantization is therefore no map of a function of phase space to a corresponding operator.

There is, however, a different, very simple rule of quantization, which attributes operators to analytic functions H , the normal order $:H:$. The normal order is linear

$$:c_1 H_1 + c_2 H_2: = c_1 :H_1: + c_2 :H_2:, \quad :1: = 1, \quad (6.33)$$

and is declared recursively on monomials in a, a^\dagger

$$:aH: = :Ha: = :H:a, \quad :a^\dagger H: = :Ha^\dagger: = a^\dagger :H: \quad \forall H. \quad (6.34)$$

The normal order of a monomial is a product of creation and annihilation operators a^\dagger and a , in which the creation operators stand on the left and the annihilation operators

stand on the right. The definition of normal order is easily extended to several creation and annihilation operators a_i^\dagger and a_j , as long as the order among the creation operators and the order among the annihilation operators is irrelevant.

$$[a_i, a_j] = 0, \quad [a_i^\dagger, a_j^\dagger] = 0, \quad [a_i, a_j^\dagger] = \delta^j_i. \quad (6.35)$$

The arguments of normal order consist of commuting variables, for if the argument is a product of factors H_1, H_2, H_3 then one has

$$:H_1 H_2 H_3: = :H_2 H_1 H_3: = :H_3 H_1 H_2:. \quad (6.36)$$

In particular the argument of normal order cannot satisfy an identity such as $XP - PX = i\hbar$ (5.29) because the normal order of $XP - PX$ vanishes.

Normal order is linear, but the normal order of a product factors is not the product of the normal order of the factors.

$$:H_1 H_2: \neq :H_1: :H_2: \quad (6.37)$$

This is welcome. Otherwise all products of normal order operators would commute by (6.36). Then quantum mechanical systems would behave like classical statistical systems because with respect to commuting operators all states Ψ are characterised by classical probability distributions given by the square modulus of the scalar products $\langle \Lambda | \Psi \rangle$.

Normal order (6.27) implies a vanishing groundstate energy. It is nevertheless an arbitrary prescription of quantization which regrettably depends on the chosen phase space coordinates. Classical systems before quantization can be cast into different, equivalent forms by canonical transformations. The quantization of the system in the different forms leads to quantum models which can be inequivalent. My disappointment about the lack of a unique quantization prescription and the arbitrariness of the quantization prescription is limited. One does not need a correspondence principle between classical and quantum systems. Otherwise one could tell the quantum properties from the classical system and two different quantum systems could not coincide in their classical regime. Canonical quantization leads the intuition which quantum system one should investigate. No matter how such a model is constructed, whether such a model is right is decided by the observations which have to agree with the theoretical conclusions from the model.

6.5 Reduction of state

Some authors postulate as axiom for quantum mechanical systems the reduction of state, that as a result of an ideal measurement with a device A the state instantaneously changes into the eigenstate corresponding to the measured value.

This axiom implies seemingly paradoxical conclusions such as the Einstein-Rosen-Podolsky paradox. The axiom can be replaced by an unspectacular physical analysis.

If in figure (1.1) we mask all rays which correspond to results a_2, a_3, \dots and only consider the ray with the result a_1 then we have prepared a new source consisting of the

original source, the beam and the device. The properties of the state which is prepared in this fashion depend on the device and should not be fixed by axioms. For example, a homogeneous magnetic field orthogonal to the outgoing beam would not influence the splitting of the beam into the partial beams and the distribution of the results \mathbf{a}_i . Therefore the homogeneous magnetic field would not change the reading of the device. However, it would rotate the spin in the outgoing beams. Without changing the results one can change the measuring device and change the states which are prepared by the measurement.

One can, of course, restrict the term “ideal measuring device” to such devices which prepare eigenstates to a repeated measurement with a similar device. But then the axiom of state reduction turns into a definition of a class of devices which happen to be ideal. For each real measuring device one has to measure whether it prepares eigenstates for a repeated measurement with a similar device.

Which state results after a measurement depends on the experimental setup. If all partial beams behind the device are combined to a new beam and if there is no fixed relative phase of the different beams and if the device prepares eigenstates to a repeated measurement then the combined beam behind the device is a mixture

$$\rho_{\mathbf{A}, \Psi} = \sum_i w(\mathbf{i}, \mathbf{A}, \Psi) |\Lambda_i\rangle \langle \Lambda_i|. \quad (6.38)$$

In such a mixture the conditional probability, to measure again the result \mathbf{a}_i if one had measured \mathbf{a}_i at first, is 1. The restriction of the set of possible events to the subset of events where \mathbf{a}_i was measured in the first measurement is the reduction of state. This reduction can be made long after the measurement if only the result of the measurement has been recorded. The reduction can take place instantaneously after the result of the first measurement is known. It can also retroact: if the result of the second measurement is known and the first is unknown then one can predict the first result and confirm this prediction later from the records.

Some devices are so simple and controllable that phase relations of the different partial beams which correspond to the results \mathbf{a}_i can be controlled. The state which one combines from the partial beams depends sensitively on the process of combination. In the extreme case, e.g. a double slit, the partial beams combine to a pure state where the parts interfere with each other. If the possible results \mathbf{a}_i can really be read off then the phase relations between the corresponding partial beams become uncontrollable and the combined beam is a mixture such as (6.38).

Numerous devices do not prepare eigenstates for a repeated measurement. This is true in particular for measurements which determine the number of photons, which cross a given area within a given time: photosensitive layers and photomultipliers count and destroy photons.

The measuring device in figure (1.1) has input and output. For each state it takes time to cross the device and to split the beam into separate partial beams. This time depends on the internal details of the device. However, for no device does this time vanish, as assumed in Zenon’s paradox. The idealization that the duration between the

begin and the end of a measurement can become infinitely short contradicts the quantum mechanical uncertainty relation (9.10) $\Delta t \Delta E \geq \hbar/2$.

6.6 Time evolution of the two state system

The time evolution of a two state system is simple enough to obtain for each time independent Hamilton operator an overview over the time dependence of the probabilities of results of all measuring devices and all states.

For our discussion we use the eigenbasis of the Hamilton operator. In this basis it is diagonal and has in all conceivable cases the form

$$H = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \quad E_1, E_2 \in \mathbb{R}. \quad (6.39)$$

We choose the eigenstates Λ_i , $i = 1, 2$, of H time independent, then the components $\psi_i(\mathbf{t}) = \langle \Lambda_i | \Psi(\mathbf{t}) \rangle$, $i = 1, 2$, of each state satisfy the following, decoupled Schrödinger equation

$$i\hbar \partial_t \psi_i(\mathbf{t}) = E_i \psi_i(\mathbf{t}), \quad i = 1, 2, \quad (6.40)$$

and have the solution

$$\psi_1(\mathbf{t}) = \psi_1(0) e^{-\frac{i}{\hbar} E_1 \mathbf{t}}, \quad \psi_2(\mathbf{t}) = \psi_2(0) e^{-\frac{i}{\hbar} E_2 \mathbf{t}}. \quad (6.41)$$

If the state Ψ is measured at time \mathbf{t} then the first result of any measuring device occurs with probability

$$w(\mathbf{t}) = |\phi_1^* \psi_1(\mathbf{t}) + \phi_2^* \psi_2(\mathbf{t})|^2. \quad (6.42)$$

In this equation ϕ_1 and ϕ_2 are the components of the first eigenvector of the measuring device. Elementary calculation shows that $w(\mathbf{t})$ has the form

$$w(\mathbf{t}) = \mathbf{a} + \mathbf{b} \cos(\omega \mathbf{t} + \alpha), \quad \mathbf{a} \geq \mathbf{b} \geq 0, \quad \omega, \alpha \in \mathbb{R} \quad (6.43)$$

with $\mathbf{a} = |\phi_1^* \psi_1(0)|^2 + |\phi_2^* \psi_2(0)|^2$, $\mathbf{b} e^{i\alpha} = 2\phi_1 \phi_2^* \psi_1(0)^* \psi_2(0)$ and

$$\omega = \frac{E_1 - E_2}{\hbar}. \quad (6.44)$$

The probability for the occurrence of the first result oscillates with the Rabi frequency $\omega/2\pi$ which is given by the energy difference. No absolute value of the energy appears in the time evolution of a measurable quantity.

The amplitude \mathbf{b} of the Rabi oscillation vanishes if $\Psi(0)$ or the eigenstate Φ of the device is an energy eigenstate.

Rabi oscillations occur in physically different situations whenever the time evolution concerns only two states. In particle physics the phenomenon is termed particle oscillation. It is observed at the neutral K-mesons and also at neutrinos. In quantum optics the phenomenon is called “quantum beat”.

If a mixture ρ with eigenvalues ρ_i and eigenstates Υ_i is measured rather than a pure state then the Rabi frequency remains unchanged. The parameters \mathbf{a} , \mathbf{b} and $\boldsymbol{\alpha}$ are $\mathbf{a} = \rho_1 \mathbf{a}_1 + \rho_2 \mathbf{a}_2$ and $\mathbf{b} e^{i\boldsymbol{\alpha}} = \rho_1 \mathbf{b}_1 e^{i\boldsymbol{\alpha}_1} + \rho_2 \mathbf{b}_2 e^{i\boldsymbol{\alpha}_2}$, where \mathbf{a}_i , \mathbf{b}_i and $\boldsymbol{\alpha}_i$ correspond to $\Psi = \Upsilon_i$. Taking into account that $\langle \Upsilon_1 | \Upsilon_2 \rangle = 0$ one confirms that the amplitude \mathbf{b} of Rabi oscillations is proportional to the difference $(\rho_1 - \rho_2)$ of the eigenvalues of the density matrix. The amplitude decreases with decreasing polarization.

6.7 Energy bands

Let us investigate the spectrum (6.17) of a Hamilton operator of a one dimensional spinless system in a periodic potential [2, chapter XIII.16]

$$V(x+l) = V(x) \quad \forall x. \quad (6.45)$$

The differential equation (6.17) with a periodic function $V(x)$ is called Hill's equation. It occurs in mechanics in the context of oscillations with time dependent, periodic frequency, for example in the calculation of the orbit of the moon.

Because of the periodicity the Hamilton operator commutes with the translation \mathbf{U}_l (5.26) by the length l

$$\begin{aligned} (\mathbf{H}\mathbf{U}_l\Psi)(x) &= \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)\right)(\mathbf{U}_l\Psi)(x) = \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)\right)\psi(x-l) \\ &= \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x-l)\right)\psi(x-l) = (\mathbf{U}_l\mathbf{H}\Psi)(x). \end{aligned} \quad (6.46)$$

Therefore, the Hamilton operator and the unitary transformation \mathbf{U}_l can jointly be diagonalized (4.24).

The translation is a unitary transformation (5.26) with complex eigenvalue with modulus 1 (4.11). We write these eigenvalues as e^{-ikl} with a real k . Then the eigenvalue equation $\mathbf{U}_{-l}\Psi_k = e^{ikl}\Psi_k$ reads

$$\psi_k(x+l) = e^{ikl}\psi_k(x). \quad (6.47)$$

This is a periodicity condition for the wave function which among physicists is called Bloch theorem, to mathematicians it is known as Floquet theorem. The condition is compatible with the eigenvalue equation and can be imposed to simplify the mathematical analysis. The condition does not state that each wave function in a periodic potential is periodic up to a phase. As a counterexample one can consider a free particle with vanishing potential. The potential $V \equiv 0$ is trivially periodic. Nevertheless the wave packets of free particles are not periodic, they are composed of periodic functions.

The eigenvalue equation (6.17) is a real, linear, homogeneous differential equation of second order for the wave function $\psi(x)$. Therefore the wave function and its derivative at $x=l$ are linearly related to the initial values at $x=0$. If we combine the wave function and its derivative to the components of a vector \mathbf{u} then we can write the linear

relation with a matrix \mathbf{A}

$$\mathbf{u}(x) = \begin{pmatrix} \psi(x) \\ \psi'(x) \end{pmatrix}, \quad (6.48)$$

$$\mathbf{u}(l) = \mathbf{A} \mathbf{u}(0). \quad (6.49)$$

The linear map \mathbf{A} of initial values $\mathbf{u}(0)$ to $\mathbf{u}(l)$ is termed return map or stroboscopic map.

The 2×2 matrix \mathbf{A} is real, because real initial conditions $\mathbf{u}(0)$ have real solutions $\mathbf{u}(x)$.

$$\mathbf{A} = \mathbf{A}^* = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}, \quad \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}. \quad (6.50)$$

The matrix elements of \mathbf{A} are differentiable functions of the energy E , because the solution $\psi(x)$ and its derivative at $x=l$ are differentiable functions of the parameter E of the differential equation.

