1   Mathematical formulation of QM

In the following pages the postulates of QM are formulated and the basic mathematical background behind them is very briefly discussed. In the Quantum Physics I course they were covered (ignoring spin) for the case of a particle in a potential, with physical states being described by wave functions. Here a more general abstract presentation is given, applicable to a generic quantum system.

■ ⟶ Formal mathematical background.
　　No discussion of existence, domains, etc. ⟵ ■

■ ⟶ Elementary.
　　May be skipped. ⟵ ■

■ ⟶ Beyond the scope of this course. Included for future use if wished.
　　The examples given are useful though. ⟵ ■

These are personal notes. Although I have revised them several times, they may contain, and I am sure they do contain, mistakes and errata. Use at your own risk!

1.1. Physical states of a quantum system.

Postulate I. At each time t the state of a physical system is described by an element $\psi(t)$ of norm one of a separable, complex Hilbert space.

■ ⟶ **Reminder on Hilbert spaces**. A **complex Hilbert space** $\mathcal{H}$ is a complex vector space equipped with a scalar product that is complete with respect to the norm induced by the scalar product. The different terms in this definition stand for the following:

- $\mathcal{H}$ is a complex vector space if for all ψ, ϕ in $\mathcal{H}$ and all a, b complex, $a\psi + b\phi$ is in $\mathcal{H}$.

- A scalar product is a mapping

$$\langle \cdot | \cdot \rangle : \ \mathcal{H} \times \mathcal{H} \to \mathbb{C}$$
$$(\phi, \psi) \mapsto \langle \phi | \psi \rangle$$

  such that

  1) $\langle \psi | \psi \rangle \geq 0$ and $\langle \psi | \psi \rangle = 0$ if and only if $\psi = 0$,
  2) $\langle \chi | (a\psi + b\phi) \rangle = a \langle \chi | \psi \rangle + b \langle \chi | \phi \rangle$, and
  3) $\langle \phi | \psi \rangle^* = \langle \psi | \phi \rangle$.

  Note that properties 2) and 3) imply that

  $$\langle (a\psi + b\phi) | \chi \rangle = \langle \chi | (a\psi + b\phi) \rangle^* = a^* \langle \chi | \psi \rangle + b^* \langle \phi | \chi \rangle . \qquad (1.6)$$

- The scalar product induces a norm
  $$\|\psi\| := \sqrt{\langle \psi | \psi \rangle} .$$
  This norm generalizes the concept of modulus and defines a distance $d(\psi, \phi)$ between two arbitrary vectors ψ, ϕ as
  $$d(\psi, \phi) = \|\psi - \phi\| .$$

- A normed vector space $(\mathcal{H}, \langle \cdot | \cdot \rangle)$ is complete if every Cauchy sequence in $\mathcal{H}$ converges in $\mathcal{H}$. In other words, if for every sequence $\{\phi_n\}$ in $\mathcal{H}$ such that $d(\phi_n, \phi_m) \to 0$ when $n, m \to \infty$ there exists χ in $\mathcal{H}$ such that $d(\phi_n, \chi) \to 0$ for $n \to \infty$. Completeness is not a very restrictive requirement in the definition of a Hilbert space, since every non-complete normed vector space can be enlarged to make it complete.

Postulate I demands the Hilbert space $\mathcal{H}$ to be **separable**. This is equivalent to saying that $\mathcal{H}$ admits a countable orthonormal basis. That is, there is a numerable (not necessarily finite) set

$$\{\phi_n\}_{n \in I} \quad \text{with} \quad \langle \phi_n | \phi_m \rangle = \delta_{nm} ,$$

such that any ψ in $\mathcal{H}$ can be written as

$$\psi = \sum_n c_n \phi_n ,$$

with $c_n = \langle \psi | \phi_n \rangle$ the Fourier coefficients.

A detailed discussion of the mathematics behind these brief remarks lies outside the scope of this course and will not be presented here. For our purposes it is enough to keep in mind that a separable Hilbert space is a vector space with a norm and an orthonormal countable basis. This is the generalization of Euclidean vector spaces $\mathbb{R}^n$ to complex vector spaces with finite or infinite dimension. $\longleftarrow$ ■

Example. The space

$$L^2(\mathbb{R}) = \left\{ f \colon \mathbb{R} \to \mathbb{C} \text{ such that } \int_\infty^\infty dx \, |f(x)|^2 < \infty \right\}$$

of square integrable functions from $\mathbb{R}$ to $\mathbb{C}$ is a separable complex Hilbert space. The scalar product of f with g, both in $L^2(\mathbb{R})$, is given by

$$\langle f|g \rangle = \int_\infty^\infty dx \, f^*(x) \, g(x) \,,$$

The integral here is understood in the Lebesgue sense, but we will not worry about this in this course. As an orthonormal basis one may choose for example the eigenfunctions of the Hamiltonian for the harmonic oscillator, but there are others.

Comment. One of the major problems in the quantum description of a system is the characterization of the Hilbert space of its physical states. For the systems that we are going to be considering, the Hilbert space will either be given or very easy to find.

1.2. Observables.

Postulate II. Every observable attribute of a physical system is represented by a linear self-adjoint operator that acts on the system's Hilbert space.

■ ⟶ The operator $A \colon \mathcal{H} \to \mathcal{H}$ $(\psi \mapsto A\psi)$ is **linear** if it satisfies

$$A\left(c_1\psi_1 + c_2\psi_2\right) = c_1 A\psi_1 + c_2 A\psi_2$$

for all ψ_1, ψ_2 in $\mathcal{H}$ and all c_1, c_2 in $\mathbb{C}$.

The **adjoint** of the operator A is an operator $A^+ \colon \mathcal{H} \to \mathcal{H}$ $(\psi \mapsto A^+\psi)$ such that

$$\langle \phi | A\psi \rangle = \langle A^+\phi | \psi \rangle \quad \text{for all} \quad \psi, \phi \in \mathcal{H}\,.$$

An operator A is **self-adjoint** if $A = A^+$, i.e. if

$$\langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle \quad \text{for all} \quad \psi, \phi \in \mathcal{H}\,. \qquad \longleftarrow ■$$

Example. The position, momentum and energy of a particle moving in one dimension in a conservative potential $V(x)$, Hilbert space $L^2(\mathbb{R})$, are represented by the self-adjoint operators

$$\text{position:} \ \to \ X, \ \text{with} \ X\psi(x) = x\psi(x),$$

$$\text{momentum:} \ \to \ P, \ \text{with} \ P\psi(x) = -\mathrm{i}\hbar\,\frac{d}{dx}\,\psi(x),$$

$$\text{energy:} \ \to \ H, \ \text{with} \ H\psi(x) = \left[-\frac{\hbar^2}{2m}\,\frac{d^2}{dx^2} + V(x)\right]\psi(x).$$

Notation. In QM it is customary to use the notation

$$\psi \ \to \ |\psi\rangle$$
$$A\psi \ \to \ A|\psi\rangle$$
$$\langle \phi | A\psi \rangle = \langle A^+\phi | \psi \rangle \ \to \ \langle \phi | A | \psi \rangle\,. \tag{1.7}$$

In eq. (1.7) it is understood that A acts on the right as A, and on the left as its adjoint A^+. In other words,

$$\langle \psi | A = \langle A^+\psi |\,.$$

We will use the same letter to refer to an observable and the linear self-adjoint operator that represents it.

