Chapter 4. Bra-ket Formalism

In this chapter we will discuss P. A. M. Dirac's bra-ket formalism of quantum mechanics, which allows to overcome some inconveniences and limitations of wave mechanics, and give the simplest examples of its applications.

4.1. Motivation

We have finished the last chapter with a complaint about wave mechanics. Moreover, we had filed several other grievances about this approach throughout the first part of the course.[1] Let us briefly summarize these complaints:

(i) Wave mechanics is focused on spatial dependence of wavefunctions. This is why attempts to do analyze temporal evolution within this approach (beyond the trivial time behavior of the eigenfunctions, described by Eq. (1.45)) run into severe technical difficulties. For example, we could obtain Eq. (2.134) describing time dynamics of the metastable state, or Eq. (2.156) for two coupled quantum well, only for the simplest potential profiles, while it is intuitively clear that such dynamics should be virtually independent of the particular form of quantum wells and the tunnel barriers providing limited quantum confinement.

(ii) Other important issues which are conceptually addressable within wave mechanics, e.g., the Feynman path integral, various perturbation theories, description of coupling to environment, etc., would lead to math so bulky that I (wisely:-) postponed them until we have got a more compact formalism on hand.

(iii) In the discussion of several key problems (the harmonic oscillator, 3D rotator, Bohr's atom) we have run into complex eigenfunctions coexisting with simple energy spectra (which imply simple background physics). It is very important to get this physics revealed!

(iv) In the wave mechanics postulates, formulated in Sec. 1.2, quantum mechanical operators are treated very unequally, with a pivotal role given to $\hat{H}$ and $\hat{p}$. On the other hand, in Sec. 3.5 we ran into operators $\hat{L}_z$ and $\hat{L}^2$ which apparently should be treated on the equal footing, but did not have a background for that. Moreover, some expressions, e.g., for the free-particle's eigenfunction

$$\exp\left\{i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}\right\}, \tag{4.1}$$

or for the harmonic oscillator's Hamiltonian function

$$H=\frac{p^2}{2m}+\frac{m\omega_0^2 r^2}{2}, \tag{4.2}$$

invite a similar treatment of momentum and coordinate, but we have not yet defined a quantum-mechanical operator corresponding to the radius-vector.

[1] All these complaints should not overshadow the fact that the wave mechanics is a great science, and is virtually sufficient for many important applications of quantum mechanics, e.g., for most semiconductor electronics.

(v) However, the strongest motivation for a new conceptual start comes from the fact that wave mechanics cannot describe elementary particle's spin and other "internal" quantum numbers, such as quark flavors or lepton numbers.

In the context of the last item, let us review the basic facts on spin (which is a very representative but experimentally the most accessible of all internal quantum numbers), so we would be aware of what the new approach we should explain - as the minimum. Figure 1 the scheme of the first experiment, carried out by O. Stern and W. Gerlach in 1922. A collimated beam of electrons is passed through the gap between poles of a strong magnet, where the magnetic field is non-uniform, so that both B_z and dB_z/dz are not equal to zero. As a result, the beam splits into two parts of equal intensity.

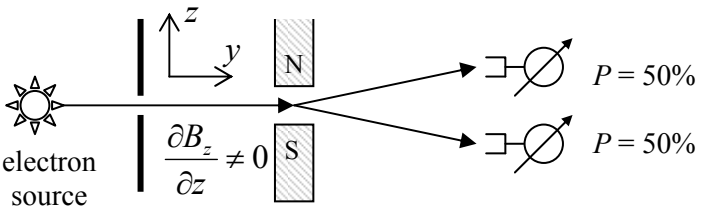

Fig. 4.1. The Stern-Gerlach experiment.

This simple experiment can be explained on classical, though phenomenological grounds by assuming that as electron has an intrinsic, spontaneous magnetic dipole moment **m**. Indeed, classical electrodynamics (see, e.g., EM Sec. 5.4) tells us that the potential energy of a localized moment in an external magnetic field equals (-**m**·**B**), so that the force acting on the particle,

$$\mathbf{F} = -\nabla U = -\nabla\left(-\mathbf{m}\cdot\mathbf{B}\right), \tag{4.3}$$

has a vertical component

$$F_z = -\frac{\partial}{\partial z}\left(-m_z \cdot B_z\right) = m_z \frac{\partial B_z}{\partial z}. \tag{4.4}$$

Hence if we postulate the existence of two possible, discrete values of $m_z = \pm\mu$, this explains the SG effect qualitatively. (A quantitative explanation of the beam splitting angle requires μ to be equal, or very close to the so-called *Bohr magneton*[2]

$$\mu_B \equiv \frac{\hbar e}{2m_e} \approx 0.9274\times10^{-23}\,\frac{\mathrm{J}}{\mathrm{T}}; \tag{4.5}$$

a convenient mnemonic rule is that it is about 1 kelvin per tesla.)

However, this classical language cannot explain the results of the following set of multi-stage Stern-Gerlach (SG for short) experiments (Fig. 2). In the first of them, the electron beam is first passed through a magnetic field oriented (together with its gradient) along axis *z*. Then one of the two resulting beams is absorbed (or otherwise filtered out of the experiment), while the other one is passed through a similar but *x*-oriented field. The experiment shows that the beam is split again into two components of equal intensity. A classical explanation of this experiment requires a very unnatural suggestion that the

[2] In Gaussian units, $\mu_B \equiv \hbar e/2m_e c \approx 0.9274\times10^{-20}$.

initial electrons had random but discrete values of the magnetic moment simultaneously in *two* directions, z and x.

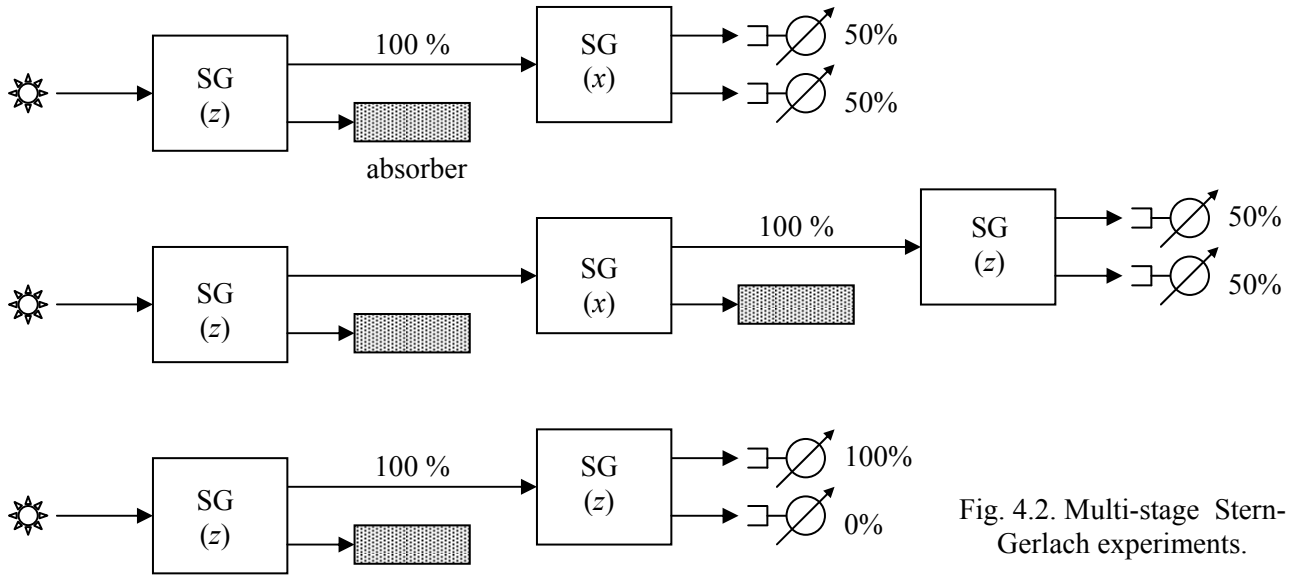

Fig. 4.2. Multi-stage Stern-Gerlach experiments.

However, even this assumption cannot explain the results of the three-state SQ experiment shown on the middle panel of Fig. 2. Here, the previous two-state setup is complemented with one more absorber and one more magnet, now with z-orientation again. Completely counter-intuitively, it again gives two beams of equal intensity, as if we have not yet filtered out the electrons with m_z corresponding to the lower beam, in the first z-experiment.

The only way to save the classical explanation venue here is to say that maybe, electrons *somehow* interact with the magnetic field that a polarized (non-absorbed) beam becomes depolarized again. But any hope for such "explanation" is ruined by the control experiment shown on the bottom panel of Fig. 2, whose results indicate that no depolarization of this sort takes place.

We will see below that all these results find a natural explanation in the "*matrix mechanics*" which was developed by W. Heisenberg, M. Born and P. Jordan in 1925. However, the matrix formalism has turned out to be inconvenient for the solution of most problems we have discussed in Chapters 2 and 3, and for a time was eclipsed by Schrödinger's wave mechanics which was put forward a few months later. However, very soon P. A. M. Dirac introduced a more general bra-ket formalism which provides a generalization of both approaches and has proves their equivalence. Let us study it.

4.2. Main postulates of quantum mechanics in bra-ket formalism

Hopefully, our previous studies have given the reader a gut feeling what the "quantum state" of a system is.[3] In particular, if the quantum state may be described by wave mechanics, this description is

[3] An attentive reader could notice my smuggling term "system" instead of "particle" which was used in the previous chapters. Indeed, the bra-ket formalism allows description of quantum systems much more complex than a single spinless particle which is a typical (though not the only possible) subject of wave mechanics.

given by the corresponding wavefunction. Notice, however, that *a wavefunction is not a state*, but just a mathematical object giving its mathematical description, just as the radius-vector and velocity are mathematical objects (vectors) describing a classical particle – see Fig. 3.

In the same spirit, in the Dirac formalism a certain quantum state α is described by either of two *state vectors*: either a "*ket-vector*" $|\alpha\rangle$ or a "*bra-vector*" $\langle\alpha|$. The terms stem from the fact that two vectors, say $|\alpha\rangle$ and $\langle\beta|$, may be considered as parts of the combination $\langle\beta|\alpha\rangle$ which reminds angle brackets. Similarly, the word "vector", in application to bra- and ket-vectors, should not be taken literally. It is used mostly to emphasize that each of these mathematical objects carries more information than a scalar. Though there are some similarities between bra- and ket-vectors and the "usual" geometric vectors, generally these are new mathematical notions, and we need new rules for them - at the first stage, postulates, though later we will derive other properties (theorems) from those postulates. Just as in Sec. 1.2, I will not try too hard to beat the number of these postulates to the smallest possible minimum, trying instead to keep their physics meaning transparent.

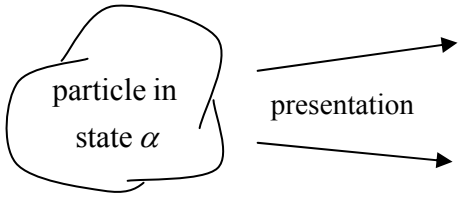

Fig. 4.3. Particle state's description.

(i) Ket-vectors. For certainty, we will start with *ket-vectors* (or just *kets*), because those of bra-vectors are almost similar. Probably, the most important property of the vectors is their *linear superposition*. Namely, if several ket-vectors $|\alpha_j\rangle$ are possible states of a quantum system, then any linear combination

$$|\alpha\rangle = \sum_j c_j |\alpha_j\rangle, \tag{4.6}$$

where c_j are any (complex) c-numbers, is also a possible state of the system. Actually, the meaning of the right-hand part of Eq. (6) becomes clear only after we have postulated the following rules of summation of the ket-vectors

$$|\alpha_j\rangle + |\alpha_{j'}\rangle = |\alpha_{j'}\rangle + |\alpha_j\rangle, \tag{4.7}$$

and their multiplication by a c-number:

$$c_j |\alpha_j\rangle = |\alpha_j\rangle c_j . \tag{4.8}$$

Notice that in the set of wave mechanics postulates, statements similar to (7) and (8) were unnecessary, because wavefunctions are the usual (albeit complex) functions, and we knew from the elementary algebra that such equations are valid. However, ket-vectors are new mathematical objects, and we should be more careful.

It is probably evident from Eq. (6) that coefficients c_j indicate the "weight" (a more common term is *amplitude*) of state α_j in superposition α. Their relation to experiment is given by the following

statement: if the particle is definitely (with probability 1) in the composite state α, the probability to find it in a particular component state α_j is

$$W_j = \left|c_j\right|^2. \tag{4.9}$$

(I will try to give a more clear definition of the term "find" later on.) One important particular case is $c_j = 0$, showing that state α_j does not participate in α. By the way, a product of any state by c-number equal 0,

$$0\left|\alpha_j\right\rangle, \tag{4.10}$$

has a special name of "null-state". It is important to avoid confusion between state (10) and the ground state of the system (which is frequently denoted by ket-vector $\left|0\right\rangle$); in some sense, the null-state does not exist, while the ground state does.

(ii) Bra-vectors and inner ("scalar") products. If a ket-vector $\left|\alpha\right\rangle$ is known, the corresponding bra-vector $\left\langle\alpha\right|$ describes the same state, i.e. does not provide any new information about the system and is only used for mathematical convenience.[4] In other words, there is a *unique dual correspondence* between $\left|\alpha\right\rangle$ and $\left\langle\alpha\right|$.The correspondence between the two vectors is provided by the following rule: is a ket-vector of a linear superposition is described by Eq. (6), then the corresponding bra-vector is

$$\left\langle\alpha\right| = \sum_j c_j^* \left\langle\alpha_j\right| = \sum_j \left\langle\alpha_j\right| c_j^* . \tag{4.11}$$

The convenience of using two types of vectors is clear from the following definition of their *inner* (or "*scalar*") *product*

$$\left\langle\beta\middle\|\alpha\right\rangle \equiv \left\langle\beta\middle|\alpha\right\rangle. \tag{4.12}$$

Since this is new object again, its properties have to be defined. First of all, the inner product is a (generally, complex)[5] *scalar*. Its main property is the linearity with respect to any of its component vectors. For example, if a linear superposition α is described by the ket-vector (6) or a bra-vector (11), then

$$\left\langle\beta\middle|\alpha\right\rangle = \sum_j c_j \left\langle\beta\middle|\alpha_j\right\rangle, \tag{4.13}$$

while

$$\left\langle\alpha\middle|\beta\right\rangle = \sum_j c_j^* \left\langle\alpha_j\middle|\beta\right\rangle. \tag{4.14}$$

In plain English, c-numbers may be moved out of the inner products.

The second key property of the inner product is

[4] Mathematicians like to say that the ket- and bra-vectors are defined in "*isomorphic Hilbert spaces*".

[5] This is one difference of bra and ket-vectors from the usual (geometrical) vectors whose scalar product is always real.

$$\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle^* . \tag{4.15}$$

This property is compatible with Eqs. (13) and (14). Indeed, the complex conjugation of both parts of Eq. (13) gives:

$$\langle \beta | \alpha \rangle^* = \sum_j c_j^* \langle \beta | \alpha_j \rangle^* = \sum_j c_j^* \langle \alpha_j | \beta \rangle = \langle \alpha | \beta \rangle . \tag{4.15}$$

One more rule: a scalar product of a state by itself (called the *norm squared*) is real and non-negative,

$$\|\alpha\|^2 \equiv \langle \alpha | \alpha \rangle \geq 0. \tag{4.16}$$

In order to give the reader some feeling about the meaning of this rule: we will show below that if state α may be described by wavefunction Ψ_α, then

$$\langle \alpha | \alpha \rangle = \int \Psi_\alpha^* \Psi_\alpha d^3 r \geq 0 . \tag{4.17}$$

(Notice that, by convention, there is no conjugation sign in the bra-part of the product.)

(iii) Linear operators. One more key notion of the Dirac formalism are quantum-mechanical *operators*. Just as the operators we have discussed in wave mechanics, the function of an operator is the "generation" of one state from another: if $|\alpha\rangle$ is a possible ket of the system, and $\hat{A}$ is a meaningful operator, then

$$|\alpha'\rangle = \hat{A}|\alpha\rangle \tag{4.18}$$

is also a ket-vector describing a possible state of the system. (The mathematicians say that these two vectors exist "in the same Hilbert space"). The main rules governing the operators is their linearity, both with respect to a superposition of vectors,

$$\hat{A}\left(\sum_j c_j |\alpha_j\rangle \right) = \sum_j c_j \hat{A} |\alpha_j\rangle , \tag{4.19}$$

and a superposition of operators:

$$\left(\sum_j c_j \hat{A}_j \right) |\alpha\rangle = \sum_j c_j \hat{A}_j |\alpha\rangle . \tag{4.20}$$

These rules are evidently similar to Eqs. (1.38) and (1.39) of wave mechanics.

The above rules imply that operator "acts" on a ket-vector on its right; however, a combination of the type $\langle \alpha | \hat{A}$ is also legitimate and presents a new bra-vector. It is important that, generally, this vector does *not* represent the same state as ket-vector (18); instead, the bra-vector corresponding to (18) is

$$\langle \alpha' | = \langle \alpha | \hat{A}^\dagger . \tag{4.21}$$

This statement serves as the definition of the *Hermitian conjugate* (or "*Hermitian adjoint*") $\hat{A}^\dagger$ of the initial operator $\hat{A}$. However, for an important class of *Hermitian operators*, the Hermitian conjugation is inconsequential; for those operators

$$\hat{A}^\dagger = \hat{A}. \tag{4.22}$$

(This equality, as well as any other operator equation below, means that these operators act similarly on any bra- or ket-vector.) [6]

To proceed further, we need to introduce an additional postulate (Dirac called it the *associative axiom of multiplication*): in *any* legitimate bra-ket expression,[7] we may insert or remove parentheses (in the usual meaning that the operation inside the parentheses is performed first) where we like. The first two examples of this postulate are given by Eqs. (19) and (20), but it is more general and says, for example:

$$\langle\beta|\hat{A}|\alpha\rangle = \langle\beta|\left(\hat{A}|\alpha\rangle\right) = \left(\langle\beta|\hat{A}\right)|\alpha\rangle. \tag{4.23}$$

This equation, when combined with the definition of the Hermitian conjugate and Eq. (15), has an interesting and important corollary:

$$\langle\beta|\hat{A}|\alpha\rangle = \langle\beta|\left(\hat{A}|\alpha\rangle\right) = \left(\langle\alpha|\hat{A}^\dagger\right)|\beta\rangle^* = \langle\alpha|\hat{A}^+|\beta\rangle^*, \tag{4.24}$$

which is most frequently rewritten as

$$\langle\alpha|\hat{A}^\dagger|\beta\rangle = \langle\beta|\hat{A}|\alpha\rangle^*. \tag{4.25}$$

The associative axiom also allows to define one more, *outer product* of bra- and ket-vectors:

$$|\beta\rangle\langle\alpha|. \tag{4.26}$$

In contrast to the inner product (12), the outer product is an *operator*. Indeed, the associative axiom allows us to remove parentheses in the following expression:

$$\left(|\beta\rangle\langle\alpha|\right)|\gamma\rangle = |\beta\rangle\langle\alpha|\gamma\rangle. \tag{4.27}$$

But the last bra-ket is just a scalar; hence object (26) acting on a ket-vector gives a new ket-vector, which is the essence of an operator. Very similarly,

$$\langle\delta|\left(|\beta\rangle\langle\alpha|\right) = \langle\delta|\beta\rangle\langle\alpha| \tag{4.28}$$

- again a typical operator action.

---

[6] *c*-numbers may be also considered a particular type of operators, whose action is limited to the change of the state amplitude. According to Eqs. (11) and (21), for them the Hermitian conjugation is equivalent to the simple complex conjugation, so that only a real c-number may be considered as a (particular case of) Hermitian operator.

[7] Some examples of *illegitimate* expressions: $|\alpha\rangle\hat{A}$, $\hat{A}\langle\alpha|, |\alpha\rangle|\beta\rangle, \langle\alpha|\langle\beta|$. Notice, however, that the last two expressions may be legitimate if α and β are states of different systems; I will give some examples of such situations in Chapters 7 and 8.

Now let us do the following calculation. We can use the parentheses insertion into the bra-ket equality, following from Eq. (15),

$$\langle\gamma|\alpha\rangle\langle\beta|\delta\rangle = \left(\langle\delta|\beta\rangle\langle\alpha|\gamma\rangle\right)^*, \tag{4.29}$$

to transform it to the following form:

$$\langle\gamma|\left(|\alpha\rangle\langle\beta|\right)|\delta\rangle = \left(\langle\delta|\left(|\beta\rangle\langle\alpha|\right)|\gamma\rangle\right)^*. \tag{4.30}$$

Since this equation should be valid for any vectors $\langle\gamma|$ and $|\beta\rangle$, its comparison with Eq. (25) gives the following operator equality

$$\left(|\alpha\rangle\langle\beta|\right)^\dagger = |\beta\rangle\langle\alpha|. \tag{4.31}$$

This is the operand swap rule for the outer product; it reminds rule (15) for the inner product, but now requires the Hermitian (rather than the usual complex) conjugation.

The associative axiom is also valid for the operator "multiplication", showing that the action of an operator product is nothing more than the sequential action of operators on state vectors:

$$\left(\hat{A}\hat{B}\right)|\alpha\rangle = \hat{A}\left(\hat{B}|\alpha\rangle\right), \quad \langle\beta|\left(\hat{A}\hat{B}\right) = \left(\langle\beta|\hat{A}\right)\hat{B}. \tag{4.32}$$

However, we have to be careful with the operator products; generally $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. This is why the *commutator*

$$\left[\hat{A},\hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \tag{4.33}$$

is a very useful option. Another similar option is *the anticommutator*,

$$\left\{\hat{A},\hat{B}\right\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}, \tag{4.34}$$

sometimes also denoted as $[\ldots]_+$.

Finally, the bra-ket formalism broadly uses two special operators: the *null operator* which "kills" *any* state, turning it into the null-state:

$$\hat{0}|\alpha\rangle \equiv 0|\alpha\rangle, \tag{4.35}$$

and the *identity operator*, which is also defined by its action (or rather non-action:-) on an arbitrary state:

$$\hat{I}|\alpha\rangle \equiv |\alpha\rangle. \tag{4.36}$$

4.3. Orthonormal basis and matrix representation

While some operations in quantum mechanics may be carried out in the general bra-ket formalism outlined above, most calculations are done in specific quantum systems which feature at least one (typically, many) *full and orthonormal* set u_j of states, also called a *basis*. This means that, first, the expansion (6) of any state vector in series over the basis vectors,

$$|\alpha\rangle = \sum_j \alpha_j |u_j\rangle \tag{4.37}$$

is unique, and, second, that

$$\langle u_j | u_{j'} \rangle = \delta_{jj'}. \tag{4.38}$$

For the systems which may be described by wave mechanics, examples of the full orthonormal bases are presented by any set of normalized eigenfunctions calculated in the previous 3 chapters, e.g., Eq. (1.60).

Due to the uniqueness of expansion (37), the full set of coefficients α_j gives a complete description of state α, just as the usual Cartesian components A_x, A_y, and A_z (in a certain coordinate system) give a complete description of a usual geometric 3D vector $\vec{A}$. This fact justifies using the term "state vectors" for quantum-mechanical bras and kets, though with two major differences:

(i) a basis set may have an infinite number of elements u_j, and

(ii) coefficients α_j may be complex.

This analogy may be pushed even further. Let us inner-multiply both parts of Eq. (37) by bra-vector $\langle u_{j'}|$ and then transform is using the rules discussed in the previous section, and Eq. (38):

$$\langle u_{j'} | \alpha \rangle = \langle u_{j'} | \sum_j \alpha_j |u_j\rangle = \sum_j \alpha_j \langle u_{j'} | u_j \rangle = \alpha_{j'}, \tag{4.39}$$

This means that any of the expansion coefficients α_j may be presented as a scalar product

$$\alpha_j = \langle u_j | \alpha \rangle, \tag{4.40}$$

which is just an analog of equalities like $A_x = \mathbf{n}_x \cdot \mathbf{A}$ in the usual vector algebra. Using this result, expansion (37) may be rewritten as

$$|\alpha\rangle = \sum_j |u_j\rangle\langle u_j|\alpha\rangle \equiv \sum_j \hat{\Lambda}_j |\alpha\rangle. \tag{4.41}$$

A comparison with Eq. (26) shows that expressions

$$\hat{\Lambda}_j \equiv |u_j\rangle\langle u_j| \tag{4.42}$$

are outer products, i.e. legitimate operators, and are called the *projection operators*. Their sense is that each of these operators separates from a full state vector (37) just one component:

$$\hat{\Lambda}_j |\alpha\rangle = |u_j\rangle\langle u_j|\alpha\rangle = \alpha_j |u_j\rangle, \tag{4.43}$$

In the vector analogy, the projection operator kills all components of the state vector but one, i.e. "projects" the state vector on certain (j-th) "direction", hence the name. Probably, the most important property of the projection operators immediately follows from Eq. (41): their sum over the full base state is equivalent to the identity operator:

$$\sum_j |u_j\rangle\langle u_j| = \hat{I}. \tag{4.44}$$

We will repeatedly use this very useful equation, starting from a few lines below.

Now let us see what expansion (37) does with all the notions introduced in the last section, starting from the "short bra-ket" (the inner product of two vectors):

$$\langle \beta | \alpha \rangle = \sum_{j,j'} \beta_j^* \alpha_{j'} \delta_{jj'} = \sum_{j'} \beta_j^* \alpha_j . \tag{4.45}$$

Besides complex conjugation, this is of course similar to the usual scalar product. Now, the full bra-ket:

$$\langle \beta | \hat{A} | \alpha \rangle = \sum_{j,j'} \beta_j^* \alpha_{j'} \langle u_j | \hat{A} | u_{j'} \rangle = \sum_{j'} \beta_j^* A_{jj'} \alpha_{j'} . \tag{4.46}$$

Here, at the last step, we have introduced a very important notion of *matrix elements* of the operator:

$$A_{jj'} \equiv \langle u_j | \hat{A} | u_{j'} \rangle . \tag{4.47}$$

Their full set completely characterizes the operator, just as the full set of expansion coefficients (40) characterizes the corresponding quantum state. The term "matrix" means, first of all, that it is convenient to present the full set of $A_{jj'}$ as a square table (matrix), with the linear dimension equal to the number of states u_j of this particular system.

