Chapter 4. Bra-ket Formalism

The objective of this chapter is to describe Dirac's "bra-ket" formalism of quantum mechanics, which not only overcomes some inconveniences of wave mechanics but also enables a natural description of such intrinsic properties of particles as their spin. In the course of the formalism's discussion, I will give only a few simple examples of its application, leaving more involved cases for the following chapters.

4.1. Motivation

As the reader could see from the previous chapters of these notes, wave mechanics gives many results of primary importance. Moreover, it is mostly sufficient for many applications, for example, 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) Attempts to analyze the *temporal* evolution of quantum systems, beyond the trivial time behavior of the stationary states, described by Eq. (1.62), run into technical difficulties. For example, we could derive Eq. (2.151) describing the metastable state's decay and Eq. (2.181) describing the 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 problems of this kind. Solving such problems for more complex potential profiles would entangle the time evolution analysis with the calculation of the spatial distribution of the evolving wavefunctions – which (as we could see in Secs. 2.9 and 3.6) may be rather complex even for time-independent potentials. Some separation of the spatial and temporal dependencies is possible using perturbation approaches (to be discussed in Chapter 6) but even those would lead, in the wavefunction language, to very cumbersome formulas.

(ii) The last statement can also be made concerning other issues that are conceptually addressable within the wave mechanics, e.g., the Feynman path integral approach, coupling to the environment, etc. Pursuing them in the wave mechanics language would lead to formulas so bulky that I had postponed their discussion until we would have 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 very 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, the quantum mechanical operators of the coordinate and momentum are treated rather unequally – see Eqs. (1.26b). However, some key expressions, e.g., for the fundamental eigenfunction of a free particle,

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

or the harmonic oscillator's Hamiltonian,

$$\hat{H}=\frac{1}{2m}\hat{p}^2+\frac{m\omega_0^2}{2}\hat{r}^2, \tag{4.2}$$

just beg for a similar treatment of coordinates and momenta.

© K. Likharev

However, the strongest motivation for a more general formalism comes from wave mechanics' conceptual inability to describe elementary particles' *spins*[1] and other internal quantum degrees of freedom, such as quarks' flavors. In this context, let us review the basic facts on spin (which is very representative and experimentally the most accessible of all internal quantum numbers), to understand what a more general formalism has to explain – as a minimum.

Figure 1 shows the conceptual scheme of the simplest spin-revealing experiment, first conceived by Otto Stern in 1921 and implemented by Walther Gerlach in 1922. A collimated beam of particles[2] from a natural source, such as a heated cathode, is passed through a gap between the poles of a strong magnet, whose magnetic field $\mathscr{B}$, (in Fig. 1, directed along the z-axis) is nonuniform, so both $\mathscr{B}_z$ and $d\mathscr{B}_z/dz$ are not equal to zero. The experiment shows that even if all particles are in the ground orbital state, the beam splits into two beams of equal intensity.

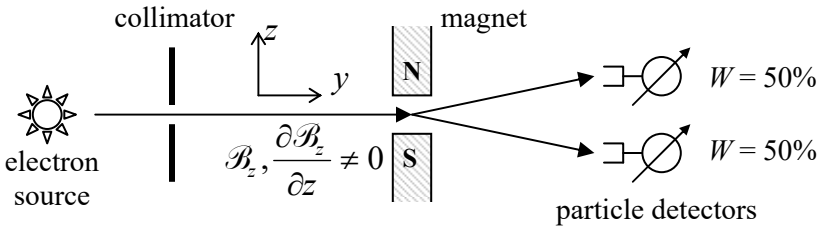


Fig. 4.1. The simplest Stern-Gerlach experiment.

This result may be semi-quantitatively explained on classical (if somewhat phenomenological) grounds by assuming that each particle has an intrinsic, permanent magnetic dipole moment **m**. Indeed, classical electrodynamics tells us[3] that the potential energy U of a magnetic dipole in an external magnetic field $\mathscr{B}$ is equal to $(-\mathbf{m}\cdot\mathscr{B})$, so the force acting on the particle,

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

has a non-zero z-component

$$F_z = -\frac{\partial}{\partial z}\left(-m_z\cdot\mathscr{B}_z\right) \equiv m_z\frac{\partial \mathscr{B}_z}{\partial z}. \tag{4.4}$$

Hence if we further assume that the particle's magnetic moment may take only two equally probable discrete values of $m_z = \pm\mu$ (though such discreteness does not follow from any classical model of the particle), this may explain the basic Stern-Gerlach effect qualitatively. The quantitative explanation of the beam splitting angle requires the magnitude of μ to be equal (or very close) to the so-called *Bohr magneton*[4]

$$\mu_{\rm B} \equiv \frac{\hbar e}{2m_{\rm e}} \approx 0.9274\times10^{-23}\,\frac{\rm J}{\rm T}. \tag{4.5}$$

Bohr magneton

[1] Reportedly, the concept of spin as a measure of the internal rotation of a particle was first suggested (though later rejected) by Ralph Kronig, then a 20-year-old student, in January 1925, a few months before two other students, George Uhlenbeck and Samuel Goudsmit, came to this idea independently. The concept was then accepted (first, rather reluctantly) and developed quantitatively by Wolfgang Pauli.

[2] The initial Stern-Gerlach experiments used silver atoms because their larger mass helps to decrease the spit beam widths. However, the discussion below is valid for any spin-½ particles including electrons.

[3] See, e.g., EM Sec. 5.4, in particular Eq. (5.100).

[4] A good mnemonic rule is that it is close to 1 K/T. In the Gaussian units, $\mu_{\rm B} \equiv \hbar e/2m_{\rm e}c \approx 0.9274\times10^{-20}$ erg/G.

However, as we will see below, this value cannot be explained by any internal motion of the particle, say its rotation about the *z*-axis. More importantly, this semi-classical phenomenology cannot explain, even qualitatively, other experimental results, for example those of the set of multistage Stern-Gerlach experiments shown in Fig. 2. In the first of the experiments, the particle beam is first passed through a magnetic field (and its gradient) oriented along the *z*-axis, just as in Fig. 1. Then one of the two resulting beams is absorbed (or removed from the setup in some other way), 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 an even more unnatural additional assumption that the initial particles had random but discrete components of the magnetic moment simultaneously in *two* directions, *z* and *x*.

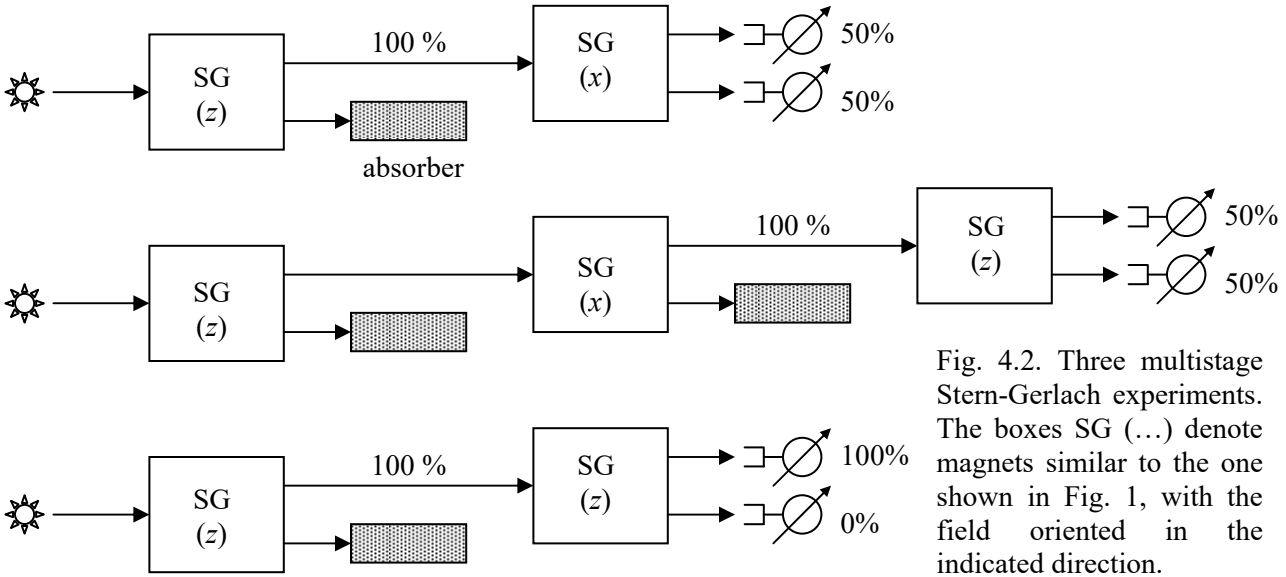


Fig. 4.2. Three multistage Stern-Gerlach experiments. The boxes SG (…) denote magnets similar to the one shown in Fig. 1, with the field oriented in the indicated direction.

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 particles with m_z corresponding to the lower beam, at the first *z*-stage. The only way to save the classical explanation here is to say that maybe, particles *somehow* interact with the magnetic field, so the *x*-polarized beam becomes spontaneously depolarized again somewhere between the two last stages. But any hope for such an 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 so-called *matrix mechanics* pioneered by Werner Heisenberg, Max Born, and Pascual Jordan in 1925. However, the matrix formalism is rather inconvenient for the solution of most problems discussed in Chapters 1-3, and for a short time, it was eclipsed by E. Schrödinger's wave mechanics, which had been put forward just a few months later. However, very soon Paul Adrien Maurice Dirac introduced a more general *bra-ket formalism* of quantum mechanics, which provides a generalization of both approaches and proves their equivalence. Let me describe it, begging for the reader's patience because (in contrast with my usual style), I will not be able to give particular examples of its application for a while – until all the basic notions of the formalism have been introduced.

4.2. States, state vectors, and linear operators

The basic notion of the general formulation of quantum mechanics is the *quantum state* of a system.[5] To get some gut feeling of this notion, if a quantum state α of a particle may be adequately described by wave mechanics, this description is given by the corresponding wavefunction $\Psi_\alpha(\mathbf{r}, t)$. Note, however, that a quantum state as such is *not* a mathematical object,[6] and can participate in mathematical formulas only as a "label" – e.g., the index of the wavefunction Ψ_α. On the other hand, such a wavefunction is *not* a state, but a mathematical object (a complex function of space and time) giving a quantitative description of the state – just as the classical radius vector $\mathbf{r}_\alpha$ and velocity $\mathbf{v}_\alpha$ as real functions of time are mathematical objects describing the motion of the particle in its classical description – see Fig. 3. Similarly, in the Dirac formalism, a certain quantum state α is described by either of two mathematical objects, called the *state vectors*: the *ket-vector* $|\alpha\rangle$ and the *bra-vector* $\langle\alpha|$,[7] whose relationship is close to that between the wavefunction Ψ_α and its complex conjugate Ψ_α^*.

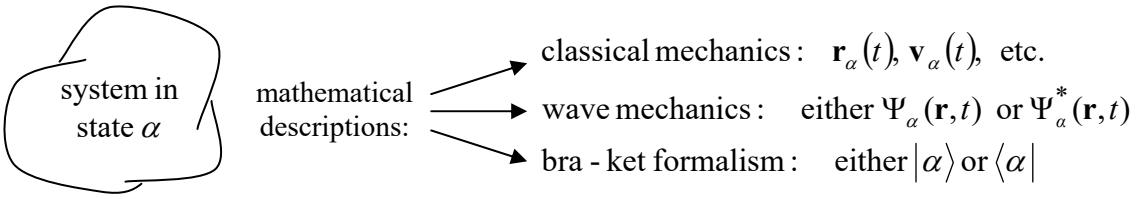


Fig. 4.3. Physical state of a system and its descriptions.

One should be cautious with the term "vector" here. The usual geometric vectors, such as $\mathbf{r}$ and $\mathbf{v}$, are defined in the usual geometric (say, Euclidean) space. In contrast, the bra- and ket-vectors are defined in a more abstract *Hilbert space* – the full set of all possible state vectors of a given system.[8] So, despite certain similarities with the geometric vectors, the bra- and ket-vectors are different mathematical objects, and we need to define the rules of their handling. The primary rules are essentially postulates and are justified only by the correct description of all experimental observations of the rules' corollaries. While there is a general consensus among physicists about what the corollaries are, there are many possible ways to carve from them the different sets of basic postulates. Just as in Sec. 1.2, I will not try too hard to beat the number of the postulates down to the minimum, trying instead to keep their physical meaning transparent.

(i) Ket-vectors. Let us start with *ket-vectors* – sometimes called just *kets* for short. Their most important property is the *linear superposition*. Namely, if several ket-vectors $|\alpha_j\rangle$ describe possible states of a quantum system, numbered by the index j, then any linear combination (*superposition*)

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

Linear superposition of ket-vectors

[5] An attentive reader could notice my smuggling the 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.

[6] As was expressed nicely by Asher Peres, one of the pioneers of the quantum information theory, "quantum phenomena do not occur in the Hilbert space, they occur in a laboratory".

[7] The terms *bra* and *ket* were suggested to reflect the fact that the pair $\langle\beta|$ and $|\alpha\rangle$ may be considered as the parts of the combinations like $\langle\beta \mid \alpha\rangle$ (see below), which remind expressions in the usual angle *brackets*.

[8] I have to confess that this is a bit loose definition; it will be refined soon.

where c_j are any (possibly complex) c-numbers, also describes a possible state of the same system.[9] Actually, since ket-vectors are new mathematical objects, the exact meaning of the right-hand side of Eq. (6) becomes clear only after we have postulated the following rules of summation of these vectors,

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

and their multiplication by an arbitrary c-number:

$$c\left|\alpha_j\right\rangle = \left|\alpha_j\right\rangle c\,. \tag{4.8}$$

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

As Eq. (6) shows, the coefficient c_j may be interpreted as the "weight" of the state α_j in the linear superposition α. One important particular case is $c_j = 0$, showing that the state α_j does not participate in the superposition α. The corresponding term of the sum (6), i.e. the product

Null-state vector

$$\boxed{0\left|\alpha_j\right\rangle}, \tag{4.9}$$

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 the ket-vector $|0\rangle$. In some sense, the null state does not exist at all, while the ground state not only does exist but frequently is the most important quantum state of the system.)

(ii) Bra-vectors and inner products. Bra-vectors $\langle\alpha|$, which obey the rules similar to Eqs. (7) and (8), are not new, independent objects: a ket-vector $|\alpha\rangle$ and the corresponding bra-vector $\langle\alpha|$ describe the same state. In other words, there is a *unique dual correspondence* between $|\alpha\rangle$ and $\langle\alpha|$,[10] very similar (though not identical) to that between a wavefunction Ψ and its complex conjugate Ψ^*.[11] 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

Linear superposition of bra-vectors

$$\boxed{\left\langle\alpha\right| = \sum_j c_j^* \left\langle\alpha_j\right| = \sum_j \left\langle\alpha_j\right| c_j^*}\,. \tag{4.10}$$

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

Short bracket (inner product)

$$\boxed{\left(\left\langle\beta\right|\right)\left(\left|\alpha\right\rangle\right) \equiv \left\langle\beta\middle|\alpha\right\rangle}, \tag{4.11}$$

which is a scalar c-number, in a certain but limited analogy with the scalar product of the usual geometric vectors. (For one difference, the product (11) may be a complex number.) The main property

[9] One may express the same statement by saying that the vector $|\alpha\rangle$ belongs to the same Hilbert space as all $|\alpha_j\rangle$.

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

[11] This analogy is not occasional: we will see very soon that the wavefunction of a quantum state is just a special ("coordinate") representation of its state vector.

of the inner product is its linearity with respect to any of its component vectors. For example, if a linear superposition α is described by the ket-vector (6), then

$$\langle \beta | \alpha \rangle = \sum_j c_j \langle \beta | \alpha_j \rangle , \tag{4.12}$$

while if Eq. (10) is true, then

$$\langle \alpha | \beta \rangle = \sum_j c_j^* \langle \alpha_j | \beta \rangle . \tag{4.13}$$

In plain English, *c*-number factors may be moved either into or out of the inner products.

The second key property of the inner product is

$$\boxed{\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle^* .} \tag{4.14}$$

Inner product:: complex conjugate

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

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

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,

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

State's norm squared

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

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

Hence the role of the bra- and ket-vectors of the same state is very similar to that of complex-conjugate pairs of its wavefunctions.

(iii) Operators. One more key notion of the Dirac formalism is quantum-mechanical *linear operators*. Just as for the operators discussed in wave mechanics, the function of an operator is to "generate" one state from another: if $|\alpha\rangle$ is a possible ket of the system, and $\hat{A}$ is a legitimate[12] operator, then the following combination,

$$\hat{A} | \alpha \rangle , \tag{4.18}$$

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 $|\alpha\rangle$. An alternative formulation of the same rule is the following refinement of the notion of the Hilbert space: for a given set of linear operators of a system, its Hilbert space includes all vectors that may be obtained from each other using the operations of the type (18). In this context, let me note that the operator set, and hence the Hilbert space of a system, usually (if not always) implies its

[12] Here the term "legitimate" means "having a clear sense in the bra-ket formalism". Some examples of "illegitimate" expressions are: $|\alpha\rangle \hat{A}$, $\hat{A} \langle \alpha|$, $|\alpha\rangle|\beta\rangle$, and $\langle \alpha|\langle \beta|$. Note, however, that the last two expressions may be legitimate if α and β are states of different systems, i.e. if their state vectors belong to different Hilbert spaces. We will run into such *direct products* of the bra- and ket-vectors (sometimes denoted, respectively, as $|\alpha\rangle \otimes |\beta\rangle$ and $\langle \alpha| \otimes \langle \beta|$) in Chapters 6-10.

certain approximate model. For example, if the coupling of orbital degrees of freedom of a particle to its spin may be ignored (as it may be for a non-relativistic particle in the absence of an external magnetic field), we may describe the dynamics of the particle using spin operators only. In this case, the set of all possible spin vectors of the particle forms a Hilbert space separate from that of the orbital-state vectors of the same particle.