1.3. Results and probabilities of measurements.

> **Postulate III**. The only possible result of the measurement of an observable is one of the eigenvalues of the corresponding self-adjoint operator A. If a measurement of the observable in a system in a physical state $|\psi\rangle$ is made, the probability of obtaining the eigenvalue a_n is
>
> $$\text{Prob}_\psi(A, a_n) = \sum_i |\langle \alpha_{ni} | \psi \rangle|^2 \,, \qquad (1.8)$$
>
> where $|\alpha_{ni}\rangle$ are all the orthonormal eigenstates of A with eigenvalue a_n,
>
> $$A|\alpha_{ni}\rangle = a_n |\alpha_{ni}\rangle \,. \qquad (1.9)$$

Example. In an hydrogen atom in a state

$$|\psi\rangle = \frac{1}{\sqrt{3}} \left(|\psi_{100}\rangle + |\psi_{210}\rangle + |\psi_{21-1}\rangle \right),$$

the only possible results for the measurement of the third component of the angular momentum are 0 and $-\hbar$, with probabilities

$$\text{Prob}_\psi(L_z, 0) = \sum_{\ell=0}^{\infty} |\langle Y_\ell^0 | \psi \rangle|^2 = \left| \frac{1}{\sqrt{3}} \right|^2 + \left| \frac{1}{\sqrt{3}} \right|^2 = \frac{2}{3} \,,$$

$$\text{Prob}_\psi(L_z, -\hbar) = \sum_{\ell=1}^{\infty} |\langle Y_\ell^{-1} | \psi \rangle|^2 = \left| \frac{1}{\sqrt{3}} \right|^2 = \frac{1}{3} \,.$$

Projection and measurement. ■ ⟶ The spectral theorem, which we give without proof, states that the eigenvectors of a (compact) self-adjoint operator A that acts on a Hilbert space form an orthonormal basis. That is, the solutions $\{|\alpha_{ni}\rangle\}$ of

$$A|\alpha_{ni}\rangle = a_n |\alpha_{ni}\rangle$$

form an orthonormal basis,

$$\langle \alpha_{ni} | \alpha_{mj} \rangle = \delta_{nm}\, \delta_{ij} \,. \qquad \longleftarrow ■$$

The index i here has been introduced to emphasize that the eigenvalue a_n may have multiplicity larger than 1, so that are several eigenvectors associated with it. The eigenvectors of a given eigenvalue a_n can always be linearly combined so as to end up with orthonormal eigenvectors, which we have denoted by $|\alpha_{ni}\rangle$. Any physical state $|\psi\rangle$ can then be written as a Fourier series

$$|\psi\rangle = \sum_{ni} \underbrace{\langle \alpha_{ni} | \psi \rangle}_{c_{ni}} |\alpha_{ni}\rangle = \sum_{ni} |\alpha_{ni}\rangle \langle \alpha_{ni} | \psi \rangle. \qquad (1.10)$$

This equation is the generalization to Hilbert spaces of the decomposition

$$\mathbf{v} = \sum_{i=1}^{N} (\mathbf{a}_i \cdot \mathbf{v})\, \mathbf{a}_i$$

of a vector $\mathbf{v}$ of $\mathbb{R}^N$ in terms of the orthonormal basis $\{\mathbf{a}_i\}$ associated to the eigenvalues of an orthogonal matrix.

In eq. (1.10), the Fourier coefficient $c_{ni} = \langle \alpha_{ni} | \psi \rangle$ is the component of $|\psi\rangle$ in the direction of $|\alpha_{ni}\rangle$. Multiplication of c_{ni} with $|\alpha_{ni}\rangle$ gives the projection of the state $|\psi\rangle$ on the direction of $|\alpha_{ni}\rangle$. So

$$P_{A,ni} = |\alpha_{ni}\rangle\langle\alpha_{ni}| \tag{1.11}$$

is actually a projection operator on the direction of $|\alpha_{ni}\rangle$. The projection operator on the subspace spanned by $\{|\alpha_{ni}\rangle\}_i$ is the sum

$$P_{A,a_n} = \sum_i |\alpha_{ni}\rangle\langle\alpha_{ni}|.$$

Let us check that P_{A,a_n} is indeed a projection operator. For this to be case it must satisfy
  (i) idempotency, and
  (ii) that the sum of all projectors be the identity.
The first one follows from

$$P_{A,a_n} P_{A,a_m} = \sum_{i,j} |\alpha_{ni}\rangle \underbrace{\langle\alpha_{ni}|\alpha_{mj}\rangle}_{\delta_{nm}\delta_{ij}} \langle\alpha_{mj}| = \delta_{nm} \sum_i |\alpha_{ni}\rangle\langle\alpha_{ni}| = \delta_{nm} P_{A,a_n}$$

As for the second one, since eq. (1.10) holds for any $|\psi\rangle$ in $\mathcal{H}$, it follows that

$$\sum_{ni} |\alpha_{ni}\rangle\langle\alpha_{ni}| = \sum_n P_{A,a_n} = \mathbb{1}. \tag{1.12}$$

The probability of obtaining a_n when A is measured can then be written as

$$\text{Prob}_\psi(A, a_n) = \sum_i |\langle\alpha_{ni}|\psi\rangle|^2$$
$$= \sum_i \langle\alpha_{ni}|\psi\rangle\langle\alpha_{ni}|\psi\rangle^* = \sum_i \langle\psi|\alpha_{ni}\rangle\langle\alpha_{ni}|\psi\rangle = \langle\psi|P_{A,a_n}|\psi\rangle.$$

Exercise. Show that

$$\left(|\psi\rangle\langle\phi| \right)^+ = |\phi\rangle\langle\psi|$$

1.4. Physical state after a measurement.

Postulate IV. If a system is in a state $|\psi\rangle$ and a measurement of an observable A is made with result a_n, immediately after the measurement the system is in the state

$$|\psi'\rangle = \frac{1}{\|P_{A,a_n}\psi\|} \, P_{A,a_n}|\psi\rangle \,. \tag{1.13}$$

This is the most controversial postulate of QM. It is commonly known as "the collapse of the wave packet" and is motivated by experimental results. If one prepares many identical systems in the same state $|\psi\rangle$ and makes a measurement of observable A on them, it is observed that different results a_n are obtained, each with probability $|\langle\psi|P_{A,a_n}|\psi\rangle|^2$. This is what Postulate III states. If, in a system in which the value a_n is obtained, immediately after the measurement a second measurement is performed, it is observed that the result is always a_n, with no probabilistic distribution. This is what Postulate IV asserts, and the equation in the box is the mathematical expression of this statement.

Indeed, if the right hand side of eq. (1.13) any basis element $|\alpha_{kj}\rangle$ with $k \neq n$ were present, the probability of obtaining in the second measurement the result a_k would be different from zero. Note finally that the sate $|\psi'\rangle$ has unit norm, since $\|P_{A,a_n}\psi\|$ is the norm of $P_{A,a_n}|\psi\rangle$.

Eq. (1.13) states that measuring observable A with result a_n is equivalent to projecting the ket $|\psi\rangle$ representing the system's physical state on the subspace associated to the eigenvalue a_n.