As two trivial examples, all matrix elements of the null-operator, defined by Eq. (35), are equal zero, and hence it is presented as a matrix of zeros:

$$0 = \begin{pmatrix} 0 & 0 & \dots \\ 0 & 0 & \dots \\ \dots & \dots & \dots \end{pmatrix}, \tag{4.48}$$

while for the identity operator $\hat{I}$ (36) we readily get

$$I_{jj'} = \langle u_j | \hat{I} | u_{j'} \rangle = \langle u_j | u_{j'} \rangle = \delta_{jj'} , \tag{4.49}$$

i.e. its matrix is diagonal:

$$\mathrm{I} = \begin{pmatrix} 1 & 0 & \dots \\ 0 & 1 & \dots \\ \dots & \dots & \dots \end{pmatrix}. \tag{4.50}$$

However, the convenience of the matrix language extends well beyond presentation of particular operators. For example, let us calculate matrix elements for the product of two operators:

$$(AB)_{jj''} = \langle u_j | \hat{A}\hat{B} | u_{j''} \rangle . \tag{4.51}$$

Here we can use Eq. (44) for the first (but not the last!) time, first inserting the identity operator between two operators, and then expressing it via a sum of projection operators:

$$\langle u_j | \hat{A}\hat{B} | u_{j''} \rangle = \langle u_j | \hat{A}\hat{I}\hat{B} | u_{j''} \rangle = \sum_{j'} \langle u_j | \hat{A} | u_{j'} \rangle \langle u_{j'} | \hat{B} | u_{j''} \rangle = \sum_{j'} A_{jj'} B_{j'j''} . \tag{4.52}$$

But the result corresponds to the regular rule of calculation of an arbitrary element of the matrix product

$$\mathrm{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots \\ B_{21} & B_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix}. \tag{4.53}$$

Hence the product of operators may be presented by that of their matrices. This is so convenient that the same language is used to present not only Eq. (46),

$$\langle \beta | \hat{A} | \alpha \rangle = \sum_{j'} \beta_j^* A_{jj'} \alpha_{j'} = \left(\beta_1^*, \beta_2^* \dots \right) \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix}, \tag{4.54}$$

but even the simpler Eq. (45):

$$\langle \beta | \alpha \rangle = \sum_{j} \beta_j^* \alpha_j = \left(\beta_1^*, \beta_2^* \dots \right) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix}, \tag{4.55}$$

though it involves the use of non-square matrices: rows of (complex-conjugate!) expansion coefficients for the presentation of bra-vectors, and columns of these coefficients for the presentation of ket-vectors. Now, this matrix mapping becomes completely general; for example, the full bra-ket of and operator product is

$$\langle \beta | \hat{A}\hat{B} | \alpha \rangle = \sum_{j',j''} \beta_j^* A_{jj'} B_{j'j''} \alpha_{j''} = \left(\beta_1^*, \beta_2^* \dots \right) \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots \\ B_{21} & B_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix}. \tag{4.56}$$

Now let us have a look at the outer product operator (26). Its matrix elements are just

$$\left(|\alpha\rangle\langle\beta| \right)_{jj'} = \langle u_j | \alpha \rangle \langle \beta | u_j{}' \rangle = \alpha_j \beta_{j'}^*. \tag{4.57}$$

This means a very special matrix, with full information provided by $2N$ (complex) numbers (where N is the basis set size), rather than N^2 numbers as for an arbitrary operator. However, a simple generalization of the outer product may present an arbitrary operator. Indeed, let us insert two identity operators (44), but with different summation indices, on both sides of the operator in the matrix element definition (47):

$$A_{jj'} = \langle u_j | \left(\sum_k |u_k\rangle\langle u_k| \right) \hat{A} \left(\sum_{k'} |u_{k'}\rangle\langle u_{k'}| \right) | u_{j'} \rangle, \tag{4.58}$$

and rewrite this expression as

$$A_{jj'} = \langle u_j | \left(\sum_{k,k'} |u_k\rangle\langle u_k|\hat{A}|u_{k'}\rangle\langle u_{k'}| \right) | u_{j'} \rangle = \langle u_j | \left(\sum_{k,k'} |u_k\rangle A_{kk'} \langle u_{k'}| \right) | u_{j'} \rangle. \tag{4.59}$$

But this means that all matrix elements of the operator in parentheses coincide with those of operator $\hat{A}$, i.e. these two operators are identical, and we may write

$$\hat{A} = \sum_{k,k'} |u_k\rangle A_{kk'} \langle u_{k'}| . \tag{4.60}$$

The reader has to agree that this formula, which is a natural generalization of Eq. (44), is extremely elegant. Also notice the following parallel: if we consider the matrix element definition (47) as some sort of analog of Eq. (40), then Eq. (60) is a similar analog of the expansion expressed by Eq. (37).

The matrix presentation is so convenient that it makes sense to move it by one hierarchical step lower – from state vector products to "bare" state vectors and operators. For example, let us present ket-vector (18) as

$$|\alpha'\rangle = \sum_j \alpha_j' |u_j\rangle . \tag{4.61}$$

But, on the other hand, we can use expansion (37) of the initial vector in the same basis:

$$|\alpha'\rangle = \hat{A}|\alpha\rangle = \hat{A}\sum_{j'} \alpha_{j'} |u_{j'}\rangle = \sum_{j'} \alpha_{j'} \hat{A}|u_{j'}\rangle . \tag{4.62}$$

Now using the same trick of inner multiplication by $\langle u_j|$, we get

$$\langle u_j|\alpha'\rangle = \sum_{j'} \alpha_{j'} \langle u_j|\hat{A}|u_{j'}\rangle = \sum_{j'} A_{jj'} \alpha_{j'} . \tag{4.63}$$

But according to Eq. (40), the left-hand part of this equation is nothing but α_j', so that

$$\alpha_j' = \sum_{j'} A_{jj'} \alpha_{j'} . \tag{4.64}$$

But this result corresponds to the usual rule of multiplication of a matrix by a column, so we can present the operator equality (18) in the similar matrix form:

$$\begin{pmatrix} \alpha_1' \\ \alpha_2' \\ \dots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix} . \tag{4.65}$$

Absolutely similarly, the bra-vector equality (21) may be presented as

$$\left(\alpha_1'^*, \alpha_2'^*, \dots\right) = \left(\alpha_1^*, \alpha_2^*, \dots\right) \begin{pmatrix} A_{11}^\dagger & A_{12}^\dagger & \dots \\ A_{21}^\dagger & A_{22}^\dagger & \dots \\ \dots & \dots & \dots \end{pmatrix} . \tag{4.66}$$

By the way, Eq. (66) naturally raises the following question: what are the elements of matrix $\hat{A}^\dagger$, or more exactly, what is the relation between the matrices of an operator and its Hermitian conjugate? The simplest way to get an answer is to use Eq. (25) with two arbitrary basis states (say, u_j and $u_{j'}$) used in the role of α and β. Together with the orthonormality relation (38), this immediately gives

$$A^{\dagger}_{jj'} = A^{*}_{j'j}. \tag{4.67}$$

In plain English, the matrix of the Hermitian conjugate operator is the *complex conjugated and transposed* matrix of the initial operator. This result exposes very clearly the difference between the Hermitian and usual (complex) conjugation. It also shows that for the Hermitian operators, defined by Eq. (22),

$$A_{jj'} = A^{*}_{j'j}, \tag{4.68}$$

i.e. any pair of their matrix elements, symmetric about the main diagonal, should be complex conjugate of each other, and hence the main-diagonal elements have to be real:

$$\mathrm{Im}\, A_{jj} = 0. \tag{4.69}$$

(The identity operator's matrix (50) evidently satisfies this condition, so that the operator is Hermitian.)

In order to appreciate the special role played by the Hermitian operators in quantum theory, let us introduce the key notions of *eigenstates* (described by *eigenvectors* $\langle a_j|$ and $|a_j\rangle$) and *eigenvalues* (*c*-numbers) A_j of an operator $\hat{A}$, defined by the equation they should satisfy:[8]

$$\hat{A}|a_j\rangle = A_j|a_j\rangle. \tag{4.70}$$

Let us prove that eigenvalues of any Hermitian operator are real,

$$A_j = A^{*}_j, \quad \text{for } j = 1, 2, ..., N, \tag{4.71}$$

while the eigenstates corresponding to different eigenvalues are orthogonal:

$$\langle a_j | a_{j'}\rangle = 0, \quad \textit{if} \quad A_j \neq A_{j'}. \tag{4.72}$$

The proof of both statements is surprisingly simple. Let us inner-multiply Eq. (70) by bra-vector $\langle a_{j'}|$. In the RHP of the result, the eigenvalue A_j, as a *c*-number, may be taken out of the bra-ket, giving

$$\langle a_{j'}|\hat{A}|a_j\rangle = A_j\langle a_{j'}|a_j\rangle. \tag{4.73}$$

This equality should hold for any pair of eigenstates, so that we may swap the indices, and complex-conjugate the result:

$$\langle a_j|\hat{A}|a_{j'}\rangle^{*} = A^{*}_{j'}\langle a_j|a_{j'}\rangle^{*}. \tag{4.74}$$

Now using Eqs. (15) and (25), together with the Hermitian operator definition (22), we transform Eq. (70) to the form

$$\langle a_{j'}|\hat{A}|a_j\rangle = A^{*}_{j'}\langle a_{j'}|a_j\rangle. \tag{4.75}$$

[8] This of course sounds familiar. We will show soon that the stationary Schrödinger equation $\hat{H}\psi_j = E_j\psi_j$, which has been the focus of our studies in the first three chapters, is just a particular (coordinate) representation of Eq. (66) for one of (many possible) quantum-mechanical operators.

Now subtracting this equation from Eq. (73), we get

$$0 = \left(A_j - A_{j'}^* \right) \langle a_{j'} | a_j \rangle. \tag{4.76}$$

There are two possibilities to satisfy this equation. If indices *j* and *j'* denote the same state, then the bra-ket is the state's norm squared, and cannot be equal to zero. Then the left parentheses (with *j* = *j'*) have to be zero, i.e. Eq. (69) is valid. On the other hand, if *j* and *j'* correspond to different eigenvalues, the parentheses cannot equal zero (we have just proved that all A_j are real!), and hence the state vectors indexed by *j* and *j'* should be orthogonal, e.g., Eq. (68) is valid.

As will be discussed below, these properties make the Hermitian operators uniquely suitable for the description of physical observables.

4.4. Change of basis, unitary operators, and diagonalization

From the discussion of last section, it may look that the matrix language is fully similar to, and in many instances more convenient than the general bra-ket formalism. In particular, Eqs. (53), (65), (66) show that any part of any bra-ket expression may be directly mapped on the similar matrix expression, with the only slight inconvenience of using not only columns, but also rows (with their elements complex-conjugated), for state vector presentation. In this context, why do we need the bra-ket language at all? The answer is that the elements of matrices, rows and columns of quantum objects depend on the particular choice of the basis set, very much like the Cartesian components of a usual vector depend on the particular choice of coordinate frame orientation (Fig. 4).

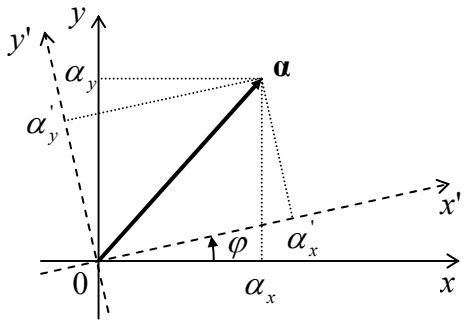

Fig. 4.4. Transformation of components of a 2D vector at coordinate frame rotation.

Let us study what happens if we change one basis, u_j, for another one, v_j (both full and orthonormal). First of all, let us prove that for each such bases pair, there exists such an operator $\hat{U}$ that, first,

$$|v_j\rangle = \hat{U}|u_j\rangle, \tag{4.77}$$

and, second,

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}. \tag{4.78}$$

(Due to the last property, operator $\hat{U}$ is called "*unitary*", and Eq. (77), a "*unitary transformation*".[9]) A very simple proof of both statements may be achieved by construction. Let us take[10]

$$\hat{U} \equiv \sum_k |v_k\rangle\langle u_k|. \tag{4.79}$$

Then

$$\hat{U}|u_j\rangle = \sum_k |v_k\rangle\langle u_k|u_j\rangle = \sum_k |v_k\rangle\delta_{kj} = |v_j\rangle, \tag{4.80}$$

so that Eq. (77) has been proven. Now, applying Eq. (31) to each term of sum (79), we get

$$\hat{U}^\dagger \equiv \sum_k |u_k\rangle\langle v_k|, \tag{4.81}$$

so that

$$\hat{U}\hat{U}^\dagger = \sum_{k,k'} |v_k\rangle\langle u_k|u_{k'}\rangle\langle v_{k'}| = \sum_{k,k'} |v_k\rangle\delta_{kk'}\langle v_{k'}| = \sum_k |v_k\rangle\langle v_k|. \tag{4.82}$$

But according to Eq. (44), the last expression is just the identity operator, QED.[11] (The proof of the second part of Eq. (78) is similar.)

As a by-product of our proof, we have also got another important expression (81). It means, in particular, that while operator $\hat{U}$ "rotates" the "old" basis u_j to the "new" basis v_j (see Eq. (77)), its Hermitian adjoint $\hat{U}^\dagger$ performs the reciprocal unitary transform:

$$\hat{U}^\dagger|v_j\rangle = \sum_k |v_k\rangle\delta_{kj} = |u_j\rangle. \tag{4.83}$$

Now, let us see how do the matrix elements of the unitary operator look. Generally, as was stated above, the elements depend on the basis we calculate them in, so we should be careful. For example, let us calculate the elements in basis u_j:

$$U_{jj'}|_{\text{in } u} \equiv \langle u_j|\hat{U}|u_{j'}\rangle = \langle u_j|\left(\sum_k |v_k\rangle\langle u_k|\right)|u_{j'}\rangle = \langle u_j|v_{j'}\rangle. \tag{4.84}$$

Now performing a similar calculation in basis v_j:

$$U_{jj'}|_{\text{in } v} \equiv \langle v_j|\hat{U}|v_{j'}\rangle = \langle v_j|\left(\sum_k |v_k\rangle\langle u_k|\right)|v_{j'}\rangle = \langle u_j|v_{j'}\rangle. \tag{4.85}$$

Surprisingly, the result is the same! This is of course true for the Hermitian conjugate as well:

$$U^\dagger_{jj'}|_{\text{in } u} = U^\dagger_{jj'}|_{\text{in } v} = \langle v_j|u_{j'}\rangle. \tag{4.86}$$

[9] This operation is a clear analog of the coordinate frame rotation in geometry.

[10] Notice that the complexity of this construct is somewhere in-between the identity operator (44) and expression (60) for an arbitrary operator.

[11] *Quod erat demonstrandum* (lat.) – what needed to be proven.

These expressions may be used, first of all, to present Eq. (77) in a more direct form. Applying to the "new" basis state $v_{j'}$ the first of Eqs. (41), we get

$$\left|v_{j'}\right\rangle = \sum_j \left|u_j\right\rangle\left\langle u_j\left|v_{j'}\right.\right\rangle = \sum_j \left|u_j\right\rangle U_{jj'}. \tag{4.87}$$

Similarly, the reciprocal transform is

$$\left|u_{j'}\right\rangle = \sum_j \left|v_j\right\rangle\left\langle v_j\left|u_{j'}\right.\right\rangle = \sum_j \left|v_j\right\rangle U^{\dagger}_{jj'}. \tag{4.88}$$

These equations are very convenient for some applications; we will use them already in this section.

Next, we may use Eqs. (85), (86) to express the effect of the unitary transform on expansion coefficients (37) of an arbitrary the state vector α. In the "old" basis u_j, they are given by Eq. (40). Similarly, in the "new" basis v_j,

$$\alpha_j\big|_{\text{in}\,v} = \left\langle v_j\left|\alpha\right.\right\rangle. \tag{4.89}$$

Again using the insertion of the identity operator (44), with the internal index *j'*, we get

$$\alpha_j\big|_{\text{in}\,v} = \left\langle v_j\left|\left(\sum_{j'}\left|u_{j'}\right\rangle\left\langle u_{j'}\right|\right)\right|\alpha\right\rangle = \sum_{j'} U^{\dagger}_{jj'}\left\langle u_{j'}\left|\alpha\right.\right\rangle = \sum_{j'} U^{\dagger}_{jj'}\alpha_{j'}\big|_{\text{in}\,u}, \tag{4.90}$$

(where for the second transition, we have used Eq. (86)). Both structurally and philosophically, this expression is similar to the transformation of components of a usual vector at coordinate frame rotation. For example, in two dimensions (Fig. 4):

$$\begin{pmatrix} \alpha_x' \\ \alpha_y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} \alpha_x \\ \alpha_y \end{pmatrix}. \tag{4.91}$$

(In this analogy, the unitary property (78) of operators $\hat{U}$ and $\hat{U}^{\dagger}$ corresponds to the fact that the determinant of the rotation matrix in Eq. (89) equals 1.) A word of caution: notice that if the transformation (77) from the "old" basis u_j to the "new" basis v_j is performed by the unitary operator $\hat{U}$, the change (90) of a state vectors components at this transformation requires its Hermitian conjugate $\hat{U}^{\dagger}$. Actually, this is also natural from the point of view of the geometric analog of the transformation (Fig. 4): if the "new" reference frame {*x', y'*} is obtained by counterclockwise rotation of the "old" frame {*x,y*} by some angle φ, for the observer rotating with the frame, vector $\vec{\alpha}$ (which is actually unchanged) rotates clockwise. The reciprocal transformation is (of course) performed by operator $\hat{U}$:

$$\alpha_j\big|_{\text{in}\,u} = \sum_{j'} U_{jj'}\alpha_{j'}\big|_{\text{in}\,v}. \tag{4.92}$$

Due to the analogy between expressions (87) and (90), and our old friend Eq. (64), it is tempting to skip indices in our new results by writing

$$\left|\alpha\right\rangle_{\text{in}\,v} = \hat{U}^{\dagger}\left|\alpha\right\rangle_{\text{in}\,u}, \quad \left|\alpha\right\rangle_{\text{in}\,u} = \hat{U}\left|\alpha\right\rangle_{\text{in}\,v}. \tag{4.93}$$

Since matrix elements of $\hat{U}$ and $\hat{U}^\dagger$ do not depend on the basis, such language is not too bad; still it should not be confused with the "genuine" (basis-independent) bra-ket equalities.

Now let us use the same trick of identity operator insertion, repeated twice, to find the transformation rule for matrix elements of an arbitrary operator:

$$A_{jj'}\Big|_{\text{in } v} \equiv \left\langle v_j \middle| \hat{A} \middle| v_{j'} \right\rangle = \left\langle v_j \middle| \left(\sum_k |u_k\rangle\langle u_k| \right) \hat{A} \left(\sum_{k'} |u_{k'}\rangle\langle u_{k'}| \right) \middle| v_{j'} \right\rangle = \sum_{k,k'} U^\dagger_{jk} A_{kk'}\Big|_{\text{in } u} U_{k'j'}; \quad (4.94)$$

absolutely similarly,

$$A_{jj'}\Big|_{\text{in } u} \equiv \sum_{k,k'} U_{jk} A_{kk'}\Big|_{\text{in } v} U^\dagger_{k'j'}. \quad (4.95)$$

In the spirit of Eq. (93), we may present these results in a compact bra-ket form:

$$\hat{A}\Big|_{\text{in } v} = \hat{U}^\dagger \hat{A}\Big|_{\text{in } u} \hat{U}, \quad \hat{A}\Big|_{\text{in } u} = \hat{U} \hat{A}\Big|_{\text{in } v} \hat{U}^\dagger. \quad (4.96)$$

These equations are of course compatible with the fact, proven earlier, that the unitary operator itself is base-independent. Another example is the identity operator:

$$\hat{I}\Big|_{\text{in } v} = \left(\hat{U}^\dagger \hat{I} \hat{U} \right)_{\text{in } u} = \left(\hat{U}^\dagger \hat{U} \right)_{\text{in } u} = \hat{I}\Big|_{\text{in } u}. \quad (4.97)$$

One more invariant of the basis change is the *trace* of any operator (or its matrix), defined as the sum of the diagonal terms of the matrix:

$$\text{Tr}\hat{A} \equiv \text{TrA} \equiv \sum_j A_{jj}. \quad (4.98)$$

The proof is left for the reader – see Problem 5.

The issue of unitary transformation is intimately connected with that of eigenstates and eigenvalues of an operator, which are both the solutions of Eq. (70). Indeed, let us assume that the eigenstates of operator $\hat{A}$ form an orthonormal set (we have already seen that this is possible at least for Hermitian operators), and find matrix elements of the operator in the basis of these states. For that, it is sufficient to inner-multiply Eq. (70), written for index *j'*, by the bra-vector of an arbitrary state of the same set:

$$\left\langle a_j \middle| \hat{A} \middle| a_{j'} \right\rangle = \left\langle a_j \middle| A_{j'} \middle| a_{j'} \right\rangle. \quad (4.99)$$

The left hand part is just the matrix element we are looking for, while the right hand part is just $A_{j'}\delta_{jj'}$. As a result, we see that the matrix is diagonal, with the diagonal consisting of eigenvalues:

$$A_{jj'} = A_j \delta_{jj'}. \quad (4.100)$$

In particular, in the eigenstate basis (only!), A_{jj} means the same as A_j. Thus the most important problem of finding the eigenvalues and eigenstates of an operator is equivalent to the *diagonalization* of its matrix, i.e. finding the basis in which it acquired form (98); than the diagonal gives us the eigenvalues, and the basis itself, the sought set of eigenstates.

Let us modify the above calculation by inner-multiplying Eq. (70) by a bra-vector of a different basis – say, the one in which we know the matrix elements $A_{jj'}$:

$$\left\langle u_k \left| \hat{A} \right| a_j \right\rangle = \left\langle u_k \left| A_j \right| a_j \right\rangle . \tag{4.101}$$

In the right hand part, we can move eigenvalue A_j out of the bra-ket, while in the left hand sign we can (as usual:-) insert the identity operator, in form (44), between the operator and the ket-vector:

$$\left\langle u_k \left| \hat{A} \sum_{k'} \right| u_{k'} \right\rangle \left\langle u_{k'} \left| a_j \right. \right\rangle = \left\langle u_k \left| A_j \right| a_j \right\rangle . \tag{4.102}$$

Moving the sum sign to the left, the right-hand part to the left hand part, and using the definition of the matrix elements we get

$$\sum_{k'} \left(A_{kk'} - A_j \delta_{kk'} \right) \left\langle u_{k'} \left| a_j \right. \right\rangle = 0 . \tag{4.103}$$

But the set of such equalities, for all N possible values of k, is just a system of linear, homogeneous equations for unknown c-numbers $\left\langle u_k \left| a_j \right. \right\rangle$ which, according to Eqs. (84) and (85), are nothing more than the matrix elements U_{kj} of a unitary matrix providing the required transformation from the initial basis u_j to the basis a_k which diagonalizes matrix **A**. The system may be presented in the matrix form

$$\begin{pmatrix} A_{11} - A_j & A_{12} & \ldots \\ A_{21} & A_{22} - A_j & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix} \begin{pmatrix} U_{1j} \\ U_{2j} \\ \ldots \end{pmatrix} = 0 , \tag{4.104}$$

and the condition of its consistency,

$$\begin{vmatrix} A_{11} - A & A_{12} & \ldots \\ A_{21} & A_{22} - A & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix} = 0, \tag{4.105}$$

plays the role of the characteristic equation of the system. We could drop index j in that equation, because it has N roots for parameter A, which we can call A_j; plugging each of them back into system (101), we can use it to find N matrix elements U_{kj} (k = 1, 2, …N) corresponding to this particular eigenvalue. [12] However, since equations (103) are homogeneous, they allow finding U_{kj} only to a constant multiplier. In order to ensure their normalization, i.e. the unitary character of matrix **U**, we may use the condition that all eigenvectors are normalized (just as the basis vectors are):

$$\left\langle a_j \left| a_j \right. \right\rangle = \sum_k \left\langle a_j \left| u_k \right. \right\rangle \left\langle u_k \left| a_j \right. \right\rangle = \sum_k \left| U_{kj} \right|^2 = 1, \tag{4.106}$$

for each k. This normalization completes the diagonalization.

As a simple example, let us diagonalize the so-called *Pauli matrices*

[12] A possible complication here is the degenerate case when characteristic equation gives several similar eigenvalues corresponding to different eigenvectors. In this the requirement of orthogonality of these states should be additionally enforced.