From the eigenvalue equation (6.17) one easily concludes that the quantum mechanical current j , the Wronski determinant, does not depend on x

$$\begin{aligned} \frac{2mi}{\hbar} j &= \psi^* \overleftrightarrow{\partial}_x \psi = \psi^* \partial_x \psi - \partial_x \psi^* \psi = \mathbf{u}^\dagger(x) \mathbf{I} \mathbf{u}(x), \quad \text{where } \mathbf{I} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ \partial_x (\mathbf{u}^\dagger(x) \mathbf{I} \mathbf{u}(x)) &= 0. \end{aligned} \quad (6.51)$$

In particular the value of $\mathbf{u}^\dagger(x) \mathbf{I} \mathbf{u}(x)$ at $x=0$ coincides with the value at $x=l$. Therefore for all initial conditions $\mathbf{u} = \mathbf{u}(0)$ one has

$$(\mathbf{A}\mathbf{u})^\dagger \mathbf{I} \mathbf{A}\mathbf{u} = \mathbf{u}^\dagger \mathbf{I} \mathbf{u} \quad \forall \mathbf{u} \quad (6.52)$$

and therefore

$$\mathbf{A}^\dagger \mathbf{I} \mathbf{A} = \mathbf{I}. \quad (6.53)$$

This matrix relation is satisfied for real 2×2 matrices \mathbf{A} if and only if the determinant has the special value 1.

$$\mathbf{ad} - \mathbf{bc} = 1 \quad (6.54)$$

The matrix \mathbf{A} is an element of the group of special linear transformations of a two dimensional real vector space.

$$\mathbf{A} \in \text{SL}(2, \mathbb{R}) \quad (6.55)$$

The eigenvalues of \mathbf{A}

$$\lambda_{1,2} = \frac{\mathbf{a} + \mathbf{d}}{2} \pm \sqrt{\left(\frac{\mathbf{a} + \mathbf{d}}{2}\right)^2 - 1} \quad (6.56)$$

are real, if $|\text{tr}\mathbf{A}| = |\mathbf{a} + \mathbf{d}| \geq 2$. Because of $\det \mathbf{A} = 1$ these eigenvalues are inverse to each other and the modulus of one of the real eigenvalues is larger or equal to 1.

$$|\text{tr}\mathbf{A}| \geq 2 \Rightarrow \lambda_1 = \lambda_1^* = \frac{1}{\lambda_2} \quad (6.57)$$

Solutions $\mathbf{u}(x)$ which correspond to eigenvalues with $|\lambda| > 1$ grow for $x \rightarrow \infty$ exponentially $\mathbf{u}(x + n\mathbf{l}) = \lambda^n \mathbf{u}(x)$. The solution corresponding to the other eigenvalue $\lambda_2 = \lambda^{-1}$ grows according to $\mathbf{u}_2(x - n\mathbf{l}) = \lambda_2^{-n} \mathbf{u}_2(x)$ for $x \rightarrow -\infty$. One cannot combine these exponentially growing solutions to normalizable wave packets.

If the modulus of the trace of A is smaller than 2, then the eigenvalues are complex and, because A is real, one eigenvalue is the complex conjugate of the other. Because their product is $\det A = 1$, the modulus of the complex eigenvalues is 1

$$|\text{tr}A| < 2 \Rightarrow \lambda_1 = \lambda_2^*, \quad |\lambda_1| = 1. \quad (6.58)$$

The periodicity condition (6.47) states that the function $\psi(x)$ belongs to eigenvectors of the matrix A with complex eigenvalues $e^{ik\mathbf{l}}$ with modulus 1.

$$(A - e^{ik\mathbf{l}}) \begin{pmatrix} \psi(0) \\ \psi'(0) \end{pmatrix} = 0. \quad (6.59)$$

This restricts the energy E to bands which in our one dimensional system is determined by $|\text{tr}A(E)| \leq 2$ such that the eigenvalues of A lie on the unit circle in the complex plane.

In the neighbourhood of the band edge, e.g. where $\text{tr}A = 2$ holds, the matrix A has the form

$$A = \begin{pmatrix} a & b \\ c & 2-a \end{pmatrix} + \delta E \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \quad (6.60)$$

Here δE is the deviation of the energy from the band edge, the matrix elements b and c are restricted by $\det A = 1$ or $bc = -(1-a)^2$, and α, β, γ and δ are the derivatives of the matrix elements a, b, c and d . If $|\text{tr}A| - 2$ changes sign and if $\frac{d}{dE} \text{tr}A = \alpha + \delta \neq 0$, then the eigenvalues of A vary in the neighbourhood of the band edge in lowest order in $(\delta E)^{\frac{1}{2}}$ as

$$\lambda_{1,2} \approx \text{sign}(\text{tr}A) \pm \sqrt{\delta E (\alpha + \delta) \text{sign}(\text{tr}A)}. \quad (6.61)$$

At the lower band edge one has $(\alpha + \delta) \text{sign}(\text{tr}A) < 0$ and energies above the lower band edge lead to complex eigenvalues $e^{\pm ik\mathbf{l}} \approx 1 \pm ik\mathbf{l}$. If here one solves for the energy as a function of k then one obtains in lowest order

$$E(k) = E(0) + \frac{\hbar^2 k^2}{2M} + \dots \quad \text{where } M = \frac{\hbar^2}{2l^2} \left| \frac{d}{dE} \text{tr}A \right|. \quad (6.62)$$

This is the energy momentum relation of a particle with an effective mass M .

If the value of $|k|$ increases together with the energy in an allowed band then for $k = \pm \frac{\pi}{\mathbf{l}}$ the upper band edge is reached. If $|\text{tr}A| - 2$ changes sign and if $\alpha + \delta \neq 0$, then the derivative $\frac{dE}{dk}$ vanishes and the second derivative $\frac{d^2E}{dk^2}$ is negative.

$$\left. \frac{dE}{dk} \right|_{k=\pm\frac{\pi}{\mathbf{l}}} = 0, \quad \left. \frac{d^2E}{dk^2} \right|_{k=\pm\frac{\pi}{\mathbf{l}}} = -2l^2 \left| \frac{d}{dE} \text{tr}A \right|^{-1}. \quad (6.63)$$

Within each band $k(E)$ is a monotonous function. To see this, one writes the eigenfunction of (6.17) and (6.47) as a product of e^{ikx} and a periodic function $\mathbf{u}_k(x + \mathbf{l}) = \mathbf{u}_k(x)$.

$$\psi_k(x) = \sqrt{\frac{\mathbf{l}}{2\pi}} e^{ikx} \mathbf{u}_k(x) \quad (6.64)$$

The eigenvalue equation then reads

$$H(k)\mathbf{u}_k = \left(E - \frac{\hbar^2 k^2}{2m} \right) \mathbf{u}_k, \quad H(k) = \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{\hbar k}{m} \left(-i\hbar \frac{d}{dx} \right) + V(x). \quad (6.65)$$

The Hamilton operator $H(k)$ acts as a hermitean operator on \mathbf{l} -periodic functions \mathbf{u} and \mathbf{v} which have a scalar product

$$\langle \mathbf{u} | \mathbf{v} \rangle = \int_0^{\mathbf{l}} dx \mathbf{u}^*(x) \mathbf{v}(x), \quad (6.66)$$

these are position wave functions on a circle with circumference \mathbf{l} .

Within a band $k(E)$ is a differentiable function, for k is a differentiable function (6.56) of the matrix elements of A which in turn depend in a differentiable way on E . If we differentiate the eigenvalue $E - \frac{\hbar^2 k^2}{2m}$ of $H(k)$ with respect to E then, because of (4.80), we obtain for normalized \mathbf{u}_k

$$P\mathbf{u}_k(x) = -i\hbar \frac{d}{dx} \mathbf{u}_k(x) \quad (6.67)$$

$$1 - \frac{\hbar^2 k}{m} \frac{dk}{dE} = \langle \mathbf{u}_k | \frac{dH(k)}{dk} \mathbf{u}_k \rangle \frac{dk}{dE} = \langle \mathbf{u}_k | P\mathbf{u}_k \rangle \frac{\hbar}{m} \frac{dk}{dE}. \quad (6.68)$$

This equation excludes zeros of $\frac{dk}{dE}$ because the matrix element $\langle \mathbf{u}_k | P\mathbf{u}_k \rangle$ is finite. Therefore $k(E)$ is invertible within each energy band and the energy is a strictly monotonous function of k between $k = 0$ and $k = \frac{\pi}{\mathbf{l}}$.

The group velocity of wave packets

$$v_{\text{group}} = \frac{\partial \omega}{\partial k} = \frac{1}{\hbar} \frac{\partial E}{\partial k} = \frac{1}{m} (\hbar k + \langle \mathbf{u}_k | P\mathbf{u}_k \rangle) \quad (6.69)$$

is a combination of the momentum carried by the factor e^{ikx} and the expectation value of the momentum within a periodicity interval. Near $k = 0$ it is proportional to k

$$\langle \mathbf{u}_k | P\mathbf{u}_k \rangle = \hbar k \left(\frac{m}{M} - 1 \right) + O(k^2), \quad (6.70)$$

at $k = \pm \frac{\pi}{\mathbf{l}}$ it compensates the momentum of e^{ikx}

$$\langle \mathbf{u}_{\frac{\pi}{\mathbf{l}}} | P\mathbf{u}_{\frac{\pi}{\mathbf{l}}} \rangle = -\langle \mathbf{u}_{-\frac{\pi}{\mathbf{l}}} | P\mathbf{u}_{-\frac{\pi}{\mathbf{l}}} \rangle = -\hbar \frac{\pi}{\mathbf{l}}. \quad (6.71)$$

The functions \mathbf{u}_k are periodic and can be written as a Fourier series (5.25)

$$\mathbf{u}_k(x) = \sum_n c_n e^{in \frac{2\pi}{\mathbf{l}} x} \quad (6.72)$$

The corresponding eigenfunctions $\psi_k(x) = \sqrt{\frac{l}{2\pi}} e^{ikx} \mathbf{u}_k(x)$ with $-\frac{\pi}{l} \leq k \leq \frac{\pi}{l}$ are therefore continuum normalized on the real line if the wave functions \mathbf{u}_k are orthonormalized in the periodicity interval.

$$\begin{aligned} \langle \Psi_k | \Psi_{k'} \rangle &= \frac{l}{2\pi} \int dx \left(e^{ikx} \sum_n c_n e^{in\frac{2\pi}{l}x} \right)^* \left(e^{ik'x} \sum_m c'_m e^{im\frac{2\pi}{l}x} \right) \\ &= l \sum_{m,n} c_n^* c'_m \delta(k' - k + (m-n)\frac{2\pi}{l}) \\ &= (l \sum_n c_n^* c'_n) \delta(k' - k) = \langle \mathbf{u}_k | \mathbf{u}'_k \rangle \delta(k' - k) . \end{aligned} \quad (6.73)$$

If the wave functions belong to the same band then \mathbf{u}_k and \mathbf{u}'_k agree and from $\langle \mathbf{u}_k | \mathbf{u}_k \rangle = 1$ one concludes

$$\langle \Psi_k | \Psi_{k'} \rangle = \delta(k' - k) . \quad (6.74)$$

If the wave functions belong to different bands then \mathbf{u}_k and \mathbf{u}'_k are orthogonal because they solve the eigenvalue equation (6.65) of an hermitean operator with different eigenvalues $E - \frac{\hbar^2 k^2}{2m}$.

If $|\text{tr}A| - 2$ changes sign at the edge of a band then a gap separates this band from the next. This gap vanishes if $|\text{tr}A| = 2$ has a local maximum.

To each energy within the band there correspond two eigenvalues e^{ikl} and e^{-ikl} of \mathbf{U}_l . The dispersion relation $E(k) = E(-k)$ is therefore an even function of k .

To each eigenvalue $e^{ikl} \neq \pm 1$ of the translation \mathbf{U}_l (6.47) there correspond denumerably many energy eigenstates with non-degenerate energies. These eigenstates belong to the different bands. At the eigenvalue $e^{ikl} = \pm 1$ the energy is degenerate if and only if $A(E) = \pm \mathbb{1}$.

7 Composite systems

7.1 Product space

Often a quantum mechanical system is composed of identifiable parts, for example of two different particles whose properties can be measured separately. Then the Hilbert space \mathcal{H} is a product space

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \quad (7.1)$$

and has an orthonormal basis $\Lambda_{i,\alpha} = \Lambda_i \otimes \Lambda_\alpha$, where i counts an orthonormal basis Λ_i of \mathcal{H}_1 and α an orthonormal basis Λ_α of \mathcal{H}_2 . States Ψ in the product space have components

$$\psi_{i,\alpha} = \langle \Lambda_{i,\alpha} | \Psi \rangle \quad (7.2)$$

If the components $\psi_{i,\alpha}$ can be written as a product $u_i v_\alpha$, i.e. if the matrix $\psi_{i,\alpha}$ has rank 1, then $\Psi = \mathbf{u} \otimes \mathbf{v}$ is a product state with $\mathbf{u} = \sum_i \Lambda_i u_i$ and $\mathbf{v} = \sum_\alpha \Lambda_\alpha v_\alpha$. More generally, states in product spaces are sums of product vectors. States that are not product states are called entangled states. All many-particle states of identical bosons or fermions which are composed of different one-particle states, such as Slater determinants (5.65) for instance, are entangled.

The index i counts possible results \mathbf{a}_i of a measurement of the first subsystem, the index α counts results \mathbf{b}_α of the second subsystem. Let the operator \mathbf{A} represent a measuring instrument which measures the first subsystem, i.e. which maps \mathcal{H}_1 onto \mathcal{H}_1 , and let \mathbf{B} represent a device to measure the second subsystem, then their tensor product $\mathbf{A} \otimes \mathbf{B}$ is defined by the matrix elements

$$(\mathbf{A} \otimes \mathbf{B})_{i,\alpha \ j,\beta} = A_{ij} B_{\alpha\beta} \quad (7.3)$$

They act on product states $\mathbf{u} \otimes \mathbf{v}$ by

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{u} \otimes \mathbf{v}) = (\mathbf{A}\mathbf{u}) \otimes (\mathbf{B}\mathbf{v}) \quad (7.4)$$

and are declared on sums of product states, i.e. on entangled states, by linearity.

