

# Notes on Quantum Mechanics and Quantum Information

M. S. Guimaraes

## Contents

|          |                                                                    |           |
|----------|--------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Principles of quantum mechanics</b>                             | <b>3</b>  |
| 1.1      | Operator algebra and representations . . . . .                     | 6         |
| 1.2      | Finite dimensional Hilbert spaces . . . . .                        | 8         |
|          | Uncertainty relations . . . . .                                    | 12        |
| 1.3      | Infinite dimensional Hilbert spaces . . . . .                      | 13        |
|          | Position and momentum operators . . . . .                          | 13        |
| 1.4      | Harmonic oscillator . . . . .                                      | 16        |
|          | Number-angle variables . . . . .                                   | 19        |
| 1.5      | Coherent and Squeezed states of the harmonic oscillator . . . . .  | 20        |
|          | Coherent states . . . . .                                          | 20        |
|          | Squeezed states . . . . .                                          | 23        |
| 1.6      | Dynamics . . . . .                                                 | 24        |
|          | Neutrino Oscillations . . . . .                                    | 25        |
|          | Two pictures of time development . . . . .                         | 25        |
|          | Time development of the Harmonic Oscillator . . . . .              | 27        |
| 1.7      | Motion of the coherent states . . . . .                            | 27        |
| <b>2</b> | <b>Propagators and Feynman path integrals</b>                      | <b>29</b> |
| 2.1      | The propagator . . . . .                                           | 29        |
| 2.2      | Free particle propagator . . . . .                                 | 30        |
| 2.3      | Path integral representation . . . . .                             | 32        |
| 2.4      | Phase-space path integral . . . . .                                | 34        |
| 2.5      | Coherent-state path integral for the harmonic oscillator . . . . . | 34        |
| 2.6      | Spin path integral and Berry phase . . . . .                       | 35        |
| <b>3</b> | <b>Open quantum systems</b>                                        | <b>37</b> |
|          | Example: Q-bits . . . . .                                          | 39        |
|          | The Bloch sphere . . . . .                                         | 41        |
| 3.1      | Quantum statistical entropy and entanglement entropy . . . . .     | 43        |
| 3.2      | Entanglement and non-locality . . . . .                            | 45        |
|          | Bell states and the Bell basis . . . . .                           | 45        |
|          | The CHSH inequality and Bell's theorem . . . . .                   | 46        |
|          | Time evolution of open systems . . . . .                           | 48        |
|          | The no-cloning theorem . . . . .                                   | 49        |
|          | Quantum teleportation . . . . .                                    | 50        |
|          | Superdense coding . . . . .                                        | 52        |
| <b>4</b> | <b>Symmetries in quantum mechanics</b>                             | <b>53</b> |

|          |                                                                     |           |
|----------|---------------------------------------------------------------------|-----------|
| 4.1      | Unitary Operators: Translations . . . . .                           | 54        |
| 4.2      | Rotations . . . . .                                                 | 54        |
| 4.3      | Lie Groups and Lie Algebras . . . . .                               | 55        |
| 4.4      | Wigner's Theorem . . . . .                                          | 56        |
| 4.5      | Eigenvalues and eigenstates of angular momentum . . . . .           | 57        |
| 4.6      | Spin-1/2 systems . . . . .                                          | 59        |
| 4.7      | Rotation matrices . . . . .                                         | 61        |
| 4.8      | Orbital angular momentum . . . . .                                  | 62        |
| <b>5</b> | <b>Addition of angular momenta and discrete symmetries</b>          | <b>63</b> |
| 5.1      | Addition of angular momenta . . . . .                               | 63        |
| 5.2      | The Wigner-Eckart theorem . . . . .                                 | 66        |
| 5.3      | Parity . . . . .                                                    | 67        |
| 5.4      | Time reversal . . . . .                                             | 68        |
| 5.5      | Symmetries and quantum information . . . . .                        | 70        |
| <b>6</b> | <b>Approximation methods</b>                                        | <b>71</b> |
| 6.1      | Time-independent perturbation theory: non-degenerate case . . . . . | 71        |
| 6.2      | Time-independent perturbation theory: degenerate case . . . . .     | 74        |
| 6.3      | The Zeeman effect . . . . .                                         | 76        |
| 6.4      | The variational method . . . . .                                    | 77        |
| 6.5      | Time-dependent perturbation theory . . . . .                        | 79        |
| 6.6      | Fermi's golden rule and interaction with radiation . . . . .        | 82        |
| <b>7</b> | <b>Scattering theory</b>                                            | <b>84</b> |
| 7.1      | The scattering problem: kinematics and cross section . . . . .      | 85        |
| 7.2      | The Lippmann-Schwinger equation . . . . .                           | 86        |
| 7.3      | The Born approximation . . . . .                                    | 88        |
| 7.4      | Partial wave analysis . . . . .                                     | 91        |
| 7.5      | The optical theorem . . . . .                                       | 93        |
| 7.6      | The $S$ -matrix . . . . .                                           | 94        |
| 7.7      | Connection with quantum information . . . . .                       | 95        |

# 1 Principles of quantum mechanics

The usual axioms of quantum mechanics define the behavior of closed quantum systems. These are systems that are perfectly isolated from everything else. Later we will discuss open systems that can be understood as subsystems of closed systems. Main references are [1, 2, 3]

The axioms defining closed quantum systems may be presented as:

1. **Axiom 1: States** *A state is the complete description of a quantum system. A state is mathematically a ray in a Hilbert space.*

A Hilbert space is vector space over the complex numbers  $\mathbb{C}$ . A vector in the Hilbert space is written in the Dirac ket notation as  $|\psi\rangle$ .

The Hilbert space has a inner product  $\langle\psi|\phi\rangle$ . This is a map that takes two states  $|\psi\rangle$  and  $|\phi\rangle$  into a number  $\langle\psi|\phi\rangle \in \mathbb{C}$ . This map satisfies the following properties:

- (a) Positivity:  $\langle\psi|\psi\rangle > 0$ , for  $|\psi\rangle \neq 0$  (note the abuse of notation: the second 0 represents the null vector in the Hilbert space and the first 0 is in  $\mathbb{C}$ )
- (b) Linearity:  $\langle\psi|(a_1|\phi_1\rangle + a_2|\phi_2\rangle) = a_1\langle\psi|\phi_1\rangle + a_2\langle\psi|\phi_2\rangle$ , where  $a_1, a_2 \in \mathbb{C}$
- (c) skew symmetry:  $\langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle$ , where “\*” denotes complex conjugation.

There is a natural one-to-one correspondence between vectors in the Hilbert space  $|\psi\rangle$  and vectors in the dual Hilbert space  $\langle\psi|$ . These spaces are isomorphic.  $|\psi\rangle$  is called a ket,  $\langle\psi|$  is called a bra.

The Hilbert space is also demanded to be complete in the norm  $\|\psi\| \equiv \sqrt{\langle\psi|\psi\rangle}$ . This means, roughly, that if one consider a sequence of states in the Hilbert space, then the limit of this sequence must be in the Hilbert space as well.

A ray in the Hilbert space is the equivalence class of vectors that differs by the multiplication of a complex number. One can take as representative of the class the vectors of unit modulus  $\langle\psi|\psi\rangle = 1$ . This leaves still the ambiguity that two vectors that differ by a complex number of modulus one are equivalent, that is,  $|\psi\rangle$  and  $e^{i\alpha}|\psi\rangle$  are both in the same class, have unit modulus and represent the same physical state.

It is important to point out that even though vectors differing by a global phase factor represent the same physical state, such that a global phase has no physical meaning, the relative phase in a superposition is physically relevant, that is  $(a_1|\phi_1\rangle + a_2|\phi_2\rangle)$  and  $e^{i\alpha}(a_1|\phi_1\rangle + a_2|\phi_2\rangle)$  are the same physical state, but  $(a_1|\phi_1\rangle + a_2|\phi_2\rangle)$  and  $(a_1|\phi_1\rangle + a_2e^{i\alpha}|\phi_2\rangle)$  are different. In the second case, the phase  $\alpha$  is observable in interference experiments.

2. **Axiom 2: Observables** *An observable is a property of a quantum physical system that can be measured, in principle. An observable is mathematically a self-adjoint<sup>1</sup> operator acting on the vectors of the Hilbert space*

An operator  $A$  is a linear map taking vectors to vectors:  $A : |\psi\rangle \rightarrow A|\psi\rangle$

$$\begin{aligned} A : |\psi\rangle &\rightarrow A|\psi\rangle \\ A(a_1|\phi_1\rangle + a_2|\phi_2\rangle) &= a_1A|\phi_1\rangle + a_2A|\phi_2\rangle \end{aligned} \tag{1}$$

The adjoint operator  $A^\dagger$  is defined as (denoting  $|A\phi\rangle = A|\phi\rangle$ )

$$\langle\psi|A\phi\rangle = \langle A^\dagger\psi|\phi\rangle \tag{2}$$

---

<sup>1</sup>physicists also use the term Hermitian, in mathematics it is more common to reserve the name Hermitian for matrices and Self-adjoint for operators

An operator is self-adjoint if  $A^\dagger = A$ , that is

$$\langle \psi | A | \phi \rangle = \langle \phi | A | \psi \rangle^* \quad (3)$$

An important property of self-adjoint operators is that they have a spectral representation<sup>2</sup>: their eigenvectors form a complete orthonormal basis in the Hilbert space and its eigenvalues are real:

$$A = \sum_n a_n |n\rangle \langle n| \quad (4)$$

Where  $\{|n\rangle\}$  is a basis of the Hilbert space,  $a_n \in \mathbb{R}$  and  $A|n\rangle = a_n|n\rangle$ . The operator  $E_n = |n\rangle \langle n|$  is called an orthogonal projection on the state  $|n\rangle$ . It satisfies

$$\begin{aligned} E_n E_m &= \delta_{nm} E_m \\ E_n^\dagger &= E_n \end{aligned} \quad (5)$$

3. **Axiom 3: Measurement** *A measurement is a process in which information about the quantum system is acquired by an observer. Mathematically, the measurement of an observable  $A$  prepares the system in an eigenstate of  $A$  and the observer learns the value of the corresponding eigenvalue with a given probability.*

If the state of the system is  $|\psi\rangle$  immediately before the measurement and the observable  $A$  is measured, the outcome of the measurement is one of the eigenvalues  $a_n$ . A particular eigenvalue  $a_n$  will be obtained with probability

$$P(a_n) = \langle \psi | E_n | \psi \rangle = \langle \psi | n \rangle \langle n | \psi \rangle = |\langle n | \psi \rangle|^2 \quad (6)$$

Immediately after the measurement, the state of the system is

$$\frac{E_n |\psi\rangle}{|\langle n | \psi \rangle|} \quad (7)$$

If the measurement is immediately repeated, the same outcome will happen with probability one.

If one consider an ensemble of copies of the system, all of them in the same state  $|\psi\rangle$  and one performs a measurement of  $A$ , the expectation value of the result of the measurement is the mean value of the possible outcomes

$$\langle A \rangle = \sum_n a_n P(a_n) = \sum_n a_n \langle \psi | E_n | \psi \rangle = \langle \psi | A | \psi \rangle \quad (8)$$

4. **Axiom 4: Dynamics** *Dynamics is the description of how a quantum system evolves over time. Mathematically, for a closed system, time evolution is described by the action of a unitary operator on the quantum state.*

The evolution of a quantum state  $|\psi\rangle$  from a time  $t_0$  to a time  $t$  is represented by

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle \quad (9)$$

where  $U(t, t_0)$  is a unitary operator.

---

<sup>2</sup>This is actually true for any normal operator. A normal operator  $A$  is defined to satisfy  $AA^\dagger = A^\dagger A$ . An operator is normal if and only if it has an spectral representation. A self-adjoint operator is a special case.

Infinitesimal time evolution is described by the Schrodinger equation

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H(t)|\psi(t)\rangle \quad (10)$$

Where  $H(t)$  is a self-adjoint operator, called the Hamiltonian, whose eigenvalues are the possible values of the energy of the system. Infinitesimally one can write

$$|\psi(t+dt)\rangle = \left( I - \frac{i}{\hbar} H(t)dt \right) |\psi(t)\rangle \quad (11)$$

Where  $I$  is the identity operator. The operator  $U(t+dt, t) = I - \frac{i}{\hbar} H(t)dt$  is thus unitary. In the case where  $H$  is time independent one can write for the finite time evolution  $U(t, t_0) = e^{-i\frac{H}{\hbar}(t-t_0)}$

5. **Axiom 5: Composite systems** *If the Hilbert space of system  $S_1$  is  $\mathcal{H}_{S_1}$  and the Hilbert space of system  $S_2$  is  $\mathcal{H}_{S_2}$ , then the Hilbert space of the composite system  $S_1 + S_2$  is the tensor product  $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$*

The tensor product means that if the state of the system  $S_1$  is prepared in  $|\psi\rangle_{S_1}$  and the state of the system  $S_2$  is prepared in  $|\psi\rangle_{S_2}$ , then the state of the composite system is  $|\psi\rangle_{S_1} \otimes |\psi\rangle_{S_2}$ , also written as  $|\psi\rangle_{S_1}|\psi\rangle_{S_2}$  or  $|\psi_{S_1}\psi_{S_2}\rangle$ . In particular, a basis of  $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$  can be constructed from a basis  $\{|i_{S_1}\rangle\}$  of  $\mathcal{H}_{S_1}$  and a basis  $\{|n_{S_2}\rangle\}$  of  $\mathcal{H}_{S_2}$  so that each element of the corresponding basis of  $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$  is of the form  $|i_{S_1}n_{S_2}\rangle$ . This means, because of linearity of the Hilbert space, that an arbitrary vector  $|\Phi_{S_1S_2}\rangle$  describing the composite system is given by

$$|\Phi_{S_1S_2}\rangle = \sum_i \sum_n C_{in} |i_{S_1}n_{S_2}\rangle \quad (12)$$

Where  $C_{in} \in \mathbb{C}$ . Note that unless one can write all the  $C_{in}$  as  $C_{in} = a_i b_n$  for some  $a_i, b_n$ , this state cannot be written as  $|\Phi_{S_1S_2}\rangle = |\Phi\rangle_{S_1} \otimes |\Phi\rangle_{S_2}$ . When a state cannot be factorized in this way, as the product of states of each system, it is called entangled.

These axioms define the theory of quantum mechanics for closed systems. An important point to note is the inherent probabilistic nature of the theory. When one speaks of a quantum state, one is effectively referring to an ensemble. For instance, suppose you have a collection of systems, all of which in the same quantum state  $|\psi\rangle$ . You want to measure an observable  $A$ . The result of the measurement in one of the systems will be one of the eigenvalues of  $A$ . But these measurements will not give the same result in every system (unless  $|\psi\rangle$  is eigenstate of  $A$ , but let's suppose it is not). Collecting the results of the measurement of  $A$  in each member of the ensemble one is able to construct the probability distribution (6). And the mean value of the result is the ensemble average given by (8).

An illuminating example is the two-slit experiment. A particle is sent to go through two slits and afterwards its position is detected in some screen. The particle can go through any one of the slits, quantum mechanically its state after passing the slits is a superposition of states:  $|\psi\rangle = \frac{1}{\sqrt{2}}(|\text{slit 1}\rangle + |\text{slit 2}\rangle)$ . When the particle hits the screen its position  $Z$  is measured along the screen. The probability of it hitting a particular position  $z$  along the screen is

$$|\langle z|\psi\rangle|^2 = \frac{1}{2}|\langle z|\text{slit 1}\rangle|^2 + \frac{1}{2}|\langle z|\text{slit 2}\rangle|^2 + \text{Re}(\langle \text{slit 1}|z\rangle\langle z|\text{slit 2}\rangle) \quad (13)$$

where  $\text{Re}(\dots)$  means the real part of the corresponding complex expression. The last term is called an “interference term”, its origin comes from the fact that a particle coming from slit 1 or slit 2 can reach the same point  $z$  and these two possibilities cause an interference effect in the distribution, much like a wave would have. In fact, the probability distribution (13) has the exact same pattern on the screen as if it were

a wave that passed through the slits. Now, in the actual experiment, one can send one particle at a time. Obviously when it hits the screen it will leave a mark on just one point (it is a particle after all). But the idea is, if one sends another particle it will hit the screen as well, probably in another point. Sending yet another particle it will also hit the screen in some other point and so on... After sending lots of particles and registering the points marked on the screen, one can see a pattern. The resulting pattern turns out to be exactly given by the probability distribution (13). The collection of particles sent through the slit in this experiment is what constitutes the ensemble and the point that a particle leaves on the screen is the measurement of one member of the ensemble. Therefore, the interesting conclusion is that each particle is described by a quantum state that naturally has an ensemble interpretation, even though one could have measured only one particle.

## 1.1 Operator algebra and representations

We have been discussing Hilbert spaces and operators acting on them in an abstract way. In order to solve problems it is better to work in a what is known as a representation of the algebra of observables. For this we have to take a look in the underlying mathematical structure of quantum mechanics. This is a vast and complex subject, but we will only study some main points.

The set of operators of a quantum system defines an algebra, which is a set  $\mathcal{A}$  of elements such that

- If  $A, B \in \mathcal{A}$ , then  $AB \in \mathcal{A}$
- If  $A, B \in \mathcal{A}$  and  $a, b \in \mathbb{C}$ , then  $aA + bB \in \mathcal{A}$
- There is a notion of norm  $\|\cdot\|$ , such that
  - $\|A\| \geq 0$ ,  $\|A\| = 0 \Leftrightarrow A = 0$ ,  $\forall A \in \mathcal{A}$
  - $\|aA\| = |a|\|A\|$ ,  $\forall A \in \mathcal{A}, \forall a \in \mathbb{C}$
  - $\|A + B\| \leq \|A\| + \|B\|$ ,  $\forall A, B \in \mathcal{A}$
  - $\|AB\| \leq \|A\|\|B\|$ ,  $\forall A, B \in \mathcal{A}$
- There exists an involution  $\dagger$ , such that
  - $(aA + bB)^\dagger = a^*A^\dagger + b^*B^\dagger$ ,  $\forall A, B \in \mathcal{A}, \forall a, b \in \mathbb{C}$
  - $(AB)^\dagger = B^\dagger A^\dagger$ ,  $\forall A, B \in \mathcal{A}$
  - $(A^\dagger)^\dagger = A$ ,  $\forall A \in \mathcal{A}$
  - $\|A^\dagger A\| = \|A\|^2$

It is tacitly assumed that all the operators are bounded, meaning that their norm is always finite. Also, this set must be complete in the norm, meaning that if we use the norm to take limits of sequences, the limits are included in  $\mathcal{A}$ . This structure defines what is known as a  $C^*$ -algebra.<sup>3</sup> There are some related concepts that add or subtract properties to the above list. For instance, in physics, the unit operator  $\mathbf{1}$  is always included in  $\mathcal{A}$  and the structure in this case is named a *von Neumann algebra*.

The norm can be given a more concrete meaning by defining the spectrum of an operator. The spectrum,  $\text{Spec}(A)$ , of an operator  $A \in \mathcal{A}$  is the set of complex numbers  $\lambda$  such that the operator  $(\mathbf{1}\lambda - A)^{-1}$  does not exist in  $\mathcal{A}$ . One can then define

$$\rho(A) \equiv \sup |\lambda|, \quad \lambda \in \text{Spec}(A) \quad (14)$$

---

<sup>3</sup>in mathematics one usually works with  $\star$  instead of  $\dagger$  and thus this is also, more commonly, known as  $C^*$ -algebra.

That is,  $\rho(A)$  is defined as the highest value (supremum) that the modulus of  $\lambda$  can take. One can then show that for a von Neumann algebra, it follows that

$$\|A\| = \sqrt{\rho(A^\dagger A)} \quad (15)$$

For the special case when  $A$  and  $A^\dagger$  commute ( $A$  is called normal in this case) one has  $\|A\| = \rho(A)$ . One can show that this uniquely determines the norm  $\|\cdot\|$ .

This provides a definition of the abstract algebra of operators. From this construction, one can also define states as linear maps, or linear forms, taking operators to complex numbers. A linear map  $\phi : \mathcal{A} \rightarrow \mathbb{C}$  is defined such that

$$\phi(aA + bB) = a\phi(A) + b\phi(B) \quad (16)$$

The linear map is bounded if

$$|\phi(A)| \leq c\|A\| \quad (17)$$

where  $c$  is a positive number. The lowest bound for  $c$  defines the norm of  $\phi$

$$\|\phi\| \equiv \sup_{A \in \mathcal{A}} \frac{|\phi(A)|}{\|A\|} \quad (18)$$

The linear map is said to be real if  $\phi(A^\dagger) = \phi(A)^*$ . It is said to be positive if  $\phi(A^\dagger A) \geq 0$  and it is said to be normalized if  $\|\phi\| = 1$ .

A state is defined as a normalized, positive linear map. In the quantum theory interpretation, a state acting on an operator gives the expectation value, or mean value, of the observable related to the operator.

A very important result in this more abstract formulation is known as the Gelfand-Naimark-Segal (GNS)-construction: Each state  $\omega$  acting on a von Neumann algebra  $\mathcal{A}$  defines a Hilbert space  $\mathcal{H}_\omega$  and a representation  $\pi_\omega$  of  $\mathcal{A}$  given by linear operators acting on  $\mathcal{H}_\omega$ . This gives a different point of view of the whole mathematical construction of quantum mechanics. In this case, the algebra of observables and linear maps are the fundamental entities. The Hilbert space, along with a representation, arise from this underlying structure.

The GNS construction proceeds as follows:

- Noting that the algebra  $\mathcal{A}$  itself is a linear space over the complex  $\mathbb{C}$ , to make it into a Hilbert space one must provide a definition of hermitian inner product. Given a state  $\omega$  one can naturally define a inner product

$$\langle A|B \rangle \equiv \omega(A^\dagger B), \quad \forall A, B \in \mathcal{A} \quad (19)$$

By the properties of  $\omega$ , it is positive semi-definite  $\langle A|A \rangle \geq 0$  and satisfies the Schwarz inequality  $\langle A|B \rangle \leq \langle A|A \rangle \langle B|B \rangle$

- The inner product has an ambiguity related to the subset  $\mathcal{J}$  of operators  $X$  that satisfy  $\omega(X^\dagger X) = 0$ . This has the following consequence: if  $X \in \mathcal{J} \subset \mathcal{A}$ , then for all  $A \in \mathcal{A}$  it follows that  $AX \in \mathcal{J}$ . This is a consequence of the Schwarz inequality:  $\omega(X^\dagger A^\dagger AX) \leq \omega(A^\dagger AX X^\dagger A^\dagger A) \omega(X^\dagger X) = 0$ . The set  $\mathcal{J}$  is called left ideal (because it is invariant under multiplication by any other operator from the left). This set depends on the state and in this context is known as the Gelfand left ideal of  $\omega$ . One then defines the equivalence class of operators  $\mathcal{A}/\mathcal{J}$  defined by the class of operators  $[A] = \{A + X, A \in \mathcal{A}, \forall X \in \mathcal{J}\}$ . It is clear that  $\omega$  only “sees” the class  $\omega([A]^\dagger[B]) = \omega(A^\dagger B)$ .

The space  $\mathcal{H}_\omega$  is thus defined as the space  $\mathcal{A}/\mathcal{J}$  completed with limits under the norm just defined. That is, a state  $\Psi$  of  $\mathcal{H}_\omega$  is defined by a element  $[A]$  of  $\mathcal{A}/\mathcal{J}$ .

- The product defined between elements of  $\mathcal{A}$  naturally defines the action of operators on  $\mathcal{H}_\omega$ . It associates with each element  $A \in \mathcal{A}$  an operator  $\pi_\omega(A)$  acting on states of  $\mathcal{H}_\omega$  in an obvious way

$$\pi_\omega(A)|B\rangle \equiv [A][B] = [AB] \quad (20)$$

The set of operators  $\pi_\omega(A)$  forms the algebra  $\mathcal{B}(\mathcal{H})$  of bounded operators acting on  $\mathcal{H}$ .

- A cyclic state  $|\Omega\rangle$  in a Hilbert space  $\mathcal{H}_\omega$  is one such that  $\pi_\omega(A)|\Omega\rangle$  is dense in  $\mathcal{H}_\omega$ , that is, any other state in  $\mathcal{H}_\omega$  can be constructed arbitrarily well from  $|\Omega\rangle$  by acting with the operators of  $\mathcal{B}(\mathcal{H})$ . In a von Neumann algebra there is natural a cyclic vector state defined by the unit of the algebra

$$|\Omega\rangle \equiv [\mathbf{1}] \quad (21)$$

It follows that

$$\omega(A) = \omega(\mathbf{1}A\mathbf{1}) = \langle\Omega|A\Omega\rangle = \langle\Omega|\pi_\omega(A)|\Omega\rangle \quad (22)$$

Which is the usual form of the expectation value. In fact one can write expectation values for any vector  $\Psi \in \mathcal{H}_\omega$ . Any such vector defines a state

$$\omega_\Psi(A) = \langle\Psi|\pi_\omega(A)|\Psi\rangle = \langle\Omega|\pi_\omega(B)^\dagger\pi_\omega(A)\pi_\omega(B)|\Omega\rangle = \omega(B^\dagger AB) \quad (23)$$

where it was used that  $\pi_\omega(B)|\Omega\rangle$  can be made arbitrarily equal to  $|\Psi\rangle$ , because  $|\Omega\rangle$  is cyclic.

The vectors such as  $|\Psi\rangle$  are called the vector states of the representation  $\pi_\omega$ . There are also more general states that cannot be written as vectors under the representation. More generally, the set of states of the form

$$\omega_\rho(A) = \text{Tr}(\rho\pi_\omega(A)) \quad (24)$$

where  $\rho$  is a positive trace class element of the algebra  $\mathcal{B}(\mathcal{H})$  are called normal states of the von Neumann algebra  $\mathcal{B}(\mathcal{H})$ .

There is a huge body of literature about this point of view. The subject becomes particularly important in the case of infinite degrees of freedom, that is, quantum field theory, see [4].

## 1.2 Finite dimensional Hilbert spaces

A quantum system that can be characterized by a finite number of states has a finite dimensional Hilbert space representation that is isomorphic to the space of matrices. An  $N$  dimensional Hilbert space has its elements represented by  $N$ -dimensional column vectors that are acted by  $N \times N$  matrices representing the operators. Therefore, the algebra of operators in this case is the same as the algebra of  $N \times N$  (generally complex) matrices.

We already used the relevant notation but let's be more systematic here. Suppose we are interested in an observable  $A$  that can have  $N$  possible values when measured. Its spectral representation is

$$A = \sum_{n=1}^N a_n |a_n\rangle\langle a_n| \quad (25)$$

Any vector  $|\psi\rangle$  in the Hilbert space can be written as a linear combination of the eigenstates  $\{|a_n\rangle\}$  of  $A$

$$|\psi\rangle = \sum_{n=1}^N c_n |a_n\rangle \quad (26)$$

It is clear that the operator and the state are completely characterized by the set of numbers  $\{a_n\}$  and  $\{c_n\}$ , respectively. These numbers can be organized as matrices

$$A \rightarrow \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{pmatrix} \quad (27)$$

$$|\psi\rangle \rightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \quad (28)$$

where

$$a_n = \langle a_n | A | a_n \rangle; \quad c_n = \langle a_n | \psi \rangle \quad (29)$$

This construction is based on the eigenstates of the operator  $A$ , but suppose there is some other operator  $B$  that don't commute with  $A$ , that is,  $[A, B] \neq 0$ . Note that if the  $\{a_n\}$  are all equal to 1, it follows that the matrix representing  $A$  becomes the identity matrix and from (25) we can infer that  $\mathbb{1} = \sum_{n=1}^N |a_n\rangle\langle a_n|$ . Therefore, for  $B$  we can write

$$B = \mathbb{1} B \mathbb{1} = \sum_{n,n'=1}^N |a_n\rangle\langle a_n| B |a_{n'}\rangle\langle a_{n'}| = \sum_{n,n'=1}^N b_{nn'} |a_n\rangle\langle a_{n'}| \quad (30)$$

where  $b_{nn'} = \langle a_n | B | a_{n'} \rangle$ . In matrix form, this becomes

$$B \rightarrow \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1N} \\ b_{21} & b_{22} & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NN} \end{pmatrix} \quad (31)$$

But note that we could just as well express  $B$  through its spectral representation

$$B = \sum_{i=1}^M b_i |b_i\rangle\langle b_i| \quad (32)$$

and in this basis the matrix is diagonal

$$B \rightarrow \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_M \end{pmatrix} \quad (33)$$

Let's explore the case  $M \neq N$ . Suppose that  $M < N$  and we are measuring observables in a generic state  $|\psi\rangle$ . If we intend to measure  $A$ , we should look at the state of the system as a superposition of  $A$  eigenstates.

$$|\psi\rangle = \sum_{n=1}^N c_n |a_n\rangle \quad (34)$$

But we could also be interested in measuring  $B$  instead and it should be possible to also express the state as a superposition of  $B$  eigenstates.

$$|\tilde{\psi}\rangle = \sum_{i=1}^M d_i |b_i\rangle \quad (35)$$

These vectors should be related somehow. Since the  $\{|b_i\rangle\}$  are a basis related to the possible eigenstates of  $B$  and there are  $M < N$  of them, they define an  $M$ -dimensional subspace of the original  $N$ -dimensional Hilbert space. There should be an operator mapping the  $\{|a_n\rangle\}$  into the  $\{|b_i\rangle\}$ , that is

$$V|a_n\rangle = \sum_{i=1}^M V_{in} |b_i\rangle \quad (36)$$

The numbers  $V_{in} = \langle b_i | a_n \rangle$  constitute a rectangular  $M \times N$  matrix. Since  $M < N$ , the vectors  $\{|b_i\rangle\}$  span only an  $M$ -dimensional subspace of  $\mathcal{H}_N$ , so  $\sum_{i=1}^M |b_i\rangle\langle b_i| = P \neq \mathbb{1}_N$ , where  $P = \sum_{i=1}^M |b_i\rangle\langle b_i|$  is the *orthogonal projector* onto the subspace. An orthogonal projector satisfies  $P^2 = P$  and  $P^\dagger = P$ ; it acts as the identity on states within the subspace and annihilates states orthogonal to it. In contrast, the eigenstates of  $A$  form a complete basis of  $\mathcal{H}_N$ , so  $\sum_{n=1}^N |a_n\rangle\langle a_n| = \mathbb{1}_N$ . Using these resolutions of identity one finds

$$(V^\dagger V)_{mn} = \sum_{i=1}^M \langle a_m | b_i \rangle \langle b_i | a_n \rangle = \langle a_m | P | a_n \rangle \equiv P_{mn}; \quad (V^\dagger V = P) \quad (37)$$

$$(VV^\dagger)_{ij} = \sum_{n=1}^N \langle b_i | a_n \rangle \langle a_n | b_j \rangle = \delta_{ij}; \quad (VV^\dagger = \mathbb{1}_M) \quad (38)$$

The map  $V^\dagger$  (an  $N \times M$  matrix embedding  $\mathcal{H}_M$  into  $\mathcal{H}_N$ ) is called an *isometry*: it preserves inner products, satisfying  $(V^\dagger)^\dagger V^\dagger = VV^\dagger = \mathbb{1}_M$ , and its image in  $\mathcal{H}_N$  is characterized by the projector  $V^\dagger(V^\dagger)^\dagger = V^\dagger V = P$ . We can also write  $V$  more conveniently in operator form

$$V = \sum_{n=1}^N \sum_{i=1}^M V_{in} |b_i\rangle\langle a_n| \quad (39)$$

It then follows that

$$V|\psi\rangle = \sum_{n=1}^N c_n V|a_n\rangle = \sum_{n=1}^N \sum_{i=1}^M c_n V_{in} |b_i\rangle \quad (40)$$

This should be the state  $|\tilde{\psi}\rangle$ , up to a phase that we consider equal to 1. We therefore obtain

$$d_i = \sum_{n=1}^N c_n V_{in} \quad (41)$$

In matrix form, the quantities discussed above are

$$|\psi\rangle \rightarrow \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{pmatrix}; \quad |\tilde{\psi}\rangle \rightarrow \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{pmatrix} \quad (42)$$

$$V \rightarrow \begin{pmatrix} V_{11} & V_{12} & V_{13} & \cdots & V_{1N} \\ V_{21} & V_{22} & V_{23} & \cdots & V_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{M1} & V_{M2} & V_{M3} & \cdots & V_{MN} \end{pmatrix} \quad (43)$$

with

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} & V_{13} & \cdots & V_{1N} \\ V_{21} & V_{22} & V_{23} & \cdots & V_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{M1} & V_{M2} & V_{M3} & \cdots & V_{MN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{pmatrix} \quad (44)$$

From the properties of  $V$ , when  $M = N$  the matrix  $V$  is square and  $VV^\dagger = \mathbb{1}_N$ . Since  $V$  is a square matrix possessing a right inverse ( $V^\dagger$ ), it is invertible, and for square matrices any one-sided inverse is also a two-sided inverse. Therefore  $V^\dagger V = \mathbb{1}_N$  as well, and  $V$  is *unitary*. This is a special feature of finite-dimensional spaces: any isometry from a finite-dimensional space to itself is automatically unitary. In infinite dimensions this fails — for example, the one-sided shift operator  $S|n\rangle = |n+1\rangle$  on  $\ell^2(\mathbb{Z}_{\geq 0})$  satisfies  $S^\dagger S = \mathbb{1}$  but  $SS^\dagger \neq \mathbb{1}$ . When  $M = N$  one therefore interprets  $V$  as a change of basis, and the two Hilbert spaces are isomorphic.