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.107}$$

(Though they have been introduced by a physicist, with a specific purpose to describe electron's spin, these matrices have a much broader significance, because they provide a linearly-independent 2×2 basis in the following sense: an arbitrary 2×2 matrix may be presented as the sum of the matrices and the 2×2 identity matrix,

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = a_0 \mathrm{I} + a_x \sigma_x + a_y \sigma_y + a_z \sigma_z, \tag{4.108}$$

with a unique set of coefficients a.) Let us start with $\boldsymbol{\sigma}_x$. For it, the characteristic equation (103) is

$$\begin{vmatrix} -A & 1 \\ 1 & -A \end{vmatrix} = 0, \tag{4.109}$$

and has two roots, $A_{1,2} = \pm 1$. Notice that the eigenvalues of this matrix are the same as of matrix $\boldsymbol{\sigma}_z$ (and matrix $\boldsymbol{\sigma}_y$). However, eigenvectors of all these matrices are different! To find them for $\boldsymbol{\sigma}_x$, let us plug is first eigenvalue, $A_1 = +1$, back into equations (103):

$$\begin{aligned} -\langle u_1 | a_1 \rangle + \langle u_2 | a_1 \rangle = 0, \\ \langle u_1 | a_1 \rangle - \langle u_2 | a_1 \rangle = 0. \end{aligned} \tag{4.110}$$

The equations are compatible (of course, because the used eigenvalue $A_1 = +1$ satisfies the characteristic equation (109)), and give

$$\langle u_1 | a_1 \rangle = \langle u_2 | a_1 \rangle, \quad \text{i.e.} \quad U_{11} = U_{21}. \tag{4.111}$$

With that, the normalization condition (106) gives

$$|U_{11}|^2 = |U_{21}|^2 = \frac{1}{2}. \tag{4.112}$$

Although the normalization is insensitive to the simultaneous multiplication of U_{11} and U_{21} by the same phase factor $\exp\{i\varphi\}$ with any real φ, it is convenient to keep the coefficients real, i.e. take

$$U_{11} = U_{21} = \frac{1}{\sqrt{2}}. \tag{4.113}$$

Performing an absolutely similar calculation for the second characteristic value, $A_2 = -1$, we get

$$U_{22} = -U_{12} = \frac{1}{\sqrt{2}}, \tag{4.114}$$

so that the whole unitary matrix for diagonalization of $\boldsymbol{\sigma}_x$ is[13]

$$\mathrm{U}_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad .\mathrm{U}_x^{\dagger} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{4.115}$$

[13] Note that these unitary operators are *not* Hermitian.

For what follows, it will be convenient to have this result in the ket-relation form – see Eqs. (87), (88):

$$|a_1\rangle = U_{11}|u_1\rangle + U_{21}|u_2\rangle = \frac{1}{\sqrt{2}}\Big(|u_1\rangle + |u_2\rangle\Big), \qquad |a_2\rangle = U_{12}|u_1\rangle + U_{22}|u_2\rangle = \frac{1}{\sqrt{2}}\Big(-|u_1\rangle + |u_2\rangle\Big), \quad (4.116)$$

$$|u_1\rangle = U_{11}^{\dagger}|a_1\rangle + U_{21}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\Big(|a_1\rangle - |a_2\rangle\Big), \qquad |u_2\rangle = U_{12}^{\dagger}|a_1\rangle + U_{22}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\Big(|a_1\rangle + |a_2\rangle\Big), \quad (4.117)$$

Acting absolutely similarly, for matrices σ_y and σ_z we get the same eigenvalues ($a_{1,2} = \pm 1$), but different diagonalizing unitary matrices – see optional Problem O.15.

These results are already sufficient to understand the Stern-Gerlach experiments described in Sec. 1, with just one additional assumption, specific for electron's spin, namely that free electron's interaction with external magnetic field may be described by a vector operator of magnetic moment

$$\hat{\mathbf{m}} = -g_e \frac{e}{2m_e}\hat{\mathbf{S}}, \qquad \hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}, \quad (4.118)$$

where $g_e = 2$ is electron's "*g*-factor,[14] the vector operator $\hat{\boldsymbol{\sigma}}$ in presented, in "*z-basis*", by the vector of 3 Pauli matrices (107):

$$\hat{\boldsymbol{\sigma}} = \mathbf{n}_x\hat{\sigma}_x + \mathbf{n}_y\hat{\sigma}_y + \mathbf{n}_z\hat{\sigma}_z, \quad (4.119)$$

and $\mathbf{n}_{x,y,z}$ are the usual Cartesian unit vectors in 3D pace. (In the quantum-mechanics sense, they are just *c*-numbers, or rather "*c*-vectors".) The *z*-basis, by definition, is an orthonormal basis of two states, typically denoted ↑ an ↓, in which the *z*-component $\hat{S}_z$ of the vector operator of spin is diagonal, with eigenvalues $+\hbar/2$ and $-\hbar/2$. Notice that we do not "know" what exactly these states are,[15] but loosely associate them with certain internal rotation of the electron about *z*-axis, with either positive or negative angular momentum component S_z. Mathematically, states ↑ an ↓ may be accepted as states called $u_{1,2}$ in the $\boldsymbol{\sigma}_x$ matrix diagonalization problem solved above, because in this basis, matrices $\boldsymbol{\sigma}_z$ and hence $\mathbf{S}_z$, are diagonal.

Actually, at this stage the description, the first (*z*-oriented) SG experiment shown in Fig. 1 is the hardest for us, because the statistical ensemble describing the unpolarized electron beam at its entrance is "*mixed*" (incoherent), and cannot be described by a coherent superposition of the type (6), which have been the subject of our studies so far. (We will discuss the mixed ensembles, and in particular the SG experiment results, in Chapter 6.) However, it is intuitively clear that its results, and in particular Eq. (5), are compatible with the description of its two output beams, of equal intensity, as sets of electrons in eigenstates ↑ an ↓, respectively. In this picture, the absorber following that experiment (Fig. 2) just takes out all spin-down electrons out of the picture, producing an output beam of polarized electrons in a pure ↑ state. This certainly explains the results of the "control" experiment (the bottom panel in Fig. 2): the repeated SG-*z* experiment just deviates the whole beam a bit more, but does not split it.

---

[14] Actually, due to quantum electrodynamics effects, is slightly higher than two: $g_e \approx 2.002319304\ldots \approx 2(1 + \alpha/2\pi + \ldots)$, where $\alpha \equiv e^2/4\pi\varepsilon_0\hbar c \approx 1/137$ is the fine structure constant.

[15] If you think about it, words "know" or "understand" typically mean that we can explain a more complex notion in terms of more simple entities, already considered "known". In our example, we cannot express the spin states by some wavefunction $\psi(\mathbf{r})$, or any other mathematical notion we are familiar with. The bra-ket formalism has been invented exactly to enable mathematical operations with such "mysterious" quantum states.

Now let us discuss the double SG experiment shown on the top panel of Fig. 2. For that, let us present the z-polarized beam in another basis of states (which we will denote, respectively, as $\rightarrow$ and $\leftarrow$) in which, by definition, matrix $S_x = (\hbar/2)\sigma_x$ is diagonal. But this is exactly the set we called $a_{1,2}$ in the problem above. Hence, in application to the spin problem, we can rewrite Eqs. (116), (117) as

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right), \qquad |\leftarrow\rangle = \frac{1}{\sqrt{2}}\left(-|\uparrow\rangle + |\downarrow\rangle\right), \tag{4.120}$$

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\left(|\rightarrow\rangle - |\leftarrow\rangle\right), \qquad |\downarrow\rangle = \frac{1}{\sqrt{2}}\left(|\rightarrow\rangle + |\leftarrow\rangle\right), \tag{4.121}$$

Currently, for us the first of Eqs. (121) is most important, because it shows that the quantum state of electrons entering the SG-*x* stage may be presented as a coherent superposition of electrons with $S_x = +\hbar/2$ and $S_x = -\hbar/2$. Notice that according to postulate (9), the beams have the same intensity. (The minus sign before the second ket-vector is of no consequence here, though may have an impact on outcome of other experiments with coherent states.)

Now, let us discuss the most mysterious (from the classical point of view) multi-stage SG experiment shown on the middle panel of Fig. 2. After the second absorber has taken out all electrons in, say, the "$\leftarrow$" state, the remaining, *x*-polarized electrons are passed to the input of the final, SG-*z*, stage. But according to the first of Eqs. (120), this state may be presented as a (coherent) superposition of the $\uparrow$ and $\downarrow$ states, with equal amplitudes. The stage separates these two states into separate beams, with an equal (50%) chance to find an electron in each of them, thus explaining the experimental results.

To conclude our discussion of the multistage SG experiment, let us notice that though is cannot be explained in terms of wave mechanics (which operates with *scalar* de Broglie waves), it has an analogy in classical theory of *vector* fields, such as classical electrodynamics – see Fig. 3.

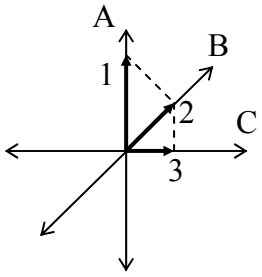

Fig. 4.3. The light polarization system similar to the 3-stage SF experiment shown in Fig. 2.

Let a plane electromagnetic wave propagate perpendicular to the plane of drawing, and pass it through polarizer A. Similarly to the initial SG-*z* stages (including absorbers) shown in Fig. 2, it produces a linearly polarized wave with amplitude 1. The wave vector has no horizontal component, as may be revealed by its full absorption in polarizer C. However, let us preliminary pass the wave through polarizer B (which is an analog of the SG-*x* experiment, again including the absorber, in the middle panel of Fig. 2). In this case, the output wave does have a horizontal component, as can be, again, revealed by passing it through polarizer C, If the angles between polarization axes of A, B, and C all equal $\pi/4$, each polarizer reduces the wave amplitude by the factor of $\sqrt{2}$, and hence intensity by a factor of 2, exactly like in the SG experiment. The only difference is that the necessary angles are $\pi/4$, rather than by $\pi/2$ for SG experiment; this is one of peculiarities of electron's half-integer spin.

4.5. Averages and the generalized uncertainty relation

Our results (120), (121) are actually quite deep, and deserve more discussion. For that we need to complement Eq. (9) by one more postulate describing the relation between the quantum mechanics formalism and experiment.[16] Namely, in quantum mechanics, each observable A is by a Hermitian operator $\hat{A}$ and the result of its "measurement" may be only one of eigenvalues A_j of its operator. (As we already know, all these eigenvalues are real).

Let us explore the corollaries of this postulate. If the eigenstates corresponding to eigenvalues A_j form a full set, we can use Eqs. (37) and (40) to write, for any coherent state,

$$\left|\alpha\right\rangle = \sum_j \alpha_j \left|a_j\right\rangle, \quad \text{with } \alpha_j = \left\langle a_j \middle| \alpha \right\rangle. \tag{4.122}$$

Then, according to Eq. (9), the probability to measure value A_j is

$$W_j = \left|\alpha_j\right|^2 = \alpha_j^* \alpha_j = \left\langle \alpha \middle| a_j \right\rangle \left\langle a_j \middle| \alpha \right\rangle. \tag{4.123}$$

(As a sanity check, if both the composite state and the basis state vectors are normalized: $\left\langle \alpha \middle| \alpha \right\rangle = 1$, $\left\langle a_j \middle| a_j \right\rangle = 1$, as we will assume below, then

$$\sum_j W_j = \sum_j \left\langle \alpha \middle| a_j \right\rangle \left\langle a_j \middle| \alpha \right\rangle = \left\langle \alpha \middle| \hat{I} \middle| \alpha \right\rangle = 1, \tag{4.124}$$

i.e. the system is certainly in *some* of the eigenstates.)

The average (expectation) value of the observable over *any* statistical ensemble may be found using the ordinary ("classical", or rather mathematical) rule (1.36):

$$\left\langle A \right\rangle = \sum_j A_j W_j. \tag{4.125}$$

For the particular case of coherent state (122), we can combine it with Eq. (123), to get

$$\left\langle A \right\rangle = \sum_j \left\langle \alpha \middle| a_j \right\rangle A_j \left\langle a_j \middle| \alpha \right\rangle = \sum_j \left\langle \alpha \middle| a_j \right\rangle \left\langle a_j \middle| \hat{A} \middle| a_j \right\rangle \left\langle a_j \middle| \alpha \right\rangle = \sum_{j,k} \left\langle \alpha \middle| a_j \right\rangle \left\langle a_j \middle| \hat{A} \middle| a_k \right\rangle \left\langle a_k \middle| \alpha \right\rangle = \left\langle \alpha \middle| \hat{A} \middle| \alpha \right\rangle. \tag{4.126}$$

This is an analog of the wave-mechanics formula (1.23); its advantage over Eq. (123) is that it does not explicitly involve the eigenvector set, which is sometimes very convenient. The similar rule may be used to calculate the uncertainty (r.m.s. deviation from the average) defined by Eq. (1.32), (1.32):

$$(\delta A)^2 \equiv \left\langle \widetilde{A}^2 \right\rangle = \left\langle A^2 \right\rangle - \left\langle A \right\rangle^2, \quad \text{where } \widetilde{A} \equiv A - \left\langle A \right\rangle. \tag{4.127}$$

Now let us use these equations for a better look at spin state $\uparrow$. Its coefficients of expansion over the z-basis with ket-vectors $\left|\uparrow\right\rangle$ and $\left|\downarrow\right\rangle$, by definition, 1 and 0, correspondingly. This fact, together with Eqs. (107), (117), and (118), enables us to calculate any averages we like. To start with, the average spin components are:

---

16 Notice, however, that just like in Chapter 1, we are still dealing with the mathematical notion of "measurements", postponing the discussion of real (physical) measurements until Chapter 6.

$$\langle S_x \rangle = \frac{\hbar}{2}\langle \uparrow | \hat{\sigma}_x | \uparrow \rangle = \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = 0, \tag{4.128}$$

and the similar result for S_y, but

$$\langle S_z \rangle = \frac{\hbar}{2}\langle \uparrow | \hat{\sigma}_z | \uparrow \rangle = \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = +\frac{\hbar}{2}. \tag{4.129}$$

(In the state ↓, the result is of course -$\hbar$/2.) This is natural, but let us have a look at the uncertainties:

$$(\delta S_x)^2 = \langle S_x^2 \rangle - \langle S_x \rangle^2 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{\sigma}_x^2 | \uparrow \rangle = \left(\frac{\hbar}{2}\right)^2 \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} = \left(\frac{\hbar}{2}\right)^2. \tag{4.130}$$

and similarly for $(\delta S_y)^2$.

In order to show that this is the maximum possible uncertainties for these observables, the direct use of Eqs. (9), (125) may be more convenient. First of all, let us use it to re-calculate the average spin components in ↑ states. Evidently, $W_\uparrow = 1$, and $W_\downarrow = 0$, while according to the first of Eqs. (121), $W_\rightarrow = W_\leftarrow = ½$. Hence,

$$\langle S_z \rangle = \left[W_\uparrow \left(+\frac{\hbar}{2}\right) + W_\downarrow \left(-\frac{\hbar}{2}\right)\right] = \frac{\hbar}{2}(1-0) = +\frac{\hbar}{2}, \tag{4.131}$$

$$\langle S_x \rangle = \left[W_\rightarrow \left(+\frac{\hbar}{2}\right) + W_\leftarrow \left(-\frac{\hbar}{2}\right)\right] = \frac{\hbar}{2}\left(\frac{1}{2}-\frac{1}{2}\right) = 0, \tag{4.132}$$

thus reproducing results (128) and (129). Now let us calculate the uncertainty of that observable, for a minute without specifying $W_\rightarrow$ and $W_\leftarrow$

$$(\delta S_x)^2 = \langle S_x^2 \rangle - \langle S_x \rangle^2 = W_\rightarrow \left(S_x^2\right)_\rightarrow + W_\leftarrow \left(S_x^2\right)_\leftarrow - \left(\frac{\hbar}{2}\right)^2 (W_\rightarrow - W_\leftarrow)^2. \tag{4.133}$$

At this point, we may use the fact (see Problem 3) that the squares of all Pauli matrices are identity matrices, so their eigenvalues in any basis are +1, so that

$$(\delta S_x)^2 = \left(\frac{\hbar}{2}\right)^2 \left[(W_\rightarrow + W_\leftarrow) - (W_\rightarrow - W_\leftarrow)^2\right] = \left(\frac{\hbar}{2}\right)^2 \left[1 - (W_\rightarrow - W_\leftarrow)^2\right]. \tag{4.134}$$

It is evident that in the case $W_\rightarrow = W_\leftarrow$, we return to Eq. (130), and it is the maximum possible value of the uncertainty possible in the system. However, an absolutely similar calculation for the uncertainty of S_z gives a completely different result:[17]

$$(\delta S_z)^2 = \left(\frac{\hbar}{2}\right)^2 \left[1 - (W_\uparrow - W_\downarrow)^2\right] = 0. \tag{4.135}$$

This result is certainly compatible with the notion of the electron being "definitely in ↑ state".

[17] We could of course obtain the same result using the direct matrix calculation, similar to Eq. (130), but the reader has to agree that the explicit expression via probabilities is physically more transparent.

Even without calculations, it is evident that in the *x*-polarized or *y*-polarized state the situation is similar, with the corresponding change of indices. Thus, we see that in neither state all 3 components of spin cannot be "measured" (or better say "known") exactly. Let us show that this is not just an occasional fact, but reflects the most profound property of quantum mechanics, the uncertainty relations.

Now let consider 2 observables, A and B, measured in the same quantum state. There are two possibilities here. If operators corresponding to the observables commute,

$$\left[\hat{A},\hat{B}\right]=0\,, \tag{4.136}$$

then all matrix elements of the commutator, in any orthogonal basis (for example, the eigenbasis a_j of operator A), are also zero. From here, we get

$$\left\langle a_j\left|\left[\hat{A},\hat{B}\right]\right|a_{j'}\right\rangle=\left\langle a_j\left|\hat{A}\hat{B}\right|a_{j'}\right\rangle-\left\langle a_j\left|\hat{B}\hat{A}\right|a_{j'}\right\rangle=0\,. \tag{4.137}$$

In the first bra-ket of the middle expression, we can act by the (Hermitian!) operator $\hat{A}$ on the bra-vector, and in the second one, on the ket-vector. The result is

$$A_j\left\langle a_j\left|\hat{B}\right|a_{j'}\right\rangle-A_{j'}\left\langle a_j\left|\hat{B}\right|a_{j'}\right\rangle=\left(A_j-A_{j'}\right)\left\langle a_j\left|\hat{B}\right|a_{j'}\right\rangle=0. \tag{4.138}$$

This means that the matrix of operator B has to be diagonal in basis a_j, i.e., the sets eigenstates of operators $\hat{A}$ and $\hat{B}$ coincide. Such pairs of observables, which share their eigenstates, are called "*compatible*". As example, in wave mechanics of a free particle, the vector of momentum an the (scalar of) kinetic energy were compatible, sharing eigenfunctions. Now we see that this is not occasional, because each component of the kinetic energy is proportional to the square of the corresponding component of the momentum, and any operator commutes with its power:

$$\left[\hat{A},\hat{A}^n\right]\equiv\left[\hat{A},\underbrace{\hat{A}\hat{A}...\hat{A}}_{n}\right]=\hat{A}\underbrace{\hat{A}\hat{A}...\hat{A}}_{n}-\underbrace{\hat{A}\hat{A}...\hat{A}}_{n}\hat{A}=0\,. \tag{4.139}$$

Now, what if operators $\hat{A}$ and $\hat{B}$ do not commute? Then the following *generalized uncertainty relation* is valid:[18]

$$\delta A\,\delta B\geq\frac{1}{2}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|. \tag{4.140}$$

The proof of Eq. (140) may be broken into two steps, the first of which proves the so-called *Schwartz inequality*:

$$\left\langle\alpha\middle|\alpha\right\rangle\left\langle\beta\middle|\beta\right\rangle\geq\left|\left\langle\alpha\middle|\beta\right\rangle\right|^2. \tag{4.141}$$

This is just the quantum-mechanics analog of the fact from vector algebra: $\alpha^2\beta^2\geq(\vec{\alpha}\cdot\vec{\beta})^2$. Though this relation is geometrically evident (because its right-hand part includes the cosine of the angle between the vectors), the bra-ket algebra is somewhat different from the vector algebra, so we still need

[18] Notice that both sides are state-specific; the uncertainty relation statement is that Eq. (140) should be valid for *any* state.

to reproduce the proof. It is started by noting that postulate (16) that the norm of any legitimate state of the system cannot be negative. Let us apply it to the state with the following ket-vector

$$|\delta\rangle = |\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}|\beta\rangle, \tag{4.142}$$

where α and β are possible, non-null states of the system. Then Eq. (16) takes the form

$$\left(\langle\alpha| - \frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}\langle\beta|\right)\left(|\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}|\beta\rangle\right) \geq 0. \tag{4.143}$$

Opening the parentheses, we get

$$\langle\alpha|\alpha\rangle + \frac{\langle\alpha|\beta\rangle\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle^2}\langle\beta|\beta\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}\langle\beta|\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}\langle\alpha|\beta\rangle \geq 0. \tag{4.144}$$

After the cancellation of one product $\langle\beta|\beta\rangle$ in the second term, it cancels with the 3rd (or 4th) term, proving the Schwartz inequality (141).

Now let us apply the inequality to states

$$|\alpha\rangle \equiv \hat{\tilde{A}}|\gamma\rangle \quad \text{and} \quad |\beta\rangle \equiv \hat{\tilde{B}}|\gamma\rangle, \tag{4.145}$$

where in both relations γ is the same (but otherwise arbitrary) possible state of the system, and the deviation operators are defined similarly to the corresponding observables, for example

$$\hat{\tilde{A}} \equiv \hat{A} - \langle A\rangle. \tag{4.146}$$

With this substitution, and taking into account that operators $\hat{A}$ and $\hat{B}$ are Hermitian, Eq. (141) yields

$$\langle\gamma|\hat{\tilde{A}}^2|\gamma\rangle\langle\gamma|\hat{\tilde{B}}^2|\gamma\rangle \geq \left|\langle\gamma|\hat{\tilde{A}}\hat{\tilde{B}}|\gamma\rangle\right|^2. \tag{4.147}$$

Since state γ is arbitrary, we may rewrite it as the operator inequality, we get

$$\delta A\,\delta B \geq \left|\left\langle\hat{\tilde{A}}\hat{\tilde{B}}\right\rangle\right|. \tag{4.148}$$

This is actually already an uncertainty relation, and in some sense even "better" (stronger) than the traditional form (140); moreover, it is more convenient in some cases. In order to continue to Eq. (140), we need a couple more steps. Let us notice that the operator product in the right-hand part of Eq. (148) may be presented as

$$\hat{\tilde{A}}\hat{\tilde{B}} = \frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\} + \frac{1}{2}\left[\hat{\tilde{A}},\hat{\tilde{B}}\right]. \tag{4.149}$$

Now, any anticommutator of Hermitian operators, including that in Eq. (149), is a Hermitian operator, and its eigenvalues are purely real. On the other hand, the commutator in that equation is just

$$\left[\hat{\tilde{A}},\hat{\tilde{B}}\right]=\left(\hat{A}-\langle A\rangle\right)\left(\hat{B}-\langle B\rangle\right)-\left(\hat{B}-\langle B\rangle\right)\left(\hat{A}-\langle A\rangle\right)=\hat{A}\hat{B}-\hat{B}\hat{A}=\left[\hat{A},\hat{B}\right], \tag{4.150}$$

and its Hermitian conjugate is

$$\left[\hat{A},\hat{B}\right]^{\dagger}=(\hat{A}\hat{B})^{\dagger}-(\hat{B}\hat{A})^{\dagger}=\hat{B}\hat{A}-\hat{A}\hat{B}=-\left[\hat{A},\hat{B}\right], \tag{4.151}$$

so it may be presented as a product of some Hermitian operator by i. Since the eigenvalues of the Hermitian operator are purely real, those of the commutator are purely imaginary. As a result, the eigenvalues of operator in the right-hand part of Eq. (148), in any orthonormal basis, may be presented as

$$\left(\hat{\tilde{A}}\hat{\tilde{B}}\right)_j=\frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\}_j-\frac{i}{2}\left(i\left[\hat{A},\hat{B}\right]\right)_j, \tag{4.152}$$

where both eigenvalues in the right-hand part are real numbers, so that

$$\left|\hat{\tilde{A}}\hat{\tilde{B}}\right|_j^2=\frac{1}{4}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\}_j^2+\frac{1}{4}\left|\left[\hat{A},\hat{B}\right]\right|_j^2\geq\frac{1}{4}\left|\left[\hat{A},\hat{B}\right]\right|_j^2, \tag{4.153}$$

and we can extend Eq. (148) as

$$\delta A\,\delta B\geq\left|\left\langle\hat{\tilde{A}}\hat{\tilde{B}}\right\rangle\right|=\left|\sum_j P_j\left|\hat{\tilde{A}}\hat{\tilde{B}}\right|_j\right|\geq\frac{1}{2}\left|\sum_j P_j\left|\left[\hat{A},\hat{B}\right]\right|_j\right|=\frac{1}{2}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|, \tag{4.154}$$

thus proving Eq. (140).

For the particular case of spin-1/2 operators, it is easy to calculate (see Problem 3):

$$\left[\hat{\sigma}_x,\hat{\sigma}_y\right]=2i\hat{\sigma}_z, \tag{4.155}$$

with similar relations for other pairs of indices. As a result, the uncertainty relations for half-integer spin particles like electrons are

$$\delta S_x\,\delta S_y\geq\frac{\hbar}{2}\left|\langle S_z\rangle\right|,\quad\text{etc.} \tag{4.156}$$

In particular, in the $\uparrow$ state, the right-hand part of this relation equals $(\hbar/2)^2$, and neither of the uncertainties δS_x, δS_y can vanish. Our direct calculation earlier in the section (see, e.g., Eq. (130)) says that they are equal, and correspond the *lowest* value allowed by the uncertainty relation (156). In the latter aspect, the spin-polarized state is similar to the Gaussian wave packet studied in Sec. 2.1.

4.6. Quantum dynamics: The Schrödinger and Heisenberg pictures

So far in this chapter, we have shied away from the discussion of system dynamics, implying that bra- and ket-vectors and operators of the system are their "snapshots" at certain instant t. Now it is a good time to examine their time dependence. One of the most peculiar (and beautiful!) features of quantum mechanics is that the time evolution may be described using either of two alternative "pictures", giving exactly the same final results.