Operators of the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ of the form $\mathbf{A} \otimes \mathbb{1}$ correspond to the measurements of the first subsystem, the operators of the form $\mathbb{1} \otimes \mathbf{B}$ to the measurements of the second subsystem.

$$(\mathbf{A} \otimes \mathbb{1})\Lambda_{i,\alpha} = \mathbf{a}_i \Lambda_{i,\alpha} \quad \forall \alpha, \quad (\mathbb{1} \otimes \mathbf{B})\Lambda_{i,\alpha} = \mathbf{b}_\alpha \Lambda_{i,\alpha} \quad \forall i . \quad (7.5)$$

The results \mathbf{a}_i and \mathbf{b}_α are degenerate. The degree of degeneracy of each eigenvalue \mathbf{a}_i of $\mathbf{A} \otimes \mathbb{1}$ is a multiple of the dimension of \mathcal{H}_2 , the degree of degeneracy of the eigenvalues \mathbf{b}_α of $\mathbb{1} \otimes \mathbf{B}$ is a multiple of the dimension of \mathcal{H}_1 .

7.2 Addition of angular momentum

Let us consider the quantum mechanical system which is formed by two spin-1/2 particles that, for simplicity, cannot move. The Hilbert space of the one-particle states is then spanned simply by basis states Λ_{\uparrow} and Λ_{\downarrow} . A basis of the product space of the two-particle states is

$$\Lambda_{\uparrow\uparrow}, \Lambda_{\uparrow\downarrow}, \Lambda_{\downarrow\uparrow}, \Lambda_{\downarrow\downarrow}. \quad (7.6)$$

The basis is chosen such that the spin operators for the first and the second particle \vec{S}_1 and \vec{S}_2 act by multiplication with the Pauli matrices.

$$\vec{S}_1 \Lambda_{ij} = \frac{\hbar}{2} \Lambda_{kj} \vec{\sigma}_{ki}, \quad \vec{S}_2 \Lambda_{ij} = \frac{\hbar}{2} \Lambda_{ik} \vec{\sigma}_{kj}. \quad (7.7)$$

All spin operators of the first particle commute with all spin operators of the second particle.

$$[S_{1a}, S_{2b}] = 0 \quad a, b \in \{1, 2, 3\} \quad (7.8)$$

Hence, the sums $S_a = S_{1a} + S_{2a}$ are components of angular momentum operators that satisfy the angular momentum algebra (4.45).

$$[S_a, S_b] = i\hbar \epsilon_{abc} S_c \quad (7.9)$$

The spectrum of $S_3 = S_{13} + S_{23}$ can be read off immediately

$$S_3 \Lambda_{\uparrow\uparrow} = \hbar \Lambda_{\uparrow\uparrow}, \quad S_3 \Lambda_{\uparrow\downarrow} = 0 \hbar \Lambda_{\uparrow\downarrow}, \quad S_3 \Lambda_{\downarrow\uparrow} = 0 \hbar \Lambda_{\downarrow\uparrow}, \quad S_3 \Lambda_{\downarrow\downarrow} = -\hbar \Lambda_{\downarrow\downarrow}. \quad (7.10)$$

Accordingly, $\Lambda_{\uparrow\uparrow}$ belongs to an angular momentum multiplet with total spin $s = 1$, because the total spin s can be read off the highest S_3 -eigenvalue (4.59). Similarly, the state $\Lambda_{\downarrow\downarrow}$ with lowest S_3 -eigenvalue $-\hbar$ belongs to total spin $s = 1$. The state with $s = 1$ and S_3 -eigenvalue 0 is obtained with a factor $\sqrt{1(1+1) - 1(1-1)} = \sqrt{2}$ (4.48) by applying the ladder operator S_- to $\Lambda_{\uparrow\uparrow}$

$$S_- \Lambda_{\uparrow\uparrow} = (S_{1-} + S_{2-}) \Lambda_{\uparrow\uparrow} = \sqrt{\frac{1}{2}(\frac{1}{2} + 1) - \frac{1}{2}(\frac{1}{2} - 1)} (\Lambda_{\downarrow\uparrow} + \Lambda_{\uparrow\downarrow}) \quad (7.11)$$

$$\Lambda_{s=1, s_3=1} = \Lambda_{\uparrow\uparrow}, \quad \Lambda_{s=1, s_3=0} = \frac{1}{\sqrt{2}} (\Lambda_{\downarrow\uparrow} + \Lambda_{\uparrow\downarrow}), \quad \Lambda_{s=1, s_3=-1} = \Lambda_{\downarrow\downarrow}. \quad (7.12)$$

These three basis vectors span a total spin-1 multiplet. As eigenstates of S^2 , further total angular momentum multiplets are orthogonal to this total spin-1 multiplet if they belong to different total spin, or can be chosen orthogonal to this total spin-1 multiplet. In our example they are therefore spanned by the state

$$\Lambda_{s=0, s_3=0} = \frac{1}{\sqrt{2}} (\Lambda_{\downarrow\uparrow} - \Lambda_{\uparrow\downarrow}). \quad (7.13)$$

It is an S_3 -eigenstate belonging to the eigenvalue 0. Since the total spin in an angular momentum multiplet can be read off the highest and the lowest S_3 -eigenvalue, this state belongs to total spin 0. This one can easily check, for S_3 , S_+ and S_- vanish on this state.

The states $\Lambda_{s=0, s_3=0}$ and $\Lambda_{s=1, s_3=0}$ are entangled.

7.3 Independent composite systems

Mixtures ρ of composite systems are composed independently, if for all measurements $A \otimes \mathbb{1}$ of the first subsystem and for all measurements $\mathbb{1} \otimes B$ of the second subsystem the probabilities factorize with which one obtains the i -th result \mathbf{a}_i of device A and the α -th result \mathbf{b}_α of device B

$$w(i, \alpha, A, B, \rho) = w_1(i, A, \hat{\rho}) \cdot w_2(\alpha, B, \tilde{\rho}). \quad (7.14)$$

Here $w_1(i, A, \hat{\rho}) = \sum_\alpha w(i, \alpha, A, B, \rho)$ is the probability of measuring the i -th result \mathbf{a}_i with device A , $w_2(\alpha, B, \tilde{\rho}) = \sum_i w(i, \alpha, A, B, \rho)$ is the probability of measuring the α -th result with device B and $w(i, \alpha, A, B, \rho)$ is the probability of measuring both \mathbf{a}_i and \mathbf{b}_α . If a mixture is not composed independently, we call the subsystems correlated.

Let us analyze what it means mathematically that systems are composed independently. The probabilities are given by main diagonal elements $\langle \Lambda | \rho \Lambda \rangle$ (3.1). The condition (7.14) must hold for all Λ that are eigenstates of an operator $A \otimes \mathbb{1}$ and an operator $\mathbb{1} \otimes B$, thus (7.14) must hold for all Λ that are product states $\mathbf{u} \otimes \mathbf{v}$.

$$\mathbf{u}_i^* \mathbf{v}_\alpha^* \rho_{i, \alpha} \mathbf{u}_j \mathbf{v}_\beta = \mathbf{u}_i^* \left(\sum_\alpha \rho_{i, \alpha} \mathbf{u}_j \right) \cdot \mathbf{v}_\alpha^* \left(\sum_i \rho_{i, \alpha} \mathbf{v}_\beta \right) \quad (7.15)$$

Both sides of the equation are bilinear forms in \mathbf{u} and in \mathbf{v} and are equal for all \mathbf{u} and \mathbf{v} if and only if the mixture is a tensor product of mixtures.

$$\rho_{\text{independent}} = \hat{\rho} \otimes \tilde{\rho} \quad (7.16)$$

$$\rho_{i, \alpha} \mathbf{u}_j \mathbf{v}_\beta = (\hat{\rho})_{ij} (\tilde{\rho})_{\alpha\beta} \quad \text{where} \quad (\hat{\rho})_{ij} = \sum_\alpha \rho_{i, \alpha} \mathbf{u}_j \quad \text{and} \quad (\tilde{\rho})_{\alpha\beta} = \sum_i \rho_{i, \alpha} \mathbf{v}_\beta \quad (7.17)$$

Only for independently composed systems there are no correlations of probabilities for results of measurements of the first and of the second subsystem. Then one can restrict oneself to one subsystem and analyze its properties isolated from the second subsystem. For instance in picture (1.1) one first has to check whether the combination of two particles in the beam to a two-particle system does not exhibit correlations which are not observed if one presupposes that the beam just contains repeatedly prepared one-particle states.

If the systems are not composed independently, the measurement results of the first and second subsystem are correlated. In the extreme case one can predict after a measurement of the first subsystem the result of the measurement of the second subsystem.

In the formulation ‘‘By measuring the first subsystem the second measurement is pre-determined’’ such a correlation seems paradoxical, in particular if the result of the first measurement is unknown at the second device and could not have any effect. The denomination ‘‘Quantum teleportation’’ suggests a non-classical faster than light transmission of quantum properties. The denomination insinuates that the first measurement is the cause of the second measured value. In fact even in retrospect, long after the first and second measurement, one can derive from the protocol of the one measurement what has been measured at the other device and vice versa.

State reduction is the transition from probabilities to conditional probabilities. It happens instantly and also in retrospect, as soon as the outcome of the first measurement is known. The first measurement does not cause the second measured value, however. The correlation of the measured values results since the two subsystems have not been prepared independently.

The claim that the first measurement determines the result of the second, contradicts quantum mechanics. Quantum mechanics does not know of a reason for the probability distribution (1.1) of the results of a measurement. If there were a reason and if the results of the measurement were predetermined as in classical physics, then the state Ψ would be an incomplete description of the physical system that is to be measured.

Among the surprising properties of correlated composite systems is the fact that a pure entangled state Ψ appears to be mixed if one measures only the first subsystem. Incomplete measurement has the same effect as ignorance of the system that is to be measured.

$$w_1(i, A, \Psi) = \sum_{\alpha} w(i, \alpha, A, B, \Psi) = \sum_{\alpha} |\langle \Lambda_{i,\alpha} | \Psi \rangle|^2 = \sum_{\alpha} |\psi_{i,\alpha}|^2 \quad (7.18)$$

So for all measurements of the first subsystem one obtains probability distributions as from the density matrix

$$(\hat{\rho})_{ij} = \sum_{\alpha} \psi_{i,\alpha} \psi_{j,\alpha}^* \quad (7.19)$$

This density matrix corresponds to a pure state \mathbf{u} of the first subsystem

$$\sum_{\alpha} \psi_{i,\alpha} \psi_{j,\alpha}^* = \mathbf{u}_i \mathbf{u}_j^* , \quad (7.20)$$

only if Ψ is a product state. This one concludes as follows. Equation (7.20) implies with the definition $\mathbf{v}_{\beta} = \sum_i \mathbf{u}_i^* \psi_{i\beta}$ that \mathbf{u} and \mathbf{v} are normalized because Ψ is normalized. In addition it follows from (7.20) that $\sum_{\beta} \psi_{i\beta} \mathbf{v}_{\beta}^* = \mathbf{u}_i$ and

$$\sum_{\beta} (\psi_{i\beta} - \mathbf{u}_i \mathbf{v}_{\beta}) (\psi_{j\beta}^* - \mathbf{u}_j^* \mathbf{v}_{\beta}^*) = 0 . \quad (7.21)$$

For $i = j$ this is a sum of squares. It vanishes only if $\psi_{i\alpha} = \mathbf{u}_i \mathbf{v}_{\alpha}$ holds, i.e. if $\Psi = \mathbf{u} \otimes \mathbf{v}$ is a product state.

7.4 Bell Inequalities

Bell inequalities are bounds for expectation values which follow from the seemingly irrefutable supposition of classical physics that the results of all measurements of single particles in a two particle system are in principle determined by hidden local parameters. In such a picture the outcome of a measurement is uncertain because one does not know the hidden parameters and the measured results of one particle do not depend on the

kind of measurement performed on the other particle because the parameters have only local effects.

We consider two spin-1/2 particles which are generated by the decay of a spin-0 particle. We measure in a series of measurements the individual spins in several directions and denote with $\frac{\hbar}{2} \mathbf{a}_{1i}$ the resulting spin of the first particle in the i -th measurement, if we measure in the direction $\vec{\mathbf{a}}$. With $\frac{\hbar}{2} \mathbf{c}_{2j}$ we denote the result of the spin measurement of the second particle in the j -th experiment, if there we measure in direction $\vec{\mathbf{c}}$. Because the particles are generated by the decay of a spin-0 particle the resulting spin of the second particle is opposite to the spin of the first particle, for all directions $\vec{\mathbf{b}}$ and in all experiments one has $\mathbf{b}_{1i} = -\mathbf{b}_{2i}$.

Because the results \mathbf{a}_{1i} , \mathbf{b}_{2i} and \mathbf{c}_{2i} take only values 1 or -1 one has in all cases, irrespective whether $\mathbf{b}_{2i} = \mathbf{c}_{2i}$ or $\mathbf{b}_{2i} = -\mathbf{c}_{2i}$,

$$\mathbf{a}_{1i} (\mathbf{b}_{2i} - \mathbf{c}_{2i}) \leq 1 - \mathbf{b}_{2i} \mathbf{c}_{2i} = 1 + \mathbf{b}_{1i} \mathbf{c}_{2i} . \quad (7.22)$$

The inequality makes only use of the supposition that for each experiment and for all directions $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ there exist the corresponding results \mathbf{a}_{1i} , \mathbf{b}_{2i} and \mathbf{c}_{2i} and that these results are do not depend on the direction which is chosen for measuring the other particle.