As the adjective "linear" in the operator definition implies, the main rule governing the operators is their linearity with respect to both any superposition of vectors:

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

and any superposition of operators:

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

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 $\langle\alpha|\hat{A}$ is also legitimate and represents a new bra-vector. It is important that, generally, this vector does *not* represent the same state as the ket-vector (18); instead, the bra-vector isomorphic to the ket-vector (18) is

Conjugate operator

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

This statement serves as the definition of the *Hermitian conjugate* (also called "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

Hermitian operator

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

(This equality, as well as any other operator equation below, means that these operators act similarly on any bra- or ket-vector of the given Hilbert space.) [13]

To proceed further, we need one more additional postulate, sometimes called the *associative axiom of multiplication*: just as an ordinary product of scalars, *any* legitimate bra-ket expression that does not include explicit summations, does not change from an insertion or removal of a pair of parentheses – meaning as usual that the operation inside them has to be performed first. The first two examples of this postulate are given by Eqs. (19) and (20), but the associative axiom is more general and means, for example, that

Long bracket: definition

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

[13] If we consider *c*-numbers as a particular type of operators (which is legitimate for any Hilbert space), then according to Eqs. (11) and (21), for them the Hermitian conjugation is equivalent to the simple complex conjugation, so only real *c*-numbers may be considered as a particular type of Hermitian operators (22).

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:

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

which is most frequently rewritten as

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

Long bracket: complex conjugate

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

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

Outer bra-ket product

In contrast to the inner product (11), which is a *scalar*, this mathematical construct is an *operator*. Indeed, the associative axiom allows us to remove parentheses in the following expression:

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

But the last short bracket 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 the operator's action. Very similarly,

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

– again a typical operator's action on a bra-vector. So, Eq. (26) defines an operator.

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

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

to transform it into the following form:

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

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

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

Outer product: Hermitian conjugate

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

The associative axiom is also valid for the *operator multiplication*:

$$\big(\hat{A}\hat{B}\big)|\alpha\rangle = \hat{A}\big(\hat{B}|\alpha\rangle\big), \qquad \langle\beta|\big(\hat{A}\hat{B}\big) = \big(\langle\beta|\hat{A}\big)\hat{B}, \tag{4.32}$$

showing that the action of an operator product on a state vector is nothing more than the sequential action of its operands. However, we have to be 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

Commutator

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

is a non-trivial and very useful notion. Another similar notion is the *anticommutator*:[14]

Anti-commutator

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

Finally, the bra-ket formalism broadly uses two special operators. The *null operator* $\hat{0}$ is defined by the following relations:

Null operator

$$\hat{0}\left|\alpha\right\rangle \equiv 0\left|\alpha\right\rangle, \qquad \left\langle\alpha\right|\hat{0} \equiv \left\langle\alpha\right|0\,, \tag{4.35}$$

where α is an arbitrary state; we may say that the null operator "kills" *any* state by turning it into the null state. Another useful notion is the *identity operator*, which is defined by the following action (or rather "inaction" :-) on an arbitrary state vector:

Identity operator

$$\hat{I}\left|\alpha\right\rangle \equiv \left|\alpha\right\rangle, \qquad \left\langle\alpha\right|\hat{I} \equiv \left\langle\alpha\right|. \tag{4.36}$$

These definitions show that the null operator and the identity operator are Hermitian.

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, many calculations are performed for quantum systems that feature a *full and orthonormal* set $\{u\} \equiv \{u_1, u_2, \ldots, u_j, \ldots\}$ of its states u_j, frequently called a *basis*. The former of these terms means that any possible state vector of the system (i.e. any vector of its Hilbert space) may be represented as a unique sum of the type (6) or (10) over its basis vectors:

Expansion over state basis

$$\left|\alpha\right\rangle = \sum_j \alpha_j \left|u_j\right\rangle, \qquad \left\langle\alpha\right| = \sum_j \alpha_j^* \left\langle u_j\right|, \tag{4.37}$$

so, in particular, if α is one of the basis states, say $u_{j'}$, then $\alpha_j = \delta_{jj'}$. The latter term means that

Basis vectors: ortho-normality

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

For the systems that may be described by wave mechanics, examples of the full orthonormal bases are represented by any full and orthonormal set of stationary functions calculated in the previous three chapters of this course – for the simplest example, see Eq. (1.87).

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

(i) a quantum state basis may have a large or even infinite number of states u_j, and
(ii) the expansion coefficients α_j may be complex.

[14] Another popular notation for the anticommutator (34) is $\left[\hat{A},\hat{B}\right]_+$; it will not be used in these notes.

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

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

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

$$\boxed{\alpha_j = \langle u_j|\alpha\rangle, \qquad \alpha_j^* = \langle \alpha|u_j\rangle;} \tag{4.40}$$

Expansion coefficients as inner products

these important equalities relations are analogs of equalities $A_j = \mathbf{n}_j \cdot \mathbf{A}$ of the usual vector algebra and will be repeatedly used in this course. With them, the expansions (37) may be rewritten as

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

where

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

Projection operator

Eqs. (41) show that $\hat{\Lambda}_j$ so defined is a legitimate linear operator. This operator, acting on any state vector of the type (37), singles out just one of its components, for example,

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

i.e. "kills" all components of the linear superposition but one. In the geometric analogy, such an operator "projects" the state vector on the 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

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

Closure relation

This means in particular that we may insert the left-hand side of Eq. (44), for any basis, into any bra-ket relation, at any place – the trick that we will use over and over again.

Now let us see how the expansions (37) transform the key notions introduced in the last section, starting with the short bracket (11), i.e. the inner product of two state vectors:

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

Besides the complex conjugation, this expression is similar to the scalar product of the usual, geometric vectors. Now, let us explore the long bracket (23):

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

Here, the last form uses the very important notion of the operator's *matrix elements* defined as

Operator's matrix elements

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

As Eq. (46) shows, the full set of the matrix elements completely characterizes the operator, just as the full set of the expansion coefficients (40) fully characterizes a quantum state. The term "matrix" means, first of all, that it is convenient to represent 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. By the way, this number (which may be infinite) is called the *dimensionality* 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 represented as a matrix of zeros (called the *null matrix*):

Null matrix

$$0 \equiv \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}, \tag{4.48}$$

while for the identity operator $\hat{I}$ defined by Eqs. (36), we readily get

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

i.e. its matrix (naturally called the *identity matrix*) is diagonal – also in any basis:

Identity matrix

$$\mathrm{I} \equiv \begin{pmatrix} 1 & 0 & \ldots \\ 0 & 1 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}. \tag{4.50}$$

The convenience of the matrix language extends well beyond the representation of particular operators. For example, let us use the definition (47) to calculate the matrix elements of a product of two operators:

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

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

Matrix element of an operator product

$$(AB)_{jj''} = \left\langle u_j \middle| \hat{A}\hat{B} \middle| u_{j''} \right\rangle = \left\langle u_j \middle| \hat{A}\hat{I}\hat{B} \middle| u_{j''} \right\rangle = \sum_{j'} \left\langle u_j \middle| \hat{A} \middle| u_{j'} \right\rangle \left\langle u_{j'} \middle| \hat{B} \middle| u_{j''} \right\rangle = \sum_{j'} A_{jj'} B_{j'j''}. \tag{4.52}$$

This result corresponds to the standard "row by column" rule of calculation of an arbitrary element of the matrix product

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

Hence a product of operators may be represented (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 represent not only long brackets,

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

Long bracket as a matrix product

but even short brackets:

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

Short bracket as a matrix product

although these equalities require the use of non-square matrices: rows of (complex-conjugate!) expansion coefficients for the representation of bra-vectors, and columns of these coefficients for the representation of ket-vectors. With that, the mapping of quantum states and operators onto matrices becomes completely general.

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

$$\left(|\alpha\rangle\langle\beta|\right)_{jj'} = \langle u_j|\alpha\rangle\langle\beta|u_{j'}\rangle = \alpha_j \beta_{j'}^*. \tag{4.56}$$

These are the elements of a very special square matrix, whose filling requires the knowledge of just $2N$ scalars (where N is the basis size) rather than N^2 scalars as for an arbitrary operator. However, a simple generalization of such an outer product may represent an arbitrary operator. Indeed, let us insert two identity operators (44), with different summation indices, on both sides of an arbitrary 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 then use the associative axiom to rewrite this expression as

$$\hat{A} = \sum_{j,j'} |u_j\rangle\left(\langle u_j|\hat{A}|u_{j'}\rangle\right)\langle u_{j'}|. \tag{4.58}$$

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

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

Operator via its matrix elements

The reader should agree that this formula, which is a natural generalization of Eq. (44), is extremely elegant.

The matrix representation is so convenient that it makes sense to extend it to one level lower – from the state vector products to the "bare" state vectors resulting from the operator's action upon a given state. For example, let us use Eq. (59) to represent the ket-vector (18) as

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

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

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

But the expression in the parentheses is just the coefficient α'_j of the expansion (37) of the resulting ket-vector (60) in the same basis, so

$$\alpha'_j = \sum_{j'} A_{jj'} \alpha_{j'} . \tag{4.62}$$

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

$$\begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \dots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix} . \tag{4.63}$$

Absolutely similarly, the operator action on the bra-vector (21), represented by its row matrix, is

$$\left(\alpha'^*_1, \alpha'^*_2, \dots\right) = \left(\alpha^*_1, \alpha^*_2, \dots\right) \begin{pmatrix} \left(A^\dagger\right)_{11} & \left(A^\dagger\right)_{12} & \dots \\ \left(A^\dagger\right)_{21} & \left(A^\dagger\right)_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} . \tag{4.64}$$

By the way, Eq. (64) naturally raises the following question: what are the elements of the matrix on its right-hand side, or more exactly, what is the relation between the matrix elements of an operator and its Hermitian conjugate? The simplest way to answer it is to use Eq. (25) with two arbitrary states (say, u_j and $u_{j'}$) of the same basis in the role of α and β. Together with the orthonormality relation (38), this immediately gives[15]

Hermitian conjugate: matrix elements

$$\boxed{\left(\hat{A}^\dagger\right)_{jj'} = \left(A_{j'j}\right)^* .} \tag{4.65}$$

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 difference between Hermitian and complex conjugation. It also shows that for the Hermitian operators defined by Eq. (22),

$$A_{jj'} = A^*_{j'j} , \tag{4.66}$$

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

$$A_{jj} = A^*_{jj} , \quad \text{i.e. } \mathrm{Im}\, A_{jj} = 0. \tag{4.67}$$

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

In order to fully appreciate the special role played by Hermitian operators in 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 $\hat{A}$, both defined by the equation they have to satisfy:[16]

$$\hat{A}\left|a_j\right\rangle = A_j\left|a_j\right\rangle. \tag{4.68}$$

Operator: eigenstates and eigenvalues

Let us prove that the eigenvalues of any Hermitian operator are real,[17]

$$A_j = A_j^*, \quad \text{for } j = 1, 2, ..., N, \tag{4.69}$$

Hermitian operator: eigenvalues

while the eigenstates corresponding to different eigenvalues are orthogonal:

$$\left\langle a_j \middle| a_{j'} \right\rangle = 0, \quad \text{if } A_j \neq A_{j'}. \tag{4.70}$$

Hermitian operator: eigenvectors

The proof of both statements is surprisingly simple. Let us inner-multiply both sides of Eq. (68) by the bra-vector $\langle a_{j'}|$. On the right-hand side of the result, the eigenvalue A_j, as a *c*-number, may be taken out of the bracket, giving

$$\left\langle a_{j'} \middle| \hat{A} \middle| a_j \right\rangle = A_j \left\langle a_{j'} \middle| a_j \right\rangle. \tag{4.71}$$

This equality has to hold for any pair of eigenstates, so we may swap the indices *j* and *j'* in Eq. (71), and write the complex-conjugate of the result:

$$\left\langle a_j \middle| \hat{A} \middle| a_{j'} \right\rangle^* = A_{j'}^* \left\langle a_j \middle| a_{j'} \right\rangle^*. \tag{4.72}$$

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

$$\left\langle a_{j'} \middle| \hat{A} \middle| a_j \right\rangle = A_{j'}^* \left\langle a_{j'} \middle| a_j \right\rangle. \tag{4.73}$$

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

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

There are two possibilities to satisfy this relation. If the indices *j* and *j'* are equal (denote the same eigenstate), then the bracket is the state's norm squared, and cannot be equal to zero. In this case, the left parentheses (with $j = j'$) have to be zero, proving Eq. (69). On the other hand, if *j* and *j'* correspond to different eigenvalues of *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, in particular, for the description of physical observables.

[16] 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. (68) for the Hamiltonian as the operator of energy.

[17] 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 $A_{kk'} = A_k\delta_{kk'}$, with $A_k^* = A_k$.

4.4. Change of basis, and matrix diagonalization

From the discussion of the last section, it may look like the matrix language is fully similar to, and in many instances more convenient than the general bra-ket formalism. In particular, Eqs. (54)-(55) and (63)-(64) show that any part of any bra-ket expression may be directly mapped onto 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 representation. This invites the question: why do we need the bra-ket language at all? The answer is that the matrix elements depend on the particular choice of the basis set, very much like the Cartesian components of a usual geometric vector depend on the particular choice of reference frame orientation (Fig. 4), and very frequently, at problem solution, it is convenient to use two or more different basis sets for the same system. (Just a bit more patience – numerous examples will follow soon.)

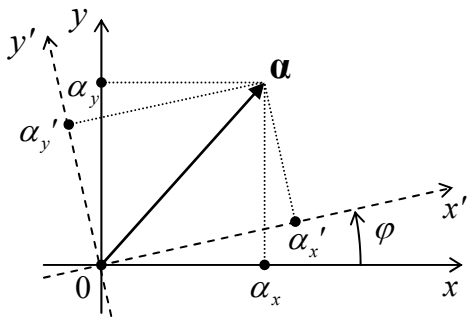


Fig. 4.4. The transformation of components of a 2D vector at a reference frame's rotation.

With this motivation, let us explore what happens at the transform 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, and an arbitrary numbering of the states of each base, there exists such an operator $\hat{U}$ that, first,

Unitary operator: definition

$$\left|v_j\right\rangle = \hat{U}\left|u_j\right\rangle, \tag{4.75}$$

and, second,

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}. \tag{4.76}$$

(Due to the last property,[18] $\hat{U}$ is called a *unitary operator*, and Eq. (75), a *unitary transformation*.)

A very simple proof of both statements may be achieved by construction. Indeed, let us take

Unitary operator: construction

$$\hat{U} \equiv \sum_{j'} \left|v_{j'}\right\rangle\left\langle u_{j'}\right|, \tag{4.77}$$

- an evident generalization of Eq. (44). Then, using Eq. (38), we obtain

$$\hat{U}\left|u_j\right\rangle = \sum_{j'} \left|v_{j'}\right\rangle\left\langle u_{j'}\middle|u_j\right\rangle = \sum_{j'} \left|v_{j'}\right\rangle \delta_{j'j} = \left|v_j\right\rangle, \tag{4.78}$$

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

Unitary operator: conjugate

$$\hat{U}^\dagger \equiv \sum_{j'} \left|u_{j'}\right\rangle\left\langle v_{j'}\right|, \tag{4.79}$$

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

so

$$\hat{U}\hat{U}^{\dagger} = \sum_{j,j'} |v_j\rangle\langle u_j|u_{j'}\rangle\langle v_{j'}| = \sum_{j,j'} |v_j\rangle\delta_{jj'}\langle v_{j'}| = \sum_j |v_j\rangle\langle v_j|. \quad (4.80)$$

But according to the closure relation (44), the last expression is just the identity operator, so one of Eqs. (76) has been proved. (The proof of the second equality is absolutely similar.) As a by-product of our proof, we have also got another important expression – Eq. (79). It implies, in particular, that while, according to Eq. (75), the operator $\hat{U}$ performs the transform from the "old" basis $\{u\}$ to the "new" basis $\{v\}$, its Hermitian adjoint $\hat{U}^{\dagger}$ performs the reciprocal transform:

$$\hat{U}^{\dagger}|v_j\rangle = \sum_{j'} |u_{j'}\rangle\delta_{j'j} = |u_j\rangle. \quad (4.81)$$

Reciprocal basis transform

Now let us see what the matrix elements of the unitary transform operators look like. Generally, as was discussed above, the operator's elements may depend on the basis we calculate them in, so let us be specific – at least initially. For example, let us calculate the desired matrix elements $U_{jj'}$ in the "old" basis $\{u\}$, by using Eq. (77):

$$U_{jj'}|_{\text{in } u} \equiv \langle u_j|\hat{U}|u_{j'}\rangle = \langle u_j|\left(\sum_{j''} |v_{j''}\rangle\langle u_{j''}|\right)|u_{j'}\rangle = \langle u_j|\sum_{j''} |v_{j''}\rangle\delta_{j''j'} = \langle u_j|v_{j'}\rangle. \quad (4.82)$$

Now performing a similar calculation in the "new" basis $\{v\}$, we get

$$U_{jj'}|_{\text{in } v} \equiv \langle v_j|\hat{U}|v_{j'}\rangle = \langle v_j|\left(\sum_{j''} |v_{j''}\rangle\langle u_{j''}|\right)|v_{j'}\rangle = \sum_{j''} \delta_{jj''}\langle u_{j''}\|v_{j'}\rangle = \langle u_j|v_{j'}\rangle. \quad (4.83)$$

