Chapter 4. Bra-ket Formalism

The objective of this chapter is a discussion of Dirac’s bra-ket formalism of quantum mechanics, which not only overcomes some inconveniences of wave mechanics, but also allows a natural description of such “internal” properties of particles as their spin. In the course of discussion of the formalism I will give several simple examples of its use, leaving more involved applications for the following chapters.

4.1. Motivation

We have seen that wave mechanics gives many results of primary importance. Moreover, it is fully (or mostly) sufficient for many applications, for example, for solid state electronics and device physics. However, in the course of our survey we have filed several grievances about this approach. Let me briefly summarize these complaints:

(i) Wave mechanics is focused on the spatial dependence of wavefunctions. On the other hand, our attempts to analyze the temporal evolution of quantum systems within this approach (beyond the trivial time behavior of the eigenfunctions, described by Eq. (1.61)), run into technical difficulties. For example, we could derive Eq. (2.159) describing time dynamics of the metastable state, or Eq. (2.185) describing quantum oscillations in coupled wells, only for the simplest potential profiles, though it is intuitively clear that these simple results should be common for all problem of this kind. Deriving the equations of such processes for arbitrary potential profiles is possible using perturbation theories (to be reviewed in Chapter 6), but that in the wave mechanics language they would require very bulky formulas.

(ii) The same is true concerning other issues that are conceptually addressable within wave mechanics, e.g., the Feynman path integral approach, description of coupling to environment, etc. Addressing them in wave mechanics would lead to formulas so bulky that I had (wisely :-) postponed them until we have got a more compact formalism on hand.

(iii) In the discussion of several key problems (for example the harmonic oscillator and spherically-symmetric potentials) we have run into rather complicated eigenfunctions coexisting with simple energy spectra - that infer some 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 of the coordinate and momentum are treated very unequally – see Eqs. (1.26b). However, some key expressions, e.g., for the fundamental eigenfunction of a free-particle,

\[
\exp\left\{i\frac{p \cdot r}{\hbar}\right\},
\]

or the harmonic oscillator’s Hamiltonian,

\[
\hat{H} = \frac{1}{2m} \hat{p}^2 + \frac{m\omega^2}{2} \hat{r}^2,
\]

invite a similar treatment of momentum and coordinate.
However, the strongest motivation for a more general formalism comes from wave mechanics’ conceptual incapability to describe elementary particles’ spins and other internal quantum degrees of freedom, such as quark flavors or lepton numbers. In this context, let us review the basic facts on spin (which is a very representative and experimentally the most accessible of all internal quantum numbers), to understand what a more general formalism should explain - as a minimum.

Figure 1 shows the conceptual scheme of the simplest spin-revealing experiment, first carried out by O. Stern and W. Gerlach in 1922. \(^1\) A collimated beam of electrons is passed through a gap between poles of a strong magnet, where the magnetic field \(\mathbf{B}\), whose orientation is taken for axis \(z\) in Fig. 1, is non-uniform, so that both \(B_z\) and \(d B_z/dz\) are not equal to zero. As a result, the beam splits into two parts of equal intensity.

This simplest experiment can be semi-quantitatively explained on classical, though somewhat phenomenological grounds by assuming that each electron has an intrinsic, permanent magnetic dipole moment \(m\). Indeed, classical electrodynamics \(^2\) tells us that the potential energy \(U\) of a magnetic dipole in an external magnetic field is equal to \((-m \cdot \mathbf{B})\), so that the force acting on the particle,

\[
F = -\nabla U = -\nabla(-m \cdot \mathbf{B}),
\]

has a nonvanishing vertical component

\[
F_z = -\frac{\partial}{\partial z}(-m_z \cdot B_z) = m_z \frac{\partial B_z}{\partial z}. \tag{4.4}
\]

Hence if we further postulate the existence of two possible, discrete values of \(m_z = \pm \mu\), this explains the Stern-Gerlach effect qualitatively, as a result of the incident electrons having a random sign, but similar magnitude of \(m_z\). A quantitative explanation of the beam splitting angle requires the magnitude of \(\mu\) to be equal (or close) to the so-called Bohr magneton \(^3\)

\[
\mu_B \equiv \frac{\hbar e}{2m_e} \approx 0.9274 \times 10^{-23} \text{ J/T}. \tag{4.5}
\]

As we will see below, this value cannot be explained by any internal motion of the electron, say its rotation about axis \(z\).

---

\(^1\) To my knowledge, the concept of spin as an internal rotation of a particle was first suggested by R. Kronig, then a 20-year-old student, in January 1925, a few months before two other students, G. Uhlenbeck and S. Goudsmit - to whom the idea is usually attributed. The concept was then accepted and developed quantitatively by W. Pauli.

\(^2\) See, e.g., EM Sec. 5.4, in particular Eq. (5.100).

\(^3\) A convenient mnemonic rule is that it is close to 1 K/T. In the Gaussian units, \(\mu_B = \hbar e/2m_e \approx 0.9274 \times 10^{-20}\).
Much more importantly, this semi-classical language cannot explain the results of the following set of multi-stage Stern-Gerlach experiments, shown in Fig. 2 - even qualitatively. In the first of the experiments, the electron beam is first passed through a magnetic field oriented (together with its gradient) along axis \( z \), just as in Fig. 1. Then one of the two resulting beams is absorbed (or otherwise removed from the setup), while the other one is passed through a similar but \( x \)-oriented field. The experiment shows that this beam is split again into two components of equal intensity. A classical explanation of this experiment would require a very unnatural suggestion that the initial electrons had random but discrete components of the magnetic moment simultaneously in two directions, \( z \) and \( x \).

However, even this assumption cannot explain the results of the three-stage Stern-Gerlach 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 the \( z \)-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, \( z \)-stage.

The only way to save the classical explanation here is to say that maybe, electrons somehow interact with the magnetic field, so that the \( x \)-polarized (non-absorbed) beam becomes spontaneously depolarized again somewhere between magnetic stages. 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 such depolarization happens.

We will see below that all these (and many more) results find a natural explanation in the matrix mechanics pioneered by W. Heisenberg, M. Born and P. Jordan in 1925. However, the matrix formalism is inconvenient for the solution of most problems discussed in Chapters 1-3, and for a time it was eclipsed by Schrödinger’s wave mechanics, which had been put forward just 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 proves their equivalence. Let me describe it.
4.2. States, state vectors, and linear operators

The basic notion of the general formulation of quantum mechanics is the \textit{quantum state} of a system.\footnote{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 the description of quantum systems much more complex than a single spinless particle that is a typical (though not the only possible) subject of wave mechanics.} To get some gut feeling of this notion, if a quantum state $\alpha$ of a particle may be adequately described by wave mechanics, this description is given by the corresponding wavefunction $\Psi_\alpha(r, t)$. Note, however, the state as such is \textit{not} a mathematical object (such as a function),\footnote{As was expressed nicely by A. Peres, one of pioneers of the quantum information theory, “quantum phenomena do not occur in the Hilbert space, they occur in a laboratory”.} and can participate in mathematical formulas only as a “pointer” – e.g., the index of function $\Psi_\alpha$. On the other hand, the wavefunction is \textit{not} a state, but a mathematical object (a complex function of space and time) giving a quantitative description of the state - just as the radius-vector as a function of time is a mathematical object describing the motion of a classical particle – see Fig. 3. Similarly, in the Dirac formalism a certain quantum state $\alpha$ is described by either of two mathematical objects, called the \textit{state vectors}: the ket-vector $|\alpha\rangle$ and bra-vector $\langle\alpha|$.\footnote{Terms \textit{bra} and \textit{ket} were suggested to reflect the fact that pair $\langle\beta|\alpha\rangle$ and $|\alpha\rangle$ may be considered as the set of parts of combination $\langle\beta|\alpha\rangle$ (see Eq. (11) below), which reminds an expression in the usual angle brackets.}

One should be cautious with the term “vector” here. Usual “geometric” vectors are defined in the usual geometric (say, Euclidean) space. In contrast, bra- and ket-vectors are defined in abstract \textit{Hilbert spaces} of a given system,\footnote{The Hilbert space of a given system is defined as the set of all its possible state vectors. As should be clear from this definition, it is not advisable to speak about a “Hilbert space of quantum states”.} and, despite certain similarities with the geometric vectors, are new mathematical objects, so that we need new rules for handling them. The primary rules are essentially postulates and are justified only the correct description/prediction of all experimental observations their corollaries. While there is a general consensus among physicists what the corollaries are, there are many possible ways to carve from them the basic postulate sets. Just as in Sec. 1.2, I will not try too hard to beat the number of the postulates to the smallest possible minimum, trying instead to keep their physical meaning transparent.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig4_3}
\caption{Particle’s state and its descriptions.}
\end{figure}

(i) \textbf{Ket-vectors}. Let us start with \textit{ket-vectors} - sometimes called just \textit{kets} for short. Perhaps the most important property of the vectors concerns their \textit{linear superposition}. Namely, if several ket-vectors $|\alpha_j\rangle$ describe possible states of a quantum system, then any linear combination (\textit{superposition})

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

\begin{tikzpicture}
\node [text=black, align=center] at (0,0) {\textbf{Linear superposition of ket-vectors}};
\end{tikzpicture}
where \( c_j \) are any (possibly complex) \( c \)-numbers, also describes a possible state of the same system. (One may say that vector \( |\alpha⟩ \) belongs to the same Hilbert space as all \( |\alpha_j⟩ \).) Actually, since ket-vectors are new mathematical objects, the exact meaning of the right-hand part of Eq. (6) becomes clear only after we have postulated the following rules of summation of these vectors,

\[
|\alpha⟩ + |\alpha'⟩ = |\alpha⟩ + |\alpha'⟩,
\]
and their multiplication by \( c \)-numbers:

\[
c_j |\alpha⟩ = |\alpha⟩ c_j.
\]

Note that in the set of wave mechanics postulates, statements parallel to (7) and (8) were unnecessary, because wavefunctions are the usual (albeit complex) functions of space and time, and we know from the usual algebra that such relations are valid.

As evident from Eq. (6), the complex coefficient \( c_j \) may be interpreted as the “weight” of state \( \alpha_j \) in the linear superposition \( \alpha \). One important particular case is \( c_j = 0 \), showing that state \( \alpha_j \) does not participate in the superposition \( \alpha \). By the way, the corresponding term of sum (6), i.e. product

\[
0|\alpha⟩,
\]
has a special name: the null-state vector. (It is important to avoid confusion between the null-state corresponding to vector (9), and the ground state of the system, which is frequently denoted by ket-vector \( |0⟩ \). In some sense, the null-state does not exist at all, while the ground state does – and frequently is the most important quantum state of the system.)

(ii) Bra-vectors and inner (“scalar”) products. Bra-vectors \( ⟨\alpha| \), which obey the rules similar to Eqs. (7) and (8), are not new, independent objects: if a ket-vector \( |\alpha⟩ \) is known, the corresponding bra-vector \( ⟨\alpha| \) describes the same state. In other words, there is a unique dual correspondence between \( |\alpha⟩ \) and \( ⟨\alpha| \), very similar (though not identical) to that between a wavefunction \( \Psi \) and its complex conjugate \( \Psi^* \). The correspondence between these vectors is described by the following rule: if a ket-vector of a linear superposition is described by Eq. (6), then the corresponding bra-vector is

\[
⟨\alpha| = \sum_j c^*_j ⟨\alpha| j = \sum_j ⟨\alpha| j c^*_j .
\]

The mathematical convenience of using two types of vectors, rather than just one, becomes clear from the notion of their inner product (also called the short bracket):

\[
⟨\beta|\alpha⟩ \equiv ⟨\beta|\alpha⟩ .
\]

This is a (generally, complex) scalar, whose main property is the linearity with respect to any of its component vectors. For example, if a linear superposition \( \alpha \) is described by the ket-vector (6), then

---

8 Mathematicians like to say that the ket- and bra-vectors of the same quantum system are defined in two isomorphic Hilbert spaces.

9 This is one of the differences of bra- and ket-vectors from the usual (geometrical) vectors whose scalar product is always a real scalar.
\[ \langle \beta | \alpha \rangle = \sum_j c_j \langle \beta | \alpha_j \rangle, \]  

while if Eq. (10) is true, then

\[ \langle \alpha | \beta \rangle = \sum_j c_j^* \langle \alpha_j | \beta \rangle. \]  

In plain English, \( c \)-numbers may be moved either into, or out of the inner products.

The second key property of the inner product is

\[ \langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle^*. \]  

It is compatible with Eq. (10); indeed, the complex conjugation of both parts of Eq. (12) gives:

\[ \langle \beta | \alpha \rangle^* = \sum_j c_j^* \langle \beta | \alpha_j \rangle^* = \sum_j c_j^* \langle \alpha_j | \beta \rangle = \langle \alpha | \beta \rangle. \]  

Finally, one more rule: the inner product of the bra- and ket-vectors describing the same state (called the norm squared) is real and non-negative,

\[ \|\alpha\|^2 = \langle \alpha | \alpha \rangle \geq 0. \]  

In order to give the reader some feeling about the meaning of this rule: we will show below that if state \( \alpha \) may be described by wavefunction \( \Psi_\alpha(r, t) \), then

\[ \langle \alpha | \alpha \rangle = \int \Psi_\alpha^* \Psi_\alpha d^3r \geq 0. \]  

Hence the role of the bra-ket is very similar to the complex conjugation of the wavefunction, and Eq. (10) emphasizes this similarity. (Note that, by convention, there is no conjugation sign in the bra-part of the inner product; its role is played by the angular bracket inversion.)

(iii) Operators. One more key notion of the Dirac formalism are quantum-mechanical linear operators. Just as for the operators 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 legitimate operator, then the following combination,

\[ \hat{A} |\alpha\rangle, \]  

is also a ket-vector describing a possible state of the system, i.e. a ket-vector in the same Hilbert space as the initial vector \( |\alpha\rangle \). As follows from the adjective “linear”, the main rules governing the operators is their linearity with respect to both any superposition of vectors:

\[ \hat{A} \left( \sum_j c_j |\alpha_j\rangle \right) = \sum_j c_j \hat{A} |\alpha_j\rangle, \]  

and any superposition of operators:

\[ \left( \sum_j c_j \hat{A}_j \right) |\alpha\rangle = \sum_j c_j \hat{A}_j |\alpha\rangle. \]
These rules are evidently similar to Eqs. (1.53)-(1.54) of wave mechanics.

The above rules imply that an operator “acts” on the ket-vector on its right; however, a combination of the type $\langle \alpha \mid \hat{A} \rangle$ 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 isomorphic to ket-vector (18) is

$$\langle \alpha \mid \hat{A}^\dagger \rangle. \quad (4.21)$$

This statement serves as the definition of the Hermitian conjugate (or “Hermitian adjoint”) $\hat{A}^\dagger$ of the initial operator $\hat{A}$. For an important class of operators, called the Hermitian operators, the conjugation is inconsequential, i.e. for them

$$\hat{A}^\dagger = \hat{A}. \quad (4.22)$$

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

To proceed further, we need an additional postulate, called the associative axiom of multiplication: into any legitimate bra-ket expression, not including an explicit summation, we may insert or remove parentheses (just in the ordinary product of scalars), meaning as usual that the operation inside the parentheses is performed first. The first two examples of this postulate are given by Eqs. (19) and (20), but the associative axiom is more general and says, for example:

$$\langle \beta \mid \hat{A} \langle \alpha \rangle \rangle = \langle \beta \mid \hat{A} \rangle \langle \alpha \rangle \equiv \langle \beta \hat{A} \rangle \langle \alpha \rangle, \quad (4.23)$$

This equality serves as the definition of the last form, called the long bracket (evidently, also a scalar), with an operator sandwiched between a bra-vector and a ket-vector. This definition, when combined with the definition of the Hermitian conjugate and Eq. (14), yields an important corollary:

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

which is most frequently rewritten as

$$\langle \alpha \hat{A} \rangle^* \beta = \langle \beta \hat{A}^\dagger \rangle \langle \alpha \rangle. \quad (4.25)$$

The associative axiom also enables to readily explore the following definition of one more, outer product of bra- and ket-vectors:

---

10 If we consider $c$-numbers as a particular type of operators, then 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 the Hermitian operator (22).