With our wave mechanics experience, the most natural of those is the *Schrödinger picture*. In this approach, the operators corresponding to time-independent observables (e.g., the Hamiltonian function H of a closed system) are also constant, while the bra- and ket-vectors change in time as

$$\left|\alpha(t)\right\rangle = \hat{\boldsymbol{u}}(t,t_0)\left|\alpha(t_0)\right\rangle, \quad \left\langle\alpha(t)\right| = \left\langle\alpha(t_0)\right|\hat{\boldsymbol{u}}^{\dagger}(t,t_0), \tag{4.157}$$

where $\hat{\boldsymbol{u}}(t,t_0)$ is the so-called *time-evolution operator* which obeys the following differential equation:

$$i\hbar\frac{\partial\hat{\boldsymbol{u}}}{\partial t} = \hat{H}\hat{\boldsymbol{u}}, \tag{4.158}$$

where $\hat{H}$ is the Hamiltonian operator of the system (which is always Hermitian: $\hat{H}^{\dagger} = \hat{H}$), and the partial derivative sign emphasizes the fact that the differentiation is over just one of two arguments of the evolution operator. While Eq. (158) is a natural replacement of the wave-mechanical equation (1.24) (and is also frequently called the *Schrödinger equation*), it still should be considered as a new, more general postulate which finds its final justification, as it is usual in physics, in the overwhelming agreement between its corollaries and experiment (and even more importantly, in having not a single reliable contradiction with experiment).[19]

Starting the discussion of Eqs. (157)-(158), let us first consider the case of time-independent Hamiltonian $\hat{H}$ whose eigenstates a_n (the "stationary states" of the system), which obey Eq. (70):

$$\hat{H}\left|a_n\right\rangle = E_n\left|a_n\right\rangle, \tag{4.159}$$

Let us see what is the time evolution of the expansion coefficients defined by Eq. (37) (or, equivalenty, by Eq. (40)):

$$\begin{aligned}\frac{\partial}{\partial t}\alpha_n(t) &\equiv \frac{\partial}{\partial t}\left\langle a_n\middle|\alpha(t)\right\rangle = \frac{\partial}{\partial t}\left\langle a_n\middle|\hat{\boldsymbol{u}}(t,t_0)\middle|\alpha(t_0)\right\rangle = \left\langle a_n\middle|\frac{\partial}{\partial t}\hat{\boldsymbol{u}}(t,t_0)\middle|\alpha(t_0)\right\rangle \\ &= \left\langle a_n\middle|\frac{1}{i\hbar}\hat{H}\hat{\boldsymbol{u}}(t,t_0)\middle|\alpha(t_0)\right\rangle = \frac{E_n}{i\hbar}\left\langle a_n\middle|\hat{\boldsymbol{u}}(t,t_0)\middle|\alpha(t_0)\right\rangle = \frac{E_n}{i\hbar}\left\langle a_n\middle|\alpha(t)\right\rangle = -\frac{i}{\hbar}E_n\alpha_n.\end{aligned} \tag{4.160}$$

This is the same (easy) equation as Eq. (1.44a), and its integration yields the same result – cf. Eq. (1.45):

$$\alpha_n(t) = \alpha_n(t_0)\exp\left\{-\frac{i}{\hbar}E_n(t-t_0)\right\}. \tag{4.161}$$

Now let us discuss the following fascinating fact: the general solution of the operator equation (158) may be presented in a similar but operator form. For the easiest case when the Hamiltonian is time-independent this equation is an exact analog of Eq. (161),

$$\hat{\boldsymbol{u}}(t,t_0) = \hat{\boldsymbol{u}}(t_0,t_0)\exp\left\{-\frac{i}{\hbar}\hat{H}(t-t_0)\right\} = \hat{I}\exp\left\{-\frac{i}{\hbar}\hat{H}(t-t_0)\right\}. \tag{4.162}$$

To start the proof we should, first of all, sort out what is a function (in this case, the exponent) of an operator. In the operator (and matrix) algebra, such functions are defined by their Taylor expansion; in particular, Eq. (162) means that

[19] Moreover, we will be able to *derive* Eq. (1.24) from Eq. (158) in the next section.

$$\hat{u}(t,t_0) = \hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2 + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^3 \hat{H}^3(t-t_0)^3 + ..., \quad (4.163)$$

where $\hat{H}^2 = \hat{H}\hat{H}$, $\hat{H}^3 = \hat{H}\hat{H}\hat{H}$, etc. Working with such infinite series of operator products is not as hard as one could imagine, due to their regular structure. For example, let us differentiate Eq. (163) over time:

$$\begin{aligned} \frac{\partial}{\partial t}\hat{u}(t,t_0) &= \hat{0} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H} + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2 2(t-t_0) + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^3 3(t-t_0)^2 + ... \\ &= \left(-\frac{i}{\hbar}\right)\hat{H}\left[\hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2\right] + ... = -\frac{i}{\hbar}\hat{H}\hat{u}(t,t_0), \end{aligned} \quad (4.164)$$

so that the differential equation (158) is indeed satisfied. On the other hand, Eq. (162) also satisfies the initial condition

$$\hat{u}(t_0,t_0) = \hat{u}^{\dagger}(t_0,t_0) = \hat{I}, \quad (4.165)$$

which follows from the definition (157) of the evolution operator.

Now let us allow operator $\hat{H}$ to be a function of time, but with the condition that its "values" (in fact, operators) at different moments commute with each other:

$$\left[\hat{H}(t'), \hat{H}(t'')\right] = 0. \quad (4.166)$$

(An important example of such a Hamiltonian is that of a particle under the effect of classical, time-dependent force:

$$\hat{H}_F = -\mathbf{F}(t)\cdot\hat{\mathbf{r}}, \quad (4.167)$$

possibly on the top of the time-independent Hamiltonian (1.47). Indeed, each operator, including $\hat{\vec{r}}$, evidently commutes with itself, as well as with the "*c*-number function" $F(t)$.) In this case it is sufficient to replace, in all formulas discussed above, product $\hat{H}(t-t_0)$ by the corresponding integral (similar to that used in the WKB approximation); in particular, Eq. (168) is generalized as

$$\hat{u}(t,t_0) = \hat{I}\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}. \quad (4.168)$$

A similar replacement should be made in Eq. (163), so it may be rewritten in the form

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\left(\int_{t_0}^{t}\hat{H}(t')dt'\right)^k. \quad (4.169)$$

Its proof is absolutely similar to that carried out above.

We may now use Eq. (173) to show that the evolution operator is unitary, even for the time-dependent Hamiltonian. (This is necessary to keep intact the system normalization, so that the total probability to find the system in its evolving state,

$$W_\alpha(t) = \langle \alpha(t) | \alpha(t) \rangle = \langle \alpha(t_0) | \hat{u}^\dagger \hat{u} | \alpha(t_0) \rangle = \langle \alpha(t_0) | \hat{I} | \alpha(t_0) \rangle = P_\alpha(t_0), \tag{4.170}$$

did not depend on time.) Indeed,

$$\hat{u}(t,t_0)\hat{u}^\dagger(t,t_0) = \hat{I} \exp\left\{ -\frac{i}{\hbar} \int_{t_0}^{t} \hat{H}(t')dt' \right\} \hat{I} \exp\left\{ +\frac{i}{\hbar} \int_{t_0}^{t} \hat{H}(t'')dt'' \right\}. \tag{4.171}$$

Since each of the exponents may be presented by the Taylor series (169), and thanks to Eq. (166) different terms of these sums may be swapped at will, expression (171) may be manipulated exactly as the product of *c*-number exponents, in particular to rewrite it as

$$\hat{u}(t,t_0)\hat{u}^\dagger(t,t_0) = \hat{I} \exp\left\{ -\frac{i}{\hbar} \left[\int_{t_0}^{t} \hat{H}(t')dt' - \int_{t_0}^{t} \hat{H}(t'')dt'' \right] \right\} = \hat{I} \exp\{\hat{0}\} = \hat{I}. \tag{4.172}$$

The most difficult case is when Eq. (166) is violated; an example is an orbital motion of a particle in the field of a classical force depending not only on time but on the particle's position as well. In that case, it may be proven that Eq. (169) should be replaced with the so-called *Dyson series*

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty} \left(-\frac{i}{\hbar} \right)^k \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{k-1}} dt_k \hat{H}(t_1)\hat{H}(t_2)\ldots\hat{H}(t_k). \tag{4.173}$$

Since we will not have time to run into such situations in our course, I will skip the proof.[20]

As a simple but important example of quantum dynamics, which defies the wave-mechanics treatment, let us consider electron spin precession in time-independent, uniform external magnetic field, taking its direction for *z* axis. To construct the system's Hamiltonian, we may use the classical expression for the energy of a magnetic moment **m** in the field[21]

$$U = -\mathbf{m} \cdot \mathbf{B}, \tag{4.174}$$

but in quantum case we should consider the moment as an operator (see Eq. (118)), so that the interaction Hamiltonian is

$$\hat{H} = -\hat{\mathbf{m}} \cdot \mathbf{B} = \frac{e}{m_e} \hat{\mathbf{S}} \cdot \mathbf{B} = \frac{eB}{m_e} \hat{S}_z, \tag{4.175}$$

where $\hat{S}_z$ is the operator of *z*-component of electron's spin. Using the *z*-basis of states ($\uparrow$ and $\downarrow$, see the previous section), this equation becomes

$$\hat{H} = \frac{\hbar e B}{2m_e} \hat{\sigma}_z = \frac{\hbar\Omega}{2} \hat{\sigma}_z, \quad \Omega \equiv \frac{eB}{m_e}, \tag{4.176}$$

where Ω may be interpreted as the classical frequency of the precession of a symmetric top, with angular moment $\vec{S}$, about axis *z*, induced by external torque $\vec{\tau} = \tau\,(\vec{n}_S \times \vec{n}_z)$:[22]

[20] It may be found, for example, in Chapter 5 of textbook J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley, 1994.

[21] Actually, we have already used this expression for the derivation of Eq. (3).

[22] See, e.g., CM Sec. 5.6, in particular Eq. (5.47).

$$\Omega = \frac{\tau}{S} = \frac{mB}{S}. \tag{4.177}$$

Generally, at this stage we would need to find the eigenstates a_n and eigenenergies E_n of our Hamiltonian; however, with our (smart:-) choice of the basis, the Hamiltonian matrix is already diagonal:

$$\mathrm{H} = \frac{\hbar\Omega}{2}\sigma_z = \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.178}$$

meaning that $\uparrow$ and $\downarrow$ are the eigenstates of the system, with eigenenergies, respectively,

$$E_\uparrow = +\frac{\hbar\Omega}{2} \quad \text{and} \quad E_\downarrow = -\frac{\hbar\Omega}{2}; \tag{4.179}$$

their difference

$$\Delta E \equiv E_\uparrow - E_\downarrow = \hbar\Omega = \frac{\hbar eB}{m_e} \tag{4.180}$$

corresponds to the classical energy $2mB$ of flipping the magnetic dipole, provided of course that we accept Eq. (118) for m, which does *not* follow from any classical model.

Now let the spin's evolution start from an arbitrary initial state $\alpha(t_0)$ which may be completely characterized by two complex coefficients $\alpha_{\uparrow\downarrow}(t_0)$ which are defined by the expansion over the basis states:

$$\left|\alpha(t_0)\right\rangle = \alpha_\uparrow(t_0)\left|\uparrow\right\rangle + \alpha_\downarrow(t_0)\left|\downarrow\right\rangle. \tag{4.181}$$

Then Eq. (161) immediately yields the time evolution equations

$$\alpha_\uparrow(t) = \alpha_\uparrow(t_0)\exp\left\{-\frac{i}{2}\Omega(t-t_0)\right\}, \quad \alpha_\downarrow(t) = \alpha_\downarrow(t_0)\exp\left\{+\frac{i}{2}\Omega(t-t_0)\right\}, \tag{4.182}$$

which allow one to predict dynamics of any observable, for example,

$$\begin{aligned} \left\langle S_z\right\rangle &= \left\langle\alpha\left|\hat{S}_z\right|\alpha\right\rangle = \left(\left\langle\uparrow\left|\alpha_\uparrow^*\right. + \left\langle\downarrow\right|\alpha_\downarrow^*\right)\hat{S}_z\left(\alpha_\uparrow\left|\uparrow\right\rangle + \alpha_\downarrow\left|\downarrow\right\rangle\right) \\ &= \alpha_\uparrow\alpha_\uparrow^*\left\langle\uparrow\left|\hat{S}_z\right|\uparrow\right\rangle + \alpha_\downarrow\alpha_\downarrow^*\left\langle\downarrow\left|\hat{S}_z\right|\downarrow\right\rangle + \alpha_\uparrow\alpha_\downarrow^*\left\langle\downarrow\left|\hat{S}_z\right|\uparrow\right\rangle + \alpha_\downarrow\alpha_\uparrow^*\left\langle\uparrow\left|\hat{S}_z\right|\downarrow\right\rangle. \end{aligned} \tag{4.183}$$

The bra-ket in this expression may be readily calculated in z-basis in which, according to Eq. (118), operator $\hat{S}_z$ is presented by matrix

$$\mathrm{S}_z = \frac{\hbar}{2}\sigma_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.184}$$

and states $\uparrow$ and $\downarrow$, by rows or columns (1, 0) and (0, 1), respectively. A simple calculation[23] yields

[23] The calculation becomes almost trivial if we use for the Pauli matrix operators the following outer-product presentations which follow from the general Eq. (60):

$$\langle S_z \rangle = \frac{\hbar}{2}\left[\alpha_\uparrow \alpha_\uparrow^* - \alpha_\downarrow \alpha_\downarrow^*\right]. \tag{4.185}$$

This result might be of course obtained simpler using probabilities $P_j = \alpha_j \alpha_j^*$:

$$\langle S_z \rangle = P_\uparrow\left(+\frac{\hbar}{2}\right) + P_\downarrow\left(-\frac{\hbar}{2}\right), \tag{4.186}$$

with just as it has been done in Eq. (130). The longer, formal way used in Eq. (183) has, however, an advantage of being applicable, without any change (besides the calculation of concrete matrix elements), for finding other observables. For example,

$$\begin{aligned} \langle S_x \rangle &= \langle \alpha | \hat{S}_x | \alpha \rangle = \left(\langle \uparrow | \alpha_\uparrow^* + \langle \downarrow | \alpha_\downarrow^* \right) \hat{S}_x \left(\alpha_\uparrow | \uparrow \rangle + \alpha_\downarrow | \downarrow \rangle \right) \\ &= \alpha_\uparrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_x | \uparrow \rangle + \alpha_\downarrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_x | \downarrow \rangle + \alpha_\uparrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_x | \uparrow \rangle + \alpha_\downarrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_x | \downarrow \rangle = \frac{\hbar}{2}\left(\alpha_\uparrow \alpha_\downarrow^* + \alpha_\downarrow \alpha_\uparrow^* \right). \end{aligned} \tag{4.187}$$

Similarly, we can calculate $\langle S_y \rangle$, as well as fluctuations of the spin components – see, e.g., Problem 8.

Now let us consider the particular case when the initial state of electron's spin at $t_0 = 0$ was $\uparrow$, i.e. $\alpha_\uparrow(0) = 1$, $\alpha_\downarrow(0) = 0$. According to Eq. (182), in this case the only effect of the time evolution is the increasing phase multiplier of $\alpha_\uparrow(t)$, with $\alpha_\downarrow(t)$ staying equal to zero. As a result, Eqs. (185) and (187) yield[24]

$$\langle S_z \rangle = \frac{\hbar}{2}, \quad \langle S_x \rangle = 0. \tag{4.188}$$

This result seems very natural, and it is tempting to interpret it as the persistent orientation of the spin along the magnetic field, corresponding to the minimum of the interaction energy (174). However, the reader should remember about our result (130) for such "fully spin-polarized" state: it has very large (indeed, the largest possible) fluctuations of S_x and S_y, in compliance with the uncertainty relation (156).

Now, what if electron's spin was initially in state $\rightarrow$, i.e. had the maximum value of component S_x (in classics, we would say "oriented in direction x")? According to the first of Eqs. (120), we can describe this state by coefficients[25]

$$\alpha_\uparrow(0) = \frac{1}{\sqrt{2}}, \quad \alpha_\downarrow(0) = \frac{1}{\sqrt{2}}, \tag{4.189}$$

so that Eqs. (182) yield

$$\hat{\sigma}_x = |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|, \quad \hat{\sigma}_y = -i|\uparrow\rangle\langle\downarrow| + i|\downarrow\rangle\langle\uparrow|, \quad \hat{\sigma}_z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|.$$

[24] Once again, sign $\langle \ldots \rangle$ means statistical averaging rather than that over time, so that a variable so averaged still can depend on time – see, e.g., Eq. (191).

[25] As you should remember from the (for most of you, sad:-) experience with final exam's Problem F.5, in the general case these coefficients may have an arbitrary (but common!) phase multiplier. However, in our current case, when the state does not have any other component, the multiplier does not have any observable consequences.

$$\alpha_{\uparrow}(t) = \frac{1}{\sqrt{2}}\exp\left\{-\frac{i}{2}\Omega t\right\}, \quad \alpha_{\downarrow}(t) = \frac{1}{\sqrt{2}}\exp\left\{+\frac{i}{2}\Omega t\right\}, \tag{4.190}$$

and the average spin components change in time as

$$\left\langle S_z \right\rangle = 0, \quad \left\langle S_x \right\rangle = \frac{\hbar}{2}\cos\Omega t, \quad \left\langle S_y \right\rangle = \frac{\hbar}{2}\sin\Omega t. \tag{4.191}$$

These formulas may be interpreted as the Cartesian components of the (average!) spin vector of length ($\hbar$/2), confined in plane [*x,y*] and performing the torque-induced precession with frequency (180) about axis *z*.

Let us now return to the general discussion of quantum dynamics to outline its alternative, *Heisenberg picture*. For that, let us recall that according to Eq. (126), in quantum mechanics the expectation value of any observable *A* is a bra-ket $\left\langle \alpha \middle| \hat{A} \middle| \alpha \right\rangle$. Below we will see that other quantities (say, the rates of quantum transition between pairs of states, say α and β) may also be measured in experiment; the most general form for all such measurable quantities is the bra-ket

$$\left\langle \alpha \middle| \hat{A} \middle| \beta \right\rangle. \tag{4.192}$$

As has been discussed above, in the Schrödinger picture the bra- and ket-vectors of the states are time-dependent, while operators stay constant (if the corresponding variables do not explicitly depend on time), so that Eq. (195) may be presented as

$$\left\langle \alpha(t) \middle| \hat{A}_{\mathrm{S}} \middle| \beta(t) \right\rangle, \tag{4.193}$$

where index "S" emphasizes the Schrödinger picture. Now apply to this expression the key Eq. (157):

$$\left\langle \alpha \middle| \hat{A} \middle| \beta \right\rangle = \left\langle \alpha(t_0) \middle| \hat{\boldsymbol{u}}^{\dagger}(t,t_0)\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}}(t,t_0) \middle| \beta(t_0) \right\rangle. \tag{4.194}$$

The right-hand part implies that if we use bra- and ket-vectors of the initial-time states, together with a time-dependent *Heisenberg operator*[26]

$$\hat{A}_{\mathrm{H}}(t,t_0) \equiv \hat{\boldsymbol{u}}^{\dagger}(t,t_0)\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}}(t,t_0) = \hat{\boldsymbol{u}}^{\dagger}(t,t_0)\hat{A}_{\mathrm{H}}(t_0,t_0)\hat{\boldsymbol{u}}(t,t_0), \tag{4.195}$$

all experimentally measurable results will remain the same:

$$\left\langle \alpha \middle| \hat{A} \middle| \beta \right\rangle = \left\langle \alpha(t_0) \middle| \hat{A}_{\mathrm{H}}(t,t_0) \middle| \beta(t_0) \right\rangle. \tag{4.196}$$

Let us see how it works for the same simple (but important!) problem of the electron spin precession in a *z*-oriented magnetic field. In the same *z*-basis as we used before, Eq. (158) for the time-evolution operator reads

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} \boldsymbol{u}_{11} & \boldsymbol{u}_{12} \\ \boldsymbol{u}_{21} & \boldsymbol{u}_{22} \end{pmatrix} = \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \boldsymbol{u}_{11} & \boldsymbol{u}_{12} \\ \boldsymbol{u}_{21} & \boldsymbol{u}_{22} \end{pmatrix} = \frac{\hbar\Omega}{2}\begin{pmatrix} \boldsymbol{u}_{11} & \boldsymbol{u}_{12} \\ -\boldsymbol{u}_{21} & -\boldsymbol{u}_{22} \end{pmatrix}. \tag{4.197}$$

[26] The reader has probably noticed the close analogy between Eqs. (96) and (194). However, it is important to remember that in the former equation describes a unitary transformation of the same operator due to a basis change, while in the latter case we speak about time evolution of an operator as such.

We see that in this simple case the equations for different matrix elements of evolution operator matrix are decoupled, and hence readily solvable (using the initial condition (160)):[27]

$$\mathrm{U}(t,0)=\begin{pmatrix} e^{-i\Omega t/2} & 0 \\ 0 & e^{+i\Omega t/2} \end{pmatrix}=\mathrm{I}\cos\frac{\Omega t}{2}-i\sigma_z\sin\frac{\Omega t}{2}. \tag{4.198}$$

Now we can use Eq. (198) to find the Heisenberg-picture operators of spin components. Dropping index "H" for brevity (time dependence clearly marks Heisenberg operators anyway), we get

$$\begin{aligned} \mathrm{S}_x(t,0)&=\mathrm{U}^\dagger(t,0)\mathrm{S}_x(0,0)\mathrm{U}(t,0)=\frac{\hbar}{2}\mathrm{U}^\dagger(t,0)\sigma_x\mathrm{U}(t,0) \\ &=\frac{\hbar}{2}\begin{pmatrix} e^{+i\Omega t/2} & 0 \\ 0 & e^{-i\Omega t/2} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} e^{-i\Omega t/2} & 0 \\ 0 & e^{+i\Omega t/2} \end{pmatrix} \\ &=\frac{\hbar}{2}\begin{pmatrix} 0 & e^{+i\Omega t} \\ e^{-i\Omega t} & 0 \end{pmatrix}=\frac{\hbar}{2}\left[\sigma_x\cos\Omega t-\sigma_y\sin\Omega t\right] \end{aligned} \tag{4.199}$$

An absolutely similar calculation for the z-component yields

$$\mathrm{S}_z(t)=\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}=\frac{\hbar}{2}\sigma_z. \tag{4.200}$$

An advantage of these formulas is that they describe the system evolution for an arbitrary initial state, and make the analysis of the initial state effects very simple. For example, if the initial state was spin-up (↑), then Eqs. (199) and (200), plugged into Eqs. (196) for the spin component expectation values, immediately yield

$$\langle S_x\rangle=\left\langle\uparrow\middle|\hat{S}_x(t)\middle|\uparrow\right\rangle=\begin{pmatrix}1 & 0\end{pmatrix}\mathrm{S}_x(t)\begin{pmatrix}1\\0\end{pmatrix}=0,\quad \langle S_z\rangle=\left\langle\uparrow\middle|\hat{S}_z(t)\middle|\uparrow\right\rangle=\begin{pmatrix}1 & 0\end{pmatrix}\mathrm{S}_z(t)\begin{pmatrix}1\\0\end{pmatrix}=\frac{\hbar}{2}. \tag{4.201}$$

On the other hand, if the initial state was spin-right (→), then with a help from Eq. (120) we get

$$\langle S_x\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\mathrm{S}_x(t)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}=\frac{\hbar}{2}\cos\Omega t,\quad \langle S_z\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\mathrm{S}_z(t)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}=0, \tag{4.202}$$

i.e. the same result in Eq. (191).

Moreover, sometimes just the very expression for operator evolution tells it all. For example, Eq. (200) clearly shows that regardless of the initial state of the system, the component of spin directed along the magnetic field does not change in time. Also, notice that the first part of expression (199) for operator S_x parallels its average (see the first of Eqs. (191) calculated for the initial state →), but its also has the second part, proportional to $\sin\Omega t$, which reflects the possible contribution of the initial y-component of the spin (if it was present in the initial state). In the next section, we will see that the Heisenberg operator dynamics is especially transparent for the orbital motion.

[27] We could of course use this result and Eq. (157) to obtain Eqs. (191) within the Schrödinger picture. In our simple case, the use of Eqs. (161) and (185) for this purpose was more straightforward, but in the general case starting from the calculation of the time-evolution matrix may be the only way to go.