Now suppose  $[A, B] = 0$ . It follows that

$$AB|a_n\rangle = BA|a_n\rangle = a_n B|a_n\rangle \quad (45)$$

This means that the state  $B|a_n\rangle$  has the same  $A$  eigenvalue as  $|a_n\rangle$ . The simplest such situation is if  $B = f(A)$ , for some arbitrary function  $f$ . In this case  $B|a_n\rangle = f(a_n)|a_n\rangle$ . More generally, we can express  $|a_n\rangle$  in the  $B$  basis

$$|a_n\rangle = \sum_{m=1}^J c_m^{(a_n)} |b_m\rangle \quad (46)$$

This means that  $A$  has in fact  $J$  eigenstates  $|b_m\rangle$ , with  $m = 1, \dots, J$ , associated with its eigenvalue  $a_n$ . This is called a degeneracy. One says that the eigenvalue  $a_n$  is degenerate, being associated with more than one eigenstate. The eigenvalues of the operator  $B$  provide then a label to differentiate the states among the degenerate subset. It is common to write the state associated with eigenvalue  $a_n$  of  $A$  and eigenvalue  $b_m$  of  $B$  in the form  $|a_n, b_m\rangle$  and rewrite (46) as

$$|a_n\rangle = \sum_{m=1}^J c_m^{(a_n)} |a_n, b_m\rangle \quad (47)$$

In matrix form, considering a simple example where only the first two eigenvalues of  $A$  are equal, one would find something like

$$A \rightarrow \begin{pmatrix} a_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & a_1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & a_2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_N \end{pmatrix} \quad (48)$$

$$B \rightarrow \begin{pmatrix} b_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & b_2 & \cdots & 0 & 0 & 0 \\ 0 & 0 & b_3 & \cdots & 0 & 0 \\ 0 & 0 & 0 & b_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & b_N \end{pmatrix} \quad (49)$$

Physically, what is happening is the following. Suppose we only have access to the measurement of  $A$ . The spectrum in the example above would be composed of  $N - 1$  eigenvalues. The states would be labeled only by the  $a_n$  and we would consider the Hilbert space of the system as having dimension  $N - 1$ . But then suppose we discover some new property of the system associated with a new observable  $B$  that does not interfere in measurements of  $A$ , that is, that commutes with  $A$ . And it happens that measuring  $B$  with the system in the state  $|a_1\rangle$  we sometimes find  $b_1$  and sometimes  $b_2$ . Therefore, we would conclude that there is one more degree of freedom and the Hilbert space is larger. We should label states by the eigenvalues of  $A$  and  $B$ , and the particular state  $|a_1\rangle$  we thought as unique splits into two physically distinct states  $|a_1, b_1\rangle$  and  $|a_1, b_2\rangle$ . The Hilbert space is now  $N$  dimensional. Obviously, this could go on and the we could find new observables enlarging the Hilbert space. Anyway, at each moment we have a particular set of commuting observables whose eigenvalues can be used to label the states, the maximum such set (in the sense of including all known commuting observables) is called “complete set of commuting observables”.

### Uncertainty relations

Returning to the case of non-commuting operators, it is clear that measurements of one operator will perturb the other one. This can be quantified by the dispersion. Define

$$\Delta_\psi A = A - \langle\psi|A|\psi\rangle \quad (50)$$

such that the dispersion is defined by

$$\langle\psi|(\Delta_\psi A)^2|\psi\rangle = \langle\psi|(A - \langle\psi|A|\psi\rangle)^2|\psi\rangle = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2 \quad (51)$$

It quantifies, on the average, how much a measured value of  $A$  in a state  $|\psi\rangle$  deviates from the mean value of  $A$  in this same state. One can use this to understand how the noncommutativity of operators is related to the disturbance on the measurement of them. First note that the inner product satisfies the Schwarz inequality. Considering two arbitrary states  $|\psi\rangle$  and  $|\phi\rangle$

$$|\langle\psi|\phi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle \quad (52)$$

This is easily proved by noting that  $(\langle\psi| + \lambda^*\langle\phi|)(|\psi\rangle + \lambda|\phi\rangle) \geq 0$  and choosing  $\lambda = \frac{-\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}$ . Therefore

$$\langle\psi|(\Delta_\psi A)^2|\psi\rangle\langle\psi|(\Delta_\psi B)^2|\psi\rangle \geq |\langle\psi|\Delta_\psi A\Delta_\psi B|\psi\rangle|^2 \geq \frac{1}{4}|\langle\psi|[\Delta_\psi A, \Delta_\psi B]|\psi\rangle|^2 = \frac{1}{4}|\langle\psi|[A, B]|\psi\rangle|^2 \quad (53)$$

where it was used that  $\langle\psi|\Delta_\psi A\Delta_\psi B|\psi\rangle = \frac{1}{2}\langle\psi|[\Delta_\psi A, \Delta_\psi B]|\psi\rangle + \frac{1}{2}\langle\psi|\{\Delta_\psi A, \Delta_\psi B\}|\psi\rangle$  implies

$$|\langle\psi|\Delta_\psi A\Delta_\psi B|\psi\rangle|^2 = \frac{1}{4}|\langle\psi|[\Delta_\psi A, \Delta_\psi B]|\psi\rangle|^2 + \frac{1}{4}|\langle\psi|\{\Delta_\psi A, \Delta_\psi B\}|\psi\rangle|^2 \geq \frac{1}{4}|\langle\psi|[\Delta_\psi A, \Delta_\psi B]|\psi\rangle|^2 \quad (54)$$

because  $\langle\psi|[\Delta_\psi A, \Delta_\psi B]|\psi\rangle$  is a pure imaginary number and  $\langle\psi|\{\Delta_\psi A, \Delta_\psi B\}|\psi\rangle$  is a pure real number. Also,  $[\Delta_\psi A, \Delta_\psi B] = [A, B]$ .

The result

$$\langle\psi|(\Delta_\psi A)^2|\psi\rangle\langle\psi|(\Delta_\psi B)^2|\psi\rangle \geq \frac{1}{4}|\langle\psi|[A, B]|\psi\rangle|^2 \quad (55)$$

known generally as the uncertainty relation, implies that there is a lower limit for the minimum uncertainties when measuring quantities whose corresponding operators don't commute.

The operators acting on a finite Hilbert space can always be represented as matrix and constitute an algebra known as type  $I_N$ , if the dimension of the Hilbert space is  $N$ .

### 1.3 Infinite dimensional Hilbert spaces

Suppose now that the Hilbert space is infinite dimensional  $N \rightarrow \infty$ . In this case the description follows the same steps as in the case of finite dimensional Hilbert spaces. But one has to be more careful because of convergence issues. In fact two new postulates must be included to make sense of a Hilbert space with infinite dimension: It must be complete in the norm and it must be separable. [3].

Complete in the norm means that every Cauchy sequence is convergent. The sequence of vectors  $|\psi_1\rangle, |\psi_2\rangle, \dots$  is Cauchy if for each  $\epsilon > 0$  there is a  $N(\epsilon)$  such that  $|||\psi_n\rangle - |\psi_m\rangle|| < \epsilon$ , for all  $n, m > N$ . In a complete space, every such sequence converges to a vector that belongs to the space.

Separable means that there is a sequence of vectors  $|\psi_1\rangle, |\psi_2\rangle, \dots$  in the space that is everywhere dense. This is equivalent of saying that every vector in the space is arbitrarily close to a countable set of vectors. In other words, there is a countable basis.

A particular example is that of the harmonic oscillator. The Hilbert space is composed of the square integrable functions of position. The basis are the Hermite polynomials that make up for a infinite countable basis. Every square integrable function can be written as a linear combination of Hermite polynomials.

But there are more subtleties related to infinite dimensional Hilbert spaces. Take the position operator, for instance. Its eigenvalues are all the real numbers. This gives rise to an uncountable set of eigenstates. Furthermore, the position operator is unbounded, meaning that acting on a vector it produces a vector that may not have a finite norm. In fact, the norm of the eigenstate of position is formally infinite. von Neumann had many reservations with these infinities and was able to formulate a theory of quantum mechanics where such problems can be avoided, paying the price of having a less intuitive approach.

Dirac's formulation [5], heavily criticized by von Neumann for its lacking of mathematical rigor, was eventually rescued by the theory of distributions. Its formulation can be made mathematically precise by the so called rigged Hilbert space formalism. It starts with a well defined Hilbert space  $\mathcal{H}$  and a subspace  $\mathcal{K}$  of states that make sense physically, that is, having finite norm and furnishing finite expectation values for any observable, even unbounded ones, such as position. Then one consider the space  $\mathcal{H}^*$ , dual to  $\mathcal{H}$ , where the "bras" live. This is the space of functions that take vectors from  $\mathcal{H}$  into numbers. This space is then augmented to a larger space  $\mathcal{K}^+$  where distributions are allowed. Noticing  $\mathcal{H}^*$  is isomorphic to  $\mathcal{H}$ , the structure becomes

$$\mathcal{K} \subset \mathcal{H} \subset \mathcal{K}^+ \tag{56}$$

this is known as the rigged Hilbert space structure or the Gelfand triple.

We will not dive into this formalism, that can be quite heavy mathematically, but the idea here is to assure that working with the Dirac formalism can be made mathematically rigorous. One just have to be careful when dealing with quantities that don't live in the physical subset  $\mathcal{K}$  of the Hilbert space, as Dirac was, even not knowing about all these future developments. For instance, the momentum eigenstate can be extremely useful even though it is not a physical state. Let's discuss this in more details.

### Position and momentum operators

Let's consider the case of one space dimension, the generalization for more dimensions is straightforward.

The position operator  $X$  can be formally regarded as an observable, thus self-adjoint, and the eigenvalue equation can be written as

$$X|x\rangle = x|x\rangle \quad (57)$$

As already mentioned, in general one would like to regard the eigenvalues taking values in whole real line  $-\infty < x < \infty$ . The main problem is that this makes the eigenvectors ill defined as vectors in a Hilbert space, because they are not normalizable. In fact, consider a eigenvalue  $x \neq x'$ . A general property of the corresponding eigenstate  $|x'\rangle$  is that it should be orthogonal to  $|x\rangle$ .

$$\langle x'|X|x\rangle = x\langle x'|x\rangle = x'\langle x'|x\rangle \Rightarrow (x - x')\langle x'|x\rangle = 0 \quad (58)$$

So if  $x \neq x'$  then  $\langle x'|x\rangle = 0$ . We would like to also conclude that when  $x = x'$  then  $\langle x|x\rangle = 1$ , but this would be inconsistent. In fact,  $x$  is a continuous variable, so  $\langle x'|x\rangle$  should define a function  $f_{x'}(x)$ , but given its properties, that is,  $f_{x'}(x) = 0$  for  $x \neq x'$  and  $f_x(x) = 1$ , it follows that

$$\int_{-\infty}^{\infty} dx |f_{x'}(x)|^2 = 0 \quad (59)$$

from where one can conclude, from the general properties of the Lebesgue integral that  $f_{x'}(x) = 0$  for every  $x$ . That is, the set of points where  $f_{x'}(x) \neq 0$  is a set of measure zero, meaning that the function is indistinguishable from zero. This conclusion holds as long as one insists in treating  $f_{x'}(x)$  as a well defined bounded function. Dirac's idea was to consider it as a singular function such that it displays a particular infinite limit as  $x = x'$ . This reinterpretation of  $f_{x'}(x)$  necessarily makes it not square integrable, in the sense that the integral of  $|f_{x'}(x)|^2$  diverges, so that  $f_{x'}(x) = \langle x'|x\rangle$  is not a valid physical quantum amplitude, because it is impossible to normalize it and thus to attach a probabilistic interpretation to this quantity. But it is possible to attach meaning  $\langle x'|x\rangle$  as a map that takes proper physical amplitudes that are functions of  $x$  and evaluates them at the point  $x'$ . This is the property of a distribution (that lives in  $\mathcal{K}^+$  as discussed above). More precisely, one demands that

$$\int_{-\infty}^{\infty} dx f_{x'}(x)g(x) = g(x') \quad (60)$$

for any function  $g$ . In particular

$$\int_{-\infty}^{\infty} dx f_{x'}(x) = 1 \quad (61)$$

An object with this property became known as Dirac delta function and represented in the form

$$\langle x'|x\rangle = \delta(x - x') \quad (62)$$

This interpretation lends a very intuitive power allowing one to work with observables with continuum spectrum as if one were dealing with the more well behaved observables with discrete spectrum. The spectral representation, for instance, is written as

$$X = \int dx x|x\rangle\langle x| \quad (63)$$

such that, even though  $|x\rangle$  is not a physical state, observable quantities related to this kind of expression are well defined. Consider for instance the expectation value of the position of a system described by the state  $|\psi\rangle$ :

$$\langle\psi|X|\psi\rangle = \int dx x\langle\psi|x\rangle\langle x|\psi\rangle = \int dx x|\langle x|\psi\rangle|^2 \quad (64)$$

This is a well defined quantity as long as the integral converge, which is true if  $\langle x|\psi\rangle = \psi(x)$  is a well behaved function. Note also that

$$\mathbb{1} = \int dx |x\rangle\langle x| \quad (65)$$

so that

$$|\psi\rangle = \mathbb{1}|\psi\rangle = \int dx |x\rangle\langle x|\psi\rangle = \int dx \psi(x)|x\rangle \quad (66)$$

This means that the state  $|\psi\rangle$  can be seen as a superposition of possible positions and the coefficients, known as “wave functions” are the amplitudes whose square modulus furnishes the probability density of finding the system in the position  $x$ . Note also that

$$\langle\psi|\psi\rangle = \langle\psi|\mathbb{1}|\psi\rangle = \int dx \langle\psi|x\rangle\langle x|\psi\rangle = \int dx |\psi(x)|^2 = 1 \quad (67)$$

Now consider the momentum operator  $P$ . We know from classical mechanics that momentum is the generator of spatial translation. So let's consider a shift in a generic wave function by an amount of  $a$

$$\psi(x) \rightarrow \psi(x+a) \quad (68)$$

This translation should be the result of a physical transformation in the system and as such must be implemented by an unitary operator. The translation operator  $U_a$  thus acts as

$$\langle x|U_a|\psi\rangle = \langle x+a|\psi\rangle = \psi(x+a) \quad (69)$$

Considering an infinitesimal translation, the unitary operator can be written as  $U_a = \mathbb{1} + i\frac{Pa}{\hbar}$ , where  $P$  is the self-adjoint operator that generates the translation and thus we would like to identify it with the momentum operator. Expanding both sides of the above equation we find

$$\langle x|\psi\rangle + \frac{ia}{\hbar}\langle x|P|\psi\rangle = \psi(x) + a\frac{\partial\psi}{\partial x} \quad (70)$$

from where we can conclude that

$$\langle x|P|\psi\rangle = \frac{\hbar}{i}\frac{\partial\psi}{\partial x} \quad (71)$$

We can also obtain an useful equation inserting  $\mathbb{1} = \int dx |x\rangle\langle x|$

$$\langle x|P\mathbb{1}|\psi\rangle = \int dx' \langle x|P|x'\rangle\psi(x') = \frac{\hbar}{i}\frac{\partial\psi}{\partial x} \quad (72)$$

from where we can infer that

$$\langle x|P|x'\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\delta(x-x') \quad (73)$$

This is the “matrix” representation of the momentum  $P$  in the basis defined by  $X$ . The fact that it is not diagonal (not proportional to a delta function, but related instead to its derivative) means that, as we already know, that these observables are not compatible. In fact, acting on an arbitrary vector state we have

$$\langle x|PX|\psi\rangle = \int dx' \langle x|P|x'\rangle x'\psi(x') = \frac{\hbar}{i}\frac{\partial}{\partial x}(x\psi) = \frac{\hbar}{i}\psi + \frac{\hbar}{i}x\frac{\partial\psi}{\partial x} \quad (74)$$

and

$$\langle x|XP|\psi\rangle = x \int dx' \langle x|P|x'\rangle \psi(x') = \frac{\hbar}{i} x \frac{\partial \psi}{\partial x} \quad (75)$$

So that

$$\langle x|(XP - PX)|\psi\rangle = -\frac{\hbar}{i} \langle x|\psi\rangle \quad (76)$$

That is, since this is valid for arbitrary  $|\psi\rangle$  and  $|x\rangle$  is a basis,

$$[X, P] = i\hbar 1 \quad (77)$$

One can also study the eigenfunctions  $|p\rangle$  of  $P$ . In the same way as it happens with the position operator, these are also not physical states because they are not normalizable. The wave function representation of the momentum eigenfunctions is easily constructed from the previous properties

$$\langle x|P|p\rangle = p\langle x|p\rangle = \frac{\hbar}{i} \frac{\partial \langle x|p\rangle}{\partial x} \quad (78)$$

therefore

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px} \quad (79)$$

the factor  $\frac{1}{\sqrt{2\pi\hbar}}$  is just a convenient factor (not a normalization, of course). This function is obviously not a physical state, because it is not normalizable, but it is a very useful quantity. Pick a general physical state  $|\psi\rangle$  and note that

$$\langle p|\psi\rangle = \int dx \langle p|x\rangle \langle x|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-\frac{i}{\hbar}px} \psi(x) \quad (80)$$

This is the Fourier transform of  $\psi(x)$ , denoted as  $\langle p|\psi\rangle = \psi(p)$

## 1.4 Harmonic oscillator

A very important example that highlights the interplay between the different kinds of infinite dimensional Hilbert spaces is the harmonic oscillator system. This system, as any quantum mechanical system, is completely defined by specifying its Hamiltonian

$$H = \frac{P^2}{2m} + \frac{m\omega^2 X^2}{2} \quad (81)$$

and the commutation relation of all the operators appearing in the Hamiltonian, that is,  $q$  and  $p$ .

$$\begin{aligned} [X, X] &= [P, P] = 0 \\ [X, P] &= i\hbar \end{aligned} \quad (82)$$

The solution of the problem consists in obtain the spectrum of the Hamiltonian. There is a very elegant way of arriving at the solution considering the algebra of the operators. One starts by defining the operator  $a$  and its corresponding conjugate  $a^\dagger$

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left( X + \frac{iP}{m\omega} \right); \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left( X - \frac{iP}{m\omega} \right) \quad (83)$$

It follows that

$$[a, a^\dagger] = 1 \quad (84)$$

Define

$$N = a^\dagger a = \frac{m\omega}{2\hbar} \left( X^2 + \frac{P^2}{m^2\omega^2} + \frac{i}{m\omega}(XP - PX) \right) = \frac{1}{\hbar\omega} \left( \frac{m\omega^2 X^2}{2} + \frac{P^2}{2m} - \frac{\hbar\omega}{2} \right) = \frac{1}{\hbar\omega} \left( H - \frac{\hbar\omega}{2} \right) \quad (85)$$

and then

$$H = \hbar\omega a^\dagger a + \frac{\hbar\omega}{2} \quad (86)$$

Thus if we know the spectrum of  $N$ , we know the spectrum of  $H$ . Suppose that  $|n\rangle$  is an eigenstate of  $N$

$$N|n\rangle = n|n\rangle \quad (87)$$

$N$  is hermitian, thus we know that  $n$  is real. We show now that it is in fact a non-negative integer. Note first that

$$[N, a^\dagger] = a^\dagger \quad (88)$$

Therefore

$$[N, a^\dagger] |n\rangle = Na^\dagger|n\rangle - na^\dagger|n\rangle = a^\dagger|n\rangle \quad (89)$$

or

$$Na^\dagger|n\rangle = (n+1)a^\dagger|n\rangle \quad (90)$$

This means that  $a^\dagger|n\rangle$  is a eigenstate of  $N$  with eigenvalue  $n+1$ . Therefore

$$a^\dagger|n\rangle = c|n+1\rangle \quad (91)$$

where  $c$  is a normalization constant that can be determined as

$$\langle n|aa^\dagger|n\rangle = \langle n|a^\dagger a|n\rangle + \langle n|n\rangle = n+1 = |c|^2 \langle n+1|n+1\rangle = |c|^2 \quad (92)$$

and therefore, up to a phase that we set equal to one, we find  $c = \sqrt{n+1}$ . Similarly, note that (taking the adjoint of (88))

$$[N, a] = -a \quad (93)$$

It follows that

$$a|n\rangle = d|n-1\rangle \quad (94)$$

and we find  $d = \sqrt{n}$ . Thus

$$\begin{aligned} a|n\rangle &= \sqrt{n}|n-1\rangle \\ a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle \end{aligned} \quad (95)$$

This is the reason that  $a$  and  $a^\dagger$  are known as lowering and raising operators, respectively, or destruction and creation operators. Note now that  $n$  is in fact positive, because it is the norm of the vector  $a|n\rangle$

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle \geq 0 \quad (96)$$

Therefore, by successively applying  $a$ , there must be a lower limit, otherwise we reach negative values of  $n$ . In fact,

$$a^M |n\rangle = \sqrt{n}\sqrt{n-1}\cdots\sqrt{n-M}|n-M\rangle \quad (97)$$

where  $M$  is a non-negative integer. This can only stop if  $n - M = 0$ . Therefore we conclude that this construction only makes sense if  $n$  is a non-negative.

This solves the problem. The spectrum of the Hamiltonian is given by the eigentstates  $|n\rangle$  and the eigenvalues are

$$E_n = \hbar\omega n + \frac{\hbar\omega}{2} \quad (98)$$

for non-negative  $n$ .

the lowest energy state is known as the ground state. It corresponds to the lowest possible  $n$  values, that is,  $n = 0$ . The corresponding energy is  $E_0 = \frac{\hbar\omega}{2}$ . From the ground state, any other state can be obtained by applying the creation operator

$$\begin{aligned} |1\rangle &= a^\dagger |0\rangle \\ |2\rangle &= \frac{1}{\sqrt{2}} a^\dagger |1\rangle \\ |3\rangle &= \frac{1}{\sqrt{3}} a^\dagger |2\rangle \\ &\vdots \end{aligned} \quad (99)$$

There is a (countably) infinite number of states and this defines an infinite dimensional Hilbert space. The Hamiltonian is a function of the operators  $P$  and  $X$  but it does not commute with neither. It is straightforward to obtain a (infinite) matrix representation of these operators. Using the fact that

$$X = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger); \quad P = i\sqrt{\frac{\hbar m\omega}{2}}(a^\dagger - a) \quad (100)$$

we obtain

$$\begin{aligned} \langle n' | X | n \rangle &= \sqrt{\frac{\hbar}{2m\omega}} \left( \sqrt{n} \delta_{n', n-1} + \sqrt{n+1} \delta_{n', n+1} \right) \\ \langle n' | P | n \rangle &= i\sqrt{\frac{\hbar m\omega}{2}} \left( -\sqrt{n} \delta_{n', n-1} + \sqrt{n+1} \delta_{n', n+1} \right) \end{aligned} \quad (101)$$

This algebraic formalism makes it easy to obtain the energy eigensates with respect to the basis defined by the eigenvectors of  $X$  or  $P$ . One can start from the ground state

$$a|0\rangle = 0 \Rightarrow \langle x|a|0\rangle = 0 \Rightarrow \langle x| \left( X + \frac{iP}{m\omega} \right) |0\rangle = x\langle x|0\rangle + \frac{i}{m\omega} \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x|0\rangle = 0 \quad (102)$$

The differential equation is trivially solved giving (including the normalization)

$$\psi_0(x) \equiv \langle x|0\rangle = \left( \frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \quad (103)$$

This is known as the wave function of the ground state. The excited states are easily obtained by acting with  $a^\dagger$  in the position basis

$$\begin{aligned}\psi_1(x) &= \langle x|1\rangle = \langle x|a^\dagger|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\langle x|\left(X - \frac{iP}{m\omega}\right)|0\rangle = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{i}{m\omega}\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\psi_0(x) \\ &\vdots \\ \psi_n(x) &= \frac{1}{\sqrt{n!}}\left(\frac{m\omega}{2\hbar}\right)^{n/2}\left(x - \frac{i}{m\omega}\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^n\psi_0(x)\end{aligned}\tag{104}$$

These are related to the Hermite polynomials.

## Number-angle variables

Let us define

$$a^\dagger = Ae^{i\theta}\tag{105}$$

where  $A$  is positive Hermitian and  $\theta$  is Hermitian.

Note that

$$a^\dagger a = Ae^{i\theta}e^{-i\theta}A = A^2\tag{106}$$

such that

$$A^2 = N.\tag{107}$$

We have also

$$[a, a^\dagger] = aa^\dagger - a^\dagger a\tag{108}$$

$$= e^{-i\theta}A^2e^{i\theta} - A^2 = 1,\tag{109}$$

which implies

$$A^2e^{i\theta} - e^{i\theta}A^2 = e^{i\theta} \Rightarrow [N, e^{i\theta}] = e^{i\theta}.\tag{110}$$

Therefore,

$$[N, e^{i\theta}]|m\rangle = Ne^{i\theta}|m\rangle - me^{i\theta}|m\rangle\tag{111}$$

$$\Rightarrow Ne^{i\theta}|m\rangle = (m+1)e^{i\theta}|m\rangle.\tag{112}$$

And we obtain

$$e^{i\theta}|m\rangle = |m+1\rangle.\tag{113}$$

Note that  $e^{i\theta}$  is an isometry but it is not unitary! This is because the vector  $|0\rangle$  is orthogonal to every vector in the range of  $e^{i\theta}$ , that is,  $e^{i\theta}$  is not onto.

Note also that  $\theta$  is not a well-defined operator, for one is led to believe that<sup>4</sup>

$$[N, \theta] = -i.\tag{116}$$

---

<sup>4</sup>This is obtained from  $[N, e^{i\theta}] = e^{i\theta}$  by expanding both sides:

$$[N, 1 + i\theta + \frac{(i\theta)^2}{2!} + \dots] = 1 + i\theta + \frac{(i\theta)^2}{2!} + \dots\tag{114}$$

$$= i[N, \theta] + \dots\tag{115}$$

But if  $[N, \theta] = -i$  holds, we would have

$$\langle m|[N, \theta]|m\rangle = (m - m)\langle m|\theta|m\rangle = -i \neq 0, \quad (117)$$

so both sides of the equation are incompatible.

The reason  $\theta$  is not well-defined as an operator is because the operator nature does not capture the periodicity of  $\theta$ , which is supposed to be an angle.

## 1.5 Coherent and Squeezed states of the harmonic oscillator

The harmonic oscillators is very well known from classical mechanics. It has a intuitive dynamics. But what we discussed in the previous section seems very far from our usual experience. What are the states that are closest to what one may perceive as a classical state? One would expect that a state with minimum uncertainty in the phase space variables is the closest one can get to a classical state. It is easy to check that the ground state is a state of minimum uncertainty. Are there other states of minimum uncertainty?

### Coherent states

The ground state  $|0\rangle$  is an eigenstate of the annihilation operator  $a$ :

$$a|0\rangle = 0 \quad (118)$$

Let's search for general eigenstate  $|z\rangle$ :

$$a|z\rangle = z|z\rangle \quad (119)$$

note that  $z$  is complex in general, because  $a$  is not hermitian. Expanding in eigenstates of energy

$$|z\rangle = \sum_{n=0}^{\infty} c_n |n\rangle \quad (120)$$

we obtain

$$a|z\rangle = \sum_{n=0}^{\infty} c_n \sqrt{n} |n-1\rangle = \sum_{n=1}^{\infty} c_n \sqrt{n} |n-1\rangle = \sum_{n=0}^{\infty} c_{n+1} \sqrt{n+1} |n\rangle = z \sum_{n=0}^{\infty} c_n |n\rangle \quad (121)$$

This gives a recurrency relation

$$c_{n+1} = \frac{z c_n}{\sqrt{n+1}} = \frac{z}{\sqrt{n+1}} \frac{z c_{n-1}}{\sqrt{n}} = \dots = \frac{z^{n+1} c_0}{\sqrt{(n+1)!}} \quad (122)$$

and therefore

$$|z\rangle = c_0 \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle \quad (123)$$

Upon normalizing

$$\langle z|z\rangle = |c_0|^2 \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{z^n z^{*m}}{\sqrt{n!} \sqrt{m!}} \langle m|n\rangle = |c_0|^2 \sum_{n=0}^{\infty} \frac{|z|^{2n}}{n!} = |c_0|^2 e^{|z|^2} \quad (124)$$

This must be equal to 1, therefore

$$|c_0| = e^{-\frac{|z|^2}{2}} \quad (125)$$

One can also choose the arbitrary phase to be zero thus defining  $c_0 = e^{-\frac{|z|^2}{2}}$  and we finally obtain

$$|z\rangle = \sum_{n=0}^{\infty} e^{-\frac{|z|^2}{2}} \frac{z^n}{\sqrt{n!}} |n\rangle \quad (126)$$

Note also that

$$|n\rangle = \frac{a^\dagger}{\sqrt{n}} |n-1\rangle = \dots = \frac{a^{\dagger n}}{\sqrt{n!}} |0\rangle \quad (127)$$

Therefore

$$|z\rangle = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{(za^\dagger)^n}{n!} |0\rangle = e^{-\frac{|z|^2}{2}} e^{za^\dagger} |0\rangle \quad (128)$$

The state  $|z\rangle$  is called a coherent state. It can be obtained from the ground state  $|0\rangle$  by acting with the operator  $e^{-\frac{|z|^2}{2}} e^{za^\dagger}$ . Note that this operator is not unitary. In order to better interpret it as a transformation one can actually define a related unitary operator, that acts in the same way on  $|0\rangle$ :

$$D(z) = e^{-\frac{|z|^2}{2}} e^{za^\dagger} e^{-z^*a}. \quad (129)$$

In fact,

$$D(z)|0\rangle = e^{-\frac{|z|^2}{2}} e^{za^\dagger} e^{-z^*a}|0\rangle = e^{-\frac{|z|^2}{2}} e^{za^\dagger}|0\rangle = |z\rangle \quad (130)$$

One can easily check that  $D(z)$  is unitary by noting first that

$$e^{za^\dagger - z^*a} = e^{za^\dagger} e^{-z^*a} e^{-\frac{1}{2}[za^\dagger, -z^*a]} = e^{za^\dagger} e^{-z^*a} e^{-\frac{|z|^2}{2}} = D(z) \quad (131)$$

where the identity  $e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}$ , valid when  $[A, B] = \text{constant}$ , was used. Then one note that

$$(za^\dagger - z^*a)^\dagger = z^*a - za^\dagger = -(za^\dagger - z^*a) \quad (132)$$

so that  $za^\dagger - z^*a$  is anti-hermitian and therefore

$$D^\dagger(z) = e^{(za^\dagger - z^*a)^\dagger} = e^{-(za^\dagger - z^*a)} = D^{-1}(z) \quad (133)$$

It is also useful to note that  $D^\dagger(z) = D(-z)$ .

Coherent states are not eigenstates of energy, but they have an interesting physical interpretation. Consider

$$\langle z|a^\dagger a|z\rangle = \langle z|N|z\rangle = |z|^2 = \bar{N} \quad (134)$$

So,  $|z|^2$  is the mean particle number  $\bar{N}$  in the state  $|z\rangle$ .

What is the probability of finding the system in the state  $|n\rangle$  if it is in  $|\alpha\rangle$ ?

$$P_m = |\langle n|\alpha\rangle|^2 = |c_n|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} = \frac{e^{-\bar{N}} \bar{N}^n}{n!} \quad (135)$$

This is the Poisson distribution.

The coherent states form a set of states that is not orthogonal to each other. If  $a|\alpha\rangle = \alpha|\alpha\rangle$  and  $a|\beta\rangle = \beta|\beta\rangle$ , then

$$\begin{aligned}\langle\beta|\alpha\rangle &= \sum_{m,n} e^{-|\alpha|^2/2} e^{-|\beta|^2/2} \frac{\alpha^m}{\sqrt{m!}} \frac{(\beta^*)^n}{\sqrt{n!}} \langle m|n\rangle \\ &= \sum_m e^{-\frac{1}{2}(|\alpha|^2+|\beta|^2)} \frac{(\alpha\beta^*)^m}{m!} \\ &= e^{-\frac{1}{2}(|\alpha|^2+|\beta|^2-2\alpha\beta^*)}.\end{aligned}\quad (136)$$

But they form an (over)complete set:

$$\begin{aligned}\int dz dz^* |z\rangle\langle z| &= \int dz dz^* \sum_{m,n} e^{-|z|^2} \frac{z^m (z^*)^n}{\sqrt{m!} \sqrt{n!}} |m\rangle\langle n| \\ &= \sum_{m,n} \frac{1}{\sqrt{m!} \sqrt{n!}} |m\rangle\langle n| \int dz dz^* e^{-|z|^2} z^m (z^*)^n\end{aligned}\quad (137)$$

Define  $z = re^{i\theta}$ ,  $dz dz^* = J dr d\theta$ , where:

$$\begin{pmatrix} \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \\ \frac{\partial z^*}{\partial r} & \frac{\partial z^*}{\partial \theta} \end{pmatrix} = \begin{pmatrix} e^{i\theta} & ire^{i\theta} \\ e^{-i\theta} & -ire^{-i\theta} \end{pmatrix} \Rightarrow J = 2r \quad (138)$$

Therefore, the last integral in (137) becomes

$$\int dr d\theta 2r e^{-r^2} r^{m+n} e^{i(m-n)\theta} = \int dr 2r e^{-r^2} r^{m+n} \int_0^{2\pi} d\theta e^{i(m-n)\theta} = 2\pi \delta_{m,n} \int_0^\infty dr 2r e^{-r^2} r^{2m} \quad (139)$$

Thus the last term in (137) can be written as:

$$\sum_m \frac{1}{m!} |m\rangle\langle m| \int_0^\infty dr 2r e^{-r^2} r^{2m} 2\pi \quad (140)$$

Note that:

$$\int_0^\infty dr r e^{-\alpha r^2} = \left[ -\frac{e^{-\alpha r^2}}{2\alpha} \right]_0^\infty = \frac{1}{2\alpha} \quad (141)$$

Taking derivatives:

$$\begin{aligned}\frac{d^m}{d\alpha^m} \int_0^\infty dr r e^{-\alpha r^2} &= (-1)^m \int_0^\infty dr r e^{-\alpha r^2} r^{2m} \\ &= \frac{d^m}{d\alpha^m} \left( \frac{1}{2\alpha} \right) = \frac{1}{2} (-1)^m \frac{m!}{\alpha^{m+1}}\end{aligned}\quad (142)$$

Therefore  $\int_0^\infty dr r e^{-r^2} r^{2m} = \frac{m!}{2}$  and we obtain

$$\sum_m \frac{1}{m!} |m\rangle\langle m| 4\pi \frac{m!}{2} = 2\pi \sum_m |m\rangle\langle m| = 2\pi \mathbb{1} \quad (143)$$

From which we conclude

$$\int \frac{dz dz^*}{2\pi} |z\rangle\langle z| = \mathbb{1} \quad (144)$$

Or, one can also write  $z = \text{Re}z + i\text{Im}z$  such that

$$\begin{pmatrix} \frac{\partial z}{\partial \text{Re}z} & \frac{\partial z}{\partial \text{Im}z} \\ \frac{\partial z^*}{\partial \text{Re}z} & \frac{\partial z^*}{\partial \text{Im}z} \end{pmatrix} = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \Rightarrow J = 2 \quad (145)$$

and thus, with  $dz dz^* = 2d\text{Re}z d\text{Im}z = 2d^2z$

$$\int \frac{d^2z}{\pi} |z\rangle\langle z| = \mathbf{1} \quad (146)$$

Finally, let's consider the uncertainty in phase space for the coherent states. We have

$$\begin{aligned} \langle z|X|z\rangle &= \sqrt{\frac{\hbar}{2m\omega}} \langle z|(a^\dagger + a)|z\rangle = \sqrt{\frac{\hbar}{2m\omega}} (z^* + z) \\ \langle z|P|z\rangle &= i\sqrt{\frac{m\hbar\omega}{2}} \langle z|(a^\dagger - a)|z\rangle = i\sqrt{\frac{m\hbar\omega}{2}} (z^* - z) \\ \langle z|X^2|z\rangle &= \frac{\hbar}{2m\omega} \langle z|(a^\dagger + a)^2|z\rangle = \frac{\hbar}{2m\omega} \langle z|(a^{\dagger 2} + a^2 + 2a^\dagger a + 1)|z\rangle = \frac{\hbar}{2m\omega} [(z^* + z)^2 + 1] \\ \langle z|P^2|z\rangle &= -\frac{m\hbar\omega}{2} \langle z|(a^\dagger - a)^2|z\rangle = -\frac{m\hbar\omega}{2} \langle z|(a^{\dagger 2} + a^2 - 2a^\dagger a - 1)|z\rangle = -\frac{m\hbar\omega}{2} [(z^* - z)^2 - 1] \end{aligned} \quad (147)$$

Therefore

$$\begin{aligned} \langle z|(\Delta X)^2|z\rangle &= \langle z|(X^2 - \langle X\rangle^2)|z\rangle = \frac{\hbar}{2m\omega} [(z^* + z)^2 + 1 - (z^* + z)^2] = \frac{\hbar}{2m\omega} \\ \langle z|(\Delta P)^2|z\rangle &= \langle z|(P^2 - \langle P\rangle^2)|z\rangle = -\frac{m\hbar\omega}{2} [(z^* - z)^2 - 1 - (z^* - z)^2] = \frac{m\hbar\omega}{2} \end{aligned} \quad (148)$$

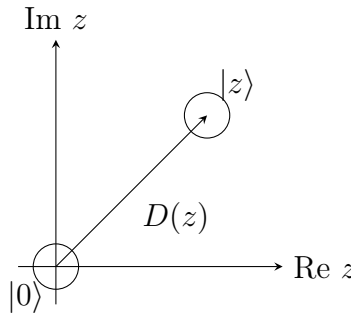
From which follows that

$$\langle z|(\Delta X)^2|z\rangle \langle z|(\Delta P)^2|z\rangle = \frac{\hbar^2}{2} \quad (149)$$

Therefore we have obtained a whole class of minimum uncertainty states: the coherent states.