Surprisingly, the result is the same! This is of course true for the Hermitian conjugate (79) as well:

$$U^{\dagger}_{jj'}|_{\text{in } u} = U^{\dagger}_{jj'}|_{\text{in } v} = \langle v_j|u_{j'}\rangle. \quad (4.84)$$

These expressions may be used, first of all, to rewrite Eq. (75) in a purely matrix form. Applying the first of Eqs. (41) to any state $v_{j'}$ of the "new" basis, and then Eq. (82), we get

$$|v_{j'}\rangle = \sum_j |u_j\rangle\langle u_j|v_{j'}\rangle = \sum_j U_{jj'}|u_j\rangle. \quad (4.85)$$

Basis transforms: matrix form

Similarly, the reciprocal transform is

$$|u_{j'}\rangle = \sum_j |v_j\rangle\langle v_j|u_{j'}\rangle = \sum_j U^{\dagger}_{jj'}|v_j\rangle. \quad (4.86)$$

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

Next, we may use Eqs. (83)-(84) to express the effect of the unitary transform on the expansion coefficients α_j of the vectors of an *arbitrary* state α, defined by Eq. (37). As a reminder, in the "old" basis $\{u\}$ they are given by Eqs. (40). Similarly, in the "new" basis $\{v\}$,

$$\alpha_j|_{\text{in } v} = \langle v_j|\alpha\rangle. \quad (4.87)$$

Again inserting the identity operator in its closure form (44) with the internal index j', and then using Eqs. (84) and (40), we get

$$\alpha_j\big|_{\text{in }v} = \langle v_j|\left(\sum_{j'}|u_{j'}\rangle\langle u_{j'}|\right)|\alpha\rangle = \sum_{j'}\langle v_j|u_{j'}\rangle\langle u_{j'}|\alpha\rangle = \sum_{j'}U^{\dagger}_{jj'}\langle u_{j'}|\alpha\rangle = \sum_{j'}U^{\dagger}_{jj'}\alpha_{j'}\big|_{\text{in }u}\,. \quad (4.88)$$

The reciprocal transform is performed by matrix elements of the operator $\hat{U}$:

$$\alpha_j\big|_{\text{in }u} = \sum_{j'}U_{jj'}\alpha_{j'}\big|_{\text{in }v}\,. \quad (4.89)$$

Per Eqs. (82)-(84), the matrix elements $U_{jj'}$ and $U^{\dagger}_{jj'}$ are the same in the bases $\{u\}$ and $\{v\}$, so Eqs. (88)-(89) may be rewritten in the following compact matrix form:

$$|\alpha\rangle_{\text{in }v} = \mathrm{U}^{\dagger}|\alpha\rangle_{\text{in }u}, \qquad |\alpha\rangle_{\text{in }u} = \mathrm{U}|\alpha\rangle_{\text{in }v}, \quad (4.90)$$

even though the reader should remember that these relations are different from the usual matrix formulas, which use the same basis for all its components.

So, if the transform (75) from the "old" basis $\{u\}$ to the "new" basis $\{v\}$ is performed by a unitary operator, the change (88) of state vector components at this transformation requires its Hermitian conjugate. This fact is similar to the transformation of components of a usual vector at coordinate frame rotation. For example, for a 2D vector whose actual position in space is fixed (Fig. 4):

$$\begin{pmatrix}\alpha_x'\\ \alpha_y'\end{pmatrix} = \begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}\alpha_x\\ \alpha_y\end{pmatrix}, \quad (4.91)$$

but the reciprocal transform is performed by a different matrix, which may be obtained from that participating in Eq. (91) by the replacement $\varphi \to -\varphi$. This replacement has a clear geometric sense: if the "new" reference frame $\{x', y'\}$ is obtained from the "old" frame $\{x, y\}$ by a *counterclockwise* rotation by angle φ, the reciprocal transformation requires such rotation with angle $-\varphi$. (In this analogy, the unitary property (76) of the unitary transform operators corresponds to the equality of the determinants of both rotation matrices to 1.)

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

$$A_{jj'}\big|_{\text{in }v} \equiv \langle v_j|\hat{A}|v_{j'}\rangle = \langle v_j|\left(\sum_{k}|u_k\rangle\langle u_k|\right)\hat{A}\left(\sum_{k'}|u_{k'}\rangle\langle u_{k'}|\right)|v_{j'}\rangle = \sum_{k,k'}U^{\dagger}_{jk}A_{kk'}\big|_{\text{in }u}U_{k'j'}\,; \quad (4.92)$$

Matrix elements' transforms

absolutely similarly, we may also get

$$A_{jj'}\big|_{\text{in }u} \equiv \sum_{k,k'}U_{jk}A_{kk'}\big|_{\text{in }v}U^{\dagger}_{k'j'}\,. \quad (4.93)$$

In the spirit of Eq. (90), we may represent these results in the similar matrix form:

$$\mathrm{A}\big|_{\text{in }v} = \mathrm{U}^{\dagger}\mathrm{A}\big|_{\text{in }u}\mathrm{U}, \qquad \mathrm{A}\big|_{\text{in }u} = \mathrm{U}\mathrm{A}\big|_{\text{in }v}\mathrm{U}^{\dagger}, \quad (4.94)$$

where, again, the matrix elements of U and $\mathrm{U}^{\dagger}$ may be calculated in any of the bases $\{u\}$ and $\{v\}$ – but not in an arbitrary basis!

As a sanity check, let us apply Eq. (93) to the identity operator:

$$\hat{I}\big|_{\text{in } v} = \left(\hat{U}^{\dagger}\hat{I}\hat{U}\right)_{\text{in } u} = \left(\hat{U}^{\dagger}\hat{U}\right)_{\text{in } u} = \hat{I}\big|_{\text{in } u} \tag{4.95}$$

– just 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:

$$\boxed{\text{Tr } \hat{A} \equiv \text{Tr A} \equiv \sum_j A_{jj} .} \tag{4.96}$$

Operator/ matrix: trace

The (easy) proof of this fact, using previous relations, is left for the reader's exercise.

So far, I have implied that both state bases $\{u\}$ and $\{v\}$ are known, and the natural question is where this information comes from in the quantum mechanics of actual physical systems. To get a partial answer to this question, let us return to Eq. (68), which defines the eigenstates and the 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 calculate the matrix elements of the operator in the basis $\{a\}$ of these states, at their arbitrary numbering. For that, it is sufficient to inner-multiply both sides of Eq. (68), written for some eigenstate $a_{j'}$, by the bra-vector of an arbitrary state a_j of the same set:

$$\left\langle a_j \middle| \hat{A} \middle| a_{j'} \right\rangle = \left\langle a_j \middle| A_{j'} \middle| a_{j'} \right\rangle . \tag{4.97}$$

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

$$\boxed{A_{jj'} = A_j \delta_{jj'} .} \tag{4.98}$$

Matrix elements in eigenstate basis

In particular, in the eigenstate basis (but not necessarily in an arbitrary basis!), A_{jj} means the same as A_j. Thus the important problem of finding the eigenvalues and eigenstates of an operator is equivalent to the *diagonalization* of its matrix,[19] i.e. finding the basis in which the operator's matrix acquires the diagonal form (98); then the diagonal elements are the eigenvalues, and the basis itself is the desirable set of eigenstates.

To see how this is done in practice, let us inner-multiply Eq. (68) by a bra-vector of the basis (say, $\{u\}$) in that we have happened to know the matrix elements $A_{jj'}$:

$$\left\langle u_k \middle| \hat{A} \middle| a_j \right\rangle = \left\langle u_k \middle| A_j \middle| a_j \right\rangle . \tag{4.99}$$

On the left-hand side, we can (as usual :-) insert the identity operator between the operator $\hat{A}$ and the ket-vector, and then use the closure relation (44) in the same basis $\{u\}$, while on the right-hand side, we can move the eigenvalue A_j (a c-number) out of the bracket, and then insert a summation over the same index as in the closure, compensating it with the proper Kronecker delta symbol:

$$\left\langle u_k \middle| \hat{A} \sum_{k'} \middle| u_{k'} \right\rangle \left\langle u_{k'} \middle| a_j \right\rangle = A_j \sum_{k'} \left\langle u_{k'} \middle| a_j \right\rangle \delta_{kk'} . \tag{4.100}$$

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

[19] Note that the expression "matrix diagonalization" is a very common but dangerous jargon. Formally, a matrix is just a table, an ordered set of c-numbers, and cannot be "diagonalized". It is OK to use this jargon (I will do this) if you remember clearly what it actually means – see the definition above.

$$\sum_{k'} \left(A_{kk'} - A_j \delta_{kk'}\right) \langle u_{k'} | a_j \rangle = 0 . \tag{4.101}$$

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

Matrix diagonalization

$$\begin{pmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} U_{1j} \\ U_{2j} \\ \dots \end{pmatrix} = 0 , \tag{4.102}$$

and the condition of its consistency,

Characteristic equation for eigenvalues

$$\begin{vmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{vmatrix} = 0, \tag{4.103}$$

plays the role of the characteristic equation of the system. This equation has N roots A_j – the eigenvalues of the operator $\hat{A}$; after they have been calculated, plugging any of them back into the system (102), we can use it to find N matrix elements U_{kj} (k = 1, 2, …N) corresponding to this particular eigenvalue. However, since the equations (102) are homogeneous, they allow finding U_{kj} only to a constant multiplier. To ensure their normalization, i.e. enforce the unitary character of the matrix U, we may use the requirement for all eigenvectors to be normalized (just as the basis vectors are):

$$\langle a_j | a_j \rangle \equiv \sum_k \langle a_j | u_k \rangle \langle u_k | a_j \rangle \equiv \sum_k \left| U_{kj} \right|^2 = 1, \tag{4.104}$$

for each j. This normalization completes the diagonalization.[20]

Now (at last!) I can give the reader some examples. As a simple but very important case, let us diagonalize each of the operators described (in a certain two-function basis $\{u\}$, i.e. in two-dimensional Hilbert space) by the so-called *Pauli matrices*

Pauli matrices

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.105}$$

Though introduced by a physicist, with a specific purpose to describe the electron's spin, these matrices have a general mathematical significance, because together with the 2×2 identity matrix, they provide a full, linearly-independent system – meaning that an arbitrary 2×2 matrix may be represented as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = b\mathrm{I} + c_x \sigma_x + c_y \sigma_y + c_z \sigma_z , \tag{4.106}$$

[20] A possible slight complication here is that the characteristic equation may give equal eigenvalues for certain groups of different eigenvectors. In such cases, the requirement of the mutual orthogonality of these *degenerate states* should be additionally enforced.

with a unique set of four c-number coefficients b, c_x, c_y, and c_z.

Since the matrix σ_z is already diagonal, with the evident eigenvalues ±1, let us start with diagonalizing the matrix σ_x. For it, the characteristic equation (103) is evidently

$$\begin{vmatrix} -A_j & 1 \\ 1 & -A_j \end{vmatrix} = 0, \qquad \text{i.e. } A_j^2 - 1 = 0, \tag{4.107}$$

and has two roots, $A_{1,2} = \pm 1$. (Again, the state numbering is arbitrary!) So the eigenvalues of the matrix σ_x are the same as those of the matrix σ_z. (The reader may readily check that the eigenvalues of the matrix σ_y are also the same.) However, the eigenvectors of the operators corresponding to these three matrices are different. To find them for σ_x, let us plug its first eigenvalue, $A_1 = +1$, back into equations (101) spelled out for this particular case ($j = 1$; $k, k' = 1,2$):

$$\begin{aligned} -\langle u_1 | a_1 \rangle + \langle u_2 | a_1 \rangle = 0, \\ \langle u_1 | a_1 \rangle - \langle u_2 | a_1 \rangle = 0. \end{aligned} \tag{4.108}$$

These two equations are compatible (of course, because the used eigenvalue $A_1 = +1$ satisfies the characteristic equation), and any of them gives

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

With that, the normalization condition (104) yields

$$|U_{11}|^2 = |U_{21}|^2 = \frac{1}{2}. \tag{4.110}$$

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

$$U_{11} = U_{21} = \frac{1}{\sqrt{2}}. \tag{4.111}$$

Performing an absolutely similar calculation for the second characteristic value, $A_2 = -1$, we get $U_{12} = -U_{22}$, and we may choose the common phase to have

$$U_{12} = -U_{22} = \frac{1}{\sqrt{2}}, \tag{4.112}$$

so the whole unitary matrix for diagonalization of the operator corresponding to σ_x is[21]

$$\mathrm{U}_x = \mathrm{U}_x^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{4.113}$$

Unitary matrix diagonalizing σ_x

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

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

[21] Though this particular unitary matrix U_x is Hermitian, this is not true for an arbitrary choice of the phases φ.

$$|u_1\rangle = U_{11}^{\dagger}|a_1\rangle + U_{21}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle + |a_2\rangle\right), \qquad |u_2\rangle = U_{12}^{\dagger}|a_1\rangle + U_{22}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle - |a_2\rangle\right). \quad (4.114b)$$

Now let me show that these results are already sufficient to understand the Stern-Gerlach experiments described in Sec. 1 – but with two additional postulates. The first of them is that the interaction of a particle with the external magnetic field, besides that due to its orbital motion, may be described by the following *operator vector* of its spin dipole magnetic moment:[22]

Spin magnetic moment

$$\hat{\mathbf{m}} = \gamma \hat{\mathbf{S}}, \quad (4.115a)$$

where the constant coefficient γ, specific for every particle type, is called the *gyromagnetic ratio*,[23] and $\hat{\mathbf{S}}$ is the *operator vector*[24] *of spin*, with three Cartesian components:

Spin vector operator

$$\hat{\mathbf{S}} = \mathbf{n}_x \hat{S}_x + \mathbf{n}_y \hat{S}_y + \mathbf{n}_z \hat{S}_z. \quad (4.115b)$$

Here $\mathbf{n}_{x,y,z}$ are the usual Cartesian unit vectors in the 3D geometric space (in the quantum-mechanics sense, they are just *c*-numbers, or rather "*c*-vectors"), while $\hat{S}_{x,y,z}$ are the "usual" (scalar) operators. For the so-called *spin-½ particles* (including the electron),[25] these components may be simply, as

$$\hat{S}_{x,y,z} = \frac{\hbar}{2}\hat{\sigma}_{x,y,z}, \quad (4.116a)$$

Pauli vector

expressed via those of the *Pauli vector* $\hat{\boldsymbol{\sigma}} \equiv \mathbf{n}_x \hat{\sigma}_x + \mathbf{n}_y \hat{\sigma}_y + \mathbf{n}_z \hat{\sigma}_z$, so we may also write

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}. \quad (4.116b)$$

In turn, in the so-called *z-basis*, each Cartesian component of the latter operator is just the corresponding Pauli matrix (105), so it may be also convenient to use the following 3D vector of these matrices:[26]

Pauli vector's matrix

$$\boldsymbol{\sigma} \equiv \mathbf{n}_x \sigma_x + \mathbf{n}_y \sigma_y + \mathbf{n}_z \sigma_z \equiv \begin{pmatrix} \mathbf{n}_z & \mathbf{n}_x - i\mathbf{n}_y \\ \mathbf{n}_x + i\mathbf{n}_y & -\mathbf{n}_z \end{pmatrix}. \quad (4.117)$$

The *z*-basis, in which such matrix representation of $\hat{\boldsymbol{\sigma}}$ is valid, is *defined* as an orthonormal basis of certain two states, commonly denoted ↑ ("spin up") an ↓ ("spin down"). In this basis, the matrix of the operator $\hat{\sigma}_z$ is diagonal, with eigenvalues, respectively, + 1 and –1, and hence the matrix $S_z \equiv (\hbar/2)\sigma_z$ of $\hat{S}_z$ is also diagonal with the eigenvalues $+\hbar/2$ and $-\hbar/2$ – see the last of Eqs. (105). Note that

[22] This was the key point in the electron spin's description, developed by W. Pauli in 1925-1927.

[23] For the 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 electron's dimensionless *g-factor*. Due to quantum-electrodynamic (relativistic) effects, this *g*-factor is slightly higher than 2: $g_e = 2(1 + \alpha/2\pi + \ldots) \approx 2.002319304\ldots$, where $\alpha \equiv e^2/4\pi\varepsilon_0\hbar c \equiv (E_H/m_e c^2)^{1/2} \approx 1/137$ is the so-called *fine structure constant*. (The origin of its name will be clear from the discussion in Sec. 6.3.)

[24] The basic rule of dealing with operator vectors is to perform all vector operations just as with the usual geometric vectors. (The vector ∇ is a good example – see the formulas for in MA Secs. 8-12.)

[25] At this point, the adjective "spin-½ " should be understood as just a name. The physical sense of this term and the generalization of the theory to other values of spin will be discussed in Sec. 5.7.

[26] Note that is some texts, the term "Pauli vector" is used for this matrix $\boldsymbol{\sigma}$ rather than for the operator $\hat{\boldsymbol{\sigma}}$.

we do not "understand" what exactly the states ↑ and ↓ are,[27] but loosely associate them with some internal rotation of a spin-½ particle about the z-axis, with either positive or negative angular momentum component S_z. However, attempts to use such classical interpretation for quantitative predictions run into fundamental difficulties – see Sec. 6 below.