It is useful to combine Eqs. (158) and (195) to form an explicit differential equation of the Heisenberg operator evolution. For that, let us differentiate its definition (195) over time, assuming that operator $\hat{A}_{\mathrm{S}}$ does not depend explicitly on time:

$$\frac{\partial \hat{A}_{\mathrm{H}}}{\partial t} = \frac{\partial \hat{\boldsymbol{u}}^{\dagger}}{\partial t}\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}^{\dagger}\hat{A}_{\mathrm{S}}\frac{\partial \hat{\boldsymbol{u}}}{\partial t}. \tag{4.203}$$

Plugging in time derivatives of the time evolution operator from Eq. (158) and Hermitian conjugate, we get

$$i\hbar\frac{\partial \hat{A}_{\mathrm{H}}}{\partial t} = -\hat{\boldsymbol{u}}^{\dagger}\hat{H}\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}^{\dagger}\hat{A}_{\mathrm{S}}\hat{H}\hat{\boldsymbol{u}}. \tag{4.204}$$

If the Hamiltonian does not depend on time (and even if it does, but condition (166) is satisfied), then, according to Eqs. (163) or (169), the operator commutes with the time evolution operator and its Hermitian conjugate,[28] and may be swapped with any of them. Hence, we may rewrite Eq. (204) as

$$i\hbar\frac{\partial \hat{A}_{\mathrm{H}}}{\partial t} = -\hat{H}\hat{\boldsymbol{u}}^{\dagger}\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}^{\dagger}\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}}\hat{H} = \left[\hat{\boldsymbol{u}}^{\dagger}\hat{A}_{\mathrm{S}}\hat{\boldsymbol{u}}, \hat{H}\right] = \left[\hat{A}_{\mathrm{H}}, \hat{H}\right] \tag{4.205}$$

This is the so-called *Heisenberg equation of motion* (which was actually derived by P. A. M. Dirac).[29]

Let us again see how does this equation look for the same problem of spin precession in a z-oriented magnetic field. In the z-basis, Eq. (205) for the vector operator of spin reads

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{21} & \mathbf{S}_{22}\end{pmatrix} = \frac{\hbar\Omega}{2}\left[\begin{pmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{21} & \mathbf{S}_{22}\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right] = \hbar\Omega\begin{pmatrix}0 & -\mathbf{S}_{12}\\ \mathbf{S}_{21} & 0\end{pmatrix}, \tag{4.206}$$

where we have used Eq. (178) for the Hamiltonian operator's matrix. Once again, the equations for different matrix elements are decoupled, so that using Eqs. (107), (118) for the initial (i.e. Schrödinger-picture) operator,

$$\mathbf{S}(0) = \frac{\hbar}{2}\left[\mathbf{n}_x\sigma_x + \mathbf{n}_y\sigma_y + \mathbf{n}_z\sigma_z\right] = \frac{\hbar}{2}\begin{pmatrix}\mathbf{n}_z & \mathbf{n}_x - i\mathbf{n}_y\\ \mathbf{n}_x + i\mathbf{n}_y & -\mathbf{n}_z\end{pmatrix}, \tag{4.207}$$

we can immediately write the solution in either of two forms:

$$\begin{aligned}\mathbf{S}(t) &= \frac{\hbar}{2}\begin{pmatrix}\mathbf{n}_z & (\mathbf{n}_x - i\mathbf{n}_y)e^{i\Omega t}\\ (\mathbf{n}_x + i\mathbf{n}_y)e^{-i\Omega t} & -\mathbf{n}_z\end{pmatrix}\\ &= \frac{\hbar}{2}\left[\mathbf{n}_x\begin{pmatrix}0 & e^{i\Omega t}\\ e^{-i\Omega t} & 0\end{pmatrix} + \mathbf{n}_y\begin{pmatrix}0 & -ie^{i\Omega t}\\ ie^{-i\Omega t} & 0\end{pmatrix} + \mathbf{n}_z\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right].\end{aligned} \tag{4.208}$$

[28] Due to the same reason, $\hat{H}_{\mathrm{H}} = \hat{\boldsymbol{u}}^{\dagger}\hat{H}_{\mathrm{S}}\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}^{\dagger}\hat{\boldsymbol{u}}\hat{H}_{\mathrm{S}} = \hat{H}_{\mathrm{S}}$, so that the index may be dropped in the time evolution equation (205).

[29] Dr. Dirac was so generous that he himself gave another person's name to this key result (the exact opposite of the behavior frequently met in science nowadays:-).

The simplicity of the first of these expressions is spectacular. (Remember, it covers *any* initial conditions, and *all* 3 spatial components of spin!) On the other hand, for some purposes the second form is more convenient; in particular, in its first and the last terms, we can readily recognize out earlier results (199) and (200).

One of advantages is that the Heisenberg picture is that it provides a closer link between the classical and quantum mechanics. Indeed, equations of classical mechanics may be used to derive the following equation[30] of time evolution of an arbitrary function $A(q_j, p_j)$ of generalized coordinates and momenta of the system:

$$\frac{dA}{dt} = -\{A, H\}, \tag{4.209}$$

where H is the Hamiltonian function of the system, and $\{\ldots\}$ is the Poisson bracket defined as

$$\{A, H\} \equiv \sum_j \left(\frac{\partial A}{\partial p_j} \frac{\partial H}{\partial q_j} - \frac{\partial A}{\partial q_j} \frac{\partial H}{\partial p_j} \right). \tag{4.210}$$

Comparing Eqs. (205) and (210), we see that the correspondence between the classical and quantum mechanics (in the Heisenberg picture) is provided by the following symbolic relation

$$\{A, H\} \leftrightarrow \frac{i}{\hbar}\left[\hat{A}, \hat{H}\right]. \tag{4.211}$$

This relation may be used, in particular, for finding appropriate operators for system's observables, if their form is not immediately evident from the correspondence principle.

We will develop this argumentation further in the next section in which we return to wave mechanics.

4.7. Revisiting wave mechanics

Now let us explore whether the bra-ket formalism may be used to describe the "orbital" motion of particle with zero spin (or disregarding its spin properties) and in particular to derive the principles of wave mechanics discussed in Chapter 1. For that, first of all, we have to either spell out or even modify some of our formulas for the case of observables with continuous spectrum of eigenvalues. (One example we already know well are momentum and kinetic energy of a free particle.) In that case, all the above expressions for states, their bra- and ket-vectors, and eigenvalues, should be stripped of discrete indices, like index j in the key equation (70) which determines eigenstates and eigenvalues of observable A. For that, Eq. (70) may be rewritten in the form

$$\hat{A}\left|a_A\right\rangle = A\left|a_A\right\rangle. \tag{4.212}$$

More essentially, sums over such continuous eigenstate sets should be replaced by integrals. For example, for a full and orthonormal set of eigenstates (215), the "completeness relation" expressed by Eq. (44) may be changed into

[30] See, e.g., CM Eq. (9.17).

$$\int dA \left| a_A \right\rangle \left\langle a_A \right| = \hat{I}\,, \tag{4.213}$$

where the integral should be taken over the whole interval of possible values of the observable. Applying this relation to the ket-vector of an arbitrary state α (generally, *not* an eigenstate of operator $\hat{A}$), we get

$$\left|\alpha\right\rangle = \hat{I}\left|\alpha\right\rangle = \int dA \left| a_A \right\rangle \left\langle a_A \left| \alpha \right.\right\rangle = \int dA \left\langle a_A \left| \alpha \right.\right\rangle \left| a_A \right\rangle . \tag{4.214}$$

For the particular case when $|\alpha\rangle = |a_{A'}\rangle$, this relation requires

$$\left\langle a_A \left| a_{A'} \right.\right\rangle = \delta(A - A'); \tag{4.215}$$

this formula replaces the orthonormality condition (38). We see that the transfer from summation to integration requires also the change of the Kronecker symbols to Dirac's delta-functions. Notice that in the contrast to the discrete spectrum case, the dimensionality of the bra-kets so normalized is different from 1; it is rather the reciprocal dimensionality of the corresponding observable.

Integral (214) replaces sum (37) for the presentation of an arbitrary ket-vector as an expansion over eigenstates of an operator, so that bra-ket $\langle a_A|\alpha\rangle$ replaces coefficient α_j which in particular determines the state probability – see Eq. (123). However, in the continuous spectrum case the probability to find the system exactly in a particular state is infinitesimal. Instead we should speak about the probability density $w(A) \propto |\langle a_A|\alpha\rangle|^2$ to find the observable close to certain value A. The coefficient in that relation may be found by making the similar change from summation to integration (without any additional coefficients) in the normalization condition (124):

$$\int dA \left\langle \alpha \left| a_A \right.\right\rangle \left\langle a_A \left| \alpha \right.\right\rangle = 1. \tag{4.216}$$

Since the total probability should be also equal to $\int w(A)dA$, this means:

$$w(A) = \left\langle \alpha \left| a_A \right.\right\rangle \left\langle a_A \left| \alpha \right.\right\rangle = \left| \left\langle \alpha \left| a_A \right.\right\rangle \right|^2 . \tag{4.217}$$

Now let us see what happens to the expectation values of observables, i.e. ensemble averages of operators. If we speak about the same observable A whose eigenstates we are using as the basis (or any compatible observable), everything is simple. Inserting Eq. (217) into the general statistical relation

$$\left\langle A \right\rangle = \int w(A) A dA \tag{4.218}$$

(which is just the continuous version of Eq. (124)), we get

$$\left\langle A \right\rangle = \int \left\langle \alpha \left| a_A \right.\right\rangle A \left\langle a_A \left| \alpha \right.\right\rangle dA. \tag{4.119}$$

Presenting this expression as a double integral,

$$\left\langle A \right\rangle = \int dA \int dA' \left\langle \alpha \left| a_A \right.\right\rangle A \delta(A - A') \left\langle a_{A'} \left| \alpha \right.\right\rangle, \tag{4.220}$$

and using the continuous version of Eq. (100),

$$\left\langle a_A \left| \hat{A} \right| a_{A'} \right\rangle = A\delta(A - A')\,, \tag{4.221}$$

we may write

$$\langle A\rangle = \int dA \int dA' \langle \alpha | a_A \rangle \langle a_A | \hat{A} | a_{A'} \rangle \langle a_{A'} | \alpha \rangle = \langle \alpha | \hat{A} | \alpha \rangle, \tag{4.222}$$

so that the general expression (126) is valid in the continuous-spectrum case without any changes.

The situation is more complicated for arbitrary operators which generally do not commute with the base-creating operator $\hat{A}$, because the matrix of such an operators in the *a*-basis may or may not be diagonal, so that Eq. (131) is generally not valid. We will consider (and overcome:-) this difficulty below, but otherwise we are ready for the discussion of wave mechanics. (For the notation simplicity I will discuss its 1D version; the generalization to 2D or 3D is trivial.)

Let us postulate the (intuitively almost evident) existence of quantum states, with ket-vectors called $|x\rangle$, corresponding to a fixed value, *x*, of the particle coordinate. Writing an evident identity

$$x|x\rangle = x|x\rangle, \tag{4.223}$$

and comparing this relation with Eq. (212), we see that they do not contradict each other if we assume that the eigenstates *x* may be considered as eigenstates the coordinate operator $\hat{x}$, if this operator's action is just the multiplication of the ket- (or bra-) vectors by *x*. This looks like a proof, but is actually a separate, independent postulate, no matter how plausible.

The bra-ket $\langle a_A | \alpha(t) \rangle$, playing the key role of the coefficient of expansion of an arbitrary state $\alpha(t)$ over the eigenstate basis (please have one more look at Eq. (214)!), now becomes $\langle x | \alpha(t) \rangle$, and this is exactly what is called the Schrödinger's wavefunction of that state:[31]

$$\langle x | \alpha(t) \rangle \equiv \Psi_\alpha(x,t). \tag{4.224}$$

This formula provides the connection between the bra-ket formalism and wave mechanics. Let us rewrite our most important formulas in this notation. Expression (217) for the probability density immediately acquires the familiar form (1.22)

$$w_\alpha(x) = \Psi_\alpha^*(x,t)\Psi_\alpha(x,t), \tag{4.225}$$

but the situation with the Schrödinger equation is a bit more complex. Indeed, let us calculate the (partial) time derivative of the wave function (multiplied by the usual coefficient $i\hbar$):

$$i\hbar \frac{\partial \Psi_\alpha}{\partial t} = i\hbar \frac{\partial}{\partial t} \langle x | \alpha(t) \rangle. \tag{4.226}$$

Since the coordinate operator $\hat{x}$ does not depend on time explicitly, its eigenstates are stationary, and we can swap the time derivative and the time-independent ket-vector and hence $\langle x|$. Making use of the

[31] I do not quite like expressions like $\langle x | \psi \rangle$, used in some textbooks. Of course, one is free to replace α with any other letter (ψ including) to denote a quantum state, but then it is better not to use the same letter to denote the wavefunction, to avoid confusion. Again, state α is something existing in Mother Nature (not on paper or in computer memory), while bra-ket $\langle x | \alpha \rangle$, and hence the wavefunction's value (in each spatial-temporal point) is just a number which may be recorded, stored, etc.

Schrödinger-picture equations (157) and (158), and then inserting the identity operator in the form (213), written for the coordinate eigenstates,

$$\int dx' \left|x'\right\rangle\left\langle x'\right| = \hat{I}\,, \tag{4.227}$$

we may continue as

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi_\alpha(x,t) &= \left\langle x\left|i\hbar\frac{\partial}{\partial t}\right|\alpha(t)\right\rangle = \left\langle x\left|i\hbar\frac{\partial}{\partial t}\hat{\boldsymbol{u}}(t,t_0)\right|\alpha(t_0)\right\rangle = \left\langle x\left|\hat{H}\hat{\boldsymbol{u}}(t,t_0)\right|\alpha(t_0)\right\rangle = \left\langle x\left|\hat{H}\right|\alpha(t)\right\rangle \\ &= \int dx'\left\langle x\left|\hat{H}\right|x'\right\rangle\left\langle x'\right|\alpha(t)\rangle = \int dx'\left\langle x\left|\hat{H}\right|x'\right\rangle\Psi_\alpha(x',t). \end{aligned} \tag{4.228}$$

This is where we have to stop for a general Hamiltonian operator, because if it does not commute with the coordinate operator[32] , its matrix in the *x*-basis is not diagonal. However, there exists a broad set of *local operators* (including all Hamiltonians which have been or will be studied in this course) whose arguments include just one coordinate, either *x* or *x'*, but not both of them. (Derivatives like either $\partial/\partial x$ or $\partial/\partial x$' fall into this class, but finite differences like $x - x'$ do not.) For example, let operator $\hat{A}$ depend only on *x* and/or $\partial/\partial x$; then in an expression like that in the last form of Eq. (228) we can move bra-vector $\left|x'\right\rangle$ to the left:

$$\left\langle x\left|\hat{A}\right|x'\right\rangle\Psi(x',t) = \left\langle x\left|x'\right\rangle\hat{A}\Psi(x',t) = \delta(x-x')\hat{A}\Psi(x',t)\right.. \tag{4.229}$$

For a local Hamiltonian, Eq. (228) immediately gives the Schrödinger equation in its wave-mechanics form (1.24):

$$i\hbar\frac{\partial\Psi_\alpha}{\partial t} = \hat{H}\Psi_\alpha(t). \tag{4.230}$$

As we saw in Sec. 1.4, if Hamiltonian is time-independent, the variable-separation procedure allows one to express the general solution to this equation in terms of its eigenstates and eigenvalues defined by Eq. (1.44b):

$$\hat{H}\psi_n = E_n\psi_n. \tag{4.231}$$

Please notice that the continuous character of coordinate *x* does *not* mean the continuity of the energy spectrum E_n. (Indeed, in the beginning of the course we have solved quite a few problems where E_n were discrete – starting from the simplest quantum well – see Eq. (1.62)). There is no contradiction at all here, because typical Hamiltonians of wave mechanics, for example the one given by the 1D version of Eq. (1.47),

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + U(x)\,, \tag{4.232}$$

do not commute with the coordinate operator $\hat{x}$, and hence do not share its continuous spectrum.

The situation with the eigenstates and eigenvalues of an arbitrary operator is similar. Please recall that in our earlier discussion of wave mechanics, I have avoided using these notions (with the only

[32] For example, the Hamiltonian which we studied so much in Chapter 2, $\hat{H} = \hat{p}^2/2m + U(x)$, has the momentum-dependent part which does not commute with $\hat{x}$ - see below.

exception of the Hamiltonian, for which the eigenstate/eigenvalue problem (230) has jumped at us all by itself from the general solution to the Schrödinger equation). For an arbitrary operator, the eigenvalue/eigenstate problem was first introduced in the bra-ket formulation – see Eq. (70). Let us see whether it may be rewritten in the wave mechanics language. Multiplying both parts of Eq. (70) by the ket-vector $\langle x|$, and inserting into the left-hand part the identity operator in form (227), we get

$$\int dx' \langle x|\hat{A}|x'\rangle \langle x'|a_j\rangle = A_j \langle x|a_j\rangle, \tag{4.233}$$

i.e., using the wavefunction definition (224),

$$\int dx' \langle x|\hat{A}|x'\rangle \Psi_j(x',t) = A_j \Psi_j(x,t). \tag{4.234}$$

If the operator A is space-local, i.e. satisfies a relation similar to Eq. (238), then this result is reduced to

$$\hat{A}\Psi_j(x,t) = A_j \Psi_j(x,t), \tag{4.235}$$

even if operator $\hat{A}$ does not commute with operator $\hat{x}$. (Again, in that case, the spectrum of eigenstates A_j may be discrete.)[33]

The operator locality also simplifies the expression for its expectation value. Indeed, applying Eq. (225) to a single eigenstate, and integrating it over x, we get the probability of finding the system in this state:

$$W_j = \int \Psi_j^* \Psi_j dx\,. \tag{4.236}$$

From here and the general Eq. (125), for an arbitrary state α, whose wavefunction Ψ_α may be presented as a sum of Ψ_j,

$$\langle A\rangle_\alpha = \sum_j P_j A_j = \int dx \sum_j \Psi_j^* A_j \Psi_j\,. \tag{4.237}$$

If eigenfunctions Ψ_j form an orthogonal set (which is typically the case), Eq. (237) is equivalent to

$$\langle A\rangle_\alpha = \int dx \sum_{j,j'} \Psi_{j'}^* A_j \Psi_j = \int dx \Psi_\alpha^* \sum_j A_j \Psi_j\,. \tag{4.238}$$

For a local operator, we can now use Eq. (235) to write

$$\langle A\rangle_\alpha = \int dx \Psi_\alpha^* \sum_j \hat{A}\Psi_j = \int dx \Psi_\alpha^* \hat{A} \sum_j \Psi_j = \int dx \Psi_\alpha^* \hat{A}\Psi_\alpha\,. \tag{4.239}$$

Thus we have recovered the (1D version of) Eq. (1.23) which we had to postulate in Chapter 1.

[33] It is interesting that in some systems of quantum mechanics postulates, the Schrödinger equation (230) itself is considered as a sort of eigenstate/eigenvalue problem for operator $i\hbar\partial/\partial t$. Notice that such construct is very close to that of the momentum operator - $i\hbar\partial/\partial x$, and similar arguments may be given for both expressions, starting from the invariance of the quantum state of a free particle with respect to translations in time and space, respectively. (See, e.g., pp. 44-46 and 70-71 in Sakurai.)

Now we are in position to start filling the blanks left in the wave-mechanics part of our course. First of all, we may now derive the Heisenberg's uncertainty relation (1.34) from the general relation (140). For that, it is sufficient to calculate the commutator of operators of the coordinate and the corresponding component of momentum:

$$[\hat{x}, \hat{p}_x]\psi = \left[x, -i\hbar\frac{\partial}{\partial x}\right]\psi = -i\hbar\left[x\frac{\partial}{\partial x}\psi - \frac{\partial}{\partial x}(x\psi)\right] = -i\hbar\left[x\frac{\partial}{\partial x}\psi - \left(x\frac{\partial}{\partial x}\psi + \psi\right)\right] = i\hbar\psi\,. \quad (4.240)$$

Since this calculation is valid for any (continuous) wavefunction, we may rewrite it as an operator relation: [34]

$$[\hat{x}, \hat{p}_x] = i\hbar, \quad (4.241)$$

so that Eq. (140) immediately gives formula (1.34):

$$\delta x \delta p_x \geq \frac{\hbar}{2}. \quad (4.242)$$

The commutation relation (241) is convenient for many other calculations, including a simple derivation of the famous *Ehrenfest theorem*. To start, let us calculate the following commutator:

$$[\hat{x}, \hat{p}_x^2] = x\hat{p}_x\hat{p}_x - \hat{p}_x\hat{p}_x x. \quad (4.243)$$

Rewriting Eq. (241) as

$$x\hat{p}_x = \hat{p}_x x + i\hbar, \quad (4.244)$$

We can use it twice in the left term of the RHP of Eq. (243) to sequentially move momentum operators to the left:

$$x\hat{p}_x\hat{p}_x = (\hat{p}_x x + i\hbar)\hat{p}_x = \hat{p}_x x\hat{p}_x + i\hbar\hat{p}_x = \hat{p}_x(\hat{p}_x x + i\hbar) + i\hbar\hat{p}_x = \hat{p}_x\hat{p}_x x + 2i\hbar\hat{p}_x. \quad (4.244)$$

The first term in the result cancels with the second term in Eq. (243), so that the commutator is rather simple:

$$[\hat{x}, \hat{p}_x^2] = 2i\hbar\hat{p}_x. \quad (4.245)$$

Let us use this result to calculate the Heisenberg-picture equation of motion for operator $\hat{x}$ from the general Eq. (205) for the orbital motion, when the Hamiltonian has the form of Eq. (232):[35]

$$\frac{d\hat{x}}{dt} = \frac{1}{i\hbar}[\hat{x}, \hat{H}] = \frac{1}{i\hbar}\left[x, \frac{\hat{p}_x^2}{2m} + U(x)\right]. \quad (4.246)$$

The potential energy function $U(x)$ may be always presented by its Taylor expansion:

$$U(x) = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{\partial^k U}{\partial x^k}x^k, \quad (4.247)$$

[34] By the same argumentation, operators corresponding to the different spatial components do commute, so that $[\hat{r}_j, \hat{p}_{j'}] = i\hbar\delta_{jj'}$ (j, j' = 1,2,3).

[35] Since this Hamiltonian is time-independent, considering the initial moment of time to is fixed, we may replace the partial derivative over time t with the full one.

where all the derivatives are evaluated at point $x = 0$. Each term of this sum evidently commutes with x, and hence the whole sum does. Thus, the right-hand part of Eq. (246) is proportional to commutator (245), and we get

$$\frac{d\hat{x}}{dt} = \frac{\hat{p}_x}{m}. \tag{4.248}$$

In that equality, we readily recognize the classical relation between the particle's momentum and is velocity.

Now it is evidently interesting what would a similar procedure give for the momentum's derivative:

$$\frac{d\hat{p}_x}{dt} = \frac{1}{i\hbar}\left[\hat{p}_x, \hat{H}\right] = \frac{1}{i\hbar}\left[\hat{p}_x, \frac{\hat{p}_x^2}{2m} + U(x)\right]. \tag{4.249}$$

The kinetic energy operator $\hat{p}_x^2/2m$ evidently commutes with operator $\hat{p}_x$ and may be dropped from the right-hand part. In order to calculate the remaining commutator, let us use expansion (247):

$$\left[\hat{p}_x, U(x)\right] = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{\partial^k U}{\partial x^k}\left[\hat{p}_x, x^k\right] = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{\partial^k U}{\partial x^k}\left(\hat{p}_x \underbrace{xx..x}_{k \text{ times}} - \underbrace{xx...x}_{k \text{ times}}\hat{p}_x\right). \tag{4.250}$$

Applying Eq. (244) k times to the last term in the parentheses, exactly as we did it in Eq. (244), we get

$$\left[\hat{p}_x, U(x)\right] = -\sum_{k=1}^{\infty} \frac{1}{k!}\frac{\partial^k U}{\partial x^k}(-ik\hbar x^{k-1}) = (-i\hbar)\sum_{k=1}^{\infty} \frac{1}{(k-1)!}\frac{\partial^k U}{\partial x^k}x^{k-1}. \tag{4.251}$$

But the last sum is just the Taylor expansion of the derivative $\partial U/\partial x$. Indeed,

$$\frac{\partial U}{\partial x} = \sum_{k'=0}^{\infty} \frac{1}{k'!}\frac{\partial^{k'}}{\partial x^{k'}}\left(\frac{\partial U}{\partial x}\right)x^{k'} = \sum_{k'=0}^{\infty} \frac{1}{k'!}\frac{\partial^{k'+1} U}{\partial x^{k'+1}}x^{k'} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!}\frac{\partial^k U}{\partial x^k}x^{k-1}, \tag{4.252}$$

(where at the last step we have replaced the notation of the summation index from k' to k - 1). As a result, Eq. (249) yields:[36]

$$\frac{d\hat{p}_x}{dt} = -\frac{\partial}{\partial x}U(x). \tag{4.253}$$

This equation again coincides with the classical equation of motion! (Discussing spin dynamics in Sec. 6, we have already seen that this is very typical of the Heisenberg picture.) However, it is important to remember that the equivalence is by form only and by no means show that quantum dynamics is equivalent to the classical one.

The degree of the similarity very much depends on the problem. In one extreme, let us consider the case when a particle's state, at any moment between t_0 and t, may be accurately represented by one, relatively narrow wave packet. Then, averaging Eqs. (248) and (253) over the initial state (as Eq. (196) prescribes),[37]

[36] The set of equations (248) and (253) constitute the Ehrenfest theorem.

[37] In the wave mechanics language, for a local operator that equation gives:

$$\frac{d\langle x\rangle}{dt} = \frac{\langle p_x\rangle}{m}, \quad \frac{d\langle p_x\rangle}{dt} = -\left\langle \frac{\partial U}{\partial x} \right\rangle, \tag{4.254}$$

we may interpret these equations as the classical equations of motion for the wave packet's center, in accordance with the correspondence principle. However, even in this case it is important to remember about the purely quantum mechanical effects of nonvanishing wave packet width and its spread in time, which were discussed in Sec. 2.1.

In the opposite extreme, Eqs. (254), though being valid, tell essentially *nothing* about the system dynamics. Maybe the most apparent example is the "leaking" quantum well which was discussed in Sec. 2.4 (see Fig. 2.18 and its discussion). Since both the potential $U(x)$ and the initial state are symmetric relative to point $x = 0$, the right-hand parts of both Eqs. (254) identically equal zero. Of course, the result (that average values of both momentum and coordinate stay equal zero at all times) is correct, but it tells us nothing about the rich dynamics of the system (the finite lifetime of the metastable state, formation of two wave packets, their waveform and propagation speed) and hence the important insight the solution may give for the quantum measurement interpretation.

Another spectacular example is the band theory (Sec. 2.5), with its allowed energy bands and forbidden gaps, of which Eq. (254) gives no clue. To summarize, the Ehrenfest theorem is important as an illustration (at to some extent the proof) of the correspondence principle, but its limitations should be well understood, and the importance must not be exaggerated.