### Squeezed states

Coherent States are discs in the phase space defined by  $\langle z|X|z\rangle \sim \text{Re}z$  and  $\langle z|P|z\rangle \sim \text{Im}z$ .



One can also consider deformations of these discs, for instance into ellipses, preserving the area and thus the minimum uncertainty. Consider the operator

$$S(r) = e^{\frac{r}{2}(a^{\dagger 2} - a^2)} \quad (150)$$

with  $r \in \mathbb{R}$ ,  $S(r)$  is a unitary operator because  $(a^{\dagger 2} - a^2)$  is anti-hermitian:  $S^{\dagger}(r) = S^{-1}(r) = S(-r)$ .

Using  $e^A B e^{-A} = B + [A, B] + \frac{1}{2!} [A, [A, B]] + \dots$ , it follows that

$$S^{\dagger}(r) a S(r) = a + r a^{\dagger} + \frac{r^2}{2!} a + \frac{r^3}{3!} a^{\dagger} = a \cosh r + a^{\dagger} \sinh r \equiv b(r) \quad (151)$$

and

$$S^{\dagger}(r) a^{\dagger} S(r) = a^{\dagger} \cosh r + a \sinh r = b^{\dagger}(r) \quad (152)$$

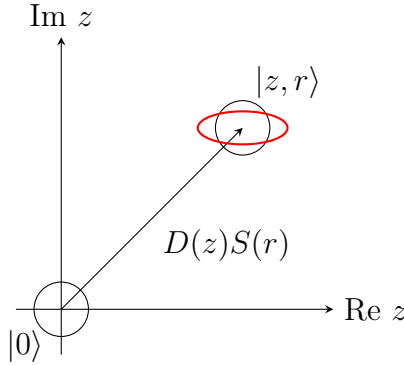
note that

$$[b, b^{\dagger}] = \cosh r^2 [a, a^{\dagger}] - \sinh r^2 [a, a^{\dagger}] = \cosh r^2 - \sinh r^2 = 1 \quad (153)$$

Now consider

$$\begin{aligned} S^{\dagger}(r) X S(r) &= \sqrt{\frac{\hbar}{2m\omega}} S^{\dagger}(r) (a^{\dagger} + a) S(r) = \sqrt{\frac{\hbar}{2m\omega}} (a^{\dagger} + a) e^r = X e^r \\ S^{\dagger}(r) P S(r) &= i \sqrt{\frac{m\hbar\omega}{2}} S^{\dagger}(r) (a^{\dagger} - a) S(r) = i \sqrt{\frac{m\hbar\omega}{2}} (a^{\dagger} - a) e^{-r} = e^{-r} P \end{aligned} \quad (154)$$

This implies that  $(\Delta X)^2 \rightarrow e^{2r} (\Delta X)^2$  and  $(\Delta P)^2 \rightarrow e^{-2r} (\Delta P)^2$ , so that the minimum uncertainty is maintained for the state  $S(r)|0\rangle$ . Furthermore, acting with  $D(z)$  we can construct the state  $|z, r\rangle = D(z)S(r)|0\rangle$ , which have the graphical representation bellow



## 1.6 Dynamics

The time development of a closed quantum system is described by Schrodinger's equation. We know that this is equivalent as the action of a unitary operator on the state

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \quad (155)$$

where  $U(t) = e^{-\frac{i}{\hbar} H t}$ . Because the unitary operator is a function of the Hamiltonian in order to study time development of a state it is useful to express it in a base of energy eigenstates. So, if one knows that

$$|\psi(0)\rangle = \sum_{n=1} c_n |E_n\rangle \quad (156)$$

then

$$|\psi(t)\rangle = \sum_{n=1} c_n e^{-\frac{i}{\hbar} E_n t} |E_n\rangle \quad (157)$$

## Neutrino Oscillations

An interesting, important, example of time evolution and the kind of effect the non-commuting nature of observables can have is the neutrino oscillation. The general idea can be understood as follows. Consider a system described by a Hamiltonian  $H$ . Suppose we are interested in measuring an observable  $A$  that does not commute with  $H$ . Let's consider a two levels system in such a way that both  $H$  and  $A$  have two eigenvectors. Since these operators don't commute, their eigenstates are different and one can be written as a linear combination of the other. For instance, if  $|E_1\rangle, |E_2\rangle$  are the eigenstates of  $H$  and  $|A_1\rangle, |A_2\rangle$  are the eigenstates of  $A$  one can write

$$\begin{aligned} |A_1\rangle &= \cos\theta|E_1\rangle - \sin\theta|E_2\rangle \\ |A_2\rangle &= \sin\theta|E_1\rangle + \cos\theta|E_2\rangle \end{aligned} \quad (158)$$

This parameterization of the coefficients by  $\sin$  and  $\cos$  can always be done (it is simply a rotation from the  $H$  basis to the  $A$  basis). Thus under time evolution, if the state is initially  $|A_1\rangle$  it will oscillate to  $|A_2\rangle$  over time.

$$|\psi(t)\rangle = e^{-i\frac{Ht}{\hbar}}|A_1\rangle = \cos\theta e^{-i\frac{E_1t}{\hbar}}|E_1\rangle - \sin\theta e^{-i\frac{E_2t}{\hbar}}|E_2\rangle \quad (159)$$

In fact, the probability of obtaining  $A_1$  after a time  $t$  is

$$\begin{aligned} P_{A_1 \rightarrow A_1}(t) &= |\langle A_1|\psi(t)\rangle|^2 = |\langle A_1|e^{-i\frac{Ht}{\hbar}}|A_1\rangle|^2 \\ &= |\cos^2\theta e^{-i\frac{E_1t}{\hbar}} + \sin^2\theta e^{-i\frac{E_2t}{\hbar}}|^2 \\ &= \cos^4\theta + \sin^4\theta + 2\cos^2\theta\sin^2\theta\cos\left(\frac{(E_2 - E_1)t}{\hbar}\right) \\ &= \cos^4\theta + \sin^4\theta + 2\cos^2\theta\sin^2\theta\left(1 - 2\sin^2\left(\frac{(E_2 - E_1)t}{2\hbar}\right)\right) \\ &= 1 - \sin^2 2\theta \sin^2\left(\frac{(E_2 - E_1)t}{2\hbar}\right) \end{aligned} \quad (160)$$

This is what happens in the phenomenon of Neutrinos oscillations. The eigenstates of energy, which are also eigenstates of mass, are not eigenstates of flavor. When detecting a neutrino we are measuring flavor through the interacting Hamiltonian with the matter in the measuring apparatus. But Neutrinos when traveling until reaching the apparatuses are moving as eigenstates of mass. Neutrinos are very light and travel at approximate light speed, so their energy obey the formula  $E = \sqrt{m^2c^4 + p^2c^2} \approx pc(1 + \frac{m^2c^4}{2p^2c^2})$ , so that energy differences are mass square differences. Therefore, by measuring the frequency of oscillations of the probability of detecting a certain flavor of neutrino, one can obtain the mass square differences of neutrinos. This is the main evidence that neutrinos are massive, instead of massless as once believed.

## Two pictures of time development

The physical output of the time development shows itself when computing time evolution of expectation values (or probability amplitudes). Consider an operator  $A$ . Its expectation value will depend on the time as

$$\langle\psi(t)|A|\psi(t)\rangle = \langle\psi(0)|U^\dagger(t)AU(t)|\psi(0)\rangle \quad (161)$$

It is clear from this equation that instead of consider the time evolution of the state,  $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ , it is equally possible to consider the time evolution of the operator (with the state frozen at  $|\psi(0)\rangle$ )

$$A_H(t) = U^\dagger(t)AU(t) \quad (162)$$

This view, or picture, is called the Heisenberg picture, as opposed to the Schrodinger picture, where the state evolves in time while the operators stay the same.

The Heisenberg picture is completely equivalent to the Schrodinger picture but it is sometimes more useful, specially when considering generalizations to quantum field theory where the notion of state is less intuitive. One just have to be careful that not all states remain time independent. Note that, from the spectral representation

$$A_S = \sum_n a_n |a_n\rangle \langle a_n| \quad (163)$$

it follows that

$$A_H(t) = \sum_n a_n U^\dagger(t) |a_n\rangle \langle a_n| U(t) \quad (164)$$

So it follows that basis eigenstates used in computing expectation values or transition amplitudes must evolve “backwards” in time in the Heisenberg picture

$$|a_n(t)\rangle_H = U^\dagger(t) |a_n\rangle \quad (165)$$

This is also consistent with the time evolution of the Schrodinger picture, because

$$\langle a_n | U(t) | \psi \rangle = \langle a_n | \psi(t) \rangle = \langle a_n(t) | \psi \rangle \quad (166)$$

the general rule to go from the Schrodinger to Heisenberg picture is to look at physical quantities and demand that they remain the same in both picture.

In the same way that the Schrodinger equation provide a differential form for the time evolution of a state, the Heisenberg equation provide a a differential form for the time evolution of an operator. Differentiating (162) we find

$$\frac{dA_H}{dt} = \frac{dU^\dagger}{dt} A_S U + U^\dagger A_S \frac{dU}{dt} \quad (167)$$

but,  $U(t) = e^{-\frac{i}{\hbar} H t}$  and therefore  $\frac{dU}{dt} = -\frac{i}{\hbar} H U$ , and therefore

$$\frac{dA_H}{dt} = -\frac{i}{\hbar} (-H U^\dagger A_S U + U^\dagger A_S U H) \quad (168)$$

that is

$$i\hbar \frac{dA_H}{dt} = A_H H - H A_H = [A_H, H] \quad (169)$$

The Heisenberg picture is useful to provide a proper definition of the time-energy uncertainty relation. We already learn about uncertainty relation between operators, but in time is not an operator in quantum mechanics. Consider an operator  $A$  that does not commute with the Hamiltonian. Physically, one can use this operator as a clock measuring time changes. The uncertainty relation is

$$\langle \psi | (\Delta_\psi A_H)^2 | \psi \rangle \langle \psi | (\Delta_\psi H)^2 | \psi \rangle \geq \frac{1}{4} |\langle \psi | [A_H, H] | \psi \rangle|^2 = \frac{\hbar^2}{4} |\langle \psi | \frac{dA_H}{dt} | \psi \rangle|^2 \quad (170)$$

this is then the uncertainty between time and energy. One can be more explicit by defining the time uncertainty as

$$(\Delta_\psi t)^2 \equiv \frac{\langle \psi | (\Delta_\psi A_H)^2 | \psi \rangle}{|\langle \psi | \frac{dA_H}{dt} | \psi \rangle|^2} \quad (171)$$

Therefore

$$(\Delta_\psi t)^2 \langle \psi | (\Delta_\psi H)^2 | \psi \rangle \geq \frac{\hbar^2}{4} \quad (172)$$

## Time development of the Harmonic Oscillator

Let's study now the example of the harmonic oscillator again. Heisenberg's equations give (we omit the  $H$  subscript)

$$i\hbar \frac{dX}{dt} = [X, H] = \frac{1}{m}P; \quad i\hbar \frac{dP}{dt} = [P, H] = -m\omega^2 X; \quad (173)$$

One can write these using  $a$  and  $a^\dagger$  instead, obtaining

$$i\hbar \frac{da}{dt} = [a, H] = \hbar\omega a; \quad i\hbar \frac{da^\dagger}{dt} = [a^\dagger, H] = -\hbar\omega a^\dagger; \quad (174)$$

whose solution can be written immediately

$$a(t) = a(0)e^{-i\omega t}; \quad a^\dagger(t) = a^\dagger(0)e^{i\omega t} \quad (175)$$

From this we can obtain the corresponding time dependence of  $X$  and  $P$

$$\begin{aligned} X(t) &= \sqrt{\frac{\hbar}{2m\omega}}(a(t) + a(t)^\dagger) = \sqrt{\frac{\hbar}{2m\omega}}(a(0)e^{-i\omega t} + a(0)^\dagger e^{i\omega t}) \\ P(t) &= i\sqrt{\frac{\hbar m\omega}{2}}(a^\dagger(t) - a(t)) = i\sqrt{\frac{\hbar m\omega}{2}}(a^\dagger(0)e^{i\omega t} - a(0)e^{-i\omega t}) \end{aligned} \quad (176)$$

or, using that  $a(0) = \sqrt{\frac{m\omega}{2\hbar}} \left( X(0) + \frac{iP(0)}{m\omega} \right)$  and  $a^\dagger(0) = \sqrt{\frac{m\omega}{2\hbar}} \left( X(0) - \frac{iP(0)}{m\omega} \right)$ , it follows

$$\begin{aligned} X(t) &= X(0) \cos \omega t + \frac{P(0)}{m\omega} \sin \omega t \\ P(t) &= P(0) \cos \omega t - m\omega X(0) \sin \omega t \end{aligned} \quad (177)$$

This can also be obtained by the Heisenberg definition

$$X(t) = e^{i\frac{Ht}{\hbar}} X(0) e^{-i\frac{Ht}{\hbar}}; \quad P(t) = e^{i\frac{Ht}{\hbar}} P(0) e^{-i\frac{Ht}{\hbar}} \quad (178)$$

and the formula BCH (Baker-Campbell-Hausdorff)

$$e^{i\lambda B} A e^{-i\lambda B} = A + i\lambda [B, A] + \frac{(i\lambda)^2}{2!} [B, [B, A]] + \cdots + \frac{(i\lambda)^n}{n!} [B, \cdots [B, [B, A]] \cdots] \quad (179)$$

## 1.7 Motion of the coherent states

Consider the Hamiltonian

$$H(t) = \hbar [\omega(t)a^\dagger a + f(t)a^\dagger + f^*(t)a + g(t)].$$

Time evolution generated by this Hamiltonian does not destroy the coherent state. The classical analogy is a forced harmonic oscillator.

Suppose that

$$|\psi(t)\rangle = e^{i\theta(t)}|z(t)\rangle.$$

We want to check that this state is the solution at time  $t$  of a state that started at  $|\psi(0)\rangle = |z\rangle$ . We are defining  $|z(t)\rangle$  as:

$$|z(t)\rangle = D(z(t))|0\rangle.$$

The solution is known once  $z(t)$  and  $\theta(t)$  are known.

It is convenient to define the operator  $\hat{a}^D(t)$ :

$$\hat{a}^D(t) \equiv D^\dagger(z(t))\hat{a}D(z(t))$$

Note that

$$D(z_1)D(z_2) = e^{z_1 a^\dagger - z_1^* a} e^{z_2 a^\dagger - z_2^* a} = e^{(z_1+z_2)a^\dagger - (z_1+z_2)^* a} e^{\frac{1}{2}[z_1 a^\dagger - z_1^* a, z_2 a^\dagger - z_2^* a]} = D(z_1 + z_2) e^{Im(z_1 z_2^*)} \quad (180)$$

and thus, for any coherent state  $|\alpha\rangle$ , we have

$$\begin{aligned} \hat{a}^D(t)|\alpha\rangle &= D^\dagger(z)\hat{a}D(z)|\alpha\rangle = e^{Im(z\alpha^*)}D^\dagger(z)\hat{a}|\alpha + z\rangle = e^{Im(z\alpha^*)}(\alpha + z)D^\dagger(z)|\alpha + z\rangle \\ &= (\alpha + z)D^\dagger(z)D(z)|\alpha\rangle = (\hat{a} + z)|\alpha\rangle \end{aligned} \quad (181)$$

But, noticing that the coherent states form a basis, this is true for any state and therefore

$$\hat{a}^D(t) = \hat{a} + z(t).$$

Then,

$$\begin{aligned} \hat{H}^D(t) &= D^\dagger(z(t))H(t)D(z(t)) \\ &= \hbar [\omega(t)\hat{a}^{D\dagger}(t)\hat{a}^D(t) + f(t)\hat{a}^{D\dagger}(t) + f^*(t)\hat{a}^D(t) + g(t)]. \end{aligned} \quad (182)$$

Heisenberg's equation gives:

$$i\hbar \frac{d\hat{a}^D(t)}{dt} = [\hat{a}^D(t), \hat{H}^D(t)] = \hbar\omega(t)\hat{a}^D(t) + \hbar f(t).$$

Taking the expectation value in the vacuum state  $|0\rangle$ :

$$i\hbar \frac{d}{dt} \langle 0|\hat{a}^D(t)|0\rangle = \hbar\omega(t)\langle 0|\hat{a}^D(t)|0\rangle + \hbar f(t)\langle 0|0\rangle.$$

Since  $\langle 0|\hat{a}^D(t)|0\rangle = z(t)$ , we obtain

$$i\dot{z}(t) = \omega(t)z(t) + f(t).$$

Similarly, for  $a^\dagger$ , we get

$$-i\dot{z}^*(t) = \omega(t)z^*(t) + f^*(t).$$

These equations determine  $z(t)$ .

Now consider Schrödinger's equation:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle.$$

Substituting  $|\psi(t)\rangle = e^{i\theta(t)}|z(t)\rangle$ :

$$i\hbar \left( i \frac{d\theta}{dt} \right) e^{i\theta} |z(t)\rangle + i\hbar e^{i\theta} \frac{d}{dt} |z(t)\rangle = \hat{H} e^{i\theta} |z(t)\rangle.$$

Multiplying by  $e^{-i\theta}\langle z(t)|$  from the left:

$$-\hbar \frac{d\theta}{dt} \langle z(t)|z(t)\rangle + i\hbar \langle z(t)| \frac{d}{dt} |z(t)\rangle = \langle z(t)| \hat{H} |z(t)\rangle.$$

Assuming  $\langle z(t)|z(t)\rangle = 1$ , we get

$$\hbar \frac{d\theta}{dt} = \langle z(t) | \left( i\hbar \frac{d}{dt} - \hat{H}(t) \right) | z(t) \rangle.$$

This equation determines  $\theta(t)$ .

We have:

$$\langle z | H | z \rangle = \hbar [\omega |z|^2 + f(t)z^* + f^*(t)z + g(t)].$$

Also,

$$\frac{d}{dt}|z(t)\rangle = \frac{d}{dt}(D(z)|0\rangle) = \frac{d}{dt}\left(e^{-|z|^2/2}e^{za^\dagger}\right)|0\rangle = \left[-\frac{1}{2}(\dot{z}^*z + z^*\dot{z})|z\rangle + \dot{z}a^\dagger|z\rangle\right].$$

Therefore,

$$\langle z | \frac{d}{dt} | z \rangle = -\frac{1}{2}(\dot{z}^*z + z^*\dot{z}) + \dot{z}z^* = \frac{1}{2}(\dot{z}z^* - \dot{z}^*z).$$

Thus,

$$\frac{d\theta}{dt} = \frac{i}{2}(\dot{z}z^* - \dot{z}^*z) - \omega |z|^2 - f(t)z^* - f^*(t)z - g(t).$$

Note that:

$$i\dot{z} = \omega z + f(t), \quad -i\dot{z}^* = \omega z^* + f^*(t).$$

Then,

$$\omega |z|^2 = \frac{i}{2}(\dot{z}z^* - \dot{z}^*z) - \frac{1}{2}[f(t)z^* + f^*(t)z].$$

Substituting back:

$$\frac{d\theta}{dt} = -\frac{1}{2}[f(t)z^* + f^*(t)z] - g(t) = -\text{Re}\{f^*(t)z(t)\} - g(t).$$

## 2 Propagators and Feynman path integrals

In the previous sections we studied the time evolution of quantum states through the Schrödinger equation and the unitary operator  $U(t, t_0)$ . We now turn to a different but equivalent formulation of quantum dynamics: the *propagator* and the *path integral* representation due to Feynman. This approach provides a direct bridge between quantum mechanics and classical mechanics through the principle of least action, and opens the door to powerful techniques in quantum field theory and quantum information. This section is based on [6].

### 2.1 The propagator

Consider the Schrödinger equation for a quantum system:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (183)$$

For a particle in one spatial dimension, the Hamiltonian takes the form

$$H = \frac{\hat{P}^2}{2m} + V(\hat{X}), \quad (184)$$

where  $\hat{P}$  and  $\hat{X}$  are the momentum and position operators, respectively, and  $V$  is the potential energy.

When  $H$  is time-independent, the time evolution operator describing finite time evolution is

$$U(t_1, t_0) = e^{-\frac{i}{\hbar}(t_1 - t_0)H}. \quad (185)$$

Suppose the system is initially prepared in a position eigenstate  $|x_0\rangle$ . The quantity

$$\langle x_1 | U(t_1, t_0) | x_0 \rangle \quad (186)$$

is the probability amplitude of finding the particle at position  $x_1$  at time  $t_1$ , given that it was at position  $x_0$  at time  $t_0$ .

This quantity defines a wave function  $\psi(x_1, t_1) = \langle x_1 | U(t_1, t_0) | x_0 \rangle$ , but since it will play a central role in what follows, let us give it a special name: we call it the **propagator** (or *Green's function* of the Schrödinger equation) and write

$$\boxed{i G(x_1, t_1; x_0, t_0) \equiv \langle x_1 | U(t_1, t_0) | x_0 \rangle}. \quad (187)$$

The factor of  $i$  is a convention that simplifies several subsequent expressions.<sup>5</sup>

Since  $U(t_1, t_0)$  satisfies the Schrödinger equation, the propagator (viewed as a function of  $x \equiv x_1$  and  $t \equiv t_1$ ) also satisfies it:

$$i\hbar \partial_t G(x, t; x_0, t_0) = H G(x, t; x_0, t_0), \quad (188)$$

with the initial condition obtained by setting  $t_1 = t_0$ :

$$\begin{aligned} G(x, t_0; x_0, t_0) &= -i \langle x | U(t_0, t_0) | x_0 \rangle \\ &= -i \langle x | x_0 \rangle \\ &= -i \delta(x - x_0). \end{aligned} \quad (189)$$

The propagator is therefore the Green's function of the time-dependent Schrödinger equation. Once  $G$  is known, the time evolution of an arbitrary initial state  $\psi(x, t_0)$  can be obtained by convolution:

$$\psi(x, t) = i \int dx_0 G(x, t; x_0, t_0) \psi(x_0, t_0). \quad (190)$$

This formula expresses a fundamental idea: the propagator contains *all* dynamical information about the quantum system. Knowledge of  $G$  is equivalent to knowledge of the full time evolution operator  $U$ .

## 2.2 Free particle propagator

Let us consider the simplest case: a free particle in one dimension, for which  $V = 0$ . The Schrödinger equation for the propagator becomes

$$i\hbar \partial_t G = -\frac{\hbar^2}{2m} \frac{\partial^2 G}{\partial x^2}. \quad (191)$$

It is convenient to solve this equation in momentum space. We write the Fourier expansion

$$G(x, t; x_0, t_0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp g(p, t) e^{\frac{i}{\hbar}p(x-x_0)}, \quad (192)$$

---

<sup>5</sup>Some references, notably Sakurai [1], define the propagator without the prefactor  $i$ , writing  $K(x_1, t_1; x_0, t_0) = \langle x_1 | U(t_1, t_0) | x_0 \rangle$ . We follow the convention that includes the  $i$  as it leads to a cleaner connection with the retarded Green's function.

where  $g(p, t)$  is the Fourier amplitude. The initial condition (189) implies

$$g(p, t_0) = \frac{-i}{\sqrt{2\pi\hbar}}, \quad (193)$$

since  $\delta(x - x_0) = \frac{1}{2\pi\hbar} \int dp e^{\frac{i}{\hbar}p(x-x_0)}$ .

Substituting the Fourier expansion into the free Schrödinger equation, we obtain

$$i\hbar \partial_t g(p, t) = \frac{p^2}{2m} g(p, t). \quad (194)$$

This is an ordinary differential equation with the solution

$$g(p, t) = g(p, t_0) e^{-\frac{ip^2(t-t_0)}{2m\hbar}} = \frac{-i}{\sqrt{2\pi\hbar}} e^{-\frac{ip^2(t-t_0)}{2m\hbar}}. \quad (195)$$

Therefore, the free-particle propagator is

$$G(x, t; x_0, t_0) = \frac{-i}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{-\frac{ip^2(t-t_0)}{2m\hbar}} e^{\frac{i}{\hbar}p(x-x_0)}. \quad (196)$$

To evaluate this Gaussian integral, we complete the square in the exponent:

$$-\frac{i(t-t_0)}{2m\hbar} \left( p - \frac{m(x-x_0)}{t-t_0} \right)^2 + \frac{im(x-x_0)^2}{2\hbar(t-t_0)}. \quad (197)$$

Using the standard Gaussian integral (obtained by analytic continuation, or equivalently by contour deformation in the complex plane — Cauchy's theorem)

$$\int_{-\infty}^{\infty} dp e^{-i\alpha p^2} = \sqrt{\frac{\pi}{i\alpha}}, \quad \alpha > 0, \quad (198)$$

with  $\alpha = \frac{t-t_0}{2m\hbar}$ , we arrive at the result

$$\boxed{G(x, t; x_0, t_0) = -i \sqrt{\frac{m}{2\pi\hbar i(t-t_0)}} \exp\left(\frac{im(x-x_0)^2}{2\hbar(t-t_0)}\right)}. \quad (199)$$

Several remarks are in order:

- This expression gives the probability amplitude of finding a free particle at position  $x$  at time  $t$ , given that it was at  $x_0$  at time  $t_0$ . The modulus squared  $|G|^2 = \frac{m}{2\pi\hbar(t-t_0)}$  is spatially uniform, reflecting the fact that a free particle initially localized at a point spreads over all space instantaneously. More precisely,  $|G|^2 \Delta x = \frac{m \Delta x}{2\pi\hbar \Delta t}$ , which shows that the probability density scales inversely with the elapsed time.
- The propagator must be understood as a distribution: it makes sense as a probability amplitude only inside integrals with well-behaved (square-integrable) wave functions. One should not interpret  $|G|^2$  at a single point as a physical probability density.
- Note the appearance of the *classical action* for a free particle in the exponent:  $S_{\text{cl}} = \frac{m(x-x_0)^2}{2(t-t_0)}$ , which is the action evaluated along the classical trajectory connecting  $(x_0, t_0)$  to  $(x, t)$ . This is not a coincidence — it is the first hint of the deep connection between the propagator and the classical action that the path integral will make manifest.

## 2.3 Path integral representation

We now derive the path integral representation of the propagator. The key observation is that the propagator satisfies a **composition rule** (also known as the Chapman–Kolmogorov or semigroup property). Using the group property  $U(t, t_0) = U(t, t_1) U(t_1, t_0)$  for any intermediate time  $t_0 < t_1 < t$ , and inserting a complete set of position states at the intermediate time, we find:

$$\begin{aligned} i G(x, t; x_0, t_0) &= \langle x | U(t, t_0) | x_0 \rangle \\ &= \langle x | U(t, t_1) U(t_1, t_0) | x_0 \rangle \\ &= \int dx_1 \langle x | U(t, t_1) | x_1 \rangle \langle x_1 | U(t_1, t_0) | x_0 \rangle \\ &= \int dx_1 [i G(x, t; x_1, t_1)] [i G(x_1, t_1; x_0, t_0)]. \end{aligned} \quad (200)$$

The physical interpretation is clear: the total amplitude for going from  $(x_0, t_0)$  to  $(x, t)$  is obtained by *summing over all possible intermediate positions*  $x_1$  at time  $t_1$ . We are effectively summing over all possible “trajectories” that pass through the intermediate time slice.

This idea can be iterated. Divide the time interval  $[t_0, t]$  into  $N$  equal segments of duration

$$\Delta t = \frac{t - t_0}{N}, \quad (201)$$

and introduce intermediate times  $t_j = t_0 + j \Delta t$ , with  $j = 0, 1, \dots, N$ . Setting  $x_N \equiv x$  and repeatedly applying the composition rule, we obtain

$$i G(x, t; x_0, t_0) = \int dx_1 \cdots dx_{N-1} \prod_{j=1}^N [i G(x_j, t_j; x_{j-1}, t_{j-1})]. \quad (202)$$

Now consider the Hamiltonian  $H = \frac{\hat{p}^2}{2m} + V(\hat{X})$ . For small  $\Delta t$ , the Lie–Trotter–Suzuki decomposition gives

$$e^{-\frac{i}{\hbar} \left( \frac{\hat{p}^2}{2m} + V(\hat{X}) \right) \Delta t} \approx e^{-\frac{i}{\hbar} \frac{\hat{p}^2}{2m} \Delta t} e^{-\frac{i}{\hbar} V(\hat{X}) \Delta t} + \mathcal{O}(\Delta t^2). \quad (203)$$

The error is of order  $\Delta t^2$  because  $[\hat{p}^2/(2m), V(\hat{X})] \neq 0$  in general, but this error vanishes in the limit  $N \rightarrow \infty$ .<sup>6</sup>

The matrix element of each infinitesimal propagator between position eigenstates is therefore

$$\langle x_j | e^{-\frac{i}{\hbar} H \Delta t} | x_{j-1} \rangle \approx \langle x_j | e^{-\frac{i}{\hbar} \frac{\hat{p}^2}{2m} \Delta t} | x_{j-1} \rangle e^{-\frac{i}{\hbar} V(x_{j-1}) \Delta t} + \mathcal{O}(\Delta t^2). \quad (204)$$

To evaluate the kinetic part, we insert a complete set of momentum states:

$$\begin{aligned} \langle x_j | e^{-\frac{i}{\hbar} \frac{\hat{p}^2}{2m} \Delta t} | x_{j-1} \rangle &= \int dp \langle x_j | p \rangle \langle p | x_{j-1} \rangle e^{-\frac{i p^2 \Delta t}{2m \hbar}} \\ &= \frac{1}{2\pi \hbar} \int dp e^{-\frac{i p^2 \Delta t}{2m \hbar}} e^{\frac{i}{\hbar} p(x_j - x_{j-1})} \\ &= \frac{1}{2\pi \hbar} \sqrt{\frac{2\pi m \hbar}{i \Delta t}} \exp\left(\frac{im(x_j - x_{j-1})^2}{2\hbar \Delta t}\right) \\ &= \sqrt{\frac{m}{2\pi \hbar i \Delta t}} \exp\left(\frac{im(x_j - x_{j-1})^2}{2\hbar \Delta t}\right). \end{aligned} \quad (205)$$

<sup>6</sup>More precisely, the total accumulated error from  $N$  factors is  $N \times \mathcal{O}(\Delta t^2) = \mathcal{O}(\Delta t)$ , which vanishes as  $\Delta t \rightarrow 0$ .

It is also convenient to note that the potential can be evaluated at the midpoint. Taylor expanding around the midpoint  $\bar{x}_j = \frac{x_j + x_{j-1}}{2}$ :

$$V(x_{j-1}) = V\left(\frac{x_j + x_{j-1}}{2}\right) + \mathcal{O}(x_j - x_{j-1}), \quad (206)$$

so that, up to terms that vanish in the continuum limit, the infinitesimal propagator becomes

$$\langle x_j | e^{-\frac{i}{\hbar} H \Delta t} | x_{j-1} \rangle \approx \sqrt{\frac{m}{2\pi\hbar i \Delta t}} \exp\left\{ \frac{i \Delta t}{\hbar} \left[ \frac{m}{2} \left( \frac{x_j - x_{j-1}}{\Delta t} \right)^2 - V\left(\frac{x_j + x_{j-1}}{2}\right) \right] \right\}. \quad (207)$$

Substituting this into the time-sliced expression (202), we obtain

$$i G(x, t; x_0, t_0) = \left( \frac{m}{2\pi\hbar i \Delta t} \right)^{N/2} \int dx_1 \cdots dx_{N-1} \exp\left\{ \frac{i}{\hbar} \sum_{j=1}^N \Delta t \left[ \frac{m}{2} \left( \frac{x_j - x_{j-1}}{\Delta t} \right)^2 - V\left(\frac{x_j + x_{j-1}}{2}\right) \right] \right\}. \quad (208)$$

In the limit  $N \rightarrow \infty$  (equivalently  $\Delta t \rightarrow 0$ ), the differences  $\frac{x_j - x_{j-1}}{\Delta t} \rightarrow \dot{x}(t')$  become velocities, the sum becomes an integral over time, and we recognize the **classical Lagrangian**  $L(x, \dot{x}) = \frac{m}{2} \dot{x}^2 - V(x)$ . We thus arrive at the **Feynman path integral** representation of the propagator:

$$\boxed{i G(x, t; x_0, t_0) = \mathcal{A} \int \mathcal{D}[x(t')] \exp\left( \frac{i}{\hbar} S[x(t')] \right)}, \quad (209)$$

where

$$S[x(t')] = \int_{t_0}^t dt' \left( \frac{m \dot{x}^2}{2} - V(x) \right) \quad (210)$$

is the **classical action** functional, and the integration measure is defined as

$$\mathcal{A} \int \mathcal{D}[x(t')] \equiv \lim_{\substack{N \rightarrow \infty \\ \Delta t \rightarrow 0}} \left( \frac{m}{2\pi\hbar i \Delta t} \right)^{N/2} \int dx_1 \cdots dx_{N-1}. \quad (211)$$

The integration is performed over *all paths*  $x(t')$  satisfying the boundary conditions  $x(t_0) = x_0$  and  $x(t) = x$ . Each path contributes to the amplitude with a phase given by  $e^{iS/\hbar}$ , where  $S$  is the classical action evaluated along that path.