11 Here “legitimate” means “having a clear sense in the bra-ket formalism”. Some examples of “illegitimate” expressions: $\langle \alpha \rangle \hat{A}, \hat{A} \langle \alpha \rangle, \langle \alpha \rangle \beta, \langle \alpha \rangle \langle \beta \rangle$. Note, however, that the last two expressions may be legitimate if $\alpha$ and $\beta$ are states of different systems, i.e. if their state vectors belong to different Hilbert spaces. We will run into such tensor products of bra- and ket vectors (sometimes denoted, respectively, as $|\alpha\rangle \otimes |\beta\rangle$ and $\langle\alpha\rangle \otimes \langle\beta\rangle$) in Chapters 6-8.
In contrast to the inner product (12), which is a scalar, this mathematical construct is an operator. Indeed, the associative axiom allows us to remove parentheses in the following expression:

\[ \langle \beta | \langle \alpha | \gamma \rangle = \langle \beta | \langle \alpha | \gamma \rangle. \]  

(4.27)

But the last bra-ket is just a scalar; hence the mathematical object (26) acting on a ket-vector (in this case, \( | \gamma \rangle \)) gives a new ket-vector, which is the essence of operator’s action. Very similarly,

\[ \langle \delta | \langle \beta | \langle \alpha | \gamma \rangle = \langle \delta | \beta | \langle \alpha | \gamma \rangle \]  

(4.28)

- again a typical operator’s action on a bra-vector.

Now let us perform the following calculation. We may use the parentheses insertion into the bra-ket equality following from Eq. (14),

\[ \langle \gamma | \langle \alpha | \beta | \delta \rangle = \langle \langle \delta | \beta | \langle \alpha | \gamma \rangle \rangle^* \]  

(4.29)

to transform it to the following form:

\[ \langle \gamma | \langle \langle \alpha | \beta | \delta \rangle = \langle \langle \delta | \beta | \langle \alpha | \gamma \rangle \rangle^* \]  

(4.30)

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

\[ \langle \langle \alpha | \beta | \rangle = \langle \beta | \langle \alpha | \rangle \]  

(4.31)

This is the conjugate rule for outer products; it reminds rule (14) for inner products, but involves the Hermitian (rather than the usual complex) conjugation.

The associative axiom is also valid for the operator “multiplication”:

\[ \langle \hat{A} \hat{B} | \alpha \rangle = \hat{A} \langle \hat{B} | \alpha \rangle, \quad \langle \beta | \langle \hat{A} \hat{B} \rangle = \langle \beta | \hat{A} \hat{B} \rangle = \langle \beta | \hat{A} \rangle \hat{B}, \]  

(4.32)

showing that the action of an operator product on a state vector is nothing more than the sequential action of the operands. However, we have to be rather careful with the operator products; generally they do not commute: \( \hat{A} \hat{B} \neq \hat{B} \hat{A} \). This is why the commutator, the operator defined as

\[ [\hat{A}, \hat{B}] = \hat{A} \hat{B} - \hat{B} \hat{A}, \]  

(4.33)

is a very useful option. Another similar notion is the anticommutator:

\[ \{\hat{A}, \hat{B}\} = \hat{A} \hat{B} + \hat{B} \hat{A}. \]  

(4.34)

Finally, the bra-ket formalism broadly uses two special operators: the null operator \( \hat{0} \) defined by the following relations:

12 Another popular notation for the anticommutator is \( [\hat{A}, \hat{B}]^+ \); it will not be used in these notes.
for an arbitrary state $\alpha$, we may say that the null operator “kills” any state, turning it into the null-state. Another elementary operator is the identity operator, which is also defined by its action (or rather “inaction” :-) on an arbitrary state vector:

$$\hat{I}\left|\alpha}\right.\equiv\left|\alpha\right.\,, \quad \langle\alpha\left|\hat{I}\equiv\langle\alpha\right|\,.$$  

(4.36)

### 4.3. State 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 for specific quantum systems that feature at least one full and orthonormal set $\{u_j\}$ of states $u_j$, frequently called a basis. These terms mean that any state vector of the system may be represented as a unique sum of the type (6) or (10) over its basis vectors:

$$\left|\alpha\right.\equiv\sum_j \alpha_j \left|u_j\right.\,,$$

(4.37)

(so that, in particular, if $\alpha$ is one of the basis states, say $u_j$, then $\alpha_j = \delta_{jj'}$), and that

$$\langle u_j \left|u_j\right.\equiv\delta_{jj'}\,.$$  

(4.38)

For the systems that may be described by wave mechanics, examples of the full orthonormal bases are represented by any orthonormal set of eigenfunctions calculated in the previous 3 chapters – as the simplest example, see Eq. (1.76).

Due to the uniqueness of expansion (37), the full set of coefficients $\alpha_j$ gives a complete description of state $\alpha$ (in a fixed basis $\{u\}$), just as the usual Cartesian components $A_x$, $A_y$, and $A_z$ give a complete description of a usual geometric 3D vector $\mathbf{A}$ (in a fixed reference frame). Still, let me emphasize some differences between the quantum-mechanical bra- and ket-vectors and the usual geometric vectors:

(i) a basis set may have a large or even infinite number of states $u_j$, and

(ii) the expansion coefficients $\alpha_j$ may be complex.

With these reservations in mind, the analogy with geometric vectors may be pushed even further. Let us inner-multiply both parts of the first of Eqs. (37) by a bra-vector $\langle u_j \left|$ and then transform the relation using the linearity rules discussed in the previous section, and Eq. (38):

$$\langle u_j \left|\alpha\right.\equiv\langle u_j \left|\sum_j \alpha_j \left|u_j\right.\equiv\sum_j \alpha_j \langle u_j \left|u_j\right.\equiv\alpha_{j'}\,.$$  

(4.39)

Together with Eq. (14), this means that any of the expansion coefficients in Eq. (37) may be presented as an inner product:

$$\alpha_j = \langle u_j \left|\alpha\right.\,, \quad \alpha_j^* = \langle\alpha\left|u_j\right.\,;$$  

(4.40)

these relations are analogs of equalities $A_i = n_i \cdot A$ of the usual vector algebra. Using these important relations (which we will use on numerous occasions), expansions (37) may be rewritten as
\[ |\alpha\rangle = \sum_j |u_j\rangle \langle u_j| \alpha\rangle \equiv \sum_j \hat{\lambda}_j |\alpha\rangle, \quad \langle \alpha| = \sum_j \langle \alpha| u_j \rangle \langle u_j| \equiv \sum_j \langle \alpha| \hat{\lambda}_j, \quad (4.41) \]

A comparison of these relations with Eq. (26) shows that the outer product defined as
\[ \hat{\lambda}_j \equiv |u_j\rangle \langle u_j|, \quad (4.42) \]
is a legitimate linear operator. Such an operator, acting on any state vector of the type (37), singles out just one of its components, for example,
\[ \hat{\lambda}_j |\alpha\rangle = |u_j\rangle \langle u_j| \alpha\rangle = \alpha_j |u_j\rangle, \quad (4.43) \]
i.e. kills all components of the linear superposition but one. In the geometric analogy, such operator “projects” the state vector on its \(j^{th}\) “direction”, hence its name – the projection operator. Probably, the most important property of the projection operators, called the closure (or completeness) relation, immediately follows from Eq. (41): their sum over the full basis is equivalent to the identity operator:
\[ \sum_j |u_j\rangle \langle u_j| = \hat{I}. \quad (4.44) \]
This means in particular that we may insert the left-hand part of Eq. (44) into any bra-ket relation, at any place – the trick that we will use again and again.

Let us see how expansions (37) transform all the notions introduced in the last section, starting from the short bra-ket (11) (the inner product of two state vectors):
\[ \langle \beta| \alpha \rangle = \sum_{j,j'} \langle u_j| \beta_j^* \alpha_{j'} |u_{j'}\rangle = \sum_{j,j'} \beta_{j'}^* \alpha_{j'} \delta_{jj'} = \sum_j \beta_j^* \alpha_j. \quad (4.45) \]
Besides the complex conjugation, this expression is similar to the scalar product of the usual vectors. Now, let us explore the long bra-ket (23):
\[ \langle \beta| \hat{A}| \alpha \rangle = \sum_{j,j'} \beta_{j'}^* \langle u_{j'}| \hat{A}|u_j \rangle \alpha_{j'} \equiv \sum_{j,j'} \beta_{j'}^* A_{jj'} \alpha_{j'}, \quad (4.46) \]
Here, the last step uses a very important notion of matrix elements of the operator, defined as
\[ A_{jj'} \equiv \langle u_j| \hat{A}|u_{j'}\rangle. \quad (4.47) \]
As evident from Eq. (46), the full set of the matrix elements completely characterizes the operator, just as the full set of expansion coefficients (40) fully characterizes a 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 basis states \(u_j\) of the system under the consideration, i.e. the size of its Hilbert space.

As two simplest examples, all matrix elements of the null-operator, defined by Eqs. (35), are evidently equal to zero (in any basis), and hence it may be presented as a matrix of zeros (the null matrix):
\[ 0 \equiv \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}. \quad (4.48) \]
while for the identity operator \( \hat{I} \), defined by Eqs. (36), we readily get
\[
I_{jj'} = \langle u_j | \hat{I} | u_{j'} \rangle = \langle u_j | u_{j'} \rangle = \delta_{jj'},
\]
(4.49)
i.e. its matrix (called the identity matrix) is diagonal – also in any basis:
\[
I = \begin{pmatrix}
1 & 0 & \cdots \\
0 & 1 & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix},
\]
(4.50)

The convenience of the matrix language extends well beyond the presentation of particular operators. For example, let us use definition (47) to calculate matrix elements for a product of two operators:
\[
(AB)_{jj'} = \langle u_j | \hat{A} \hat{B} | u_{j'} \rangle.
\]
Here we can use Eq. (44) for the first (but not the last!) time, inserting the identity operator between the two operators, and then expressing it via a sum of projection operators:
\[
(AB)_{jj'} = \langle u_j | \hat{A} \hat{B} | u_{j'} \rangle = \sum_j \langle u_j | \hat{A} | u_{j'} \rangle \langle u_{j'} | \hat{B} | u_{j'} \rangle = \sum_j A_{jj'} B_{j'j'}.
\]
(4.52)

This result corresponds to the standard “row by column” rule of calculation of an arbitrary element of the matrix product
\[
AB = \begin{pmatrix}
A_{11} & A_{12} & \cdots \\
A_{21} & A_{22} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix}
\begin{pmatrix}
B_{11} & B_{12} & \cdots \\
B_{21} & B_{22} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix},
\]
(4.53)
Hence the product of operators may be presented (in a fixed basis!) by that of their matrices (in the same basis). This is so convenient that the same language is often used to present not only the long bracket,
\[
\langle \beta | \hat{A} | \alpha \rangle = \sum_j \beta_j^* A_{jj'} \alpha_{j'} = \begin{pmatrix}
\beta_1^* & \beta_2^* & \cdots
\end{pmatrix}
\begin{pmatrix}
A_{11} & A_{12} & \cdots \\
A_{21} & A_{22} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots
\end{pmatrix},
\]
(4.54)
but even the simpler short bracket:
\[
\langle \beta | \alpha \rangle = \sum_j \beta_j^* \alpha_j = \begin{pmatrix}
\beta_1^* & \beta_2^* & \cdots
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots
\end{pmatrix},
\]
(4.55)
although these equalities require 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. With that, the mapping of states and operators on matrices becomes completely general.

Now let us have a look at the outer product operator (26). Its matrix elements are just
\[
(\langle \alpha | \beta \rangle)_{jj'} = \langle u_j | \alpha \rangle \langle \beta | u_{j'} \rangle = \alpha_j \beta_{j'}^*. \tag{4.56}
\]

These are elements of a very special square matrix, whose filling requires the knowledge of just \(2N\) scalars (where \(N\) is the basis set size), rather than \(N^2\) scalars as for an arbitrary operator. However, a simple generalization of such outer product may present an arbitrary operator. Indeed, let us insert two identity operators (44), with different summation indices, on both sides of any operator:

\[
\hat{A} = \hat{I} \hat{A} \hat{I} = \left( \sum_j |u_j\rangle \langle u_j| \right) \hat{A} \left( \sum_{j'} |u_{j'}\rangle \langle u_{j'}| \right), \tag{4.57}
\]

and use the associative axiom to rewrite this expression as

\[
\hat{A} = \sum_{j,j'} |u_j\rangle \langle u_j| \hat{A} |u_{j'}\rangle \langle u_{j'}|. \tag{4.58}
\]

But the expression in the middle long bracket is just the matrix element (47), so that we may write

\[
\hat{A} = \sum_{j,j'} |u_j\rangle \langle u_j| A_{jj'} |u_{j'}\rangle \langle u_{j'}|. \tag{4.59}
\]

The reader has to agree that this formula, which is a natural generalization of Eq. (44), is extremely elegant. Also note the following parallel: if we consider the matrix element definition (47) as some sort of analog of Eq. (40), then Eq. (59) 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 level lower – from state vector products to “bare” state vectors resulting from operator’s action upon a given state. For example, let us use Eq. (59) to present the ket-vector (18) as

\[
|\alpha'\rangle \equiv \hat{A} |\alpha\rangle = \left( \sum_{j,j'} |u_j\rangle A_{jj'} |u_{j'}\rangle \right) |\alpha\rangle = \sum_{j,j'} |u_j\rangle A_{jj'} \langle u_{j'}| |\alpha\rangle. \tag{4.60}
\]

According to Eq. (40), the last short bracket is just \(\alpha_{j'}\), so that

\[
|\alpha'\rangle = \sum_{j,j'} |u_j\rangle A_{jj'} \alpha_{j'} = \sum_j \left( \sum_{j'} A_{jj'} \alpha_{j'} \right) |u_j\rangle \tag{4.61}
\]

But expression in middle parentheses is just the coefficient \(\alpha'_{j'}\) of expansion (37) of the resulting ket-vector (60) in the same basis, so that

\[
\alpha'_{j'} = \sum_{j'} A_{jj'} \alpha_{j'}. \tag{4.62}
\]

This result corresponds to the usual rule of multiplication of a matrix by a column, so that we may represent any ket-vector by its column matrix, with the operator action looking like

\[
\begin{pmatrix}
\alpha'_1 \\
\alpha'_2 \\
\vdots
\end{pmatrix} =
\begin{pmatrix}
A_{11} & A_{12} & \ldots & \alpha_1 \\
A_{21} & A_{22} & \ldots & \alpha_2 \\
\vdots & \vdots & \ddots & \vdots
\end{pmatrix}
\tag{4.63}
\]

Absolutely similarly, the operator action on the bra-vector (21), represented by its row-matrix, is
By the way, Eq. (64) naturally raises the following question: what are the elements of the matrix in its right-hand part, or more exactly, what is the relation between the matrix elements of an operator and its Hermitian conjugate? The simplest way to get an answer is to use Eq. (25) with two arbitrary states (say, \( u_j \) and \( u'_j \)) of the same basis in the role of \( \alpha \) and \( \beta \). Together with the orthonormality relation (38), this immediately gives

\[
\left( \hat{A}^\dagger \right)_{jj'} = \left( A^*_{jj'} \right).
\]

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

\[
A_{jj'} = A^*_{jj'},
\]

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

\[
A_{jj} = A^*_{jj}, \quad \text{i.e.} \quad \text{Im} A_{jj} = 0.
\]

(Matrix (50) evidently satisfies Eq. (66), so that the identity operator is Hermitian.)