Next, in Chapter 2 I promised you to discuss appropriate normalization of monochromatic waves (2.5),

$$\psi_p(x) = C_p \exp\left\{ i\frac{px}{\hbar} \right\}, \tag{4.253}$$

which are (the spatial parts of) the Schrödinger equation's solution for a free 1D particle. Since operators of momentum and kinetic energy commute and hence share a common set of eigenstates (253), these solutions may be understood as the eigenstates of either operator. For our current purposes, let us consider them as eigenstates p of momentum, so we can use Eq. (224) to write

$$\psi_p(x) = \langle x | p \rangle . \tag{4.254}$$

Now, identifying states a_A in the general Eq. (215) with p-states, we may write the normalization condition as

$$\langle p | p' \rangle = \delta(p - p'). \tag{4.255}$$

Inserting the identity operator in the form (227) (with the integration variable x' replaced by x) into the left-hand side of this equation, we can translate this normalization rule into the wavefunction language:

$$\int dx \langle p | x \rangle \langle x | p' \rangle \equiv \int dx \psi_p^*(x) \psi_{p'}(x) = \delta(p - p'). \tag{4.256}$$

Now using Eq. (253), this requirement becomes

$$\langle A \rangle = \int dx \Psi^*(x, t_0) \hat{A}_{\mathrm{H}}(t, t_0) \Psi(x, t_0) .$$

$$C_p C_{p'}^* \int_{-\infty}^{+\infty} \exp\left\{ i \frac{(p-p')x}{\hbar} \right\} dx \equiv \left| C_p \right|^2 2\pi\hbar\,\delta(p-p') = \delta(p-p'), \tag{4.257}$$

so that, finally, $C_p = \exp\{i\varphi\}/\sqrt{2\pi\hbar}$, and Eq. (253) becomes

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} \exp\left\{ i\left(\frac{px}{\hbar} + \varphi \right) \right\}, \tag{4.258}$$

where φ is an arbitrary (real) constant phase.[38]

Notice that with this "*delta-normalization*", integral

$$\int_{-\infty}^{+\infty} \psi_p(x) \psi_p^*(x) dx \tag{4.259}$$

diverges, and it seems that we do not need it for finite-width wave packets (Sec. 2.1), where we may return to the "regular" normalization, with integral (259) required to equal 1. However, for some purposes it makes sense to keep the coefficient in Eq. (258) even for wave packets, because of the following reason. Let us form a wave packet (2.6) based on wavefunctions (258) (taking $\varphi = 0$ for the notation brevity):

$$\psi(x) \equiv \frac{1}{\sqrt{2\pi\hbar}} \int \varphi(p) \exp\left\{ i \frac{px}{\hbar} \right\} dp\,. \tag{4.260}$$

This is just the equation of a 1D Fourier spatial transform, and its reciprocal is

$$\varphi(p) \equiv \frac{1}{\sqrt{2\pi\hbar}} \int \psi(x) \exp\left\{ -i \frac{px}{\hbar} \right\} dx\,. \tag{4.261}$$

These expressions are completely symmetrical, and present the same wave packet; this is why they are frequently called, respectively, the *x*- and *p-representations* of the system's state.

We already know that in the *x*-representation, i.e. in the usual wave mechanics, the coordinate operator $\hat{x}$ is presented by multiplication by *x*, and the momentum operator is proportional to a derivative over *x*:

$$\hat{x}\Big|_{\text{in } x} = x, \quad \hat{p}\Big|_{\text{in } x} = -i\hbar \frac{\partial}{\partial x}. \tag{4.262}$$

It is natural to guess that in the *p*-representation, the expressions for operators would be reciprocal:

$$\hat{x}\Big|_{\text{in } p} = i\hbar \frac{\partial}{\partial p}, \quad \hat{p}\Big|_{\text{in } p} = p, \tag{4.263}$$

with the difference in one sign only, due to the different signs of the Fourier exponents in Eqs. (260) and (261). The proof of Eqs. (263) is straightforward; for example, the differentiation of the both sides of Eq. (260) over *x* yields[39]

---

[38] Repeating the calculation for each Cartesian component of a plane monochromatic wave of arbitrary dimensionality *d*, we get $\psi_{\mathbf{p}} = (2\pi\hbar)^{-d/2} \exp\{i(\mathbf{p}\cdot\mathbf{r}/\hbar + \varphi)\}$.

$$\hat{p}\psi(x) \equiv -i\hbar\frac{\partial}{\partial x}\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int \varphi(p)\left(-i\hbar\frac{\partial}{\partial x}\exp\left\{i\frac{px}{\hbar}\right\}\right)dp$$
$$= \frac{1}{\sqrt{2\pi\hbar}}\int p\varphi(p)\exp\left\{i\frac{px}{\hbar}\right\}dp = \frac{1}{\sqrt{2\pi\hbar}}\int \left(\hat{p}\varphi(p)\right)\exp\left\{i\frac{px}{\hbar}\right\}dp. \quad (4.264)$$

However, it is interesting to have a different, more general look at the same coordinate-to-momentum duality. For that, notice that according to Eq. (84) we may consider the bra-ket $\langle x|p\rangle$ as an element (84) of the (infinite) matrix U_{xp} of the unitary transform from the x-basis to p-basis. Now let us derive the operator transform rule which would be a continuous version of Eq. (94). Say, we want to calculate the matrix element of some operator in p-representation:

$$\langle p|\hat{A}|p'\rangle. \quad (4.265)$$

Inserting two identity operators (227), and then using Eq. (258) and its complex conjugate, and Eq. (229) valid for local operators, we get

$$\langle p|\hat{A}|p'\rangle = \int dx\int dx'\langle p|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|p'\rangle = \int dx\int dx'\psi_p^*(x)\langle x|\hat{A}|x'\rangle\psi_{p'}(x')$$
$$= \frac{1}{2\pi\hbar}\int dx\int dx'\exp\left\{-i\frac{px}{\hbar}\right\}\delta(x-x')\hat{A}\exp\left\{i\frac{p'x'}{\hbar}\right\} = \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\hat{A}\exp\left\{i\frac{p'x}{\hbar}\right\}. \quad (4.266)$$

For example, for the momentum operator this equation yields:

$$\langle p|\hat{p}|p'\rangle = \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\left(-i\hbar\frac{\partial}{\partial x}\right)\exp\left\{i\frac{p'x}{\hbar}\right\} = \frac{p'}{2\pi\hbar}\int_{-\infty}^{+\infty}\exp\left\{i\frac{(p'-p)x}{\hbar}\right\}dx = p'\delta(p'-p). \quad (4.267)$$

Due to Eq. (229), this result is equivalent to the second of Eqs. (263).

A natural question arises: why isn't the momentum representation used much less frequently than the coordinate representation (i.e., wave mechanics)? The answer is purely practical: the orbital motion Hamiltonian (232) is usually not $x \leftrightarrow p$ symmetric, with the potential energy $U(x)$ typically a more complex function of x than the kinetic energy which is quadratic in momentum. Because of that, it is easier to keep operator x being just a wave function multiplier, as it is in the coordinate representation.

The most significant exceptions of this rule (besides the free particle case) are:

(i) the harmonic oscillator whose Hamiltonian is a quadratic in both $\hat{x}$ and $\hat{p}$ (see Sec. 9 below for more discussion), and

(ii) a particle in a periodic potential, in the presence of additional external force $F(t)$, which may result, for example, the Bloch oscillations, Zener tunneling etc.

Indeed, as we have discussed in Sec. 2.5, in the latter case the dispersion relation $E_n(q)$, typically rather involved, plays the role of the effective kinetic energy, while the effective potential energy $U_{ef} = -F(t)x$ in the field of the additional force is a simple function of x. This is why the discussions of the

---

39 In some systems of quantum mechanics axioms, this calculation is considered *the proof* of Eq. (262) for the momentum operator. Notice, however, that it is based on accepting Eq. (253) for the wavefunction, which in this case has to be postulated (e.g., on the basis of experimental facts discussed in Sec. 1.1) rather than derived as we did in Chapter 1.

listed and more complex issues of the band theory (such as quasiparticle scattering, mobility, excitation, etc., usually discussed in solid state physics courses), are mostly using the momentum representation.

4.8. Feynman's path integral

Even within wave mechanics, the bra-ket language allows some calculations which would be very bulky using the traditional notation used in the beginning of this course. Probably the best example in the famous "path integral" formulation of quantum mechanics developed in 1948 by R. Feynman (on the basis of much earlier remark in the pioneering 1930 textbook by P. Dirac).[40]

Let us multiply both parts of Eq. (157), which gives the definition of the time-evolution operator, by the bra-vector of state x,

$$\langle x|\alpha(t)\rangle = \langle x|\hat{u}(t,t_0)|\alpha(t_0)\rangle, \tag{4.268}$$

and insert the identity operator in form (227), with x' replaced by x_0, before the ket-vector in the right-hand part:

$$\langle x|\alpha(t)\rangle = \int dx_0 \langle x|\hat{u}(t,t_0)|x_0\rangle\langle x_0|\alpha(t_0)\rangle. \tag{4.269}$$

According to the wavefunction definition (224), this equality may be presented as

$$\Psi_\alpha(x,t) = \int dx \langle x|\hat{u}(t,t_0)|x_0\rangle \Psi_\alpha(x_0,t_0)\,. \tag{4.270}$$

Comparing this expression with Eq. (2.21), we see that the bra-ket in it is nothing more that the 1D propagator:

$$\langle x|\hat{u}(t,t_0)|x_0\rangle = G(x,t;x_0,t_0)\,. \tag{4.271}$$

In Sec. 2.1, we have calculated it for a free particle – see Eq. (2.23) which we will use very soon.

Now let us break the time segment $[t_0, t]$ into N (for the time being, not necessarily equal) parts by inserting $(N-1)$ intermediate points (Fig. 4)

$$t_0 < t_1 < ... < t_k < ... < t_{N-1} < t\,, \tag{4.272}$$

and rewrite the time evolution operator in the form

$$\hat{u}(t,t_0) = \hat{u}(t,t_{N-1})\hat{u}(t_{N-1},t_{N-2})...\hat{u}(t_2,t_1)\hat{u}(t_1,t_0)\,, \tag{4.273}$$

whose correctness is evident from the very definition (157) of the operator. Now let us plug this expression into Eq. (271), and then insert the identity operator in the form (227),

$$\hat{I} = \int dx_k |x_k\rangle\langle x_k|\,, \tag{4.274}$$

between each two partial evolution operators depending, in particular, on argument t_k. The result is

$$G(x,t;x_0,t_0) = \int dx_{N-1}\int dx_{N-2}...\int dx_1 \langle x|\hat{u}(t,t_{N-1})|x_{N-1}\rangle\langle x_{N-1}|\hat{u}(t_{N-1},t_{N-2})|x_{N-2}\rangle...\langle x_1|\hat{u}(t_1,t_0)|x_0\rangle. \tag{4.275}$$

[40] The standard text on this topic is R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, 1965.

The physical sense of each integration variable x_k is the wavefunction argument at time t_k (Fig. 4).

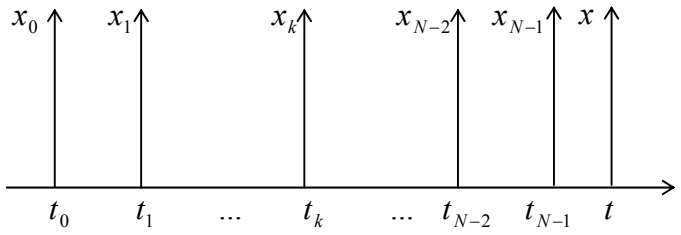

Fig. 4.4. Time partition and coordinate notation at the initial stage of the Feynman's path integral derivation.

Feynman has realized that if all intervals are similar and sufficiently small, $t_k - t_{k-1} = d\tau \to 0$, all the partial bra-kets participating in Eq. (275) may be readily expressed via Eq. (2.23), even if particle is not free, but moves in potential $U(x)$. To show that, let us use either Eq. (162) or Eq. (168) which give the same result for a small time interval $d\tau$:

$$\hat{u}(\tau + d\tau, \tau) = \exp\left\{-\frac{i}{\hbar}\hat{H}d\tau\right\} = \exp\left\{-\frac{i}{\hbar}\left(\frac{\hat{p}^2}{2m}d\tau + Ud\tau\right)\right\}. \tag{4.276}$$

Generally, an exponent of a sum of two operators can be treated as that of *c*-number arguments, in particular, factored into a product of two exponents, only if the operators commute. (Indeed, in that case we can use all the algebra done for exponents of c-number arguments.) In our case, this is not so, because operators $\hat{p}$ and x (and hence $U(x)$) do not commute – see, e.g., Eq. (251). However, it may be argued[41] that for infinitesimal time interval the nonvanishing commutator

$$\left[\frac{\hat{p}^2}{2m}d\tau, U(x)d\tau\right] \neq 0, \tag{4.277}$$

which is proportional to $(d\tau)^2$, is so small that in the first approximation in $d\tau$ its effects may be ignored. As a result, we may factor the exponent by writing

$$\hat{u}(\tau + d\tau, \tau)_{d\tau\to 0} \to \exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}\exp\left\{-\frac{i}{\hbar}U(x)d\tau\right\}. \tag{4.278}$$

(This approximation is very much similar in spirit to the rectangle-formula approximation for a usual 1D integral, which in also impeachable asymptotically.)

Since the second exponent in the right-hand part of Eq. (278) commutes with the coordinate operator, we can move it out of the partial bra-ket:

$$\left\langle x_{\tau+d\tau}\left|\hat{u}(\tau + d\tau, \tau)\right|x_\tau\right\rangle = \left\langle x_{\tau+d\tau}\left|\exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}\right|x_{\tau+d\tau}\right\rangle \exp\left\{-\frac{i}{\hbar}U(x)d\tau\right\}. \tag{4.279}$$

But the remaining bra-ket is just the propagator of a free particle, and we can use Eq. (2.23) for it:

[41] I should confess that a strict proof of this intuitively evident statement would take more space and time than I can afford in this course. However, even with this omission, this derivation of the Feynman's result is more strict than those (of comparable length) in many textbooks.

$$\left\langle x_{\tau+d\tau}\left|\exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}\right|x_\tau\right\rangle=\left(\frac{m}{2\pi i\hbar d\tau}\right)^{1/2}\exp\left\{i\frac{m(dx)^2}{2\hbar d\tau}\right\}. \quad (4.280)$$

As the result, the full propagator (275) takes the form

$$G(x,t;x_0,t_0)=\lim_{\substack{d\tau\to 0\\ N\to\infty}}\int dx_{N-1}\int dx_{N-2}..\int dx_1\left(\frac{m}{2\pi i\hbar d\tau}\right)^{N/2}\exp\left\{\sum_{k=1}^{N}\left[i\frac{m(dx)^2}{2\hbar d\tau}-i\frac{U(x)}{\hbar}d\tau\right]\right\}. \quad (4.281)$$

At $d\tau=(t-t_0)/N\to 0$, the exponent in this expression tends to

$$\sum_{k=1}^{N}\frac{i}{\hbar}\left[\frac{m}{2}\left(\frac{dx}{dt}\right)^2-U(x)\right]dt\to\frac{i}{\hbar}\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2-U(x)\right]d\tau, \quad (4.282)$$

so that the expression in square brackets is just the particle's Lagrangian function L (see, e.g., CM Sec. 2.1). The integral of the function over time is the classical action S calculated along a particular "path" $x(\tau)$ – see CM, Sec. 9.2. As a result, defining the *path integral* as

$$\int(...)D[x(\tau)]\equiv\lim_{\substack{d\tau\to 0\\ N\to\infty}}\left(\frac{m}{2\pi i\hbar d\tau}\right)^{N/2}\int dx_{N-1}\int dx_{N-2}..\int dx_1(...), \quad (4.283)$$

we can bring our result to an extremely simple form

$$G(x,t;x_0,t_0)=\int\exp\left\{\frac{i}{\hbar}S[x(\tau)]\right\}D[x(\tau)]. \quad (4.284)$$

The term "path integral" for the mathematical construct (283) may be readily explained if we keep the number N of time intervals large but finite, and also approximate each of the enclosed integrals in Eq. (283) by M discrete points (Fig. 5a).

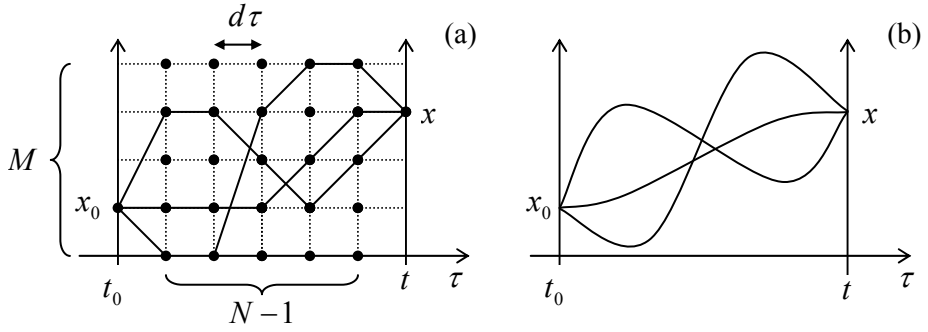

Fig. 4.5. Several 1D classical paths in (a) the discrete approximation and (b) the continuous limit.

Then the path integral is a product of (N - 1) sums corresponding to different values of time τ, each of them with M terms, each representing the function under the integral in a particular spatial point. Multiplying those (N – 1) sums, we get a single sum of (N - 1)M terms, each evaluating the function in a specific spatial-temporal point [x, τ]. These terms may be now grouped to present all possible and different classical paths $x[\tau]$ from the initial point [x_0,t_0] to the finite point [x,t]. It is evident that the last interpretation remains true in the continuous limit (Fig. 5b).

Now let us explore Eq. (284) – unfortunately, not in much detail due to our time restrictions.[42] First, from classics we know that at certain initial conditions, the particle would follow just one path, corresponding to the minimum of action S. Let us see what is the value of the propagator on that path. Leaving alone the normalization issues (which are not very important because of the dominating role of the exponent under the integral at quasiclassical motion, when $S_{cl} >> \hbar$), we get

$$G_{cl} \propto \exp\left\{\frac{i}{\hbar} S_{cl}\right\} = \exp\left\{\frac{i}{\hbar}\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 - U(x)\right]d\tau\right\}. \tag{4.285}$$

The sum of the kinetic and potential energies is the full energy E of the particle, which remains constant for motion in a stationary potential $U(x)$, so that we can present the expression under the integral as

$$\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 - U(x)\right]d\tau = \left[m\left(\frac{dx}{d\tau}\right)^2 - E\right]d\tau = m\frac{dx}{d\tau}dx - Ed\tau. \tag{4.286}$$

With that replacement, Eq. (285) yields

$$G_{cl} \propto \exp\left\{\frac{i}{\hbar}\int_{x_0}^{x} m\frac{dx}{d\tau}dx\right\}\exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\} = \exp\left\{\frac{i}{\hbar}\int_{x_0}^{x} p(x)dx\right\}\exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\}, \tag{4.287}$$

where p is the classical momentum of the particle. But this (at least, leaving the pre-exponential factor alone) is exactly the WKB approximation result which was derived and studied in detail in Chapter 2!

It is natural to ask what would be the effect of other paths in the quasiclassical limit. Actually, this issue may be answered in detail, though I will not exhibit the corresponding math. The key here is to focus on a bundle of paths which are very close to the classical one – so close that the deviation of action S on each path from its classical (minimum) value S_{cl} is of the order of $\hbar$. In that case, the difference $(S - S_{cl})$ may be presented as a quadratic functional of deviations from the classical trajectory. (Linear terms vanish because S_{cl} is the variational minimum of action) As a result, the path integral may be presented as a product of the classical-path result (287) by a multidimensional Gaussian integral which may be worked out exactly and turns into a pre-exponential coefficient which we have already found in Chapter 2 when refining the WKB approximation.

One can question the value of a result, which could be obtained readily from Schrödinger's wave mechanics, by such complex means as path integration. However, Feynman's approach has merits. First, it has an important philosophical value. Indeed, Eq. (284) may be interpreted by saying that quantum mechanics is the exploration, by the system, of all possible classical trajectories, with the resulting partial propagators added up coherently to form the final propagator, and via it, the probability $w \propto |G|^2$ of the particle propagation from $[x_0,t_0]$ to $[x,t]$. Of course, as the action (i.e. of the energy-by-time product) scale of the effect decreases and approaches $\hbar$, more and more paths produce substantial contribution to this sum, and hence to w, ensuring a larger and larger difference between the quantum and classical properties of the system.

[42] For a more thorough review of path integral applications, the reader may be referred to the 1965 Feynman-Hibbs book cited above, and more recent texts, for example by L. S. Schulman, *Techniques and Applications of Path Integration*, Wiley, 1981.

Second, the path integral provides a justification for some simple explanations of quantum phenomena. A typical example is the quantum interference effects discussed in Sec. 3.2 – see, e.g., Fig. 3.1 and the corresponding text. At that discussion, we conjectured that at the two-slit interference, the WKB approximation might be restricted of effects by two path which pass through different slits, but otherwise consisting of straight-line segments. To prove that assumption, we should first generalize the path integral to multi-dimensional geometries. Fortunately, the simple structure of Eq. (284) makes such generalization evident:

$$G(\mathbf{r},t;\mathbf{r}_{0,}t_0)=\int\exp\left\{\frac{i}{\hbar}S[\mathbf{r}(\tau)]\right\}D[\mathbf{r}(\tau)],\qquad S=\int_{t_0}^{t}L(\mathbf{r},\frac{d\mathbf{r}}{d\tau})d\tau=\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{d\mathbf{r}}{d\tau}\right)^2-U(\mathbf{r})\right]d\tau. \tag{4.288}$$

though of course definition (283) of the path integral should be also modified appropriately. Most importantly, in the quasiclassical limit $S_{cl} >> \hbar$, Eq. (288) may be also reduced to multi-dimensional Gaussian integral, with the same main result: besides the pre-exponential factor, the propagator is reduced to the WKB exponent taken along the classical trajectory (or trajectories). For the Young-type experiment (Fig. 3.1), conditioned by particle passage through *some* slit, these trajectories are the straight-line segments shown in Fig. 3.1, validating the simple analysis of the interference in Sec. 3.2.

Finally, the path integral allows simple solutions of some problems, which would be hard to get by other ways. Let us consider the problem of tunneling in more than one dimension, which is depicted in Fig. 6 for the 2D case (just for graphics simplicity). Here, potential $U(x,y)$ has the "saddle" shape. (Another helpful image is a mountain path between two summits, in Fig. 6 located on the top and bottom of the drawing.) Particle may move classically in the left and right regions with $E > U(x,y)$, but can pass from one of these region to another one only via the quantum-mechanical tunneling under the pass. Let us calculate the transparency of this tunnel barrier in the WKB approximation, ignoring the possible pre-exponential factor.

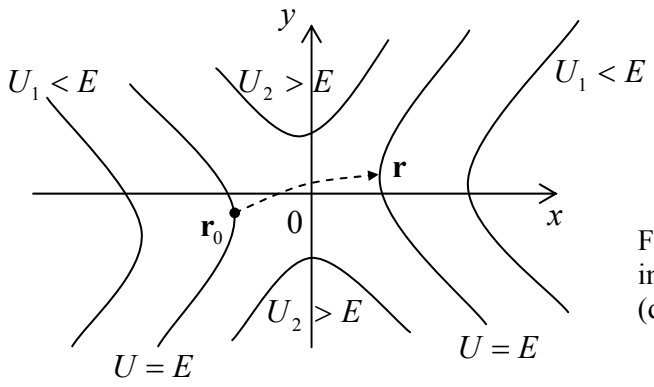

Fig. 4.6. Saddle-type 2D potential profile and the instanton trajectory of a particle of energy E (dashed line, schematically).

Repeating the simple calculation which has led to Eq. (287), for the classically region forbidden region where $E < U(x,y)$, the contribution of a path to propagator becomes

$$G\propto\exp\left\{-\int_{\vec{r}_0}^{\vec{r}}\boldsymbol{\kappa}(\mathbf{r})\cdot d\mathbf{r}\right\}\exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\}, \tag{4.289}$$

where the magnitude of vector $\boldsymbol{\kappa}$ should be calculated just in the 1D case (see, e.g., Eq. (2.38)),

$$\frac{\hbar^2\kappa^2(\mathbf{r})}{2m} = U(\mathbf{r}) - E, \tag{4.290}$$

while its direction should be tangential to the path trajectory in space. Now the path integral is actually much simpler than in the classically-allowed region, because the exponents are purely real and there is no complex interference like that in Eqs. (284) or (288). The largest contribution evidently comes from the trajectory (or rather a narrow bundle of trajectories) for which functional

$$\int_{\mathbf{r}_0}^{\mathbf{r}} \boldsymbol{\kappa}(\mathbf{r}') \cdot d\mathbf{r}' \tag{4.291}$$

has the smallest value, and the transmission coefficient can be calculated as

$$T \approx |G|^2 \approx \exp\left\{-2\int_{\mathbf{r}_0}^{\mathbf{r}} \boldsymbol{\kappa}(\mathbf{r}') \cdot d\mathbf{r}'\right\}. \tag{4.292}$$

(One can argue that the pre-exponential coefficient in that equation should be close to 1, just like in Eq. (2.92), especially if the potential is smooth in the sense of Eq. (2.82)).