This result is remarkable for several reasons:

- It provides a formulation of quantum mechanics that makes no reference to operators or Hilbert spaces. Instead, the dynamical content is entirely captured by the classical action functional and a sum over histories.
- The classical limit  $\hbar \rightarrow 0$  emerges naturally: when  $\hbar$  is small, the phase  $S/\hbar$  oscillates rapidly and the dominant contribution to the integral comes from paths near the *stationary phase* point, i.e., paths that satisfy  $\delta S = 0$ . These are precisely the *classical trajectories*. This is the content of the *principle of stationary phase* and provides a direct derivation of the principle of least action from quantum mechanics.
- Quantum effects (interference, tunneling) arise from the contributions of paths far from the classical trajectory, whose phases do not cancel out.

## 2.4 Phase-space path integral

It is worth noting that before performing the Gaussian integration over momenta in equation (205), one can write the propagator in **phase-space** form. Returning to the time-sliced expression and keeping the momentum integrals explicit, we obtain

$$i G(x, t; x_0, t_0) = \int \mathcal{D}[x(t')] \int \mathcal{D}[p(t')] \exp \left\{ \frac{i}{\hbar} \int_{t_0}^t dt' [p \dot{x} - H(x, p)] \right\}, \quad (212)$$

where  $H(x, p) = \frac{p^2}{2m} + V(x)$  is the classical Hamiltonian. The integrand in the exponent,  $p \dot{x} - H$ , is precisely the *phase-space Lagrangian* familiar from Hamiltonian mechanics.

This form is more general than the configuration-space path integral (209) because it applies even when the Hamiltonian is not of the standard kinetic-plus-potential form. For Hamiltonians quadratic in momentum, the Gaussian integration over  $p$  can be performed analytically, recovering the Lagrangian path integral.

## 2.5 Coherent-state path integral for the harmonic oscillator

The path integral can also be formulated using the coherent states studied in Section 1.5. Consider the harmonic oscillator Hamiltonian

$$H = \hbar\omega a^\dagger a + \frac{\hbar\omega}{2}, \quad (213)$$

and recall that the coherent states  $|z\rangle$  satisfy  $a|z\rangle = z|z\rangle$  and form an overcomplete set:

$$\int \frac{d^2 z}{\pi} |z\rangle \langle z| = \mathbb{1}, \quad d^2 z = d(\operatorname{Re} z) d(\operatorname{Im} z). \quad (214)$$

We wish to compute the coherent-state propagator

$$i G(z, t; z_0, t_0) \equiv \langle z| U(t, t_0) |z_0\rangle. \quad (215)$$

We proceed by time-slicing, inserting the coherent-state resolution of the identity at each intermediate step. The key building block is the matrix element of the infinitesimal evolution operator between two coherent states. For  $H = \hbar\omega a^\dagger a$  (dropping the constant  $\hbar\omega/2$  for the moment), we use the normal-ordered form:

$$\langle z'| e^{-\frac{i}{\hbar} \hbar\omega a^\dagger a dt} |z\rangle = e^{-i\omega z^* z dt} \langle z'| z\rangle. \quad (216)$$

This follows because  $a^\dagger a$  acts on a coherent state by multiplication when sandwiched:  $\langle z'| a^\dagger a |z\rangle = z'^* z \langle z'| z\rangle$ .

To evaluate  $\langle z'| z\rangle$ , recall that

$$\langle z'| z\rangle = e^{-\frac{1}{2}|z|^2 - \frac{1}{2}|z'|^2 + z'^* z}. \quad (217)$$

This can be verified by inserting the expansion  $|z\rangle = e^{-|z|^2/2} \sum_n \frac{z^n}{\sqrt{n!}} |n\rangle$  and performing the sum:

$$\begin{aligned} \langle z'| z\rangle &= e^{-\frac{1}{2}(|z|^2 + |z'|^2)} \sum_{m,n} \frac{z'^{*m} z^n}{\sqrt{m! n!}} \langle m| n\rangle \\ &= e^{-\frac{1}{2}(|z|^2 + |z'|^2)} \sum_n \frac{(z'^* z)^n}{n!} \\ &= e^{-\frac{1}{2}(|z|^2 + |z'|^2 - 2z'^* z)}. \end{aligned} \quad (218)$$

For nearby coherent states,  $z' = z + \dot{z} dt$ , we have to first order in  $dt$ :

$$\langle z' | z \rangle \approx \exp\left(\frac{1}{2}(z\dot{z}^* - z^*\dot{z}) dt\right). \quad (219)$$

Combining with (216), the matrix element of the infinitesimal evolution becomes

$$\langle z' | e^{-\frac{i}{\hbar} H dt} | z \rangle \approx \exp\left\{\frac{i}{\hbar} \left[ \frac{i\hbar}{2}(z\dot{z}^* - z^*\dot{z}) - \hbar\omega z^* z \right] dt\right\}. \quad (220)$$

Iterating over  $N$  time slices and taking the limit  $N \rightarrow \infty$ , we arrive at the **coherent-state path integral**:

$$\boxed{i G(z, t; z_0, t_0) = \mathcal{A} \int \mathcal{D}^2[z(t')] \exp\left\{\frac{i}{\hbar} \int_{t_0}^t dt' \left[ \frac{i\hbar}{2}(z\dot{z}^* - z^*\dot{z}) - \hbar\omega z^* z - \frac{\hbar\omega}{2} \right]\right\}}, \quad (221)$$

where the “Lagrangian” in the exponent is

$$L = \frac{i\hbar}{2}(z\dot{z}^* - z^*\dot{z}) - \hbar\omega z^* z. \quad (222)$$

Note that the first term,  $\frac{i\hbar}{2}(z\dot{z}^* - z^*\dot{z})$ , plays the role of the “ $p\dot{q}$ ” symplectic term in the phase-space path integral, while  $\hbar\omega z^* z$  is the Hamiltonian expressed in terms of coherent-state labels.

## 2.6 Spin path integral and Berry phase

The coherent-state path integral can be generalized to spin systems, where it leads naturally to a geometric (Berry) phase. Consider a spin- $\frac{1}{2}$  particle and a general spin-component operator  $\hat{n} \cdot \vec{S}$ , where  $\hat{n}$  is a unit vector on the sphere parameterized by angles  $(\alpha, \beta)$ :

$$\hat{n} = (\sin \beta \cos \alpha, \sin \beta \sin \alpha, \cos \beta). \quad (223)$$

The eigenstate of  $\hat{n} \cdot \vec{S}$  with eigenvalue  $+\hbar/2$  is

$$|\hat{n}\rangle = \begin{pmatrix} \cos \frac{\beta}{2} \\ e^{i\alpha} \sin \frac{\beta}{2} \end{pmatrix}. \quad (224)$$

One can verify that  $\hat{n} = \langle \hat{n} | \vec{\sigma} | \hat{n} \rangle$ , where  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  are the Pauli matrices. Equivalently, writing  $\hat{n} = \hat{z}^\dagger \vec{\sigma} \hat{z}$  with the spinor  $\hat{z} = (\cos \frac{\beta}{2}, e^{i\alpha} \sin \frac{\beta}{2})^T$ , we see that the spin coherent states provide a natural parameterization of the sphere  $S^2$ .

Note, however, that there is a  $U(1)$  ambiguity: the transformation  $\hat{z} \rightarrow e^{i\theta} \hat{z}$  leaves  $\hat{n}$  invariant. This gauge freedom will be crucial in what follows.

The spin coherent states satisfy a resolution of the identity:

$$\int \frac{d\Omega}{2\pi} |\hat{n}\rangle \langle \hat{n}| = \mathbb{1}, \quad (225)$$

where  $d\Omega = \sin \beta d\beta d\alpha$  is the solid angle element on  $S^2$ . This can be verified by explicit integration.

**Derivation of the spin path integral.** Consider a time evolution with  $H = 0$ , so that  $U(t, t_0) = \mathbb{1}$ . The amplitude  $\langle \hat{n}(t) | \hat{n}(0) \rangle$  is simply the overlap between two spin coherent states. Inserting  $N$  resolutions of the identity at intermediate times, we write

$$\langle \hat{n}(t) | \hat{n}(0) \rangle = \prod_{i=1}^N \int \frac{d\Omega(t_i)}{2\pi} \langle \hat{n}(t) | \hat{n}(t_N) \rangle \langle \hat{n}(t_N) | \hat{n}(t_{N-1}) \rangle \cdots \langle \hat{n}(t_1) | \hat{n}(0) \rangle. \quad (226)$$

For an infinitesimal time step, the overlap between neighboring spin coherent states is

$$\begin{aligned} \langle \hat{n}(t + \delta t) | \hat{n}(t) \rangle &= 1 - \langle \hat{n}(t + \delta t) | (|\hat{n}(t + \delta t)\rangle - |\hat{n}(t)\rangle) \\ &\approx \exp\left(-\langle \hat{n} | \frac{d}{dt} | \hat{n} \rangle dt\right). \end{aligned} \quad (227)$$

In the continuum limit  $N \rightarrow \infty$ , this yields

$$\langle \hat{n}(t) | \hat{n}(0) \rangle = \mathcal{A} \int \mathcal{D}\Omega \exp\left(-\int_0^t dt' \langle \hat{n}(t') | \frac{d}{dt'} | \hat{n}(t') \rangle\right). \quad (228)$$

**Berry phase and Berry connection.** The phase factor  $e^{-\int_0^t dt' \langle \hat{n} | \frac{d}{dt'} | \hat{n} \rangle}$  is a pure quantum effect known as the **Berry phase** [7]. Let us compute the Berry connection explicitly. Using  $|\hat{n}\rangle = (\cos \frac{\beta}{2}, e^{i\alpha} \sin \frac{\beta}{2})^T$ :

$$\begin{aligned} \langle \hat{n} | \frac{d}{dt} | \hat{n} \rangle &= (\cos \frac{\beta}{2}, e^{-i\alpha} \sin \frac{\beta}{2}) \begin{pmatrix} -\frac{\dot{\beta}}{2} \sin \frac{\beta}{2} \\ i\dot{\alpha} e^{i\alpha} \sin \frac{\beta}{2} + e^{i\alpha} \frac{\dot{\beta}}{2} \cos \frac{\beta}{2} \end{pmatrix} \\ &= -\frac{\dot{\beta}}{2} \sin \frac{\beta}{2} \cos \frac{\beta}{2} + i\dot{\alpha} \sin^2 \frac{\beta}{2} + \frac{\dot{\beta}}{2} \sin \frac{\beta}{2} \cos \frac{\beta}{2} \\ &= i\dot{\alpha} \sin^2 \frac{\beta}{2}. \end{aligned} \quad (229)$$

Therefore, the Berry phase accumulated along a path in parameter space is

$$i S_B = -i \int_0^T dt \dot{\alpha} \sin^2 \frac{\beta}{2}. \quad (230)$$

**Gauge dependence and closed paths.** It is important to note that the Berry phase is *not* gauge-invariant for open paths. Under the reparameterization  $|\hat{n}(t)\rangle \rightarrow e^{i\theta(t)} |\hat{n}(t)\rangle$ , the Berry connection transforms as

$$\langle \hat{n} | \frac{d}{dt} | \hat{n} \rangle \rightarrow \langle \hat{n} | \frac{d}{dt} | \hat{n} \rangle + i\dot{\theta}, \quad (231)$$

so that

$$i S_B \rightarrow i S_B + i(\theta(T) - \theta(0)). \quad (232)$$

The Berry phase is therefore only well-defined (gauge-invariant, up to  $2\pi$ ) for **closed paths**,  $\hat{n}(T) = \hat{n}(0)$ , where the boundary terms cancel.

**Geometric interpretation.** For a closed path, we can rewrite the Berry phase using  $\sin^2 \frac{\beta}{2} = \frac{1}{2}(1 - \cos \beta)$ :

$$i S_B = -i \int dt \dot{\alpha} \sin^2 \frac{\beta}{2}$$

$$= -\frac{i}{2} \oint d\alpha + \frac{i}{2} \int dt \dot{\alpha} \cos \beta. \quad (233)$$

The first term,  $-\frac{i}{2} \oint d\alpha$ , contributes  $-i\pi n$  for a path winding  $n$  times and amounts to a sign  $(-1)^n$ . Applying Stokes' theorem to the second term:

$$\frac{i}{2} \oint d\alpha \cos \beta = \frac{i}{2} \int d\alpha d\beta \frac{\partial \cos \beta}{\partial \beta} = \frac{i}{2} \int d\alpha d\beta \sin \beta = \frac{i}{2} \Omega, \quad (234)$$

where  $\Omega$  is the **solid angle** subtended by the closed path on the unit sphere. We therefore arrive at the elegant result

$$\boxed{S_B = \frac{\Omega}{2}}. \quad (235)$$

The Berry phase for a spin- $\frac{1}{2}$  particle traversing a closed loop on the sphere is *half the solid angle* enclosed by the loop. This is a purely geometric quantity — it depends only on the shape of the path, not on how fast it is traversed.

This result admits a beautiful geometric interpretation: the solid angle  $\Omega$  is the area enclosed by the path on  $S^2$ . The Berry phase is therefore directly related to the curvature of the parameter space. More precisely, the Berry connection  $\mathcal{A} = i\langle \hat{n} | d | \hat{n} \rangle$  defines a  $U(1)$  connection (gauge field) on  $S^2$ , and the Berry phase is the holonomy of this connection. The curvature (Berry curvature) is proportional to the area form on  $S^2$ , which is precisely the statement  $S_B = \Omega/2$ .

This is analogous to the Aharonov–Bohm phase  $e^{i\frac{e}{\hbar} \oint \vec{A} \cdot d\vec{r}}$ , but here the “gauge field” lives in parameter space rather than physical space. In both cases, the phase is a topological/geometric quantity associated with the curvature of a connection.

For a general spin  $s$  (not just  $s = \frac{1}{2}$ ), the Berry phase generalizes to  $S_B = s\Omega$ , and the spin path integral involves the action

$$S = \int dt [-\hbar s \dot{\alpha} \cos \beta - H(\hat{n})], \quad (236)$$

where the first term is the Wess–Zumino–Witten topological term that encodes the quantum nature of spin.

### 3 Open quantum systems

Now let's add another level to our ignorance (besides the inherent quantum one). Suppose we are interested in a subsystem  $S_1$  of a larger system  $S$ . Suppose further that  $S$  is in a definite quantum state  $|\Phi_{S_1 S_2}\rangle$ . We know from (12) that the general form of this state is

$$|\Phi_{S_1 S_2}\rangle = \sum_i \sum_n C_{in} |i_{S_1} n_{S_2}\rangle \quad (237)$$

where  $S_2$  is the complement of system  $S_1$  in  $S$ . If  $C_{in} = a_i b_n$  then we can write

$$|\Phi_{S_1 S_2}\rangle = \sum_i a_i |i_{S_1}\rangle \sum_n b_n |n_{S_2}\rangle = |\Phi_{S_1}\rangle |\Phi_{S_2}\rangle \quad (238)$$

And the systems are completely decoupled. The quantum state of the system  $S_1$  is simply  $|\Phi_{S_1}\rangle$ . So the factorization  $C_{in} = a_i b_n$  corresponds then to the very special situation where the systems are independent from each other and even though  $S_1$  is part of a larger system it behaves as if it were closed.

Now the interesting case is when  $C_{in} \neq a_i b_n$ . One says then that systems  $S_1$  and  $S_2$  are entangled. What is the state of system  $S_1$  in this case? Suppose you want to measure an observable of system  $S_1$  and you

have no way of accessing system  $S_2$ . The observable has therefore the form  $A \otimes \mathbb{1}$  meaning that  $A$  acts on vectors of the Hilbert space  $\mathcal{H}_{S_1}$  and then identity  $\mathbb{1}$  acts on the vectors of the Hilbert space  $\mathcal{H}_{S_2}$ . Let's compute the expectation value of this operator on the state (237):

$$\begin{aligned}\langle \Phi_{S_1 S_2} | A \otimes \mathbb{1} | \Phi_{S_1 S_2} \rangle &= \sum_{ij} \sum_{nm} C_{in} C_{jm}^* \langle j_{S_1} | A | i_{S_1} \rangle \langle m_{S_2} | n_{S_2} \rangle \\ &= \sum_{ij} \sum_n C_{in} C_{jn}^* \langle j_{S_1} | A | i_{S_1} \rangle\end{aligned}\quad (239)$$

Define the operator  $\rho_{S_1}$

$$\rho_{S_1} = \sum_{kl} \sum_n C_{kn} C_{ln}^* |k_{S_1}\rangle \langle l_{S_1}| \quad (240)$$

Note that it acts on the space  $\mathcal{H}_{S_1}$  and

$$\begin{aligned}Tr_{S_1}(\rho_{S_1} A) &= \sum_i \langle i_{S_1} | \left( \sum_{kl} \sum_n C_{kn} C_{ln}^* |k_{S_1}\rangle \langle l_{S_1}| A \right) | i_{S_1} \rangle \\ &= \sum_{il} \sum_n C_{in} C_{ln}^* \langle l_{S_1} | A | i_{S_1} \rangle = \langle \Phi_{S_1 S_2} | A \otimes \mathbb{1} | \Phi_{S_1 S_2} \rangle\end{aligned}\quad (241)$$

Where  $Tr_{S_1}(\dots) = \sum_i \langle i_{S_1} | (\dots) | i_{S_1} \rangle$  is the trace operation in the space  $\mathcal{H}_{S_1}$ . So we were able to write the expectation value of the operator  $A \otimes \mathbb{1}$  in terms of quantities that only depend on  $S_1$ . But this doesn't seem like the usual expectation value with respect to a state in  $\mathcal{H}_{S_1}$ . That is, it seems that it is not possible to write this result as  $\langle \Phi_{S_1} | A | \Phi_{S_1} \rangle$  for some  $|\Phi_{S_1}\rangle \in \mathcal{H}_{S_1}$ . This conclusion is correct: if  $S_1$  and  $S_2$  are entangled the state of  $S_1$  is not described by a ket in  $\mathcal{H}_{S_1}$ . The system is said to be in a *mixed state* and its description is provided by the so called density operator, or density matrix,  $\rho_{S_1}$  defined in (240). In contrast, a state that can be described by a ket is called a *pure state*.

Therefore open systems are described by a density operator  $\rho$ . From its definition (240) one can easily see that a general density operator acting on a Hilbert space  $\mathcal{H}$  satisfies the following properties:

1. It is self-adjoint:  $\rho^\dagger = \rho$
2. It has unity trace:  $Tr \rho = 1$
3. It is nonnegative:  $\langle \psi | \rho | \psi \rangle \geq 0$ , for all  $|\psi\rangle \in \mathcal{H}$
4. A state is pure if and only if  $\rho$  is a projector: pure state  $\Leftrightarrow \rho^2 = \rho$

These properties can be taken as the definition of a density operator. Since  $\rho$  is self-adjoint and positive, one can find a basis that diagonalizes it with positive eigenvalues and write  $\rho$  in its most general and convenient form:

$$\rho = \sum_n p_n |\psi_n\rangle \langle \psi_n| \quad (242)$$

The properties above then mean that  $p_n$  are real (property 1), positive (property 3) and can be less or equal to 1 (property 2). It also follows from property 2 that if one of the  $p_n$  is equal to 1 then all the others are zero and the state is pure (property 4).

This form of the density operator allows for a much clearer interpretation of the expectation value of an operator in an open system. We found in (241) that the expectation value is given by

$$Tr(\rho A) = \sum_m \langle \psi_m | \rho A | \psi_m \rangle$$

$$\begin{aligned}
&= \sum_{mn} p_n \langle \psi_m | \psi_n \rangle \langle \psi_n | A | \psi_m \rangle \\
&= \sum_n p_n \langle \psi_n | A | \psi_n \rangle
\end{aligned} \tag{243}$$

Note that  $\langle \psi_n | A | \psi_n \rangle$  is the quantum expectation value of the operator  $A$  in the state  $|\psi_n\rangle$ . The expression above is then a mean of mean values.

What is happening here is a generalization of the idea of ensemble that we discussed for closed systems. In that case, one could interpret a state  $|\psi\rangle$  as representing an ensemble of systems all in the same state  $|\psi\rangle$ . In the case of open systems, one can interpret  $\rho$  as an ensemble of systems all in possibly different quantum states  $|\psi_n\rangle$ , the distribution of systems in these different states, that is, the probability of find one member of the ensemble in the state  $|\psi_n\rangle$  is given by  $p_n$ . Note that this is not a “quantum distribution”, in the sense that there is no interference (coherence) term between states defining  $\rho$ .

This is a generalization because the closed system case can be obtained as a special case. If all the  $p_n$  are zero except for one of them the expression reduces to the usual quantum expectation value. That is, if  $\rho = |\psi\rangle\langle\psi|$  then

$$\text{Tr}(\rho A) = \langle \psi | A | \psi \rangle \tag{244}$$

The conclusion from this discussion is that Axiom 1 must be generalized to encompass also the case of open systems as well:

1. **Axiom 1': States** *A state is the complete description of a quantum system. A state is mathematically a density operator, which is a nonnegative, self-adjoint operator with unit trace.*

Another point of view of the same situation is the following: Suppose you have a quantum system, but you don't have complete knowledge of the state. Complete knowledge would be equivalent to know in which pure state the system is in, this would be the knowledge of the microstate. But it may be the case that you don't have access to this refined information and only know broadly in which state the system may be in (you only know a macrostate). The best you can do is to assign probabilities for the possible states the system may be in compatible with the macrostate. Again, these are classical probabilities quantifying your classical ignorance. Then the description of the system will be in terms of an ensemble of possible microstates  $\{|\psi_n\rangle\}$  with probabilities  $p_n$ . That is the information contained in  $\rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|$ .

### Example: Q-bits

A particularly simple example that illustrates a bit more some of the points raised above is the system of quantum bits. A quantum bit is a two-level system that can be represented by a spin-1/2 particle or a atom transitioning between any two levels of energy, etc... Mathematically it is a system described by a Hilbert space of dimension 2. The general pure state of a q-bit is

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{245}$$

where  $a, b \in \mathbb{C}$ . For the sake of discussion, let's consider that the states are representing spin 1/2 so that  $|0\rangle = |\uparrow\rangle$  and  $|1\rangle = |\downarrow\rangle$  where  $S_z |\uparrow\rangle = \frac{\hbar}{2} |\uparrow\rangle$  and  $S_z |\downarrow\rangle = -\frac{\hbar}{2} |\downarrow\rangle$ . The state (245) describes a system such that if one measures  $S_z$  there is  $|a|^2$  probability of finding  $\frac{\hbar}{2}$  and  $|b|^2$  probability of finding  $-\frac{\hbar}{2}$  and the expectation value of  $S_z$  is

$$\langle \psi | S_z | \psi \rangle = \frac{\hbar}{2} (|a|^2 - |b|^2) \tag{246}$$

If  $a = b = \frac{1}{\sqrt{2}}$ , then there is a 50% chance of finding either one of the eigenvalues and the expectation value is zero.

Suppose now that you want to prepare a state that is completely unpolarized, in the sense that the expectation value of the spin in any direction is zero. Consider for instance the expectation value of  $S_x$ , whose action on the spin states is  $S_x|0\rangle = \frac{\hbar}{2}|1\rangle$  and  $S_x|1\rangle = \frac{\hbar}{2}|0\rangle$

$$\langle\psi|S_x|\psi\rangle = \frac{\hbar}{2}(ab^* + a^*b) \quad (247)$$

and of  $S_y$ , whose action on the spin states is  $S_y|0\rangle = i\frac{\hbar}{2}|1\rangle$  and  $S_y|1\rangle = -i\frac{\hbar}{2}|0\rangle$

$$\langle\psi|S_y|\psi\rangle = i\frac{\hbar}{2}(ab^* - a^*b) \quad (248)$$

Therefore, for all of them to be zero we should have  $|a|^2 = |b|^2$ ,  $ab^* = -a^*b$  and  $ab^* = a^*b$ . But this set of conditions is impossible to satisfy for nonzero  $a$  and  $b$ , because the last two equations imply that  $ab^* = 0$  and this implies that  $a = 0$  or  $b = 0$ , but then the first equation implies that both  $a$  and  $b$  are zero.

Therefore, a completely unpolarized state cannot be realized in a pure ensemble. But what about a mixed ensemble? The most general state describing a mixed ensemble can be written as

$$\rho = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1| \quad (249)$$

where  $p_0$  and  $p_1$  are real positive numbers that are less or equal to one, subjected to the condition  $p_0 + p_1 = 1$ . Let's compute the expectation values

$$\begin{aligned} \langle S_z \rangle &= \text{Tr}(\rho S_z) = \frac{\hbar}{2}(p_0 - p_1) \\ \langle S_x \rangle &= \text{Tr}(\rho S_x) = 0 \\ \langle S_y \rangle &= \text{Tr}(\rho S_y) = 0 \end{aligned} \quad (250)$$

Therefore, if  $p_0 = p_1 = \frac{1}{2}$  the expectation value of the components of spins along the three directions vanishes. In fact it is easy to show that the expectation value of the spin  $\hat{n} \cdot \vec{S}$  in any direction  $\hat{n}$  is zero for such a state noting that  $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\mathbb{1}$

$$\langle \hat{n} \cdot \vec{S} \rangle = \text{Tr}(\rho \hat{n} \cdot \vec{S}) = \frac{1}{2} \text{Tr}(\hat{n} \cdot \vec{S}) = 0 \quad (251)$$

because the trace of the spin1/2 operator along any direction is zero, since it is the sum of its eigenvalues that are always  $\pm\frac{\hbar}{2}$ .

The interesting thing about quantum mechanics is that the mixed ensemble may be realized as pure ensemble of a larger system. This is called purification. Consider two q-bits in a pure state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (252)$$

Let's compute the expectation value of  $\hat{n} \cdot \vec{S} \otimes \mathbb{1}$  in this state

$$\begin{aligned} \langle \Psi | \hat{n} \cdot \vec{S} \otimes \mathbb{1} | \Psi \rangle &= \frac{1}{2} \left( \langle 0 | \hat{n} \cdot \vec{S} | 0 \rangle \langle 0 | \mathbb{1} | 0 \rangle + \langle 1 | \hat{n} \cdot \vec{S} | 1 \rangle \langle 1 | \mathbb{1} | 1 \rangle \right) \\ &= \frac{1}{2} \left( \langle 0 | \hat{n} \cdot \vec{S} | 0 \rangle + \langle 1 | \hat{n} \cdot \vec{S} | 1 \rangle \right) = 0 \end{aligned} \quad (253)$$

And we see that, from the point of view of one of the q-bits this is a completely unpolarized state.

**Obs:** starting from a larger system, it is possible to obtain directly the mixed state of one of the subsystems. This is done through the partial trace operation. Consider the state described by the density operator

$|0_{S_1}0_{S_2}\rangle\langle 0_{S_1}0_{S_2}| = |0_{S_1}\rangle\langle 0_{S_1}| \otimes |0_{S_2}\rangle\langle 0_{S_2}|$ . The partial trace operation on this state over the system  $S_2$  is defined as

$$\begin{aligned} \text{Tr}_{S_2}(|0_{S_1}0_{S_2}\rangle\langle 0_{S_1}0_{S_2}|) &= |0_{S_1}\rangle\langle 0_{S_1}| \text{Tr}_{S_2}(|0_{S_2}\rangle\langle 0_{S_2}|) \\ &= |0_{S_1}\rangle\langle 0_{S_1}| (\langle 0_{S_2}|0_{S_2}\rangle\langle 0_{S_2}|0_{S_2}\rangle + \langle 1_{S_2}|0_{S_2}\rangle\langle 0_{S_2}|1_{S_2}\rangle) \\ &= |0_{S_1}\rangle\langle 0_{S_1}| \end{aligned} \quad (254)$$

Note that the partial trace returns a operator acting on the subsystem. This can obviously be generalized for a general bipartite state:  $|\Phi_{S_1S_2}\rangle = \sum_i \sum_n C_{in} |i_{S_1} n_{S_2}\rangle$

$$\begin{aligned} \text{Tr}_{S_2}(|\Phi_{S_1S_2}\rangle\langle \Phi_{S_1S_2}|) &= \sum_{ij} \sum_{nm} C_{in} C_{jm}^* |i_{S_1}\rangle\langle j_{S_1}| \text{Tr}_{S_2}(|n_{S_2}\rangle\langle m_{S_2}|) \\ &= \sum_{ij} \sum_{nm} C_{in} C_{jm}^* |i_{S_1}\rangle\langle j_{S_1}| \sum_p (\langle p_{S_2}|n_{S_2}\rangle\langle m_{S_2}|p_{S_2}\rangle) \\ &= \sum_{ij} \sum_n C_{in} C_{jn}^* |i_{S_1}\rangle\langle j_{S_1}| = \rho_{S_1} \end{aligned} \quad (255)$$

which is the same density operator obtained in (240) from a different route. Applying to our system of two q-bits (252) discussed above we find

$$\text{Tr}_{S_2}(|\Psi\rangle\langle \Psi|) = \frac{1}{2} \mathbb{1}_{S_1} \quad (256)$$

## The Bloch sphere

For a single qubit, the most general density operator can be expressed in terms of a real three-dimensional vector called the *Bloch vector*. This representation provides a powerful geometric picture of qubit states that we will use repeatedly throughout this course.

Recall that any  $2 \times 2$  Hermitian matrix can be expanded in the basis  $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ , since the Pauli matrices together with the identity form a basis for the space of  $2 \times 2$  Hermitian matrices. Since the density operator is Hermitian and has unit trace, we can write the most general qubit density operator as

$$\rho = \frac{1}{2} (\mathbb{1} + \vec{r} \cdot \vec{\sigma}) = \frac{1}{2} (\mathbb{1} + r_x \sigma_x + r_y \sigma_y + r_z \sigma_z) \quad (257)$$

where  $\vec{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$  and  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ . The factor  $1/2$  and the absence of a term multiplying  $\mathbb{1}$  (beyond the  $1/2$ ) follow from the trace condition: since each Pauli matrix is traceless,  $\text{Tr}(\rho) = \frac{1}{2}(\text{Tr}(\mathbb{1}) + 0) = 1$  for any  $\vec{r}$ . Hermiticity  $\rho^\dagger = \rho$  is also automatic because each Pauli matrix is Hermitian and  $\vec{r}$  is real.

The remaining requirement is that  $\rho$  be a *positive* operator. This is equivalent to requiring that all eigenvalues of  $\rho$  be non-negative. The eigenvalues of  $\rho$  are

$$\lambda_{\pm} = \frac{1}{2} (1 \pm |\vec{r}|) \quad (258)$$

which can be verified by direct diagonalization (this is left as an exercise). Both eigenvalues are non-negative if and only if  $|\vec{r}| \leq 1$ .

This means that all physical qubit states correspond to vectors  $\vec{r}$  lying within or on the unit sphere in  $\mathbb{R}^3$ . This is the *Bloch sphere*.

Let us now characterize pure and mixed states through the *purity*  $\text{Tr}(\rho^2)$ . Using the identity  $(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b}) \mathbb{1} + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma}$ , which follows from the Pauli algebra  $\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i\epsilon_{ijk} \sigma_k$ , we compute:

$$\begin{aligned} \rho^2 &= \frac{1}{4}(\mathbb{1} + \vec{r} \cdot \vec{\sigma})^2 = \frac{1}{4}(\mathbb{1} + 2\vec{r} \cdot \vec{\sigma} + (\vec{r} \cdot \vec{\sigma})^2) \\ &= \frac{1}{4}(\mathbb{1} + 2\vec{r} \cdot \vec{\sigma} + |\vec{r}|^2 \mathbb{1}) = \frac{1}{4}((1 + |\vec{r}|^2)\mathbb{1} + 2\vec{r} \cdot \vec{\sigma}) \end{aligned} \quad (259)$$

where we used  $\vec{r} \times \vec{r} = 0$ . Taking the trace:

$$\text{Tr}(\rho^2) = \frac{1}{2}(1 + |\vec{r}|^2) \quad (260)$$

Therefore:

- *Pure states*:  $\text{Tr}(\rho^2) = 1$  requires  $|\vec{r}| = 1$ . Pure states lie on the *surface* of the Bloch sphere.
- *Mixed states*:  $\text{Tr}(\rho^2) < 1$  corresponds to  $|\vec{r}| < 1$ , i.e., the *interior* of the sphere. The more mixed the state, the shorter the Bloch vector.
- *Maximally mixed state*:  $\rho = \mathbb{1}/2$  corresponds to  $\vec{r} = 0$  (the center of the sphere), with purity  $1/2$ . This is the completely depolarized state we encountered in the previous discussion.

Now let us connect this parametrization to the standard representation of pure qubit states. The most general pure state (up to global phase) can be written as

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \quad (261)$$

with  $\theta \in [0, \pi]$  and  $\phi \in [0, 2\pi)$ . The corresponding density matrix is

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2 \frac{\theta}{2} & e^{-i\phi} \cos \frac{\theta}{2} \sin \frac{\theta}{2} \\ e^{i\phi} \cos \frac{\theta}{2} \sin \frac{\theta}{2} & \sin^2 \frac{\theta}{2} \end{pmatrix} \quad (262)$$

Comparing with (257), which in matrix form reads

$$\rho = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix} \quad (263)$$

we can extract the Bloch vector components:

$$r_x = \sin \theta \cos \phi, \quad r_y = \sin \theta \sin \phi, \quad r_z = \cos \theta \quad (264)$$

This is precisely the unit vector in spherical coordinates  $(\theta, \phi)$ . Therefore the angles  $\theta$  and  $\phi$  in the parametrization of the pure state (261) are the polar and azimuthal angles on the Bloch sphere. The geometric interpretation is:

- *North pole* ( $\theta = 0$ ,  $\vec{r} = \hat{z}$ ): the state  $|0\rangle$ , eigenstate of  $\sigma_z$  with eigenvalue  $+1$ .
- *South pole* ( $\theta = \pi$ ,  $\vec{r} = -\hat{z}$ ): the state  $|1\rangle$ , eigenstate of  $\sigma_z$  with eigenvalue  $-1$ .
- *Equator* ( $\theta = \pi/2$ ): equal superpositions of  $|0\rangle$  and  $|1\rangle$  with varying relative phase  $\phi$ . For example, the states  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$  and  $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$  are at  $\phi = 0$  and  $\phi = \pi$ , respectively.
- *Antipodal points* on the sphere correspond to *orthogonal* states: if  $\vec{r}$  represents  $|\psi\rangle$ , then  $-\vec{r}$  represents the state orthogonal to  $|\psi\rangle$ .