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

\[
\hat{A} | a_j \rangle = A_j | a_j \rangle.
\]

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

\[
A_j = A^*_j, \quad \text{for} \quad j = 1, 2, ..., N,
\]

---

13 For the sake of formula compactness, below I will use the shorthand notation in which the operands of this equality are just \( A^\dagger_{jj'} \) and \( A^*_{jj} \). I believe that it leaves little chance for confusion, because the Hermitian conjugation sign \( \dagger \) may pertain only to an operator (or its matrix), while the complex conjugation sign \( * \) to a scalar – say a matrix element.

14 This equation should look familiar to the reader – see the stationary Schrödinger equation (1.60), which was the focus of our studies in the first three chapters. We will see soon that that equation is just a particular (coordinate) representation of Eq. (66) for the Hamiltonian as the operator of energy.

15 The reciprocal statement is also true: if all eigenvalues of an operator are real, it is Hermitian (in any basis). This statement may be readily proved by applying Eq. (93) below to the case when \( A_{kk} = \delta_{kk} \), with \( A^*_k = A_k \).
while the eigenstates corresponding to different eigenvalues are orthogonal:

$$\langle a_j | a_{j'} \rangle = 0, \quad \text{if} \quad A_j \neq A_{j'}.$$  \hfill (4.70)

The proof of both statements is surprisingly simple. Let us inner-multiply both sides of Eq. (68) by bra-vector $\langle a_j' |$. In the right-hand part of the result, the eigenvalue $A_j$, as a $c$-number, may be taken out of the bra-ket, giving

$$\langle a_j' | \hat{A} | a_j \rangle = A_j \langle a_j' | a_j \rangle.$$ \hfill (4.71)

This equality should hold for any pair of eigenstates, so that we may swap the indices in Eq. (71), and complex-conjugate the result:

$$\langle a_j | \hat{A} | a_{j'} \rangle^* = A_{j'}^* \langle a_j | a_{j'} \rangle^*.$$ \hfill (4.72)

Now using Eqs. (14) and (25), together with the Hermitian operator definition (22), we may transform Eq. (72) to the following form:

$$\langle a_j | \hat{A} | a_{j'} \rangle = A_{j'}^* \langle a_j | a_{j'} \rangle.$$ \hfill (4.73)

Subtracting this equation from Eq. (71), we get

$$0 = \left( A_j - A_{j'}^* \right) \langle a_j | a_{j'} \rangle.$$ \hfill (4.74)

There are two possibilities to satisfy this equation. If indices $j$ and $j'$ are equal (denote the same eigenstate), 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 of $\hat{A}$, 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. (70) is valid.

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

### 4.4. Change of basis and matrix 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. (52), (54), (55) 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 the matrices 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 reference frame orientation (Fig. 4), and very frequently it is convenient to use two or more different basis sets for the same system.

With this motivation, let us study what happens if we change from one basis, $\{ u \}$, to another one, $\{ v \}$ - both full and orthonormal. First of all, let us prove that for each such pair of bases, there exists such an operator $\hat{U}$ that, first,
and, second,

\[ \hat{U} \hat{U}^\dagger = \hat{U}^\dagger \hat{U} = \hat{I}. \]  

(4.76)

(Due to the last property,\(^{16}\) \(\hat{U}\) is called a \textit{unitary operator}, and Eq. (75), a \textit{unitary transformation}.)

\[ \sum_{j'} |j'\rangle \langle j'| = \sum_{j'} v_{j'}^* u_{j'} = \sum_{j} v_{j}^* \delta_{j} = |v_{j}\rangle , \]  

(4.77)

- an evident generalization of Eq. (44). Then

\[ \hat{U} |u_{j}\rangle = \sum_{j'} |v_{j'}\rangle \langle u_{j'}| = \sum_{j} |v_{j}\rangle \delta_{j} = |v_{j}\rangle , \]  

(4.78)

so that Eq. (75) has been proved. Now, applying Eq. (31) to each term of sum (77), we get

\[ \hat{U}^\dagger = \sum_{j'} |u_{j'}\rangle \langle v_{j'}| , \]  

(4.79)

so that

\[ \hat{U} \hat{U}^\dagger = \sum_{j,j'} |v_{j}\rangle \langle u_{j'}| v_{j}\rangle = \sum_{j,j'} |v_{j}\rangle \delta_{j} |v_{j}\rangle = \sum_{j} |v_{j}\rangle |v_{j}\rangle . \]  

(4.80)

But according to the closure relation (44), the last expression is just the identity operator, q.e.d.\(^{17}\) (The proof of the second equality in Eq. (76) is absolutely similar.)

As a by-product of our proof, we have also got another important expression (79). It implies, in particular, that while, according to Eq. (77), operator \(\hat{U}\) performs the transform from the “old” basis \(u_{j}\) to the “new” basis \(v_{j}\), its Hermitian adjoint \(\hat{U}^\dagger\) performs the reciprocal unitary transform:

\[ \hat{U}^\dagger |v_{j}\rangle = \sum_{j'} |u_{j'}\rangle \delta_{j} = |u_{j}\rangle . \]  

(4.81)

---

\(^{16}\) An alternative way to express Eq. (76) is to write \(\hat{U}^\dagger = \hat{U}^{-1}\), but I will try to avoid this language.

\(^{17}\) Quod erat demonstrandum (Lat.) – what needed to be proved.
Now, let us see how do the matrix elements of the unitary transform operators look like.
Generally, as was stated above, operator’s elements depend on the basis we calculate them in, so we
should be careful - initially. For example, let us calculate the elements in basis \( \{ u \} \): 

\[
U_{j'j} \big|_{\text{in } u} = \langle u_j | \hat{U} | u_{j'} \rangle = \langle u_j \left| \sum_k |v_k \rangle \langle u_k | \right| u_{j'} \rangle = \langle u_j | v_{j'} \rangle.
\]

(4.82)

Now performing a similar calculation in basis \( \{ v \} \), we get 

\[
U_{j'j} \big|_{\text{in } v} = \langle v_j | \hat{U} | v_{j'} \rangle = \langle v_j \left| \sum_k |v_k \rangle \langle u_k | \right| v_{j'} \rangle = \langle u_j | v_{j'} \rangle.
\]

(4.83)

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

\[
U^\dagger_{j'j} \big|_{\text{in } u} = U^\dagger_{j'j} \big|_{\text{in } v} = \langle v_j | u_{j'} \rangle.
\]

(4.84)

These expressions may be used, first of all, to rewrite Eq. (75) in a more direct form. Applying
the first of Eqs. (41) to state \( v_{j'} \) of the “new” basis, we get 

\[
\begin{align*}
| v_{j'} \rangle &= \sum_j |u_j \rangle \langle u_j | v_{j'} \rangle = \sum_j U_{j'j} | u_j \rangle .
\end{align*}
\]

(4.85)

Similarly, the reciprocal transform is 

\[
\begin{align*}
| u_{j'} \rangle &= \sum_j |v_j \rangle \langle v_j | u_{j'} \rangle = \sum_j U^\dagger_{j'j} | v_j \rangle .
\end{align*}
\]

(4.86)

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

Next, we may use Eqs. (83), (84) to express the effect of the unitary transform on expansion
coefficients (37) of vectors of an arbitrary state \( \alpha \). In the “old” basis \( \{ u \} \), they are given by Eq. (40).
Similarly, in the “new” basis \( \{ v \} \),

\[
\alpha_j \big|_{\text{in } v} = \langle v_j | \alpha \rangle.
\]

(4.87)

Again inserting the identity operator in the form of closure (44), with internal index \( j' \), and then using
Eq. (84), we get 

\[
\begin{align*}
\alpha_j \big|_{\text{in } v} &= \langle v_j \left| \sum_{j'} |u_{j'} \rangle \langle u_{j'} | \right| \alpha \rangle = \sum_{j'} \langle v_j | u_{j'} \rangle \langle u_{j'} | \alpha \rangle = \sum_{j'} U^\dagger_{j'j} \langle u_{j'} | \alpha \rangle = \sum_{j'} U^\dagger_{j'j} \alpha_j \big|_{\text{in } u}.
\end{align*}
\]

(4.88)

The reciprocal transform is (of course) performed by matrix elements of operator \( \hat{U} \) :

\[
\begin{align*}
\alpha_j \big|_{\text{in } u} &= \sum_{j'} U_{j'j} \alpha_j \big|_{\text{in } v}.
\end{align*}
\]

(4.89)

Both structurally and philosophically, these expressions are similar to the transformation of
components of a usual vector at coordinate frame rotation. For example, in two dimensions (Fig. 4):
\[
\begin{pmatrix}
\alpha_x \\
\alpha_y
\end{pmatrix} = \begin{pmatrix}
\cos \varphi & \sin \varphi \\
-\sin \varphi & \cos \varphi
\end{pmatrix}
\begin{pmatrix}
\alpha_x \\
\alpha_y
\end{pmatrix}.
\]

(In this analogy, the equality to 1 of the determinant of the rotation matrix in Eq. (90) corresponds to the unitary property (76) of the unitary transform operators.) Please pay attention here: while the transform (75) from the “old” basis \{u\} to the “new” basis \{v\} is performed by the unitary operator, the change (88) of a state vectors components at this transformation requires its Hermitian conjugate. Actually, this is also natural from the point of view of the geometric analog of the unitary transform (Fig. 4): if the “new” reference frame \{x’, y’\} is obtained by a counterclockwise rotation of the “old” frame \{x, y\} by some angle \(\varphi\), for the observer rotating with the frame, vector \(\alpha\) (which is itself unchanged) rotates clockwise. Due to the analogy between expressions (88) and (89) on one hand, and our old friend Eq. (62) on the other hand, it is tempting to skip indices in our new results by writing

\[
|\alpha\rangle_{in,v} = \hat{U}^\dagger |\alpha\rangle_{in,u}, \quad |\alpha\rangle_{in,u} = \hat{U} |\alpha\rangle_{in,v}.
\]

Since matrix elements of \(\hat{U}\) and \(\hat{U}^\dagger\) do not depend on basis, such language is not too bad; still, the symbolic Eq. (91) should not be confused with 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:

\[
A_{ij}\big|_{in,v} \equiv \langle v_j | \hat{A} | v_j \rangle = \langle v_j | \left( \sum_k |u_k\rangle \langle u_k| \right) \hat{A} \left( \sum_k |u_k\rangle \langle u_k| \right) \rangle |v_j\rangle = \sum_{k,k'} U_{jk}^\dagger A_{kk'} |u_k\rangle \langle u_k| U_{kj} |v_j\rangle;
\]

absolutely similarly, we can get

\[
A_{ij}\big|_{in,u} \equiv \sum_{k,k'} U_{jk} A_{kk'} |v_k\rangle \langle v_k| U_{kj}^\dagger |i_j\rangle.
\]

In the spirit of Eq. (91), we may present these results symbolically as well, in a compact bra-ket form:

\[
\hat{A}|_{in,v} = \hat{U}^\dagger \hat{A}|_{in,u} \hat{U}, \quad \hat{A}|_{in,u} = \hat{U} \hat{A}|_{in,v} \hat{U}^\dagger.
\]

As a sanity check, let us apply this result to the identity operator:

\[
|\hat{I}|_{in,v} = \left( \hat{U}^\dagger \hat{I} \hat{U} \right)|_{in,u} = \left( \hat{U}^\dagger \hat{U} \right)|_{in,v} = |\hat{I}|_{in,u}
\]

- as it should be. One more invariant of the basis change is the trace of any operator, defined as the sum of the diagonal terms of its matrix in a certain basis:

\[
\text{Tr} \hat{A} \equiv \text{Tr} A \equiv \sum_j A_{jj}.
\]

The (easy) proof of this fact, using the relations we have already discussed, is left for reader’s exercise.

So far, I have implied that both state bases \{u\} and \{v\} are known, and the natural question is where does this information comes from in quantum mechanics of actual physical systems. To get a partial answer to this question, let us return to Eq. (68) that defines eigenstates and eigenvalues of an
operator. Let us assume that the eigenstates \( a_j \) of a certain operator \( \hat{A} \) form a full and orthonormal set, and find the matrix elements of the operator in the basis of these states. For that, it is sufficient to inner-multiply both sides of Eq. (68), written for index \( j' \), by the bra-vector of an arbitrary state \( a_j \) of the same set:

\[
\langle a_j | \hat{A} | a_{j'} \rangle = \langle a_j | A_{j'} | a_{j'} \rangle.
\]  
(4.97)

The left-hand part is just the matrix element \( A_{jj'} \), 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:

\[
A_{jj'} = A_{j'} \delta_{jj'}.
\]  
(4.98)

In particular, in the eigenstate basis (but not necessarily in an arbitrary basis!), \( A_{jj} \) 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,\(^\text{18} \) i.e. finding the basis in which the corresponding operator acquires the diagonal form (98); then the diagonal elements are the eigenvalues, and the basis itself is the desirable set of eigenstates.