1.5. Time evolution of a physical state.

Postulate V. In the time interval between two consecutive measurements time evolution is described by the Scrödinger equation

$$\mathrm{i}\hbar \frac{d}{dt} \left|\psi(t)\right\rangle = H(t) \left|\psi(t)\right\rangle,$$

where $H(t)$ is a self-adjoint operator, called the Hamiltonian of the system.

Comments

1. The observables of a system are represented by operators that are constant in time, unless the experimental devices that measure them explicitly change in time. If this is the case, the operators representing them should contain the change.

2. When $H(t)$ does not depend on time, it is usually called the energy operator and is the observable representing the energy of the system.

3. The Schrödinger equation is deterministic. That is, given the quantum state $\left|\psi(t_0)\right\rangle$ at time t_0, its solution gives the quantum state $\left|\psi(t)\right\rangle$ at any other time t. This does not contradict the probabilistic interpretation of QM, since any measurement at any time t is governed by postulates III and IV.

4. In contrast with position, time in QM is not an observable but a parameter.

5. Using Scrödinger's equation it is trivial to show that the norm of physical states does not change with time,

$$\frac{d}{dt} \left\langle \psi(t) | \psi(t) \right\rangle = 0 \,.$$

6. The Schrödinger equation admits stationary solutions

$$\left|\psi(t)\right\rangle = e^{-\mathrm{i}Et/\hbar} \left|\phi\right\rangle,$$

with time-independent $\left|\phi\right\rangle$, if $\left|\phi\right\rangle$ and E solve the eigenvalue problem

$$H|\phi\rangle = E|\phi\rangle \,.$$

Since H is self-adjoint, its eigenvalues E_n are real and its eigenstates $\{|\phi_n\rangle\}$ form a basis of the Hilbert space.

7. It is very easy to check (do it as an exercise) that the expectation value in $\left|\psi(t)\right\rangle$ of an observable $A(t)$ satisfies

$$\mathrm{i}\hbar \frac{d}{dt} \left\langle \psi(t)\right| A(t) \left|\psi(t)\right\rangle = \left\langle \psi(t)\right| [A(t), H(t)] \left|\psi(t)\right\rangle + \mathrm{i}\hbar \left\langle \psi(t)\right| \frac{dA(t)}{dt} \left|\psi(t)\right\rangle. \tag{1.14}$$

The observable $A(t)$ is a said to be a constant of motion if

$$[A(t), H(t)] + \mathrm{i}\hbar \frac{d}{dt} A(t) = 0 \,.$$

Note that then the expectation value of $A(t)$ in any $\left|\psi(t)\right\rangle$, given by the left hand side of eq. (1.14), vanishes. Hence the name.

---

8. The necessary and sufficient condition for two self-adjoint operators A and B to commute is that there exists a basis of vectors which are simultaneous eigenvectors of both of them. If this is the case the corresponding observables are said to be **compatible**. Denote this basis by $\{|ab\rangle\}$. If every element $|ab\rangle$ spans a one-dimensional subspace, we say that A and B form **complete set of compatible observables (CSCO)**. If not, we look for a third observable C, compatible with A and B, such that every $|abc\rangle$ spans a one-dimensional subspace. And so on.

We say that the observables A, B, C, ... form a complete set of compatible observables if

  i) The corresponding operators commute.

 ii) The basis formed by the simultaneous eigenvectors $\{|abc\ldots\rangle\}$ is unique up to phases.

iii) Condition ii) fails if one observable is removed.

9. In conservative systems the Hamiltonian does not depend on time, so it is itself a constant of motion. Every other operator that commutes with it and does not depend on time is also a constant of motion. In such systems one usually looks for complete sets of compatible observables taking as starting point the Hamiltonian.

Examples. $\{\mathbf{L}^2, L_z\}$ is a CSCO on the Hilbert space of square integrable functions defined on the sphere. $\{H, \mathbf{L}^2, L_z\}$ is a CSCO on the Hilbert space of the Hydrogen atom.

1.6. Canonical quantization.

Postulate VI. For a physical system with Cartesian coordinates $q_q, \ldots, q_N$ and corresponding conjugate momenta $p_1, \ldots, p_N$, the operators $X_1, \ldots, X_N$ and $P_1, \ldots, P_N$ that represent them must satisfy the commutation relations

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = \mathrm{i}\hbar\, \delta_{ij}.$$

We already know that for a particle in a one-dimensional potential in a state described by the wave function $\psi(x, t)$ position and momentum are represented by the self-adjoint operators $X = x$ (multiplication with x) and $P = -\mathrm{i}\hbar\dfrac{d}{dx}$. It follows that $[X, P] = \mathrm{i}\hbar$, which provides a realization of the postulate.

Things in general are not that simple and this postulate gives a general law for observables having a classical analogue. Let us see an example. If a system has an observable with classical expression $A(q_1, \ldots, q_N, p_1, \ldots, p_N; t)$, the corresponding operator is obtained by replacing the variables q_i and p_j with the operators X_i and P_j. This, however, may not give a unique result and, moreover, may lead to operators that are not self-adjoint. Think for example of the observable xp for a particle moving in a one dimension. Classically, there is no distinction between xp and px, but quantum-mechanically XP and PX are different, since $XP - PX = \mathrm{i}\hbar$. Moreover, neither XP nor PX is self-adjoint. A practical prescription, justified by experimental evidence, to deal with classical observables of this form is

$$\sum_j f_j(q_i; t)\, p_j = \frac{1}{2} \sum_j \left[\, f_j(q_i; t)\, p_j + p_j\, f_j(q_i; t)\,\right] \to \frac{1}{2} \sum_j \left[\, f_j(X_i)\, P_j + P_j\, f_j(X_i)\,\right].$$

This rule yields operators that are formally self-adjoint. If applied to xp, it gives

$$xp \to \frac{1}{2}(XP + PX),$$

which is clearly self-adjoint.

In classical mechanics, the generalized coordinates and their canonical conjugate momenta satisfy the Poisson brackets

$$\{q_i, p_j\}_{\mathrm{PB}} = \delta_{ij}$$

The prescription

$$\text{Classical mechanics} \quad \to \quad \text{Quantum mechanics}$$
$$\{A, B\}_{\mathrm{PB}} \quad \to \quad \frac{1}{\mathrm{i}\hbar}\,[A, B]$$

leads to the commutation rules in the postulate and is known as **canonical quantization**.

1.7. Uncertainty relations.

If many identical systems are prepared in the same physical state $|\psi\rangle$ and the observable A is measured, according to Postulate III, different results are obtained, each with a certain probability. The **uncertainty** or standard deviation that characterizes the dispersion of the results is

$$\Delta_\psi A := \sqrt{\langle\psi|\left(A - \langle A\rangle_\psi\right)^2|\psi\rangle}\,.$$

This can be rewritten as

$$\Delta_\psi A := \sqrt{\langle A^2\rangle_\psi - \langle A\rangle_\psi^2}\,,$$

Exercise. For a system in a quantum state $|\psi\rangle$ and two observables A and B show that

$$\Delta_\psi A \, \Delta_\psi B \,\geq\, \frac{1}{2}\left|\langle\psi|\left[A, B\right]|\psi\rangle\right|\,.$$

1.8. Unbounded operators

Equations (1.10), (1.11) and (1.12) only make sense for bounded self-adjoint operators, whose eigenvalues form a discrete set. There are relevant observables whose self-adjoint operators are unbounded and have a non-countable set of eigenvalues and eigenfunctions. For such operators, eqs. (1.10), (1.11) and (1.12) as they currently stand do not make sense. In what follows we discuss how to modify the latter.