Note that the components of the Bloch vector have a direct physical meaning:  $r_i = \text{Tr}(\rho \sigma_i) = \langle \sigma_i \rangle$  is the expectation value of the  $i$ -th Pauli matrix. Measuring the spin along the three Cartesian directions therefore determines the Bloch vector completely — a procedure known as *quantum state tomography* for a single qubit.

### 3.1 Quantum statistical entropy and entanglement entropy

We already know that a mixed ensemble is a result of our ignorance about which pure state the system is in. This lack of knowledge has its origin in a limited information about the system. If we consider the system as a subsystem of a larger system, we saw that the ignorance comes from not knowing or not having access to the other complementary subsystems.

Entropy is a quantity that quantifies information (or ignorance). Let's construct the corresponding expression for quantum systems.

Let's start by considering a general mixed state

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (265)$$

where the  $|\psi_i\rangle$  are orthonormalized vectors and  $\sum_i p_i = 1$ . The number of microstates in this ensemble is obtained as follows. Suppose there are  $N$  copies of this system. There is a probability  $p_i$  of finding a particular copy in the state  $|\psi_i\rangle$ . By the very definition of probability, the number of microstates in the state  $|\psi_i\rangle$  is  $N_i = p_i N$ , such that  $\sum_i N_i = N$ . Thus, the total number of microstates in this ensemble is

$$\Omega = \frac{N!}{N_1! N_2! \cdots N_N!} \quad (266)$$

The statistical entropy is defined as

$$\begin{aligned} S_N = k \ln \Omega &= k \ln \frac{N!}{N_1! N_2! \cdots N_N!} = k \ln N! - k \sum_i^N \ln (p_i N)! \\ &\approx kN \ln N - kN - k \sum_i^N (p_i N \ln (p_i N) - p_i N) \\ &= -kN \sum_i^N p_i \ln p_i \end{aligned} \quad (267)$$

Therefore we can associate an entropy to each system given by

$$S = S_N/N = -k \sum_i^N p_i \ln p_i \quad (268)$$

Or, written equivalently, in terms of the density matrix

$$S = -k \text{Tr}(\rho \ln \rho) \quad (269)$$

This form of the entropy is known as von Neumann entropy. As we already know it quantifies our ignorance about the state. In the context of quantum theory, it is also known as the entanglement entropy and it is a way to quantify the amount of entanglement in a system.

The von Neumann entropy satisfy many interesting properties:

1.  $S \geq 0$  and it is zero only if all the probabilities  $p_i$  are zero, except for one of them which is equal to 1. This is the case of a pure state. Therefore pure states have zero entropy, which makes sense because a pure state is the maximum amount of information one can have about a quantum system.
2. The entropy attains its maximum values when the state is maximally mixed, that is, when all probabilities  $p_i$  are equal. If there are  $N$  states in the ensemble,  $p_i = 1/N$  and  $S = k \ln N$ .

3. The entropy satisfies an additivity property under combination of two independent systems. That is, if  $\rho_{S_1 S_2} = \rho_{S_1} \otimes \rho_{S_2}$ , then one can write in diagonal basis

$$\rho_{S_1 S_2} = \rho_{S_1} \otimes \rho_{S_2} = \sum_{i,n} p_{S_1 i} p_{S_2 n} |i, S_1\rangle \langle i, S_1| \otimes |n, S_2\rangle \langle n, S_2| \quad (270)$$

and therefore

$$\begin{aligned} S(\rho_{S_1 S_2}) &= -k \text{Tr} \rho_{S_1 S_2} \ln \rho_{S_1 S_2} = -k \sum_{i,n} \langle i, S_1; n, S_2 | \rho_{S_1 S_2} \ln \rho_{S_1 S_2} | i, S_1; n, S_2 \rangle \\ &= -k \sum_{i,n} p_{S_1 i} p_{S_2 n} \ln (p_{S_1 i} p_{S_2 n}) \\ &= -k \sum_i p_{S_1 i} \ln (p_{S_1 i}) + -k \sum_n p_{S_2 n} \ln (p_{S_2 n}) \\ &= S(\rho_{S_1}) + S(\rho_{S_2}) \end{aligned} \quad (271)$$

4. On the other hand, if the systems are not independent, the entropy obeys a subadditivity property.

$$S(\rho_{S_1 S_2}) \leq S(\rho_{S_1}) + S(\rho_{S_2}) \quad (272)$$

The proof can be left as an exercise. The meaning of this relation is that when there are correlations among the parts the total entropy is less than it would be without correlations. This makes sense because correlations is related to order in the system.

5. If the probability distribution don't depend on time, the entropy is stationary. This is easy to see because in this case the time evolution of the density matrix follows immediately from Schrodinger's equation (Axiom 4). Given a quantum state of a closed system we know it evolves unitarily as  $|\psi(t)\rangle = U(t, t_0)|\psi(0)\rangle$  and it follows then, from linearity, that  $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$  evolves as

$$\rho(t) = \sum_i p_i U(t, t_0) |\psi_i\rangle \langle \psi_i| U^\dagger(t, t_0) = U(t, t_0) \rho(t_0) U^\dagger(t, t_0) \quad (273)$$

Under a general unitary transformation of the density matrix, the entropy transforms as

$$\begin{aligned} S(U\rho U^\dagger) &= -k \text{Tr} (U\rho U^\dagger \ln (U\rho U^\dagger)) \\ &= -k \text{Tr} (U\rho U^\dagger U \ln (\rho) U^\dagger) \\ &= -k \text{Tr} (\rho \ln \rho) \end{aligned} \quad (274)$$

Where in the second line we used the property  $\ln(UAU^\dagger) = U \ln(A) U^\dagger$ <sup>7</sup> and in the third line the ciclicity of the trace was used. Therefore:  $S(t) = S(t_0)$  and the entropy is stationary

This is an interesting result and in fact it can be viewed in the light of the quantum version of the Liouville theorem. Recall that in classical mechanics Liouville's theorem says that the volume of phase space remains constant during time evolution and this means that the entropy of the system remains the same, since no information is lost. Only when we coarse grain (so that the volume of phase space changes) the entropy increases. The quantum version of Liouville's theorem is due to von Neumann and it is simply the infinitesimal version of the unitarity evolution, that is, Schrodinger equation for the density matrix. In fact consider taking the time derivative of the density matrix  $\rho(t) = \sum_i p_i |\psi_i(t)\rangle \langle \psi_i(t)|$ :

$$\frac{\partial \rho(t)}{\partial t} = \sum_i p_i \frac{\partial |\psi_i(t)\rangle}{\partial t} \langle \psi_i(t)| + \sum_i p_i |\psi_i(t)\rangle \frac{\partial \langle \psi_i(t)|}{\partial t}$$

---

<sup>7</sup>Note that, if  $A = e^B$ , then  $UAU^\dagger = Ue^B U^\dagger = \sum_n \frac{UB^n U^\dagger}{n!} = \sum_n \frac{(UBU^\dagger)^n}{n!} = e^{UBU^\dagger}$  Therefore  $\ln(UAU^\dagger) = UBU^\dagger = U(\ln A)U^\dagger$

$$\begin{aligned}
&= \sum_i p_i \frac{1}{i\hbar} H |\psi_i(t)\rangle \langle \psi_i(t)| + \sum_i p_i \left( \frac{-1}{i\hbar} \right) |\psi_i(t)\rangle \langle \psi_i(t)| H \\
&= \frac{1}{i\hbar} [H, \rho]
\end{aligned} \tag{275}$$

where Schrodinger's equation was used and we see that this is in fact similar to (??), with the density matrix playing the role of probability density. So, in an equilibrium situation we must have  $[H, \rho] = 0$  and therefore in this case the density matrix must be a function of conserved quantities. But note that the entropy is invariant under any unitary transformation of the density matrix, without assuming equilibrium. This is because unitary evolution conserves the information content. The actual equilibrium situation is attained by maximizing the entropy.

### 3.2 Entanglement and non-locality

Entanglement is not merely a mathematical curiosity of the tensor product structure — it has observable physical consequences. A composite quantum system can exhibit correlations between spatially separated subsystems that are *stronger* than any correlations producible by classical systems. This remarkable feature, known as *non-locality*, is captured precisely by Bell's theorem and its experimental verification.

#### Bell states and the Bell basis

For a system of two qubits with Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_B$ , we define the four *Bell states*:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \tag{276}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \tag{277}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \tag{278}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \tag{279}$$

where  $|ij\rangle \equiv |i\rangle_A \otimes |j\rangle_B$ . It is straightforward to verify that these four states are orthonormal:

$$\langle \Phi^\pm | \Phi^\pm \rangle = 1, \quad \langle \Psi^\pm | \Psi^\pm \rangle = 1, \quad \text{and all cross terms vanish.} \tag{280}$$

Since  $\mathcal{H}_A \otimes \mathcal{H}_B$  is four-dimensional, these four states constitute a complete orthonormal basis — the *Bell basis*. Any two-qubit state can be expanded in terms of Bell states.

Each Bell state is *maximally entangled*. We verify this for  $|\Phi^+\rangle$  by computing the reduced density matrix for subsystem A:

$$\begin{aligned}
\rho_A &= \text{Tr}_B (|\Phi^+\rangle \langle \Phi^+|) = \text{Tr}_B \left[ \frac{1}{2} (|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|) \right] \\
&= \frac{1}{2} \left( |0\rangle \langle 0| \underbrace{\langle 0|0\rangle}_1 + |0\rangle \langle 1| \underbrace{\langle 1|0\rangle}_0 + |1\rangle \langle 0| \underbrace{\langle 0|1\rangle}_0 + |1\rangle \langle 1| \underbrace{\langle 1|1\rangle}_1 \right) \\
&= \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{1}{2} \mathbb{1}
\end{aligned} \tag{281}$$

The reduced state is the maximally mixed state. On the Bloch sphere it corresponds to  $\vec{r} = 0$ , the center of the sphere. The entanglement entropy is  $S(\rho_A) = -\text{Tr}(\rho_A \ln \rho_A) = \ln 2$ , the maximum possible value for a two-level system. The same calculation yields  $\rho_A = \mathbb{1}/2$  for all four Bell states (this is left as an exercise).

Note that the state  $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  is precisely the purified state (252) we encountered in the qubits discussion, where we showed that from the perspective of one qubit the state is completely unpolarized. The Bell basis provides a systematic framework for understanding this phenomenon.

The Bell states also have a neat algebraic structure: they are related to each other by local Pauli operations on one of the qubits. Starting from  $|\Phi^+\rangle$ :

$$|\Phi^-\rangle = (\sigma_z \otimes \mathbb{1})|\Phi^+\rangle, \quad |\Psi^+\rangle = (\sigma_x \otimes \mathbb{1})|\Phi^+\rangle, \quad |\Psi^-\rangle = (i\sigma_y \otimes \mathbb{1})|\Phi^+\rangle \quad (282)$$

This property will be essential for the superdense coding protocol discussed later.

## The CHSH inequality and Bell's theorem

The question of whether quantum mechanics is a complete description of physical reality was raised in a famous 1935 paper by Einstein, Podolsky, and Rosen (EPR). They argued that quantum mechanics might be incomplete and that there could exist *hidden variables* — additional physical quantities not described by the quantum state — that predetermine the outcomes of measurements. In 1964, John Bell showed that any such *local hidden variable* (LHV) theory makes predictions that differ from those of quantum mechanics, and that the difference is experimentally testable. The most experimentally convenient form of Bell's result is the *CHSH inequality*, due to Clauser, Horne, Shimony, and Holt [8, 2].

**Setup.** Consider two spatially separated observers, Alice and Bob, who share a pair of qubits. Alice can choose to measure one of two observables on her qubit:  $\hat{A}_1$  or  $\hat{A}_2$ . Similarly, Bob can choose between  $\hat{B}_1$  or  $\hat{B}_2$ . Each observable has eigenvalues  $\pm 1$  (for instance, spin measurements along chosen directions:  $\hat{A}_i = \vec{a}_i \cdot \vec{\sigma}$  and  $\hat{B}_j = \vec{b}_j \cdot \vec{\sigma}$  for unit vectors  $\vec{a}_i, \vec{b}_j$ ).

Define the *CHSH correlator*:

$$S = \langle \hat{A}_1 \otimes \hat{B}_1 \rangle + \langle \hat{A}_1 \otimes \hat{B}_2 \rangle + \langle \hat{A}_2 \otimes \hat{B}_1 \rangle - \langle \hat{A}_2 \otimes \hat{B}_2 \rangle \quad (283)$$

**Classical bound from local hidden variables.** In a local hidden variable theory, the measurement outcomes are predetermined by a hidden variable  $\lambda$ , distributed according to some probability density  $p(\lambda)$ . Locality means that Alice's outcome  $a_i(\lambda) \in \{-1, +1\}$  depends only on her measurement choice and  $\lambda$ , not on Bob's choice, and vice versa. The correlator is then

$$S_{\text{LHV}} = \int d\lambda p(\lambda) [a_1(\lambda)b_1(\lambda) + a_1(\lambda)b_2(\lambda) + a_2(\lambda)b_1(\lambda) - a_2(\lambda)b_2(\lambda)] \quad (284)$$

For any fixed  $\lambda$ , consider the quantity inside the brackets:

$$\begin{aligned} s(\lambda) &= a_1b_1 + a_1b_2 + a_2b_1 - a_2b_2 \\ &= a_1(b_1 + b_2) + a_2(b_1 - b_2) \end{aligned} \quad (285)$$

where we suppressed the  $\lambda$  dependence for brevity. Since  $b_1, b_2 \in \{-1, +1\}$ , either  $b_1 = b_2$  or  $b_1 = -b_2$ . If  $b_1 = b_2$ , then  $(b_1 + b_2) = \pm 2$  and  $(b_1 - b_2) = 0$ , so  $|s(\lambda)| = 2|a_1| = 2$ . If  $b_1 = -b_2$ , then  $(b_1 + b_2) = 0$  and  $(b_1 - b_2) = \pm 2$ , so  $|s(\lambda)| = 2|a_2| = 2$ . In either case:

$$|s(\lambda)| = 2 \quad \text{for all } \lambda \quad (286)$$

Averaging over  $\lambda$ :

$$|S_{\text{LHV}}| \leq \int d\lambda p(\lambda) |s(\lambda)| = 2 \quad (287)$$

This is the *CHSH inequality*:

$$\boxed{|S_{\text{LHV}}| \leq 2} \quad (288)$$

Any local hidden variable theory must satisfy this bound.

**Quantum violation.** Now let us compute the quantum prediction for a specific state and specific measurement choices. Consider the singlet state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \quad (289)$$

For spin measurements along directions  $\vec{a}$  (Alice) and  $\vec{b}$  (Bob), the correlation function for the singlet state is

$$E(\vec{a}, \vec{b}) \equiv \langle \Psi^- | (\vec{a} \cdot \vec{\sigma}) \otimes (\vec{b} \cdot \vec{\sigma}) | \Psi^- \rangle = -\vec{a} \cdot \vec{b} \quad (290)$$

Let us derive this result. Write  $\vec{a} \cdot \vec{\sigma} = a_x \sigma_x + a_y \sigma_y + a_z \sigma_z$  and similarly for  $\vec{b} \cdot \vec{\sigma}$ . Then

$$E(\vec{a}, \vec{b}) = \sum_{i,j} a_i b_j \langle \Psi^- | \sigma_i \otimes \sigma_j | \Psi^- \rangle \quad (291)$$

A direct computation shows that the only nonzero contributions come from the diagonal terms  $i = j$ :

$$\langle \Psi^- | \sigma_i \otimes \sigma_j | \Psi^- \rangle = -\delta_{ij} \quad (292)$$

This can be verified by explicit calculation for each Pauli matrix. For example, with  $\sigma_z$ :

$$\begin{aligned} \langle \Psi^- | \sigma_z \otimes \sigma_z | \Psi^- \rangle &= \frac{1}{2} (\langle 01| - \langle 10|) (\sigma_z \otimes \sigma_z) (|01\rangle - |10\rangle) \\ &= \frac{1}{2} (\langle 01| - \langle 10|) (-|01\rangle + |10\rangle) \\ &= \frac{1}{2} (-1 - 1) = -1 \end{aligned} \quad (293)$$

where we used  $\sigma_z|0\rangle = |0\rangle$ ,  $\sigma_z|1\rangle = -|1\rangle$ . Similar calculations for  $\sigma_x$  and  $\sigma_y$  also yield  $-1$  (this is left as an exercise). Therefore  $E(\vec{a}, \vec{b}) = -\sum_i a_i b_i = -\vec{a} \cdot \vec{b}$ .

Now choose measurement directions in the  $xz$ -plane (we can always orient our coordinate system so that all directions lie in this plane). Let:

$$\vec{a}_1 = \hat{z}, \quad \vec{a}_2 = \hat{x} \quad (294)$$

and

$$\vec{b}_1 = \frac{1}{\sqrt{2}}(\hat{z} + \hat{x}), \quad \vec{b}_2 = \frac{1}{\sqrt{2}}(\hat{z} - \hat{x}) \quad (295)$$

That is, Bob's measurement directions are rotated by  $\pm\pi/4$  from Alice's directions. The four correlations are:

$$\begin{aligned} E(\vec{a}_1, \vec{b}_1) &= -\vec{a}_1 \cdot \vec{b}_1 = -\frac{1}{\sqrt{2}} \\ E(\vec{a}_1, \vec{b}_2) &= -\vec{a}_1 \cdot \vec{b}_2 = -\frac{1}{\sqrt{2}} \end{aligned}$$

$$\begin{aligned}
E(\vec{a}_2, \vec{b}_1) &= -\vec{a}_2 \cdot \vec{b}_1 = -\frac{1}{\sqrt{2}} \\
E(\vec{a}_2, \vec{b}_2) &= -\vec{a}_2 \cdot \vec{b}_2 = +\frac{1}{\sqrt{2}}
\end{aligned} \tag{296}$$

The CHSH correlator is therefore:

$$S_{\text{QM}} = E(\vec{a}_1, \vec{b}_1) + E(\vec{a}_1, \vec{b}_2) + E(\vec{a}_2, \vec{b}_1) - E(\vec{a}_2, \vec{b}_2) = -\frac{4}{\sqrt{2}} = -2\sqrt{2} \tag{297}$$

so that

$$|S_{\text{QM}}| = 2\sqrt{2} \approx 2.828 > 2 \tag{298}$$

The quantum prediction *violates* the CHSH inequality! The value  $2\sqrt{2}$  is, in fact, the maximum achievable by quantum mechanics for any state and any choice of measurements. This upper bound is known as the *Tsirelson bound*.

**Bell's theorem and its significance.** Bell's theorem can be stated as follows: *No local hidden variable theory can reproduce all the predictions of quantum mechanics.* The CHSH inequality provides a concrete, experimentally testable criterion. Experiments testing Bell inequalities — starting with the pioneering work of Aspect, Dalibard, and Roger in 1982, through the “loophole-free” tests by Hensen *et al.* (2015) and others — have consistently confirmed the quantum mechanical predictions and ruled out local hidden variable theories.

The physical content of this result is profound: quantum correlations between entangled particles cannot be explained by assuming that each particle carries predetermined values for all observables (realism) and that measurements on one particle do not instantaneously affect the other (locality). At least one of these assumptions must be abandoned. Most interpretations of quantum mechanics give up local realism while maintaining the impossibility of faster-than-light *signaling*, which is guaranteed by the no-signaling theorem (the reduced density matrix of one party is independent of the measurement choice of the other).

For a thorough discussion see [8], Chapter 2; [2], Chapter 4; and [1], §3.10.

## Time evolution of open systems

In order to understand how the entropy can grow with time one has to consider more fully what it means for a system to be open. A nice way to approach the problem is to start with a closed system and investigate what happens to a subsystem. Consider a separable (non-entangled) state at  $t_0$ :  $|\Phi_{S_1 S_2}(t_0)\rangle = |\Phi_{S_1}(t_0)\rangle|\Phi_{S_2}(t_0)\rangle$ . Suppose the whole system  $S_1 + S_2$  is closed. Therefore the state at a later time  $t$  is obtained by applying the time evolution unitary operator

$$|\Phi_{S_1 S_2}(t)\rangle = U(t, t_0)|\Phi_{S_1 S_2}(t_0)\rangle \tag{299}$$

Now, suppose we don't have access to system  $S_2$  or we simply don't care about it, how would we describe the time evolution from the point of view of system  $S_1$ ? What we want is a description of how the reduced density matrix  $\rho_{S_1} = \text{Tr}_{S_2} \rho$  evolves with time given the unitary evolution of the density matrix  $\rho$  of the whole system. Thus we have

$$\begin{aligned}
\rho_{S_1}(t) &= \text{Tr}_{S_2} (|\Phi_{S_1 S_2}(t)\rangle\langle\Phi_{S_1 S_2}(t)|) \\
&= \text{Tr}_{S_2} (U(t, t_0)|\Phi_{S_1 S_2}(t_0)\rangle\langle\Phi_{S_1 S_2}(t_0)|U^\dagger(t, t_0)) \\
&= \sum_n \langle n_{S_2} | U(t, t_0) |\Phi_{S_1 S_2}(t_0)\rangle \langle\Phi_{S_1 S_2}(t_0) | U^\dagger(t, t_0) | n_{S_2} \rangle
\end{aligned} \tag{300}$$

where  $\{|n_{S_2}\rangle\}$  is a base for  $\mathcal{H}_2$ . Note that  $\langle n_{S_2}|U(t, t_0)|\Phi_{S_1 S_2}(t_0)\rangle$  is a vector in  $\mathcal{H}_1$ , in fact

$$\langle n_{S_2}|U(t, t_0)|\Phi_{S_1 S_2}(t_0)\rangle = \langle n_{S_2}|U(t, t_0)(|\Phi_{S_1}(t_0)\rangle|\Phi_{S_2}(t_0)\rangle) = M_n|\Phi_{S_1}(t_0)\rangle \quad (301)$$

Therefore

$$\rho_{S_1}(t) = \sum_n M_n|\Phi_{S_1}(t_0)\rangle\langle\Phi_{S_1}(t_0)|M_n^\dagger = \sum_n M_n\rho_{S_1}(t_0)M_n^\dagger \quad (302)$$

where the operator  $M_n : \mathcal{H}_1 \rightarrow \mathcal{H}_1$  is defined by  $M_n = \langle n_{S_2}|U(t, t_0)|\Phi_{S_2}(t_0)\rangle$ . From the point of view of  $S_1$  then the time evolution is not unitary because the operators  $M_n$  are not unitary. Note that the  $M_n$  depend on the chosen basis for  $S_2$ , the initial state of  $S_2$  and on the time evolution unitary operator of the whole system, but if we don't know anything about system  $S_2$  we have to construct these operators by other means. Even though  $M_n$  are not unitary they satisfy an important relation

$$\sum_n M_n^\dagger M_n = \mathbb{1} \quad (303)$$

The operators  $M_n$  are known as Kraus operators. A map  $\mathcal{E}(\rho) = \sum_n M_n \rho M_n^\dagger$  such as the one defined in (302) with operators  $M_n$  satisfying the relation (303) is called a quantum channel.

A quantum channel maps density matrices to density matrices and has the following properties (try to verify them):

1. Linearity:  $\mathcal{E}(\alpha\rho_1 + \beta\rho_2) = \alpha\mathcal{E}(\rho_1) + \beta\mathcal{E}(\rho_2)$ .
2. Preserves Hermiticity:  $\rho = \rho^\dagger$  implies  $\mathcal{E}(\rho) = \mathcal{E}(\rho)^\dagger$ .
3. Preserves positivity:  $\rho \geq 0$  implies  $\mathcal{E}(\rho) \geq 0$ .
4. Preserves trace:  $Tr(\mathcal{E}(\rho)) = Tr(\rho)$ .

A unitary transformation is a special case of this. But in general a quantum channel is not unitary. A immediate consequence of this is the fact the reversibility is lost. In fact it can be shown (try) that a quantum channel is reversible if and only if it is unitary. Therefore, under a general quantum channel transformation the entropy can change. It is possible to show that it will grow or stay the same, staying the same if the channel is unitary.

The conclusion is that we have to update the Axiom 4 to encompass also open systems:

**Axiom 4': Dynamics** *Time evolution is described by a quantum channel.*

### The no-cloning theorem

A fundamental distinction between classical and quantum information is that classical bits can be freely copied, but quantum states cannot. This is the content of the *no-cloning theorem*, first proved by Wootters and Zurek (1982) and independently by Dieks (1982).

**Statement.** There is no unitary operator  $U$  acting on  $\mathcal{H} \otimes \mathcal{H}$  such that

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \quad (304)$$

for all  $|\psi\rangle \in \mathcal{H}$ , where  $|0\rangle$  is a fixed initial state of the “target” qubit.

**Proof.** Suppose, for the sake of contradiction, that such a  $U$  exists. Consider two arbitrary states  $|\psi\rangle$  and  $|\phi\rangle$  in  $\mathcal{H}$ . By hypothesis:

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \quad (305)$$

$$U |\phi\rangle|0\rangle = |\phi\rangle|\phi\rangle \quad (306)$$

Take the inner product of the left-hand sides:

$$\langle\psi|\langle 0| U^\dagger U |\phi\rangle|0\rangle = \langle\psi|\phi\rangle \langle 0|0\rangle = \langle\psi|\phi\rangle \quad (307)$$

where we used the unitarity of  $U$  (so  $U^\dagger U = \mathbb{1}$ ) and the normalization  $\langle 0|0\rangle = 1$ . Now take the inner product of the right-hand sides:

$$\langle\psi|\langle\psi||\phi\rangle|\phi\rangle = \langle\psi|\phi\rangle \langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2 \quad (308)$$

Since the inner products must be equal, we obtain:

$$\langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2 \quad (309)$$

Setting  $c = \langle\psi|\phi\rangle$ , this gives  $c = c^2$ , which implies  $c(c - 1) = 0$ , so  $c = 0$  or  $c = 1$ . Therefore:

- If  $\langle\psi|\phi\rangle = 0$ : the states are orthogonal, and cloning is consistent.
- If  $\langle\psi|\phi\rangle = 1$ : the states are identical (up to phase), and cloning is trivially satisfied.

For any pair of states with  $0 < |\langle\psi|\phi\rangle| < 1$  (i.e., non-orthogonal and non-identical), the assumed cloning operation leads to a contradiction. Since an arbitrary unknown state may be non-orthogonal to any other state, no single unitary can clone all states.  $\square$

**Physical significance.** The no-cloning theorem has far-reaching consequences:

1. It establishes a fundamental asymmetry between classical and quantum information: classical bits (0 or 1) can be copied because they are always in one of two orthogonal states, and cloning orthogonal states is allowed.
2. It ensures the security of *quantum key distribution* (QKD) protocols such as BB84: an eavesdropper cannot intercept a quantum state, copy it, and forward the original without being detected.
3. It is deeply connected to the linearity of quantum mechanics. If cloning were possible, one could use it to distinguish non-orthogonal states perfectly, to signal faster than light, and to violate the uncertainty principle.

Note that the no-cloning theorem does not forbid *approximate* cloning (with reduced fidelity) nor cloning of states drawn from a *known* orthogonal set.

See [8], §12.1 and [2], §4.1 for further discussion.

## Quantum teleportation

*Quantum teleportation* is a protocol that allows Alice to transmit an unknown quantum state to Bob using only a pre-shared entangled pair and two bits of classical communication. It was proposed by Bennett, Brassard, Crépeau, Jozsa, Peres, and Wootters in 1993 [8].

**Setup.** Alice possesses a qubit in an unknown state

$$|\psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1 \quad (310)$$

with  $|\alpha|^2 + |\beta|^2 = 1$ . Alice and Bob also share a Bell pair in the state  $|\Phi^+\rangle_{23}$ , where qubit 2 belongs to Alice and qubit 3 belongs to Bob. The total state of the three qubits is:

$$|\Psi\rangle_{123} = |\psi\rangle_1 \otimes |\Phi^+\rangle_{23} = (\alpha|0\rangle_1 + \beta|1\rangle_1) \otimes \frac{1}{\sqrt{2}} (|00\rangle_{23} + |11\rangle_{23})$$

$$= \frac{1}{\sqrt{2}} (\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \quad (311)$$

where in the last line the three slots refer to qubits 1, 2, 3 in order.

**Rewriting in the Bell basis.** The key step is to rewrite the state of qubits 1 and 2 (both held by Alice) in the Bell basis. Using the inverse relations

$$\begin{aligned} |00\rangle &= \frac{1}{\sqrt{2}} (|\Phi^+\rangle + |\Phi^-\rangle), & |11\rangle &= \frac{1}{\sqrt{2}} (|\Phi^+\rangle - |\Phi^-\rangle) \\ |01\rangle &= \frac{1}{\sqrt{2}} (|\Psi^+\rangle + |\Psi^-\rangle), & |10\rangle &= \frac{1}{\sqrt{2}} (|\Psi^+\rangle - |\Psi^-\rangle) \end{aligned} \quad (312)$$

we substitute into (311). The four terms become:

$$\begin{aligned} \alpha|0\rangle_1|0\rangle_2|0\rangle_3 &= \frac{\alpha}{\sqrt{2}} (|\Phi^+\rangle_{12} + |\Phi^-\rangle_{12}) |0\rangle_3 \\ \alpha|0\rangle_1|1\rangle_2|1\rangle_3 &= \frac{\alpha}{\sqrt{2}} (|\Psi^+\rangle_{12} + |\Psi^-\rangle_{12}) |1\rangle_3 \\ \beta|1\rangle_1|0\rangle_2|0\rangle_3 &= \frac{\beta}{\sqrt{2}} (|\Psi^+\rangle_{12} - |\Psi^-\rangle_{12}) |0\rangle_3 \\ \beta|1\rangle_1|1\rangle_2|1\rangle_3 &= \frac{\beta}{\sqrt{2}} (|\Phi^+\rangle_{12} - |\Phi^-\rangle_{12}) |1\rangle_3 \end{aligned} \quad (313)$$

Collecting terms by Bell state of qubits 1 and 2:

$$\begin{aligned} |\Psi\rangle_{123} &= \frac{1}{2} \left[ |\Phi^+\rangle_{12} (\alpha|0\rangle_3 + \beta|1\rangle_3) + |\Phi^-\rangle_{12} (\alpha|0\rangle_3 - \beta|1\rangle_3) \right. \\ &\quad \left. + |\Psi^+\rangle_{12} (\alpha|1\rangle_3 + \beta|0\rangle_3) + |\Psi^-\rangle_{12} (\alpha|1\rangle_3 - \beta|0\rangle_3) \right] \end{aligned} \quad (314)$$

This can be written more compactly as:

$$|\Psi\rangle_{123} = \frac{1}{2} \left[ |\Phi^+\rangle_{12} |\psi\rangle_3 + |\Phi^-\rangle_{12} \sigma_z |\psi\rangle_3 + |\Psi^+\rangle_{12} \sigma_x |\psi\rangle_3 + |\Psi^-\rangle_{12} (-i\sigma_y) |\psi\rangle_3 \right] \quad (315)$$

where  $\sigma_z |\psi\rangle = \alpha|0\rangle - \beta|1\rangle$ ,  $\sigma_x |\psi\rangle = \alpha|1\rangle + \beta|0\rangle$ , and  $-i\sigma_y |\psi\rangle = \alpha|1\rangle - \beta|0\rangle$  (which equals  $\sigma_x \sigma_z |\psi\rangle$ ).

### Protocol.

1. **Alice's Bell measurement.** Alice performs a projective measurement on qubits 1 and 2 in the Bell basis. Each of the four outcomes occurs with probability 1/4.
2. **Classical communication.** Alice sends her measurement result (2 classical bits encoding which of the four Bell states she obtained) to Bob through a classical channel.
3. **Bob's correction.** Depending on Alice's message, Bob applies a unitary correction to qubit 3:

| Alice's result        | Bob's state                        | Bob's correction                |
|-----------------------|------------------------------------|---------------------------------|
| $ \Phi^+\rangle_{12}$ | $\alpha 0\rangle + \beta 1\rangle$ | $\mathbb{1}$ (do nothing)       |
| $ \Phi^-\rangle_{12}$ | $\alpha 0\rangle - \beta 1\rangle$ | $\sigma_z$                      |
| $ \Psi^+\rangle_{12}$ | $\alpha 1\rangle + \beta 0\rangle$ | $\sigma_x$                      |
| $ \Psi^-\rangle_{12}$ | $\alpha 1\rangle - \beta 0\rangle$ | $i\sigma_y = \sigma_x \sigma_z$ |

After applying the correction, Bob's qubit is in the state  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ .

### Key observations.

1. *No faster-than-light communication:* Before receiving Alice's classical message, the state of Bob's qubit (from his perspective) is the maximally mixed state  $\rho_3 = \mathbb{1}/2$ , regardless of  $|\psi\rangle$ . This can be seen by tracing out qubits 1 and 2 from (314). Bob cannot extract any information about  $|\psi\rangle$  without Alice's classical bits. The protocol therefore respects causality.
2. *No violation of no-cloning:* Alice's Bell measurement destroys the original state  $|\psi\rangle$  in qubit 1. After the measurement, qubits 1 and 2 are in one of the Bell states and no longer carry information about  $|\psi\rangle$ . The state has been *transferred*, not copied.
3. *Entanglement is consumed:* The Bell pair shared between Alice and Bob is used up in the process. After the protocol, qubits 2 and 3 are no longer entangled.
4. *Resource accounting:* Teleporting one qubit requires 1 shared Bell pair (1 ebit of entanglement) and 2 classical bits of communication.

Quantum teleportation has been demonstrated experimentally in numerous settings, including photonic systems, trapped ions, and even over satellite links. It is a central building block for quantum repeaters (which extend the range of quantum communication) and for distributed quantum computing architectures. See [8], §1.3.7 and [2], §4.2.