If we sum the inequalities for N experiments and if we divide by N then in the limit $N \rightarrow \infty$ we obtain a Bell inequality for the expectation values

$$\langle \mathbf{A}_1 \mathbf{B}_2 \rangle - \langle \mathbf{A}_1 \mathbf{C}_2 \rangle \leq 1 + \langle \mathbf{B}_1 \mathbf{C}_2 \rangle . \quad (7.23)$$

Here \mathbf{A}_1 is the device with reading 1 if the spin of the first particle is up in direction $\vec{\mathbf{a}}$ and -1 otherwise. The device \mathbf{C}_2 shows 1 if the spin of the second particle, measured in direction $\vec{\mathbf{c}}$, is up and so on.

In quantum mechanics the individual results \mathbf{a}_{1i} and \mathbf{b}_{1i} are not determined in each case. They have to be measured in different experiments \mathbf{a}_{1i} and \mathbf{b}_{1j} with $i \neq j$. The quantum mechanical expectation values are therefore not subject by Bell inequalities.

In fact, in the spin-0-state (7.13)

$$\Psi = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \quad (7.24)$$

the result that the spin of the first particle is up and the spin of the second particle is up in direction (θ, φ) (4.67) occurs with probability

$$w(\uparrow\uparrow_{(\theta, \varphi)}) = |\langle \uparrow\uparrow_{(\theta, \varphi)} | \Psi \rangle|^2 = |\cos \frac{\theta}{2} e^{i\frac{\varphi}{2}} \langle \uparrow\uparrow | \Psi \rangle + \sin \frac{\theta}{2} e^{-i\frac{\varphi}{2}} \langle \uparrow\downarrow | \Psi \rangle|^2 = \frac{1}{2} \sin^2 \frac{\theta}{2} . \quad (7.25)$$

More generally both spins are found up in direction $\vec{\mathbf{a}}$ for the first particle and $\vec{\mathbf{b}}$ for the second particle with probability

$$w(\uparrow_{\vec{\mathbf{a}}} \uparrow_{\vec{\mathbf{b}}}) = \frac{1}{2} \sin^2 \frac{\beta}{2} , \quad (7.26)$$

where β is the angle enclosed by \vec{a} and \vec{b} . For the state Ψ is invariant under rotations and $|\uparrow_{\vec{a}}\uparrow_{\vec{b}}\rangle$ can be rotated by a unitary transformation into the state $|\uparrow_{(\theta=\beta, \varphi=0)}\rangle$ in which the first spin points up in z -direction and the second in direction $(\theta = \beta, \varphi = 0)$.

This probability yields the expectation value

$$\begin{aligned} \langle A_1 B_2 \rangle &= w(\uparrow_{\vec{a}}\uparrow_{\vec{b}}) - w(\downarrow_{\vec{a}}\uparrow_{\vec{b}}) - w(\uparrow_{\vec{a}}\downarrow_{\vec{b}}) + w(\downarrow_{\vec{a}}\downarrow_{\vec{b}}) = \\ &= \frac{1}{2} \sin^2 \frac{\beta}{2} - \frac{1}{2} \cos^2 \frac{\beta}{2} - \frac{1}{2} \cos^2 \frac{\beta}{2} + \frac{1}{2} \sin^2 \frac{\beta}{2} = -\cos \beta . \end{aligned} \quad (7.27)$$

The combination of expectation values

$$\langle A_1 B_2 \rangle - \langle A_1 C_2 \rangle - \langle B_1 C_2 \rangle = -\cos \beta + \cos \gamma + \cos(\beta - \gamma) , \quad (7.28)$$

where γ denotes the angle between \vec{a} and \vec{c} and $\beta - \gamma$ the angle between \vec{b} and \vec{c} , takes the value $\frac{3}{2}$ for $\beta = \frac{2\pi}{3}$ and $\gamma = \frac{\pi}{3}$ and violates the Bell inequality (7.23). Quantum mechanics is not compatible with the supposition that the results of each measurement is determined in principle by unknown parameters with local effects.

That quantum mechanical expectation values coincide with experimental results and violate Bell inequalities shatters classical views of physics. The physical probability distributions cannot be caused by hidden, unknown parameters which have only local effects, i.e., which determine each individual result and do not depend on which kind of measurement is performed elsewhere.

8 Basics of Thermodynamics

8.1 Entropy

Lack of polarization (4.74) or the variable $1 - (\text{tr } \rho^2)$ can be used as a measure for how much the prepared state differs from a pure state. A better suited measure of the lack of knowledge about the prepared state is the entropy S . The entropy is additive, if two independent systems are combined, it stays constant during the time evolution according to the Schrödinger equation and increases by mixing and by random perturbations.

Lack of knowledge or entropy is defined as a function of the eigenvalues ρ_i of the density matrix

$$S = - \sum_i \rho_i \ln \rho_i = - \text{tr } \rho \ln \rho . \quad (8.1)$$

The function $x \ln x$ is continuously completed for $x = 0$ and vanishes there. The entropy is non-negative because the eigenvalues ρ_i of the density matrix vary between 0 and 1. The entropy of a pure state vanishes.

If a mixture is composed of two independent parts

$$\rho = \hat{\rho} \otimes \tilde{\rho} , \quad (8.2)$$

then the eigenstates are product states $\hat{Y}_i \otimes \tilde{Y}_j$ of eigenstates of the density matrices $\hat{\rho}$ and $\tilde{\rho}$. The eigenvalues of the composite density matrix are products of the eigenvalues of the separate density matrices.

$$\rho_{ij} = \hat{\rho}_i \tilde{\rho}_j , \quad \sum_i \hat{\rho}_i = 1 , \quad \sum_j \tilde{\rho}_j = 1 . \quad (8.3)$$

Therefore the entropy of independently composed systems is the sum of the entropy of its parts.

$$\begin{aligned} S &= - \sum_{ij} \hat{\rho}_i \tilde{\rho}_j \ln(\hat{\rho}_i \tilde{\rho}_j) = - \sum_{ij} (\hat{\rho}_i \tilde{\rho}_j \ln \hat{\rho}_i + \hat{\rho}_i \tilde{\rho}_j \ln \tilde{\rho}_j) \\ &= - (\sum_j \tilde{\rho}_j) \sum_i \hat{\rho}_i \ln \hat{\rho}_i - (\sum_i \hat{\rho}_i) \sum_j \tilde{\rho}_j \ln \tilde{\rho}_j = \hat{S} + \tilde{S} \end{aligned} \quad (8.4)$$

The density matrix changes in the course of time (6.6), its eigenvalues ρ_i , however, do not change on account of the Schrödinger equation. If one adapts the notation then (4.80) and (6.6) and the eigenvalue equation $\rho \Upsilon_i = \rho_i \Upsilon_i$ imply $\partial_t \rho_i(t) = 0$, because in

eigenstates of an hermitean operator ρ the expectation value of each commutator $[H, \rho]$ vanishes.

$$i\hbar\partial_t\rho_i(t) = i\hbar\langle\Upsilon_i|\partial_t\rho(t)\Upsilon_i\rangle = \langle\Upsilon_i|(H\rho - \rho H)\Upsilon_i\rangle = \langle\Upsilon_i|(H\rho_i - \rho_i H)\Upsilon_i\rangle = 0 \quad (8.5)$$

So the entropy does not change in the course of the time evolution of the states which is generated by the Hamilton operator.

The entropy increases by true mixing. If a mixture $\rho(\lambda)$ is mixed out of different mixtures $\hat{\rho}$ and $\tilde{\rho}$, $\hat{\rho} \neq \tilde{\rho}$

$$\rho(\lambda) = \lambda\hat{\rho} + (1-\lambda)\tilde{\rho} \quad \text{with} \quad 0 < \lambda < 1, \quad (8.6)$$

then the entropy $S(\rho(\lambda))$ is larger than the proportionate sum by the entropy generated by mixing

$$S(\rho(\lambda)) > \lambda S(\hat{\rho}) + (1-\lambda)S(\tilde{\rho}). \quad (8.7)$$

Before we prove this assertion we make two remarks: If $\hat{\rho} \neq \tilde{\rho}$ and $\lambda \neq \lambda'$ then one also has $\rho(\lambda) \neq \rho(\lambda')$. In the range $0 < \lambda < 1$ the kernel, the space which is mapped to 0, of $\rho(\lambda)$ does not depend on λ .

The second remark holds because $\hat{\rho}$ and $\tilde{\rho}$ and therefore also $\rho(\lambda')$ vanish in the kernel of $\rho(\lambda)$ for each chosen λ .

$$\begin{aligned} \langle\Lambda|(\lambda\hat{\rho} + (1-\lambda)\tilde{\rho})\Lambda\rangle = 0 &\Rightarrow \langle\Lambda|\hat{\rho}\Lambda\rangle = 0 \text{ and } \langle\Lambda|\tilde{\rho}\Lambda\rangle = 0 \\ &\Rightarrow \hat{\rho}\Lambda = 0 \text{ and } \tilde{\rho}\Lambda = 0 \end{aligned}$$

The first implication follows because λ and $(1-\lambda)$ are positive and main diagonal elements of density matrices are non-negative (3.5). The second implication holds because main diagonal elements $\langle\Lambda|\rho\Lambda\rangle$ of a density matrix ρ vanish only if $\rho\Lambda$ vanishes (3.6).

After these preparatory remarks we prove for different density matrices ρ and ρ' , $\rho \neq \rho'$, if ρ' has only non-vanishing eigenvalues, the inequality

$$\text{tr } \rho \ln \rho' < \text{tr } \rho \ln \rho. \quad (8.8)$$

To show the inequality we evaluate the trace in the eigenbasis Υ_i of ρ and insert a resolution of the identity with the eigenbasis Υ'_i of ρ' .

$$\begin{aligned} \text{tr } \rho (\ln \rho' - \ln \rho) &= \sum_{ij} \langle\Upsilon_i|\rho\Upsilon'_j\rangle \langle\Upsilon'_j|(\ln \rho' - \ln \rho)\Upsilon_i\rangle \\ &= \sum_{ij} |\langle\Upsilon_i|\Upsilon'_j\rangle|^2 \rho_i (\ln \rho'_j - \ln \rho_i) \\ &= \sum_{ij} |\langle\Upsilon_i|\Upsilon'_j\rangle|^2 \rho_i \ln \frac{\rho'_j}{\rho_i}. \end{aligned}$$

For positive x the inequality $\ln x \leq (x-1)$ is strict if $x \neq 1$.

If the matrices ρ and ρ' are different then there exists at least one pair of eigenvalues ρ'_j and ρ_i , $\rho'_j \neq \rho_i$, with eigenvectors Υ'_j and Υ_i which have a non-vanishing scalar product. This implies the inequality we want to show.

$$\begin{aligned} \text{tr } \rho (\ln \rho' - \ln \rho) &< \sum_{ij} |\langle\Upsilon_i|\Upsilon'_j\rangle|^2 \rho_i \left(\frac{\rho'_j}{\rho_i} - 1\right) = \sum_{ij} |\langle\Upsilon_i|\Upsilon'_j\rangle|^2 (\rho'_j - \rho_i) = \\ &= \sum_{ij} \langle\Upsilon'_j|\rho'\Upsilon_i\rangle \langle\Upsilon_i|\Upsilon'_j\rangle - \sum_{ij} \langle\Upsilon'_j|\rho\Upsilon_i\rangle \langle\Upsilon_i|\Upsilon'_j\rangle = \text{tr } \rho' - \text{tr } \rho = 0. \end{aligned}$$

If $\hat{\rho} \neq \tilde{\rho}$ and $0 < \lambda < 1$ holds then one also has $\hat{\rho} \neq \rho(\lambda) \neq \tilde{\rho}$. We use the inequality (8.8) for the density matrix $\rho(\lambda)$ in place of ρ' and $\hat{\rho}$ and $\tilde{\rho}$ in place of ρ . We take the trace over the space orthogonal to the kernel of $\rho(\lambda)$.

The inequality (8.8) implies

$$\text{tr } \tilde{\rho} \ln(\lambda\hat{\rho} + (1-\lambda)\tilde{\rho}) < \text{tr } \tilde{\rho} \ln \tilde{\rho}, \quad (8.9)$$

$$\text{tr } \hat{\rho} \ln(\lambda\hat{\rho} + (1-\lambda)\tilde{\rho}) < \text{tr } \hat{\rho} \ln \hat{\rho}. \quad (8.10)$$

If we multiply (8.9) with λ , $0 < \lambda < 1$ and (8.10) with $(1-\lambda)$ and if we add the resulting inequalities we obtain

$$\text{tr}(\lambda\hat{\rho} + (1-\lambda)\tilde{\rho}) \ln(\lambda\hat{\rho} + (1-\lambda)\tilde{\rho}) < \lambda \text{tr } \hat{\rho} \ln \hat{\rho} + (1-\lambda) \text{tr } \tilde{\rho} \ln \tilde{\rho} \quad (8.11)$$

Finally we change the sign and obtain for the entropy of a mixture of $\hat{\rho}$ with $\tilde{\rho}$, $\hat{\rho} \neq \tilde{\rho}$, for $0 < \lambda < 1$

$$S(\rho(\lambda)) > \lambda S(\hat{\rho}) + (1-\lambda)S(\tilde{\rho}). \quad (8.12)$$

The entropy of a mixture is larger than the proportionate sum of the entropies of its parts. Entropy increases by true mixing.

A mixture $\rho(\lambda)$ can be mixed from $\rho(\lambda_1)$ and $\rho(\lambda_2)$ with neighbouring mixing parameter $0 \leq \lambda_1 < \lambda < \lambda_2 \leq 1$.

$$\rho(\lambda) = \frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1} \rho(\lambda_1) + \frac{\lambda - \lambda_1}{\lambda_2 - \lambda_1} \rho(\lambda_2) \quad (8.13)$$

$$S(\rho(\lambda)) > \frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1} S(\rho(\lambda_1)) + \frac{\lambda - \lambda_1}{\lambda_2 - \lambda_1} S(\rho(\lambda_2)) \quad (8.14)$$

So the entropy $S(\rho(\lambda))$ is a convex function of the mixing parameter λ .