To fix the ideas, consider the position and momentum of a particle moving in one dimension. If we use wave functions $\psi(x)$, their eigenvalues and eigenfunctions are the solutions to the equations

$$(X\psi)(x) = x\psi(x) = \lambda\psi(x) , \qquad (P\psi)(x) = -\mathrm{i}\hbar\,\frac{\psi(x)}{dx} = p\psi(x) .$$

Modulo integration constants, their solutions read

$$\psi_\lambda(x) = \delta(x - \lambda) , \qquad \psi_p(x) = \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} , \tag{1.15}$$

with real arbitrary eigenvalues λ and p. These solutions exist in the sense of distributions but not as functions. Note that there is a continuum (non-countable) set of eigenvalues and "eigenfunctions".

Denote by $|x\rangle$ an eigenstate of X with eigenvalue x, and by $|p\rangle$ an eigenstate of P with eigenvalue p,

$$X|x\rangle = x|x\rangle, \qquad P|p\rangle = p|p\rangle.$$

A natural generalization of the orthogonality condition $\langle\alpha_n|\alpha_m\rangle = \delta_{nm}$ is

$$\langle x|x'\rangle = \delta(x - x') , \qquad \langle p|p'\rangle = \delta(p - p') .$$

For the projection operators on the subspaces associated to the eigenvalues x and p, we write the continuum analogues of eqs. (1.11),

$$P_x = |x\rangle\langle x| \quad \text{and} \quad P_p = |p\rangle\langle p| ,$$

and of eqs. (1.12),

$$\int dx\, |x\rangle\langle x| = 1 \quad \text{and} \quad \int dp\, |p\rangle\langle p| = 1 .$$

If the particle is in a physical state $|\psi\rangle$, its expansion in terms of the eigenfunctions of the position operator, i.e. the analogue of the Fourier expansion (1.10), is

$$|\psi\rangle = \int_{-\infty}^{\infty} dy\, c(y)\, |y\rangle .$$

Multiplying with $\langle y'|$ and using $\langle y'|y\rangle = \delta(y' - y)$ gives

$$\langle y'|\psi\rangle = c(y') .$$

In the same way that the set of Fourier coefficients $\{c_n\}$ determine the physical state in the discrete case, the function $c(y)$ determines the physical state when this is written in terms of the eigenstates of the position operator. Let us now look at the probability that a measurement of the position gives a

---

value in $[x, x + dx]$. On the one hand, Postulate III in eq. (1.8), or rather its continuum generalization, gives

$$\mathrm{Prob}(X, dx) = \int_x^{x+dx} dy\, |\langle y|\psi\rangle|^2 \approx |\langle x|\psi\rangle|^2\, dx\,.$$

On the other, if instead of a generic $|\psi\rangle$, a square integrable wave function is used to represent the physical state, the probability is given by $|\psi(x)|^2\, dx$. Hence we conclude

$$\langle x|\psi\rangle = \psi(x)\,. \tag{1.16}$$

Note that the probability of finding the particle is 1, since

$$1 = \int_{-\infty}^{\infty} dx\, |\psi(x)|^2 = \int_{-\infty}^{\infty} dx\, |\langle x|\psi\rangle|^2 = \langle\psi|\left[\int_{-\infty}^{\infty} dx\, |x\rangle\langle x|\right]|\psi\rangle\,.$$

Take now an eigenstate $|p\rangle$ of the momentum operator P with eigenvalue p and form $\langle p|\psi\rangle$. Inserting the identity $\int dx\, |x\rangle\langle x| = 1$, one has

$$\langle p|\psi\rangle = \int dx\, \langle p|x\rangle\, \langle x|\psi\rangle \tag{1.17}$$

Now, we have already argued that $\langle x|\psi\rangle = \psi(x)$. As for $\langle p|x\rangle$ we have that, according to eq. (1.16), $\langle x|p\rangle$ is the wave function of the eigenstate $|p\rangle$ of the momentum opertator P with eigenvalue p, given in the second equation in (1.15),

$$\langle p|x\rangle = \langle x|p\rangle^* = \frac{e^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\,.$$

All in all,

$$\langle p|\psi\rangle = \int dx\, \frac{e^{-\mathrm{i}px/\hbar}}{\sqrt{2\pi\hbar}}\, \psi(x) = \text{Fourier transform } \hat{\psi}(p) \text{ of } \psi(x)\,. \tag{1.18}$$

As a consistency check, note that

$$\langle p|p'\rangle = \int_{-\infty}^{\infty} dx\, \langle p|x\rangle\langle x|p'\rangle = \int_{-\infty}^{\infty} \frac{dx}{2\pi\hbar}\, e^{-\mathrm{i}(p-p')x/\hbar} = \delta(p - p')\,.$$

Comment. These statements can be more rigorously made using spectral theory.

1.9. Density matrix

Consider a state described by an element $|\psi\rangle$ of a basis specified by the simultaneous eigenstates of a complete set of compatible observables. This state is unique up to a phase, and we have maximal information on it. States of this type are called **pure states**. Examples of them are the eigenstates $|\psi_{n\ell m}\rangle$ of the Hydrogen atom.

It may well occur, however, that we only have partial information on the state. For example when we only know the probabilities $p_1, \ldots, p_i, \ldots$ that the system be found in the pure states $|\psi_1\rangle, \ldots, |\psi_i\rangle, \ldots$, with $\sum_i p_i = 1$ and $0 \le p_i \le 1$. A state of this type is called **mixed state**, since different pure states participate in it.

Remark. The pure states $|\psi_i\rangle$ that enter a mixed state need not be orthogonal; they may be pure with respect to different complete sets of compatible observables. Think of an electron in a state of which it is known, one, that the probability of finding it with spin $1/2$ in the x-direction is $1/2$, and, two, that the probability of finding it with spin $1/2$ in the z-direction is $1/2$. The eigenstates of σ_x and σ_z are both pure states, yet they are not orthogonal to each other.

Comment. Another way to understand this distinction between pure and mixed states is the following. To make statements on probabilities, an ensemble of N identical systems is needed. If all of them are prepared in the same state, the ensemble is said to be prepared in a pure state. If however they are not all in the same state but N_i of them are in the state $|\psi_i\rangle$, with $\sum_i N_i = N$, the ensemble is in a mixed state.

Since the projector $|\psi_i\rangle\langle\psi_i|$ picks the component along $|\psi_i\rangle$ it seems plausible to describe a mixed state by the operator

$$\varrho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad p_i \ge 0, \quad \sum_{ni} p_i = 1, \tag{1.19}$$

called **density matrix** or density operator. This by itself does not say much. We must supplement it with a prescription to compute in terms of ϱ the expectation value of any observable. To do this, we note that the expectation value of an observable A in a mixed state is the statistical average, with weights p_i, of the expectation values of A in the pure states $|\psi_i\rangle$,

$$\langle A \rangle_{\text{mixed state}} = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \, .$$

Now we choose a basis $\{|\phi_n\rangle\}$ of the Hilbert space and introduce in this equation the identity as $1 = \sum_n |\phi_n\rangle\langle\phi_n|$. This gives

$$\langle A \rangle_{\text{mixed state}} = \sum_{ni} p_i \langle\psi_i|A|\phi_n\rangle\langle\phi_n|\psi_i\rangle = \sum_n p_i \langle\phi_n|\psi_i\rangle\langle\psi_i|A\phi_n\rangle = \text{tr}\,(\varrho A) \, . \tag{1.20}$$

We thus conclude that ϱ, together with $\langle A \rangle_\varrho = \text{tr}\,(\varrho A)$, represents a mixed stated.