## Superdense coding

*Superdense coding* (or *dense coding*) is a protocol, proposed by Bennett and Wiesner (1992), in which Alice can transmit 2 classical bits to Bob by sending only 1 qubit, provided that they share a pre-prepared entangled pair. It is, in a precise sense, the “dual” of quantum teleportation.

**Setup.** Alice and Bob share the Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (316)$$

where Alice holds qubit 1 and Bob holds qubit 2. Alice wishes to communicate a 2-bit classical message  $m \in \{00, 01, 10, 11\}$  to Bob.

**Protocol.** Depending on the message  $m$ , Alice applies one of four unitary operations to *her* qubit (qubit 1):

| Message $m$ | Alice's operation | Resulting state  |
|-------------|-------------------|------------------|
| 00          | $\mathbb{1}$      | $ \Phi^+\rangle$ |
| 01          | $\sigma_x$        | $ \Psi^+\rangle$ |
| 10          | $\sigma_z$        | $ \Phi^-\rangle$ |
| 11          | $i\sigma_y$       | $ \Psi^-\rangle$ |

Let us verify each case explicitly. For  $m = 01$ :

$$\begin{aligned}
(\sigma_x \otimes \mathbb{1})|\Phi^+\rangle &= \frac{1}{\sqrt{2}} (\sigma_x|0\rangle \otimes |0\rangle + \sigma_x|1\rangle \otimes |1\rangle) \\
&= \frac{1}{\sqrt{2}} (|1\rangle|0\rangle + |0\rangle|1\rangle) = |\Psi^+\rangle
\end{aligned} \quad (317)$$

For  $m = 10$ :

$$\begin{aligned}
(\sigma_z \otimes \mathbb{1})|\Phi^+\rangle &= \frac{1}{\sqrt{2}} (\sigma_z|0\rangle \otimes |0\rangle + \sigma_z|1\rangle \otimes |1\rangle) \\
&= \frac{1}{\sqrt{2}} (|0\rangle|0\rangle - |1\rangle|1\rangle) = |\Phi^-\rangle
\end{aligned} \quad (318)$$

For  $m = 11$ :

$$\begin{aligned}
(i\sigma_y \otimes \mathbb{1})|\Phi^+\rangle &= \frac{1}{\sqrt{2}}(i\sigma_y|0\rangle \otimes |0\rangle + i\sigma_y|1\rangle \otimes |1\rangle) \\
&= \frac{1}{\sqrt{2}}(|1\rangle|0\rangle - |0\rangle|1\rangle) = |\Psi^-\rangle
\end{aligned} \tag{319}$$

where we used  $i\sigma_y|0\rangle = |1\rangle$  and  $i\sigma_y|1\rangle = -|0\rangle$ .

The crucial observation is that the four resulting states are precisely the four Bell states (276)–(279), which are mutually orthogonal. This is exactly the algebraic structure noted in (282).

After applying her operation, Alice sends her qubit to Bob (this is the only quantum communication in the protocol). Bob now has *both* qubits and performs a *Bell measurement* — a complete projective measurement in the Bell basis. Since the four possible states are orthogonal, the measurement outcome uniquely identifies the message  $m$  with certainty.

**Resource accounting.** Superdense coding transmits 2 classical bits using:

- 1 qubit sent from Alice to Bob, and
- 1 pre-shared Bell pair (1 ebit of entanglement), which is consumed.

**Duality with teleportation.** There is an elegant duality between superdense coding and quantum teleportation, which reflects a deep relationship between classical and quantum information resources:

| Protocol          | Input   | Resources consumed    | Output         |
|-------------------|---------|-----------------------|----------------|
| Teleportation     | 1 qubit | 1 ebit + 2 cbits      | 1 qubit at Bob |
| Superdense coding | 2 cbits | 1 ebit + 1 qubit sent | 2 cbits at Bob |

Both protocols consume exactly one Bell pair. Teleportation uses 2 classical bits to send 1 qubit; superdense coding uses 1 qubit to send 2 classical bits. This duality is a manifestation of the interconvertibility of quantum and classical resources in the presence of entanglement.

See [8], §2.3 and [2], §4.2 for further discussion.

## 4 Symmetries in quantum mechanics

A symmetry in physics is a transformation that does not change the observable results. In quantum mechanics (QM), all transformations are implemented by operators that preserve the norm of the state vector.

Let  $|\psi\rangle$  and  $|\phi\rangle$  be two states. Under a transformation implemented by the operator  $U$ :

$$|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle \tag{320}$$

$$|\phi\rangle \rightarrow |\phi'\rangle = U|\phi\rangle \tag{321}$$

The inner product must be preserved:

$$\langle\psi'|\phi'\rangle = \langle\psi|U^\dagger U|\phi\rangle = \langle\psi|\phi\rangle \tag{322}$$

This implies  $U^\dagger U = I = U U^\dagger$ , so  $U$  is unitary.

One can adopt a weaker condition where only the absolute value of the inner product is preserved:

$$|\langle \psi' | \phi' \rangle| = |\langle \psi | \phi \rangle| \quad (323)$$

This allows for anti-unitary operators  $A$ , satisfying:

$$\langle \psi | A^\dagger A | \phi \rangle = \langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle \quad (324)$$

A general anti-unitary operator can be decomposed as  $A = UK$ , where  $K$  is the complex conjugation operator and  $U$  is unitary. This is more conveniently expressed as  $(A\psi, A\phi) = (UK\psi, UK\phi) = (U\psi^*, U\phi^*) = (\psi^*, \phi^*) = (\psi, \phi)^* = (\phi, \psi)$ .

## 4.1 Unitary Operators: Translations

Let's first consider unitary operators, starting with translations.

For a translation by  $a$  in position:

$$\langle \psi | \hat{X} | \psi \rangle \rightarrow \langle \psi | \hat{X}' | \psi \rangle = \langle \psi | \hat{X} | \psi \rangle + a \quad (325)$$

The translation operator is  $e^{-i\hat{p}a/\hbar}$ , where  $\hat{p}$  is the momentum operator. This satisfies

$$[\hat{X}, e^{i\hat{p}a/\hbar}] = ae^{i\hat{p}a/\hbar} \quad (326)$$

and thus

$$e^{i\hat{p}a/\hbar} \hat{X} e^{-i\hat{p}a/\hbar} = \hat{X} + a \quad (327)$$

From the point of view of the states, it follows that the transformed state is:

$$|\psi'\rangle = e^{-i\hat{p}a/\hbar} |\psi\rangle \quad (328)$$

Thus one can either think of the transformation as acting on the state or on the operator, but not both at the same time.

## 4.2 Rotations

Now consider an infinitesimal rotation in the  $xy$ -plane by an angle  $\theta$ .

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \approx x + \theta y \\ y' &= y \cos \theta - x \sin \theta \approx y - \theta x \end{aligned} \quad (329)$$

We are looking for a unitary operator of the form  $e^{-i\theta\hat{G}/\hbar}$  that implements this transformation. It can be shown that the generator  $\hat{G}$  is the  $z$ -component of angular momentum operator  $\hat{L}_z = \hat{X}\hat{P}_y - \hat{Y}\hat{P}_x$ . Let's verify this by considering the transformation of the position operators  $\hat{X}$  and  $\hat{Y}$  under the rotation given in (329). We have

$$\begin{aligned} e^{i\theta\hat{G}/\hbar} \hat{X} e^{-i\theta\hat{G}/\hbar} &= \hat{X} + \theta \hat{Y} \\ e^{i\theta\hat{G}/\hbar} \hat{Y} e^{-i\theta\hat{G}/\hbar} &= \hat{Y} - \theta \hat{X} \end{aligned} \quad (330)$$

For infinitesimal  $\theta$ :

$$(1 + i\theta\hat{G}/\hbar)\hat{X}(1 - i\theta\hat{G}/\hbar) = \hat{X} + i(\theta/\hbar)[\hat{G}, \hat{X}] \quad (331)$$

Therefore comparing coefficients we get:

$$[\hat{G}, \hat{X}] = -i\hbar\hat{Y}, \quad [\hat{G}, \hat{Y}] = i\hbar\hat{X} \quad (332)$$

It is easy to verify that the solution for  $\hat{G}$  is  $\hat{G} = -\hat{L}_z = -\hat{X}\hat{P}_y + \hat{Y}\hat{P}_x$ .

Similarly for other planes:

- Rotation in  $xz$ -plane generated by  $\hat{L}_y = \hat{Z}\hat{P}_x - \hat{X}\hat{P}_z$

- Rotation in  $yz$ -plane generated by  $\hat{L}_x = \hat{Y}\hat{P}_z - \hat{Z}\hat{P}_y$

In general, a rotation by angle  $\phi$  around direction  $\hat{n}$  is:

$$U = e^{-i\phi\hat{n}\cdot\hat{\mathbf{L}}/\hbar} \quad (333)$$

The angular momentum operators satisfy:

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k \quad (334)$$

This implies  $\mathbf{L}$  behaves as a vector under rotation. For example:

$$e^{i\theta\hat{L}_z/\hbar}\hat{L}_xe^{-i\theta\hat{L}_z/\hbar} = \hat{L}_x\cos\theta + \hat{L}_y\sin\theta \quad (335)$$

In general, a vector operator  $\mathbf{V}$  satisfies:

$$[\hat{L}_i, V_j] = i\hbar\epsilon_{ijk}V_k \quad (336)$$

This holds, for instance, for  $\mathbf{V} = \mathbf{X}$  (position) or  $\mathbf{V} = \mathbf{P}$  (momentum).

### 4.3 Lie Groups and Lie Algebras

The underlying structure is the theory of Lie groups and Lie algebras. A Lie group is a manifold where each point is a group element (e.g., rotations). The Lie algebra is the tangent space at the identity, defined by tangents to curves through the identity (e.g., angular momentum commutation relations). To every Lie group corresponds a unique Lie algebra, but a Lie algebra can correspond to more than one group. The simply connected group is unique.

The group of 3D rotations is represented by orthogonal 3x3 matrices with determinant 1:

$$U^\dagger\mathbf{X}U = R\mathbf{X}, \quad R \in \text{SO}(3) \quad (337)$$

$\text{SO}(3)$  is called the special orthogonal group in 3 dimensions ("S" for  $\det R = 1$ ).

However,  $\text{SO}(3)$  is not simply connected. A simply connected space is one where every closed loop can be continuously contracted to a point.

The representation space of  $\text{SO}(3)$  can be visualized as a ball of radius  $\pi$ : the direction from the origin defines the axis of rotation, and the magnitude defines the angle.

Note that rotations by  $\theta$  and  $2\pi - \theta$  around opposite axes are equivalent. Thus, antipodal points on the surface are identified.

Paths on this ball may form loops that cannot be contracted due to this identification.

The Lie algebra of  $\text{SO}(3)$  is:

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$$

The simply connected group associated with this Lie algebra is  $\text{SU}(2)$ , the special unitary group in 2 dimensions, which is the double cover of  $\text{SO}(3)$ .

For spin-1/2 particles, the rotation operators are 2x2 unitary matrices:

$$U = e^{-i(\theta/2)\hat{n}\cdot\boldsymbol{\sigma}}$$

where  $\boldsymbol{\sigma}$  are the Pauli matrices, satisfying:

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$$

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$$

These commute with position  $\mathbf{X}$ , as they act on spin degrees of freedom.

For integer spin (like orbital angular momentum), the representations are odd-dimensional, while for half-integer spin, they are even-dimensional.

Note that for  $\theta = 2\pi$ :

$$e^{-i(2\pi/2)\hat{n}\cdot\boldsymbol{\sigma}} = e^{-i\pi\hat{n}\cdot\boldsymbol{\sigma}} = -\mathbb{I}$$

So the state picks up a minus sign under a  $2\pi$  rotation, and a full  $4\pi$  rotation is needed to return to the original state. This is characteristic of fermions.

In the 3x3 representation for orbital rotations,  $U = e^{-i\phi\mathbf{L}/\hbar}$ , where  $\mathbf{L}$  are infinite-dimensional differential operators acting on wave functions in Hilbert space.

## 4.4 Wigner's Theorem

We have argued that symmetry transformations must preserve the absolute value of inner products. Wigner's theorem makes this statement precise: it characterizes the most general maps on the Hilbert space that are compatible with the probabilistic structure of quantum mechanics.

**Theorem (Wigner, 1931).** *Let  $\mathcal{H}$  be a Hilbert space with  $\dim \mathcal{H} \geq 2$ . Any map  $T$  on the set of rays of  $\mathcal{H}$  that preserves transition probabilities, i.e.,*

$$|\langle T\psi | T\phi \rangle|^2 = |\langle \psi | \phi \rangle|^2 \quad \text{for all } |\psi\rangle, |\phi\rangle \in \mathcal{H} \quad (338)$$

*can be represented by an operator  $U$  on  $\mathcal{H}$  that is either unitary or anti-unitary.*

The proof of this theorem, while not exceedingly difficult, is somewhat lengthy and technical. It can be found in [9], Chapter 2, Appendix A, or in [10]. We content ourselves with understanding the content and the implications.

We already discussed unitary operators. Let us say more about anti-unitary operators. Recall that a unitary operator  $U$  is linear and satisfies  $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle$ . An anti-unitary operator  $A$  is instead *anti-linear*:

$$A(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha^* A|\psi\rangle + \beta^* A|\phi\rangle \quad (339)$$

and satisfies

$$\langle A\psi|A\phi\rangle = \langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle \quad (340)$$

Any anti-unitary operator can be decomposed as  $A = UK$ , where  $K$  is the complex conjugation operator (defined with respect to some chosen basis) and  $U$  is unitary.

The physical importance of this distinction is the following. For continuous symmetries — those connected to the identity by continuous parameter variation — the corresponding operators must be unitary, since the identity is unitary and continuity prevents a “jump” to anti-unitarity. Anti-unitary symmetries are necessarily *discrete*. The paradigmatic example is time reversal, which we discuss in §5.4.

**Connection with Quantum Information.** In the language of quantum information, Wigner’s theorem states that the physically allowed transformations of quantum states are either unitary (quantum gates) or anti-unitary. Since anti-unitary operations cannot be implemented by quantum circuits (they correspond to “complex conjugation” of the density matrix, which is not a completely positive map), the set of implementable symmetry operations coincides with the group of unitary operators. This is one of the reasons why quantum computation is formulated in terms of unitary gates. However, anti-unitary symmetries play a crucial role in constraining which quantum channels are physically realizable and in the classification of topological phases of matter.

## 4.5 Eigenvalues and eigenstates of angular momentum

In §3.2 we showed that rotations are generated by angular momentum operators satisfying the algebra

$$[\hat{J}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{J}_k \quad (341)$$

We now determine the spectrum of angular momentum operators using purely algebraic methods, in close analogy with the treatment of the harmonic oscillator in §1.4. This is one of the most elegant applications of operator algebra in quantum mechanics.

### The Casimir operator $\hat{J}^2$

Define the operator

$$\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \quad (342)$$

A direct computation using (341) shows that  $\hat{J}^2$  commutes with each component:

$$[\hat{J}^2, \hat{J}_i] = 0 \quad \text{for } i = x, y, z \quad (343)$$

To see this, consider for instance  $[\hat{J}^2, \hat{J}_z] = [\hat{J}_x^2, \hat{J}_z] + [\hat{J}_y^2, \hat{J}_z]$ . Using the identity  $[AB, C] = A[B, C] + [A, C]B$  one finds

$$[\hat{J}_x^2, \hat{J}_z] = \hat{J}_x[\hat{J}_x, \hat{J}_z] + [\hat{J}_x, \hat{J}_z]\hat{J}_x = -i\hbar(\hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x) \quad (344)$$

and similarly  $[\hat{J}_y^2, \hat{J}_z] = i\hbar(\hat{J}_x\hat{J}_y + \hat{J}_y\hat{J}_x)$ , so the two terms cancel, giving  $[\hat{J}^2, \hat{J}_z] = 0$ . The same argument applies to the other components by cyclic permutation.

Since  $\hat{J}^2$  and  $\hat{J}_z$  commute, they can be simultaneously diagonalized. We denote their common eigenstates as  $|j, m\rangle$ :

$$\hat{J}^2|j, m\rangle = \lambda|j, m\rangle, \quad \hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \quad (345)$$

where we have parametrized the  $\hat{J}_z$  eigenvalue as  $m\hbar$  with  $m$  dimensionless, and  $\lambda$  has dimensions of  $\hbar^2$ . Our goal is to determine the allowed values of  $\lambda$  and  $m$ .

## Raising and lowering operators

Define the operators

$$\hat{J}_+ \equiv \hat{J}_x + i\hat{J}_y, \quad \hat{J}_- \equiv \hat{J}_x - i\hat{J}_y \quad (346)$$

Note that  $\hat{J}_+^\dagger = \hat{J}_-$ . The commutation relations that follow from (341) are

$$[\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm, \quad [\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_z \quad (347)$$

and also  $[\hat{J}^2, \hat{J}_\pm] = 0$ , which follows from (343).

The first relation in (347) shows that  $\hat{J}_\pm$  act as raising and lowering operators for  $\hat{J}_z$ . Indeed, consider the state  $\hat{J}_+|j, m\rangle$ :

$$\hat{J}_z(\hat{J}_+|j, m\rangle) = (\hat{J}_+\hat{J}_z + \hbar\hat{J}_+)|j, m\rangle = (m+1)\hbar(\hat{J}_+|j, m\rangle) \quad (348)$$

Therefore  $\hat{J}_+|j, m\rangle$  is an eigenstate of  $\hat{J}_z$  with eigenvalue  $(m+1)\hbar$ . Since  $[\hat{J}^2, \hat{J}_+] = 0$ , this state still has the same  $\hat{J}^2$  eigenvalue  $\lambda$ . We conclude that

$$\hat{J}_+|j, m\rangle = c_+^{(m)}|j, m+1\rangle \quad (349)$$

for some constant  $c_+^{(m)}$ . Similarly,  $\hat{J}_-|j, m\rangle = c_-^{(m)}|j, m-1\rangle$ .

## Bounds on $m$ and quantization

The key observation is that  $m$  must be bounded. Since  $\hat{J}^2 - \hat{J}_z^2 = \hat{J}_x^2 + \hat{J}_y^2$  is a sum of squares of Hermitian operators, it is nonnegative:

$$\langle j, m | (\hat{J}^2 - \hat{J}_z^2) | j, m \rangle = \lambda - m^2\hbar^2 \geq 0 \quad (350)$$

Therefore  $m^2 \leq \lambda/\hbar^2$ . This means there must exist a maximum value  $m_{\max}$  and a minimum value  $m_{\min}$  of  $m$ .

Let  $m_{\max} \equiv j$ . By definition,  $\hat{J}_+$  cannot raise the  $m$ -value beyond  $j$ , so

$$\hat{J}_+|j, j\rangle = 0 \quad (351)$$

Applying  $\hat{J}_-$  to this equation from the left and using  $\hat{J}_-\hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z$  (which follows by direct computation from the definitions), we obtain

$$\hat{J}_-\hat{J}_+|j, j\rangle = (\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|j, j\rangle = (\lambda - j^2\hbar^2 - j\hbar^2)|j, j\rangle = 0 \quad (352)$$

Since  $|j, j\rangle \neq 0$ , we conclude

$$\lambda = \hbar^2 j(j+1) \quad (353)$$

Similarly, let  $m_{\min} \equiv j'$ . Then  $\hat{J}_-|j, j'\rangle = 0$  and using  $\hat{J}_+\hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar\hat{J}_z$  one finds  $\lambda = \hbar^2 j'(j'-1)$ . Combined with (353) this gives  $j(j+1) = j'(j'-1)$ , whose solutions are  $j' = -j$  or  $j' = j+1$ . Since  $j' \leq j$  by construction, we must have

$$m_{\min} = -j \quad (354)$$

Starting from  $|j, j\rangle$  and applying  $\hat{J}_-$  successively, we must reach  $|j, -j\rangle$  after an integer number of steps. This means  $j - (-j) = 2j$  must be a non-negative integer, so

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \quad (355)$$

and for each  $j$ , the magnetic quantum number  $m$  takes the  $2j+1$  values

$$m = -j, -j+1, \dots, j-1, j \quad (356)$$

**Result.** The eigenvalues of the angular momentum operators are

$$\boxed{\hat{J}^2|j, m\rangle = \hbar^2 j(j+1)|j, m\rangle, \quad \hat{J}_z|j, m\rangle = m\hbar|j, m\rangle} \quad (357)$$

with  $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$  and  $m = -j, -j+1, \dots, j$ .

The half-integer values of  $j$  are absent for orbital angular momentum (as we shall see in §4.8) but arise naturally for spin. The integer values correspond to single-valued representations of the rotation group  $\text{SO}(3)$ , while the half-integer values correspond to double-valued representations, or equivalently, to representations of  $\text{SU}(2)$  that are not representations of  $\text{SO}(3)$ .

### Matrix elements of $\hat{J}_\pm$

The normalization constants in (349) can be determined as follows. The norm of  $\hat{J}_+|j, m\rangle$  is

$$\begin{aligned} \|\hat{J}_+|j, m\rangle\|^2 &= \langle j, m|\hat{J}_-\hat{J}_+|j, m\rangle \\ &= \langle j, m|(\hat{J}^2 - \hat{J}_z^2 - \hbar\hat{J}_z)|j, m\rangle \\ &= \hbar^2[j(j+1) - m^2 - m] = \hbar^2[j(j+1) - m(m+1)] \end{aligned} \quad (358)$$

Choosing the phases by convention (the Condon-Shortley convention), we obtain

$$\boxed{\hat{J}_\pm|j, m\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}|j, m\pm 1\rangle} \quad (359)$$

Note that this expression correctly gives zero when  $m = j$  (for  $\hat{J}_+$ ) or  $m = -j$  (for  $\hat{J}_-$ ).

### Finite-dimensional matrix representations

For a given value of  $j$ , the operators  $\hat{J}_z$  and  $\hat{J}_\pm$  are represented by  $(2j+1) \times (2j+1)$  matrices in the basis  $\{|j, m\rangle\}_{m=-j}^j$ . Let us write them explicitly for the simplest cases.

$j = 1/2$ : The basis is  $\{|1/2, 1/2\rangle, |1/2, -1/2\rangle\}$ , which we also denote  $\{|\uparrow\rangle, |\downarrow\rangle\}$  or  $\{|0\rangle, |1\rangle\}$  in the computational basis of quantum information<sup>8</sup>. The angular momentum operators are  $\hat{J}_i = \frac{\hbar}{2}\sigma_i$  where  $\sigma_i$  are the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (360)$$

and  $\hat{J}^2 = \frac{3\hbar^2}{4}\mathbb{1}$ , consistent with  $j(j+1) = 3/4$ .

$j = 1$ : The basis is  $\{|1, 1\rangle, |1, 0\rangle, |1, -1\rangle\}$ , and

$$J_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad J_+ = \hbar\sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_- = \hbar\sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (361)$$

from which  $J_x = \frac{1}{2}(J_+ + J_-)$  and  $J_y = \frac{1}{2i}(J_+ - J_-)$ .

## 4.6 Spin-1/2 systems

The simplest non-trivial angular momentum is the case  $j = 1/2$ , which describes the intrinsic angular momentum (spin) of electrons, protons, neutrons, and many other elementary particles. This is also the system that constitutes a *qubit* in quantum information.

<sup>8</sup>The identification  $|0\rangle \equiv |j = 1/2, m = +1/2\rangle = |\uparrow\rangle$  and  $|1\rangle \equiv |j = 1/2, m = -1/2\rangle = |\downarrow\rangle$  is a convention widely used in quantum information. We adopt it throughout.

## Properties of Pauli matrices

The Pauli matrices (360) satisfy the following identities, already partly mentioned in the existing notes:

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i\epsilon_{ijk} \sigma_k \quad (362)$$

from which follow the commutation and anti-commutation relations

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk} \sigma_k, \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij} \mathbb{1} \quad (363)$$

Other useful properties are:  $\sigma_i^\dagger = \sigma_i$ ,  $\text{Tr}(\sigma_i) = 0$ ,  $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$ , and  $\det(\sigma_i) = -1$ .

A particularly important identity is

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = (\mathbf{a} \cdot \mathbf{b}) \mathbb{1} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}) \quad (364)$$

which is a direct consequence of (362) and holds for any two vectors  $\mathbf{a}$  and  $\mathbf{b}$  (commuting with the  $\sigma_i$ ). As a corollary,  $(\boldsymbol{\sigma} \cdot \hat{n})^2 = \mathbb{1}$  for any unit vector  $\hat{n}$ .

## Rotation operators for spin-1/2

The operator that rotates a spin-1/2 state by an angle  $\phi$  about the direction  $\hat{n}$  is

$$U(\hat{n}, \phi) = e^{-i\frac{\phi}{2}\hat{n} \cdot \boldsymbol{\sigma}} \quad (365)$$

Using the identity  $(\hat{n} \cdot \boldsymbol{\sigma})^2 = \mathbb{1}$ , the Taylor expansion of the exponential can be summed in closed form:

$$e^{-i\frac{\phi}{2}\hat{n} \cdot \boldsymbol{\sigma}} = \cos \frac{\phi}{2} \mathbb{1} - i \sin \frac{\phi}{2} \hat{n} \cdot \boldsymbol{\sigma} \quad (366)$$

This is an extremely useful formula. It shows that any  $2 \times 2$  unitary matrix with unit determinant can be written in this form for some  $\hat{n}$  and  $\phi$ , providing a concrete parametrization of  $\text{SU}(2)$ .

**Example.** Rotation by  $\phi$  about the  $z$ -axis:

$$U(\hat{z}, \phi) = \cos \frac{\phi}{2} \mathbb{1} - i \sin \frac{\phi}{2} \sigma_z = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix} \quad (367)$$

**The  $2\pi$  rotation.** Setting  $\phi = 2\pi$  in (366):

$$U(\hat{n}, 2\pi) = \cos \pi \mathbb{1} - i \sin \pi \hat{n} \cdot \boldsymbol{\sigma} = -\mathbb{1} \quad (368)$$

This means that a  $2\pi$  rotation does not return a spin-1/2 state to itself; instead, it acquires a minus sign:  $|\psi\rangle \rightarrow -|\psi\rangle$ . A full  $4\pi$  rotation is needed to return to the original state. This is the hallmark of fermions and reflects the fact that  $\text{SU}(2)$  is a double cover of  $\text{SO}(3)$ : the two elements  $+\mathbb{1}$  and  $-\mathbb{1}$  of  $\text{SU}(2)$  both correspond to the identity rotation in  $\text{SO}(3)$ .

Since overall phases are not observable for pure states (recall Axiom 1), one might think this sign is unphysical. However, it becomes observable in interference experiments where a spin-1/2 system traverses two paths, one of which involves a  $2\pi$  rotation. This has been verified experimentally with neutron interferometry [11].

## Spin-1/2 in an arbitrary direction

The eigenstate of  $\hat{n} \cdot \hat{\mathbf{S}} = \frac{\hbar}{2} \hat{n} \cdot \boldsymbol{\sigma}$  with eigenvalue  $+\hbar/2$ , for the direction  $\hat{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ , is

$$|\hat{n}, +\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \quad (369)$$

and with eigenvalue  $-\hbar/2$ :

$$|\hat{n}, -\rangle = -e^{-i\phi} \sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle \quad (370)$$

These can be obtained by rotating  $|0\rangle$  and  $|1\rangle$  using (365) with appropriate angles. They are orthonormal,  $\langle \hat{n}, + | \hat{n}, - \rangle = 0$ , and form a complete basis for the qubit Hilbert space.

## 4.7 Rotation matrices

### Euler angles

Any rotation in three dimensions can be parametrized by three Euler angles  $(\alpha, \beta, \gamma)$  through the decomposition

$$R(\alpha, \beta, \gamma) = R_z(\alpha) R_y(\beta) R_z(\gamma) \quad (371)$$

with  $0 \leq \alpha < 2\pi$ ,  $0 \leq \beta \leq \pi$ ,  $0 \leq \gamma < 2\pi$ . The corresponding quantum mechanical rotation operator in the subspace of angular momentum  $j$  is

$$\hat{D}^{(j)}(\alpha, \beta, \gamma) = e^{-i\alpha \hat{J}_z/\hbar} e^{-i\beta \hat{J}_y/\hbar} e^{-i\gamma \hat{J}_z/\hbar} \quad (372)$$

### Wigner $D$ -matrices and $d$ -matrices

The matrix elements of the rotation operator in the  $|j, m\rangle$  basis are

$$D_{m'm}^{(j)}(\alpha, \beta, \gamma) = \langle j, m' | \hat{D}^{(j)}(\alpha, \beta, \gamma) | j, m \rangle = e^{-im'\alpha} d_{m'm}^{(j)}(\beta) e^{-im\gamma} \quad (373)$$

where the *Wigner small-d matrix* is defined as

$$d_{m'm}^{(j)}(\beta) = \langle j, m' | e^{-i\beta \hat{J}_y/\hbar} | j, m \rangle \quad (374)$$

The separation in (373) occurs because  $\hat{J}_z$  is diagonal in this basis.

For  $j = 1/2$ , the  $d$ -matrix is

$$d^{(1/2)}(\beta) = \begin{pmatrix} \cos \frac{\beta}{2} & -\sin \frac{\beta}{2} \\ \sin \frac{\beta}{2} & \cos \frac{\beta}{2} \end{pmatrix} \quad (375)$$

For  $j = 1$ :

$$d^{(1)}(\beta) = \begin{pmatrix} \frac{1+\cos \beta}{2} & -\frac{\sin \beta}{\sqrt{2}} & \frac{1-\cos \beta}{2} \\ \frac{\sin \beta}{\sqrt{2}} & \cos \beta & -\frac{\sin \beta}{\sqrt{2}} \\ \frac{1-\cos \beta}{2} & \frac{\sin \beta}{\sqrt{2}} & \frac{1+\cos \beta}{2} \end{pmatrix} \quad (376)$$

A general closed-form expression for  $d_{m'm}^{(j)}(\beta)$  exists (Wigner's formula) but is rather cumbersome. For practical purposes, the explicit matrices for small  $j$  and recursion relations are more useful. The  $D$ -matrices satisfy important orthogonality relations:

$$\int d\Omega D_{m'_1 m_1}^{(j')*}(\Omega) D_{m'_2 m_2}^{(j)}(\Omega) = \frac{8\pi^2}{2j+1} \delta_{jj'} \delta_{m'_1 m'_2} \delta_{m_1 m_2} \quad (377)$$

where  $d\Omega = d\alpha \sin \beta d\beta d\gamma$  is the Haar measure on  $\text{SO}(3)$ .

## 4.8 Orbital angular momentum

We now specialize the general theory to the case of orbital angular momentum, where the angular momentum arises from the spatial degrees of freedom of a particle.

### Angular momentum in position representation

Recall from §3.2 that the orbital angular momentum operator is  $\hat{\mathbf{L}} = \hat{\mathbf{X}} \times \hat{\mathbf{P}}$ , with components

$$\hat{L}_i = \epsilon_{ijk} \hat{X}_j \hat{P}_k \quad (378)$$

In the position representation,  $\hat{P}_k = \frac{\hbar}{i} \frac{\partial}{\partial x_k}$ , and in spherical coordinates  $(r, \theta, \phi)$  the angular momentum operators take the form:

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \quad (379)$$

$$\hat{L}_{\pm} = \hbar e^{\pm i\phi} \left( \pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \quad (380)$$

$$\hat{L}^2 = -\hbar^2 \left[ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \quad (381)$$

These are purely angular operators, acting only on  $\theta$  and  $\phi$ .

### Spherical harmonics

The simultaneous eigenfunctions of  $\hat{L}^2$  and  $\hat{L}_z$  are the *spherical harmonics*  $Y_{\ell}^m(\theta, \phi)$ :

$$\hat{L}^2 Y_{\ell}^m = \hbar^2 \ell(\ell+1) Y_{\ell}^m, \quad \hat{L}_z Y_{\ell}^m = m\hbar Y_{\ell}^m \quad (382)$$

The requirement that  $Y_{\ell}^m(\theta, \phi)$  be single-valued under  $\phi \rightarrow \phi + 2\pi$  imposes  $e^{im(\phi+2\pi)} = e^{im\phi}$ , which forces  $m$  to be an integer. Since  $m$  ranges from  $-\ell$  to  $\ell$  in integer steps,  $\ell$  must also be a non-negative integer:

$$\ell = 0, 1, 2, 3, \dots; \quad m = -\ell, -\ell+1, \dots, \ell \quad (383)$$

This eliminates the half-integer values of  $j$  that are allowed by the abstract algebra. The physical reason is clear: orbital angular momentum is associated with a wave function  $\psi(\mathbf{x})$  defined on ordinary (simply connected) space, which requires single-valuedness. In contrast, spin has no classical spatial analog and can take half-integer values.

The spherical harmonics have the explicit form

$$Y_{\ell}^m(\theta, \phi) = (-1)^m \sqrt{\frac{(2\ell+1)(\ell-m)!}{4\pi(\ell+m)!}} P_{\ell}^m(\cos \theta) e^{im\phi} \quad (384)$$

where  $P_{\ell}^m$  are the associated Legendre polynomials, defined for  $m \geq 0$  by

$$P_{\ell}^m(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_{\ell}(x) \quad (385)$$

and  $P_{\ell}(x) = \frac{1}{2^{\ell} \ell!} \frac{d^{\ell}}{dx^{\ell}} (x^2-1)^{\ell}$  are the Legendre polynomials (Rodrigues' formula). The phase  $(-1)^m$  in (384) is the Condon-Shortley convention, consistent with the phase choice made in (359).

The first few spherical harmonics are:

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}, \quad Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y_1^{\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \quad (386)$$

They satisfy the orthonormality condition

$$\int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta Y_{\ell'}^{m'*}(\theta, \phi) Y_\ell^m(\theta, \phi) = \delta_{\ell\ell'} \delta_{mm'} \quad (387)$$

and form a complete basis for square-integrable functions on the sphere: any  $f(\theta, \phi) \in L^2(S^2)$  can be expanded as  $f(\theta, \phi) = \sum_{\ell=0}^\infty \sum_{m=-\ell}^\ell c_{\ell m} Y_\ell^m(\theta, \phi)$ .