Let us modify the above calculation by inner-multiplying Eq. (68) by a bra-vector of a different basis – say, the one, denoted \( \{ u \} \), in which we know the matrix elements \( A_{jj'} \). The multiplication gives

\[
\langle u_k | \hat{A} | a_j \rangle = \langle u_k | A_{j} | a_j \rangle.
\]  
(4.99)

In the left-hand part we can (as usual :-) insert the identity operator, between the operator and the ket-vector, and then use the closure relation (44), while in the right-hand part, we can move the eigenvalue \( A_{j} \) out of the bra-ket, and then insert a summation over a new index, compensating it with the proper Kronecker delta symbol:

\[
\langle u_k | \hat{A} \sum_{k'} | u_{k'} \rangle \langle u_{k'} | a_j \rangle = A_{j} \sum_{k'} \langle u_{k'} | a_j \rangle \delta_{kk'}.
\]  
(4.100)

Moving out the sign of summation over \( k' \), and using definition (47) of the matrix elements, we get

\[
\sum_{k'} (A_{kk'} - A_{j} \delta_{kk'}) \langle u_{k'} | a_j \rangle = 0.
\]  
(4.101)

But the set of such equalities, for all \( N \) possible values of index \( k \), is just a system of linear, homogeneous equations for unknown \( c \)-numbers \( \langle u_k | a_j \rangle \). But according to Eqs. (82)-(84), these numbers are nothing else than the matrix elements \( U_{kj} \) of a unitary matrix providing the required transformation from the initial basis \( \{ u \} \) to the basis \( \{ a \} \) that diagonalizes matrix A. The system may be presented in the matrix form:

\[
\begin{pmatrix}
A_{11} - A_{j} & A_{12} & \cdots \\
A_{21} & A_{22} - A_{j} & \cdots \\
\vdots & \vdots & \ddots
\end{pmatrix}
\begin{pmatrix}
U_{1j} \\
U_{2j} \\
\vdots
\end{pmatrix} = 0,
\]  
(4.102)

\(^{18}\) Note that expression “matrix diagonalization” is a common and convenient, but dangerous jargon. (A matrix is just a matrix, an ordered set of \( c \)-numbers, and cannot be diagonalized.) It is OK to use this jargon if you remember clearly what it actually means – see the definition above.
and the usual condition of its consistency,

\[
\begin{vmatrix}
A_{11} - A_j & A_{12} & \cdots \\
A_{21} & A_{22} - A_j & \cdots \\
\vdots & \vdots & \ddots
\end{vmatrix} = 0,
\]

(4.103)

plays the role of the characteristic equation of the system. This equation has \( N \) roots \( A_j \); plugging each of them back into system (102), we can use it to find \( N \) matrix elements \( U_{kj} (k = 1, 2, \ldots N) \) corresponding to this particular eigenvalue. 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):

\[
\langle a_j | a_j \rangle = \sum_k \langle a_j | u_k \rangle \langle u_k | a_j \rangle = \sum_k |U_{kj}|^2 = 1,
\]

(4.104)

for each \( j \). This normalization completes the diagonalization.\(^{19}\)

Now (at last!) I can give the reader some examples. As a simple but very important case, let us diagonalize the operators described (in a certain 2-function basis \{\( u \)\}) by the so-called Pauli matrices

\[
\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

(4.105)

Though introduced by a physicist, with a specific purpose to describe electron’s spin, these matrices have a general mathematical significance, because together with the 2×2 identity matrix \( I \), they provide a full, linearly-independent 2×2 basis - meaning that an arbitrary 2×2 matrix may be presented as

\[
\begin{pmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{pmatrix} = a_0 I + a_x \sigma_x + a_y \sigma_y + a_z \sigma_z,
\]

(4.106)

with a unique set of 4 coefficients \( a \).

Let us start with diagonalizing matrix \( \sigma_x \). For it, the characteristic equation (103) is evidently

\[
\begin{vmatrix}
-A_j & 1 \\
1 & -A_j
\end{vmatrix} = 0,
\]

(4.107)

and has two roots, \( A_{1,2} = \pm 1 \). (Again, the numbering is arbitrary!) The reader may readily check that the eigenvalues of matrices \( \sigma_y \) and \( \sigma_z \) are similar. However, the eigenvectors of the operators corresponding to all these matrices are different. To find them for \( \sigma_x \), let us plug its first eigenvalue, \( A_1 = +1 \), back into equations (101), written for this particular case:

\[
-\langle u_1 | a_1 \rangle + \langle u_2 | a_1 \rangle = 0,
\]

\[
\langle u_1 | a_1 \rangle - \langle u_2 | a_1 \rangle = 0.
\]

(4.108)

\(^{19}\) A possible slight complication here are degenerate cases when characteristic equation gives certain equal eigenvalues corresponding to different eigenvectors. In this case the requirement of the mutual orthogonality of these states should be additionally enforced.
The equations are compatible (of course, because the used eigenvalue $A_1 = +1$ satisfies the characteristic equation), and any of them gives

$$\langle u_1 | a_1 \rangle = \langle u_2 | a_1 \rangle, \quad \text{i.e.} \quad U_{11} = U_{21}. \quad (4.109)$$

With that, the normalization condition (104) yields

$$|U_{11}|^2 = |U_{21}|^2 = \frac{1}{2}. \quad (4.110)$$

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 $\varphi$, it is convenient to keep the coefficients real, for example taking $\varphi = 0$, i.e. to get

$$U_{11} = U_{21} = \frac{1}{\sqrt{2}}. \quad (4.111)$$

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

$$U_{12} = -U_{22}, \quad (4.112)$$

so that the whole unitary matrix for diagonalization of the operator corresponding to $\sigma_x$ is\(^{20}\)

$$U_x = U_x^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (4.113)$$

For what follows, it will be convenient to have this result expressed in the ket-relation form – see Eqs. (85)-(86):

$$|a_1\rangle = U_{11} |u_1\rangle + U_{21} |u_2\rangle = \frac{1}{\sqrt{2}} (|u_1\rangle + |u_2\rangle), \quad |a_2\rangle = U_{12} |u_1\rangle + U_{22} |u_2\rangle = \frac{1}{\sqrt{2}} (|u_1\rangle - |u_2\rangle), \quad (4.114)$$

$$|u_1\rangle = U_{11}^\dagger |a_1\rangle + U_{21}^\dagger |a_2\rangle = \frac{1}{\sqrt{2}} (|a_1\rangle + |a_2\rangle), \quad |u_2\rangle = U_{12}^\dagger |a_1\rangle + U_{22}^\dagger |a_2\rangle = \frac{1}{\sqrt{2}} (|a_1\rangle - |a_2\rangle). \quad (4.115)$$

These results are already sufficient to understand the Stern-Gerlach experiments described in Sec. 1 - with two additional postulates. The first of them is that particle’s interaction with external magnetic field may be described by the following vector operator of the dipole magnetic moment:\(^{21}\)

$$\hat{m} = \gamma \hat{S}, \quad (4.116)$$

where the coefficient $\gamma$, specific for every particle type, is called the gyromagnetic ratio,$^{22}$ and $\hat{S}$ is the vector operator of spin. For the so-called spin-$\frac{1}{2}$ particles (including the electron), this operator may be represented, in the so-called $z$-basis, by the following 3D vector of the Pauli matrices (105):

---

\(^{20}\) Note that though this particular unitary matrix is Hermitian, this is not true for an arbitrary choice of phases $\varphi$.

\(^{21}\) This is the key point in the electron’s spin description, developed by W. Pauli in 1925-1927.

\(^{22}\) For an electron, with its negative charge $q = -e$, the gyromagnetic ratio is negative: $\gamma_e = -g_e e/2m_e$, where $g_e \approx 2$ is the dimensionless $g$-factor. Due to quantum electrodynamics effects, the factor is slightly higher than 2; $g_e = 2(1 + \alpha/2\pi + \ldots) \approx 2.002319304\ldots$, where $\alpha \equiv e^2/4\pi\varepsilon_0\hbar c \equiv (E_0/m_e c^2)^{1/2} \approx 1/137$ is the fine structure constant. (The origin of its name will be clear from the discussion in Sec. 6.3.)
S = \frac{\hbar}{2} \left( n_x \sigma_x + n_y \sigma_y + n_z \sigma_z \right) = \frac{\hbar}{2} \left( \begin{array}{c} n_z \ n_x - i n_y \\ n_y + i n_x \\ -n_z \end{array} \right), \tag{4.117}

and \( n_{x,y,z} \) are the usual Cartesian unit vectors in 3D space. (In the quantum-mechanics sense, they are just \( c \)-numbers, or rather “\( c \)-vectors”.) The \( z \)-basis, in which Eq. (177) is valid, is defined as an orthonormal basis of two states, frequently denoted \( \uparrow \) and \( \downarrow \), in which the \( z \)-component of the vector operator of spin is diagonal, with eigenvalues \( +h/2 \) and \( -h/2 \). Note that we do not “understand” what exactly these states are, but loosely associate them with a certain internal rotation of the electron about \( z \)-axis, with either positive or negative angular momentum component \( S_z \). However, any attempt to use such classical interpretation for quantitative predictions runs into fundamental difficulties – see Sec. 5.7 below.

The second new postulate describes the general relation between the bra-ket formalism and experiment. Namely, in quantum mechanics, each real observable \( A \) is represented by a Hermitian operator \( \hat{A} = \hat{A}^\dagger \), and a result of its measurement in a quantum state \( \alpha \), described by a linear superposition of the eigenstates \( a_j \) of the operator,

\[
|\alpha\rangle = \sum_j \alpha_j |a_j\rangle, \quad \text{with} \quad \alpha_j = \langle a_j | \alpha \rangle, \tag{4.118}
\]

may be only one of corresponding eigenvalues \( A_j \). If state (118) and all eigenstates \( a_j \) are normalized to unity,

\[
\langle \alpha | \alpha \rangle = 1, \quad \langle a_j | a_j \rangle = 1, \tag{4.119}
\]

then the probability of outcome \( A_j \) is

\[
W_j = |\alpha_j|^2 = \alpha_j^* \alpha_j = \langle \alpha | a_j \rangle \langle a_j | \alpha \rangle, \tag{4.120}
\]

This relation is evidently a generalization of Eq. (1. 22) in wave mechanics. As a sanity check, let us assume that the set of eigenstates \( a_j \) is full, and calculate the sum of all the probabilities:

\[
\sum_j W_j = \sum_j \langle \alpha | a_j \rangle \langle a_j | \alpha \rangle = \langle \alpha | \alpha \rangle = 1. \tag{4.121}
\]

Now returning to the Stern-Gerlach experiment, conceptually the description of the first (\( z \)-oriented) experiment shown in Fig. 1 is the hardest for us, because the statistical ensemble describing the unpolarized electron beam at its input is mixed (“incoherent”), and cannot be described by a pure
(“coherent”) superposition of the type (6) that have been the subject of our studies so far. (We will
discuss the mixed ensembles in Chapter 7.) However, it is intuitively clear that its results, and in
particular Eq. (6), are compatible with the description of its two output beams as sets of electrons in pure
states $\uparrow$ and $\downarrow$, respectively. The absorber following that first stage (Fig. 2) just takes all spin-down
electrons out of the picture, producing an output beam of polarized electrons in a pure $\uparrow$ state. For such
beam, probabilities (120) are $W_{\uparrow} = 1$ and $W_{\downarrow} = 0$. This is certainly compatible with the result of the
“control” experiment shown on the bottom panel of Fig. 2: the repeated SG ($z$) stage does not split such
a beam, keeping the probabilities the same.

Now let us discuss the double Stern-Gerlach experiment shown on the top panel of Fig. 2. For
that, let us present the $z$-polarized beam in another basis of two states (I will denote them as $\rightarrow$ and $\leftarrow$)
in which, by definition, the matrix of operator $\hat{S}_z$ is diagonal. But this is exactly the set we called $u_{1,2}$ in
the $\sigma$, matrix diagonalization problem solved above. On the other hand, states $\uparrow$ and $\downarrow$ are exactly what
we called $u_{1,2}$ in that problem, because in this basis, matrices $\sigma_z$ and hence $S_z$ are diagonal. Hence, in
application to the electron spin problem, we may rewrite Eqs. (114)-(115) as

$$
|\rightarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle), \quad |\leftarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle) \quad (4.122)
$$
$$
|\uparrow\rangle = \frac{1}{\sqrt{2}} (|\rightarrow\rangle + |\leftarrow\rangle), \quad |\downarrow\rangle = \frac{1}{\sqrt{2}} (|\rightarrow\rangle - |\leftarrow\rangle) \quad (4.123)
$$

Currently, for us the first of Eqs. (123) 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 $S_x = +h/2$ and $S_x = -h/2$. Notice that the beams have equal probability amplitude moduli, so that
according to Eq. (122), the split beams $\rightarrow$ and $\leftarrow$ have equal intensities, in accordance with experiment.
(The minus sign before the second ket-vector is of no consequence here, though it may have an impact
on outcome of other experiments – for example if the $\rightarrow$ and $\leftarrow$ beams are brought together again.)

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 electrons in state $\rightarrow$ are passed to the final, SG ($z$), stage. But according
to the first of Eqs. (122), this state may be presented as a (coherent) linear superposition of the $\uparrow$ and $\downarrow$
states, with equal amplitudes. The stage separates these two states into separate beams, with equal
probabilities $W_{\uparrow} = W_{\downarrow} = \frac{1}{2}$ to find an electron in each of them, thus explaining the experimental results.

To conclude our discussion of the multistage Stern-Gerlach experiment, let me note that though
it cannot be explained in terms of wave mechanics (which operates with scalar de Broglie waves), it has
an analogy in classical theories of vector fields, such as the classical electrodynamics. Let a plane
electromagnetic wave propagate perpendicular to the plane of drawing in Fig. 5, and pass through linear
polarizer 1. Similarly to the initial SG ($z$) stages (including the following absorbers) shown in Fig. 2, the
polarizer produces a wave linearly polarized in one direction – the vertical direction in Fig. 5. Its electric
field vector has no horizontal component, as may be revealed by wave’s full absorption in a
perpendicular polarizer 3. However, let us pass the wave through polarizer 2 first. In this case, the
output wave does acquire a horizontal component, as can be, again, revealed by passing it through
polarizer 3. If angles between polarization direction 1 and 2, and between 2 and 3, are both equal $\pi/4$,
each polarizer reduces the wave amplitude by a factor of $\sqrt{2}$, and hence intensity by a factor of 2, exactly
like in the multistage SG experiment, with polarizer 2 playing the role of the SG (x) stage. The “only” difference is that the necessary angle is $\pi/4$, rather than by $\pi/2$ for the Stern-Gerlach experiment. In quantum electrodynamics (see Chapter 9 below), which confirms the classical predictions for this experiment, this difference is explained by that between the integer spin of the electromagnetic field quanta, photons, and the half-integer spin of electrons.

Fig. 4.5. Light polarization sequence similar to the 3-stage Stern-Gerlach experiment shown on the middle panel of Fig. 2.

4.5. Observables: Expectation values and uncertainties

After this particular (and hopefully very inspiring) example, let us discuss the general relation between the Dirac formalism and experiment in more detail. The expectation value of an observable over any statistical ensemble (not necessarily coherent) may be always calculated using the general rule (1.37). For the particular case of a coherent superposition (118), we can combine that definition with Eq. (120) and the second of Eqs. (118), and then use Eqs. (59) and (98) to write

$$\langle A \rangle = \sum_j A_j W_j = \sum_j \alpha_j^* A_j \alpha_j = \sum_j \langle \alpha | a_j \rangle A_j \langle a_j | \alpha \rangle = \sum_{j,j'} \langle \alpha | a_j \rangle \langle a_j | \hat{A} | a_j \rangle \langle a_j | \alpha \rangle.$$  (4.124)

Now using the completeness relation (44) twice, with indices $j$ and $j'$, we arrive at a very simple and important formula

$$\langle A \rangle = \langle \alpha | \hat{A} | \alpha \rangle.$$  (4.125)

This is a clear analog of the wave-mechanics formula (1.23) – and as we will see in the next chapter, may be used to derive it. A huge advantage of Eq. (125) is that it does not explicitly involve the eigenvector set of the corresponding operator, and allows the calculation to be performed in any convenient basis.28

For example, let us consider an arbitrary state $\alpha$ of spin-$1/2$, and calculate the expectation values of its components. The calculations are easiest in the z-basis, because we know the operators of the components in that basis – see Eq. (117). Representing the ket- and bra-vectors of our state as linear superpositions of vectors of the basis states $\uparrow$ and $\downarrow$,

$$|\alpha\rangle = \alpha_\uparrow |\uparrow\rangle + \alpha_\downarrow |\downarrow\rangle,$$

$$\langle \alpha | = \langle \uparrow | \alpha_\uparrow^* + \langle \downarrow | \alpha_\downarrow^*.$$  (4.126)

27 This equality reveals the full beauty of Dirac’s notation. Indeed, initially the quantum-mechanical brackets just reminded the angular brackets used for statistical averaging. Now we see that in this particular (but most important) case, the angular brackets of these two types may be indeed equal to each other!