■ $\longrightarrow$ In the last equality sign in eq. (1.20) we have used that the **trace of an operator B** is given by

$$\text{tr}\,(B) = \sum_n \langle\phi_n|B|\phi_n\rangle \, .$$

Note that this definition is independent of the basis used, since if $\{|\chi_m\rangle\}$ is a different basis, we have

$$\operatorname{tr}(B) = \sum_n \langle\phi_n|B|\phi_n\rangle = \sum_{nm}\langle\phi_n|B|\chi_m\rangle\langle\chi_m|\phi_n\rangle = \sum_{nm}\langle\chi_m|\phi_n\rangle\langle\phi_n|B|\chi_m\rangle$$

$$= \sum_m \langle\chi_m|B|\chi_m\rangle . \qquad \longleftarrow \blacksquare$$

Properties of the density matrix.

(i) $\operatorname{tr}(\varrho) = 1$, since

$$\operatorname{tr}(\varrho) = \sum_{ni}\langle\phi_n|p_i|\psi_i\rangle\langle\psi_i|\phi_n\rangle$$

$$= \sum_{ni}p_i\langle\psi_i|\phi_n\rangle\langle\phi_n|\psi_i\rangle = \sum_i p_i\langle\psi_i|\psi_i\rangle = \sum_i p_i = 1.$$

(ii) $\varrho^2 \neq \varrho$ and $\operatorname{tr}(\varrho^2) < 1$. Indeed, in general we have

$$\varrho^2 = \sum_{ij}p_i p_j|\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j| \neq \varrho . \qquad (1.21)$$

Note that if ϱ is a pure state, all the p_i vanish except of one i, say i_0, for which $p_{i_0} = 1$. In this case $\varrho^2 = \varrho$ trivially. If the states $|\psi_i\rangle$ form an orthonormal set, then $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ and again $\varrho^2 = \varrho$. This is however another realization of pure state, as we discuss below. As for the trace of ϱ^2, it is clear that

$$\operatorname{tr}(\varrho^2) = \sum_{nij}p_i p_j\langle\phi_n|\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j|\phi_n\rangle = \sum_{nij}p_i p_j\langle\psi_j|\phi_n\rangle\langle\phi_n|\psi_i\rangle\langle\psi_i|\psi_j\rangle$$

$$= \sum_{ij}p_i p_j\langle\psi_j|\psi_i\rangle\langle\psi_i|\psi_j\rangle = \sum_{ij}p_i p_j\,|\langle\psi_i|\psi_j\rangle|^2$$

$$\leq \sum_{ij}p_i p_j\,|\langle\psi_i|\psi_i\rangle|\,|\langle\psi_j|\psi_j\rangle| = \left(\sum_i p_i\right)^2 = 1 ,$$

where we have used the Schwarz inequality

$$|\langle\alpha|\beta\rangle|^2 \leq \|\alpha\|^2\|\beta\|^2 .$$

(iii) $\varrho^+ = \varrho$, since $|\psi_i\rangle\langle\psi_i|$ is self-adjoint and the p_i are real.

(iv) $\langle\chi|\varrho|\chi\rangle = \sum p_i|\langle\chi|\psi_i\rangle|^2 \geq 0$ for all $|\chi\rangle$.

Criterion for purity. As already mentioned, for a pure state, $\varrho^2 = \varrho$. This criterion make it possible to distinguish pure from mixed states.

Example. Fourier exapansions as density matrices. Assume that $\{|\phi_n\rangle\}$ are the elements of a basis associated to a CSCO. The linear combination

$$|\psi\rangle = \sum_n c_n|\phi_n\rangle , \quad c_n = \langle\phi_n|\psi\rangle , \quad \sum_n |c_n|^2 = 1$$

is a pure state since we have maximal inofrmation on it, modulo a phase. Its realization as a density matrix is

$$\varrho = \begin{pmatrix} c_1\,c_1^* & c_1\,c_2^* & \cdots & c_1\,c_n^* & \cdots \\ c_2\,c_1^* & c_2\,c_2^* & \cdots & c_2\,c_n^* & \cdots \\ \vdots & \vdots & & \vdots & \\ c_n\,c_1^* & c_n\,c_2^* & \cdots & c_n\,c_n^* & \cdots \end{pmatrix} .$$

First we check that $\varrho^2 = \varrho$, so that, according to the purity ctiterion, ϱ describes a pure state:

$$(\varrho^2)_{ij} = ij\text{-matrix element of } \varrho^2 = \sum_k c_i c_k^* c_k c_j^* = c_i c_j^* = ij\text{-matrix element of } \varrho$$

And now let us show that the expectation value $\langle \psi | A | \psi \rangle$ of any observable A in $|\psi\rangle$ can be written as $\mathrm{tr}\,(\varrho A)$. Using that the matrix elements of A in the basis $\{|\phi_n\rangle\}$ are $A_{nm} = \langle \phi_m | A | \phi_n \rangle$ we have

$$(\varrho A)_{ij} = ij\text{-matrix element of } \varrho A = \sum_k c_i c_k^* A_{kj}$$

$$= \sum_k \langle \phi_i | \psi \rangle \langle \psi | \phi_k \rangle \langle \phi_k | A | \phi_j \rangle = \left[\text{use } \sum_k |\phi_k\rangle\langle\phi_k| = 1 \right] = \langle \psi | A | \phi_j \rangle \langle \phi_i | \psi \rangle .$$

Taking now the trace, we conclude the argument

$$\mathrm{tr}\,(\varrho A) = \sum_i (\varrho A)_{ii} = \sum_i \langle \psi | A | \phi_i \rangle \langle \phi_i | \psi \rangle = \langle \psi | A | \psi \rangle .$$

Note that ϱ can be written as $\varrho = |\psi\rangle\langle\psi|$.

■ $\longrightarrow$ Any self-adjoint operator B on a complex Hilbert space such that $\langle \chi | B | \chi \rangle \geq 0$ for all $|\chi\rangle$ in $\mathcal{H}$ is called positive semidefinite. If B is positive semidefinite, all its eigenvalues are ≥ 0. Indeed, for b an eigenvalue of B with eigenvector $|\beta\rangle$, we have

$$B|\beta\rangle = b|\beta\rangle \quad \Rightarrow \quad 0 \leq \langle \beta | B | \beta \rangle = b\langle \beta | \beta \rangle = b . \qquad \longleftarrow ■$$

To include mixed sates in Postulate I, this is reformulated as

---

Postulate I'. To every system there corresponds a separable, complex Hilbert space $\mathcal{H}$. Every physical state is described by a linear operator ϱ, called density matrix, that satisfies the following properties:
$$\varrho^+ = \varrho , \qquad \mathrm{tr}\,(\varrho) = 1 , \qquad \langle \psi | \varrho | \psi \rangle \geq 0 \text{ for all } |\psi\rangle \text{ in } \mathcal{H} .$$