## The hydrogen atom: separation of variables

The time-independent Schrödinger equation for a particle in a central potential  $V(r)$  is

$$\left[ -\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right] \psi(\mathbf{x}) = E \psi(\mathbf{x}) \quad (388)$$

where  $\mu$  is the reduced mass. Writing the Laplacian in spherical coordinates,

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) - \frac{\hat{L}^2}{\hbar^2 r^2} \quad (389)$$

we see that  $\hat{L}^2$  appears naturally. Since  $[H, \hat{L}^2] = 0$  and  $[H, \hat{L}_z] = 0$  for central potentials, the energy eigenstates can be chosen as simultaneous eigenstates of  $H$ ,  $\hat{L}^2$ , and  $\hat{L}_z$ :

$$\psi_{n\ell m}(r, \theta, \phi) = R_{n\ell}(r) Y_\ell^m(\theta, \phi) \quad (390)$$

Substituting into (388) yields the radial equation

$$-\frac{\hbar^2}{2\mu} \left[ \frac{d^2 u_{n\ell}}{dr^2} - \frac{\ell(\ell+1)}{r^2} u_{n\ell} \right] + V(r) u_{n\ell} = E u_{n\ell} \quad (391)$$

where  $u_{n\ell}(r) \equiv r R_{n\ell}(r)$ . The term  $\frac{\hbar^2 \ell(\ell+1)}{2\mu r^2}$  is the centrifugal barrier, a quantum mechanical version of the classical effective potential arising from angular momentum conservation.

## 5 Addition of angular momenta and discrete symmetries

We now turn to two central topics that complete the theory of symmetries in quantum mechanics: the problem of combining angular momenta (relevant whenever a system has more than one source of angular momentum), and the discrete symmetries of parity and time reversal.

### 5.1 Addition of angular momenta

Consider a system composed of two subsystems with angular momenta  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$ . These could be two spins, a spin and an orbital angular momentum, or two orbital angular momenta. The total angular momentum of the composite system is

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2 \quad (392)$$

Since  $\hat{\mathbf{J}}_1$  and  $\hat{\mathbf{J}}_2$  act on different Hilbert spaces ( $\mathcal{H}_1$  and  $\mathcal{H}_2$ ), their components commute:

$$[\hat{J}_{1i}, \hat{J}_{2j}] = 0 \quad \text{for all } i, j \quad (393)$$

It is straightforward to verify from (341) and (393) that  $\hat{\mathbf{J}}$  also satisfies the angular momentum algebra:  $[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k$ .

The Hilbert space of the composite system is the tensor product  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ . Two natural bases exist for this space:

**Uncoupled basis:**  $\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\} \equiv \{|j_1, m_1; j_2, m_2\rangle\}$ , which diagonalizes  $\{\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}\}$ . This basis has dimension  $(2j_1 + 1)(2j_2 + 1)$ .

**Coupled basis:**  $\{|j_1, j_2; J, M\rangle\}$ , which diagonalizes  $\{\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z\}$ . Here  $J$  is the quantum number associated with the total angular momentum  $\hat{J}^2$ , and  $M$  the one for  $\hat{J}_z$ .

Both bases span the same space and are related by a unitary transformation. The components of this transformation are the *Clebsch-Gordan coefficients*.

### Allowed values of $J$

Since  $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$ , we have  $M = m_1 + m_2$ . The maximum value of  $M$  is  $j_1 + j_2$  (when  $m_1 = j_1$  and  $m_2 = j_2$ ), and there is exactly one such state. Therefore, the maximum value of  $J$  is  $j_1 + j_2$ .

The next value,  $M = j_1 + j_2 - 1$ , can be achieved in two ways:  $(m_1 = j_1, m_2 = j_2 - 1)$  and  $(m_1 = j_1 - 1, m_2 = j_2)$ . One linear combination belongs to  $J = j_1 + j_2$  (obtained by applying  $\hat{J}_-$  to the state  $|J = j_1 + j_2, M = j_1 + j_2\rangle$ ); the orthogonal combination must be the highest state of  $J = j_1 + j_2 - 1$ .

Continuing this argument (a counting exercise: verify that both bases have the same total number of states), one finds that  $J$  takes all values

$$J = |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 - 1, j_1 + j_2 \quad (394)$$

This is the *triangle rule*:  $J$  ranges from  $|j_1 - j_2|$  to  $j_1 + j_2$  in integer steps. The total number of states is

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} (2J+1) = (2j_1+1)(2j_2+1) \quad (395)$$

which confirms the dimension count.

### Clebsch-Gordan coefficients

The change of basis between the coupled and uncoupled representations is

$$|j_1, j_2; J, M\rangle = \sum_{m_1, m_2} \langle j_1, m_1; j_2, m_2 | J, M \rangle |j_1, m_1; j_2, m_2\rangle \quad (396)$$

where the sum runs over all  $m_1, m_2$  with  $m_1 + m_2 = M$ . The coefficients  $\langle j_1, m_1; j_2, m_2 | J, M \rangle$  are the *Clebsch-Gordan (CG) coefficients*. They are real by the Condon-Shortley phase convention.

The CG coefficients can be computed by the following recursive procedure:

1. Start with the “stretched state”:  $|J = j_1 + j_2, M = j_1 + j_2\rangle = |j_1, j_1\rangle |j_2, j_2\rangle$ , so the CG coefficient is 1.
2. Apply  $\hat{J}_- = \hat{J}_{1-} + \hat{J}_{2-}$  to generate states with lower  $M$  within the same  $J$ .
3. For each new value of  $M$ , construct the state with the next lower  $J$  by requiring orthogonality to the states already found.

Let us work out the most important example in detail.

**Example: addition of two spin-1/2**

Consider  $j_1 = j_2 = 1/2$ . The uncoupled basis has four states:  $|+, +\rangle$ ,  $|+, -\rangle$ ,  $|-, +\rangle$ ,  $|-, -\rangle$ , where we abbreviate  $|\pm\rangle \equiv |1/2, \pm 1/2\rangle$ .

The total angular momentum takes values  $J = 0$  and  $J = 1$ . The  $J = 1$  subspace is three-dimensional (the *triplet*) and  $J = 0$  is one-dimensional (the *singlet*).

**Triplet ( $J = 1$ ):**

Start with  $|J = 1, M = 1\rangle = |+, +\rangle$ . Apply  $\hat{J}_- = \hat{J}_{1-} + \hat{J}_{2-}$ :

$$\hat{J}_-|1, 1\rangle = \hbar\sqrt{2}|1, 0\rangle \quad (397)$$

On the right-hand side:

$$(\hat{J}_{1-} + \hat{J}_{2-})|+, +\rangle = \hbar(|-, +\rangle + |+, -\rangle) \quad (398)$$

Therefore

$$|1, 0\rangle = \frac{1}{\sqrt{2}}(|+, -\rangle + |-, +\rangle) \quad (399)$$

Applying  $\hat{J}_-$  once more gives  $|1, -1\rangle = |-, -\rangle$ .

**Singlet ( $J = 0$ ):**

The state  $|0, 0\rangle$  must have  $M = 0$  and be orthogonal to (399):

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(|+, -\rangle - |-, +\rangle) \quad (400)$$

In summary, the decomposition of the tensor product is

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0 \quad (401)$$

or, in terms of dimensions,  $\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1}$ .

**Connection with Quantum Information.** The singlet state (400) and the three triplet states form the *Bell basis*:

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|0, 0\rangle + |1, 1\rangle) = |1, 1\rangle + |1, -1\rangle / \sqrt{2} \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|0, 0\rangle - |1, 1\rangle) \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|0, 1\rangle + |1, 0\rangle) = |1, 0\rangle \\ |\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|0, 1\rangle - |1, 0\rangle) = |0, 0\rangle \end{aligned} \quad (402)$$

(with appropriate relabeling of the computational basis). The Bell states are maximally entangled and play a fundamental role in quantum teleportation, superdense coding, and entanglement-based quantum key distribution. The decomposition (401) has a direct information-theoretic interpretation: the entangled Hilbert space of two qubits decomposes into subspaces labeled by total angular momentum, and the singlet subspace is particularly distinguished by its invariance under *collective* rotations  $U \otimes U$  (it is a “decoherence-free” subspace).

**General example: coupling  $j_1 = 1$  and  $j_2 = 1/2$**

This is the physically important case of spin-orbit coupling. The triangle rule gives  $J = 1/2$  and  $J = 3/2$ :  $\mathbf{3} \otimes \mathbf{2} = \mathbf{4} \oplus \mathbf{2}$ .

The CG coefficients for this case can be computed by the recursive method described above. For instance, the stretched state is  $|3/2, 3/2\rangle = |1, 1\rangle|1/2, 1/2\rangle$ . Applying  $\hat{J}_-$ :

$$|3/2, 1/2\rangle = \sqrt{\frac{2}{3}}|1, 0\rangle|1/2, 1/2\rangle + \sqrt{\frac{1}{3}}|1, 1\rangle|1/2, -1/2\rangle \quad (403)$$

The orthogonal combination gives

$$|1/2, 1/2\rangle = \sqrt{\frac{1}{3}}|1, 0\rangle|1/2, 1/2\rangle - \sqrt{\frac{2}{3}}|1, 1\rangle|1/2, -1/2\rangle \quad (404)$$

These coefficients are tabulated in standard references (e.g., [1], Appendix C).

## 5.2 The Wigner-Eckart theorem

The Wigner-Eckart theorem is one of the most powerful results in the theory of angular momentum. It allows one to extract the dependence of matrix elements on the magnetic quantum numbers  $m, m'$  without performing any explicit integral, relying only on the rotational symmetry of the system.

### Spherical tensor operators

We first need the concept of a *spherical tensor operator*. A set of  $2k + 1$  operators  $\hat{T}_q^{(k)}$  (with  $q = -k, -k + 1, \dots, k$ ) is called a spherical tensor operator of rank  $k$  if it transforms under rotations as

$$\hat{D}^\dagger(\mathcal{R}) \hat{T}_q^{(k)} \hat{D}(\mathcal{R}) = \sum_{q'} D_{q'q}^{(k)}(\mathcal{R}) \hat{T}_{q'}^{(k)} \quad (405)$$

Equivalently, in terms of commutation relations with  $\hat{J}_z$  and  $\hat{J}_\pm$ :

$$[\hat{J}_z, \hat{T}_q^{(k)}] = q\hbar \hat{T}_q^{(k)} \quad (406)$$

$$[\hat{J}_\pm, \hat{T}_q^{(k)}] = \hbar\sqrt{k(k+1) - q(q \pm 1)} \hat{T}_{q \pm 1}^{(k)} \quad (407)$$

Note the structural similarity with the transformation of angular momentum eigenstates — the tensor components  $\hat{T}_q^{(k)}$  transform among themselves exactly as the states  $|k, q\rangle$  do.

### Examples.

- A scalar operator ( $k = 0$ ):  $\hat{T}_0^{(0)} = \hat{S}$ , which commutes with all  $\hat{J}_i$ . Examples:  $\hat{H}$  (for a rotationally invariant system),  $\hat{\mathbf{J}}^2$ .
- A vector operator ( $k = 1$ ): from a vector operator  $\hat{\mathbf{V}} = (\hat{V}_x, \hat{V}_y, \hat{V}_z)$  satisfying  $[\hat{J}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}\hat{V}_k$ , the spherical components are

$$\hat{T}_0^{(1)} = \hat{V}_z, \quad \hat{T}_{\pm 1}^{(1)} = \mp \frac{1}{\sqrt{2}}(\hat{V}_x \pm i\hat{V}_y) \quad (408)$$

Examples:  $\hat{\mathbf{X}}, \hat{\mathbf{P}}, \hat{\mathbf{L}}, \hat{\mathbf{S}}$ .

## Statement of the theorem

**Theorem (Wigner-Eckart).** *The matrix element of a spherical tensor operator  $\hat{T}_q^{(k)}$  between angular momentum eigenstates is*

$$\langle j', m' | \hat{T}_q^{(k)} | j, m \rangle = \langle j, m; k, q | j', m' \rangle \frac{\langle j' || \hat{T}^{(k)} || j \rangle}{\sqrt{2j' + 1}} \quad (409)$$

where  $\langle j, m; k, q | j', m' \rangle$  is a Clebsch-Gordan coefficient and  $\langle j' || \hat{T}^{(k)} || j \rangle$  is the **reduced matrix element**, which is independent of  $m, m'$ , and  $q$ .

The power of this result lies in the factorization: all the  $m$ -dependence is contained in the Clebsch-Gordan coefficient (which is known, tabulated, or easily computed), while the dynamical information specific to the operator and the system is encoded entirely in the reduced matrix element. This has immediate consequences:

1. **Selection rules:** The matrix element vanishes unless the CG coefficient is nonzero, which requires  $m' = m + q$  and  $|j - k| \leq j' \leq j + k$  (triangle rule). For vector operators ( $k = 1$ ), this gives  $\Delta m = 0, \pm 1$  and  $\Delta j = 0, \pm 1$  (excluding  $j = 0 \rightarrow j' = 0$ ).
2. **Proportionality:** All matrix elements of a given rank- $k$  tensor between fixed  $j$  and  $j'$  are proportional to each other, with proportionality constants given by CG coefficients. This means that to compute all  $(2j + 1)(2j' + 1)(2k + 1)$  matrix elements, it suffices to compute *one* reduced matrix element.
3. **Ratio relations:** Ratios of matrix elements of different operators of the same rank, between the same states, equal ratios of their reduced matrix elements — and are thus independent of  $m, m', q$ .

*Proof sketch.* The key idea is that the state  $\hat{T}_q^{(k)} | j, m \rangle$  transforms under rotations in the same way as  $| k, q \rangle \otimes | j, m \rangle$ . More precisely, equations (406) and (407) imply that the states  $\hat{T}_q^{(k)} | j, m \rangle$  span a subspace of  $\mathcal{H}$  that carries the representation  $D^{(k)} \otimes D^{(j)}$  of  $SU(2)$ . Projecting onto the irreducible component  $D^{(j')}$  of this tensor product, the overlap is precisely a CG coefficient, up to a factor that depends only on  $j$  and  $j'$ . This factor is the reduced matrix element. A complete proof can be found in [1], Chapter 3, §3.11.  $\square$

## 5.3 Parity

We now turn to the discrete symmetry of parity (spatial inversion). Unlike rotations, parity is not connected to the identity and cannot be generated by angular momentum. It provides independent quantum numbers and selection rules.

### The parity operator

The parity transformation inverts the spatial coordinates:  $\mathbf{x} \rightarrow -\mathbf{x}$ . The corresponding quantum operator  $\hat{\Pi}$  acts on position eigenstates as

$$\hat{\Pi} | \mathbf{x} \rangle = | -\mathbf{x} \rangle \quad (410)$$

It follows immediately that:

$$\hat{\Pi} \hat{\mathbf{X}} \hat{\Pi}^\dagger = -\hat{\mathbf{X}} \quad (411)$$

$$\hat{\Pi} \hat{\mathbf{P}} \hat{\Pi}^\dagger = -\hat{\mathbf{P}} \quad (412)$$

$$\hat{\Pi} \hat{\mathbf{L}} \hat{\Pi}^\dagger = +\hat{\mathbf{L}} \quad (413)$$

The last relation follows from  $\hat{\mathbf{L}} = \hat{\mathbf{X}} \times \hat{\mathbf{P}}$ , so both factors change sign and the product is invariant. Quantities that change sign under parity (like  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{P}}$ ) are called *polar vectors*; those that are invariant (like  $\hat{\mathbf{L}}$ ) are called *axial vectors* or *pseudovectors*.

### Properties of $\hat{\Pi}$

Since applying parity twice returns to the original configuration:

$$\hat{\Pi}^2 = \mathbb{1} \quad (414)$$

This means  $\hat{\Pi}$  is both Hermitian and unitary, and its eigenvalues are  $\pm 1$ . A state  $|\psi\rangle$  with definite parity satisfies  $\hat{\Pi}|\psi\rangle = \pm|\psi\rangle$ ; we say the state has *even parity* (+1) or *odd parity* (−1).

### Parity of angular momentum eigenstates

For the spherical harmonics, the parity transformation  $\mathbf{x} \rightarrow -\mathbf{x}$  corresponds to  $(\theta, \phi) \rightarrow (\pi - \theta, \phi + \pi)$ . Under this transformation:

$$Y_\ell^m(\pi - \theta, \phi + \pi) = (-1)^\ell Y_\ell^m(\theta, \phi) \quad (415)$$

Therefore, orbital angular momentum eigenstates have definite parity  $(-1)^\ell$ .

### Selection rules from parity

If the Hamiltonian commutes with parity,  $[\hat{H}, \hat{\Pi}] = 0$ , then the energy eigenstates can be chosen to have definite parity. This leads to important selection rules:

Consider a matrix element  $\langle\psi'|\hat{O}|\psi\rangle$  where  $|\psi\rangle$  and  $|\psi'\rangle$  have parities  $\pi$  and  $\pi'$ , and  $\hat{O}$  has a definite transformation under parity:  $\hat{\Pi}\hat{O}\hat{\Pi}^\dagger = \pi_O \hat{O}$ . Then

$$\langle\psi'|\hat{O}|\psi\rangle = \langle\psi'|\hat{\Pi}^\dagger\hat{\Pi}\hat{O}\hat{\Pi}^\dagger\hat{\Pi}|\psi\rangle = \pi' \pi_O \pi \langle\psi'|\hat{O}|\psi\rangle \quad (416)$$

This implies

$$\langle\psi'|\hat{O}|\psi\rangle = 0 \quad \text{unless} \quad \pi' \pi_O \pi = 1 \quad (417)$$

**Example: electric dipole transitions.** The electric dipole operator is  $\hat{\mathbf{d}} = q\hat{\mathbf{X}}$ , which is odd under parity ( $\pi_O = -1$ ). Therefore, electric dipole transitions can only occur between states of opposite parity:  $\pi' = -\pi$ , or equivalently  $\Delta\ell = \pm 1$  (odd). Combined with the Wigner-Eckart selection rule ( $\Delta\ell = 0, \pm 1$  for a rank-1 tensor), this gives the familiar  $\Delta\ell = \pm 1$  selection rule for electric dipole radiation.

### Intrinsic parity

Elementary particles can carry an *intrinsic parity* in addition to the spatial parity  $(-1)^\ell$ . The total parity of a state is then  $(-1)^\ell$  times the intrinsic parity. For instance, the photon has intrinsic parity  $-1$  (it is a “vector particle”), while the pion has intrinsic parity  $-1$  as well (it is a pseudoscalar). Protons and neutrons have intrinsic parity  $+1$ . These assignments are determined experimentally and play a crucial role in particle physics.

## 5.4 Time reversal

Time reversal is the most subtle of the discrete symmetries. It is the transformation  $t \rightarrow -t$ , and Wigner’s theorem tells us it must be implemented by either a unitary or anti-unitary operator. We shall see that it must be *anti-unitary*.

## Why anti-unitary?

Consider the Schrödinger equation for a free particle. If  $|\psi(t)\rangle$  is a solution, then under time reversal we expect  $|\psi(-t)\rangle$  to also be a solution. The Schrödinger equation is  $i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle$ . Replacing  $t \rightarrow -t$ :

$$-i\hbar\frac{\partial|\psi(-t)\rangle}{\partial(-t)} = \hat{H}|\psi(-t)\rangle \implies i\hbar\frac{\partial|\psi(-t)\rangle}{\partial t} = \hat{H}|\psi(-t)\rangle \quad (418)$$

This is the same Schrödinger equation, so time reversal of a solution is automatically a solution for the free particle. However, this used  $\hat{H} = \hat{P}^2/2m$ , which is real in the position representation.

For a general Hamiltonian, the time-reversed state must involve complex conjugation to compensate for the factor of  $i$  in the Schrödinger equation. If the time reversal operator were unitary ( $\hat{T} = U$ , linear), then acting on  $i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle$  would give  $i\hbar\partial_t(U|\psi\rangle) = U\hat{H}|\psi\rangle$ . For this to look like a valid Schrödinger equation with  $\hat{H}$  (not  $-\hat{H}$ ), we would need  $U\hat{H} = \hat{H}U$ , but that would not account for the sign change  $t \rightarrow -t$ . Instead, if  $\hat{T}$  is anti-unitary, the factor of  $i$  gets conjugated, providing the necessary sign flip.

More precisely, let  $\hat{\Theta}$  denote the time reversal operator. We require:

$$\hat{\Theta} i \hat{\Theta}^{-1} = -i \quad (419)$$

This is exactly the property of anti-linear (and hence anti-unitary) operators.

## Action on fundamental operators

Classically, time reversal reverses momenta and angular momenta but leaves positions unchanged:

$$\hat{\Theta}\hat{\mathbf{X}}\hat{\Theta}^{-1} = \hat{\mathbf{X}}, \quad \hat{\Theta}\hat{\mathbf{P}}\hat{\Theta}^{-1} = -\hat{\mathbf{P}}, \quad \hat{\Theta}\hat{\mathbf{J}}\hat{\Theta}^{-1} = -\hat{\mathbf{J}} \quad (420)$$

The last relation applies to *all* angular momenta (orbital and spin). One can verify consistency:  $[\hat{X}_i, \hat{P}_j] = i\hbar\delta_{ij}$ . Applying  $\hat{\Theta}$ :  $\hat{\Theta}[\hat{X}_i, \hat{P}_j]\hat{\Theta}^{-1} = [\hat{X}_i, -\hat{P}_j] = -i\hbar\delta_{ij}$ , and  $\hat{\Theta}(i\hbar\delta_{ij})\hat{\Theta}^{-1} = -i\hbar\delta_{ij}$  (since  $\hat{\Theta}$  is anti-linear). Both sides agree.

## Time reversal for spin-1/2

For a spin-1/2 system,  $\hat{\Theta}$  must satisfy  $\hat{\Theta}\hat{\mathbf{S}}\hat{\Theta}^{-1} = -\hat{\mathbf{S}}$ . Writing  $\hat{\Theta} = UK$  (with  $U$  unitary and  $K$  complex conjugation in the  $\{|0\rangle, |1\rangle\}$  basis), and using  $K\sigma_x K = \sigma_x$ ,  $K\sigma_y K = -\sigma_y$ ,  $K\sigma_z K = \sigma_z$ , we need:

$$U\sigma_x U^\dagger = -\sigma_x, \quad U(-\sigma_y)U^\dagger = -(-\sigma_y) = \sigma_y, \quad U\sigma_z U^\dagger = -\sigma_z \quad (421)$$

The middle equation gives  $U\sigma_y U^\dagger = -\sigma_y$ . All three conditions are satisfied by  $U = e^{i\delta}(-i\sigma_y)$  for arbitrary phase  $\delta$ . Choosing  $\delta = 0$ :

$$\hat{\Theta} = -i\sigma_y K = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} K \quad (422)$$

One can verify that  $\hat{\Theta}^2 = -\mathbb{1}$  for spin-1/2. This is a general result: for half-integer spin,  $\hat{\Theta}^2 = -\mathbb{1}$ ; for integer spin,  $\hat{\Theta}^2 = +\mathbb{1}$ .

## Kramers' theorem

The fact that  $\hat{\Theta}^2 = -\mathbb{1}$  for half-integer spin systems has a profound consequence.

**Theorem (Kramers, 1930).** *For a system with half-integer total angular momentum and a time-reversal invariant Hamiltonian ( $[\hat{H}, \hat{\Theta}] = 0$ ), every energy eigenvalue is at least doubly degenerate.*

*Proof.* Let  $|E\rangle$  be an energy eigenstate. Since  $[\hat{H}, \hat{\Theta}] = 0$ , the state  $\hat{\Theta}|E\rangle$  is also an eigenstate with the same energy  $E$ . We claim that  $|E\rangle$  and  $\hat{\Theta}|E\rangle$  are linearly independent (hence the degeneracy is at least two-fold).

Suppose, for contradiction, that  $\hat{\Theta}|E\rangle = c|E\rangle$  for some  $c \in \mathbb{C}$ . Apply  $\hat{\Theta}$  again:

$$\hat{\Theta}^2|E\rangle = \hat{\Theta}(c|E\rangle) = c^*\hat{\Theta}|E\rangle = c^*c|E\rangle = |c|^2|E\rangle \quad (423)$$

But  $\hat{\Theta}^2 = -1$ , so  $|c|^2 = -1$ , which is impossible. Therefore  $\hat{\Theta}|E\rangle$  cannot be proportional to  $|E\rangle$ .  $\square$

The pair  $\{|E\rangle, \hat{\Theta}|E\rangle\}$  is called a *Kramers pair*. Kramers' degeneracy can only be lifted by breaking time-reversal symmetry, for instance by applying an external magnetic field. This is the fundamental reason why the energy levels of atoms with an odd number of electrons remain at least doubly degenerate even in the presence of crystal fields or other perturbations that reduce spatial symmetry, as long as no magnetic field is applied.

Note that for integer spin,  $\hat{\Theta}^2 = +1$ , and the proof fails:  $|c|^2 = 1$  is perfectly consistent, so there is no guaranteed degeneracy.

## 5.5 Symmetries and quantum information

We close this chapter by highlighting the deep connections between symmetries and quantum information theory, which have become increasingly important in modern physics.

### Quantum gates as symmetry operations

The standard single-qubit gates of quantum computation are intimately related to the symmetry operations we have studied:

- The Pauli gates  $X, Y, Z$  are (up to global phases) rotations by  $\pi$  about the  $x, y, z$  axes of the Bloch sphere:  $X = -i e^{-i\pi\sigma_x/2} = \sigma_x$ , etc.
- The Hadamard gate  $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$  is a rotation by  $\pi$  about the axis  $(\hat{x} + \hat{z})/\sqrt{2}$ .
- The phase gate  $S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$  and the  $T$  gate  $= \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$  are rotations about  $\hat{z}$  by angles  $\pi/2$  and  $\pi/4$ , respectively.

More generally, any single-qubit gate is an element of  $\text{SU}(2)$  (up to a global phase), which is the rotation group of the Bloch sphere. The statement that a finite set of gates is *universal* — meaning any unitary can be approximated to arbitrary precision — is related to the density of the subgroup they generate in  $\text{SU}(2)$ .

### Entanglement and symmetry representations

The decomposition of tensor product representations (addition of angular momenta) has a natural quantum information interpretation. The Hilbert space of  $n$  qubits carries a representation of  $\text{SU}(2)^{\otimes n}$  (independent rotations of each qubit) and also a representation of the collective  $\text{SU}(2)$  (simultaneous rotation of all qubits). States that are invariant under the collective  $\text{SU}(2)$  — such as the singlet (400) — are decoherence-free subspaces for collective noise, which is valuable for quantum error correction.

## Time reversal and the transpose

In the language of quantum channels (Chapter 2), time reversal is closely related to the *transpose* operation on density matrices. Specifically, in a basis where  $\hat{\Theta} = UK$ , the action of  $\hat{\Theta}$  on a density matrix is  $\rho \rightarrow U\rho^T U^\dagger$  (where  $\rho^T$  is the transpose). The partial transpose — transposing only one subsystem of a bipartite state — is the basis of the Peres-Horodecki criterion for detecting entanglement [12, 13]: a state  $\rho_{AB}$  is entangled if its partial transpose  $\rho_{AB}^{TB}$  has at least one negative eigenvalue. This provides a concrete link between the symmetry of time reversal and the theory of entanglement.

## 6 Approximation methods

Exact solutions of the Schrödinger equation are available only for a handful of systems: the harmonic oscillator, the hydrogen atom, a few other idealized potentials. Any realistic physical system — an atom in an external field, a molecule, a many-body system — requires methods that go beyond exact solutions. In this chapter we develop the main approximation techniques of quantum mechanics: perturbation theory (both time-independent and time-dependent) and the variational method.

The philosophy is simple. One starts from a problem whose solution is known exactly (the “unperturbed” system) and treats the additional complication as a small correction. The mathematical framework then provides systematic expansions for the energies, states, and transition rates of the full system. The question of what counts as “small” will be made precise in each case.

Main references for this chapter are [1], Chapters 5 and 6; [14], Volume 2, Chapters XI–XIII; and [Ballentine], Chapters 11–12.

### 6.1 Time-independent perturbation theory: non-degenerate case

#### Setup

Consider a Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} \quad (424)$$

where  $\hat{H}_0$  is the unperturbed Hamiltonian whose eigenvalues and eigenstates are known:

$$\hat{H}_0 |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle \quad (425)$$

and  $\hat{V}$  is the perturbation. The real parameter  $\lambda$  controls the strength of the perturbation; at the end of the calculation one typically sets  $\lambda = 1$ , but keeping it explicit helps organize the expansion. The superscript (0) denotes unperturbed quantities.

The assumption of non-degeneracy means that all eigenvalues  $E_n^{(0)}$  are distinct. The degenerate case requires a separate treatment (see §6.2).

We seek the eigenvalues and eigenstates of the full Hamiltonian:

$$\hat{H} |n\rangle = E_n |n\rangle \quad (426)$$

as power series in  $\lambda$ :

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (427)$$

$$|n\rangle = |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots \quad (428)$$

## First-order corrections

Substituting (427)–(428) into (426) and collecting terms of order  $\lambda^1$ :

$$\hat{H}_0|n^{(1)}\rangle + \hat{V}|n^{(0)}\rangle = E_n^{(0)}|n^{(1)}\rangle + E_n^{(1)}|n^{(0)}\rangle \quad (429)$$

Taking the inner product with  $\langle n^{(0)}|$ :

$$\langle n^{(0)}|\hat{H}_0|n^{(1)}\rangle + \langle n^{(0)}|\hat{V}|n^{(0)}\rangle = E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle + E_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle \quad (430)$$

Since  $\hat{H}_0$  is Hermitian, the first term on the left equals  $E_n^{(0)}\langle n^{(0)}|n^{(1)}\rangle$ , which cancels with the first term on the right. Using  $\langle n^{(0)}|n^{(0)}\rangle = 1$ , we obtain

$$\boxed{E_n^{(1)} = \langle n^{(0)}|\hat{V}|n^{(0)}\rangle \equiv V_{nn}} \quad (431)$$

The first-order energy correction is simply the expectation value of the perturbation in the unperturbed state. This is a physically intuitive result: to first approximation, the energy shift is the “average” of the perturbation as seen by the unperturbed state.

For the first-order state correction, take the inner product of (429) with  $\langle m^{(0)}|$  for  $m \neq n$ :

$$E_m^{(0)}\langle m^{(0)}|n^{(1)}\rangle + V_{mn} = E_n^{(0)}\langle m^{(0)}|n^{(1)}\rangle + E_n^{(1)}\underbrace{\langle m^{(0)}|n^{(0)}\rangle}_{=0} \quad (432)$$

where  $V_{mn} = \langle m^{(0)}|\hat{V}|n^{(0)}\rangle$ . Solving:

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \quad (433)$$

and therefore

$$\boxed{|n^{(1)}\rangle = \sum_{m \neq n} \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle} \quad (434)$$

This expression makes the non-degeneracy assumption explicit: the denominator  $E_n^{(0)} - E_m^{(0)}$  must be nonzero for all  $m \neq n$ . If two unperturbed levels are close in energy, the correction becomes large and the perturbative expansion may break down. This is the signal that degenerate perturbation theory is needed.

We adopt the convention  $\langle n^{(0)}|n^{(1)}\rangle = 0$ , which amounts to choosing the normalization  $\langle n|n\rangle = 1$  order by order in  $\lambda$ . This is always possible and simplifies the formulas.

## Second-order energy correction

Collecting terms of order  $\lambda^2$  in the eigenvalue equation:

$$\hat{H}_0|n^{(2)}\rangle + \hat{V}|n^{(1)}\rangle = E_n^{(0)}|n^{(2)}\rangle + E_n^{(1)}|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle \quad (435)$$

Taking the inner product with  $\langle n^{(0)}|$  and using  $\langle n^{(0)}|n^{(1)}\rangle = 0$ :

$$\langle n^{(0)}|\hat{V}|n^{(1)}\rangle = E_n^{(2)} \quad (436)$$

Substituting (434):

$$\boxed{E_n^{(2)} = \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}}} \quad (437)$$

This formula has important physical content:

- The second-order correction involves all other states  $m \neq n$ , weighted by the square of the matrix element  $|V_{mn}|^2$  (the “coupling strength”) divided by the energy difference (the “energy cost” of virtual transitions).
- For the *ground state* ( $n = 0$ ), all denominators  $E_0^{(0)} - E_m^{(0)}$  are negative (since  $E_0^{(0)}$  is the lowest energy), so  $E_0^{(2)} \leq 0$ . That is, the second-order correction always lowers the ground state energy. This makes physical sense: the perturbation allows the system to explore configurations of lower energy through virtual transitions.
- States that are closer in energy contribute more to the sum, confirming that near-degenerate levels are the most strongly affected by perturbations.

### Validity of the perturbative expansion

The expansion is useful only if the corrections are small compared to the unperturbed quantities. A rough criterion for the validity of first-order perturbation theory is

$$\left| \frac{V_{mn}}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1 \quad \text{for all } m \neq n \quad (438)$$

When this condition is violated for some particular  $m$ , it typically indicates that the states  $|n^{(0)}\rangle$  and  $|m^{(0)}\rangle$  are strongly mixed by the perturbation, and the problem should be reformulated — either by treating the near-degenerate levels together (degenerate perturbation theory) or by choosing a better unperturbed Hamiltonian.

### Example: anharmonic oscillator

Consider a harmonic oscillator with a quartic perturbation:

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2 \hat{X}^2 + \lambda \hat{X}^4 \quad (439)$$

Here  $\hat{H}_0$  is the harmonic oscillator and  $\hat{V} = \lambda \hat{X}^4$ . Using the expression  $\hat{X} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger)$  from §1.4, we have

$$\hat{X}^4 = \left( \frac{\hbar}{2m\omega} \right)^2 (\hat{a} + \hat{a}^\dagger)^4 \quad (440)$$

To compute  $E_n^{(1)} = \langle n | \hat{X}^4 | n \rangle$ , we need to expand  $(\hat{a} + \hat{a}^\dagger)^4$  and keep only terms with equal numbers of  $\hat{a}$  and  $\hat{a}^\dagger$  (since only these have nonzero diagonal matrix elements). Using the commutation relation  $[\hat{a}, \hat{a}^\dagger] = 1$  and normal ordering, the relevant terms are:

$$\langle n | (\hat{a} + \hat{a}^\dagger)^4 | n \rangle = 6n^2 + 6n + 3 \quad (441)$$

This can be verified by direct computation using  $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$  and  $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$ , expanding the fourth power and collecting terms. The combinatorics reduces to counting pairings (this is related to Wick’s theorem in quantum field theory).

Therefore, the first-order energy correction is

$$E_n^{(1)} = \lambda \left( \frac{\hbar}{2m\omega} \right)^2 (6n^2 + 6n + 3) \quad (442)$$

Note that the correction grows with  $n$ : higher excited states are more affected by the anharmonicity, as expected physically since they explore larger values of  $x$  where the quartic term dominates.

For the ground state ( $n = 0$ ):  $E_0^{(1)} = \frac{3\lambda\hbar^2}{4m^2\omega^2}$ .