28 Note that Eq. (120) may be rewritten in the form similar to Eq. (125): $W_j = \langle \alpha | \hat{A}_j | \alpha \rangle$, where $\hat{A}_j \equiv | a_j \rangle \langle a_j |$ is the operator (42) of projection upon the $j^{th}$ eigenstate $a_j$. 
and plugging these expressions to Eq. (125) written for observable $S_z$, we get

$$
\langle S_z \rangle = \left( \langle \uparrow | \alpha_\uparrow^* + \langle \downarrow | \alpha_\downarrow^* \rangle S_z \langle \alpha_\uparrow \uparrow \rangle + \alpha_\downarrow \downarrow \rangle \right)
$$

$$
= \alpha_\uparrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_z \uparrow \rangle + \alpha_\downarrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_z \downarrow \rangle + \alpha_\uparrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_z \uparrow \rangle + \alpha_\downarrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_z \downarrow \rangle.
$$

(4.127)

Now there are two equivalent ways (both very simple :-) to calculate the bra-kets in this expression. The first one is to represent each of them in the matrix form in the $z$-basis, in which bra- and ket-vectors of states $\uparrow$ and $\downarrow$ are, respectively, matrix-rows $(1, 0)$ and $(0, 1)$, or the similar matrix-columns. Another (perhaps more elegant) way is to use the general Eq. (59), for the $z$-basis, to write

$$
\langle \uparrow | \hat{S}_z \uparrow \rangle = i \frac{\hbar}{2} \left( \langle \uparrow | \hat{S}_z \uparrow \rangle - \langle \downarrow | \hat{S}_z \downarrow \rangle \right),
$$

$$
\langle \downarrow | \hat{S}_z \downarrow \rangle = -i \frac{\hbar}{2} \left( \langle \uparrow | \hat{S}_z \uparrow \rangle - \langle \downarrow | \hat{S}_z \downarrow \rangle \right).
$$

(4.128)

For our particular calculation, we may plug the last of these expressions into Eq. (127), and to use the orthonormality conditions (119):

$$
\langle \uparrow | \uparrow \rangle = \langle \downarrow | \downarrow \rangle = 1, \quad \langle \uparrow | \downarrow \rangle = \langle \downarrow | \uparrow \rangle = 0.
$$

(4.129)

Both calculations give (of course) the same result:

$$
\langle S_z \rangle = \frac{\hbar}{2} \left( \alpha_\uparrow \alpha_\uparrow^* - \alpha_\downarrow \alpha_\downarrow^* \right).
$$

(4.130)

This particular result might be also obtained using Eq. (120) for probabilities $W_\uparrow = \alpha_\uparrow \alpha_\uparrow^*$ and $W_\downarrow = \alpha_\downarrow \alpha_\downarrow^*$:

$$
\langle S_z \rangle = W_\uparrow \left( \frac{\hbar}{2} \right) + W_\downarrow \left( -\frac{\hbar}{2} \right) = \alpha_\uparrow \alpha_\uparrow^* \left( \frac{\hbar}{2} \right) + \alpha_\downarrow \alpha_\downarrow^* \left( -\frac{\hbar}{2} \right).
$$

(4.131)

The formal way (127), based on using Eq. (125), has, however, an advantage of being applicable, without any change, to finding the observables whose operators are not diagonal in the $z$-basis, as well. In particular, absolutely similar calculations give

$$
\langle S_x \rangle = \alpha_\uparrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_x \uparrow \rangle + \alpha_\downarrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_x \uparrow \rangle + \alpha_\uparrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_x \uparrow \rangle + \alpha_\downarrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_x \uparrow \rangle = \frac{\hbar}{2} \left( \alpha_\uparrow \alpha_\uparrow^* + \alpha_\downarrow \alpha_\downarrow^* \right),
$$

(4.132)

$$
\langle S_y \rangle = \alpha_\uparrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_y \uparrow \rangle + \alpha_\downarrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_y \uparrow \rangle + \alpha_\uparrow \alpha_\downarrow^* \langle \downarrow | \hat{S}_y \uparrow \rangle + \alpha_\downarrow \alpha_\uparrow^* \langle \uparrow | \hat{S}_y \uparrow \rangle = \frac{i}{2} \left( \alpha_\uparrow \alpha_\uparrow^* - \alpha_\downarrow \alpha_\downarrow^* \right),
$$

(4.133)

Similarly, we can express, via the same coefficients $\alpha_\uparrow$ and $\alpha_\downarrow$, the r.m.s. fluctuations of all spin components. For example, let us have a good look at the spin state $\uparrow$. According to Eq. (126), in this state $\alpha_\uparrow = 1$ and $\alpha_\downarrow = 0$, so that Eqs. (130)-(133) yield:

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

(4.134)

Now let us use the same Eq. (125) to calculate the spin component uncertainties. According to Eqs. (105) and (117), operators of spin component squared are equal to $(\hbar/2)^2 \hat{I}$, so that the general Eq. (1.33) yields
While Eqs. (134) and (135a) are compatible with the classical notion of the spin being “definitely in the \( \uparrow \) state”, this correspondence should not be overstretched to the interpretation of this state as a certain \( (z) \) orientation of electron’s magnetic moment \( \mathbf{m} \), because such classical picture cannot explain Eqs. (135b) and (135c). The best (but still imprecise!) classical image I can offer is the magnetic moment \( \mathbf{m} \) oriented, on the average, in the \( z \)-direction, but still having \( x \)- and \( y \)-components strongly “wobbling” about their zero average values.

It is straightforward to verify that in the \( x \)-polarized and \( y \)-polarized states the situation is similar, with the corresponding change of indices. Thus, in neither state may all 3 components of the spin have exact values. Let me show that this is not just an occasional fact, but reflects the most profound property of quantum mechanics, the uncertainty relations. Consider 2 observables, \( A \) and \( B \), that may be measured in the same quantum state. There are two possibilities here. If operators corresponding to the observables commute,

\[
[\hat{A}, \hat{B}] = 0,
\]

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

\[
\langle a_j | [\hat{A}, \hat{B}] | a_j \rangle = \langle a_j | \hat{A} \hat{B} | a_j \rangle - \langle a_j | \hat{B} \hat{A} | a_j \rangle = 0.
\]

In the first bra-ket of the middle expression, let us act by operator \( \hat{A} \) on the bra-vector, while in the second one, on the ket-vector. According to Eq. (68), such action turns operators into the corresponding eigenvalues, so that we get

\[
A_j \langle a_j | \hat{B} | a_j \rangle - A_j \langle a_j | \hat{B} | a_j \rangle = (A_j - A_j) \langle a_j | \hat{B} | a_j \rangle = 0.
\]

This means that if eigenstates of operator \( \hat{A} \) are non-degenerate (i.e. \( A_j \neq A_j \); if \( j \neq j' \)), the matrix of operator \( \hat{B} \) has to be diagonal in basis \( a_j \), i.e., the eigenstate sets of operators \( \hat{A} \) and \( \hat{B} \) coincide. Such pairs of observables, that share their eigenstates, are called compatible. For example, in wave mechanics of a particle, momentum (1.26) and the kinetic energy (1.27) are compatible, sharing eigenfunctions (1.29). Now we see that this is not occasional, because each Cartesian component of the kinetic energy is proportional to the square of the corresponding component of the momentum, and any operator commutes with an arbitrary power of itself:

\[
[\hat{A}, \hat{A}^n] = [\hat{A}, \hat{A} \hat{A} \ldots \hat{A}] = \hat{A} \hat{A} \hat{A} 
\]

(4.139)
Now, what if operators $\hat{A}$ and $\hat{B}$ do not commute? Then the following general uncertainty relation is valid:\(^{29}\)

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

The proof of Eq. (140) may be divided into two steps, the first of which proves the so-called Schwartz inequality:\(^{30}\)

\[
\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq \left| \langle \alpha | \beta \rangle \right|^2 .
\]  (4.141)

The proof may be started by using postulate (16) - that the norm of any legitimate state of the system cannot be negative. Let us apply this postulate to the state with the following ket-vector:

\[
| \delta \rangle \equiv | \alpha \rangle - \frac{\langle \beta | \alpha \rangle}{\langle \beta | \beta \rangle} | \beta \rangle ,
\]  (4.142)

where $\alpha$ and $\beta$ are possible, non-null states of the system, so that the denominator in Eq. (142) is not equal to zero. For this case, Eq. (16) gives

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

Opening the parentheses, we get

\[
\langle \alpha | \alpha \rangle - \frac{\langle \alpha | \beta \rangle}{\langle \beta | \beta \rangle} \langle \beta | \alpha \rangle - \frac{\langle \beta | \alpha \rangle}{\langle \beta | \beta \rangle} \langle \alpha | \beta \rangle + \frac{\langle \alpha | \beta \rangle \langle \beta | \alpha \rangle}{\langle \beta | \beta \rangle} \geq 0 .
\]  (4.144)

After the cancellation of one inner product $\langle \beta | \beta \rangle$ in the numerator and denominator of the last term, it cancels with the 2nd (or 3rd) term, proving the Schwartz inequality (141).

Now let us apply this inequality to states

\[
| \alpha \rangle \equiv \hat{A} | \gamma \rangle \quad \text{and} \quad | \beta \rangle \equiv \hat{B} | \gamma \rangle ,
\]  (4.145)

where, in both relations, $\gamma$ is the same (but otherwise arbitrary) possible state of the system, and the deviations operators are defined similarly to observable deviations (see Sec. 1.2), for example,

\[
\hat{A} \equiv \hat{A} - \langle \hat{A} \rangle .
\]  (4.146)

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

\[
\langle \gamma | \hat{A}^2 | \gamma \rangle \langle \gamma | \hat{B}^2 | \gamma \rangle \geq \left| \langle \gamma | \hat{A} \hat{B} | \gamma \rangle \right|^2 .
\]  (4.147)

\(^{29}\) Note that both sides of Eq. (140) are state-specific; the uncertainty relation statement is that this inequality should be valid for any possible quantum state of the system.

\(^{30}\) This inequality is the quantum-mechanical analog of the usual vector algebra result $\hat{A} \hat{B} \geq |\hat{A} \cdot \hat{B}|^2$.  

Since state $\gamma$ is arbitrary, we may use Eq. (125) to rewrite this relation as an operator inequality:

$$\delta A \delta B \geq \left| \frac{\langle \hat{A} \hat{B} \rangle}{\frac{1}{2}} \right|.$$  \hspace{1cm} (4.148)

Actually, this is already an uncertainty relation, even “better” (stronger) than its standard form (140); moreover, it is more convenient in some cases. In order to proceed to Eq. (140), we need a couple more steps. First, let us notice that the operator product in Eq. (148) may be recast as

$$\hat{A} \hat{B} = \frac{1}{2} \left( \hat{A} \hat{B} - i \hat{C} \right), \quad \text{where} \quad \hat{C} = i \left[ \hat{A}, \hat{B} \right].$$  \hspace{1cm} (4.149)

Any anticommutator of Hermitian operators, including that in Eq. (149), is a Hermitian operator, and its eigenvalues are purely real, so that its expectation value (in any state) is also purely real. On the other hand, the commutator part of Eq. (149) is just

$$\hat{C} = i \left[ \hat{A}, \hat{B} \right] = i \left( \hat{A} - \langle \hat{A} \rangle \right) \left( \hat{B} - \langle \hat{B} \rangle \right) - i \left( \hat{B} - \langle \hat{B} \rangle \right) \left( \hat{A} - \langle \hat{A} \rangle \right) = i \left( \hat{A} \hat{B} - \hat{B} \hat{A} \right) = i \left[ \hat{A}, \hat{B} \right].$$  \hspace{1cm} (4.150)

Second, according to Eqs. (52) and (65), the Hermitian conjugate of any product of Hermitian operators $\hat{A}$ and $\hat{B}$ is just the product of swapped operators. Using the fact, we may write

$$\hat{C} = \left( i \left[ \hat{A}, \hat{B} \right] \right)^\dagger = -i \left( \hat{A} \hat{B} \right)^\dagger + i \left( \hat{B} \hat{A} \right)^\dagger = -i \hat{A} \hat{B} = i \hat{B} \hat{A} = i \left[ \hat{A}, \hat{B} \right] = \hat{C},$$  \hspace{1cm} (4.151)

so that operator $\hat{C}$ is also Hermitian, i.e. its eigenvalues are also real, and thus its average is purely real as well. As a result, the square of the average of the operator product (149) may be presented as

$$\langle \hat{A} \hat{B} \rangle^2 = \left( \frac{1}{2} \left( \hat{A} \hat{B} \right) \right)^2 + \frac{1}{2} \left( \hat{C} \right)^2.$$  \hspace{1cm} (4.152)

Since the first term in the right-hand part of this equality cannot be negative,

$$\langle \hat{A} \hat{B} \rangle^2 \geq \left( \frac{1}{2} \hat{C} \right)^2 = \left( \frac{1}{2} \left[ \hat{A}, \hat{B} \right] \right)^2,$$  \hspace{1cm} (4.153)

and we can continue Eq. (148) as

$$\delta A \delta B \geq \left| \frac{\langle \hat{A} \hat{B} \rangle}{\frac{1}{2}} \right| \geq \frac{1}{2} \left| \left[ \hat{A}, \hat{B} \right] \right|,$$  \hspace{1cm} (4.154)

thus proving Eq. (140).

For the particular case of operators $\hat{x}$ and $\hat{p}_x$ (or a similar pair of operators for another Cartesian coordinate), we can readily combine Eq. (140) with Eq. (2.14b) and to prove the original Heisenberg’s uncertainty relation (2.13). For the spin-1/2 operators defined by Eq. (117), it is straightforward (and highly recommended to the reader) to show that

$$\left[ \hat{S}_x, \hat{S}_y \right] = i \hbar \hat{S}_z,$$  \hspace{1cm} (4.155)

with similar relations for other pairs of indices taken in the “correct” order (from $x$ to $y$ to $z$ to $x$, etc.). As a result, the uncertainty relations (140) for spin-1/2 particles, notably including electrons, are
\[
\delta S_x \delta S_y \geq \frac{\hbar}{2} |\langle S_z \rangle|, \text{ etc.} \tag{4.156}
\]

In particular, in the \(\uparrow\) state, the right-hand part of this relation equals \((\hbar/2)^2\), and neither of the uncertainties \(\delta S_x, \delta S_y\) can equal zero. As a reminder, our direct calculation earlier in this section has shown that each of these uncertainties is equal to \(\hbar/2\), i.e. their product equals to the lowest value allowed by the uncertainty relation (156). In this aspect, the spin-polarized states are similar to the Gaussian wave packets studied in Sec. 2.2.

### 4.6. Quantum dynamics: Three pictures

So far in this chapter, I shied away from the discussion of system dynamics, implying that the bra- and ket-vectors of the system are their “snapshots” at a certain instant \(t\). Now we are sufficiently prepared to examine their time dependence. One of the most beautiful features of quantum mechanics is that the time evolution may be described using either of three alternative “pictures”, giving exactly the same final results for expectation values of all observables.