Eigenvalue problem for the density matrix,

$$\varrho|\eta_k\rangle = q_k|\eta_k\rangle . \tag{1.22}$$

Since ϱ is self-adjoint, its eigenvectors $\{|\eta_k\rangle\}$ form an orthonormal basis. Furthermore, since ϱ is positive semidefinite [see property (iv)], its eigenvalues are $q_k \geq 0$ and satisfy

$$1 = \mathrm{tr}\,(\varrho) = \sum_k \langle \eta_k | \varrho | \eta_k \rangle = \sum_k q_k \langle \eta_k | \eta_k \rangle = \sum_k q_k. \tag{1.23}$$

Using now $1 = \sum_i |\eta_i\rangle\langle\eta_i|$ and eq. (1.22), we conclude

$$\varrho = \varrho \sum_i |\eta_i\rangle\langle\eta_i| = \sum_i q_i |\eta_i\rangle\langle\eta_i| , \quad q_i \geq 0, \quad \sum_i q_i = 1. \tag{1.24}$$

This is analogous to eq. (1.19) but with pairwise orthogonal projections:

$$Q_i = |\eta_i\rangle\langle\eta_i|, \quad Q_i Q_j = \delta_{ij}, \quad \text{whereas}$$
$$P_i = |\psi_i\rangle\langle\psi_i|, \quad P_i P_j \text{ not necessarily} = \delta_{ij} .$$

Let us illustrate our discussion of density matrix with two simple examples.

Example 1. Same density matrix, different mixtures of pure states. Let us give an example of the ideas just explained. Consider a system with two-dimensional Hilbert space $\mathbb{C}^2$. A mixed state is prepared with weights

$$p_1 = \frac{3}{7}, \qquad p_2 = \frac{4}{7},$$

relatives to the pure states

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad |\psi_2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The corresponding density matrix is, see eq. (1.19),

$$\varrho = \frac{3}{7} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} , & \frac{1}{\sqrt{2}} \end{pmatrix} + \frac{4}{7} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 , & 0 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 11 & 3 \\ 3 & 3 \end{pmatrix}.$$

This matrix ϱ has eigenvalues and eigenvectors

$$q_1 = \frac{1}{7}, \quad |\eta_1\rangle = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}; \qquad q_2 = \frac{6}{7}, \quad |\eta_2\rangle = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix}.$$

hence can be recast [see eq. (1.24)] as

$$\varrho = \frac{1}{7} \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{10}} , & \frac{-3}{\sqrt{10}} \end{pmatrix} + \frac{6}{7} \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{10}} , & \frac{1}{\sqrt{10}} \end{pmatrix}.$$

We thus have two mixtures of pure states for the same density matrix. This is turn confirms that we do not have maximal information on the mixed state, since we do not know with certainty the participating states.

Example 2. Consider in the two-dimensional Hilbert space $\mathbb{C}^2$ the eigenstate $|x_+\rangle$ of $S_x = \frac{\hbar}{2}\sigma_x$ with eigenvalue $\hbar/2$, given by

$$|x_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

This is a pure state and can be represented by the density matrix

$$\varrho_1 = |x_+\rangle\langle x_+| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

Consider now the eigenstates $|z_+\rangle$ and $|z_-\rangle$ of $S_z = \frac{\hbar}{2}\sigma_z$ with eigenvalues $\pm\hbar/2$, given by

$$|z_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |z_-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and form the mixed state

$$\varrho_2 = \frac{1}{2} \left(|z_+\rangle\langle z_+| + |z_-\rangle\langle z_-| \right) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For these densities matrices we have:

- The same probability of obtaining $\pm\hbar/2$ in a measurement of S_z. Indeed, noting that the latter is the statistical average, with the corresponding weights, of the probabilities of obtaining $\pm\hbar/2$ in the participating pure states, we obtain

$$\mathrm{Prob}_{\varrho_1}\left(S_z, \pm\frac{\hbar}{2}\right) = \frac{1}{2}, \qquad \mathrm{Prob}_{\varrho_2}\left(S_z, \pm\frac{\hbar}{2}\right) = \frac{1}{2}.$$

- The same expectation values of S_z and S_z^2, since

$$\langle S_z \rangle_{\varrho_1} = \mathrm{tr}\left(\varrho_1 S_z\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] = 0,$$

$$\langle S_z \rangle_{\varrho_2} = \mathrm{tr}\left(\varrho_2 S_z\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] = 0$$

$$\langle S_z^2 \rangle_{\varrho_1} = \mathrm{tr}\left(\varrho_1 S_z^2\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2\right] = \frac{\hbar^2}{4},$$

$$\langle S_z^2 \rangle_{\varrho_2} = \mathrm{tr}\left(\varrho_2 S_z^2\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2\right] = \frac{\hbar^2}{4},$$

This in turn gives the same uncertainty in a measurement of S_z,

$$\Delta_{\varrho_1} S_z = \frac{\hbar}{2}, \qquad \Delta_{\varrho_2} S_z = \frac{\hbar}{2}.$$

- Different probabilities of obtaining $\pm\hbar/2$ in a measurement of S_x. In particular, we have

$$\mathrm{Prob}_{\varrho_1}\left(S_x, +\frac{\hbar}{2}\right) = 1, \qquad \mathrm{Prob}_{\varrho_1}\left(S_x, -\frac{\hbar}{2}\right) = 0,$$

$$\mathrm{Prob}_{\varrho_2}\left(S_x, +\frac{\hbar}{2}\right) = \frac{1}{2}, \qquad \mathrm{Prob}_{\varrho_1}\left(S_x, -\frac{\hbar}{2}\right) = \frac{1}{2}.$$

This shows that ϱ_1 has a definite orientation along S_x, whereas ϱ_2 does not.

- In fact for the uncertainty in S_x we have

$$\langle S_x \rangle_{\varrho_1} = \mathrm{tr}\left(\varrho_1 S_x\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] = \frac{\hbar}{2},$$

$$\langle S_x \rangle_{\varrho_2} = \mathrm{tr}\left(\varrho_2 S_x\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] = 0$$

$$\langle S_x^2 \rangle_{\varrho_1} = \mathrm{tr}\left(\varrho_1 S_x^2\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2\right] = \frac{\hbar^2}{4},$$

$$\langle S_x^2 \rangle_{\varrho_2} = \mathrm{tr}\left(\varrho_2 S_x^2\right) = \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2\right] = \frac{\hbar^2}{4},$$

and

$$\Delta_{\varrho_1} S_x = 0, \qquad \Delta_{\varrho_2} S_x = \frac{\hbar}{2}.$$

The state ϱ_1 represents a **polarized** along the x-direction beam of spins, whilst ϱ_2 represents and **unpolarized** beam.

■ ⟶ Other (non-examinable!) issues on density matrices.