Random perturbation of the time evolution is a mixing process and increases the entropy. We consider in figure (1.1) a beam in a vacuum which is not perfect and take into account only the two possibilities that with probability λ no atom of the rest gas perturbs the beam and with probability $(1-\lambda)$ the case that an atom perturbs the beam. If we denote the mixtures which develop with and without perturbation as $\hat{\rho}$ and $\tilde{\rho}$, then the input to the measuring device is the mixture $\lambda\hat{\rho} + (1-\lambda)\tilde{\rho}$ if the perturbation by the rest atom occurs at random, independent from the preparation in the source.

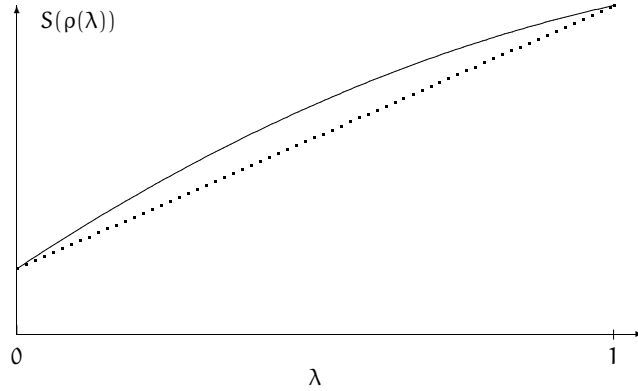


Figure 8.1: Entropy as convex function of the mixing parameter

8.2 Equilibrium

If the state in figure (1.1) is repeatedly and randomly perturbed before the measurement then ultimately the probability distribution of the results $w(i, \mathcal{A}, \rho(t))$ no longer depends on the time t at which the measurement is performed. By the von-Neumann equation (6.6) the density matrix ρ commutes in such a situation with the Hamilton operator

$$i\hbar \partial_t \rho(t) = [H, \rho] = 0 \quad (8.15)$$

and both operators have joint eigenstates Λ_i

$$H\Lambda_i = E_i\Lambda_i, \quad \rho\Lambda_i = \rho_i\Lambda_i. \quad (8.16)$$

Because with each random perturbation the entropy has increased one expects to find a mixture ρ with an entropy which is as large as possible. Time independent mixtures with maximal entropy define thermodynamic equilibrium.

If for example the mixture exchanges energy with its environment in such a way that the mean value has the constant value $\langle E \rangle$ then in thermodynamical equilibrium the entropy is maximal as a function of the eigenvalues ρ_i which are subject to the constraints $\sum_i \rho_i = 1$ and $\sum_i \rho_i E_i = \langle E \rangle$. We take these constraints into account by Lagrange multipliers α and β and determine the maximum of

$$S = - \sum_j \rho_j \ln \rho_j + \alpha(1 - \sum_j \rho_j) + \beta(\langle E \rangle - \sum_j \rho_j E_j). \quad (8.17)$$

The derivatives with respect to α and β are zero if the constraints hold. Differentiating with respect to ρ_i gives

$$0 = -(\ln \rho_i + 1) - \alpha - \beta E_i. \quad (8.18)$$

So ρ_i is a Boltzmann distribution.

$$\rho_{\text{Boltzmann } i} = \frac{e^{-\beta E_i}}{Z} \quad (8.19)$$

The factor Z is determined by the constraint $\sum_i \rho_i = 1$, Z is the partition sum.

$$Z(\beta) = \sum_i e^{-\beta E_i} \quad (8.20)$$

The partition sum is the Laplace transformed density of energy eigenstates. The parameter β is the inverse temperature

$$\beta = \frac{1}{T}, \quad (8.21)$$

which is determined implicitly by the constraint $\sum_i \rho_i E_i = \langle E \rangle$ as a function of the mean value $\langle E \rangle$ of the energy.

$$\langle E \rangle = \sum_i \rho_i E_i = \frac{1}{Z} \sum_i e^{-\beta E_i} E_i = -\frac{1}{Z} \partial_\beta \sum_i e^{-\beta E_i} = -\partial_\beta \ln Z(\beta) \quad (8.22)$$

The logarithm of the partition sum is a function of β , its derivative determines the mean energy.

The entropy of the Boltzmann distribution is closely related to the partition sum and the mean energy.

$$S = - \sum_i \rho_i \ln \rho_i = - \sum_i \frac{e^{-\beta E_i}}{Z} (-\beta E_i - \ln Z) = \beta \langle E \rangle + \ln Z \quad (8.23)$$

In terms of the free energy $F = \langle E \rangle - TS$ one has

$$Z = e^{-\beta F} \quad (8.24)$$

and the Boltzmann distribution can be written as

$$\rho_i = e^{-\beta(E_i - F)}. \quad (8.25)$$

The probabilities ρ_i only depend on differences of energy, not on the absolute value of the groundstate energy. The claim ‘‘Cold helium does not become solid because the groundstate energy is not zero’’ is wrong.

If no heat bath guarantees the mean energy of the mixture ρ then in (8.17) there is no Lagrange factor β , the other algebraic steps to determine the maximal entropy can be taken over as if $\beta = 0$. The entropy becomes maximal for equal distribution $\rho_i = 1/N$, where N is the dimension of the Hilbert space. The maximal entropy is

$$S = \ln N. \quad (8.26)$$

If the dimension N is infinite then there is no mixture with maximal entropy.

Important examples of thermodynamical systems in thermal equilibrium are the harmonic oscillator and the two state system. If one chooses the groundstate energy to vanish then the energies of the harmonic oscillator are given by

$$E_n = n\mathcal{E} \quad n = 0, 1, 2, \dots \quad (8.27)$$

In the two state system the energies E_n are of the same form, but n can take the values 0 and 1 only. The energies are the energies of free, identical bosons and fermions. $\mathcal{E} = \hbar\omega$ is the energy of one particle. A state with several particles has a multiple of the one particle energy because the particles are free and have no interaction. The Pauli exclusion principle forbids $n \geq 2$ for fermions.

The partition sum of the harmonic oscillator is a geometric series,

$$Z_{\text{boson}} = \sum_{n=0}^{\infty} e^{-\beta\mathcal{E}n} = \frac{1}{1 - e^{-\beta\mathcal{E}}} \quad (8.28)$$

the partition sum of the two state system is as simple as simple can be.

$$Z_{\text{fermion}} = 1 + e^{-\beta\mathcal{E}} \quad (8.29)$$

The energy expectation value is determined by (8.22).

$$\langle E \rangle_{\text{boson}} = \frac{\mathcal{E}}{e^{\beta\mathcal{E}} - 1} \quad \langle E \rangle_{\text{fermion}} = \frac{\mathcal{E}}{e^{\beta\mathcal{E}} + 1} \quad (8.30)$$

More complicated systems often consist of several, different bosons and fermions, for example of photons with different wave vector \vec{k} , with no interaction among each other.

More precisely we call a system freely composite if the Hilbert space is a product space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and if the Hamilton operator $H = H_1 \otimes \mathbb{1} + \mathbb{1} \otimes H_2$ is the sum of the single Hamilton operators. Then there are energy eigenstates $\Lambda_{i,\alpha}$, where i enumerates a basis of \mathcal{H}_1 and α a basis of \mathcal{H}_2 , and the energy is a sum of the single energies.

$$E_{i,\alpha} = E_1(i) + E_2(\alpha) \quad (8.31)$$

In such a case the partition sum of the composite system is the product of the single partition sums

$$Z = \sum_{i,\alpha} e^{-\beta(E_1(i) + E_2(\alpha))} = \sum_i e^{-\beta E_1(i)} \sum_{\alpha} e^{-\beta E_2(\alpha)} = Z_1 Z_2, \quad (8.32)$$

and the expectation value of the energy is the sum of the single expectation values.

$$\langle E \rangle = \langle E_1 \rangle + \langle E_2 \rangle \quad (8.33)$$

9 Decay of an unstable particle

9.1 Lorentz resonance

We consider a Hamiltonian with continuous energies and continuum normalized energy eigenstates $\Lambda_{E,p}$

$$H\Lambda_{E,p} = E\Lambda_{E,p} \quad \langle \Lambda_{E,p} | \Lambda_{E',p'} \rangle = \delta(E - E')\delta(p - p'). \quad (9.1)$$

The variable p distinguishes states with degenerate energy. If we represent states $\Psi(t)$ that satisfy the Schrödinger equation as a linear combination of this continuous basis, then $\Psi(t)$ has the following form

$$\Psi(t) = \int dE dp \Lambda_{E,p} \psi(E, p) e^{-\frac{i}{\hbar}Et}, \quad \psi(E, p) = \langle \Lambda_{E,p} | \Psi(0) \rangle. \quad (9.2)$$

A measurement device which examines whether the normalized initial state $\Psi(0)$ is present, finds this state at the instant t with the probability

$$w(t) = |\alpha(t)|^2, \quad (9.3)$$

$$\alpha(t) = \langle \Psi(0) | \Psi(t) \rangle = \int dE F(E)^2 e^{-\frac{i}{\hbar}Et}, \quad (9.4)$$

where we use the notation

$$F(E)^2 = \int dp |\langle \Lambda_{E,p} | \Psi(0) \rangle|^2. \quad (9.5)$$

The probability $w(t)$ decreases exponentially with the time t if the state is a Lorentz resonance.

$$F_{\text{Lorentz}}(E) = \left| \sqrt{\frac{\Gamma}{2\pi}} \frac{1}{(E - E_0) + i\frac{\Gamma}{2}} \right| \quad (9.6)$$

E_0 is the resonance energy and $|\Gamma| \geq 0$ the width of the resonance curve. However, the resonance energy and width are not defined as the energy expectation value $\langle H \rangle$ and energy uncertainty ΔH . These quantities diverge, since the function $F(E)$ does not decrease fast enough (cf. section (4.2)). $\Psi(t) = \exp(-iHt/\hbar)\Psi(0)$ satisfies the integrated Schrödinger equation, but $\Psi(t)$ is at no instant differentiable, $\|H\Psi(t)\|$ diverges.

One can evaluate the amplitude

$$\alpha_{\text{Lorentz}}(t) = \frac{\Gamma}{2\pi} \int dE \frac{e^{-\frac{i}{\hbar}Et}}{(E - E_0)^2 + \frac{\Gamma^2}{4}} \quad (9.7)$$

with the theorem of residues, since for positive (negative) times one may close the integration path in the lower (upper) complex half-plane, and obtains

$$\mathbf{a}_{\text{Lorentz}}(\mathbf{t}) = e^{-\frac{i}{\hbar}E_0\mathbf{t}} e^{-\frac{|\Gamma|\mathbf{t}}{2\hbar}}. \quad (9.8)$$

So the probability $w(\mathbf{t})$ of still finding the Lorentz resonance at the time $\mathbf{t} > 0$ decreases exponentially.

$$w_{\text{Lorentz}}(\mathbf{t}) = e^{-\mathbf{t}/\tau} \quad \text{for } \mathbf{t} > 0 \quad (9.9)$$

The lifetime is the inverse width $\tau = \hbar/|\Gamma|$, which is read as the uncertainty relation

$$\Delta\mathbf{t}\Delta E \geq \hbar/2. \quad (9.10)$$

This uncertainty relation is problematic, however: ΔE diverges and one does not measure \mathbf{t} but the property of still being the Lorentz resonance $\Psi(0)$ at the time \mathbf{t} .

The width $|\Gamma|$ is the decay rate of the exponentially decaying state

$$|\Gamma| = -\hbar \frac{1}{w(\mathbf{t})} \frac{d}{d\mathbf{t}} w(\mathbf{t}). \quad (9.11)$$

From equation (9.4) one infers that the phase of $\psi(E, \mathbf{p})$ is irrelevant. This is understandable, because any continuum normalized basis $\Lambda_{E, \mathbf{p}} = e^{i\varphi(E, \mathbf{p})} \Lambda_{E, \mathbf{p}}$ with arbitrary real function $\varphi(E, \mathbf{p})$ could have been used as well. In contrast to position and momentum there is no operator \mathbf{T} conjugate to the Hamiltonian with commutation relation $[\mathbf{T}, \mathbf{H}] = i\hbar$. Such an operator would be $\mathbf{T} = i\hbar\partial_E$ in an energy basis and would fix the relative phases of the basis $\Lambda_{E, \mathbf{p}}$.

Similarly, the sign of Γ is without significance, as (9.7) shows. The decay rate is $|\Gamma|$.

9.2 Deviation from exponential decay

Strictly speaking, there is no Lorentz resonance, since every realistic Hamiltonian has a spectrum that is bounded from below ($\psi(E, \mathbf{p}) = 0$ for $E < E_{\min}$), and since the energy expectation value must be finite. So the Lorentz resonance is unrealistic for small and large energies. Strictly speaking, there is no exponential decay, either.

There are no states which satisfy the differential Schrödinger equation $i\hbar\partial_t\Psi = \mathbf{H}\Psi$ and which satisfy a law of exponential decay for all times. The probability $w(\mathbf{t})$ is differentiable if $\Psi(\mathbf{t})$ is differentiable and is maximal $w(0) = 1$ at the time $\mathbf{t} = 0$. Therefore its time derivative vanishes then

$$\frac{d}{d\mathbf{t}} w(\mathbf{t})|_{\mathbf{t}=0} = 0 \quad (9.12)$$

and exponential decay is impossible for small times.