From the standpoint of our wave mechanics experience, the *Schrödinger picture* is the most natural. In this picture, the operators corresponding to time-independent observables (e.g., to the Hamiltonian function \(H\) of an isolated system) are also constant, while the bra- and ket-vectors of the quantum state of the system evolve in time as

\[
|\alpha(t)\rangle = |\alpha(t_0)\rangle |\hat{u}(t,t_0)\rangle, \quad |\alpha(t)\rangle = \hat{u}(t,t_0) |\alpha(t_0)\rangle, \tag{4.157}
\]

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

\[
i\hbar \dot{\hat{u}} = \hat{H}\hat{u}, \tag{4.158}
\]

where \(\hat{H}\) is the Hamiltonian operator of the system (that is always Hermitian, \(\hat{H}^\dagger = \hat{H}\)), and the dot means the differentiation is over argument \(t\), but not \(t_0\). While this equation is a very natural replacement of the wave-mechanical equation (1.25), and is also frequently called the *Schrödinger equation*,\(^{31}\) it still should be considered as a new, more general postulate, which finds its final justification (as it is usual in physics) in the agreement between its corollaries with experiment - more exactly, in having not a single credible contradiction with experiment.

Starting the discussion of Eqs. (157)-(158), let us first consider the case of a system described by a time-independent Hamiltonian, whose eigenstates \(a_n\) and eigenvalues \(E_n\) obey Eq. (68),\(^{32}\)

\[
\hat{H}|a_n\rangle = E_n|a_n\rangle, \tag{4.159}
\]

and hence are also time-independent. (Similarly to the wavefunctions \(\psi_n\) defined by Eq. (1.60), \(a_n\) are called the *stationary states* of the system.) Let us use Eqs. (157)-(159) to calculate the law of time evolution of the expansion coefficients \(\alpha_n\), defined by Eq. (118), in the stationary state basis:

---

31 Moreover, we will be able to derive Eq. (1.25) from Eq. (154) – see Sec. 5.2.

32 Here I intentionally use index \(n\) rather than \(j\), to emphasize the special role played by the stationary eigenstates \(a_n\) in quantum dynamics.
\[ \dot{\alpha}_n(t) = \frac{d}{dt} \langle a_n | \alpha(t) \rangle = \frac{d}{dt} \langle a_n | \hat{u}(t,t_0) \rangle \alpha(t_0) \rangle = \langle a_n | \hat{u}(t,t_0) | \alpha(t_0) \rangle \]

\[ = \langle a_n | \frac{1}{i \hbar} \hat{\mathcal{H}}(t,t_0) | \alpha(t_0) \rangle \rangle = \frac{E_n}{i \hbar} \langle a_n | \hat{u}(t,t_0) | \alpha(t_0) \rangle \rangle = \frac{E_n}{i \hbar} \langle a_n | \alpha(t) \rangle = -\frac{i}{\hbar} E_n \alpha_n. \]  

(4.160)

This is the same simple equation as Eq. (1.59), and its integration yields a similar result – cf. Eq. (1.61), just with the initial time \( t_0 \) rather than 0:

\[ \alpha_n(t) = \alpha_n(t_0) \exp \left( -\frac{i}{\hbar} E_n (t-t_0) \right). \]  

(4.161)

In order to illustrate how does this result work, let us consider spin-\( \frac{1}{2} \) dynamics in a time-independent, uniform external magnetic field \( \mathcal{B} \), taking its direction for axis \( z \). To construct the system’s Hamiltonian, we may apply the correspondence principle to the classical expression for the energy of a magnetic moment \( \mathbf{m} \) in the external magnetic field \( \mathcal{B} \), 33

\[ U = -\mathbf{m} \cdot \mathbf{B}. \]  

(4.162)

In quantum mechanics, the operator corresponding to the moment \( \mathbf{m} \) is given by Eq. (116) (suggested by W. Pauli), so that the spin-field interaction is described by the so-called Pauli Hamiltonian:

\[ \hat{\mathcal{H}} = -\hat{\mathbf{m}} \cdot \mathbf{B} = -\gamma \hat{\mathbf{S}} \cdot \mathbf{B} = -\gamma \mathcal{B} \hat{S}_z, \]  

(4.163)

where \( \hat{S}_z \) is the operator of the \( z \)-component of electron’s spin. According to Eq. (117), in the \( z \)-basis of states \( \uparrow \) and \( \downarrow \), the matrix of operator (163) is

\[ \mathbf{H} = -\frac{\gamma \hbar}{2} \sigma_z = \frac{\hbar \Omega}{2} \sigma_z, \text{ with } \Omega \equiv -\gamma \mathcal{B}. \]  

(4.164)

The constant \( \Omega \) so defined coincides with the classical frequency of the precession of a symmetric top, with an angular momentum \( \mathbf{S} \) and magnetic moment \( \mathbf{m} = \gamma \mathbf{S} \), about axis \( z \), induced by external torque \( \mathbf{\tau} = \mathbf{m} \times \mathbf{B} \): 34

\[ \Omega = -\frac{\mathbf{\tau}}{\mathbf{S}} = -\frac{m \mathcal{B}}{S} = -\gamma \mathcal{B}. \]  

(4.165a)

For an electron, with its negative gyromagnetic ratio \( \gamma_e = -g_e e/2m_e \), neglecting the minor difference between factors \( g_e \) and 2, we get

\[ \Omega = \frac{e}{m_e} \mathcal{B}, \]  

(4.165b)

i.e. the frequency’s magnitude coincides with that of the cyclotron frequency \( \omega_c \) – see Eq. (3.48).

In order to apply the general Eq. (161), 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 direction of axis \( z \), the Hamiltonian matrix is already diagonal:

---

33 See, e.g., EM Eq. (5.100). As a reminder, we have already used this expression for the derivation of Eq. (3).
34 See, e.g., CM Sec. 6.5, in particular Eq. (6.72), and EM Sec. 5.5, in particular Eq. (5.114) and its discussion.
\[
H = -\frac{\hbar \Omega}{2} \sigma_z = \frac{\hbar \Omega}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},
\]
(4.166)

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

\[
E_\uparrow = +\frac{\hbar \Omega}{2} \quad \text{and} \quad E_\downarrow = -\frac{\hbar \Omega}{2}.
\]
(4.167)

(Note that their difference,

\[
\Delta E \equiv |E_\uparrow - E_\downarrow| = \hbar|\Omega| = \hbar|\gamma B|,
\]
(4.168)
corresponds to the classical energy \(2|mB|\) of flipping the magnetic dipole with moment \(m = \gamma \hbar / 2\), oriented along the direction of field \(B\). 35) With that, Eq. (161) immediately yields following expressions for the time evolution of the expansion coefficients:

\[
\alpha_\uparrow(t) = \alpha_\uparrow(t_0) \exp \left\{-\frac{i}{2} \Omega (t - t_0) \right\}, \quad \alpha_\downarrow(t) = \alpha_\downarrow(t_0) \exp \left\{\frac{i}{2} \Omega (t - t_0) \right\},
\]
(4.169)
allowing a ready calculation of time evolution of the expectation values of any observable.

In particular, we can calculate the expectation value of \(S_z\) as a function of time by applying Eq. (130) to an arbitrary time moment \(t\):

\[
\langle S_z \rangle(t) = \frac{\hbar}{2} \left[ \alpha_\uparrow(t) \alpha_\uparrow^*(t) - \alpha_\downarrow(t) \alpha_\downarrow^*(t) \right] = \frac{\hbar}{2} \left[ \alpha_\uparrow(0) \alpha_\uparrow^*(0) - \alpha_\downarrow(0) \alpha_\downarrow^*(0) \right] = \langle S_z \rangle(0).
\]
(4.170)

Thus the expectation value of the spin component parallel to the applied magnetic field remains constant, regardless of the initial state of the system. However, this is not true for the components perpendicular to the field. For example, Eq. (132), applied to moment \(t\), gives

\[
\langle S_x \rangle(t) = \frac{\hbar}{2} \left[ \alpha_\uparrow(t) \alpha_\downarrow^*(t) + \alpha_\downarrow(t) \alpha_\uparrow^*(t) \right] = \frac{\hbar}{2} \left[ \alpha_\uparrow(0) \alpha_\downarrow^*(0) e^{-i\Omega(t-t_0)} + \alpha_\downarrow(0) \alpha_\uparrow^*(0) e^{i\Omega(t-t_0)} \right].
\]
(4.171)

Clearly, this expression describes sinusoidal oscillations with frequency (165). The amplitude and phase of these oscillations depend on initial conditions. Indeed, solving Eqs. (132)-(133) for the expansion coefficient products, we get relations

\[
\hbar \alpha_\uparrow(t) \alpha_\downarrow^*(t) = \langle S_x \rangle(t) + i \langle S_y \rangle(t), \quad \hbar \alpha_\downarrow(t) \alpha_\uparrow^*(t) = \langle S_x \rangle(t) - i \langle S_y \rangle(t)
\]
(4.172)
valid for any time \(t\). Plugging their values for \(t = 0\) into Eq. (171), we get

\[
\langle S_x \rangle(t) = \frac{1}{2} \left[ \langle S_x \rangle(0) + i \langle S_y \rangle(0) \right] e^{i\Omega(t-t_0)} + \frac{1}{2} \left[ \langle S_x \rangle(0) - i \langle S_y \rangle(0) \right] e^{-i\Omega(t-t_0)}
\]
(4.173)
\[
= \langle S_x \rangle(0) \cos \Omega t - \langle S_y \rangle(0) \sin \Omega t.
\]

An absolutely similar calculation using Eq. (133) gives

35 Note also that if the product \(\gamma B\) is positive, so is \(\Omega\), so that \(E_\uparrow\) is negative, while \(E_\downarrow\) is positive. This is in the correspondence with the classical picture of a magnetic dipole \(m\) having negative potential energy when it is aligned with the external magnetic field \(B\) – see Eq. (162).
\[ \langle S_y \rangle(t) = \langle S_y \rangle(0) \cos \Omega t + \langle S_x \rangle(0) \sin \Omega t. \] (4.174)

These formulas show, for example, if at moment \( t = 0 \) the spin’s state was \( \uparrow \), i.e. \( \langle S_x \rangle(0) = \langle S_y \rangle(0) = 0 \), then the amplitude of oscillation of the both “lateral” component of spin vanishes. On the other hand, if the spin was initially in state \( \rightarrow \), i.e. had the definite, maximum possible value of \( S_x \), equal to \( \hbar/2 \) (in classics, we would say “the spin \( \hbar/2 \) was oriented in direction \( x \)”), then both expectation values \( \langle S_x \rangle \) and \( \langle S_y \rangle \) oscillate in time\(^{36} \) with this amplitude, with the phase shift \( \pi/2 \) between them. These formulas may be interpreted as the torque-induced precession of the Cartesian components of the spin vector of length \( S = \hbar/2 \), confined in plane \([x, y]\), with classical frequency \( \Omega = \gamma B \) about axis \( z \) (counterclockwise if \( \gamma B > 0 \)).

Thus, the gyromagnetic ratio is just the angular frequency of the torque-induced precession of spin (about field’s direction) per unit magnetic field; for electrons, \( |\gamma| \approx 1.761 \times 10^{11} \text{ s}^{-1}\text{T}^{-1} \); for protons, the ratio is much smaller because of their larger mass: \( \gamma_p \approx 2.675 \times 10^8 \text{ s}^{-1}\text{T}^{-1} \), and for larger spin-\( \frac{1}{2} \) nuclei, \( \gamma \) may be much smaller still – e.g., \( 8.681 \times 10^6 \text{ s}^{-1}\text{T}^{-1} \) for the \(^{57}\text{Fe} \) nucleus.\(^{37} \)

Note, however, that this classical language does not describe large quantum-mechanical uncertainties of these observables, which are absent in the classical picture of the precession – at least when it starts from a definite orientation of the angular momentum vector.

Now let us return to the discussion of the general Schrödinger equation (158) and prove the following fascinating fact: it is possible to write the general solution of this operator equation. In the easiest case when the Hamiltonian is time-independent, this solution is an exact analog of Eq. (161),

\[ \hat{u}(t_0, t_0) = \hat{u}(t_0, t_0) \exp \left\{ -\frac{i}{\hbar} \hat{H}(t - t_0) \right\} = \hat{I} \exp \left\{ -\frac{i}{\hbar} \hat{H}(t - t_0) \right\}. \] (4.175)

To start its proof we should, first of all, understand what does a function (in this case, the exponent) of an operator mean. In the operator (and matrix) algebra, such functions are defined by their Taylor expansions; in particular, Eq. (175) means that

\[
\hat{u}(t, t_0) = \hat{I} + \sum_{k=1}^{\infty} \frac{1}{k!} \left( -\frac{i}{\hbar} \hat{H} \right)^k.
\] (4.176)

where \( \hat{H}^2 \equiv \hat{H} \hat{H} \), \( \hat{H}^3 \equiv \hat{H} \hat{H} \hat{H} \), etc. Working with such series of operator products is not as hard as one could imagine, due to their regular structure. For example, let us differentiate Eq. (176) over \( t \):

\(^{36} \)This is one more (hopefully, redundant :-) illustration of the difference between averaging over the statistical ensemble and over time: in Eqs. (170), (173)-(174), and quite a few relations below, only the former averaging has been performed, so the results are still functions of time.

\(^{37} \)Such composite particles as nuclei (and, from the point of view of high-energy physics, even such hadrons as protons) may be characterized by a certain net spin (and hence by certain \( \gamma \)) only if during the considered process their internal degrees of freedom remain in a certain (usually, ground) quantum state.
\[
\hat{u}(t,t_0) = \hat{0} + \frac{1}{1!} \left( -\frac{i}{\hbar} \right) \hat{H} + \frac{1}{2!} \left( -\frac{i}{\hbar} \right)^2 \hat{H}^2 2(t-t_0) + \frac{1}{3!} \left( -\frac{i}{\hbar} \right)^2 \hat{H}^2 3(t-t_0)^2 + ...
\]

so that the differential equation (158) is indeed satisfied. On the other hand, Eq. (175) also satisfies the initial condition

\[
\hat{u}(t_0,t_0) = \hat{u}^\dagger(t_0,t_0) = \hat{I},
\]

which immediately follows from the definition (157) of the evolution operator, so it is indeed the (unique) solution for the time evolution operator – in the Schrödinger picture.

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 instants commute with each other:

\[
[H(t'), H(t'')] = 0, \quad \text{for any } t', t''.
\]  

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

\[
\hat{H}_F = -\nabla F(t) \cdot \hat{\mathbf{r}}.
\]

Indeed, the radius-vector operator \( \hat{\mathbf{r}} \) does not depend explicitly on time and hence commutes with itself, as well as with \( c \)-numbers \( F(t') \) and \( F(t'') \).) In this case it is sufficient to replace, in all above formulas, product \( \hat{H}(t-t_0) \) with the corresponding integral over time; in particular, Eq. (175) is generalized as

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

This replacement means that the first form of Eq. (176) should be replaced with

\[
\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^\infty \frac{1}{k!} \left( -\frac{i}{\hbar} \right)^k \int_{t_0}^t \hat{H}(t') dt' = \hat{I} + \sum_{k=1}^\infty \frac{1}{k!} \int_{t_0}^t \int_{t_0}^{t'} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_k} dt_k \hat{H}(t_1)\hat{H}(t_2)\ldots\hat{H}(t_k).
\]  

The proof that the first form of Eq. (182) satisfies Eq. (158) is absolutely similar to the one carried out above.