Once the density matrix eigenvalue problem has been solved, one may now go further and consider an orthonormal basis $\{|\phi_k\rangle\}$ different from $\{|\eta_k\rangle\}$. These two basis are related by a transformation

$$|\phi_k\rangle = \sum_n u_{kn}|\eta_n\rangle\,, \tag{1.25}$$

where the coefficients u_{kn} cannot be arbitrary but must satisfy

$$\delta_{kl} = \langle\phi_k|\phi_l\rangle = \sum_{nm}\langle u_{kn}\eta_n|u_{lm}\eta_m\rangle = \sum_{nm}u_{kn}^*\,u_{lm}\langle\eta_n|\eta_m\rangle = \sum_n u_{ln}\,u_{kn}^*. \tag{1.26}$$

Let us arrange the u_{nk} in an infinite-dimensional matrix U,

$$U = \text{matrix, with matrix elements } u_{nk}\,.$$

The product AB of two matrices A and B has matrix elements $(AB)_{lk} = \sum_n a_{ln}b_{nk}$. Set $B = A^+$ and use that the adjoint of a matrix is the complex conjugate of its transpose, so that $b_{nk} = a_{kn}^*$. This gives $(AA^+)_{lk} = \sum_n a_{ln}a_{kn}^*$, which for $A = U$ is the right hand side in eq. (1.26), so condition (1.26) can then be written as

$$1 = UU^+\,.$$

Multiplication with U^{-1} from the left and with U from the right yields

$$1 = UU^+ \quad\Leftrightarrow\quad U^{-1} = U^+ \quad\Leftrightarrow\quad 1 = U^+U\,. \tag{1.27}$$

Analogous arguments to those given above show that eq (1.27) can be recast as

$$\delta_{kl} = \sum_n u_{ln}\,u_{kn}^* \quad\Leftrightarrow\quad (U^{-1})_{ln} = u_{nl}^* \quad\Leftrightarrow\quad \delta_{lk} = \sum_n u_{nl}^*\,u_{nk}\,. \tag{1.28}$$

To invert eq. (1.25), we multiply it with u_{kl}^*, sum over k and use eq. (1.28), thus obtaining

$$|\eta_l\rangle = \sum_k u_{kl}^*|\phi_k\rangle\,. \tag{1.29}$$

An operator U satisfying condition (1.27 is called **unitary**, and the transformation (1.25), whose inverse is given by (1.29), is called unitary transformation.

Substituting the change (1.29) in the expression (1.24) of the density matrix ϱ, we obtain

$$\varrho = \sum_k q_k|\eta_k\rangle\langle\eta_k| = \sum_{knm} q_k|u_{nk}^*\phi_n\rangle\langle u_{mk}^*\phi_m| = \sum_{nm} p_{nm}|\phi_n\rangle\langle\phi_m|\,.$$

where p_{nm} are given by

$$p_{nm} := \sum_k u_{nk}^*q_k u_{mk} = \langle\phi_n|\varrho|\phi_m\rangle\,.$$

For $n = m$ we have

$$\left. \begin{array}{l} q_k \geq 0 \;\Rightarrow\; p_{nn} = \sum_k q_k |u_{nk}|^2 \geq 0 \\[2mm] 1 = \mathrm{tr}\,(\varrho) = \sum_n p_{nn} \end{array} \right\} \;\Rightarrow\; 0 \leq p_{nn} \leq 1 \,.$$

The quantity p_{nn} is the probability of finding the system in the state pure $|\phi_n\rangle$ and is called the population of the state $|\phi_n\rangle$. For $n \neq m$, p_{nm} is in general a complex number and is known as coherence. It accounts for the interference effects between the states $|\eta_n\rangle$ and $|\eta_m\rangle$ when these are expressed as linear combinations of states $|\phi_n\rangle$.

Probabilities of measurements in mixed states. The probability of obtaining an eigenvalue a of an observable A in a measurement is the statistical average, with weights p_i, of the probabilities of obtaining a in each participating state in the mixture,

$$\begin{aligned} \mathrm{Prob}_\varrho(A, a) = \sum_i p_i \langle \psi_i | P_{A,a} | \psi_i \rangle &= \sum_i p_i \langle \psi_i | P_{A,a}^2 | \psi_i \rangle \\[2mm] &= \sum_i p_i \sum_I \langle \psi_i | P_{A,a} | \alpha_I \rangle \langle \alpha_I | P_{A,a} | \psi_i \rangle \\[2mm] &= \sum_I \sum_i p_i \langle \alpha_I | P_{A,a} | \psi_i \rangle \langle \psi_i | P_{A,a} | \alpha_{NI} \rangle \\[2mm] &= \sum_I \langle \alpha_I | P_{A,a} \varrho P_{A,a} | \alpha_I \rangle \\[2mm] &= \mathrm{tr}\,(P_{A,a} \varrho P_{A,a}) \\[2mm] &= \mathrm{tr}\,(\varrho P_{A,a}) \,, \end{aligned}$$

where we have denoted by $|\alpha_I\rangle$ the orthonormal eigenvectors of A with eigenvalue a, and have used the cyclic property of the trace, $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$.

Mixed state after a measurement. Consider a mixed ensamble described by the density matrix ϱ. Recall that this a statistical ensemble formed by identical systems prepared in, in gerneral, different states. Suppose that observable A is measured and **select** those systems which give as outcome of the measurement the eigenvalue a. We may now ask what is the density matrix describing the ensemble that results from this filtering measurement.

To answer this question, look at a system in the ensemble in the pure state $|\psi_i\rangle$. In accordance with Postulate III, the probability of obtaining a from this system is

$$p_{a_N} := \mathrm{Prob}\,(A, a) = \langle \psi_i | P_{A,a} | \psi_i \rangle = \| P_{A,a} | \psi_i \rangle \|^2 \,.$$

If p_a is zero, the system does not participate in the emerging ensemble, but if it is nonzero, it participates and in accordance with Postulate IV, after the measurement the system collapses to a state $|\psi_i'\rangle$,

$$|\psi_i\rangle \rightarrow |\psi_i'\rangle = \frac{1}{\|P_{A,a}\|} \, P_{A,a} |\psi_i\rangle \,.$$

After the measurement, we thus have the following ensemble

$$\varrho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \longrightarrow \sum_i p_i \, p_a \, \frac{P_{A,a}|\psi_i\rangle}{\|P_{A,a}|\psi_i\rangle\|} \, \frac{\langle\psi_i|P_{A,a}}{\|P_{A,a}|\psi_i\rangle\|} \,.$$

The p_a in the numerator cancels the denominator and we are left with

$$\varrho \longrightarrow P_{A,a} \left(\sum_i p_i |\psi_i\rangle\langle\psi_i| \right) P_{A,a} = P_{A,a}\varrho P_{A,a} \,.$$

Finally, to have a physical state we must normalize, i.e. divide by the norm $\text{tr}(P_{A,a}\varrho P_{A,a}) = \text{tr}(\varrho P_{A,a})$. We conclude that the mixed state that results from the measurement is

$$\varrho \longrightarrow \varrho_{A,a} := \frac{P_{A,a}\varrho P_{A,a}}{\text{tr}(\varrho P_{A,a})} \,.$$

For mixed states, Postulates III, IV and V also change their form. From our discussion above, Postulates III and IV now read

---

Postulate III'. If a physical system is in a state described by the density matrix ϱ, the probability of obtaining in a measurement of observable A one of its eigenvalues a is

$$\text{Prob}_\varrho(A, a) = \text{tr}\big(\varrho P_{A,a}\big) \,.$$

Postulate IV'. If a physical system is in a state described by the density matrix ϱ, and a filtering measurement of an observable A for its eigenvalue a is made, the system after the measurement is in a mixed state with density matrix

$$\varrho_{A,a} = \frac{P_{A,a}\varrho P_{A,a}}{\text{tr}(\varrho P_{A,a})} \,.$$