The second necessary 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 the result of its measurement,[28] in a quantum state α described by a linear superposition of the eigenstates a_j of the operator,

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

may be only one of the corresponding eigenvalues A_j.[29] Specifically, if the ket (118) and all eigenkets $|a_j\rangle$ are normalized to 1,

$$\langle\alpha|\alpha\rangle = 1, \qquad \langle a_j|a_j\rangle = 1, \tag{4.119}$$

then the probability of a certain measurement outcome A_j is[30]

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

Quantum measurement postulate

This relation is evidently a generalization of Eq. (1.22) in wave mechanics. As a sanity check, let us assume that the set of the eigenstates a_j is full, and calculate the sum of the probabilities to find the system in each of these states:

$$\sum_j W_j = \sum_j \langle\alpha|a_j\rangle\langle a_j|\alpha\rangle = \langle\alpha|\hat{I}|\alpha\rangle = 1. \tag{4.121}$$

Now returning to the Stern-Gerlach experiment, conceptually the description of the first (z-oriented) experiment shown in Fig. 1 is formally the hardest for us, because the statistical ensemble describing the unpolarized particle 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 mixed ensembles in Chapter 7.) However, it is intuitively clear that its results are compatible with the description of the two output beams as sets of particles in the pure states ↑ and ↓, respectively. The absorber following that first stage (Fig. 2) just takes all spin-down particles out of the picture, producing an output beam of polarized particles in the definite ↑ state. For such a beam, the

[27] If you think about it, the word "understand" typically means that we can express a new notion in terms of those discussed earlier and thus considered "known". (For example, in our current case, we cannot describe the spin states by any wavefunction $\psi(\mathbf{r})$, or any other mathematical notion discussed in the previous three chapters and hence considered "known".) The bra-ket formalism was invented exactly to enable mathematical analyses of such "new" quantum states we do not initially "understand". Gradually, as we learn more and more about their properties and get *accustomed* to these notions, we start treating them as "known" ones.

[28] Here again, just like in Sec. 1.2, the statement implies the abstract notion of "ideal experiments", deferring the discussion of real (physical) measurements until Chapter 10.

[29] As a reminder, at the end of Sec. 3 we have already proved that such eigenstates corresponding to different values A_j are orthogonal. If any of these values is degenerate, i.e. corresponds to several different eigenstates, they should be also selected orthogonal, in order for Eq. (118) to be valid.

[30] This relation, in particular, explains the most common term for the (generally, complex) coefficients α_j, which was already mentioned several times earlier: the *probability amplitudes*.

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 represent the z-polarized beam in another basis – of the two states (I will denote them as $\rightarrow$ and $\leftarrow$) in that, by definition, the matrix S_x is diagonal. But this is exactly the set we called $a_{1,2}$ in the σ_x matrix diagonalization problem solved above. On the other hand, the states $\uparrow$ and $\downarrow$ are exactly what we called $u_{1,2}$ in that problem because in this basis, we know the matrix σ explicitly – see Eq. (117). Hence, in the application to the particle spin problem, we may rewrite Eqs. (114) as

Relation between eigenvectors of S_x and S_z

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

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

Currently for us the first of Eqs. (123) is most important, because it shows that the quantum state of particles entering the SG (x) stage may be represented as a coherent superposition of particles with $S_x = +\hbar/2$ and $S_x = -\hbar/2$. Notice that the beams have equal probability amplitude moduli, so according to Eq. (120), the split beams $\rightarrow$ and $\leftarrow$ have equal intensities, in accordance with experimental results.

Now, let us discuss the most mysterious (from the classical point of view) multistage SG experiment shown on the middle panel of Fig. 2. After the second absorber has taken out all particles in, say, the $\leftarrow$ state, the remaining particles, all in the state $\rightarrow$, are passed to the final, SG (z), stage. But according to the first of Eqs. (122), this state may be represented as a (coherent) linear superposition of the $\uparrow$ and $\downarrow$ states, with equal probability amplitudes. The final stage separates particles in these two states into separate beams, with equal probabilities $W_\uparrow = W_\downarrow = ½$ to find an particle 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. Indeed, let a plane electromagnetic wave propagate normally to the plane of the drawing in Fig. 5, and pass through the linear polarizer 1.

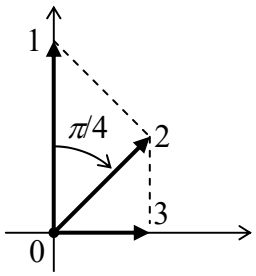


Fig. 4.5. A light polarization sequence similar to the three-stage Stern-Gerlach experiment shown on the middle panel of Fig. 2.

Similarly to the output of the initial SG (z) stages (including the absorbers) shown in Fig. 2, the output wave is linearly polarized in one direction – the vertical direction in Fig. 5. Now its electric field vector has no horizontal component – as may be revealed by the 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 the angles between the polarization directions 1 and 2, and between 2 and 3, are both equal to $\pi/4$, each polarizer reduces the wave amplitude by a factor of $\sqrt{2}$, and hence the intensity by a factor of 2, exactly like in the multistage SG experiment, with the polarizer 2 playing the role of the SG (x) stage. The "only" difference is that the necessary angle between the polarizer orientations is $\pi/4$, rather than $\pi/2$ for the Stern-Gerlach experiment. In quantum electrodynamics (see Chapter 9 below), which confirms classical predictions for this experiment, this difference may be explained by that between the *integer* spin of electromagnetic field quanta (photons) and the *half-integer* spin of electrons.

4.5. Observables: Expectation values and uncertainties

After this particular (and hopefully 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 a coherent one) may be always calculated using the general statistical rule (1.37). For the particular case of a coherent superposition (118), we can combine that rule with Eq. (120) and the second of Eqs. (118):

$$\langle A\rangle_\alpha = \sum_j A_j W_j = \sum_j \alpha_j^* A_j \alpha_j = \sum_j \langle \alpha | a_j \rangle A_j \langle a_j | \alpha \rangle \equiv \langle \alpha | \left(\sum_j |a_j\rangle A_j \langle a_j| \right) |\alpha\rangle . \tag{4.124}$$

Now using Eq. (59) for the particular case of the eigenstate basis $\{a\}$, for which Eq. (98) is valid, we arrive at a very simple and important formula[31]

$$\boxed{\langle A\rangle_\alpha = \langle \alpha | \hat{A} | \alpha \rangle .} \tag{4.125}$$

Expectation value as a long bracket

This is a clear analog of the wave-mechanics formula (1.23) – and as we will see soon, may be used to derive it.[32] A great convenience 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.

For example, let us consider an arbitrary coherent state α of spin-½,[33] and calculate the expectation values of its components. The calculations are easier in the z-basis because we know the matrix elements of the spin operator components in that basis. Representing the ket- and bra-vectors of the given state as linear superpositions of the corresponding vectors of the basis states $\uparrow$ and $\downarrow$,

$$|\alpha\rangle = \alpha_\uparrow |\uparrow\rangle + \alpha_\downarrow |\downarrow\rangle, \qquad \langle\alpha| = \langle\uparrow|\alpha_\uparrow^* + \langle\downarrow|\alpha_\downarrow^* . \tag{4.126}$$

and plugging these expressions into Eq. (125) written for the observable S_z, we get

---

[31] This equality reveals the full beauty of Dirac's notation. Indeed, initially in this chapter, 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!

[32] Note also that Eq. (120) may be rewritten in a form similar to Eq. (125): $W_j = \langle \alpha | \hat{\Lambda}_j | \alpha \rangle$, where $\hat{\Lambda}_j$ is the operator (42) of the state's projection upon the j^{th} eigenstate a_j.

[33] For clarity, the noun "spin-½" is used, here and below, to denote the spin degree of freedom of a spin-½ particle, independent of its orbital motion.

$$\begin{aligned}\langle S_z\rangle &= \left(\langle\uparrow|\alpha_\uparrow^* + \langle\downarrow|\alpha_\downarrow^*\right)\hat{S}_z\left(\alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle\right)\\ &= \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_z|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_z|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_z|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_z|\downarrow\rangle.\end{aligned} \tag{4.127}$$

Now there are two equivalent ways (both very simple) to calculate the long brackets in this expression. The first one is to represent each of them in the matrix form in the *z*-basis, in which the bra- and ket-vectors of states ↑ and ↓ are the matrix rows (1, 0) and (0, 1), or similar matrix columns – the exercise highly recommended to the reader. Another (perhaps more elegant) way is to use the general Eq. (59), in the *z*-basis, together with the spin-½-specific Eqs. (116a) and (105) to write

Spin-½ component operators

$$\hat{S}_x = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right), \quad \hat{S}_y = -i\frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|\right), \quad \hat{S}_z = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right). \tag{4.128}$$

For our particular calculation, we may plug the last of these expressions into Eq. (127), and use the orthonormality conditions (38):

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1, \quad \langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0. \tag{4.129}$$

Both approaches give (of course) the same result:

$$\langle S_z\rangle = \frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\uparrow^* - \alpha_\downarrow\alpha_\downarrow^*\right). \tag{4.130}$$

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

$$\langle S_z\rangle = W_\uparrow\left(+\frac{\hbar}{2}\right) + W_\downarrow\left(-\frac{\hbar}{2}\right) = \alpha_\uparrow\alpha_\uparrow^*\left(+\frac{\hbar}{2}\right) + \alpha_\downarrow\alpha_\downarrow^*\left(-\frac{\hbar}{2}\right). \tag{4.131}$$

The formal way (127), based on the general Eq. (125), has, however, the advantage of being applicable to finding the observables whose operators are *not* diagonal in the *z*-basis, as well. In particular, absolutely similar calculations give

$$\langle S_x\rangle = \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_x|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_x|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_x|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_x|\downarrow\rangle = \frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\downarrow^* + \alpha_\downarrow\alpha_\uparrow^*\right), \tag{4.132}$$

$$\langle S_y\rangle = \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_y|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_y|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_y|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_y|\downarrow\rangle = i\frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\downarrow^* - \alpha_\downarrow\alpha_\uparrow^*\right), \tag{4.133}$$

Let us have a good look at a particular spin state, for example the spin-up state ↑. According to Eq. (126), in this state $\alpha_\uparrow = 1$ and $\alpha_\downarrow = 0$, so Eqs. (130)-(133) yield:

$$\langle S_z\rangle = \frac{\hbar}{2}, \qquad \langle S_x\rangle = \langle S_y\rangle = 0. \tag{4.134}$$

Now let us use the same Eq. (125) to calculate the spin component uncertainties. According to Eqs. (105) and (116)-(117), the operator of each spin component squared is equal to $(\hbar/2)^2\hat{I}$, so the general Eq. (1.33) yields

$$(\delta S_z)^2 = \langle S_z^2 \rangle - \langle S_z \rangle^2 = \langle \uparrow | \hat{S}_z^2 | \uparrow \rangle - \left(\frac{\hbar}{2}\right)^2 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle - \left(\frac{\hbar}{2}\right)^2 = 0, \tag{4.135a}$$

$$(\delta S_x)^2 = \langle S_x^2 \rangle - \langle S_x \rangle^2 = \langle \uparrow | \hat{S}_x^2 | \uparrow \rangle - 0 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle = \left(\frac{\hbar}{2}\right)^2, \tag{4.135b}$$

$$(\delta S_y)^2 = \langle S_y^2 \rangle - \langle S_y \rangle^2 = \langle \uparrow | \hat{S}_y^2 | \uparrow \rangle - 0 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle = \left(\frac{\hbar}{2}\right)^2. \tag{4.135c}$$

While Eqs. (134) and (135a) are compatible with the classical notion of the angular momentum of magnitude $\hbar/2$ being directed exactly along the z-axis, this correspondence should not be overstretched, because such a classical picture cannot explain Eqs. (135b) and (135c). The best (but still imprecise!) classical image I can offer is the spin vector **S** oriented, on average, in the z-direction, but still having its x- and y-components strongly "wobbling" (fluctuating) 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 axis indices. Thus, in neither of these states, all three spin components have definite values. Let me show that this is not just an occasional fact, but reflects one of the most profound properties of quantum mechanics, the *uncertainty relations*. For that, let us consider two measurable observables, A and B, of the same quantum system. There are two possibilities here. If the operators corresponding to these observables commute,

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

then all matrix elements of the commutator in any orthogonal basis (in particular, in the basis of eigenstates a_j of the operator $\hat{A}$) have to equal zero:

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

In the first bracket of the middle expression, let us act by the (Hermitian!) operator $\hat{A}$ on the bra-vector, while in the second one, on the ket-vector. According to Eq. (68), such action turns the operators into the corresponding eigenvalues, which may be taken out of the long brackets, so we get

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

This means that if all eigenstates of the operator $\hat{A}$ are non-degenerate (i.e. $A_j \neq A_{j'}$ if $j \neq j'$), the matrix of the operator $\hat{B}$ has to be diagonal in the basis $\{a\}$, i.e., the operators $\hat{A}$ and $\hat{B}$ have common eigenstates. Such pairs of observables (and their operators) that can share their eigenstates are called *compatible*. For example, in the wave mechanics of a particle, its momentum (1.26) and kinetic energy (1.27) are compatible, sharing their 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 integer power of itself:

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

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

General uncertainty relation

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

where all expectation values are for the same but arbitrary state of the system. The proof of Eq. (140) may be divided into two steps, the first one proving the so-called *Schwartz inequality* for any two possible states, say α and β:[34]

Schwartz inequality

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

Its proof may be readily achieved by applying the postulate (16) – that the norm of any legitimate state of the system cannot be negative – to the state with the following ket-vector:

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

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

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

Opening the parentheses, we get

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

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

Now let us apply this inequality to states

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

where, in both relations, γ is the same possible state of the system, and the deviation operators are defined similarly to the deviations of the observables (see Sec. 1.2):

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

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

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

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

[34] This inequality is the quantum-mechanical analog of the usual vector algebra's result $\alpha^2\beta^2 \geq |\boldsymbol{\alpha}\cdot\boldsymbol{\beta}|^2$.

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

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

$$\hat{\tilde{A}}\hat{\tilde{B}} = \frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\} - \frac{i}{2}\hat{C}, \quad \text{where } \hat{C} \equiv i\left[\hat{\tilde{A}},\hat{\tilde{B}}\right]. \tag{4.149}$$

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

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

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

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

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

$$\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle^2 = \left\langle \frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\} \right\rangle^2 + \left\langle \frac{1}{2}\hat{C} \right\rangle^2. \tag{4.152}$$

Since the first term on the right-hand side of this equality cannot be negative, we may write

$$\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle^2 \geq \left\langle \frac{1}{2}\hat{C} \right\rangle^2 \equiv \left\langle \frac{i}{2}\left[\hat{A},\hat{B}\right] \right\rangle^2, \tag{4.153}$$

and hence continue Eq. (148) as

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

thus proving Eq. (140).

For the particular case of operators $\hat{x}$ and $\hat{p}_x$ (or a similar pair of operators for another Cartesian coordinate), we may readily combine Eq. (140) with Eq. (2.14b) to prove the original Heisenberg's uncertainty relation (2.13). For the spin-½ operators defined by Eq. (116)-(117), it is very simple (and highly recommended to the reader) to show that

$$\left[\hat{\sigma}_j,\hat{\sigma}_{j'}\right] = 2i\sum_{j''=1}^{3}\varepsilon_{jj'j''}\hat{\sigma}_{j''}, \qquad \text{i.e. } \left[\hat{S}_j,\hat{S}_{j'}\right] = i\hbar\sum_{j''=1}^{3}\varepsilon_{jj'j''}\hat{S}_{j''}, \tag{4.155}$$

Spin- ½: commutation relations

where $\varepsilon_{jj'j''}$ is the Levi-Civita permutation symbol.[35] As a result, the uncertainty relations (140) for all Cartesian components of spin-½ systems are similar, for example

[35] See, e.g., MA Eq. (13.2).

Spin-½: uncertainty relations

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

In particular, as we already know, in the ↑ state the right-hand side of this relation equals $(\hbar/2)^2 > 0$, so neither of the uncertainties δS_x, δS_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 is equal to the lowest value allowed by the uncertainty relation (156) – just as the Gaussian wave packets (2.16) provide the lowest possible value of the product $\delta x \delta p_x$, allowed by the Heisenberg relation (2.13).

4.6. Quantum dynamics: Three pictures

So far in this chapter, I shied away from the discussion of the system's dynamics, implying that the bra- and ket-vectors were just their "snapshots" at a certain instant t. Now we are sufficiently prepared to examine their evolution in time. One of the most beautiful features of quantum mechanics is that this evolution may be described using either of three alternatives (called *pictures*), giving exactly the same final results for the expectation values of all observables.

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

$$\langle \alpha(t)| = \langle \alpha(t_0)|\, \hat{u}^\dagger(t,t_0), \qquad |\alpha(t)\rangle = \hat{u}(t,t_0)\,|\alpha(t_0)\rangle. \tag{4.157a}$$

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

$$i\hbar \frac{\partial}{\partial t}\hat{u} = \hat{H}\hat{u}, \tag{4.157b}$$

where $\hat{H}$ is the Hamiltonian operator of the system – which is always Hermitian: $\hat{H}^\dagger = \hat{H}$, and t_0 is the initial moment of time. (Note that Eqs. (157) remain valid even if the Hamiltonian depends on time explicitly.) Differentiating the second of Eqs. (157a) over time t, and then using Eq. (157b) twice, we can merge these two relations into a single equation, without explicit use of the time-evolution operator:

Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}|\alpha(t)\rangle = \hat{H}|\alpha(t)\rangle, \tag{4.158}$$

which is frequently more convenient. (However, for some purposes the notion of the time-evolution operator, together with Eq. (157b), are useful – as we will see in a minute.) While Eq. (158) is a very natural generalization of the wave-mechanical equation (1.25), and is also frequently called the *Schrödinger equation*,[36] 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 of its corollaries with experiment – more exactly, in the absence of a single credible contradiction to an experiment.