We may now use Eq. (181) to show that the time-evolution operator is unitary at any moment, even for the time-dependent Hamiltonian. Indeed, from that formula,

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

Since each of the exponents may be presented with the Taylor series (182), and, thanks to Eq. (179), different components of these sums may be swapped at will, expression (183) may be manipulated exactly as the product of \( c \)-number exponents, in particular rewritten it as
\[ \hat{u}(t, t_0) \hat{u}^\dagger(t, t_0) = \hat{I} \exp \left\{ -\frac{i}{\hbar} \left[ \int_{t_0}^{t} \hat{H}(t') dt' - \int_{t_0}^{t} \hat{H}(t'') dt'' \right] \right\} = \hat{I} \exp \{ \hat{0} \} = \hat{I}. \quad (4.184) \]

This property ensures, in particular, that the system state’s normalization does not depend on time:

\[ \langle \alpha(t) | \alpha(t) \rangle = \langle \alpha(t_0) | \hat{u}^\dagger(t, t_0) \hat{u}(t, t_0) | \alpha(t_0) \rangle = \langle \alpha(t_0) | \alpha(t_0) \rangle. \quad (4.185) \]

The most difficult cases for the explicit solution of Eq. (158) are those when Eq. (179) is violated. It may be proven that in these cases the integral limits in the last form of Eq. (182) should be truncated, giving the so-called Dyson series

\[ \hat{u}(t, t_0) = \hat{I} + \sum_{k=1}^\infty \frac{1}{k!} \left( -\frac{i}{\hbar} \right)^k \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{k-1}} dt_k \hat{H}(t_1) \hat{H}(t_2) \cdots \hat{H}(t_k). \quad (4.186) \]

Since we would not have time to use this relation in our course, I will skip its proof.

Let me now return to the general discussion of quantum dynamics to outline its alternative, *Heisenberg picture*. For that, let us recall that according to Eq. (125), in quantum mechanics the expectation value of any observable \( A \) is a long bra-ket. Below we will see that other quantities (say, the rates of quantum transitions between pairs of different states, say \( \alpha \) and \( \beta \)) may also be measured in experiment; the most general form for all such measurable quantities is the following long bracket:

\[ \langle \alpha | \hat{A} | \beta \rangle. \quad (4.187) \]

As has been discussed above, in the Schrödinger picture the bra- and ket-vectors of the states are time-dependent, while the variable operators stay constant (if the corresponding variables do not explicitly depend on time), so that Eq. (187), applied to moment \( t \), may be presented as

\[ \langle \alpha(t) | \hat{A}_S | \beta(t) \rangle, \quad (4.188) \]

where index “S” emphasizes the Schrödinger picture. Let us apply to the bra- and ket-vectors in this expression the evolution law (157):

\[ \langle \alpha(t) | \hat{A} | \beta(t) \rangle = \langle \alpha(t_0) | \hat{u}^\dagger(t, t_0) \hat{A}_S \hat{u}(t, t_0) | \beta(t_0) \rangle. \quad (4.189) \]

This equality means that if we form a long bracket with bra- and ket-vectors of the initial-time states, together with the following time-dependent *Heisenberg operator* 40

\[ \hat{A}_H(t) \equiv \hat{u}^\dagger(t, t_0) \hat{A}_S \hat{u}(t, t_0) = \hat{u}^\dagger(t, t_0) \hat{A}_H(t_0) \hat{u}(t, t_0), \quad (4.190) \]

all experimentally measurable results will remain the same as in the Schrödinger picture:

\[ \langle \alpha | \hat{A} | \beta \rangle = \langle \alpha(t_0) | \hat{A}_H(t, t_0) | \beta(t_0) \rangle. \quad (4.191) \]

---

38 We will run into such situations in Chapter 7, but will not need to apply Eq. (186).
39 It may be found, for example, in Chapter 5 of J. Sakurai’s textbook – see References.
40 Note this relation is similar in structure to the symbolic Eqs. (94).
Let us see how does the Heisenberg picture work for the same simple (but very important!) problem of the spin-½ precession in a $z$-oriented magnetic field, described (in the $z$-basis) by the Hamiltonian matrix (164). In that basis, Eq. (158) for the time-evolution operator reads

$$\begin{pmatrix}
    1 & 0 \\
    0 & -1
\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}
    u_{11} & u_{12} \\
    u_{21} & u_{22}
\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}
    u_{11} & u_{12} \\
    -u_{21} & -u_{22}
\end{pmatrix}.$$

We see that in this simple case the equations for different matrix elements of the evolution operator matrix are decoupled, and readily solvable, using the universal initial condition (178):

$$u(t,0) = \begin{pmatrix}
    e^{\frac{i\Omega t}{2}} & 0 \\
    0 & e^{-\frac{i\Omega t}{2}}
\end{pmatrix} \equiv 1 \cos \frac{\Omega t}{2} - i\sigma_z \sin \frac{\Omega t}{2}. \quad (4.193)$$

Now we can use Eq. (190) to find the Heisenberg-picture operators of spin components. Dropping index "H" for brevity (the Heisenberg-picture operators are clearly marked by their dependence on time anyway), we get

$$S_x(t) = u(t,0)^\dagger S_x(0)u(t,0) = \frac{\hbar}{2}u(t,0)^\dagger \sigma_x u(t,0) = \frac{\hbar}{2} \begin{pmatrix}
    e^{i\Omega t / 2} & 0 \\
    0 & e^{-i\Omega t / 2}
\end{pmatrix} \begin{pmatrix}
    0 & 1 \\
    1 & 0
\end{pmatrix} \begin{pmatrix}
    e^{-i\Omega t / 2} & 0 \\
    0 & e^{i\Omega t / 2}
\end{pmatrix} \equiv \frac{\hbar}{2} \left[ \sigma_x \cos \Omega t - \sigma_y \sin \Omega t \right] \equiv S_x(0) \cos \Omega t - S_y(0) \sin \Omega t. \quad (4.194)$$

Absolutely similar calculations of the other spin components yield

$$S_y(t) = \frac{\hbar}{2} \begin{pmatrix}
    0 & -ie^{i\Omega t} \\
    ie^{-i\Omega t} & 0
\end{pmatrix} = \frac{\hbar}{2} \left[ \sigma_y \cos \Omega t + \sigma_z \sin \Omega t \right] \equiv S_y(0) \cos \Omega t + S_z(0) \sin \Omega t, \quad (4.195)$$

$$S_z(t) = \frac{\hbar}{2} \begin{pmatrix}
    1 & 0 \\
    0 & -1
\end{pmatrix} = \frac{\hbar}{2} \sigma_z = S_z(0). \quad (4.196)$$

A practical advantage of these formulas is that they describe system’s evolution for arbitrary initial conditions, thus making the analysis of the initial state effects very simple. Indeed, since in the Heisenberg picture the expectation values of observables are calculated using Eq. (191) (with $\beta = \alpha$), with time-independent bra- and ket vectors, such averaging of Eqs. (194)-(196) immediately returns us to Eqs. (170), (173), and (174), obtained in the Schrödinger picture. Moreover, these equations for the Heisenberg operators formally coincide with the classical equations of the torque-induced precession for $c$-number variables. (In the next chapter, we will see that the same exact mapping is valid for the Heisenberg picture of the orbital motion.)

---

41 We could of course use this equation result, together with Eq. (157), to obtain all the above results for this system within the Schrödinger picture. In our simple case, the use of Eqs. (161) for this purpose was more straightforward, but in some cases (e.g., for time-dependent Hamiltonians) an explicit calculation of the time-evolution matrix may be the only practicable way to proceed.
In order to see that the last fact is by no means a coincidence, let us combine Eqs. (158) and (190) to form an explicit differential equation of the Heisenberg operator evolution. For that, let us differentiate Eq. (190) over time:

\[
\frac{d}{dt} \hat{A}_{II} = \frac{\partial \hat{u}^\dagger}{\partial t} \hat{A}_S \hat{u} + \hat{u}^\dagger \frac{\partial \hat{A}_S}{\partial t} \hat{u} + \hat{u}^\dagger \hat{A}_S \frac{\partial \hat{u}}{\partial t}.
\]  

(4.197)

Plugging in the derivatives of the time evolution operator from Eq. (158) and its Hermitian conjugate, and multiplying both parts of the equation by \(i\hbar\), we get

\[
i\hbar \frac{d}{dt} \hat{A}_{II} = -\hat{u}^\dagger \hat{H} \hat{A}_S \hat{u} + \hat{u}^\dagger \frac{\partial \hat{A}_S}{\partial t} \hat{u} + \hat{u}^\dagger \hat{A}_S \hat{H} \hat{u}.
\]

(4.198a)

If for the Schrödinger-picture Hamiltonian the condition similar to Eq. (179) is satisfied, then, according to Eqs. (177) or (182), the Hamiltonian commutes with the time evolution operator and its Hermitian conjugate, and may be swapped with any of them.\(^{42}\) Hence, we may rewrite Eq. (198a) as

\[
i\hbar \frac{d}{dt} \hat{A}_{II} = -\hat{H} \hat{u} + \hat{u}^\dagger \frac{\partial \hat{A}_S}{\partial t} \hat{u} + \hat{u}^\dagger \hat{A}_S \hat{H} \hat{u} + \left[ \hat{u}^\dagger \hat{A}_S \hat{u}, \hat{H} \right].
\]

(4.198b)

Now using the definition (190) again, for both terms in the right-hand part, we may write

\[
i\hbar \frac{d}{dt} \hat{A}_{II} = i\hbar \left( \frac{\partial \hat{A}}{\partial t} \right)_{II} + [\hat{A}_{II}, \hat{H}].
\]

(4.199)

This is the so-called \textit{Heisenberg equation of motion}.\(^{43}\)

Let us see how does this equation look for the same problem of spin-\(\frac{1}{2}\) precession in a \(z\)-oriented, time-independent magnetic field, described in the \(z\)-basis by the Hamiltonian matrix (164), which does not depend on time. In this basis, Eq. (199) for the vector operator of spin reads\(^{44}\)

\[
i\hbar \begin{pmatrix} \hat{S}_{11} & \hat{S}_{12} \\ \hat{S}_{21} & \hat{S}_{22} \end{pmatrix} = \frac{\hbar \Omega}{2} \begin{pmatrix} \hat{S}_{11} & \hat{S}_{12} \\ \hat{S}_{21} & \hat{S}_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hbar \Omega \begin{pmatrix} 0 & -\hat{S}_{12} \\ \hat{S}_{21} & 0 \end{pmatrix}.
\]

(4.200)

Once again, the equations for different matrix elements are decoupled, and their solution is elementary:

\[
\begin{align*}
\hat{S}_{11}(t) &= \hat{S}_{11}(0) = \text{const}, \\
\hat{S}_{22}(t) &= \hat{S}_{22}(0) = \text{const}, \\
\hat{S}_{12}(t) &= \hat{S}_{12}(0)e^{i\Omega t}, \\
\hat{S}_{21}(t) &= \hat{S}_{21}(0)e^{-i\Omega t}.
\end{align*}
\]

(4.201)

\(^{42}\) Due to the same reason, \(\hat{H}_{II} \equiv \hat{u}^\dagger \hat{H} \hat{S} \hat{u} = \hat{u}^\dagger \hat{H} \hat{u} = \hat{H}_S\); this is why the index of the Hamiltonian operator may be dropped in Eqs. (198)-(199).

\(^{43}\) Reportedly, this equation was derived by P. A. M. Dirac, who was so generous that he himself gave the name of his colleague to this key result, because “Heisenberg was saying something like this”.

\(^{44}\) Using commutation relations (155), this equation may be readily generalized to the case of arbitrary magnetic field \(\mathbf{B}(t)\) and arbitrary state basis – the exercise highly recommended to the reader.
According to Eq. (190), the initial “values” of the Heisenberg-picture matrix elements are just the Schrödinger-picture ones, so that using Eq. (117) we may rewrite this solution in either of two forms:

\[
S(t) = \frac{\hbar}{2} \left[ n_x \begin{pmatrix} 0 & e^{i\Omega t} \\ e^{-i\Omega t} & 0 \end{pmatrix} + n_y \begin{pmatrix} 0 & -ie^{i\Omega t} \\ ie^{-i\Omega t} & 0 \end{pmatrix} + n_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right]
\]

\[
= \frac{\hbar}{2} \left[ \begin{pmatrix} n_z & n_z e^{i\Omega t} \\ n_z e^{-i\Omega t} & -n_z \end{pmatrix} \right], \quad \text{with } n_z \equiv n_x \pm in_y.
\] (4.202)

The simplicity of the last expression is spectacular. (Remember, it covers any initial conditions, and all 3 spatial components of spin!) On the other hand, for some purposes the former expression may be more convenient; in particular, its Cartesian components immediately give our earlier results (194)-(196).

One of advantages is that the Heisenberg picture is that it provides a more clear link between the classical and quantum mechanics. Indeed, analytical classical mechanics may be used to derive the following equation of time evolution of an arbitrary function \(A(q_j, p_j, t)\) of generalized coordinates and momenta of the system, and time: \(^{45}\)

\[
\frac{dA}{dt} = \frac{\partial A}{\partial t} - \{A, H\},\] (4.203)

where \(H\) is the classical Hamiltonian function of the system, and \(\{...,\}\) is the so-called Poisson bracket defined, for two arbitrary functions \(A(q_j, p_j, t)\) and \(B(q_j, p_j, t)\), as

\[
\{A, B\} \equiv \sum_j \left( \frac{\partial A}{\partial q_j} \frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j} \frac{\partial B}{\partial q_j} \right).
\] (4.204)

Comparing Eq. (203) with Eq. (199), we see that the correspondence between the classical and quantum mechanics (in the Heisenberg picture) is provided by the following symbolic relation\(^{46}\)

\[
\{A, B\} \leftrightarrow \frac{i}{\hbar} [\hat{A}, \hat{B}].
\] (4.205)

\(^{45}\) See, e.g., CM Eq. (10.17). Also, please excuse my use, for the Poisson bracket, the same (traditional) symbol \{………\} as for the anticommutator. We will not run into the Poisson brackets again in the course, leaving very little chance for confusion.

\(^{46}\) Since we have run into the commutator of Heisenberg-picture operators, let me note emphasize again that the “values” of such an operator at different moments of time often do not commute. Perhaps the simplest example is the operator \(\hat{x}\) of coordinate of a free 1D particle, with Hamiltonian \(\hat{H} = \hat{p}^2 / 2m\). Indeed, in this case Eq. (199) yields equations \(i\hbar \frac{d}{dt} = [\hat{x}, \hat{H}] = i\hbar \hat{p} / m\) and \(i\hbar \frac{d}{dt} = [\hat{p}, \hat{H}] = 0\), with simple solutions (similar to those for classical motion of the corresponding observables): \(\hat{p}(t) = \text{const} = \hat{p}(0), \quad \hat{x}(t) = \hat{x}(0) + \hat{p}(0)t / m\), so that \([\hat{x}(0), \hat{x}(t)] = [\hat{x}(0), \hat{p}(0)]t / m = [\hat{x}_s, \hat{p}_s]t / m = i\hbar t / m \neq 0, \text{ if } t \neq 0\).
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 chapter where we revisit the wave mechanics, and also in Chapter 9.