All deviations from exponential decay occur because the decay products regenerate the original state. This one sees with the following decomposition of the amplitude

$$\mathbf{a}(\mathbf{t} + \mathbf{t}') = \langle \Psi(0) | \Psi(\mathbf{t} + \mathbf{t}') \rangle = \langle \Psi(0) | e^{-\frac{i}{\hbar}\mathbf{H}\mathbf{t}} e^{-\frac{i}{\hbar}\mathbf{H}\mathbf{t}'} \Psi(0) \rangle. \quad (9.13)$$

If one inserts in between the e-functions a resolution of the identity

$$\mathbb{1} = |\Psi(0)\rangle\langle\Psi(0)| + \sum_{\mathbf{n}} |\Upsilon_{\mathbf{n}}\rangle\langle\Upsilon_{\mathbf{n}}|, \quad (9.14)$$

where the states $\Upsilon_{\mathbf{n}}$ complement the initial state $\Psi(0)$ to an orthonormal basis and represent the decay products, one sees

$$\mathbf{a}(\mathbf{t} + \mathbf{t}') = \mathbf{a}(\mathbf{t})\mathbf{a}(\mathbf{t}') + \sum_{\mathbf{n}} \langle \Psi(0) | e^{-\frac{i}{\hbar}\mathbf{H}\mathbf{t}} \Upsilon_{\mathbf{n}} \rangle \langle \Upsilon_{\mathbf{n}} | e^{-\frac{i}{\hbar}\mathbf{H}\mathbf{t}'} \Psi(0) \rangle. \quad (9.15)$$

The last term is the amplitude for the decay and reproduction of the initial state. If this last term were zero, the relation $\mathbf{a}(\mathbf{t} + \mathbf{t}') = \mathbf{a}(\mathbf{t})\mathbf{a}(\mathbf{t}')$ would follow and would imply exponential decay. Since the decay products leave the site and thereby reduce their density, one should guess for local interactions that deviations from exponential decay are observed only for time resolutions that are small as compared to the time of flight in which the decay products are diluted so that the reproduction can be neglected.

There must be deviations from the exponential decay also for large times if the decaying state was created by the action of the Hamiltonian \mathbf{H} during times $\mathbf{t} < 0$. The requirement that the Schrödinger equation was valid also at negative times is not mandatory, however: what happens at the preparation and measurement depends on the setup of the source and the measurement device and is not necessarily described by the Hamiltonian (cf. section (6.5)).

If however we accept (9.2) for all times, the modulus of the amplitude $\mathbf{a}(\mathbf{t})$ decreases exponentially for positive times and is bounded by $Ce^{-\mathbf{t}/\tau}$ with positive constants C and τ , if and only if $|\mathbf{a}(\mathbf{t})| < Ce^{-|\mathbf{t}|/\tau}$ holds for all times. In this case the Fourier transform

$$\tilde{\mathbf{a}}(E) = \int \frac{d\mathbf{t}}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}E\mathbf{t}} \mathbf{a}(\mathbf{t}) \quad (9.16)$$

is an analytic function of complex energies E in the range $|\Im(E)| < \hbar/\tau$, which in addition vanishes for real energies $E = E^* < E_{\min}$ below the minimal energy, if the energy is bounded from below. Thus $\tilde{\mathbf{a}}(E)$ must vanish. This, however, is in contradiction to $\mathbf{a}(\mathbf{t}) = \int \frac{dE}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}E\mathbf{t}} \tilde{\mathbf{a}}(E)$ and $\mathbf{a}(0) = 1$. Therefore the modulus of the amplitude $\mathbf{a}(\mathbf{t})$ cannot be bounded exponentially but has to decrease more slowly for large times.

9.3 Golden Rule

An energy eigenstate cannot decay, since from $\Psi(\mathbf{t}) = e^{-\frac{i}{\hbar}E\mathbf{t}}\Psi(0)$ it follows for real energy E that $w(\mathbf{t}) = 1$. So derivations of the law of exponential decay for a state which is assigned a definite energy are self contradictory. The energy E cannot have a negative imaginary part $\Gamma/2$, either. While this would imply $w(\mathbf{t}) = e^{-\frac{\Gamma}{\hbar}\mathbf{t}}$ and is being used to parameterize decaying states without having to describe the decay products, a Hamiltonian which completely describes the time evolution of a decaying state has to describe also the decay products and their time evolution. It has to be hermitean, so that $\Psi(\mathbf{t})$ remains normalized for all times, and can have only real eigenvalues.

Let us consider a Hilbert space which is spanned by a normalized state Υ and continuum normalized basis states $\Lambda_{E,p}$ orthogonal to it with $E \geq E_{\min}$

$$\langle \Upsilon | \Upsilon \rangle = 1, \quad \langle \Upsilon | \Lambda_{E,p} \rangle = 0, \quad \langle \Lambda_{E,p} | \Lambda_{E',p'} \rangle = \delta(E - E') \delta(p - p'). \quad (9.17)$$

Let this basis be adjusted to the decomposition

$$H = H_0 + H_{\text{int}} \quad (9.18)$$

of the Hamiltonian and be chosen such that Υ is a normalized H_0 -eigenstate with eigenvalue $E_0 > E_{\min}$ in the continuum and that $\Lambda_{E,p}$ are generalized eigenstates with energy E in the continuous spectrum of H_0 and with an index p to discriminate states with degenerate energy

$$H_0 \Upsilon = E_0 \Upsilon, \quad H_0 \Lambda_{E,p} = E \Lambda_{E,p}, \quad E \geq E_{\min}. \quad (9.19)$$

For example H_0 can contain the rest energy of the decaying particle and the kinetic energies including the rest energy of the decay products with continuous kinetic energies of the relative motion. The variable p can be taken in the example as angles under which the decay products are emitted.

We consider the amplitude

$$\langle \Lambda_{E,p} | e^{-\frac{i}{\hbar} H t} \Upsilon \rangle \quad (9.20)$$

for the transition of the normalized H_0 -eigenstate Υ into the orthogonal continuum normalized H_0 -eigenstates $\Lambda_{E,p}$ to lowest order in the interaction H_{int} and expand for this purpose $e^{-\frac{i}{\hbar} H t}$ into a Taylor series in $t H_{\text{int}}$. We extract the coefficients of the series by repeated differentiation from the relation

$$\partial_\lambda e^{A(\lambda)} = \int_0^1 dz e^{z A(\lambda)} \partial_\lambda A e^{(1-z) A(\lambda)}. \quad (9.21)$$

One proves this relation by expansion of both sides

$$\sum_n \frac{1}{n!} \sum_{l=0}^{n-1} A^{n-l-1} (\partial_\lambda A) A^l = \sum_k \frac{1}{k!} \sum_l \frac{1}{l!} \int dz z^k (1-z)^l A^k (\partial_\lambda A) A^l \quad (9.22)$$

with the combinatorial formula

$$\int_0^1 dz z^k (1-z)^l = \frac{k! l!}{(k+l+1)!}. \quad (9.23)$$

If we expand the time evolution operator $e^{-\frac{i}{\hbar} H t}$ with respect to $t H_{\text{int}}$ using (9.21), the transition amplitude vanishes to lowest order and is given up to terms of higher order by

$$\langle \Lambda_{E,p} | e^{-\frac{i}{\hbar} H t} \Upsilon \rangle = -\frac{i}{\hbar} t \langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle \int_0^1 dz e^{-\frac{i}{\hbar} z t E} e^{-\frac{i}{\hbar} (1-z) t E_0} + \dots \quad (9.24)$$

Here we have employed the H_0 -eigenvalue equations.

In this order of approximation the reproduction of Υ from the decay products and the interaction of the decay products is not taken into account.

The z -integration gives

$$\langle \Lambda_{E,p} | e^{-\frac{i}{\hbar} H t} \Upsilon \rangle = \langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle \frac{e^{-\frac{i}{\hbar} E t} - e^{-\frac{i}{\hbar} E_0 t}}{E - E_0} + \dots \quad (9.25)$$

The probability of finding a state $\Lambda_{E,p}$ in the H_0 -energy range Δ at the time t is determined by the modulus squared of this amplitude and given to lowest order by

$$w(\Delta, t) = \int_\Delta dE dp |\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle|^2 \frac{\sin^2\left(\frac{t(E-E_0)}{2}\right)}{\left(\frac{E-E_0}{2}\right)^2} \quad (9.26)$$

Due to (B.3) one has for sufficiently large times approximately

$$\frac{w(\Delta, t)}{t} = \frac{2\pi}{\hbar} \int dE dp |\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle|^2 \delta(E - E_0). \quad (9.27)$$

However, the limit $t \rightarrow \infty$ must not be taken, since otherwise higher powers of $t H_{\text{int}}$ cannot be neglected anymore. In particular the time t has to remain small as compared to the lifetime of the decaying state: for larger times the probability of finding decay products does not increase linearly with time. For small times t that are large enough for the approximation (B.3) we interpret $\frac{w(\Delta, t)}{t}$ as the derivative $-\frac{dw}{dt}$ of the probability of still finding the decaying state and read off the decay rate.

$$\Gamma_{\text{Golden Rule}} = 2\pi \int dE dp |\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle|^2 \delta(E - E_0) \quad (9.28)$$

The decay rate is composed additively of partial decay rates $d\Gamma$ of various processes

$$d\Gamma = 2\pi \delta(E - E_0) |\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle|^2 dE dp. \quad (9.29)$$

For this standard derivation of the Golden Rule the time t is sufficiently large, because no quantum mechanical system can decay exponentially for small times (see section (9.2)). In addition this time t is small against the lifetime τ . It is remarkable how obediently this textbook derivation of the Golden Rule is accepted by students and recited by lecturers. The assumptions on t exclude each other in the limit and errors that one makes for intermediate times, which are both sufficiently large and sufficiently small, are not easily estimated.

9.4 Decay into the continuum

One can solve the time evolution of the decaying state exactly in terms of integrals without approximations. The Golden Rule arises in the limit of small and non-resonant transition amplitudes.

A general state is written with an expansion coefficient $\psi_0 = \langle \Upsilon | \Psi \rangle$ and a wave function $\psi(E, \mathbf{p}) = \langle \Lambda_{E, \mathbf{p}} | \Psi \rangle$ in the basis (9.17) as a linear combination

$$\Psi = \Upsilon \psi_0 + \int dE d\mathbf{p} \Lambda_{E, \mathbf{p}} \psi(E, \mathbf{p}), \quad \psi(E, \mathbf{p}) = 0 \text{ for } E < E_{\min}. \quad (9.30)$$

The probability $w_0(\Delta, \Psi)$ of finding the energy belonging to H_0 in the open interval Δ upon measuring Ψ amounts to

$$w_0(\Delta, \Psi) = \begin{cases} \int_{\Delta} dE d\mathbf{p} |\psi(E, \mathbf{p})|^2 & \text{if } E_0 \notin \Delta \\ |\psi_0|^2 + \int_{\Delta} dE d\mathbf{p} |\psi(E, \mathbf{p})|^2 & \text{if } E_0 \in \Delta \end{cases}. \quad (9.31)$$

So the part $\Upsilon \psi_0$ contributes to the probability density of the energy a sharp line at E_0 , which is sharper than any detector resolution Δ and whose area amounts to $|\psi_0|^2$.

The interaction $H_{\text{int}} = H_{\text{int}}^\dagger$ causes transitions from the normalized H_0 -eigenstate Υ into the continuum and vice versa

$$H_{\text{int}} \Upsilon = \int dE d\mathbf{p} \Lambda_{E, \mathbf{p}} v(E, \mathbf{p}), \quad H_{\text{int}} \Lambda_{E, \mathbf{p}} = v^*(E, \mathbf{p}) \Upsilon, \quad (9.32)$$

$$v(E, \mathbf{p}) = \langle \Lambda_{E, \mathbf{p}} | H_{\text{int}} \Upsilon \rangle. \quad (9.33)$$

The matrix elements $\langle \Upsilon | H_{\text{int}} \Upsilon \rangle$ and $\langle \Lambda_{E', \mathbf{p}'} | H_{\text{int}} \Lambda_{E, \mathbf{p}} \rangle$ vanish in our computation. This is no essential restriction, we can imagine such a part of the interaction to be absorbed already in H_0 . Similarly, $v(E, \mathbf{p}) = \langle \Lambda_{E, \mathbf{p}} | H_{\text{int}} \Upsilon \rangle$ vanishes for $E < E_{\min}$.

The operator H_{int} is defined on Υ only if $\|H_{\text{int}} \Upsilon\| < \infty$.

$$\int dE d\mathbf{p} |v(E, \mathbf{p})|^2 < \infty \quad (9.34)$$

No matter how small the amplitude $v(E, \mathbf{p})$ is for transitions into the continuum, if at an energy E_1 it does not vanish for all \mathbf{p} , there is no normalizable H -eigenstate with this energy. For the equation $(H - E_1)\Psi = 0$ determines the energy wave function

$$\psi(E, \mathbf{p}) = -\frac{v(E, \mathbf{p})}{E - E_1} \psi_0 \quad (9.35)$$

and the energy E_1 by means of the gap equation

$$E_1 - E_0 = -\int_{E_{\min}}^{\infty} dE d\mathbf{p} \frac{|v(E, \mathbf{p})|^2}{E - E_1}. \quad (9.36)$$

$\psi(E, \mathbf{p})$ is a square integrable function only if $\int d\mathbf{p} |v(E_1, \mathbf{p})|^2$ vanishes.