Starting the discussion of Eq. (158), let us first consider the case of a time-independent Hamiltonian, whose eigenstates a_n and eigenvalues E_n obey Eq. (68) for this operator:[37]

[36] Moreover, we will be able to *derive* Eq. (1.25) from Eq. (158) – see below.

[37] I have switched the state index notation from j to n, which was used for numbering stationary states in Chapter 1, to emphasize the special role played by the stationary states a_n in quantum dynamics.

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

and hence are also time-independent. (Similarly to the wavefunctions ψ_n defined by Eq. (1.60), a_n are called the *stationary states* of the system.) Let us use Eqs. (158)-(159) to calculate the law of time evolution of the expansion coefficients α_n (i.e. the probability amplitudes) defined by Eq. (118), in a stationary state basis, using Eq. (158):

$$\dot{\alpha}_n(t) = \frac{d}{dt}\langle a_n|\alpha(t)\rangle = \langle a_n|\frac{d}{dt}|\alpha(t)\rangle = \langle a_n|\frac{1}{i\hbar}\hat{H}|\alpha(t)\rangle = \frac{E_n}{i\hbar}\langle a_n|\alpha(t)\rangle = -\frac{i}{\hbar}E_n\alpha_n. \tag{4.160}$$

This is the same simple equation as Eq. (1.61), and its integration, with the initial moment t_0 taken for 0, yields a similar result – cf. Eq. (1.62):

$$\boxed{\alpha_n(t) = \alpha_n(0)\exp\left\{-\frac{i}{\hbar}E_n t\right\}.} \tag{4.161}$$

Time evolution of probability amplitudes

In order to illustrate how this result works, let us consider the dynamics of a spin-½ in a time-independent, uniform external magnetic field $\mathscr{B}$. To construct the system's Hamiltonian, we may apply the correspondence principle to the classical expression for the energy of a magnetic moment **m** in the external magnetic field $\mathscr{B}$, [38]

$$U = -\mathbf{m}\cdot\mathscr{B}. \tag{4.162}$$

In quantum mechanics, the operator corresponding to the moment **m** is given by Eq. (115) (suggested by W. Pauli), so the spin-field interaction is described by the so-called *Pauli Hamiltonian*, which may be, due to Eqs. (116)-(117), represented in several equivalent forms:

$$\boxed{\hat{H} = -\hat{\mathbf{m}}\cdot\mathscr{B} \equiv -\gamma\hat{\mathbf{S}}\cdot\mathscr{B} = -\gamma\frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\cdot\mathscr{B}.} \tag{4.163a}$$

Pauli Hamiltonian: operator

If the z-axis is aligned with the field's direction, this expression is reduced to

$$\hat{H} = -\gamma\mathscr{B}\hat{S}_z \equiv -\gamma\mathscr{B}\frac{\hbar}{2}\hat{\sigma}_z. \tag{4.163b}$$

According to Eq. (117), in the z-basis of the spin states ↑ and ↓, the matrix of the operator (163b) is

$$\boxed{\mathrm{H} = -\frac{\gamma\hbar\mathscr{B}}{2}\sigma_z \equiv \frac{\hbar\Omega}{2}\sigma_z, \quad \text{where } \Omega \equiv -\gamma\mathscr{B}.} \tag{4.164}$$

Pauli Hamiltonian: z-basis matrix

The constant Ω so defined coincides with the classical frequency of the precession, about the z-axis, of an axially-symmetric rigid body (the so-called *symmetric top*), with an angular momentum **S** and the magnetic moment $\mathbf{m} = \gamma\mathbf{S}$, induced by the external torque $\boldsymbol{\tau} = \mathbf{m}\times\mathscr{B}$.[39] (For an electron, with its negative gyromagnetic ratio $\gamma_e = -g_e e/2m_e$, neglecting the tiny difference of the g_e-factor from 2, we get

$$\Omega = \frac{e}{m_e}\mathscr{B}, \tag{4.165}$$

so according to Eq. (3.48), the frequency Ω coincides with the electron's cyclotron frequency ω_c.)

[38] See, e.g., EM Eq. (5.100). As a reminder, we have already used this expression for the derivation of Eq. (3).
[39] See, e.g., CM Sec. 4.5, in particular Eq. (4.72), and EM Sec. 5.5, in particular Eq. (5.114) and its discussion.

In order to apply the general Eq. (161) to this case, we need to find the eigenstates a_n and eigenenergies E_n of our Hamiltonian. However, with our (smart :-) choice of the z-axis, the Hamiltonian matrix is already diagonal:

$$\hat{H} = \frac{\hbar\Omega}{2}\sigma_z \equiv \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.166}$$

meaning that the states ↑ and ↓ are the eigenstates of this system, with the eigenenergies, respectively, [40]

Spin-½ in magnetic field: eigenenergies

$$E_\uparrow = +\frac{\hbar\Omega}{2} \quad \text{and} \quad E_\downarrow = -\frac{\hbar\Omega}{2}. \tag{4.167}$$

Note that their difference,

$$\Delta E \equiv \left|E_\uparrow - E_\downarrow\right| = \hbar\left|\Omega\right| = \hbar\left|\gamma\mathscr{B}\right|, \tag{4.168}$$

corresponds to the classical energy $2|m\mathscr{B}|$ of flipping a magnetic dipole with the moment's magnitude $m = \gamma\hbar/2$, oriented along the direction of the field $\mathscr{B}$. Note also that if the product $\gamma\mathscr{B}$ is positive, then Ω is negative, so $E_\uparrow$ is negative, while $E_\downarrow$ is positive. This is in agreement with the classical picture of a magnetic dipole **m** having negative potential energy when it is aligned with the external magnetic field $\mathscr{B}$ – see Eq. (162) again.

So, for the time evolution of the probability amplitudes of these states, Eq. (161) immediately yields the following expressions:

$$\alpha_\uparrow(t) = \alpha_\uparrow(0)\exp\left\{-\frac{i}{2}\Omega t\right\}, \qquad \alpha_\downarrow(t) = \alpha_\downarrow(0)\exp\left\{+\frac{i}{2}\Omega t\right\}, \tag{4.169}$$

allowing a ready calculation of the 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 the (arbitrary) time moment t:

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

Thus the expectation value of the spin component parallel to the applied magnetic field remains constant in time, 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 the moment t, gives

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

Clearly, this expression describes sinusoidal oscillations with frequency (164). The amplitude and the phase of these oscillations depend on initial conditions. Indeed, solving Eqs. (132)-(133) for the probability amplitude products, we get the following relations:

$$\hbar\alpha_\downarrow(t)\alpha_\uparrow^*(t) = \langle S_x\rangle(t) + i\langle S_y\rangle(t), \qquad \hbar\alpha_\uparrow(t)\alpha_\downarrow^*(t) = \langle S_x\rangle(t) - i\langle S_y\rangle(t), \tag{4.172}$$

valid for any time t. Plugging their values for $t = 0$ into Eq. (171), we get

[40] So, spin-½ gives one more example of two-level systems whose discussion was started in Sec. 2.6. The fact that all quantum two-level systems are isomorphic (see Sec. 5.1) adds importance to our current discussion.

$$\begin{aligned}\langle S_x\rangle(t) &= \frac{1}{2}\left[\langle S_x\rangle(0)+i\langle S_y\rangle(0)\right]e^{+i\Omega t}+\frac{1}{2}\left[\langle S_x\rangle(0)-i\langle S_y\rangle(0)\right]e^{-i\Omega t}\\ &\equiv \langle S_x\rangle(0)\cos\Omega t-\langle S_y\rangle(0)\sin\Omega t\,.\end{aligned} \tag{4.173}$$

An absolutely similar calculation using Eq. (133) gives

$$\langle S_y\rangle(t)=\langle S_y\rangle(0)\cos\Omega t+\langle S_x\rangle(0)\sin\Omega t\,. \tag{4.174}$$

These formulas show, for example, that if at moment $t = 0$ the spin's state was ↑, i.e. $\langle S_x\rangle(0) = \langle S_y\rangle(0) = 0$, then the oscillation amplitudes of both "lateral" components of the spin vanish. On the other hand, if the spin was initially in the state →, i.e. had the definite, largest possible value of S_x equal to $\hbar/2$ (in classics, we would say "the spin-½ was oriented in the x-direction"), then both expectation values $\langle S_x\rangle$ and $\langle S_y\rangle$ oscillate in time[41] with this amplitude, and with the phase shift $\pi/2$ between them.

So, the quantum-mechanical results for the expectation values of the Cartesian components of spin-½ are indistinguishable from the classical results for the precession, with the frequency $\Omega = -\gamma\mathscr{B}$,[42] of a symmetric top with the angular momentum $\mathbf{L}$ of magnitude $\hbar/2$, about the field's direction (our axis z), under the effect of an external torque $\boldsymbol{\tau} = \mathbf{m}\times\mathscr{B}$ exerted by the field $\mathscr{B}$ on the magnetic moment $\mathbf{m} = \gamma\mathbf{L}$. Note, however, that the classical language does not describe the large quantum-mechanical uncertainties of the components, obeying Eqs. (156), which are absent in the classical picture – at least when the precession starts from a definite orientation of the angular momentum vector.[43]

Recall also that at the stationary orbital motion of a particle, the component L_z of its angular momentum is always a multiple of $\hbar$ – see, e.g., Eq. (3.139). As a result, the angular momentum of a spin-½ particle, with its stationary values $S_z = \pm\hbar/2$, cannot be explained by the summation of orbital moments of its hypothetical components, i.e. by any internal rotation of the particle about its axis.

After this illustration, let us return to the discussion of the general Schrödinger equation (157b) 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 turns out to be an exact analog of Eq. (161),

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

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

---

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

[42] Note that according to this relation, the gyromagnetic ratio γ may be interpreted as the angular frequency of the spin precession in a unit magnetic field – hence the name. In particular, for electrons, $|\gamma_e| \approx 1.761\times10^{11}$ s^{-1}T^{-1}; for protons, the ratio is much smaller, $\gamma_p \equiv g_p e/2m_p \approx 2.675\times10^{8}$ s^{-1}T^{-1} – mostly because of their larger mass m_p, at a g-factor of the same order as for the electron: $g_p \approx 5.586$. For heavier spin-½ particles, e.g., atomic nuclei with such spin, the values of γ are correspondingly smaller – e.g., $\gamma \approx 8.681\times10^{6}$ s^{-1}T^{-1} for the ^{57}Fe nucleus.

[43] If the initial conditions are random, the classical motion is stochastic even if its laws are deterministic.

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty} \frac{1}{k!}\left[-\frac{i}{\hbar}\hat{H}(t-t_0)\right]^k$$
$$\equiv \hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2 + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^3 \hat{H}^3(t-t_0)^3 + ..., \quad (4.176)$$

where $\hat{H}^2 \equiv \hat{H}\hat{H}$, $\hat{H}^3 \equiv \hat{H}\hat{H}\hat{H}$, etc. Working with such a series of operator products is not as hard as one could imagine, due to their regular structure. For example, let us differentiate both sides of Eq. (176) over t, at constant t_0, at the last step using this equality again – that time, backward:

$$\frac{\partial}{\partial t}\hat{u}(t,t_0) = \hat{0} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H} + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2 2(t-t_0) + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^3 3(t-t_0)^2 + ...$$
$$\equiv \left(-\frac{i}{\hbar}\right)\hat{H}\left[\hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2\right] + ... \equiv -\frac{i}{\hbar}\hat{H}\hat{u}(t,t_0), \quad (4.177)$$

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

$$\hat{u}(t_0,t_0) = \hat{u}^{\dagger}(t_0,t_0) = \hat{I} \quad (4.178)$$

that immediately follows from the definition (157a) of the evolution operator. Thus, Eq. (175) indeed gives the (unique) solution for the time evolution operator – in the Schrödinger picture.

Now let us allow the 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:

$$\left[\hat{H}(t'), \hat{H}(t'')\right] = 0, \quad \text{for any } t', t''. \quad (4.179)$$

(A good example is the Pauli Hamiltonian (4.163) for a spin in a classical magnetic field $\mathscr{B}$ even if it depends on time. Indeed, the spin operator $\hat{\mathbf{S}}$ does not depend explicitly on time and hence commutes with itself as well as with the c-numbers $\mathscr{B}(t')$ and $\mathscr{B}(t'')$. Note, however, that a similar operator describing the effect of a classical position-independent force $\mathbf{F}(t)$ on the orbital motion of a particle,

$$\hat{H}_F = -\mathbf{F}(t)\cdot\hat{\mathbf{r}}, \quad (4.180)$$

may be deceiving: though it satisfies Eq. (179), this relation is invalid for the particle's *full* Hamiltonian including its kinetic energy.) In this case, it is sufficient to replace, in all the above formulas, the product $\hat{H}(t-t_0)$ with the corresponding integral over time; in particular, Eq. (175) is generalized as

Evolution operator: explicit expression

$$\hat{u}(t,t_0) = \exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}. \quad (4.181)$$

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

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\left(\int_{t_0}^{t}\hat{H}(t')dt'\right)^k \equiv \hat{I} + \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2...\int_{t_0}^{t}dt_k\hat{H}(t_1)\hat{H}(t_2)...\hat{H}(t_k). \quad (4.182)$$

The proof that 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 remains unitary at any moment, even for a time-dependent Hamiltonian, if it satisfies Eq. (179). Indeed, Eq. (181) yields

$$\hat{u}(t,t_0)\hat{u}^{\dagger}(t,t_0)=\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}\exp\left\{+\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t'')dt''\right\}. \tag{4.183}$$

Since each of these exponents may be represented with the Taylor series (182), and, thanks to Eq. (179), different components of these sums may be swapped at will, the expression (183) may be manipulated exactly as the product of *c*-number exponents, for example rewritten as

$$\hat{u}(t,t_0)\hat{u}^{\dagger}(t,t_0)=\exp\left\{-\frac{i}{\hbar}\left[\int_{t_0}^{t}\hat{H}(t')dt'-\int_{t_0}^{t}\hat{H}(t'')dt''\right]\right\}=\exp\{\hat{0}\}=\hat{I}. \tag{4.184}$$

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

$$\langle\alpha(t)|\alpha(t)\rangle=\langle\alpha(t_0)|\hat{u}^{\dagger}(t,t_0)\hat{u}(t,t_0)|\alpha(t_0)\rangle=\langle\alpha(t_0)|\alpha(t_0)\rangle. \tag{4.185}$$

The most difficult cases for the explicit solution of Eq. (158) are those where Eq. (179) is violated.[44] It may be proved that in these cases, Eqs. (181)-(182) should be replaced with the following *Dyson series* using the so-called *time-ordering operator* $\hat{\mathcal{T}}$:

$$\hat{u}(t,t_0)=\hat{\mathcal{T}}\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}=\hat{I}+\sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^{k}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\ldots\int_{t_0}^{t}dt_k\,\hat{\mathcal{T}}\left[\hat{H}(t_1)\hat{H}(t_2)\ldots\hat{H}(t_k)\right],$$
$$\text{where }\hat{\mathcal{T}}\left[\hat{H}(t_1)\hat{H}(t_2)\right]\equiv\begin{cases}\hat{H}(t_1)\hat{H}(t_2), & \text{for } t_2\le t_1,\\ \hat{H}(t_2)\hat{H}(t_1), & \text{for } t_1\le t_2.\end{cases} \tag{4.186}$$

Since we would not have time/space to use this relation in this course, I will skip its proof.[45]

Let me now return to the general discussion of quantum dynamics to outline its alternative, the *Heisenberg picture*. For its introduction, let us recall that according to Eq. (125), in quantum mechanics the expectation value of any observable A is a long bracket. Let us explore the even more general form of such a bracket:

$$\langle\alpha|\hat{A}|\beta\rangle, \tag{4.187}$$

because in some applications, the states α and β may be different. As was discussed above, in the Schrödinger picture the bra- and ket-vectors of the states evolve in time, while the operators of observables remain time-independent (if they do not explicitly depend on time). As a result, Eq. (187) applied to the moment t, may be represented as

$$\langle\alpha(t)|\hat{A}_{\rm S}|\beta(t)\rangle, \tag{4.188}$$

where the index "S" is added to emphasize the Schrödinger picture. Let us apply the evolution law (157a) to the bra- and ket-vectors in this expression:

$$\langle\alpha(t)|\hat{A}_{\rm S}|\beta(t)\rangle=\langle\alpha(t_0)|\hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0)|\beta(t_0)\rangle. \tag{4.189}$$

[44] We will run into such situations in Chapter 7, but will not need to apply Eq. (186) there.

[45] It may be found, for example, in Chapter 5 of J. Sakurai's textbook – see *References*.