Thus the tunneling problem is reduced to finding the trajectory which minimizes functional (291). This is of course a well-known problem of the calculus of variations,[43] but it is interesting that the path integral provides a simple alternative way to solve it. Let us consider an auxiliary problem, with U inverted relative to E:

$$U_{\text{inv}}(\mathbf{r}) - E = E - U(\mathbf{r}). \tag{4.293}$$

As was discussed above, the path integral for the WKB motion in the classically allowed region of potential $U_{\text{inv}}(x,y)$ (which coincides with the classically forbidden region of the original problem) is dominated by the classical trajectory which corresponds to the minimum of

$$S_{\text{cl}} = \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{p}_{\text{inv}}(\mathbf{r}') \cdot d\mathbf{r}' = \hbar \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{k}_{\text{inv}}(\mathbf{r}') \cdot d\mathbf{r}, \tag{4.294}$$

where k_i should be determined from the relation

$$\frac{\hbar^2 k_{inv}^2(\mathbf{r})}{2m} \equiv E - U_{\text{inv}}(\mathbf{r}). \tag{4.295}$$

But comparing Eqs. (290), (293), and (295), we see that $\mathbf{k}_{\text{inv}} = \boldsymbol{\kappa}$ at each point of space! This means that the tunneling path corresponds to the "*instanton trajectory*"[44] of the same particle in the inverted potential U_i. This trajectory may be readily found by means of classical mechanics if its initial ($\mathbf{r}_0$) and final ($\mathbf{r}$) points (see Fig. 6) are known. Thus the problem is reduced to much simpler task of minimizing functional (294) over the positions of these two points on the equipotential lines $U(x,y) = E$. For many

[43] See, e.g., I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Dover, 2000.

[44] In our context, the term "instanton" is used for the solution of this auxiliary classical problem. In quantum field theory, the instanton concept may be formulated differently, and has much more complex applications (see, e.g. R. Rajaraman, *Solitons and Instantons*, North Holland, 1987) but for our simple problem, it is reduced to the idea described above.

symmetric potentials (such as one in Problem 13), the points may be readily found from the problem symmetry, thus completing the calculation of the barrier transparency (292).

4.9. Revisiting harmonic oscillator

Let us return to the 1D harmonic oscillator described by Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2 \hat{x}^2}{2}. \tag{4.296}$$

In Sec. 2.6 we have already used the "brute-force" (wave-mechanics) approach to find the eigenfunctions $\psi_n(x)$ and eigenvalues E_n of this Hamiltonian, and seen that, unfortunately, that approach required relatively complex math which obscures the physics of these stationary ("Fock") states. Now let us use the bra-ket formalism to make the properties of these states much more transparent, with very simple calculations.

Let us start from introducing normalized (dimensionless) operators of coordinates and momentum:

$$\hat{\xi} \equiv \frac{\hat{x}}{x_0}, \quad \hat{\zeta} \equiv \frac{\hat{p}}{m\omega_0 x_0}, \tag{4.297}$$

where x_0 is the natural scale of the spread of ground-state wavefunction, which already was discussed in Sec. 2.6:

$$x_0 \equiv \left(\frac{\hbar}{m\omega_0}\right)^{1/2}. \tag{4.298}$$

In these variables, the Hamiltonian becomes very simple and symmetric:

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\hat{\xi}^2 + \hat{\zeta}^2\right). \tag{4.299}$$

Now, let us introduce a new operator:

$$\hat{a} \equiv \frac{1}{\sqrt{2}}\left(\hat{\xi} + i\hat{\zeta}\right) = \left(\frac{m\omega_0}{2\hbar}\right)^{1/2}\left(\hat{x} + i\frac{\hat{p}}{m\omega_0}\right), \tag{4.300}$$

and its Hermitian conjugate $\hat{a}^\dagger$. Since $\hat{\xi}$ and $\hat{\zeta}$ are Hermitian (self-adjoint), we get

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\hat{\xi} - i\hat{\zeta}\right) = \left(\frac{m\omega_0}{2\hbar}\right)^{1/2}\left(\hat{x} - i\frac{\hat{p}}{m\omega_0}\right), \tag{4.301}$$

The reciprocal relation is

$$\hat{\xi} = \frac{1}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right), \quad \hat{\zeta} = \frac{1}{\sqrt{2}i}\left(\hat{a} - \hat{a}^\dagger\right), \tag{4.302}$$

so that

$$\hat{\xi}^2 = \frac{1}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right), \quad \hat{\zeta}^2 = -\frac{1}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\right). \tag{4.303}$$

When deriving these expressions, we had to be careful to avoid swapping $\hat{a}$ and $\hat{a}^\dagger$ in their products, because these operators do not commute. Indeed, using Eqs. (300) and (301) together with Eq. (241), we readily get

$$\left[\hat{a}, \hat{a}^\dagger\right] = \frac{m\omega_0}{2\hbar}\left(-\frac{i}{m\omega_0}\left[\hat{x}, \hat{p}\right] + \frac{i}{m\omega_0}\left[\hat{p}, \hat{x}\right]\right) = \hat{I}. \tag{4.304}$$

Now we can plug Eqs. (303) into (299) to express the Hamiltonian in terms of our new operators:

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right). \tag{4.305}$$

It is more convenient to recast the commutation relation (304) into the form

$$\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + \hat{I} \tag{4.306}$$

which enables us to rewrite Eq. (305) as

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(2\hat{a}^\dagger\hat{a} + \hat{I}\right) \equiv \hbar\omega_0\left(\hat{N} + \frac{1}{2}\hat{I}\right), \tag{4.307}$$

where we have introduced one more (evidently, Hermitian) operator

$$\hat{N} \equiv \hat{a}^\dagger\hat{a}. \tag{4.308}$$

The most important conclusion one can do from Eq. (307) is that that operators $\hat{H}$ and $\hat{N}$ do commute, and thus share the set of stationary (Fock) eigenstates n, and write

$$\hat{N}|n\rangle = N_n|n\rangle, \tag{4.309}$$

where N_n are some eigenvalues which, according to Eq. (307), determine also the energy spectrum of the oscillator:

$$E_n = \hbar\omega_0\left(N_n + \frac{1}{2}\right). \tag{4.310}$$

So far, we know only that all N_n are real, but not much more. In order to find these eigenvalues, let us carry out the following calculation (splendid in its simplicity and efficiency). Consider the result of action of operator $\hat{N}$ on ket-vector $\hat{a}^\dagger|n\rangle$. Using the associative rule to remove and insert parentheses wherever we like, we may write

$$\hat{N}\left(\hat{a}^\dagger|n\rangle\right) = \left(\hat{a}^\dagger\hat{a}\right)\left(\hat{a}^\dagger|n\rangle\right) = \hat{a}^\dagger\hat{a}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger\left(\hat{a}\hat{a}^\dagger\right)|n\rangle. \tag{4.311}$$

Using the commutation relation (306), and then Eq. (309), we may continue as

$$\hat{a}^{\dagger}\left(\hat{a}\hat{a}^{\dagger}\right)|n\rangle = \hat{a}^{\dagger}\left(\hat{a}^{\dagger}\hat{a}+\hat{I}\right)|n\rangle = \hat{a}^{\dagger}\left(\hat{N}+\hat{I}\right)|n\rangle = \hat{a}^{\dagger}\left(N_n+1\right)|n\rangle = \left(N_n+1\right)\left(\hat{a}^{\dagger}|n\rangle\right). \quad (4.312)$$

Let us summarize our result:

$$\hat{N}\left(\hat{a}^{\dagger}|n\rangle\right) = \left(N_n+1\right)\left(\hat{a}^{\dagger}|n\rangle\right). \quad (4.313)$$

Acting absolutely similarly, we can get

$$\hat{N}\left(\hat{a}|n\rangle\right) = \left(N_n-1\right)\left(\hat{a}|n\rangle\right). \quad (4.314)$$

It is time to stop and translate our results into plain English: if $|n\rangle$ is an eigenket of operator $\hat{N}$ with eigenvalue N_n, then $\hat{a}^{\dagger}|n\rangle$ and $\hat{a}|n\rangle$ are also eigenkets of that operator, with eigenvalues $(N_n + 1)$, and $(N_n - 1)$, respectively. This statement may be presented by the "ladder diagram" shown in Fig. 7. This diagram shows very clearly that operator $\hat{a}^{\dagger}$ moves the system a step up the ladder, while operator $\hat{a}$ brings it one step down. In other words, the former operator creates a new *excitation* of the system (for electromagnetic field oscillators, called the *photon*; for mechanical wave field oscillators, called the *phonon*, etc.), whiles the latter operator kills (annihilates) such excitation. This is why $\hat{a}^{\dagger}$ is called the *creation operator*, and $\hat{a}$, the *annihilation operator*. In its turn, according to Eq. (309), operator $\hat{N}$ does not change the state of the system, but "counts" its position on the ladder:

$$\langle n|\hat{N}|n\rangle = \langle n|N_n|n\rangle = N_n. \quad (4.315)$$

This is why it is called the "*number operator*", in our current context meaning the number of excitations in the system.

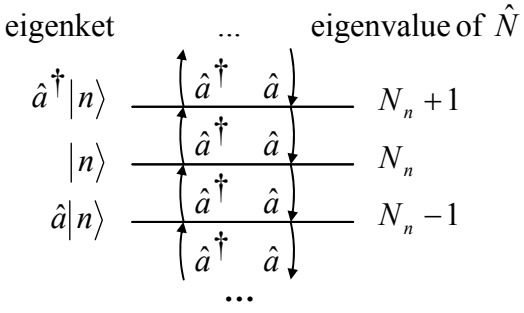

Fig. 4.7. The hierarchy ("ladder") of eigenstates of the 1D harmonic oscillator. Arrows show the action of the creation and annihilation operators on the eigenstates.

We still need to complete our calculation. Indeed, we still do not know (i) whether the ladder shown in Fig. 7 shows all eigenstates of the oscillator, and (ii) what exactly are numbers N_n. Fascinating enough, both questions may be answered by exploring one paradox. Let us start with some state of the ladder, and keep going down it, applying operator $\hat{a}$ again and again. Each time, eigenvalue N_n is decreased by one, so that eventually it should become negative. However, this cannot be, because < 0, because any real state, including states presented by kets $|d\rangle \equiv \hat{a}|n\rangle$ and $|n\rangle$, should have a positive norm defined by Eq. (16). Calculating them,

$$\|n\|^2 = \langle n|n\rangle, \quad \|d\|^2 = \langle n|\hat{a}^\dagger a|n\rangle = \langle n|\hat{N}|n\rangle = N_n\langle n|n\rangle, \tag{4.316}$$

we see that the both norms cannot be positive if N_n is negative.

The only resolution of this paradox is to notice that the action of the creation and annihilation operators on Fock states may consist in not only their promotion to the next step, but also by their multiplication by some *c*-number:

$$\hat{a}|n\rangle = A_n|n-1\rangle, \quad \hat{a}^\dagger|n\rangle = B_n|n+1\rangle. \tag{4.317}$$

(Linear relations (313) and (314) clearly allow that.) In order to find A_n, we should require that if state n has been normalized, so is state (n -1):

$$\langle n|n\rangle = 1, \quad \langle n|\hat{a}^\dagger A_n^* A_n a|n\rangle = \langle n|A_n^* A_n \hat{N}|n\rangle = A_n^* A_n N_n\langle n|n\rangle = 1, \tag{4.318}$$

Selecting the phase of the complex coefficient A_n equal zero for convenience we get $A_n = (N_n)^{1/2}$, i.e.

$$\hat{a}|n\rangle = N_n^{1/2}|n-1\rangle. \tag{4.319}$$

Now let us assume that all numbers N_n are integers. (Because of the definition of N_n, given by Eq. (309), it is convenient to call these integers n, i.e. by the same letter as the corresponding eigenstate.) Then when we have come down to state with $n = 0$, an attempt to make one more step down gives

$$\hat{a}|0\rangle = 0|-1\rangle. \tag{4.320}$$

But in accordance with Eq. (10), the state in the right-hand part of this equation is the "null-state" **0**, i.e. does not exist in nature. (Please notice the radical difference between that state and the state described by ket-vector $|0\rangle$ in the left-hand side of this equation. The latter state *does* exist and, moreover, presents the most important, ground state of the system; its wavefunction is given by a simple Gaussian – see Eq. (2.230)). This gives the (only known:-) resolution of the state ladder paradox: the ladder has the lowest step with $N_n = n = 0$.

As a by-product of our discussion we have obtained a very important relation $N_n = n$ which means in particular, that the state ladder includes *all* eigenstates of the oscillator. Plugging this relation into Eq. (310), we see that the full spectrum of eigenenergies of the harmonic oscillator is described by the simple formula

$$E_n = \hbar\omega_0\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2... \tag{4.321}$$

which we already cited in Sec. 2.6. It is of course remarkable that the bra-ket formalism has allowed us to derive it without calculation of the corresponding wavefunctions $\psi_n(x)$ (2.241). Moreover, the formalism also allows one to calculate virtually any observable, and even virtually any Fock-state bra-ket, without using $\psi_n(x)$. Indeed, let us rewrite Eq. (319) with $N_n = n$, and the (similarly derived) expression for the creation operator:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{4.322}$$

Now we can use these formulas to calculate, for example, an arbitrary Fock-state bra-ket[45] of operator $\hat{x}$, plugging it in the form given by Eq. (302):

$$\langle n'|\hat{x}|n\rangle = x_0\langle n'|\hat{\xi}|n\rangle = \frac{x_0}{\sqrt{2}}\langle n'|\left(\hat{a}+\hat{a}^\dagger\right)|n\rangle = \frac{x_0}{\sqrt{2}}\left(\langle n'|\hat{a}|n\rangle + \langle n'|\hat{a}^\dagger|n\rangle\right). \tag{4.323}$$

To complete the calculation, we may now use Eqs. (322) and the Fock state orthonormality:

$$\langle n'|n\rangle = \delta_{n'n}. \tag{4.324}$$

The result is

$$\langle n'|\hat{x}|n\rangle = \frac{x_0}{\sqrt{2}}\left(\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right) = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2}\left(\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right). \tag{4.324}$$

The best image of this result is the matrix of the coordinate operator in the Fock-state basis has only two diagonals, adjacent to the main diagonal; all other elements (including the diagonal ones) are zeros. Acting absolutely similarly, for momentum bra-kets we can get a similar expression:

$$\langle n'|\hat{p}|n\rangle = i\frac{m\omega_0 x_0}{\sqrt{2}}\left(-\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right) = i\left(\frac{\hbar m\omega_0}{2}\right)^{1/2}\left(-\sqrt{n}\delta_{n',n-1} + \sqrt{n+1}\delta_{n',n+1}\right) \tag{4.325}$$

Matrix elements of higher powers of the operators, as well as their products, may be handled similarly. For example, to calculate an arbitrary bra-ket of $\hat{x}^2$ we can use either its direct expression (303) via the creation-annihilator operators, or the previous result (324):[46]

$$\langle n'|\hat{x}^2|n\rangle = \langle n'|\hat{x}\hat{x}|n\rangle = \sum_{n''=0}^{\infty}\langle n'|\hat{x}|n''\rangle\langle n''|\hat{x}|n\rangle$$

$$= \frac{x_0^2}{2}\sum_{n''=0}^{\infty}\left(\sqrt{n''}\delta_{n',n''-1} + \sqrt{n''+1}\delta_{n',n''+1}\right)\left(\sqrt{n}\delta_{n'',n-1} + \sqrt{n+1}\delta_{n'',n+1}\right) \tag{4.326}$$

$$= \frac{x_0^2}{2}\left[\sqrt{n(n-1)}\delta_{n',n-2} + \sqrt{(n+1)(n+2)}\delta_{n',n+2} + (2n+1)\delta_{n',n}\right]$$

The most important matrix elements are those on its main diagonal:

$$\langle n|\hat{x}^2|n\rangle = \frac{x_0^2}{2}(2n+1). \tag{4.327}$$

This expression gives, in particular, the expectation value of oscillator's potential energy in *n*-th Fock state:

$$\langle U\rangle = \frac{m\omega_0^2}{2}\langle x^2\rangle = \frac{m\omega_0^2 x_0^2}{2}\left(n+\frac{1}{2}\right) = \frac{\hbar\omega_0}{2}\left(n+\frac{1}{2}\right). \tag{4.328}$$

[45] Note that such bra-ket is nothing else that the matrix element of the operator in the Fock state basis.

[46] Of course, the higher the operator power the more cumbersome such calculations become – see Problem 15.

This is exactly ½ of the total energy (321) of the oscillator. As a sanity check, an absolutely similar calculation shows that

$$\left\langle \frac{p^2}{2m} \right\rangle = \frac{1}{2m}\langle n|\hat{p}^2|n\rangle = \frac{\hbar\omega_0}{2}\left(n+\frac{1}{2}\right), \tag{4.329}$$

i.e. also equals $E_n/2$, just as in a classical oscillator.

4.10. Glauber and squeezed states

There is evidently a huge difference between a quantum stationary (Fock) state of the oscillator and its classical state. Indeed, let us write the classical equations of motion of the oscillator (using capital letters to distinguish them from the arguments used at quantum description of the system):

$$\begin{aligned} \frac{dX}{dt} &= \frac{P}{m}, \\ \frac{dP}{dt} &= -\frac{\partial U}{\partial x} = -m\omega_0^2 X. \end{aligned} \tag{4.330}$$

On the "phase plane" with Cartesian coordinates x and p (Fig. 8), these equations describe a clockwise rotation of the "phase point" $\{X(t), P(t)\}$ along an elliptic trajectory specified by the initial conditions $\{X(0), P(0)\}$. (Normalization of momentum by $m\omega_0$ makes the trajectory pleasantly circular.)

On the other hand, the time dependence of the Fock state, as of any stationary state, is exponential,

$$\Psi_n(x,t) = \langle x|n\rangle \exp\left\{-i\frac{E_n}{\hbar}t\right\}, \tag{4.331}$$

and a result, gives time-independent expectation values of x, p, or any function thereof. The best classical image for such a state on the phase plane is a circle of radius $r = x_0(2n + 1)^{1/2}$, along which the wavefunction is uniformly spread as a standing wave.

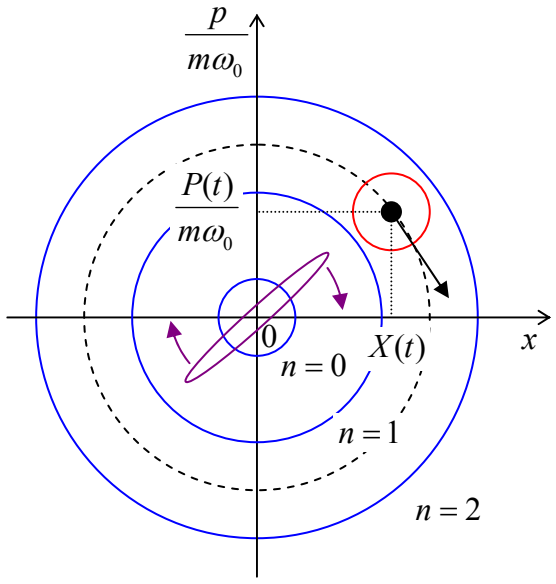

Fig. 4.8. Various states of a harmonic oscillator (schematically only!). Solid dot shows a classical state, with the dashed lines showing its trajectory. Red lines are images of stationary (Fock) states, while the blue-line circle, the "coherent" (Glauber) state closest to the classical state. Finally, a magenta ellipse shows a squeezed state (for the simplest case $\alpha = 0$), and the arrow, the direction of its time evolution.

It is natural to ask how to form a quantum state whose properties would be closer to the classical ones. Such states, with the center in the classical point $\{X(t), P(t)\}$, and the smallest possible product of quantum uncertainties of coordinate and momentum, are called *Glauber states*.[47] Conceptually the simplest way to present the Glauber state $|\alpha\rangle$ is as the ground Fock state $|0\rangle$ with the center shifted from the origin to the classical point $\{X(t), P(t)\}$. (After such a shift, the state automatically rotates, following the classical motion.) Let us study how this shift may be implemented (or rather described) in quantum mechanics. The mechanism for such shifts are called the *translation operators*.

Let us start with finding such operator $\hat{\mathcal{I}}_X$ for the desirable shift of coordinate by X. Say, we had a 1D wave packet in the standard form (260)

$$\psi(x) \equiv \frac{1}{\sqrt{2\pi\hbar}} \int \varphi(p) \exp\left\{ i \frac{px}{\hbar} \right\} dp \,. \tag{4.332}$$

Evidently, the result of its shift by X is

$$\hat{\mathcal{I}}_X \psi(x) = \psi(x - X) = \frac{1}{\sqrt{2\pi\hbar}} \int \varphi(p) \exp\left\{ i \frac{p(x - X)}{\hbar} \right\} dp \,. \tag{4.333}$$

Hence the shift is achieved by the multiplication of each Fourier (fixed-momentum) component of the packet by $\exp\{-ipX/\hbar\}$. This gives us a hint (not a proof yet!) that the general form of the shift operator, valid in any representation, is

$$\hat{\mathcal{I}}_X = \exp\left\{ -i \frac{\hat{p}X}{\hbar} \right\}. \tag{4.334}$$

The proof may be readily achieved by at least two ways, both using the fact that (for example, see Eq. (60)), each operator is uniquely determined by the set of its matrix elements in any full and orthogonal basis. For example, the analog of Eq. (229) for the p-representation, applied to the translation operator, is

$$\left\langle p \left| \hat{\mathcal{I}}_X \right| p' \right\rangle = \delta(p - p') \exp\left\{ -i \frac{pX}{\hbar} \right\} \varphi(p). \tag{4.335}$$

Comparing it with Eq. (333) we see that this operator does exactly the job we need it to.

Providing a shift by P is absolutely similar (with the opposite sign, due to the opposite sign of the exponent in the reciprocal Fourier transform (261)), so that the shift by both X and P may be achieved by operator

$$\hat{\mathcal{I}}_\alpha = \exp\left\{ i \frac{P\hat{x} - \hat{p}X}{\hbar} \right\} = \exp\left\{ \alpha \hat{a}^\dagger - \alpha^* \hat{a} \right\}, \quad \hat{\mathcal{I}}_\alpha{}^\dagger = \exp\left\{ \alpha^* \hat{a} - \alpha \hat{a}^\dagger \right\}, \tag{4.336}$$

[47] After R. J. Glauber who studied these states in detail in the mid-1965, though they were known to E. Schrödinger as early as in 1928. Another popular name, "*coherent states*", does not make much sense, because all the quantum states we have studied so far (including the Fock states) may be presented as coherent superpositions.

where α is the (normalized) complex amplitude of the classical oscillations we are trying to approximate:

$$\alpha \equiv \frac{1}{\sqrt{2}x_0}\left(X + i\frac{P}{m\omega_0}\right), \tag{4.337}$$

so that the Glauber state is

$$|\alpha\rangle = \hat{\mathfrak{I}}_\alpha |0\rangle. \tag{4.338}$$

Working directly with the shift operator is not too convenient because of its exponential form, but turns out that a much simpler presentation for the Glauber state is possible. To show than, let us start with the following general property of exponential operators: if

$$\left[\hat{A}, \hat{B}\right] = \mu\hat{I}, \tag{4.339}$$

(where $\hat{A}$ and $\hat{B}$ are arbitrary operators, and μ is an arbitrary *c*-number), then

$$\exp\{+\hat{A}\}\hat{B}\exp\{-\hat{A}\} = \hat{B} + \mu\hat{I}. \tag{4.340}$$

(A proof is the statement, with is the subject of Problem 17, may be readily achieved by expanding operator

$$\hat{f}(\lambda) \equiv \exp\{+\lambda\hat{A}\}\hat{B}\exp\{-\lambda\hat{A}\} \tag{4.341}$$

into the Taylor series in *c*-number parameter λ, and then evaluation it at $\lambda = 1$.)

Now let us apply Eq. (440) to two cases, both with

$$\hat{A} = \alpha^* \hat{a} - \alpha\hat{a}^\dagger, \quad \text{i.e. } \exp\{+\hat{A}\} = \hat{\mathfrak{I}}_\alpha^\dagger, \quad \exp\{-\hat{A}\} = \hat{\mathfrak{I}}_\alpha. \tag{4.342}$$

First, let us take $\hat{B} = \hat{I}$, then from Eq. (339) we have $\mu = 0$, and Eq. (440) yields

$$\hat{\mathfrak{I}}_\alpha^\dagger \hat{\mathfrak{I}}_\alpha = \hat{I}, \tag{4.343}$$

This merely means that the shift operator is unitary – not a big surprise, because if we shift the phase point by $(+\alpha)$ and then by $(-\alpha)$, we certainly come back to the initial position. Equation (443) just means that this is true for a quantum state as the whole.

Second, let us take $\hat{B} = \hat{a}$; then we need to calculate commutator (339), at the due moment using Eq. (304):

$$\left[\hat{A}, \hat{B}\right] = \left[\alpha^* \hat{a} - \alpha\hat{a}^\dagger, \hat{a}\right] = -\alpha\left[\hat{a}^\dagger, \hat{a}\right] = \alpha\hat{I}, \tag{4.344}$$

so that $\mu = \alpha$, and Eq. (440) yields

$$\hat{\mathfrak{I}}_\alpha^\dagger \hat{a} \hat{\mathfrak{I}}_\alpha = \hat{a} + \alpha\hat{I}. \tag{4.345}$$

Now let us consider operator

$$\hat{\mathfrak{I}}_\alpha \hat{\mathfrak{I}}_\alpha^\dagger \hat{a} \hat{\mathfrak{I}}_\alpha . \tag{4.346}$$

Applying to it Eq. (443), we get $\hat{a}\hat{\mathfrak{I}}_\alpha$, while the application of Eq. (445) yields $\hat{\mathfrak{I}}_\alpha \hat{a} + \alpha \hat{\mathfrak{I}}_\alpha$. Hence, we get the following operator equality

$$\hat{a}\hat{\mathfrak{I}}_\alpha = \hat{\mathfrak{I}}_\alpha \hat{a} + \alpha \hat{\mathfrak{I}}_\alpha . \tag{4.347}$$

Now applying this equality to the ground state $|0\rangle$ and using the facts that $\hat{a}|0\rangle = \mathbf{0}$, while $\hat{\mathfrak{I}}_\alpha |0\rangle \equiv |\alpha\rangle$, we finally get a very simple and elegant result:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle . \tag{4.348}$$

Thus *the Glauber state is an eigenstate of the annihilation operator*, corresponding to the eigenvalue α, i.e. to the (normalized) complex amplitude of the classical process approximated by the state.[48] This fact makes the calculations of the Glauber state properties much simpler.