If the interaction $v(E, \mathbf{p})$ is sufficiently large, then there is a normalizable H -eigenstate below the continuous energies, here $v(E, \mathbf{p})$ vanishes, since the gap equation (9.36) has for $E_1 < E_{\min}$ exactly one solution, if

$$\lim_{\epsilon \rightarrow 0^+} \int_{E_{\min}}^{\infty} dE d\mathbf{p} \frac{|v(E, \mathbf{p})|^2}{E - E_{\min} + \epsilon} \geq E_0 - E_{\min} \quad (9.37)$$

Namely, the left side of (9.36) is a monotonically increasing function of E_1 and the right side decreases monotonically. In addition for strongly negative E_1 the left side of (9.36) is smaller than the right, for $E_1 = E_{\min}$ the left side is larger than the right, if the interaction $v(E, \mathbf{p})$ is sufficiently large.

We now investigate the time evolution of the state $\Psi(t) = e^{-\frac{i}{\hbar} H t} \Upsilon$, which was prepared at the time $t = 0$ as a normalized eigenstate Υ of the unperturbed Hamiltonian H_0 . The amplitude $\mathbf{a}(t)$ for finding the initial state in a measurement at the time t is the matrix element

$$\mathbf{a}(t) = \langle \Upsilon | e^{-\frac{i}{\hbar} H t} \Upsilon \rangle. \quad (9.38)$$

We use the theorem of residues and represent $e^{-\frac{i}{\hbar} H t}$ as an integral along a curve Γ which surrounds the spectrum in the complex plane in the counter-clockwise direction.

$$e^{-\frac{i}{\hbar} H t} = \frac{1}{2\pi i} \oint_{\Gamma} dz e^{-\frac{i}{\hbar} t z} \frac{1}{z - H} \quad (9.39)$$

The formula can easily be verified with the theorem of residues in the case of a discrete spectrum $H \Lambda_n = E_n \Lambda_n$ with $\Psi(0) = \sum_n \Lambda_n \psi_n$

$$\begin{aligned} & \frac{1}{2\pi i} \sum_n \oint_{\Gamma} dz e^{-\frac{i}{\hbar} z t} \frac{1}{z - H} \Lambda_n \psi_n = \\ & = \frac{1}{2\pi i} \sum_n \oint_{\Gamma} dz e^{-\frac{i}{\hbar} z t} \frac{1}{z - E_n} \Lambda_n \psi_n = \sum_n \Lambda_n \psi_n e^{-\frac{i}{\hbar} E_n t} \end{aligned}$$

and holds also for a continuous spectrum.

The negative resolvent $(z - H)^{-1}$ can be written as a geometric series. Namely, one has for operators A and V , provided A and $A - V$ are invertible,

$$(A - V)^{-1} = (A(\mathbb{1} - A^{-1}V))^{-1} = (\mathbb{1} - A^{-1}V)^{-1} A^{-1} = \sum_{n=0}^{\infty} (A^{-1}V)^n A^{-1} \quad (9.40)$$

on states on which the series converges. We thus write

$$\frac{1}{z - H} = \frac{1}{z - H_0 - H_{\text{int}}} = \sum_{n=0}^{\infty} \left(\frac{1}{z - H_0} H_{\text{int}} \right)^n \frac{1}{z - H_0}. \quad (9.41)$$

The powers of $(z - H_0)^{-1} H_{\text{int}}$ are easily applied to Υ , since Υ is eigenstate of $((z - H_0)^{-1} H_{\text{int}})^2$. With the notation

$$|V(E)|^2 = \int d\mathbf{p} |v(E, \mathbf{p})|^2 \quad (9.42)$$

one has

$$\frac{1}{z - H_0} H_{\text{int}} \Upsilon = \int dE d\mathbf{p} \Lambda_{E, \mathbf{p}} \frac{v(E, \mathbf{p})}{z - E} \quad (9.43)$$

$$\left(\frac{1}{z - H_0} H_{\text{int}} \right)^2 \Upsilon = \frac{1}{z - E_0} \int dE \frac{|V(E)|^2}{z - E} \Upsilon. \quad (9.44)$$

Hence, to the matrix element $\langle \Upsilon | (z - H)^{-1} \Upsilon \rangle$ only the even powers of $(z - H_0)^{-1} H_{\text{int}}$ contribute.

$$\begin{aligned} \langle \Upsilon | \frac{1}{z - H} \Upsilon \rangle &= \sum_n \left(\frac{1}{z - E_0} \int dE \frac{|V(E)|^2}{z - E} \right)^n \frac{1}{z - E_0} \\ &= (z - E_0 - \int dE \frac{|V(E)|^2}{z - E})^{-1} \end{aligned} \quad (9.45)$$

For $\alpha(t)$ it then follows due to (9.38) and (9.39)

$$\alpha(t) = \frac{1}{2\pi i} \oint_{\Gamma} dz e^{-\frac{i}{\hbar} z t} \frac{1}{z - E_0 + \int dE \frac{|V(E)|^2}{E - z}}. \quad (9.46)$$

We choose the path of integration Γ around the spectrum in the counter-clockwise direction such that we traverse with fixed imaginary part $\epsilon > 0$ the points $z = x + i\epsilon$ from $x = \infty$ to $x = -\infty$ and then the points $z = x - i\epsilon$ from $x = -\infty$ to $x = \infty$. Then the complex path integral is the difference of two integrals along the real axis.

The integral does not depend on ϵ and we evaluate it in the limit $\epsilon \rightarrow 0+$.

$$\alpha(t) = \lim_{\epsilon \rightarrow 0+} \frac{i}{2\pi} \int_{-\infty}^{\infty} dx (e^{-\frac{i}{\hbar} t(x+i\epsilon)} \frac{1}{x + i\epsilon - E_0 + \int dE \frac{|V(E)|^2}{E - x - i\epsilon}} - (\epsilon \rightarrow -\epsilon)) \quad (9.47)$$

As shown in appendix A, the denominator

$$f(x) = x + i\epsilon - E_0 + \int dE \frac{|V(E)|^2}{E - x - i\epsilon} \quad (9.48)$$

tends in the limit $\epsilon \rightarrow 0+$ to

$$f(x) = x - E_0 + \int dE' \frac{|V(E' + x)|^2}{E'} + i\pi |V(x)|^2, \quad (9.49)$$

where \int denotes the principal value integral. The denominator f does not vanish if $V(x)$ is small enough such that no bound state exists, which solves the gap-equation (9.36).

The second contribution to $\alpha(t)$ which is obtained with ϵ replaced by $-\epsilon$ yields the complex conjugate denominator f^* . Therefore the integrand is of the form

$$e^{-\frac{i}{\hbar} t x} \left(\frac{1}{f} - \frac{1}{f^*} \right) = e^{-\frac{i}{\hbar} t x} \left(\frac{f^* - f}{ff^*} \right) \quad (9.50)$$

and $\alpha(t)$ is of the form

$$\alpha(t) = \int dE e^{-\frac{i}{\hbar} E t} |F(E)|^2 \quad (9.51)$$

$$F(E) = \frac{V(E)}{E - E_0 + \int dE' \frac{|V(E'+E)|^2}{E'} + i\pi |V(E)|^2}. \quad (9.52)$$

Due to the coupling to the continuum the decaying state is no longer an energy eigenstate, but a resonance similar to the Lorentz resonance, because for small transition amplitude $V(E)$, $F(E)$ is nearly the function $F_{\text{Lorentz}}(E)$ given in (9.6). The deviations from the Lorentz resonance lead to deviations from the exponential decay.

In the denominator of $F(E)$ the real part dominates

$$f(E) = E - E_0 + \int dE' \frac{|V(E' + E)|^2}{E'} \quad (9.53)$$

except in the small neighbourhood of the root \hat{E} of $f(E)$. The numerator $V(E)$ eliminates the unphysical aspects of the Lorentz resonance. It ensures that the energy wave function vanishes for $E < E_{\text{min}}$ and improves the high energy behaviour, such that $H\Upsilon$ has finite norm. Then there exist the expectation value $\langle H \rangle$ and the uncertainty Δ_H , however, these quantities depend on the behaviour of $V(E)$ for $E \neq \hat{E}$ and need not coincide with the resonance energy and the width.

If in the neighbourhood of the root of $f(E)$ the imaginary part of the denominator of $F(E)$ does not vary strongly, we can approximate it there by $\pi |V(\hat{E})|^2$ and obtain approximately

$$F(E) \approx \frac{V(\hat{E})}{(E - \hat{E})(1 + \partial_{\hat{E}} \int dE' \frac{|V(E'+\hat{E})|^2}{E'} + i\pi |V(\hat{E})|^2)} \quad (9.54)$$

with the resonance energy \hat{E} , which is defined implicitly by the gap equation $f(\hat{E}) = 0$, and the width

$$\Gamma = \frac{2\pi \int dp |v(\hat{E}, p)|^2}{1 + \partial_{\hat{E}} \int dE' dp \frac{|v(E'+\hat{E}, p)|^2}{E'}}. \quad (9.55)$$

To lowest order this is the Golden Rule.

$$\Gamma_{\text{Golden Rule}} = 2\pi \int dE dp \delta(E - \hat{E}) |\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle|^2 \quad (9.56)$$

The decay rate of a state, which decays by transitions into a continuum of energy states, is a 2π -multiple of the integral of the square modulus of the transition amplitude $v(\hat{E}, p)$ at the resonance energy \hat{E} times a delta function for the conservation of energy. More exactly, equation (9.52) shows that the decaying state is not an energy eigenstate but a resonance similar to the Lorentz resonance.

This derivation of the Golden Rule shows that an exact computation of the transition amplitude $\langle \Lambda_{E,p} | H_{\text{int}} \Upsilon \rangle$ at higher order in perturbation theory is to be supplemented by an exact computation of the time behaviour of the decaying state. The Golden Rule is valid to lowest order of the transition amplitude.

The approximation (9.54) assumes not only that $V(E)$ is small, but also that $V(E)$ is smooth and is not by itself a Lorentz resonance with resonance energy E_1 and width Γ_1 . Otherwise the denominator $f(E_1) \approx (E_1 - E_0)$ in (9.52) suppresses the contribution of the resonance at E_1 to be sure, but if the width Γ_1 is small against $2\pi |V(\hat{E})|^2$, first the state Υ decays quickly as a resonance with energy \hat{E} up to a small, longer living remnant of the resonance with energy E_1 and width Γ_1 .

9.5 General validity

In (9.51,9.52) we have computed in full generality the amplitude for some normalized state $\Psi(\mathbf{t})$, which as a wave packet is composed of continuous energy eigenstates, to coincide with the state $\Psi(0)$. Namely, let the normalized initial state $\Upsilon = \Psi(0)$ be composed of continuous energy eigenstates of the Hamiltonian H . The projector

$$P = |\Upsilon\rangle\langle\Upsilon|, \quad P^2 = P, \quad \mathbb{1} = P + (\mathbb{1} - P), \quad (9.57)$$

is used in the definition of the unperturbed Hamiltonian

$$H_0 = PHP + (\mathbb{1} - P)H(\mathbb{1} - P). \quad (9.58)$$

The state $\Upsilon = P\Upsilon$ is a normalized eigenstate of H_0

$$H_0\Upsilon = PHP\Upsilon = \Upsilon\langle\Upsilon|H\Upsilon\rangle = E_0\Upsilon \quad (9.59)$$

with eigenvalue $E_0 = \langle\Upsilon|H\Upsilon\rangle$, which lies in the continuum of the eigenvalues of H_0 .

The interaction

$$H_{\text{int}} = H - H_0 = PH(\mathbb{1} - P) + (\mathbb{1} - P)HP \quad (9.60)$$

makes transitions from Υ to states orthogonal to it.

So every state Υ and every Hamiltonian H with continuous spectrum satisfy the assumptions we made in section (9.4). By choice of the arbitrary state Υ the function $F(E)$ in (9.51) with the restrictions $F(E) = 0$ for $E < E_{\text{min}}$ and $\int dE |F(E)|^2 = 1$ can be chosen arbitrarily. Therefore the amplitude usually does not decrease exponentially.

The state decreases exponentially if the amplitude $v(E, \mathbf{p})$ for the decay into the continuum of the decay products is small and does not show resonant behaviour by itself.

9.6 Decay of moving particles

In relativistic quantum mechanics the Hamiltonian $H = cP^0$ is a component of the four-momentum. To Lorentz transformations Λ , which satisfy $\Lambda^0_0 \geq 1$ and thus do not reverse the time direction, there correspond unitary operators $U(\Lambda)$ which represent the Lorentz transformations on states with integral spin.

$$U(\Lambda_2\Lambda_1) = U(\Lambda_2)U(\Lambda_1) \quad (9.61)$$

For half-integral spin and for Lorentz transformations which reverse the time direction the situation is more involved [3, chapter 2]: time reversal can only be realized as an antiunitary transformation and on states with half integer spin the universal covering group $SL(2, \mathbb{C})$ of the Lorentz group is represented. These complications do not matter here, however.