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*[46]

Heisenberg operator

$$\hat{A}_{\rm H}(t) \equiv \hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0) = \hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm H}(t_0)\hat{u}(t,t_0)\,, \tag{4.190}$$

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

Heisenberg picture

$$\langle\alpha(t)|\hat{A}|\beta(t)\rangle = \langle\alpha(t_0)|\hat{A}_{\rm H}(t,t_0)|\beta(t_0)\rangle. \tag{4.191}$$

For full clarity, let us see how the Heisenberg picture works 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. (157b) for the time-evolution operator becomes

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22}\end{pmatrix} = \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22}\end{pmatrix} \equiv \frac{\hbar\Omega}{2}\begin{pmatrix} u_{11} & u_{12} \\ -u_{21} & -u_{22}\end{pmatrix}. \tag{4.192}$$

We see that in this simple case, the differential equations for different matrix elements of the evolution operator matrix are decoupled, and readily solvable by using the universal initial conditions (178):[47]

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

Now let us use them in Eq. (190) to calculate the Heisenberg-picture operators of spin components – still in the *z*-basis. Dropping the index "H" for the notation brevity (the Heisenberg-picture operators are clearly marked by their dependence on time anyway), we get

$$\begin{aligned} \mathrm{S}_x(t) &= \mathrm{u}^{\dagger}(t,0)\mathrm{S}_x(0)\mathrm{u}(t,0) = \frac{\hbar}{2}\mathrm{u}^{\dagger}(t,0)\sigma_x\mathrm{u}(t,0) \\ &= \frac{\hbar}{2}\begin{pmatrix} e^{i\Omega t/2} & 0 \\ 0 & e^{-i\Omega t/2}\end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}\begin{pmatrix} e^{-i\Omega t/2} & 0 \\ 0 & e^{i\Omega t/2}\end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} 0 & e^{i\Omega t} \\ e^{-i\Omega t} & 0\end{pmatrix} = \frac{\hbar}{2}\left(\sigma_x\cos\Omega t - \sigma_y\sin\Omega t\right) \equiv \mathrm{S}_x(0)\cos\Omega t - \mathrm{S}_y(0)\sin\Omega t\,. \end{aligned} \tag{4.194}$$

Absolutely similar calculations of the other spin components yield

$$\mathrm{S}_y(t) = \frac{\hbar}{2}\begin{pmatrix} 0 & -ie^{i\Omega t} \\ ie^{-i\Omega t} & 0\end{pmatrix} = \frac{\hbar}{2}\left(\sigma_y\cos\Omega t + \sigma_x\sin\Omega t\right) \equiv \mathrm{S}_y(0)\cos\Omega t + \mathrm{S}_x(0)\sin\Omega t\,, \tag{4.195}$$

[46] Note that this relation is similar in structure to the first of Eqs. (94), with the state bases $\{v\}$ and $\{u\}$ loosely associated with the time moments, respectively, t and t_0.

[47] We could of course use this solution, 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 some time-dependent Hamiltonians, an explicit calculation of the time-evolution matrix may be the best (or even the only practicable) way to proceed.

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

One practical advantage of these formulas is that they describe the system's evolution for *arbitrary* initial conditions, thus making the analysis of 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), which were obtained above 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. (Below we will see that the same exact correspondence is valid for the Heisenberg picture of the orbital motion.)

In order to see that the last fact is by no means a coincidence, let us combine Eqs. (157b) and (190) to form an explicit differential equation of the Heisenberg operator's evolution. For that, let us differentiate Eq. (190) over time:

$$\frac{d}{dt}\hat{A}_{\rm H} = \frac{\partial \hat{u}^\dagger}{\partial t}\hat{A}_{\rm S}\hat{u} + \hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\frac{\partial \hat{u}}{\partial t}. \tag{4.197}$$

Plugging in the derivatives of the time evolution operator from Eq. (157b) and its Hermitian conjugate, and multiplying both sides of the equation by $i\hbar$, we get

$$i\hbar\frac{d}{dt}\hat{A}_{\rm H} = -\hat{u}^\dagger\hat{H}\hat{A}_{\rm S}\hat{u} + i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\hat{H}\hat{u}. \tag{4.198a}$$

If for the Schrödinger-picture's Hamiltonian, the condition (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.[48] Hence, we may rewrite Eq. (198a) as

$$i\hbar\frac{d}{dt}\hat{A}_{\rm H} = -\hat{H}\hat{u}^\dagger\hat{A}_{\rm S}\hat{u} + i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\hat{u}\hat{H} \equiv i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \left[\hat{u}^\dagger\hat{A}_{\rm S}\hat{u}, \hat{H}\right]. \tag{4.198b}$$

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

$$\boxed{i\hbar\frac{d}{dt}\hat{A}_{\rm H} = i\hbar\left(\frac{\partial \hat{A}}{\partial t}\right)_{\rm H} + \left[\hat{A}_{\rm H}, \hat{H}\right].} \tag{4.199}$$

Heisenberg equation of motion

This is the so-called *Heisenberg equation of motion*.

Let us see how this equation looks for the same problem of the spin-½ 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 operator vector of spin reads[49]

[48] Due to the same reason, $\hat{H}_{\rm H} \equiv \hat{u}^\dagger\hat{H}_{\rm S}\hat{u} = \hat{u}^\dagger\hat{u}\hat{H}_{\rm S} = \hat{H}_{\rm S}$; this is why the Hamiltonian operator's index may be dropped in Eqs. (198)-(199).

[49] Using the commutation relations (155), this equation may be readily generalized to the case of an arbitrary magnetic field $\mathscr{B}(t)$ and an arbitrary state basis – the exercise highly recommended to the reader.

$$i\hbar\begin{pmatrix}\dot{\mathbf{S}}_{11} & \dot{\mathbf{S}}_{12}\\ \dot{\mathbf{S}}_{21} & \dot{\mathbf{S}}_{22}\end{pmatrix}=\frac{\hbar\Omega}{2}\left[\begin{pmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{21} & \mathbf{S}_{22}\end{pmatrix},\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]=\hbar\Omega\begin{pmatrix}0 & -\mathbf{S}_{12}\\ \mathbf{S}_{21} & 0\end{pmatrix}. \tag{4.200}$$

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

$$\begin{aligned}&\mathbf{S}_{11}(t)=\mathbf{S}_{11}(0)=\text{const}, \quad &&\mathbf{S}_{22}(t)=\mathbf{S}_{22}(0)=\text{const},\\ &\mathbf{S}_{12}(t)=\mathbf{S}_{12}(0)e^{+i\Omega t}, \quad &&\mathbf{S}_{21}(t)=\mathbf{S}_{21}(0)e^{-i\Omega t}.\end{aligned} \tag{4.201}$$

According to Eq. (190), the initial values of the Heisenberg-picture matrix elements are just the Schrödinger-picture ones, so using Eq. (117) we may rewrite this solution in either of two forms:

$$\begin{aligned}\mathbf{S}(t)&=\frac{\hbar}{2}\left[\mathbf{n}_x\begin{pmatrix}0 & e^{+i\Omega t}\\ e^{-i\Omega t} & 0\end{pmatrix}+\mathbf{n}_y\begin{pmatrix}0 & -ie^{+i\Omega t}\\ ie^{-i\Omega t} & 0\end{pmatrix}+\mathbf{n}_z\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]\\ &\equiv\frac{\hbar}{2}\begin{pmatrix}\mathbf{n}_z & \mathbf{n}_-e^{+i\Omega t}\\ \mathbf{n}_+e^{-i\Omega t} & -\mathbf{n}_z\end{pmatrix}, \qquad \text{where } \mathbf{n}_\pm\equiv\mathbf{n}_x\pm i\mathbf{n}_y\ .\end{aligned} \tag{4.202}$$

The simplicity of the last expression is spectacular. (Remember, it covers *any* initial conditions and *all* three spatial components of spin!) On the other hand, for some purposes the previous form may be more convenient; in particular, its Cartesian components give our earlier results (194)-(196).[50]

One of the advantages of the Heisenberg picture is that it provides a more clear link between classical and quantum mechanics, found by P. Dirac. 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 the generalized coordinates q_j and momenta p_j of the system, and time t: [51]

$$\frac{dA}{dt}=\frac{\partial A}{\partial t}-\{A,H\}_{\rm P}, \tag{4.203}$$

where H is the classical Hamiltonian function of the system, and $\{..,..\}_{\rm P}$ 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

Poisson bracket

$$\{A,B\}_{\rm P}\equiv\sum_j\left(\frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j}-\frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j}\right). \tag{4.204}$$

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

[50] Note that the "values" of the same Heisenberg operator at different moments of time may or may not commute. For example, consider a free 1D particle, with the time-independent Hamiltonian $\hat{H}=\hat{p}^2/2m$. In this case, Eq. (199) yields the following equations: $i\hbar\dot{\hat{x}}=[\hat{x},\hat{H}]=i\hbar\hat{p}/m$ and $i\hbar\dot{\hat{p}}=[\hat{p},\hat{H}]=0$, with simple solutions (similar to those for the classical motion): $\hat{p}(t)=\text{const}=\hat{p}(0)$ and $\hat{x}(t)=\hat{x}(0)+\hat{p}(0)t/m$, so $[\hat{x}(0),\hat{x}(t)]=[\hat{x}(0),\hat{p}(0)]t/m\equiv[\hat{x}_{\rm S},\hat{p}_{\rm S}]t/m=i\hbar t/m\neq 0$, for $t\neq 0$.

[51] See, e.g., CM Eq. (10.17). The notation there does not use the subscript "P" that is employed in Eqs. (203)-(205) to distinguish the classical Poisson bracket (204) from the quantum anticommutator (34).

$$\{A,B\}_{\rm P} \leftrightarrow \frac{i}{\hbar}\left[\hat{A},\hat{B}\right]. \tag{4.205}$$

Classical vs. quantum mechanics

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

Finally, let us discuss one more alternative picture of quantum dynamics. It is 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 *perturbative* (approximate) approaches to systems whose Hamiltonians may be partitioned into two parts,

$$\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}, \tag{4.206}$$

where $\hat{H}_0$ is the sum of relatively simple Hamiltonians of the component subsystems, while the second term in Eq. (206) represents their weak interaction.[52] (Note, however, that all relations in the balance of this section are exact and not directly based on the interaction weakness.) In this case, it is natural to consider, together with the full operator $\hat{u}(t,t_0)$ of the system's evolution, which obeys Eq. (157b), a similarly defined unitary operator $\hat{u}_0(t,t_0)$ of the "unperturbed" evolution described by $\hat{H}_0$ alone:

$$i\hbar\frac{\partial}{\partial t}\hat{u}_0 = \hat{H}_0\hat{u}_0, \tag{4.207}$$

and also the following *interaction evolution operator*,

$$\hat{u}_{\rm I} \equiv \hat{u}_0^\dagger\hat{u}\,. \tag{4.208}$$

Interaction evolution operator

The motivation for these definitions becomes more clear if we insert the reciprocal relation,

$$\hat{u} \equiv \hat{u}_0\,\hat{u}_0^\dagger\hat{u} = \hat{u}_0\hat{u}_{\rm I}, \tag{4.209}$$

and its Hermitian conjugate,

$$\hat{u}^\dagger = \left(\hat{u}_0\hat{u}_{\rm I}\right)^\dagger = \hat{u}_{\rm I}^\dagger\hat{u}_0^\dagger, \tag{4.210}$$

into the basic Eq. (189):

$$\begin{aligned}\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle &= \left\langle\alpha(t_0)\middle|\hat{u}^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0)\middle|\beta(t_0)\right\rangle \\ &= \left\langle\alpha(t_0)\middle|\hat{u}_{\rm I}^\dagger(t,t_0)\hat{u}_0^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}_0(t,t_0)\hat{u}_{\rm I}(t,t_0)\middle|\beta(t_0)\right\rangle.\end{aligned} \tag{4.211}$$

This relation shows that any long bracket (187), i.e. any experimentally verifiable result of quantum mechanics, may be expressed as

$$\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle = \left\langle\alpha_{\rm I}(t)\middle|\hat{A}_{\rm I}(t)\middle|\beta_{\rm I}(t)\right\rangle, \tag{4.212}$$

if we assume that both the state vectors and the operators depend on time, with the vectors evolving only due to the *interaction operator* $\hat{u}_{\rm I}$,

$$\left\langle\alpha_{\rm I}(t)\right| \equiv \left\langle\alpha(t_0)\right|\hat{u}_{\rm I}^\dagger(t,t_0), \qquad \left|\beta_{\rm I}(t)\right\rangle \equiv \hat{u}_{\rm I}(t,t_0)\left|\beta(t_0)\right\rangle, \tag{4.213}$$

Interaction picture: state vectors

[52] This picture may also useful in more standard problems of the perturbation theory (see Ch. 6 below) where $\hat{H}_{\rm int}$ describes a weak perturbation of a *single* system described by a relatively simple Hamiltonian $\hat{H}_0$.

while the operators' evolution being governed by the *unperturbed operator* $\hat{u}_0$:

Interaction picture: operators

$$\hat{A}_{\rm I}(t) \equiv \hat{u}_0^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}_0(t,t_0). \tag{4.214}$$

These relations describe the interaction picture of quantum dynamics. Let me defer an example of its use until the perturbative analysis of open quantum systems in Sec. 7.6, and end this section with proof that the interaction evolution operator (208) satisfies the following natural equation,

$$i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} = \hat{H}_{\rm I}\hat{u}_{\rm I}, \tag{4.215}$$

where $\hat{H}_{\rm I}$ is the interaction Hamiltonian formed from $\hat{H}_{\rm int}$ in accordance with the same rule (214):

$$\hat{H}_{\rm I}(t) \equiv \hat{u}_0^\dagger(t,t_0)\hat{H}_{\rm int}\hat{u}_0(t,t_0). \tag{4.216}$$

The proof is very straightforward: first using the definition (208), and then Eqs. (157b) and the Hermitian conjugate of Eq. (207), we may write

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} &\equiv i\hbar\frac{\partial}{\partial t}\left(\hat{u}_0^\dagger\hat{u}\right) \equiv i\hbar\frac{\partial\hat{u}_0^\dagger}{\partial t}\hat{u} + \hat{u}_0^\dagger i\hbar\frac{\partial\hat{u}}{\partial t} = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\hat{H}\hat{u} = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\left(\hat{H}_0 + \hat{H}_{\rm int}\right)\hat{u} \\ &\equiv -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\hat{H}_0\hat{u} + \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u} \equiv \left(-\hat{H}_0\hat{u}_0^\dagger + \hat{u}_0^\dagger\hat{H}_0\right)\hat{u} + \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}. \end{aligned} \tag{4.217}$$

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

$$i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} = \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}_0 u_{\rm I} \equiv \left(\hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}_0\right)\hat{u}_{\rm I}, \tag{4.218}$$

which is clearly equivalent to the combination of Eqs. (215) and (216).

As Eq. (215) shows, if the energy scale of the interaction $H_{\rm int}$ is much smaller than that of the background Hamiltonian H_0, the interaction evolution operators $\hat{u}_{\rm I}$ and $\hat{u}_{\rm I}^\dagger$, and hence the state vectors (213) evolve relatively slowly, without fast background oscillations. This is very 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. Coordinate and momentum representations

Now let me show that in application to the orbital motion of a particle, the bra-ket formalism naturally reduces to the notions and postulates of wave mechanics, which were discussed in Chapter 1. For that, we first have to modify some of the above formulas for the case of a basis with a continuous spectrum of eigenvalues. In that case, it is more appropriate to replace discrete indices, such as *j*, *j'*, etc. broadly used above, with the corresponding eigenvalue – just as it was done earlier for functions of the wave vector – see, e.g., Eqs. (1.88), (2.20), etc. For example, the key Eq. (68), defining the eigenkets and eigenvalues of an operator, may be conveniently rewritten in the form

$$\hat{A}|a_A\rangle = A|a_A\rangle. \tag{4.219}$$

More substantially, all sums over such continuous eigenstate sets should be replaced with integrals. For example, for a full and orthonormal set of the continuous eigenstates $|a_A\rangle$, the closure relation (44) should be replaced with

Continuous spectrum: closure relation

$$\int dA\,|a_A\rangle\langle a_A| = \hat{I}, \tag{4.220}$$

where the integral is over the whole interval of possible eigenvalues of the observable A.[53] Applying this relation to the ket-vector of an arbitrary state α, we get the following replacement of Eq. (37):

$$|\alpha\rangle \equiv \hat{I}|\alpha\rangle = \int dA\,|a_A\rangle\langle a_A|\alpha\rangle = \int dA\,\langle a_A|\alpha\rangle|a_A\rangle. \tag{4.221}$$

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

Continuous spectrum: state orthonormality

$$\langle a_A|a_{A'}\rangle = \delta(A - A'); \tag{4.222}$$

this formula replaces the orthonormality condition (38).

According to Eq. (221), in the continuous case the bracket $\langle a_A|\alpha\rangle$ still plays the role of probability amplitude, i.e. a complex c-number whose modulus squared determines the state a_A's probability – see the last form of Eq. (120). However, for a continuous observable, the probability of finding the system exactly in a particular state is infinitesimal; instead (as was already discussed in Sec. 1.2), we should speak about the probability $dW = w(A)dA$ of finding the observable within a small interval $dA << A$ near the value A, with *probability density* $w(A) \propto |\langle a_A|\alpha\rangle|^2$. The coefficient of proportionality in this relation may be found by making a similar change from the summation to integration in the normalization condition (121):

$$\int dA\langle\alpha|a_A\rangle\langle a_A|\alpha\rangle = 1. \tag{4.223}$$

Since the total probability of the system being in *some* state should be equal to $\int w(A)dA$, this means that

Continuous spectrum: probability density

$$w(A) = \langle\alpha|a_A\rangle\langle a_A|\alpha\rangle = |\langle\alpha|a_A\rangle|^2. \tag{4.224}$$

Now let us see how we can calculate the expectation values of continuous observables, i.e. their ensemble averages. If we speak about the same observable A whose eigenstates are used as the continuous basis (or any compatible observable), everything is simple. Indeed, inserting Eq. (224) into the general statistical relation

$$\langle A\rangle = \int w(A)AdA \tag{4.225}$$

that is the obvious continuous version of Eq. (1.37), we get

$$\langle A\rangle = \int\langle\alpha|a_A\rangle A\langle a_A|\alpha\rangle dA. \tag{4.226}$$

Inserting a delta function to represent this expression formally as a double integral,

$$\langle A\rangle = \int dA\int dA'\langle\alpha|a_A\rangle A\delta(A - A')\langle a_{A'}|\alpha\rangle, \tag{4.227}$$

[53] The generalization to cases when the eigenvalue spectrum consists of both a continuous interval plus some set of discrete values, is straightforward, though leads to somewhat bulky formulas.

and using the continuous-spectrum version of Eq. (98),

$$\langle a_A|\hat{A}|a_{A'}\rangle = A\delta(A-A'), \tag{4.228}$$

we may write

$$\langle A\rangle = \int dA\int dA'\langle\alpha|a_A\rangle\langle a_A|\hat{A}|a_{A'}\rangle\langle a_{A'}|\alpha\rangle \equiv \langle\alpha|\hat{A}|\alpha\rangle, \tag{4.229}$$

so Eq. (4.125) remains valid in the continuous-spectrum case without any changes. This formula is very convenient for applications because it does not require the calculation of the eigenstates a_A, and its matrix form is valid in any basis.