First of all, let us find coefficients α_n in the presentation of the Glauber state as a superposition of Fock states:

$$|\alpha\rangle = \sum_{n=0}^{\infty} \alpha_n |n\rangle . \tag{4.349}$$

Plugging this expression into both sides of Eq. (448), and using the first of Eqs. (322), we get the following recurrent relation for the coefficients:

$$\alpha_{n+1} = \frac{\alpha}{\sqrt{n+1}} \alpha_n . \tag{4.350}$$

Assuming some value of α_0, and applying the relation sequentially for $n = 1, 2$, etc., we get

$$\alpha_n = \frac{\alpha^n}{(n!)^{1/2}} \alpha_0 . \tag{4.351}$$

Now we can find α_0 from the normalization requirement $\langle \alpha | \alpha \rangle = 1$, getting

$$|\alpha_0|^2 \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = 1. \tag{4.352}$$

In the sum we can readily recognize the Taylor expansion of function $\exp\{|\alpha|^2\}$, so that the final result is

$$|\alpha\rangle = \exp\left\{-\frac{|\alpha|^2}{2}\right\} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle . \tag{4.353}$$

[48] Notice that the eigenvalue spectrum of Glauber states is continuous. Also, in the limit $\alpha = 0$, the Glauber state coincides with the (Fock) ground state.

It means in particular that the probability $P_n \equiv \alpha_n \alpha_n^*$ of finding the system energy on a certain level (321) obeys the well-known Poisson distribution (Fig. 9),

$$W_n = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}, \tag{4.354}$$

where in our case

$$\langle n \rangle = |\alpha|^2. \tag{4.355}$$

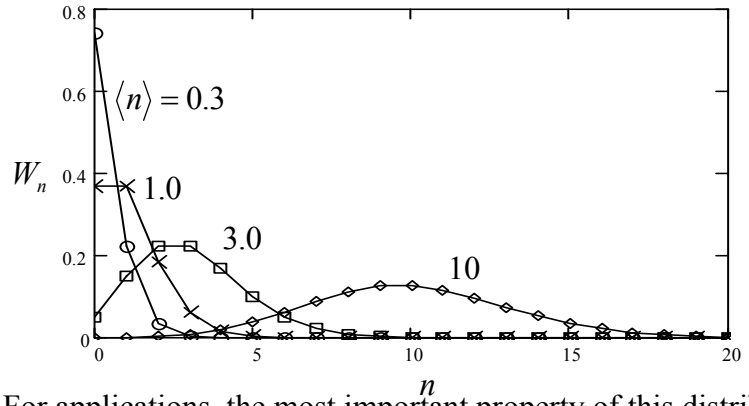

Fig. 4.9. The Poisson distribution for several values of $\langle n \rangle$. Note that W_n are defined only for integer values of n; lines are only guides for the eye.

For applications, the most important property of this distribution is

$$\delta n = \langle n \rangle^{1/2}; \tag{4.356}$$

notice also that at $\langle n \rangle >> 1$, and hence $\delta n << \langle n \rangle$ the Poisson distribution approaches the "normal" (Gaussian) one.

Time dynamics of Glauber states may be readily described in the Schrödinger picture, by replacing α for its time-dependent value

$$\alpha(t) \equiv \frac{1}{\sqrt{2}x_0}\left(X(t) + i\frac{P(t)}{m\omega_0} \right), \tag{4.357}$$

where $\{X(t), P(t)\}$ is the solution to the classical equations of motion (330). By the way, it is even easier to merge them into a single equation for $\alpha(t)$:

$$\frac{d\alpha(t)}{dt} = -i\omega_0 \alpha(t), \tag{4.358}$$

with the evident solution

$$\alpha(t) = \alpha(0)\exp\{-i\omega_0 t\}. \tag{4.357}$$

An alternative is to use the Heisenberg equation of motion (whose derivation is the subject of Problem 14), which are absolutely similar to the classical equation (458):

$$\frac{d\hat{a}_{\rm H}(t)}{dt} = -i\omega_0 \hat{a}_{\rm H}(t), \quad \frac{d\hat{a}_{\rm H}^\dagger(t)}{dt} = +i\omega_0 \hat{a}_{\rm H}^\dagger(t), \tag{4.358}$$

and hence same similar solutions:

$$\hat{a}_{\mathrm{H}}(t) = \hat{a}_{\mathrm{H}}(0)\exp\{-i\omega_0 t\}, \quad \hat{a}_{\mathrm{H}}^{\dagger}(t) = \hat{a}_{\mathrm{H}}^{\dagger}(0)\exp\{+i\omega_0 t\}, \tag{4.359}$$

where $\hat{a}_{\mathrm{H}}(0)$ and $\hat{a}_{\mathrm{H}}^{\dagger}(0)$ are the Schrödinger-picture creation and annihilation operators used above.

Either of these approaches allows one to calculate any expectation values we want as the function of time. In particular, the average coordinate and momentum evolve exactly as in classics, while their fluctuations are time-independent:

$$\delta x = \frac{x_0}{\sqrt{2}} = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2}, \quad \delta p = \frac{m\omega_0 x_0}{\sqrt{2}} = \left(\frac{\hbar m\omega_0}{2}\right)^{1/2}. \tag{4.360}$$

In the quantum theory of measurements these expressions are known as the "standard quantum limit". Notice that their product

$$\delta x \delta p = \frac{\hbar}{2} \tag{4.361}$$

corresponds to the lower bound of the Heisenberg's uncertainty relation (242).

To show that this is not an occasional coincidence, let us calculate the Glauber state's wavefunction. Multiplying both sides of Eq. (448) by bra-vector of coordinate, and using definition (300) of the annihilation operator, we get

$$\frac{1}{\sqrt{2}x_0}\left\langle x\right|\left(\hat{x} + i\frac{\hat{p}}{m\omega_0}\right)\left|\alpha\right\rangle = \alpha\left\langle x\middle|\alpha\right\rangle. \tag{4.362}$$

Since $\langle x|$ is the vector of the eigenstate of the Hermitian operator $\hat{x}$, they may be swapped, while for the (local!) operator of momentum, we can use the same trick as in Eqs. (228)-(229). As a result, we get

$$\frac{1}{\sqrt{2}x_0}\left(\hat{x}\left\langle x\middle|\alpha\right\rangle + i\frac{\hat{p}}{m\omega_0}\left\langle x\middle|\alpha\right\rangle\right) = \alpha\left\langle x\middle|\alpha\right\rangle. \tag{4.363}$$

But $\langle x|\alpha\rangle$ is nothing else as the Glauber state's wavefunction Ψ_α, so that using x-representation for the coordinate and momentum operators, we get for it a first-order differential equation

$$\frac{1}{\sqrt{2}x_0}\left(x\Psi_\alpha + \frac{\hbar}{m\omega_0}\frac{\partial}{\partial x}\Psi_\alpha\right) = \alpha\Psi_\alpha. \tag{4.364}$$

Chasing Ψ_α and x into the opposite sides of the equation, and using the definition of α, we bring this equation to a form

$$\frac{\partial\Psi_\alpha}{\Psi_\alpha} = \frac{m\omega_0}{\hbar}\left[-x + \left(X + i\frac{P}{m\omega_0}\right)\right]\partial x \tag{4.365}$$

which allows for an easy integration, giving[49]

[49] Time participates in this function through its dependence on the classically-evolving $X(t)$ and $P(t)$.

$$\Psi_\alpha = C\exp\left\{-\frac{m\omega_0}{2\hbar}(x-X)^2 + i\frac{Px}{\hbar}\right\}. \tag{4.366}$$

In this expression we readily recognize a Gaussian wave packet centered to coordinate X, with average momentum P, and we know from Sec. 2.1 that such packets allow one to reach the minimum value of the uncertainty product.

Notice that such packet forms a Glauber state only if the packet's width is chosen in a very specific way. A natural question is: what happens if we form a Gaussian packet with a different width δx, say with $\delta x < x_0/\sqrt{2}$, and let it "go" (evolve). Such state is called *squeezed*; it is easy to get convinced that, just like the coherent state, it "rotates" clockwise with angular frequency ω_0 about the origin, keeping its relative "shape" (meaning, for example, the distribution of probability density to find the system in a specific point of the phase plane) intact –see Fig. 8. This means, in particular, that the coordinate uncertainty is an oscillating function of time, reaching its initial (minimum) value twice during each period. Such states, which may be conveniently formed, e.g., by parametric amplification, have important applications for the measurement of weak forces.[50]

4.11. Revisiting spherically-symmetric problems

One more blank spot we should fill has been left in our study of wave mechanics of spherically-3D symmetric systems (Sec. 3.5). Indeed, while the eigenfunctions describing axially-symmetric 2D systems, and the azimuthal component of the spherically-symmetric 3D systems, are very simple,

$$\psi_m = \frac{1}{\sqrt{2\pi}}e^{im\varphi}, \quad m = 0, \pm 1, \pm 2, ... \tag{4.367}$$

the polar components of the eigenfunctions in the latter case include the associate Legendre functions $P_n^l(\cos\theta)$ which may be expressed via elementary functions only indirectly. This makes all the calculations less than transparent and, in particular, do not allow a clear insight on the origin of a very simple eigenvalue spectrum (3.199). The bra-ket formalism, applied to the angular momentum operator, allows one to get such insight, and also produces a very convenient tool for many calculations involving spherically-symmetric potentials.

We start from using the correspondence principle to generalize the classical formula (3.156) for angular momentum to quantum-mechanical operators:

$$\hat{\mathbf{L}} = \hat{\mathbf{r}}\times\hat{\mathbf{p}} = \begin{vmatrix} \mathbf{n}_x & \mathbf{n}_y & \mathbf{n}_z \\ \hat{x} & \hat{y} & \hat{z} \\ \hat{p}_x & \hat{p}_y & \hat{p}_z \end{vmatrix}, \quad \text{i.e.,} \quad \hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \text{ etc.} \tag{4.368}$$

From this definition we can readily calculate all essential commutation relations, for example,

$$\left[\hat{L}_x, \hat{y}\right] = \left[\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{y}\right] = -\hat{z}\left[\hat{p}_y, \hat{y}\right] = i\hbar\hat{z}, \tag{4.369}$$

[50] For an early review see, e.g., D. Walls, *Nature* **306**, 141 (1983); spectacular measurements of squeezed states have been reported by G. Breitenbach *et al*., *Nature* **387**, 471 (1997).

etc. Using indices 1, 2, 3 to number the Cartesian component of vector operators $\hat{\vec{r}}$ and $\hat{\vec{p}}$ (e.g., $\hat{x} = \hat{r}_1$, etc.), the summary of these calculations may be presented as

$$\left[\hat{L}_j, \hat{r}_k\right] = i\hbar \hat{r}_l \varepsilon_{jkl}, \qquad \left[\hat{L}_j, \hat{p}_k\right] = i\hbar \hat{p}_l \varepsilon_{jkl}, \qquad \left[\hat{L}_j, \hat{L}_k\right] = i\hbar \hat{L}_l \varepsilon_{jkl}, \tag{4.370}$$

where ε_{jkl} is the well-known "*Levi-Civita symbol*" (or "*permutation symbol*") which is the natural generalization of the Kronecker delta to the case of three indices:

$$\varepsilon_{jkl} = \begin{cases} +1, & \text{if indices } jkl \text{ form any of the "correct" sequences: 123, or 231, or 312,} \\ -1, & \text{if the sequence } jkl \text{ is "wrong", such as 213, etc.,} \\ 0, & \text{if any two of the three indices coincide.} \end{cases} \tag{4.371}$$

Also introducing in the natural way a (scalar!) operator of L^2,

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \tag{4.372}$$

it is straightforward to prove that this operator commutes with each of the Cartesian components:

$$\left[\hat{L}^2, \hat{L}_j\right] = 0. \tag{4.373}$$

A result, the set of eigenfunctions of any component of momentum is a subset of all eigenfunctions of $\hat{L}^2$.

With a focus on the common set of eigenstates of operators $\hat{L}^2$ and $\hat{L}_z$, i.e. at the usual spherical functions (Sec. 3.5), we can introduce the so-called "*ladder operators*" (or "*raising*" and "*lowering*" operators)

$$\hat{L}_\pm \equiv \hat{L}_x \pm i\hat{L}_y. \tag{4.373}$$

(Notice a substantial similarity between this definition and Eqs. (300)-(301) which define the creation-annihilation operators.) It is straightforward to use these definition and the last of Eqs. (470) to show that

$$\left[\hat{L}_+, \hat{L}_-\right] = 2\hbar \hat{L}_z, \text{ and } \left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar \hat{L}_\pm; \tag{4.374}$$

another useful relation (see Problem 20) is

$$\hat{L}^2 = \hbar \hat{L}_z + \hat{L}_z^2 + \hat{L}_- \hat{L}_+. \tag{4.375}$$

Now let us apply the last of Eqs. (474), in the form

$$\hat{L}_z \hat{L}_\pm = \hat{L}_\pm \hat{L}_z \pm \hbar \hat{L}_\pm, \tag{4.376}$$

to the ket-vector $|l,m\rangle$ of a common eigenstate of $\hat{L}^2$ and $\hat{L}_z$:

$$\hat{L}_z \hat{L}_\pm |l,m\rangle = \hat{L}_\pm \hat{L}_z |l,m\rangle \pm \hbar \hat{L}_\pm |l,m\rangle. \tag{4.377}$$

From Sec. 3.5, we already know that the eigenvalues of operator $\hat{L}_z$ equal $\hbar m$ (where m is the same integer as in Eq. (467)), so that in the first term of the right-hand part,

$$\hat{L}_z|l,m\rangle = \hbar m|l,m\rangle. \tag{4.378}$$

With that, Eq. (477) may be recast as

$$\hat{L}_z\left(\hat{L}_\pm|l,m\rangle\right) = \hbar(m\pm 1)\left(\hat{L}_\pm|l,m\rangle\right) \tag{4.379}$$

In a spectacular similarity with Eqs. (313)-(314) for the harmonic oscillator, Eq. (450) means that states $\hat{L}_\pm|l,m\rangle$ are also the eigenstates of operator $\hat{L}_z$, but corresponding to its eigenvalues ($m \pm 1$). Thus the ladder operators act exactly as the creation-annihilation operator in the oscillator, moving the system up or down the ladder of eigenstates (Fig. 10).

The most significant difference is that now the ladder must end in both directions, because an infinite increase of $|m|$, with whatever sign, would cause the norm of eigenstates of operator

$$\hat{L}_x^2 + \hat{L}_y^2 = \hat{L}^2 - \hat{L}_z^2 \tag{4.380}$$

become negative, which cannot happen. Hence there should be two states on both ends of the ladder, $|l,m_{\max}\rangle$ and $|l,m_{\min}\rangle$, for whom

$$\hat{L}_+|l,m_{\max}\rangle = 0, \quad \hat{L}_-|l,m_{\min}\rangle = 0. \tag{4.381}$$

Due to the symmetry of the problem with respect to the replacement $m \to -m$, we should have $m_{\min} = -m_{\max}$; this is exactly the quantum number called l:

$$m_{\max} = -m_{\min} \equiv l. \tag{4.382}$$

(Earlier in this section we have used l just as a symbol for the designation of eigenstates of operator $\hat{L}^2$, not as a particular integer.)

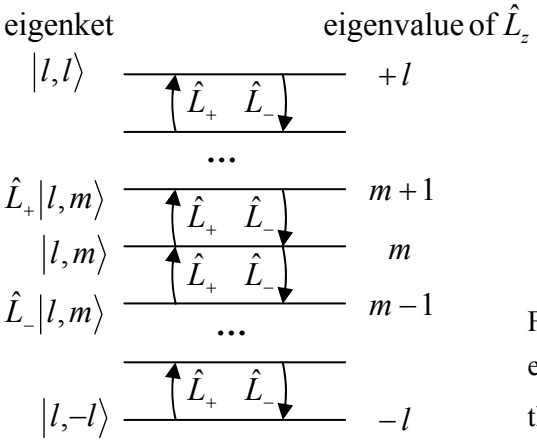

Fig. 4.10. The hierarchy ("ladder") of common eigenstates of operators $\hat{L}^2$ and $\hat{L}_z$. Arrows show the action of the ladder operators $\hat{L}_\pm$.

Now, applying the operator equality (475) to the top state $|l,m_{\max}\rangle = |l,l\rangle$, we get the equality

$$\hat{L}^2|l,l\rangle = \hbar\hat{L}_z|l,l\rangle + \hat{L}_z^2|l,l\rangle + \hat{L}_-\hat{L}_+|l,l\rangle = \hbar^2 l|l,l\rangle + \hbar^2 l^2|l,l\rangle + 0 = \hbar^2 l(l+1)|l,l\rangle \tag{4.383}$$

which shows that the eigenvalues of operator $\hat{L}^2$ are indeed equal to $\hbar^2 l(l+1)$, in agreement with our earlier brute force approach – see, e.g., Eq. (3.186). The deviation of this result from the maximum eigenvalue of operator $\hat{L}_z^2$, $\hbar^2 l^2$, may be viewed as the result of unavoidable fluctuations of the *x*- and *y*-components of the angular momentum, which give a finite positive contribution to L^2 even if the angular momentum vector is aligned in the best possible way with the *z*-axis.

4.12. Problems

4.1. Let α and β be two possible quantum states of the same system, and $\hat{A}$ be a legitimate linear operator. Which of the following expressions are meaningful in the bra-ket formalism?

(i) $\langle\alpha\rangle$ (ii) $\langle\alpha|\beta\rangle^2$ (iii) $|\alpha\rangle\langle\beta|$ (iv) $\langle\hat{A}|$

(v) $\langle\alpha|\hat{A}$ (vi) $\alpha|\hat{A}\rangle$ (vii) $|\alpha\rangle^2$ (viii) $\hat{A}^2$

4.2. Prove that for any linear operators $\hat{A}, \hat{B}, \hat{C}, \hat{D}$,

$$\left[\hat{A}\hat{B}, \hat{C}\hat{D}\right] = \hat{A}\left\{\hat{B}, \hat{C}\right\}\hat{D} - \hat{A}\hat{C}\left\{\hat{B}, \hat{D}\right\} + \left\{\hat{A}, \hat{C}\right\}\hat{D}\hat{B} - \hat{C}\left\{\hat{A}, \hat{D}\right\}\hat{B}.$$

4.3. Calculate all possible binary products $\sigma_j\sigma_k$ ($j, k = 1, 2, 3$) of the Pauli matrices

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and their commutators and anticommutators (defined similarly to those of the corresponding operators).

4.4. Trace (denoted as either Tr or Sp) of a square matrix is the sum of all diagonal elements. Prove that for arbitrary square matrices A and B,

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

Is each diagonal element $(AB)_{jj}$ necessarily equal to $(BA)_{jj}$?

4.5. Trace of an operator defined as the trace of its matrix. Prove that the trace of an arbitrary operator does not change at an arbitrary unitary transformation.

4.6. Prove that for any two full and orthonormal bases u_j, v_j of the same system,

$$\mathrm{Tr}\left(|u_j\rangle\langle v_k|\right) = \langle v_k|u_j\rangle.$$

4.7. Calculate $\langle \sigma_z \rangle$ in a quantum state with the following ket-vector:

$$|\alpha\rangle = \text{const} \times \left(|\uparrow\rangle + |\downarrow\rangle + |\rightarrow\rangle + |\leftarrow\rangle \right),$$

where $(\uparrow, \downarrow)$ and $(\rightarrow, \leftarrow)$ are eigenstates of the Pauli matrices $\boldsymbol{\sigma}_z$ and $\boldsymbol{\sigma}_x$, respectively. Check whether the solution you are giving is general.

4.8. At $t = 0$, the spin of an electron, whose interaction with an external field is described by Hamiltonian

$$\hat{H} = \mathbf{a} \cdot \hat{\boldsymbol{\sigma}} \equiv a_x \hat{\sigma}_x + a_y \hat{\sigma}_y + a_z \hat{\sigma}_z ,$$

(where $a_{x,y,z}$ are real and constant c-numbers, and operators $\hat{\sigma}_{x,y,z}$ in the z-basis are presented by the Pauli matrices $\boldsymbol{\sigma}_{x,y,z}$), was in state $\uparrow$, an eigenstate of operator $\hat{\sigma}_z$. Use the Schrödinger picture equations to calculate the time evolution of:

(i) the ket-vector $|\alpha\rangle$ of the system (in any stationary basis you like),

(ii) probabilities to find the system in states $\uparrow$ and $\downarrow$,

(iii) expectation values of all 3 spatial components ($\hat{S}_x$, etc.) of the spin vector operator $\hat{\mathbf{S}} = (\hbar/2)\hat{\boldsymbol{\sigma}}$, and

(iv) r.m.s. uncertainties (fluctuations) of these components.

Analyze and interpret the results for the simple case $a_y = a_z = 0$.

4.9. For the previous problem, use the Heisenberg picture equations to calculate the time evolution of:

(i) all three spatial components ($\hat{S}_x$, etc.) of the spin operator $\hat{\mathbf{S}}_{\mathrm{H}}(t)$,

(ii) expectation values of the spatial components.

Compare the latter results with that of Problem 8.

4.10. An electron is in a constant vertical field,

$$\hat{H} = \frac{\hbar\Omega}{2}\hat{\sigma}_z ,$$

but its spin's initial state corresponds to eigenvalue of the Hamiltonian of Problem 8. Use any approach you like to calculate the time evolution of the expectation values of the spin components. Interpret the results.

4.11. Prove the following identity

$$\left(\mathbf{a} \cdot \boldsymbol{\sigma}\right)^{2n} = a^{2n}\mathbf{I}$$

for the scalar product of the Pauli matrix vector $\boldsymbol{\sigma} = \mathbf{n}_x\sigma_x + \mathbf{n}_y\sigma_y + \mathbf{n}_z\sigma_z$ by any *c*-number vector **a**, where $n \geq 0$ is an integer.

4.12. A two-level system is in a quantum state described by ket-vector $|\alpha\rangle = \alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle$, with given (generally, complex) *c*-number coefficients $\alpha_{\uparrow\downarrow}$. Prove that we can always select a 3-component vector $\mathbf{a} = \{a_x, a_y, a_z\}$ of real *c*-numbers, such that $|\alpha\rangle$ is an eigenket of operator $\mathbf{a}\cdot\hat{\boldsymbol{\sigma}}$, where $\hat{\boldsymbol{\sigma}}$ is the Pauli matrix vector. Find all possible values of **a** satisfying this condition, and the second eigenket of operator $\mathbf{a}\cdot\hat{\boldsymbol{\sigma}}$.

4.13. Use the WKB approximation to calculate the transmission coefficient *T* for tunneling of a 2D particle with energy $E < U_0$ through a saddle-shaped potential "pass"

$$U(\mathbf{r}) = U_0\left(1 + \frac{xy}{a^2}\right),$$

where $U_0 > 0$ and *a* are real constants.

4.14. In the Heisenberg picture of quantum dynamics, calculate time evolution of:

(i) the creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ for a harmonic oscillator of mass *m* and frequency ω_0, and

(ii) their expectation values for a stationary ("Fock") state *n*.

4.15. For a 1D harmonic oscillator with mass *m* and frequency ω_0, calculate:

(i) all matrix elements $\langle n|\hat{x}^3|n'\rangle$, and

(ii) diagonal matrix elements $\langle n|\hat{x}^4|n\rangle$,

where *n* are Fock states.

4.16. Find the expectation value of energy, and time evolution of expectation values of the coordinate and momentum of a 1D harmonic oscillator, provided that in the initial moment ($t = 0$) it was in state

$$|\alpha\rangle = \frac{1}{\sqrt{2}}\left(|7\rangle + |8\rangle\right),$$

where $|n\rangle$ are its stationary (Fock) states with energy $E_n = \hbar\omega_0(n + 1/2)$.

4.17. Discussing the Glauber state properties in Sec. 10, we have used the following general statement: if

$$\left[\hat{A},\hat{B}\right]=\mu\hat{I},$$

where $\hat{A}$ and $\hat{B}$ are arbitrary operators, and μ is an arbitrary *c*-number, then

$$\exp\left\{\hat{A}\right\}\hat{B}\exp\left\{-\hat{A}\right\}=\hat{B}+\mu\hat{I}.$$

Prove the statement.

Hint: One (of many) ways is to expand operator

$$\hat{f}(\lambda)=\exp\left\{\lambda\hat{A}\right\}\hat{B}\exp\left\{-\lambda\hat{A}\right\}$$

into the Taylor series in *c*-number parameter λ, and then evaluate it at $\lambda = 1$.

4.18. Prove that:

(i) the uncertainties of coordinate and momentum in any Glauber state do not depend on time;

(ii) the product of the uncertainties corresponds to the minimum allowed by Heisenberg's uncertainty relation.

4.19. Prove the following formula from the propagator of the 1D harmonic oscillator:

$$G(x,t;x_0,t_0)=\left(\frac{m\omega_0}{2\pi i\hbar\sin[\omega_0(t-t_0)]}\right)^{1/2}\exp\left\{\frac{im\omega_0}{2\hbar\sin[\omega_0(t-t_0)]}\left[\left(x^2+x_0^2\right)\cos[\omega_0(t-t_0)]-2xx_0\right]\right\}.$$

Discuss the relation between this formula and the propagator of a free 1D particle.

4.20. Prove the following relations for operators of the angular momentum:

$$\hat{L}^2=\hat{L}_z^2+\hat{L}_+\hat{L}_--\hbar\hat{L}_z=\hat{L}_z^2+\hat{L}_-\hat{L}_++\hbar\hat{L}_z.$$

4.21. A particle is in a state α with an orbital wavefunction proportional to the spherical harmonic $Y_1^1(\theta,\varphi)$. Find the angular dependence of the states described by the following ket-vectors:

(i) $\hat{L}_x|\alpha\rangle$, (ii) $\hat{L}_y|\alpha\rangle$, (iii) $\hat{L}_z|\alpha\rangle$, and (iv) $\hat{L}^2|\alpha\rangle$.