Finally, let us discuss one more alternative picture of quantum dynamics. It is also attributed to P. A. M. Dirac, and is called either the “Dirac picture”, or (more frequently) the interaction picture. The last name stems from the fact that this picture is very useful for the perturbative (approximate) approaches to systems whose Hamiltonians may be partitioned into two parts,

\[ \hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}, \]  

where \( \hat{H}_0 \) is the sum of relatively simple Hamiltonians of non-interacting component sub-systems, while their second term in Eq. (206) represents their weak interaction. (Note, however, that the relations in the balance of this section are exact and not based on these assumptions.) In this case, it is natural to consider, together with the genuine unitary operator \( \hat{u}(t,t_0) \) of the time evolution of the system, which obeys Eq. (158), a similarly defined unitary operator of evolution of the “unperturbed system” described by Hamiltonian \( \hat{H}_0 \) alone:

\[ i\hbar \hat{u}_0 = \hat{H}_0 \hat{u}_0, \]  

and also the following interaction evolution operator,

\[ \hat{u}_I \equiv \hat{u}_0^\dagger \hat{u}. \]  

The sense of this definition becomes more clear if we insert the reciprocal relation,

\[ \hat{u} \equiv \hat{u}_0 \hat{u}_0^\dagger \hat{u} = \hat{u}_0 \hat{u}_I, \]  

and its Hermitian conjugate,

\[ \hat{u}^\dagger = \left(\hat{u}_0 \hat{u}_I\right)^\dagger = \hat{u}_I^\dagger \hat{u}_0^\dagger, \]  

into the basic Eq. (190) – which is valid in any picture:

\[ \langle \alpha | \hat{A} | \beta \rangle = \langle \alpha(t_0) | \hat{u}_0^\dagger (t,t_0) \hat{A}_I \hat{u}(t,t_0) | \beta(t_0) \rangle = \langle \alpha(t_0) | \hat{u}_I^\dagger (t,t_0) \hat{u}_0^\dagger (t,t_0) \hat{A}_I \hat{u}_0(t,t_0) \hat{u}_I(t,t_0) | \beta(t_0) \rangle. \]

This relation shows that all calculations of the observable expectation values and transition rates (i.e. all the results of quantum mechanics that may be experimentally verified) are expressed by the following formula, with the standard bra-ket structure (187),

\[ \langle \alpha | \hat{A} | \beta \rangle = \langle \alpha_I(t) | \hat{A}_I(t) | \beta_I(t) \rangle, \]

if we assume that both the state vectors and operators evolve in time, with the vectors evolving due to the interaction operator \( \hat{u}_I \),

\[ \langle \alpha_I(t) | = \langle \alpha(t_0) | \hat{u}_I^\dagger (t,t_0), \quad | \beta_I(t) \rangle = \hat{u}_I(t,t_0) | \beta(t_0) \rangle, \]

while the operators’ evolution being governed by the unperturbed operator \( \hat{u}_0 \):

\[ \hat{A}_I(t) \equiv \hat{u}_0^\dagger (t,t_0) \hat{A}_I \hat{u}_0(t,t_0) \].
These relations describe the interaction picture of quantum dynamics. Let me defer an example of its convenience until the perturbative analysis of open quantum systems in Sec. 7.6, and here end the discussion with a proof that the interaction evolution operator satisfies the Schrödinger equation,

\[ i\hbar \dot{\hat{u}}_t = \hat{H}_I \hat{u}_t, \]  

(4.215)
in which \( \hat{H}_I \) is the interaction Hamiltonian transformed in accordance with rule (214):

\[ \hat{H}_I(t) = \hat{u}_0(\tau, t)\hat{H}_{\text{int}}\hat{u}_0(\tau, t). \]  

(4.216)
The proof is very straightforward: first using definition (208), and then Eqs. (158) and the Hermitian conjugate of Eq. (207), we may write

\[ i\hbar \dot{\hat{u}}_t = i\hbar \frac{d}{dt}(\hat{u}_0^\dagger\hat{u}) = i\hbar(\hat{\dot{u}}_0^\dagger\hat{u} + \hat{u}_0^\dagger\hat{\dot{u}}) = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{\dot{u}}_0 + \hat{\dot{u}}_0^{\dagger}(\hat{H}_0 + \hat{H}_{\text{int}})\hat{u} \]

\[ \dot{\hat{u}}_t = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{\dot{u}}_0 + \hat{\dot{u}}_0^{\dagger}\hat{H}_{\text{int}}\hat{u} = \left(-\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{\dot{u}}_0 + \hat{\dot{u}}_0^{\dagger}\hat{H}_{\text{int}}\right)\hat{u} + \hat{\dot{u}}_0^{\dagger}\hat{H}_{\text{int}}\hat{u}. \]  

(4.217)
Since \( \hat{u}_0^\dagger \) may be presented as an integral of \( \hat{H}_0 \) (similar to Eq. (181) relating \( \hat{u} \) and \( \hat{H} \), these operators commute, so that the parentheses in the last form of Eq. (217) vanish. Now plugging \( \hat{u} \) from Eq. (209), we get the equation,

\[ i\hbar \dot{\hat{u}}_t = \hat{u}_0^\dagger\hat{H}_{\text{int}}\hat{u}_0\hat{u}_t = \left(\hat{u}_0^\dagger\hat{H}_{\text{int}}\hat{u}_0\right)\hat{u}_t, \]

(4.218)
that is equivalent to the combination of Eqs. (215) and (216).

Equation (215) shows that if the energy scale of interaction \( H_{\text{int}} \) is much weaker than the background energy \( H_0 \), operators \( \hat{u} \) and \( \hat{u}_t \), and hence the state vectors (213) evolve relatively slowly. Such an exclusion of fast background oscillations is especially convenient for the perturbative approaches to complex interacting systems, in particular to the open quantum systems that weakly interact with their environment – see Sec. 7.6.

4.7. Exercise problems

4.1. Let \( \alpha \) and \( \beta \) be two possible quantum states of the same system, and \( \hat{A} \) be a linear operator. Which of the following expressions are legitimate (i.e. have a well-defined meaning) within the bra-ket formalism?

1. \( \langle \alpha | \rangle \)  
2. \( \langle \alpha | \beta \rangle^2 \)  
3. \( |\alpha\rangle\langle \beta | \)  
4. \( \hat{A}^* \)  
5. \( \langle \hat{A} | \rangle \)

6. \( \langle \alpha | \hat{A} \rangle \)  
7. \( \alpha\langle \hat{A} \rangle \)  
8. \( |\alpha\rangle^2 \)  
9. \( \hat{A}^2 \)  
10. \( \langle \alpha | \hat{A}^\dagger \rangle \)

4.2. Prove that if \( \hat{A} \) and \( \hat{B} \) are linear operators, then:

(i) \( \left(\hat{A}^\dagger\right)^\dagger = \hat{A} \);  
(ii) \( i\hat{\dot{A}}^\dagger = -i\hat{A}^\dagger \);  
(iii) \( \left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger \);

(iv) operators \( \hat{A}\hat{A}^\dagger \) and \( \hat{A}^\dagger\hat{A} \) are Hermitian.
4.3. Prove that for any linear operators $\hat{A}, \hat{B}, \hat{C}, \hat{D}$,

$$[\hat{A}\hat{B}, \hat{C}\hat{D}] = \hat{A}[\hat{B}, \hat{C}]\hat{D} - \hat{A}[\hat{C}, \hat{D}]\hat{B} + \{\hat{A}, \hat{C}\}\hat{D}\hat{B} - \{\hat{A}, \hat{D}\}\hat{C}\hat{B}.$$ 

4.4. Calculate all possible binary products $\sigma_j \sigma_{j'}$ (for $j, j' = x, y, z$) of the Pauli matrices (105),

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

and their commutators and anticommutators (defined similarly to those of the corresponding operators). Present the results using the Kronecker delta and Levi-Civita permutation symbols.\(^{47}\)

4.5. Calculate the following expressions,

(i) $(c \cdot \sigma)^n$, and then

(ii) $(bI + c \cdot \sigma)^n$,

for the scalar product $c \cdot \sigma$ of the Pauli matrix vector $\sigma \equiv n_x \sigma_x + n_y \sigma_y + n_z \sigma_z$ by an arbitrary $c$-number vector $c$, where $n \geq 0$ is an integer, and $b$ is an arbitrary scalar $c$-number.

*Hint:* For task (ii), you may like to use the binomial theorem,\(^{48}\) and then transform the result in a way enabling you to use the same theorem backwards.

4.6. * Use the results of the previous problem to derive Eqs. (2.165)-(2.166) for the transparency $T$ of a system of $N$ similar, equidistant, delta-functional tunnel barriers.

4.7. Use result of Problem 5 to spell out the following matrix: $\exp\{i \theta n \cdot \sigma\}$, where $\sigma$ is the vector of Pauli matrices, $n$ is a $c$-number vector of unit length, and $\theta$ is a $c$-number scalar.

4.8. Use the result of Problem 5(ii) to calculate $\exp\{A\}$, where $A$ is an arbitrary $2 \times 2$ matrix.

4.9. Express elements of matrix $B = \exp\{A\}$ explicitly via those of the $2 \times 2$ matrix $A$. Spell out your result for the following matrices:

$$A = \begin{pmatrix} a & a \\ a & a \end{pmatrix}, \quad A' = \begin{pmatrix} i\varphi & i\varphi \\ i\varphi & i\varphi \end{pmatrix},$$

with real $a$ and $\varphi$.

4.10. Prove that for arbitrary square matrices $A$ and $B$,

$$\text{Tr} (AB) = \text{Tr} (BA).$$

Is each diagonal element $(AB)_{jj}$ necessarily equal to $(BA)_{jj}$?

---

\(^{47}\) See, e.g., MA Eqs. (13.1) and (13.2).

\(^{48}\) See, e.g., MA Eq. (2.9).
4.11. Prove that the matrix trace of an arbitrary operator does not change at an arbitrary unitary transformation.

4.12. Prove that for any two full and orthonormal bases \( u_j, v_j \) of the same Hilbert space,
\[
\text{Tr} \left( |u_j\rangle \langle v_j| \right) = \langle v_j | u_j \rangle.
\]

4.13. Is the 1D scattering matrix \( S \), defined by Eq. (133), unitary? What about the 1D transfer matrix \( T \) defined by Eq. (134)?

4.14. Calculate the trace of the following matrix:
\[
\exp[i \mathbf{a} \cdot \mathbf{\sigma}] \exp[i \mathbf{b} \cdot \mathbf{\sigma}],
\]
where \( \mathbf{\sigma} \) is the Pauli matrix vector, while \( \mathbf{a} \) and \( \mathbf{b} \) are usual \((c\text{-number})\) geometric vectors.

4.15. Let \( A_j \) be eigenvalues of some operator \( \hat{A} \). Express the following two sums,
\[
\Sigma_1 = \sum_j A_j, \quad \Sigma_2 = \sum_j A_j^2,
\]
via the matrix elements \( A_{jj'} \) of this operator in an arbitrary basis.

4.16. Calculate \( \langle \sigma_z \rangle \) of a two-level system 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 \( \sigma_z \) and \( \sigma_x \), respectively.

*Hint:* Double-check whether the solution you are giving is general.

4.17. An electron is fully polarized in the positive \( z \)-direction. Calculate the probabilities of the alternative outcomes of a perfect Stern-Gerlach experiment with the magnetic field \( \mathbf{B} \) oriented in the direction of some axis \( \mathbf{n} \), performed on this electron.

4.18. A perfect Stern-Gerlach instrument makes a single-shot measurement of the following combination, \((S_x + S_y)/\sqrt{2}\), of two spin components of a \( z \)-polarized electron; after that, component \( S_z \) of the same particle is measured. What are the possible outcomes of these measurements and their probabilities?

4.19. In a certain basis, the Hamiltonian of a spin-\(\frac{1}{2}\) (two-level) system is described by matrix
\[
H = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad \text{with} \ E_1 \neq E_2,
\]
and the operator of some observable \( \hat{A} \), by matrix
\[
A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.
\]
For the system’s state with the energy equal exactly to \( E_1 \), find the possible results of measurements of observable \( A \) and the probabilities of the corresponding measurement outcomes.

4.20. States \( u_{1,2,3} \) form an orthonormal basis of a system with Hamiltonian

\[
\hat{H} = -\delta (\langle u_1 | u_2 \rangle + \langle u_2 | u_3 \rangle + \langle u_3 | u_1 \rangle) + \text{h.c.},
\]

where \( \delta \) is a real constant, and h.c. means the Hermitian conjugate of the previous expression. Calculate its stationary states and energy levels. Can you relate this system with any other(s) discussed earlier in the course?

4.21. Suggest a Hamiltonian describing particle’s dynamics in an infinite 1D set of similar quantum wells in the tight-binding approximation, in the bra-ket formalism, and verify that is yields the correct dispersion relation (2.206).

4.22. Calculate eigenvectors and eigenvalues of the following matrices:

\[
A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.
\]

4.23. Find eigenvalues of the following matrix:

\[
A = a \cdot \sigma = a_x \sigma_x + a_y \sigma_y + a_z \sigma_z,
\]

where \( a_{x,y,z} \) are real \( c \)-numbers (scalars), and \( \sigma_{x,y,z} \) are the Pauli matrices. Sketch the dependence of the eigenvalues on parameter \( a_z \), with \( a_x \) and \( a_y \) fixed. Compare the result with Fig. 29.

4.24. Derive a differential equation for the time evolution of the expectation value of an observable, using both the Schrödinger picture and the Heisenberg picture of quantum mechanics.

4.25. At \( t = 0 \), a spin-\( \frac{1}{2} \) particle, whose interaction with an external field is described by Hamiltonian

\[
\hat{H} = a \cdot \hat{\sigma} = a_x \hat{\sigma}_x + a_y \hat{\sigma}_y + a_z \hat{\sigma}_z,
\]

(where \( a_{x,y,z} \) are real and constant \( c \)-numbers, and \( \hat{\sigma}_{x,y,z} \) are the operators that, in the \( z \)-basis, are represented by the Pauli matrices \( \sigma_{x,y,z} \)), was in state \( \uparrow \), one of two eigenstates 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) the probabilities to find the system in states \( \uparrow \) and \( \downarrow \), and

(iii) the expectation values of all 3 spatial components (\( \hat{S}_x \), etc.) of the spin vector operator \( \hat{S} = (h/2)\hat{\sigma} \).

Analyze and interpret the results for the particular case \( a_y = a_z = 0 \).
4.26. For the same system as in 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 $\hat{\mathbf{S}}_H(t)$,
(ii) the expectation values of the spin components.

Compare the latter results with those of the previous problem.

4.27. For the same system as in two last problems, calculate the matrix elements of operator $\hat{\sigma}_z$ in the basis of eigenstates $a_1, a_2$.

Hint: In contrast to the cited problems, the answer evidently does not depend on the initial conditions.

4.28. In the Schrödinger picture of quantum mechanics, three operators satisfy the following commutation relation:

$$[\hat{A}, \hat{B}] = \hat{C}.$$ 

What is their relation in the Heisenberg picture (at the same time instant)?

4.29. A spin-$\frac{1}{2}$ particle is placed into a magnetic field $\mathbf{B}(t)$, which is an arbitrary function of time. Derive the differential equations describing the time evolution of:

(i) the vector operator $\hat{\mathbf{S}}$ of particle’s spin (in the Heisenberg picture), and
(ii) the expectation value $\langle \mathbf{S} \rangle$ of spin’s vector.

Contemplate the relative merits of the latter equation for the description of a single spin and of a large collection of similar, non-interacting spins.

4.30. * Prove the Bloch theorem given by either Eq. (3.107) or Eq. (3.108).

Hint: Consider the translation operator $\hat{T}_R$, defined by the following result of its action on an arbitrary function $f(\mathbf{r})$:

$$\hat{T}_R f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}),$$

where $\mathbf{R}$ is an arbitrary vector of the Bravais lattice (3.106). In particular, analyze the commutation properties of the operator, and apply them to an eigenfunction $\psi(\mathbf{r})$ of the stationary Schrödinger equation for a particle in a 3D periodic potential described by Eq. (3.105).