The unitary transformations generate time direction preserving Lorentz transformations of the four-momenta

$$U^{-1}(\Lambda)P^m U(\Lambda) = \Lambda^m_n P^n. \quad (9.62)$$

Hence, applied to a four-momentum eigenstate Φ_p with $P^m\Phi_p = p^m\Phi_p$, $U(\Lambda)$ gives an eigenstate with Lorentz transformed four-momentum.

$$P^m U(\Lambda)\Phi_p = U(\Lambda)\Lambda^m_n P^n \Phi_p = \Lambda^m_n p^n U(\Lambda)\Phi_p \quad (9.63)$$

For decaying particles which move with velocity \mathbf{v} it follows from this that their lifetime τ_v is increased by time dilation.

$$\tau_v = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tau_0 \quad (9.64)$$

The relation is not mathematically exact due to quantum mechanical reasons. Strictly speaking, a particle at rest cannot be prepared, to this there would correspond a constant and thus non-normalizable position wave function. If in order to keep the localization energy small one works with wave functions which are constant in a large domain and go to zero outside the domain, a Lorentz transformed observer sees the state in this large domain before and after a long period of time. If the idealization that the state decays exponentially has already been made, then for a Lorentz transformed observer there corresponds to this decay a wave function which increases exponentially in the opposite direction of motion of the particle. One encounters similar difficulties if one wants to describe a decaying state as a momentum eigenstate and as an eigenstate of a non-hermitean Hamiltonian. If the energy has a negative imaginary part, the Lorentz transformed state has a complex momentum eigenvalue. The corresponding wave function then increases exponentially in some direction.

If one considers wave packets and works with normalized states, the amplitude $\mathbf{a}(\mathbf{t})$ (9.4) is time-dependent already for stable particles. This is because wave packets of free massive particles spread in time since they are composed of parts with different momenta and thus different velocities. Apart from this the amplitude

$$\langle\Psi_v(0)|\Psi_v(\mathbf{t})\rangle \quad (9.65)$$

of a state moving with velocity \mathbf{v} is time-dependent simply because it moves with velocity \mathbf{v} and thus overlaps less and less with the wave packet at time \mathbf{t} . In order to determine the amplitude of the state to still be present at the time \mathbf{t} for a nearly monochromatic wave packet, $\Psi_v(\mathbf{t})$, one therefore has to compare with the state $e^{-\frac{i}{\hbar}\vec{P}\vec{v}\mathbf{t}}\Psi_v(0)$ shifted by $\mathbf{x} = \mathbf{v}\mathbf{t}$ (5.26).

$$\mathbf{a}_v(\mathbf{t}) = \langle e^{-\frac{i}{\hbar}\vec{P}\vec{v}\mathbf{t}}\Psi_v(0)|\Psi_v(\mathbf{t})\rangle = \langle\Psi_v(0)|e^{\frac{i}{\hbar}\vec{P}\vec{v}\mathbf{t}}e^{-\frac{i}{\hbar}H\mathbf{t}}\Psi_v(0)\rangle \quad (9.66)$$

One obtains the states Ψ_v moving with velocity \mathbf{v} from states Ψ_0 at rest, they are eigenstates of the spatial momentum $\vec{P}\Psi_0 = 0$, by means of the unitary transformation

$$\Psi_v = U(\Lambda(\mathbf{v}))\Psi_0, \quad (9.67)$$

which belongs to the rotation-free Lorentz transformation $\Lambda(\mathbf{v})$, for instance

$$\Lambda(\mathbf{v}) = \begin{pmatrix} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} & \begin{pmatrix} 1 & \frac{v}{c} \\ \frac{v}{c} & 1 \end{pmatrix} & \\ & & 1 \\ & & & 1 \end{pmatrix} \quad (9.68)$$

for a state moving in x -direction. Ψ_v is no momentum eigenstate if Ψ_0 is unstable, since Ψ_0 is no energy eigenstate.

If we insert (9.67) into (9.66) and employ (9.62), it follows with $\mathbf{u}_m = (ct, -vt, 0, 0)$

$$\langle \Psi_0 | \mathbf{U}^{-1} e^{-\frac{i}{\hbar} \mathbf{u}_m \cdot \mathbf{P}^m} \mathbf{U} \Psi_0 \rangle = \langle \Psi_0 | e^{-\frac{i}{\hbar} \mathbf{u}_m \cdot \mathbf{U}^{-1} \mathbf{P}^m \mathbf{U}} \Psi_0 \rangle = \langle \Psi_0 | e^{-\frac{i}{\hbar} \mathbf{u}_m \cdot \Lambda^m \mathbf{P}^n} \Psi_0 \rangle \quad (9.69)$$

From $\mathbf{u}_m \cdot \Lambda^m \mathbf{P}^n = \sqrt{1 - \frac{v^2}{c^2}} \mathbf{t} \cdot \mathbf{P}^0$ and $c\mathbf{P}^0 = H$ we finally obtain

$$\mathbf{a}_v(\mathbf{t}) = \langle \Psi_0 | e^{-\frac{i}{\hbar} \sqrt{1 - \frac{v^2}{c^2}} \mathbf{t} H} \Psi_0 \rangle = \mathbf{a}_0 \left(\sqrt{1 - \frac{v^2}{c^2}} \mathbf{t} \right). \quad (9.70)$$

So the moving particle decays more slowly than the one at rest. If the probability of survival of the particle at rest decreases exponentially with a lifetime τ , i.e. if $|\mathbf{a}_0(\mathbf{t})|^2 = e^{-\frac{t}{\tau}}$, then one has for the moving particle $|\mathbf{a}_v(\mathbf{t})|^2 = e^{-\frac{\sqrt{1 - \frac{v^2}{c^2}} t}{\tau}}$. So it has a lifetime of

$$\tau_v = \frac{\tau}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (9.71)$$

There is no universal answer to the question whether the acceleration influences the lifetime or whether quantum mechanical particles are ideal clocks and measure the length of the world line. One has to expect that the kind of acceleration is essential. For example a magnetic field changes the energy differences of atomic levels. Therefore acceleration in a magnetic field changes the rates of clocks which measure the time by counting atomic transitions. If one brings to mind that the notion of proper time makes use of the localization of the clock on a world line, but that such a localization at different times contradicts quantum mechanics, one realizes that already the question whether quantum mechanical clocks measure the length along a world line is problematic.

Appendix

$$\mathbf{A} \quad \lim_{\epsilon \rightarrow 0^+} \frac{1}{x+i\epsilon} = \mathbf{PV} \frac{1}{x} - i\pi\delta(x)$$

In order to examine the limit $\lim_{\epsilon \rightarrow 0^+} \frac{1}{x+i\epsilon}$ we apply $\frac{1}{x+i\epsilon} = \frac{x-i\epsilon}{x^2+\epsilon^2}$ to a real test function $t(x)$.

The imaginary part is identified by substitution of variables.

$$\int dx t(x) \frac{-\epsilon}{x^2+\epsilon^2} = - \int \frac{dx}{\epsilon} t\left(\frac{x}{\epsilon}\right) \frac{1}{1+\left(\frac{x}{\epsilon}\right)^2} = -\text{sign}(\epsilon) \int dx t(\epsilon x) \frac{1}{1+x^2}$$

For continuous, bounded test functions $t(x)$ this tends for $\epsilon \rightarrow 0^+$ to

$$-t(0) \int \frac{dx}{1+x^2} = -\pi t(0) = - \int dx t(x) \pi\delta(x) . \quad (\text{A.1})$$

The real part $\frac{x}{x^2+\epsilon^2}$ is an odd function of x . Hence, applied to a test function $t(x)$ which decreases fast enough for large x , only the odd part $x\hat{t}(x)$ of the test function contributes.

$$t(x) - t(-x) = 2x \hat{t}(x) \quad (\text{A.2})$$

$$\begin{aligned} \int dx t(x) \frac{x}{x^2+\epsilon^2} &= \int dx \frac{1}{2} \frac{xt(x) - xt(-x)}{x^2+\epsilon^2} = \int dx \frac{\frac{x^2}{\epsilon^2}}{1+\frac{x^2}{\epsilon^2}} \hat{t}(x) \\ &= \int dx \hat{t}(x) - |\epsilon| \int d\left(\frac{x}{\epsilon}\right) \frac{1}{1+\frac{x^2}{\epsilon^2}} \hat{t}\left(\frac{x}{\epsilon}\right) \\ &= \int dx \hat{t}(x) - |\epsilon| \int dx \frac{\hat{t}(\epsilon x)}{1+x^2} \xrightarrow{\epsilon \rightarrow 0} \int dx \hat{t}(x) \end{aligned} \quad (\text{A.3})$$

For differentiable test functions t , \hat{t} is continuously completed at $x=0$ and assumes the value $\frac{dt}{dx}$ there. The integral over \hat{t} is the principal value $\mathbf{PV} \frac{1}{x}$ integrated with a test function $t(x)$.

$$\int dx \frac{1}{2} \frac{t(x) - t(-x)}{x} = \lim_{\epsilon \rightarrow 0^+} \left(\int_{-\infty}^{-\epsilon} dx \frac{t(x)}{x} + \int_{+\epsilon}^{\infty} dx \frac{t(x)}{x} \right) = \int dx \frac{t(x)}{x} \quad (\text{A.4})$$

This proves the assertion.

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x+i\epsilon} = \mathbf{PV} \frac{1}{x} - i\pi\delta(x) \quad (\text{A.5})$$

$$\mathbf{B} \quad \lim_{t \rightarrow \infty} \frac{\sin^2(tx)}{tx^2} = \pi\delta(x)$$

The integral with a test function $f(x)$ is written as

$$\int dx f(x) \frac{\sin^2(tx)}{tx^2} = \int t dx f\left(\frac{tx}{t}\right) \frac{\sin^2 tx}{(tx)^2} = \int dx f\left(\frac{x}{t}\right) \frac{\sin^2 x}{x^2} \quad (\text{B.1})$$

and, if f is a continuous and bounded test function, it tends for $t \rightarrow \infty$ to

$$f(0) \int dx \frac{\sin^2 x}{x^2} = f(0)\pi . \quad (\text{B.2})$$

So one has

$$\lim_{t \rightarrow \infty} \frac{\sin^2(tx)}{tx^2} = \pi\delta(x) . \quad (\text{B.3})$$

C Remark on Fourier transformation

Fourier transformations map square integrable functions ψ unitarily to square integrable functions $\tilde{\psi}$

$$\tilde{\psi}(x) = \int \frac{dy}{\sqrt{2\pi}} e^{ixy} \psi(y), \quad \psi(x) = \int \frac{dy}{\sqrt{2\pi}} e^{-ixy} \tilde{\psi}(y), \quad (\text{C.1})$$

because the scalar product

$$\int dx \tilde{\phi}^*(x) \tilde{\psi}(x) = \int dy dy' \frac{dx}{2\pi} e^{-ixy} e^{ixy'} \phi^*(y) \psi(y') \quad (\text{C.2})$$

coincides due to (5.50) with the original scalar product $\int dy \phi^*(y) \psi(y)$. Because of $\tilde{\tilde{f}}(x) = f(-x)$ a Fourier transformation repeated four times leads back to the original function f . Since the Fourier transformation is unitary, it can only have eigenvalues λ with absolute value 1 which moreover satisfy $\lambda^4 = 1$. So the possible eigenvalues are $\pm 1, \pm i$. If one decomposes a function $f = g + u$ into an even $g(x) = g(-x)$ and an odd $u(x) = -u(-x)$ part, one can write f as a sum of four eigenfunctions under Fourier transformations with eigenvalues $\pm 1, \pm i$: $g = 1/2(g + \tilde{g}) + 1/2(g - \tilde{g})$ and $u = 1/2(u - i\tilde{u}) + 1/2(u + i\tilde{u})$.

It is therefore not exceptional that Fourier transformations map a function such as the Gauß function to itself.

D Derivative of the determinant

The determinant of a matrix M is a polynomial function of the matrix elements M^i_j . From its definition

$$\det M = \epsilon_{i_1 i_2 \dots i_n} M^{i_1}_1 M^{i_2}_2 \dots M^{i_n}_n \quad (\text{D.1})$$

it follows by differentiation that

$$\frac{\partial \det M}{\partial M^i_j} = \epsilon_{i_1, \dots, i_{j-1}, i, i_{j+1}, \dots, i_n} M^{i_1}_1 \dots M^{i_{j-1}}_{j-1} M^{i_{j+1}}_{j+1} \dots M^{i_n}_n. \quad (\text{D.2})$$

If one multiplies this derivative with M^i_l and sums over i , one obtains again the determinant, if $l = j$. In all other cases $l \neq j$ one obtains zero, since in the sum with the ϵ -tensor $M^{i_1}_l$ already occurs and since ϵ is totally antisymmetric. These observations suffice to identify the derivative $\frac{\partial \det M}{\partial M^i_j}$.

$$\frac{\partial \det M}{\partial M^i_j} = \det M (M^{-1})^j_i \quad (\text{D.3})$$

Together with the chain rule one obtains the derivative of the determinant of matrices $M(\alpha)$ which depend on a parameter

$$\partial_\alpha \det M(\alpha) = \det M (M^{-1})^j_i \partial_\alpha M^i_j. \quad (\text{D.4})$$

If for $\alpha = 0$ the matrix is $M(0) = \mathbb{1}$, as in one-parameter matrix groups, then the derivative of the determinant at $\alpha = 0$ is the trace of the differentiated matrix $\partial_\alpha M|_{\alpha=0}$

$$\partial_\alpha \det M|_{\alpha=0} = 1 \cdot \delta^i_j \partial_\alpha M^i_j|_{\alpha=0} = \partial_\alpha M^i_i|_{\alpha=0} = \text{tr} \partial_\alpha M|_{\alpha=0}. \quad (\text{D.5})$$

Bibliography

- [1] A. Galindo and P. Pascual, Quantum Mechanics I and II, Springer Verlag, Heidelberg, 1990
- [2] M. Reed and B. Simon, Analysis of Operators, Academic Press, New York, 1978
- [3] Steven Weinberg, The Quantum Theory of Fields, Volume 1, Cambridge University Press, Cambridge, 1995