Now we are ready for a discussion of the relationship between the bra-ket formalism and wave mechanics. (For the notation simplicity I will discuss its 1D version; its generalization to 2D and 3D cases is straightforward.) Let us start with postulating the (intuitively, almost evident) existence of a quantum state basis, whose ket-vectors will be called $|x\rangle$, corresponding to a certain definite value x of the particle's coordinate. Writing the trivial identity $x|x\rangle = x|x\rangle$ and comparing it with Eq. (219), we see that they do not contradict each other if we assume that x on the left-hand side of this relation is the Hermitian operator $\hat{x}$ of the particle's coordinate, in a specific representation when its action on a ket- (or bra-) vector is just the multiplication by the *c*-number x:

$$\hat{x}|x\rangle = x|x\rangle. \tag{4.230}$$

In this way, we consider vectors $|x\rangle$ to be the eigenstates of the operator $\hat{x}$. (This looks like a proof, but is actually a separate, independent postulate, no matter how plausible.)

Let me hope that the reader will excuse me if I do not pursue here strict proof that the set of all x-states is full and orthogonal,[54] so we may apply Eq. (222) to it:

$$\langle x|x'\rangle = \delta(x-x'). \tag{4.231}$$

Using this basis is called the *coordinate representation* – the term which was already mentioned several times in this course, but without explanation. In the basis of the x-states, the inner product $\langle a_A|\alpha(t)\rangle$ becomes $\langle x|\alpha(t)\rangle$, and Eq. (223) takes the following form:

$$w(x,t) = \langle\alpha(t)|x\rangle\langle x|\alpha(t)\rangle \equiv \langle x|\alpha(t)\rangle^*\langle x|\alpha(t)\rangle. \tag{4.232}$$

Comparing this formula with the basic postulate (1.22) of wave mechanics, we see that they coincide if the wavefunction of a time-dependent state α is identified with that short bracket:[55]

Wave-function as inner product

$$\boxed{\Psi_\alpha(x,t) \equiv \langle x|\alpha(t)\rangle}. \tag{4.233}$$

This key formula provides the desired connection between the bra-ket formalism and the wave mechanics, and should not be too surprising for the (thoughtful :-) reader. Indeed, Eq. (45) shows that any inner product of two state vectors describing two states is a measure of their similarity – just as the scalar product of two geometric vectors is; the orthonormality condition (38) is a particular

[54] Such proof is rather involved mathematically, but physically this fact should be evident.

[55] I do not quite like expressions like $\langle x|\Psi\rangle$ used in some papers and even textbooks. Of course, one is free to replace α with any other letter (Ψ including) to denote a quantum state, but then it is better not to use the same letter to denote the wavefunction, i.e. an inner product of two state vectors, to avoid confusion.

manifestation of this fact. In this language, the particular value (233) of a wavefunction Ψ_α at some point x and moment t characterizes "how much of a particular coordinate x" the state α contains at time t. (Of course, this informal language is too crude to reflect the fact that $\Psi_\alpha(x, t)$ is a complex function, which has not only a modulus but also an argument – the quantum-mechanical phase.)

Now let us rewrite the most important formulas of the bra-ket formalism in the wave mechanics notation. Inner-multiplying both parts of Eq. (219), written for an arbitrary operator, by the ket-vector $\langle x|$, and then inserting into the left-hand side of that relation the identity operator in the form (220) for coordinate x', we get

$$\int dx' \langle x|\hat{A}|x'\rangle\langle x'|a_A\rangle = A\langle x|a_A\rangle, \tag{4.234}$$

i.e., using the wavefunction's definition (233),

$$\int dx' \langle x|\hat{A}|x'\rangle \Psi_A(x') = A\,\Psi_A(x), \tag{4.235}$$

where, for the notation brevity, the time dependence of the wavefunction is just implied (with the capital Ψ serving as a reminder of this fact), and will be restored when needed. For a general operator, we would have to stop here, because if it does not commute with the coordinate operator, its matrix in the x-basis is not diagonal, and the integral on the left-hand side of Eq. (235) cannot be worked out explicitly. However, virtually all quantum-mechanical operators discussed in this course[56] are (*space*-) *local*: they depend on only one spatial coordinate, say x. For such an operator, we may define its coordinate representation by the following equality (valid for an arbitrary wavefunction, not only Ψ_A):

$$\hat{A}\big|_{\text{in } x}\Psi(x) \equiv \int \langle x|\hat{A}|x'\rangle\Psi(x')dx'. \tag{4.236}$$

Operator: coordinate representation

The explicit form of the coordinate representation still needs to be determined for each operator type. Let us consider, for example, the 1D version of the Hamiltonian (1.41),

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + U(\hat{x}), \tag{4.237}$$

which was the basis of all our discussions in Chapter 2. Its potential-energy part U (even if it is time-dependent as well) commutes with the operator $\hat{x}$, i.e. its matrix in the x-basis is diagonal. For such an operator, the long bracket in Eq. (236) may be transformed using Eq. (231): $\langle x|U|x'\rangle = U(x)\delta(x-x')$, so the right-hand part of this equality becomes just $U(x)\Psi(x)$. Comparing it with the left-hand part, we see that the coordinate representation of such an operator is given merely by the c-number function $U(x)$. (Eq. (230) may be viewed as just a particular manifestation of this rule.)

The situation with the momentum operator $\hat{p}_x$ (and hence the kinetic energy $\hat{p}_x^2/2m$), which do not commute with $\hat{x}$, is less evident. Let me show that its coordinate representation is given by the 1D version of Eq. (1.26), if we *postulate* that the commutation relation (2.14),

$$[\hat{x}, \hat{p}] = i\hbar\hat{I}, \qquad \text{i.e.} \quad \hat{x}\hat{p}_x - \hat{p}_x\hat{x} = i\hbar\hat{I}, \tag{4.238}$$

[56] The only substantial exception is the statistical operator $\hat{w}(x, x')$, to be discussed separately in Chapter 7.

is valid in *any* representation.[57] For that, let us consider the following matrix element: $\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle$. On one hand, we may use Eq. (238), and then Eq. (231), to write

$$\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle = \langle x|i\hbar\hat{I}|x'\rangle = i\hbar\langle x|x'\rangle = i\hbar\delta(x-x'). \tag{4.239}$$

On the other hand, since $\hat{x}|x'\rangle = x'|x'\rangle$ and $\langle x|\hat{x} = \langle x|x$, we may represent the same matrix element as

$$\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle = \langle x|x\hat{p}_x - \hat{p}_x x'|x'\rangle = (x-x')\langle x|\hat{p}_x|x'\rangle. \tag{4.240}$$

Comparing Eqs. (239) and (240), we get

$$\langle x|\hat{p}_x|x'\rangle = i\hbar\frac{\delta(x-x')}{x-x'}. \tag{4.241}$$

As it follows from the definition of the delta function,[58] all expressions involving it acquire final sense only at their integration, in our current case, as described by Eq. (236). Plugging Eq. (241) into the right-hand side of that relation, we get

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = i\hbar\int\frac{\delta(x-x')}{x-x'}\Psi(x')dx'. \tag{4.242}$$

Since the right-hand-part integral is contributed only by an infinitesimal vicinity of the point $x' = x$, we may calculate it by expanding the continuous wavefunction $\Psi(x')$ into the Taylor series in small $(x' - x)$, and keeping only two leading terms of the series, so Eq. (242) is reduced to

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = i\hbar\left[\Psi(x)\int\frac{\delta(x-x')}{x-x'}dx' - \int\delta(x-x')\frac{\partial\Psi(x')}{\partial x'}\Big|_{x'=x}dx'\right]. \tag{4.243}$$

Since the delta function may be always understood as an even function of its argument, in our case of $(x - x')$, the first term on the right-hand side is proportional to an integral of an odd function in symmetric limits and is equal to zero, and we get[59]

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = -i\hbar\frac{\partial\Psi}{\partial x}. \tag{4.244}$$

Comparing this expression with the left-hand side of Eq. (236) with $\hat{A} = \hat{p}_x$, we see that in the coordinate representation, we indeed get the 1D version of Eq. (1.26), which was used so much in Chapter 2,[60]

$$\hat{p}_x\big|_{\text{in } x} = -i\hbar\frac{\partial}{\partial x}. \tag{4.245}$$

---

[57] Another possible approach to the axiomatics of wave mechanics is to *derive* Eg. (238) by *postulating* the form, $\hat{\mathcal{T}}_X = \exp\{-i\hat{p}_x X/\hbar\}$, of the operator that shifts any wavefunction by distance X along the x-axis. In my approach, this expression will be *derived* when we need it (in Sec. 5.5), while Eq. (238) is *postulated*.

[58] If necessary, please revisit MA Sec. 14.

[59] One more useful expression of this type, which may be proved similarly, is $(\partial/\partial x)\delta(x - x') = \delta(x - x')\partial/\partial x'$.

[60] This means, in particular, that in the sense of Eq. (236), the operator of differentiation is local, despite the fact that its action on a function f may be interpreted as the limit of the fraction $\Delta f/\Delta x$, involving *two* points. (In some axiomatic systems, local operators are *defined* as arbitrary polynomials of functions and their derivatives.)

It is virtually evident (and straightforward to prove by using the Taylor expansion just as in Sec. 6) that the coordinate representation of any operator function $f(\hat{p}_x)$ is

$$f\left(-i\hbar\frac{\partial}{\partial x}\right). \tag{4.246}$$

In particular, this pertains to the kinetic energy operator in Eq. (237), so the coordinate representation of this Hamiltonian also takes the very familiar form:

$$\hat{H}\big|_{\text{in } x} = \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}\right)^2 + U(x,t) \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + U(x,t). \tag{4.247}$$

Now returning to the discussion of the general Eq. (235), and comparing its last form with that of Eq. (236), we see that for a local operator in the coordinate representation, the eigenproblem (219) takes the form

$$\hat{A}\big|_{\text{in } x}\Psi_A(x) = A\Psi_A(x), \tag{4.248}$$

Eigenproblem in *x*-representation

even if the operator $\hat{A}$ does not commute with the operator $\hat{x}$. The most important case of this coordinate-representation form of the eigenproblem (68) is the familiar Eq. (1.60) for the eigenvalues E_n of the energy of a system with a time-independent Hamiltonian.

The operator locality also simplifies the expression for its expectation value. Indeed, plugging the closure relation in the form (231) into the general Eq. (125) twice (written in the first case for x and in the second case for x'), we get

$$\langle A\rangle = \int dx \int dx' \langle\alpha(t)|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|\alpha(t)\rangle = \int dx \int dx' \Psi_\alpha^*(x,t)\langle x|\hat{A}|x'\rangle\Psi_\alpha(x',t). \tag{4.249}$$

Now, Eq. (236) reduces this result to just

$$\langle A\rangle = \int dx \int dx' \Psi_\alpha^*(x,t)\hat{A}\big|_{\text{in } x}\Psi_\alpha(x,t)\delta(x-x') \equiv \int \Psi_\alpha^*(x,t)\hat{A}\big|_{\text{in } x}\Psi_\alpha(x,t)dx. \tag{4.250}$$

i.e. to Eq. (1.23), which had to be postulated in Chapter 1 where the x-representation of the operators was just implied.

Finally, let us discuss the time evolution of the wavefunction, in the Schrödinger picture. For that, we may use Eq. (233) to calculate the (partial) time derivative of the wavefunction of some state α:

$$i\hbar\frac{\partial\Psi_\alpha}{\partial t} = i\hbar\frac{\partial}{\partial t}\langle x|\alpha(t)\rangle. \tag{4.251}$$

Since the coordinate operator $\hat{x}$ does not depend on time explicitly, its eigenstates x are stationary, and we can swap the time derivative and the time-independent bra-vector $\langle x|$. Now using the Schrödinger-picture equation (158), and then inserting the identity operator in the continuous form (220) of the closure relation, written for the coordinate eigenstates,

$$\int dx'|x'\rangle\langle x'| = \hat{I}, \tag{4.252}$$

we may continue to develop the right-hand side of Eq. (251) as

$$\langle x|i\hbar\frac{\partial}{\partial t}|\alpha(t)\rangle = \langle x|\hat{H}|\alpha(t)\rangle = \int dx'\langle x|\hat{H}|x'\rangle\langle x'|\alpha(t)\rangle = \int dx'\,\langle x|\hat{H}|x'\rangle\Psi_\alpha(x'), \tag{4.253}$$

If the Hamiltonian operator is local, we may apply Eq. (236) to the last expression, to get the familiar form (1.28) of the Schrödinger equation:

$$i\hbar\frac{\partial\Psi_\alpha}{\partial t} = \hat{H}\big|_{\text{in }x}\Psi_\alpha. \tag{4.254}$$

So, for the local operators that obey Eq. (236), we have been able to derive all the basic notions and postulates of the wave mechanics from the bra-ket formalism. Moreover, the formalism has allowed us to get a very useful equation (248) for an arbitrary local operator, which will be repeatedly used below. (In the first three chapters of this course, we have only used its particular case (1.60) for the Hamiltonian operator.)

Now let me deliver on my promise to develop a more balanced view of the de Broglie wave (4.1), which would be more respectful to the evident **r** ↔ **p** symmetry of the coordinate and momentum. Let us discuss the 1D case when the wave may be represented as

$$\psi_p(x) = a_p\exp\left\{i\frac{px}{\hbar}\right\}, \qquad \text{for all} -\infty < x < +\infty. \tag{4.255}$$

(For the sake of brevity, from this point to the end of the section, I am dropping the index x in the notation of the momentum – just as it was done in Chapter 2.) Let us have a good look at this function. Since it satisfies Eq. (248) for the 1D momentum operator (245),

$$\hat{p}\big|_{\text{in }x}\psi_p = p\psi_p, \tag{4.256}$$

ψ_p is an eigenfunction of that operator. But this means that we can also write Eq. (219) for the corresponding ket-vector:

$$\hat{p}|p\rangle = p|p\rangle, \tag{4.257}$$

and according to Eq. (233), the wavefunction (255) may be represented as

$$\psi_p(x) = \langle x|p\rangle, \qquad \text{so } \psi_p^*(x) = \langle p|x\rangle. \tag{4.258}$$

These expressions are quite remarkable in their $x \leftrightarrow p$ symmetry – which may be pursued further on. Before doing that, however, we have to discuss the normalization of such wavefunctions. Indeed, in this case, the probability density $w(x)$ of the wave (255) is constant, so its integral

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

diverges if $a_p \neq 0$. Earlier in the course, we discussed two ways to avoid this divergence. One is to use a very large but finite integration volume – see Eq. (1.31). Another way is to work with wave packets of the type (2.20), possibly of a very large length and hence a very narrow spread of the momentum values. Then the integral (259) may be required to equal 1 without any conceptual problem.

However, both these methods, while being convenient for the solution of many particular problems, violate the $x \leftrightarrow p$ symmetry and hence are unfit for our current conceptual discussion.