With very little extra work it is easy to arrive at Postulate V for mixed states, which replaces the Schrödinger equation with the von Neumann equation to account for time evolution. This will be covered in the Quantum Mechanics course next semester.　　←— ∎

1.10. Composite systems and entanglement

Two systems 1 and 2, with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, can be put together to form a **composite system**. The Hilbert space of the resulting system is the tensor product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, formed by ordered pairs

$$|\psi_1\rangle \otimes |\psi_2\rangle := |\psi_1\rangle |\psi_2\rangle := |\psi_1 \psi_2\rangle \quad \text{with} \quad |\psi_1\rangle \in \mathcal{H}_1, \quad |\psi_2\rangle \in \mathcal{H}_2.$$

Let us recall some properties of tensor products. For any complex number a, one has

$$a\left(|\psi_1\rangle \otimes |\psi_2\rangle\right) = \left(a\,|\psi_1\rangle\right) \otimes |\psi_2\rangle = |\psi_1\rangle \otimes \left(a\,|\psi_2\rangle\right),$$
$$\left(|\chi_1\rangle + |\phi_1\rangle\right) \otimes |\psi_2\rangle = |\chi_1\rangle \otimes |\psi_2\rangle + |\phi_1\rangle \otimes |\psi_2\rangle$$
$$|\psi_1\rangle \otimes \left(|\chi_2\rangle + |\phi_2\rangle\right) = |\psi_1\rangle \otimes |\chi_2\rangle + |\psi_1\rangle \otimes |\phi_2\rangle,$$

The scalar product of two states $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ and $|\Phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle$ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ is

$$\langle\Phi|\Psi\rangle := \left\langle |\phi_1\rangle \otimes |\phi_2\rangle, \, |\psi_1\rangle \otimes |\psi_2\rangle \right\rangle = \langle\phi_1|\psi_1\rangle_1 \langle\phi_2|\psi_2\rangle_2,$$

where $\langle\cdot|\cdot\rangle_1$ and $\langle\cdot|\cdot\rangle_2$ are the scalar products in $\mathcal{H}_1$ and $\mathcal{H}_2$.

If $\{|\phi_{1n}\rangle\}$ and $\{|\phi_{2i}\rangle\}$ are orthonormal bases of $\mathcal{H}_1$ and $\mathcal{H}_2$, any element Ψ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as

$$|\Psi\rangle := \sum_{ni} c_{ni} \, |\phi_{1n}\rangle \otimes |\phi_{2i}\rangle.$$

States that can be written as the product of one $|\psi_1\rangle$ in $\mathcal{H}_1$ and one $|\psi_2\rangle$ in $\mathcal{H}_2$ are called **separable states**. Linear combinations of two or more separable states are called **entangled states**.

If A_1 and A_2 are operators acting on $\mathcal{H}_1$ and $\mathcal{H}_2$, the operator $A_1 \otimes A_2$ acts on $\mathcal{H}_1 \otimes \mathcal{H}_2$ as

$$\left(A_1 \otimes A_2\right)\left(|\psi_1\rangle \otimes |\psi_2\rangle\right) = \left(A_1|\psi_1\rangle\right) \otimes \left(A_2|\psi_2\rangle\right).$$

Every linear operator C acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as

$$C = \sum_{ij} c_{ij} \, A_{1i} \otimes A_{2j},$$

with A_{1i} and A_{2j} operators acting on $\mathcal{H}_1$ and $\mathcal{H}_2$.

An observable A_1 acting **only** on subsystem 1, is represented by $A_1 \otimes 1$, with

$$\left(A_1 \otimes 1\right)\left(|\psi_1\rangle \otimes |\psi_2\rangle\right) = \left(A_1|\psi_1\rangle\right) \otimes |\psi_2\rangle,$$

and similarly for $A_2 = 1 \otimes A_2$.

Consider a density matrix ϱ of the composite system and an observable A_1 of subsystem 1. The expectation value of A_1 in the state ϱ is

$$\operatorname{tr}\left(\varrho A_1\right) = \sum_{ni} \langle\phi_{2i}\phi_{1n}|\varrho A_1|\phi_{1n}\phi_{2i}\rangle = \sum_n \langle\phi_{1n}|\left(\sum_i \langle\phi_{2i}|\varrho|\phi_{2i}\rangle\right) A_1|\phi_{1n}\rangle|, . \tag{1.30}$$

We introduce partial traces of ϱ over $\mathcal{H}_1$ and $\mathcal{H}_2$ as

$$\text{tr}_{\mathcal{H}_1}(\varrho) = \sum_n \langle \phi_{1n}|\varrho|\phi_{1n}\rangle =: \varrho_{\mathcal{H}_2}, \qquad \text{tr}_{\mathcal{H}_2}(\varrho) = \sum_i \langle \phi_{2i}|\varrho|\phi_{2i}\rangle =: \varrho_{\mathcal{H}_1}.$$

The two resulting quantities are called **reduced density matrices**. Note that taking the trace over a subsystem amounts to ignoring the information on that subsystem. The reduced density matrix $\varrho_{\mathcal{H}_1}$ thus describes the state of subsystem 1 when the information about the rest of the system is ignored. Coming back to $\text{tr}(\varrho A_1)$ in eq. (1.30) we have

$$\text{tr}(\varrho A_1) = \text{tr}_{\mathcal{H}_1}\left(\varrho_{\mathcal{H}_1} A_1\right).$$

Because of the loss of information, if the density matrix ϱ that we start with describes a pure state of the composite system, its reduced density matrices may be a mixed state. This happens in particular when the state of the composite system is an entangled state. Let us see an example.

Example. Consider a composite system with Hilbert space $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$. For the two subsystems consider the bases formed by the eigenstates of σ_z and f σ_x

$$\left\{ |z_+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |z_-\rangle\} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \qquad \left\{ |x_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, |x_-\rangle\} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

With the convention

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \qquad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \qquad \mathbf{v} \otimes \mathbf{w} = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ v_2 w_1 \\ v_2 w_2 \end{pmatrix},$$

the tensor product basis is formed by

$$|z_+ x_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |z_+ x_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad |z_- x_+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad |z_- x_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Assume that the composite system is in the entangled state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|z_+ x_+\rangle + |z_- x_-\rangle\right) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

The density matrix that describes this state is

$$\varrho = |\psi\rangle\langle\psi| = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 , & 1 , & 1 , & -1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}.$$

Note that $\varrho^2 = \varrho$, so that, according to the purity criterion given earlier, ϱ is a pure state. We already knew this since we have maximal information on $|\psi\rangle$. The reduced density matrix of subsystem 1 is the

partial trace

$$\varrho_{\mathcal{H}_1} = \text{tr}_{\mathcal{H}_2}(\varrho) = \langle x_+|\varrho|x_+\rangle + \langle x_-|\varrho|x_-\rangle = \langle x_+|\psi\rangle\langle\psi|x_+\rangle + \langle x_-|\psi\rangle\langle\psi|x_-\rangle$$

$$= \frac{1}{\sqrt{2}}\left(|z_+\rangle\langle z_+| + |z_-\rangle\langle z_-|\right).$$

This does not describe a pure state, since $\varrho_{\mathcal{H}_1}^2 \neq \varrho_{\mathcal{H}_1}$. Half of the times subsystem 1 is the pure state $|z_+\rangle$ and the other half is in $|z_-\rangle$.