Instead, let us continue to identify the eigenvectors $\langle p|$ and $|p\rangle$ of the momentum with the bra- and ket-vectors $\langle a_A|$ and $|a_A\rangle$ of the general theory described at the beginning of this section. Then the normalization condition (222) becomes

$$\langle p|p'\rangle = \delta(p-p'). \tag{4.260}$$

Inserting the identity operator in the form (252), with the integration variable x' replaced by x, into the left-hand side of this equation, and using Eq. (258), we can translate this normalization rule to the wavefunction language:

$$\int dx \langle p|x\rangle\langle x|p'\rangle \equiv \int dx \psi_p^*(x)\psi_{p'}(x) = \delta(p-p'). \tag{4.261}$$

For the particular wavefunction (255), this requirement turns into the following condition:

$$a_p^* a_{p'} \int_{-\infty}^{+\infty} \exp\left\{i\frac{(p'-p)x}{\hbar}\right\}dx \equiv |a_p|^2 2\pi\hbar\delta(p-p') = \delta(p-p'), \tag{4.262}$$

so, finally, $a_p = e^{i\phi}/(2\pi\hbar)^{1/2}$, where ϕ is an arbitrary (real) phase, and Eq. (255) becomes[61]

$$\psi_p(x) = \langle x|p\rangle = \frac{1}{(2\pi\hbar)^{1/2}}\exp\left\{i\left(\frac{px}{\hbar}+\phi\right)\right\}. \tag{4.263}$$

Now let us represent an arbitrary wavefunction $\psi(x)$ as a wave packet of the type (2.20), based on the wavefunctions (263), taking $\phi = 0$ for the notation brevity, because the phase may be incorporated into the (generally, complex) envelope function $\varphi(p)$:

$$\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}}\int \varphi(p)\exp\left\{i\frac{px}{\hbar}\right\}dp\,. \tag{4.264}$$

x-representation: wavefunctions

From the mathematical point of view, this is just a 1D Fourier spatial transform, and its reciprocal is

$$\varphi(p) \equiv \frac{1}{(2\pi\hbar)^{1/2}}\int \psi(x)\exp\left\{-i\frac{px}{\hbar}\right\}dx\,. \tag{4.265}$$

p-representation: wavefunctions

These expressions are completely symmetric, and represent the same wave packet; this is why the functions $\psi(x)$ and $\varphi(p)$ are frequently called the *reciprocal representations* of a quantum state of the particle: respectively, its coordinate (x-) and momentum (p-) representations. Using Eq. (258), and Eq. (263) with $\phi = 0$, they may be recast into simpler forms,

$$\psi(x) = \int \varphi(p)\langle x|p\rangle dp, \qquad \varphi(p) = \int \psi(x)\langle p|x\rangle dx\,, \tag{4.266}$$

in which the inner products satisfy the basic postulate (14) of the bra-ket formalism:

$$\langle p|x\rangle = \frac{1}{(2\pi\hbar)^{1/2}}\exp\left\{-i\frac{px}{\hbar}\right\} = \langle x|p\rangle^*. \tag{4.267}$$

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

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

x-representation: operators

$$\hat{x}\Big|_{\text{in } x} = x, \qquad \hat{p}\Big|_{\text{in } x} = -i\hbar\frac{\partial}{\partial x}. \tag{4.268}$$

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

p-representation: operators

$$\hat{x}\Big|_{\text{in } p} = +i\hbar\frac{\partial}{\partial p}, \qquad \hat{p}\Big|_{\text{in } p} = p, \tag{4.269}$$

with the only difference of one sign, which is due to the opposite signs of the Fourier exponents in Eqs. (264) and (265). The proofs of Eqs. (269) are straightforward; for example, acting by the momentum operator on the arbitrary wavefunction (264), we get

$$\hat{p}\psi(x) = -i\hbar\frac{\partial}{\partial x}\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}}\int\varphi(p)\left(-i\hbar\frac{\partial}{\partial x}\exp\left\{i\frac{px}{\hbar}\right\}\right)dp = \frac{1}{(2\pi\hbar)^{1/2}}\int p\varphi(p)\exp\left\{i\frac{px}{\hbar}\right\}dp, \tag{4.270}$$

and similarly for the operator $\hat{x}$ acting on the function $\varphi(p)$. Comparing the final form of Eq. (270) with the initial Eq. (264), we see that the action of the operators (268) on the wavefunction ψ (i.e. the state's x-representation) gives the same results as the action of the operators (269) on the function φ (i.e. its p-representation).

It is also illuminating to have a different look at this coordinate-momentum duality. For that, notice that according to Eqs. (82)-(84), we may consider the bracket $\langle x|p\rangle$ as an element of the (infinite-size) matrix U_{xp} of the unitary transform from the x-basis to the p-basis. Let us use this fact to derive the general operator transform rule that would be a continuous version of Eq. (92). Say, we want to calculate the general matrix element of some operator known in the x-representation, in the p-representation:

$$\langle p|\hat{A}|p'\rangle. \tag{4.271}$$

Inserting two identity operators (252) written for x and x' into this bracket, and then using Eq. (258) and its complex conjugate, and also Eq. (236) (again, valid only for space-local operators!), we get

$$\begin{aligned}\langle p|\hat{A}|p'\rangle &= \int dx\int dx'\langle p|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|p'\rangle = \int dx\int dx'\psi_p^*(x)\langle x|\hat{A}|x'\rangle\psi_{p'}(x') \\ &= \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\hat{A}\Big|_{\text{in } x}\exp\left\{+i\frac{p'x}{\hbar}\right\}.\end{aligned} \tag{4.272}$$

As a sanity check, for the momentum operator itself, this relation yields:

$$\langle p|\hat{p}|p'\rangle = \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\left(-i\hbar\frac{\partial}{\partial x}\right)\exp\left\{i\frac{p'x}{\hbar}\right\} = \frac{p'}{2\pi\hbar}\int_{-\infty}^{+\infty}\exp\left\{i\frac{(p'-p)x}{\hbar}\right\}dx = p'\delta(p'-p). \tag{4.273}$$

Due to Eq. (257), this result is equivalent to the second of Eqs. (269).

From a thoughtful reader, I anticipate the following natural question: why is the momentum representation used much less often than the coordinate representation – i.e. wave mechanics? The answer is purely practical: with an important exception of the 1D harmonic oscillator (to be revisited in

Sec. 5.4), in most systems, the orbital-motion Hamiltonian (237) is not $x \leftrightarrow p$ symmetric, with the potential energy $U(\mathbf{r})$ typically being a more complex function than the kinetic energy $p^2/2m$. Because of that, it is easier to analyze such systems treating the potential energy operator just as a *c*-number multiplier, as it is in the coordinate representation – and as this was done in Chapters 1-3.

The most significant exception from this practice is the motion in a periodic potential in the presence of a coordinate-independent external force $\mathbf{F}(t)$. As was discussed in Secs. 2.7 and 3.4, in such periodic systems the eigenenergies $E_n(\mathbf{q})$, playing the role of the effective kinetic energy of the particle, may be rather involved functions of its quasimomentum $\hbar\mathbf{q}$, while its effective potential energy $U_{\text{ef}} = -\mathbf{F}(t)\cdot\mathbf{r}$ due to the additional force $\mathbf{F}(t)$ is a very simple function of coordinates. This is why detailed analyses of the quantum effects that were briefly discussed in Sec. 2.8 (the Bloch oscillations, etc.) and also such statistical phenomena as drift, diffusion, etc.[62] in solid-state theory are typically based on the momentum (or rather quasimomentum) representation.

4.8. Exercise problems

4.1. Prove that if $\hat{A}$ and $\hat{B}$ are linear operators, and C is a *c*-number, then:

(i) $\left(\hat{A}^\dagger\right)^\dagger = \hat{A}$; (ii) $\left(C\hat{A}\right)^\dagger = C^*\hat{A}^\dagger$; (iii) $\left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger$;

(iv) the operators $\hat{A}\hat{A}^\dagger$ and $\hat{A}^\dagger\hat{A}$ are Hermitian.

4.2. Prove that for any linear operators $\hat{A}, \hat{B}, \hat{C}$, and $\hat{D}$,

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

4.3. Calculate all possible binary products $\sigma_j\sigma_{j'}$ (for $j, j' = x, y, z$) of the Pauli matrices defined by Eqs. (105), and their commutators and anticommutators (defined similarly to those of the corresponding operators). Summarize the results by using the Kronecker delta and Levi-Civita permutation symbols.[63]

4.4. Calculate the following expressions,

(i) $(\mathbf{c}\cdot\boldsymbol{\sigma})^n$, and then

(ii) $(b\mathrm{I} + \mathbf{c}\cdot\boldsymbol{\sigma})^n$,

for the scalar product $\mathbf{c}\cdot\boldsymbol{\sigma}$ of the Pauli vector's matrix $\boldsymbol{\sigma} \equiv \mathbf{n}_x\sigma_x + \mathbf{n}_y\sigma_y + \mathbf{n}_z\sigma_z$ by an arbitrary *c*-number geometric vector $\mathbf{c}$, where n is a non-negative integer *c*-number and b is an arbitrary scalar *c*-number.

Hint: For Task (ii), you may like to use the binomial theorem[64] and then transform the result to a form enabling you to use the same theorem backward.

4.5. Use the solution of the previous problem to derive Eqs. (2.191) for the transparency $\mathscr{T}$ of the Dirac comb – a system of N similar, equidistant, delta-functional potential barriers.

---

[62] In this series, a brief discussion of these effects may be found in SM Chapter 6.

[63] See, e.g., MA Eqs. (13.1) and (13.2).

[64] See, e.g. MA Eq. (2.9).

4.6. Use the solution of Problem 4(i) to spell out the following matrix: $\exp\{i\theta\,\mathbf{n}\cdot\boldsymbol{\sigma}\}$, where $\boldsymbol{\sigma}$ is the 3D vector (117) of the Pauli matrices, $\mathbf{n}$ is a *c*-number geometric vector of unit length, and θ is a *c*-number scalar.

4.7. Use the solution of Problem 4(ii) to calculate $\exp\{\mathrm{A}\}$, where A is an arbitrary 2×2 matrix.

4.8. Express all elements of the matrix $\mathrm{B} \equiv \exp\{\mathrm{A}\}$ explicitly via those of the 2×2 matrix A. Spell out your result for the following matrices:

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

with real a and φ.

4.9. Prove that for arbitrary square matrices A and B,

$$\mathrm{Tr}\,(\mathrm{AB}) = \mathrm{Tr}\,(\mathrm{BA})\,.$$

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

4.10. Calculate the trace of the following 2×2 matrix:

$$\mathrm{A} \equiv (\mathbf{a}\cdot\boldsymbol{\sigma})(\mathbf{b}\cdot\boldsymbol{\sigma})(\mathbf{c}\cdot\boldsymbol{\sigma}),$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, while $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are arbitrary *c*-number vectors.

4.11. Prove that the matrix trace of an arbitrary operator does not change at its unitary transformation.

4.12. Prove that for any two full and orthonormal bases $\{u\}$ and $\{v\}$ of the same Hilbert space,

$$\mathrm{Tr}\left(|u_j\rangle\langle v_{j'}|\right) = \langle v_{j'}|u_j\rangle.$$

4.13. Is the 1D scattering matrix S, defined by Eq. (2.124), unitary? What about the 1D transfer matrix T defined by Eq. (2.125)?

4.14. Calculate the trace of the following matrix:

$$\exp\{i\mathbf{a}\cdot\boldsymbol{\sigma}\}\exp\{i\mathbf{b}\cdot\boldsymbol{\sigma}\},$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, while $\mathbf{a}$ and $\mathbf{b}$ are *c*-number geometric vectors.

4.15. Prove the following operator-vector identity:

$$(\boldsymbol{\sigma}\cdot\hat{\mathbf{r}})(\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}) = \mathrm{I}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} + i\boldsymbol{\sigma}\cdot(\hat{\mathbf{r}}\times\hat{\mathbf{p}}),$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, and I is the 2×2 identity matrix.

Hint: Take into account that the operator vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ are defined in the orbital-motion Hilbert space, different from that of the Pauli vector $\hat{\boldsymbol{\sigma}}$, and hence commute with it – even though they do not commute with each other.

4.16. Let A_j be the eigenvalues of some operator $\hat{A}$. Express the following two sums,

$$\Sigma_1 \equiv \sum_j A_j \quad \text{and} \quad \Sigma_2 \equiv \sum_j A_j^2 ,$$

via the matrix elements $A_{jj'}$ of this operator in an arbitrary basis.

4.17. Calculate $\langle \sigma_z \rangle$ of a spin–½ in the 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 the eigenstates of the Pauli matrices σ_z and σ_x, respectively.

Hint: Double-check whether your solution is general.

4.18. A spin-½ 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 oriented in an arbitrarily different direction.

4.19. In a certain basis, the Hamiltonian of a two-level system is described by the matrix

$$\mathrm{H} = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \qquad \text{with } E_1 \neq E_2 ,$$

while the operator of some observable A of this system, by the matrix

$$\mathrm{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

For the system's state with the energy definitely equal to E_1, find the possible results of measurements of the observable A and the probabilities of the corresponding measurement outcomes.

4.20. Three states $u_{1,2,3}$ form a full and orthonormal basis of a system with the following Hamiltonian

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

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

4.21. Guided by Eq. (2.203), and by the solutions of the previous problem and also of Problem 3.15, suggest a Hamiltonian describing particle's dynamics in an infinite 1D chain of similar potential wells within the tight-binding approximation, in the bra-ket formalism. Verify that its eigenstates and eigenvalues correspond to those discussed in Sec. 2.7.

4.22. In a certain full and orthonormal basis of three states $u_{1,2,3}$, operators $\hat{A}$ and $\hat{B}$ are defined by the following equalities:

$$\hat{A}|u_1\rangle = |u_3\rangle, \quad \hat{A}|u_2\rangle = |u_2\rangle, \quad \hat{A}|u_3\rangle = |u_1\rangle; \qquad \hat{B}|u_1\rangle = |u_1\rangle, \quad \hat{B}|u_2\rangle = 0, \quad \hat{B}|u_3\rangle = -|u_3\rangle .$$

(i) Prove that the operators $\hat{A}^2$ and $\hat{B}$ commute and form an orthonormal basis of their common eigenstates.

(ii) Give the most general expression for the matrix (in the *u*-basis) of an operator that would commute with $\hat{B}$.

4.23. Calculate the eigenvectors and eigenvalues of the following matrices:

$$\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

4.24. A certain state γ is an eigenstate of each of two operators, $\hat{A}$ and $\hat{B}$. What can be said about the corresponding eigenvalues *a* and *b*, if the operators anticommute?

4.25. An operator $\hat{A}$ commutes with each of two other operators $\hat{B}$ and $\hat{C}$, but these two operators do not commute: $\left[\hat{B},\hat{C}\right] \neq 0$. Prove that the full set of eigenvalues of the operator $\hat{A}$ includes some degenerate ones.

4.26. Derive the differential equation for the time evolution of the expectation value of an observable, by using (i) the Schrödinger picture and (ii) the Heisenberg picture of quantum dynamics.

4.27. At $t = 0$, a spin-½ whose interaction with an external field is described by the Hamiltonian

$$\hat{H} = \mathbf{c}\cdot\hat{\boldsymbol{\sigma}} \equiv c_x\hat{\sigma}_x + c_y\hat{\sigma}_y + c_z\hat{\sigma}_z$$

(where $c_{x,y,z}$ are real *c*-number constants, and $\hat{\sigma}_{x,y,z}$ are the Pauli operators) was in the state $\uparrow$, one of the two eigenstates of $\hat{\sigma}_z$. In the Schrödinger picture, calculate the time evolution of:

(i) the ket-vector $|\alpha\rangle$ of the spin (in any time-independent basis you like),
(ii) the probabilities to find the spin in the states $\uparrow$ and $\downarrow$, and
(iii) the expectation values of all three Cartesian components of the spin vector.

Analyze and interpret the results for the particular case $c_y = c_z = 0$.

Hint: Think about the best basis to use for the solution.

4.28. For the same system as in the previous problem, use the Heisenberg picture to calculate the time evolution of:

(i) all three Cartesian components of the spin operator $\hat{\mathbf{S}}_{\mathrm{H}}(t)$, and
(ii) the expectation values of the spin components.

Compare the latter results with those of the previous problem.

4.29. For the same system as in the two previous problems, calculate the matrix elements of the operator $\hat{\sigma}_z$ in the basis of the stationary states of the system.

4.30. In the Schrödinger picture of quantum dynamics, certain three operators satisfy the following commutation relation:

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

What is their relation in the Heisenberg picture, at a certain time instant t?

4.31. Prove the Bloch theorem given by either Eq. (3.107) or Eq. (3.108), where **R** is an arbitrary vector of the Bravais lattice (3.106).

Hint: Analyze the commutation properties of the so-called *translation operator* $\hat{\mathcal{T}}_{\mathbf{R}}$, defined by the following result of its action on an arbitrary function $f(\mathbf{r})$:

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

and apply them to an eigenfunction $\psi(\mathbf{r})$ of the stationary Schrödinger equation for a particle moving in the periodic potential described by Eq. (3.105).

4.32. A constant force F is applied to an (otherwise free) 1D particle of mass m. Calculate the stationary wavefunctions of the particle in:

(i) the coordinate representation, and
(ii) the momentum representation.

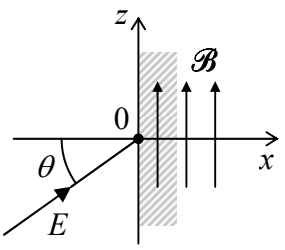
Discuss the relation between the results.

4.33. Use the momentum representation to re-solve the problem discussed at the beginning of Sec. 2.6, i.e. calculate the eigenenergy of a 1D particle of mass m, localized in a very short potential well of "weight" $\mathcal{W}$.

4.34. The momentum representation of a certain operator of orbital 1D motion is p^{-1}. Use two different approaches to find its coordinate representation.

4.35.* For a particle moving in a 3D periodic potential, develop the bra-ket formalism for the **q**-representation, in which a complex amplitude similar to a_q in Eq. (2.234) (but generalized to 3D and all energy bands) plays the role of the wavefunction. In particular, calculate the operators **r** and **v** in this representation, and use the result to prove Eq. (2.237) for the 1D case in the low-field limit.

4.36. A uniform, time-independent magnetic field $\mathcal{B} = \mathbf{n}_z\mathcal{B}$ is induced in one semi-space, while the other semi-space is field-free, with a sharp plane boundary $x = 0$ between these two regions – see figure on the right. A monochromatic beam of non-relativistic, electrically-neutral spin-½ particles with a gyromagnetic ratio $\gamma \neq 0$,[65] in a certain spin state and with a kinetic energy E, propagating within the $[x, z]$ plane, is incident on this boundary from the field-free side under angle θ. Calculate the coefficient of particle reflection from the boundary.



[65] The fact that γ may be different from zero even for electrically-neutral particles such as neutrons, is explained by the Standard Model of the elementary particles, in which a neutron "consists" (in a broad sense of this word) of three electrically-charged quarks with zero net charge.