### Example: quadratic Stark effect

The Stark effect is the shift of atomic energy levels in an external electric field  $\mathcal{E}$ . For the hydrogen atom in a uniform field  $\boldsymbol{\mathcal{E}} = \mathcal{E}\hat{z}$ , the perturbation is

$$\hat{V} = -e\mathcal{E}\hat{Z} = -e\mathcal{E}r\cos\theta \quad (443)$$

where  $-e$  is the electron charge.

For the ground state  $|1s\rangle = |n=1, \ell=0, m=0\rangle$ , the first-order correction vanishes:

$$E_{1s}^{(1)} = \langle 1s | \hat{V} | 1s \rangle = -e\mathcal{E} \langle 1s | r \cos\theta | 1s \rangle = 0 \quad (444)$$

because the integrand is odd under parity ( $r\cos\theta = z$ ), while the  $1s$  state has even parity. This is an example of the parity selection rule discussed in §5.3: the electric dipole operator is odd ( $\pi_O = -1$ ), and  $\pi' \cdot \pi_O \cdot \pi = (+1)(-1)(+1) = -1 \neq 1$ .

The leading correction is therefore second-order:

$$E_{1s}^{(2)} = \sum_{n\ell m \neq 1s} \frac{|\langle n, \ell, m | e\mathcal{E}r\cos\theta | 1s \rangle|^2}{E_1^{(0)} - E_n^{(0)}} \quad (445)$$

The selection rules from the Wigner-Eckart theorem (rank-1 tensor,  $\Delta\ell = \pm 1$ ,  $\Delta m = 0$ ) restrict the sum to states with  $\ell = 1$ ,  $m = 0$ . The result is proportional to  $\mathcal{E}^2$ , hence the name *quadratic* Stark effect. This shift can be written as

$$E_{1s}^{(2)} = -\frac{1}{2}\alpha_d\mathcal{E}^2 \quad (446)$$

where  $\alpha_d$  is the static electric dipole polarizability of the atom. For the hydrogen ground state, the exact calculation (which requires summing over the discrete spectrum and integrating over the continuum) gives  $\alpha_d = \frac{9}{2} \frac{a_0^3}{4\pi\epsilon_0}$ , where  $a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2}$  is the Bohr radius.

The physical interpretation is clear: the external field polarizes the atom, inducing a dipole moment proportional to  $\mathcal{E}$ , and the interaction of this induced dipole with the field gives an energy that is quadratic in  $\mathcal{E}$ .

## 6.2 Time-independent perturbation theory: degenerate case

### The problem with degeneracy

Suppose the unperturbed energy level  $E_n^{(0)}$  is  $g$ -fold degenerate, with orthonormal eigenstates  $|n_\alpha^{(0)}\rangle$ ,  $\alpha = 1, \dots, g$ :

$$\hat{H}_0 |n_\alpha^{(0)}\rangle = E_n^{(0)} |n_\alpha^{(0)}\rangle \quad (447)$$

The formula (434) for the first-order state correction contains terms

$$\frac{\langle n_\beta^{(0)} | \hat{V} | n_\alpha^{(0)} \rangle}{E_n^{(0)} - E_n^{(0)}} = \frac{V_{\beta\alpha}}{0} \quad (448)$$

which diverge whenever  $V_{\beta\alpha} \neq 0$ . The perturbative expansion in the original basis is therefore ill-defined.

The physical issue is that any linear combination  $\sum_\alpha c_\alpha |n_\alpha^{(0)}\rangle$  is also an eigenstate of  $\hat{H}_0$  with the same energy, and the perturbation will select a particular linear combination — the “correct” zeroth-order states — which are not determined by  $\hat{H}_0$  alone.

### Resolution: diagonalize $\hat{V}$ within the degenerate subspace

The strategy is to find linear combinations  $|\tilde{n}_\alpha^{(0)}\rangle = \sum_\beta c_{\alpha\beta} |n_\beta^{(0)}\rangle$  such that  $\hat{V}$  is diagonal in this new basis within the degenerate subspace:

$$\langle \tilde{n}_\alpha^{(0)} | \hat{V} | \tilde{n}_\beta^{(0)} \rangle = E_{n,\alpha}^{(1)} \delta_{\alpha\beta} \quad (449)$$

This amounts to solving the  $g \times g$  eigenvalue problem

$$\sum_{\beta=1}^g V_{\alpha\beta} c_{\gamma\beta} = E_{n,\gamma}^{(1)} c_{\gamma\alpha} \quad (450)$$

or equivalently, diagonalizing the  $g \times g$  matrix  $\mathbf{V}^{(n)}$  with elements  $V_{\alpha\beta} = \langle n_\alpha^{(0)} | \hat{V} | n_\beta^{(0)} \rangle$ .

The eigenvalues  $E_{n,\alpha}^{(1)}$  are the first-order energy corrections, and the eigenvectors determine the correct zeroth-order states. Once these are found, the non-degenerate formulas (434) and (437) can be applied with the divergent terms removed (the sum over  $m$  excludes states in the degenerate subspace):

$$|\tilde{n}_\alpha^{(1)}\rangle = \sum_{m \notin \mathcal{D}_n} \frac{\langle m^{(0)} | \hat{V} | \tilde{n}_\alpha^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle \quad (451)$$

where  $\mathcal{D}_n$  denotes the set of degenerate states at energy  $E_n^{(0)}$ .

If some eigenvalues of  $\mathbf{V}^{(n)}$  are still degenerate (that is, the perturbation does not completely lift the degeneracy), one must go to higher order or use additional operators that commute with both  $\hat{H}_0$  and  $\hat{V}$  to resolve the remaining degeneracy.

### Application: fine structure of hydrogen

The hydrogen atom provides the most important application of degenerate perturbation theory. The unperturbed Hamiltonian is the non-relativistic Coulomb Hamiltonian, whose energy levels

$$E_n^{(0)} = -\frac{me^4}{2(4\pi\epsilon_0)^2\hbar^2} \frac{1}{n^2} \equiv -\frac{E_R}{n^2} \quad (452)$$

(where  $E_R \approx 13.6$  eV is the Rydberg energy) depend only on the principal quantum number  $n$ , making each level  $n^2$ -fold degenerate (considering all values of  $\ell$  and  $m$ , and a further factor of 2 from spin).

The *fine structure* corrections arise from three effects, all of order  $\alpha^2 E_R$  (where  $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx 1/137$  is the fine-structure constant):

**(i) Relativistic kinetic energy correction.** The relativistic energy-momentum relation  $E = \sqrt{p^2 c^2 + m^2 c^4}$ , expanded to next order beyond  $p^2/2m$ , gives:

$$\hat{H}_{\text{rel}} = -\frac{\hat{P}^4}{8m^3 c^2} \quad (453)$$

This is a correction to the kinetic energy. Using  $\hat{P}^2 = 2m(\hat{H}_0 - \hat{V}(r))$  and first-order perturbation theory within each degenerate subspace, one finds

$$\langle \hat{H}_{\text{rel}} \rangle_{n\ell} = -\frac{E_R \alpha^2}{n^2} \left( \frac{n}{\ell + 1/2} - \frac{3}{4} \right) \quad (454)$$

Note that this depends on  $\ell$ , partially lifting the degeneracy.

**(ii) Spin-orbit coupling.** In the electron's rest frame, the nucleus orbits the electron, creating a magnetic field that interacts with the electron's magnetic moment. The resulting interaction is (after including the Thomas precession factor of  $1/2$ ):

$$\hat{H}_{\text{SO}} = \frac{1}{2m^2c^2} \frac{1}{r} \frac{dV}{dr} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \quad (455)$$

The operator  $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$  does not commute with either  $\hat{L}_z$  or  $\hat{S}_z$  individually, but it commutes with  $\hat{J}^2$ ,  $\hat{J}_z$ ,  $\hat{L}^2$ , and  $\hat{S}^2$ , where  $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$  is the total angular momentum. Using  $\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ , we find

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2}(\hat{J}^2 - \hat{L}^2 - \hat{S}^2) \quad (456)$$

so that in the coupled basis  $|n, \ell, j, m_j\rangle$  (with  $j = \ell \pm 1/2$ ):

$$\langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle = \frac{\hbar^2}{2}[j(j+1) - \ell(\ell+1) - 3/4] \quad (457)$$

The coupled basis  $\{|n, \ell, j, m_j\rangle\}$  is the correct zeroth-order basis for this perturbation — it is the basis that diagonalizes  $\hat{H}_{\text{SO}}$  within each degenerate subspace. This is an example where knowing the symmetry of the perturbation (it commutes with  $\hat{J}^2$  and  $\hat{J}_z$ ) immediately identifies the correct basis, without having to solve the secular equation (450) explicitly.

The expectation value of (455) in the hydrogen atom gives

$$\langle \hat{H}_{\text{SO}} \rangle_{n\ell j} = \frac{E_R \alpha^2}{2n^3} \frac{j(j+1) - \ell(\ell+1) - 3/4}{\ell(\ell+1/2)(\ell+1)} \quad (458)$$

for  $\ell \neq 0$ . For  $\ell = 0$  there is no spin-orbit coupling since  $\hat{\mathbf{L}} = 0$ .

**(iii) Darwin term.** This is a relativistic correction that affects only  $s$ -states ( $\ell = 0$ ). It arises from the “Zitterbewegung” (trembling motion) of the electron in the Dirac theory:

$$\hat{H}_{\text{D}} = \frac{\pi \hbar^2}{2m^2c^2} \frac{Ze^2}{4\pi\epsilon_0} \delta^3(\hat{\mathbf{X}}) \quad (459)$$

Its contribution is  $\langle \hat{H}_{\text{D}} \rangle_{n,\ell=0} = \frac{E_R \alpha^2}{n^3}$ .

**Combined result.** Remarkably, the three corrections combine into a single formula that depends only on  $n$  and  $j$  (not on  $\ell$  separately):

$$E_{nj}^{(1)} = -\frac{E_R \alpha^2}{n^2} \left( \frac{n}{j+1/2} - \frac{3}{4} \right) \quad (460)$$

This means that states with the same  $n$  and  $j$  but different  $\ell$  (such as  $2s_{1/2}$  and  $2p_{1/2}$ ) remain degenerate at this order. This accidental degeneracy is a special property of the Coulomb potential and is broken at higher order (the Lamb shift, which requires quantum electrodynamics to explain).

### 6.3 The Zeeman effect

When a hydrogen-like atom is placed in an external magnetic field  $\mathbf{B} = B\hat{z}$ , the additional interaction is

$$\hat{H}_B = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B} = \frac{eB}{2mc}(\hat{L}_z + 2\hat{S}_z) = \frac{eB}{2mc}(\hat{J}_z + \hat{S}_z) \quad (461)$$

where we used  $\hat{\boldsymbol{\mu}} = -\frac{e}{2mc}(\hat{\mathbf{L}} + 2\hat{\mathbf{S}})$  (the factor of 2 multiplying  $\hat{\mathbf{S}}$  is the electron  $g$ -factor,  $g_s \approx 2$ , a result that follows from the Dirac equation).

The relative magnitude of  $\hat{H}_B$  compared to  $\hat{H}_{\text{SO}}$  determines the appropriate treatment.

## Weak-field (anomalous) Zeeman effect

When  $\hat{H}_B \ll \hat{H}_{\text{SO}}$  (typical for laboratory magnetic fields  $B \lesssim 1$  T), the fine structure is established first and the magnetic field is treated as a perturbation on top of it. The good quantum numbers are  $n, \ell, j, m_j$ , and the energy shift is

$$E_B^{(1)} = \langle n, \ell, j, m_j | \hat{H}_B | n, \ell, j, m_j \rangle = \frac{eB}{2mc} \langle n, \ell, j, m_j | (\hat{J}_z + \hat{S}_z) | n, \ell, j, m_j \rangle \quad (462)$$

The expectation value of  $\hat{J}_z$  is simply  $m_j \hbar$ . The expectation value of  $\hat{S}_z$  requires the Wigner-Eckart theorem: since  $\hat{S}_z$  is the  $q = 0$  component of a rank-1 tensor operator, its matrix element within a fixed- $j$  multiplet is proportional to  $\hat{J}_z$ . The result (which can also be derived by a projection argument) is

$$\langle j, m_j | \hat{S}_z | j, m_j \rangle = \frac{\langle j, m_j | \hat{\mathbf{J}} \cdot \hat{\mathbf{S}} | j, m_j \rangle}{j(j+1)\hbar^2} \cdot m_j \hbar \quad (463)$$

Using  $\hat{\mathbf{J}} \cdot \hat{\mathbf{S}} = \frac{1}{2}(\hat{J}^2 + \hat{S}^2 - \hat{L}^2)$ , we obtain

$$E_B^{(1)} = g_j \mu_B m_j B \quad (464)$$

where  $\mu_B = \frac{e\hbar}{2mc}$  is the Bohr magneton and

$$g_j = 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \quad (465)$$

is the *Landé g-factor*. For a single electron ( $s = 1/2$ ),  $g_j$  depends on both  $j$  and  $\ell$ : for  $j = \ell + 1/2$ ,  $g_j = 1 + \frac{1}{2\ell+1}$ ; for  $j = \ell - 1/2$ ,  $g_j = 1 - \frac{1}{2\ell+1}$ . The fact that  $g_j \neq 1$  (and is not even the same for different  $j$  values) is why this is called the “anomalous” Zeeman effect — it puzzled physicists before the discovery of electron spin.

Each level with quantum numbers  $(n, \ell, j)$  splits into  $2j + 1$  sublevels, equally spaced by  $g_j \mu_B B$ .

## Strong-field (Paschen-Back) effect

When  $\hat{H}_B \gg \hat{H}_{\text{SO}}$ , the magnetic field dominates and the good quantum numbers are  $n, \ell, m_\ell, m_s$ . The energy is

$$E = E_n^{(0)} + \mu_B B(m_\ell + 2m_s) + \text{fine structure corrections} \quad (466)$$

The spin-orbit coupling is now treated as a perturbation on top of the magnetic splitting. In this regime the splitting pattern simplifies: it is called the “normal” Zeeman effect because the splitting is independent of the angular momentum quantum numbers (at zeroth order in spin-orbit).

The crossover between the weak-field and strong-field regimes provides an instructive example of how the appropriate choice of basis (and of the unperturbed Hamiltonian) depends on the physical situation.

## 6.4 The variational method

### The variational principle

Perturbation theory requires a known, exactly solvable “nearby” problem. The *variational method* makes no such assumption: it provides an upper bound on the ground state energy of *any* Hamiltonian, given only the ability to compute expectation values.

**Theorem (Rayleigh-Ritz).** For any normalized state  $|\psi\rangle$ ,

$$\langle\psi|\hat{H}|\psi\rangle \geq E_0 \quad (467)$$

where  $E_0$  is the exact ground state energy. Equality holds if and only if  $|\psi\rangle$  is the exact ground state.

*Proof.* Expand  $|\psi\rangle$  in the exact energy eigenbasis:  $|\psi\rangle = \sum_n c_n |n\rangle$  with  $\sum_n |c_n|^2 = 1$ . Then

$$\langle\psi|\hat{H}|\psi\rangle = \sum_n |c_n|^2 E_n \geq E_0 \sum_n |c_n|^2 = E_0 \quad (468)$$

since  $E_n \geq E_0$  for all  $n$ . □

The method consists of choosing a trial wave function  $|\psi(\boldsymbol{\alpha})\rangle$  that depends on a set of parameters  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots)$ , and minimizing

$$E(\boldsymbol{\alpha}) = \frac{\langle\psi(\boldsymbol{\alpha})|\hat{H}|\psi(\boldsymbol{\alpha})\rangle}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle} \quad (469)$$

over  $\boldsymbol{\alpha}$ . The minimum value is the best upper bound on  $E_0$  within the chosen family of trial states. The quality of the bound depends on the physical insight used to choose the trial family.

### Example: ground state of helium

The helium atom has two electrons in the Coulomb field of a nucleus with charge  $Z = 2$ :

$$\hat{H} = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{Ze^2}{4\pi\epsilon_0} \left( \frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (470)$$

The electron-electron repulsion term  $e^2/(4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|)$  makes exact solution impossible.

If we ignore the repulsion, the ground state is a product of two hydrogen-like 1s wave functions with  $Z = 2$ :

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_1+r_2)/a_0} \quad (471)$$

This gives  $E = -2Z^2 E_R = -8E_R = -108.8$  eV, while the experimental value is  $-78.975$  eV — the discrepancy of 38% shows that the electron-electron interaction is far from negligible.

The variational approach is to use the same functional form but with an *effective nuclear charge*  $Z_{\text{eff}}$  as a variational parameter, replacing  $Z = 2$  in (471) by  $Z_{\text{eff}}$ :

$$\psi_{\text{trial}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z_{\text{eff}}^3}{\pi a_0^3} e^{-Z_{\text{eff}}(r_1+r_2)/a_0} \quad (472)$$

The idea is that each electron partially screens the nuclear charge as seen by the other electron, so the effective charge should be less than  $Z = 2$ .

Computing  $E(Z_{\text{eff}}) = \langle\psi_{\text{trial}}|\hat{H}|\psi_{\text{trial}}\rangle$  (the calculation involves a somewhat lengthy but standard evaluation of the electron-electron repulsion integral, see [1], §6.1, or [15], §8.2):

$$E(Z_{\text{eff}}) = \left( Z_{\text{eff}}^2 - 2Z Z_{\text{eff}} + \frac{5}{8} Z_{\text{eff}} \right) E_R \quad (473)$$

Minimizing:  $\frac{dE}{dZ_{\text{eff}}} = 0$  gives

$$Z_{\text{eff}} = Z - \frac{5}{16} = 2 - 0.3125 = 1.6875 \quad (474)$$

and the corresponding energy is

$$E_{\min} = -\left(Z - \frac{5}{16}\right)^2 \cdot 2E_R = -(1.6875)^2 \times 27.2 \text{ eV} = -77.5 \text{ eV} \quad (475)$$

This is within 2% of the experimental value of  $-78.975 \text{ eV}$  — a remarkable improvement from the unscreened value, achieved with a single variational parameter and a physically motivated trial function. The value  $Z_{\text{eff}} = 1.6875 < 2$  confirms the physical expectation of screening.

## Connection with Quantum Information: the Variational Quantum Eigensolver

The variational principle has a natural quantum computational extension. In the *Variational Quantum Eigensolver* (VQE) algorithm [16], a quantum computer prepares a parametrized trial state  $|\psi(\boldsymbol{\theta})\rangle$  using a circuit with tunable parameters  $\boldsymbol{\theta}$ , measures the expectation value  $\langle \hat{H} \rangle$ , and a classical optimizer adjusts  $\boldsymbol{\theta}$  to minimize the energy. This hybrid quantum-classical algorithm is one of the leading candidates for achieving useful quantum advantage on near-term quantum hardware, particularly for quantum chemistry applications where the Hilbert space dimension grows exponentially with the number of electrons.

The key insight is that while the trial state space of a quantum circuit can be exponentially large (exploring entangled states that are classically intractable to represent), the optimization itself is a classical variational problem — exactly the same Rayleigh-Ritz principle we discussed above, but with a much richer space of trial states.

## 6.5 Time-dependent perturbation theory

We now consider perturbations that vary in time. The physical setting is a system initially in a known eigenstate of  $\hat{H}_0$ , subjected to a time-dependent perturbation  $\hat{V}(t)$  that is turned on at  $t = 0$ :

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t), \quad \hat{V}(t) = 0 \text{ for } t < 0 \quad (476)$$

The central question is: what is the probability of finding the system in a different eigenstate of  $\hat{H}_0$  at a later time  $t > 0$ ? This is the problem of *transition probabilities*.

### The interaction picture

We have already encountered the Schrödinger picture (states evolve, operators fixed) and the Heisenberg picture (operators evolve, states fixed) in §1.6. For time-dependent perturbation theory, a third picture is most convenient: the *interaction picture* (also called the Dirac picture).

The idea is to separate the “trivial” time evolution due to  $\hat{H}_0$  from the “interesting” evolution due to  $\hat{V}(t)$ . Define the interaction-picture state:

$$|\psi_I(t)\rangle = e^{i\hat{H}_0 t/\hbar} |\psi_S(t)\rangle \quad (477)$$

where  $|\psi_S(t)\rangle$  is the Schrödinger-picture state. Similarly, operators in the interaction picture are:

$$\hat{O}_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{O}_S e^{-i\hat{H}_0 t/\hbar} \quad (478)$$

The equation of motion for  $|\psi_I(t)\rangle$  is obtained by differentiating (477) and using the Schrödinger equation  $i\hbar\partial_t|\psi_S\rangle = (\hat{H}_0 + \hat{V})|\psi_S\rangle$ :

$$i\hbar\frac{d}{dt}|\psi_I(t)\rangle = -\hat{H}_0 e^{i\hat{H}_0 t/\hbar} |\psi_S\rangle + e^{i\hat{H}_0 t/\hbar} (\hat{H}_0 + \hat{V}) |\psi_S\rangle$$

$$\begin{aligned}
&= e^{i\hat{H}_0 t/\hbar} \hat{V}(t) e^{-i\hat{H}_0 t/\hbar} e^{i\hat{H}_0 t/\hbar} |\psi_S\rangle \\
&= \hat{V}_I(t) |\psi_I(t)\rangle
\end{aligned} \tag{479}$$

where  $\hat{V}_I(t) = e^{i\hat{H}_0 t/\hbar} \hat{V}(t) e^{-i\hat{H}_0 t/\hbar}$  is the perturbation in the interaction picture. The key feature is that only the perturbation drives the evolution: if  $\hat{V} = 0$ , then  $|\psi_I(t)\rangle$  is constant.

### The Dyson series

The formal solution of (479) is the time-ordered exponential:

$$|\psi_I(t)\rangle = \hat{U}_I(t, 0) |\psi_I(0)\rangle \tag{480}$$

where the interaction-picture evolution operator is

$$\hat{U}_I(t, 0) = \mathcal{T} \exp \left( -\frac{i}{\hbar} \int_0^t dt' \hat{V}_I(t') \right) \tag{481}$$

and  $\mathcal{T}$  denotes time ordering. Expanding the exponential gives the *Dyson series*:

$$\begin{aligned}
\hat{U}_I(t, 0) &= \mathbb{1} + \left( -\frac{i}{\hbar} \right) \int_0^t dt_1 \hat{V}_I(t_1) \\
&\quad + \left( -\frac{i}{\hbar} \right)^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \hat{V}_I(t_1) \hat{V}_I(t_2) + \dots
\end{aligned} \tag{482}$$

Note the nested time ordering: in the second-order term,  $t_2 \leq t_1$ , ensuring that operators are applied in chronological order from right to left.

### Transition amplitudes

Suppose the system starts in an eigenstate  $|i\rangle$  of  $\hat{H}_0$  at  $t = 0$ . The amplitude to find it in a different eigenstate  $|f\rangle$  at time  $t$  is

$$c_{f \leftarrow i}(t) = \langle f | \hat{U}_I(t, 0) | i \rangle \tag{483}$$

and the transition probability is  $P_{f \leftarrow i}(t) = |c_{f \leftarrow i}(t)|^2$ .

To first order in  $\hat{V}$ :

$$c_{f \leftarrow i}^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle f | \hat{V}_I(t') | i \rangle = -\frac{i}{\hbar} \int_0^t dt' e^{i\omega_{fi}t'} \langle f | \hat{V}(t') | i \rangle \tag{484}$$

where  $\omega_{fi} = (E_f - E_i)/\hbar$  is the Bohr frequency. The exponential factor  $e^{i\omega_{fi}t'}$  arises from the interaction-picture transformation.

### Constant perturbation turned on abruptly

Consider a perturbation that is turned on suddenly at  $t = 0$  and remains constant:  $\hat{V}(t) = \hat{V} \Theta(t)$ , where  $\Theta(t)$  is the Heaviside step function. Then:

$$c_{f \leftarrow i}^{(1)}(t) = -\frac{i}{\hbar} V_{fi} \int_0^t dt' e^{i\omega_{fi}t'} = -\frac{V_{fi}}{\hbar\omega_{fi}} (e^{i\omega_{fi}t} - 1) \tag{485}$$

The transition probability is

$$P_{f \leftarrow i}(t) = \frac{|V_{fi}|^2}{\hbar^2} \frac{4 \sin^2(\omega_{fi}t/2)}{\omega_{fi}^2} = |V_{fi}|^2 \frac{t^2}{\hbar^2} \text{sinc}^2\left(\frac{\omega_{fi}t}{2}\right) \tag{486}$$

where  $\text{sinc}(x) = \sin(x)/x$ .

The function  $\text{sinc}^2(\omega_{fi}t/2)$  is sharply peaked around  $\omega_{fi} = 0$  (i.e.,  $E_f \approx E_i$ ), with a width of order  $2\pi/t$ . This has a clear physical interpretation: the energy conservation condition  $E_f = E_i$  is enforced more and more precisely as  $t$  increases. For short times, the system can make “virtual” transitions to states with  $E_f \neq E_i$ , but these become suppressed at long times. This is intimately connected with the time-energy uncertainty relation discussed in §1.6.

### Harmonic perturbation and resonance

A particularly important case is a perturbation that oscillates harmonically in time:

$$\hat{V}(t) = \hat{V}_0 e^{-i\omega t} + \hat{V}_0^\dagger e^{i\omega t} \quad (487)$$

(with  $\hat{V}_0$  possibly non-Hermitian). The first-order transition amplitude from  $|i\rangle$  to  $|f\rangle$  is

$$c_{f \leftarrow i}^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \left[ V_{fi} e^{i(\omega_{fi} - \omega)t'} + V'_{fi} e^{i(\omega_{fi} + \omega)t'} \right] \quad (488)$$

where  $V_{fi} = \langle f | \hat{V}_0 | i \rangle$  and  $V'_{fi} = \langle f | \hat{V}_0^\dagger | i \rangle$ .

The first term is resonant when  $\omega_{fi} \approx \omega$ , i.e.,  $E_f \approx E_i + \hbar\omega$  (absorption of energy  $\hbar\omega$ ). The second term is resonant when  $\omega_{fi} \approx -\omega$ , i.e.,  $E_f \approx E_i - \hbar\omega$  (stimulated emission of energy  $\hbar\omega$ ). Near resonance, only the resonant term contributes significantly (this is the *rotating wave approximation*), and the transition probability becomes

$$P_{f \leftarrow i}(t) \approx \frac{|V_{fi}|^2}{\hbar^2} \frac{4 \sin^2[(\omega_{fi} - \omega)t/2]}{(\omega_{fi} - \omega)^2} \quad (489)$$

### Rabi oscillations in a two-level system

For a two-level system driven by a harmonic perturbation near resonance, perturbation theory breaks down at long times (the transition probability exceeds 1). The exact solution can be obtained.

Consider two states  $|1\rangle$  and  $|2\rangle$  with energies  $E_1$  and  $E_2$  ( $E_2 > E_1$ ), coupled by a perturbation  $\langle 2 | \hat{V}(t) | 1 \rangle = V_0 e^{-i\omega t}$ . Define  $\omega_0 = (E_2 - E_1)/\hbar$  and the detuning  $\delta = \omega - \omega_0$ . Writing  $|\psi(t)\rangle = c_1(t)e^{-iE_1 t/\hbar}|1\rangle + c_2(t)e^{-iE_2 t/\hbar}|2\rangle$ , Schrödinger's equation in the rotating wave approximation gives:

$$i\hbar \dot{c}_1 = V_0^* e^{-i\delta t} c_2 \quad (490)$$

$$i\hbar \dot{c}_2 = V_0 e^{i\delta t} c_1 \quad (491)$$

These coupled equations can be solved exactly. With initial conditions  $c_1(0) = 1$ ,  $c_2(0) = 0$ , the result is

$$P_{2 \leftarrow 1}(t) = |c_2(t)|^2 = \frac{\Omega_R^2}{\Omega^2} \sin^2\left(\frac{\Omega t}{2}\right) \quad (492)$$

where  $\Omega_R = |V_0|/\hbar$  is the *Rabi frequency* and  $\Omega = \sqrt{\Omega_R^2 + \delta^2}$  is the *generalized Rabi frequency*.

At exact resonance ( $\delta = 0$ ), the system oscillates completely between the two states:

$$P_{2 \leftarrow 1}(t) = \sin^2\left(\frac{\Omega_R t}{2}\right) \quad (493)$$

At time  $t_\pi = \pi/\Omega_R$  (a “ $\pi$ -pulse”), the population is completely transferred to state  $|2\rangle$ . At time  $t_{\pi/2} = \pi/(2\Omega_R)$  (a “ $\pi/2$ -pulse”), the system is in an equal superposition  $\frac{1}{\sqrt{2}}(|1\rangle + e^{i\phi}|2\rangle)$ .

This is directly connected to quantum information: the  $\pi$ -pulse implements a NOT gate, and the  $\pi/2$ -pulse implements a Hadamard-like gate. The control of Rabi oscillations through precise pulse timing is the fundamental mechanism by which quantum gates are implemented in many physical platforms (trapped ions, superconducting qubits, NV centers in diamond).

## 6.6 Fermi's golden rule and interaction with radiation

### Transition rate to a continuum

Consider transitions from a discrete state  $|i\rangle$  to a continuum of final states  $|f\rangle$  with a density of states  $\rho(E_f)$ . The total transition probability is obtained by integrating (486) over the final states:

$$P_{\text{total}}(t) = \int dE_f \rho(E_f) P_{f \leftarrow i}(t) = \frac{1}{\hbar^2} \int dE_f \rho(E_f) |V_{fi}|^2 \frac{4 \sin^2(\omega_{fi}t/2)}{\omega_{fi}^2} \quad (494)$$

For large  $t$ , the function  $\frac{\sin^2(\omega_{fi}t/2)}{\omega_{fi}^2}$  becomes very sharply peaked around  $\omega_{fi} = 0$ . More precisely, using the distributional identity

$$\lim_{t \rightarrow \infty} \frac{\sin^2(\omega t/2)}{\pi \omega^2/2} \cdot \frac{1}{t} = \delta(\omega) \quad (495)$$

one obtains  $P_{\text{total}}(t) \approx \Gamma \cdot t$ , where the transition *rate* (probability per unit time) is given by *Fermi's golden rule*<sup>9</sup>:

$$\Gamma_{f \leftarrow i} = \frac{2\pi}{\hbar} |\langle f | \hat{V} | i \rangle|^2 \rho(E_f) \Big|_{E_f=E_i} \quad (496)$$

This is one of the most widely used formulas in quantum mechanics. Note the following features:

- The transition rate is proportional to the *square* of the matrix element — this is characteristic of first-order perturbation theory and accounts for many order-of-magnitude estimates in atomic, nuclear, and particle physics.
- The factor  $\rho(E_f)$  evaluated at  $E_f = E_i$  enforces energy conservation in the long-time limit.
- The formula is valid in the regime  $\Gamma^{-1} \ll t \ll$  (higher-order correction time), i.e., the transition probability is much less than 1 but the sinc function has already narrowed to approximate a delta function.

For a harmonic perturbation (487), the golden rule becomes

$$\Gamma_{f \leftarrow i} = \frac{2\pi}{\hbar} |V_{fi}|^2 \rho(E_f) \Big|_{E_f=E_i \pm \hbar\omega} \quad (497)$$

with the  $+$  sign for absorption and  $-$  for stimulated emission.

### Interaction with classical electromagnetic radiation

The most important application of Fermi's golden rule is the interaction of atoms with light. The Hamiltonian of a charged particle in an electromagnetic field is obtained by the minimal coupling prescription  $\hat{\mathbf{P}} \rightarrow \hat{\mathbf{P}} - \frac{e}{c} \mathbf{A}$ :

$$\hat{H} = \frac{1}{2m} \left( \hat{\mathbf{P}} - \frac{e}{c} \mathbf{A} \right)^2 + V(r) = \hat{H}_0 - \frac{e}{mc} \mathbf{A} \cdot \hat{\mathbf{P}} + \frac{e^2}{2mc^2} |\mathbf{A}|^2 \quad (498)$$

<sup>9</sup>The name is due to Fermi, who called it “Golden Rule No. 2” in his lectures at the University of Chicago. The result was originally derived by Dirac (1927).

where we used the Coulomb gauge  $\nabla \cdot \mathbf{A} = 0$  (so that  $\mathbf{A} \cdot \hat{\mathbf{P}} = \hat{\mathbf{P}} \cdot \mathbf{A}$ ). For weak fields, the  $|\mathbf{A}|^2$  term is negligible and the perturbation is

$$\hat{V}(t) = -\frac{e}{mc} \mathbf{A}(\mathbf{x}, t) \cdot \hat{\mathbf{P}} \quad (499)$$

For a monochromatic plane wave,  $\mathbf{A} = A_0 \hat{\mathbf{e}} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)$ , where  $\hat{\mathbf{e}}$  is the polarization vector. In the *electric dipole approximation*, valid when the wavelength of the radiation is much larger than the atomic size ( $|\mathbf{k}| a_0 \ll 1$ , which holds for optical and UV light), we set  $e^{i\mathbf{k} \cdot \mathbf{x}} \approx 1$ . The perturbation simplifies to

$$\hat{V}(t) \approx -\frac{eA_0}{2mc} \hat{\mathbf{e}} \cdot \hat{\mathbf{P}} (e^{-i\omega t} + e^{i\omega t}) \quad (500)$$

Using the operator identity  $\langle f | \hat{\mathbf{P}} | i \rangle = \frac{im}{\hbar} (E_f - E_i) \langle f | \hat{\mathbf{X}} | i \rangle$  (which follows from  $[\hat{H}_0, \hat{\mathbf{X}}] = -\frac{i\hbar}{m} \hat{\mathbf{P}}$ ), the matrix element can be rewritten as

$$\langle f | \hat{V} | i \rangle \propto \omega_{fi} \hat{\mathbf{e}} \cdot \langle f | \hat{\mathbf{X}} | i \rangle = \omega_{fi} \hat{\mathbf{e}} \cdot \mathbf{d}_{fi} \quad (501)$$

where  $\mathbf{d}_{fi} = \langle f | e \hat{\mathbf{X}} | i \rangle$  is the *electric dipole matrix element*. Applying Fermi's golden rule (497):

$$\Gamma_{f \leftarrow i}^{(\text{abs})} = \frac{\pi e^2}{m^2 \omega^2 c^2} |A_0|^2 \omega_{fi}^2 |\hat{\mathbf{e}} \cdot \mathbf{d}_{fi}|^2 \rho(E_f) \quad (502)$$

The electric dipole selection rules follow directly from the properties of  $\mathbf{d}_{fi}$ : using the Wigner-Eckart theorem and the parity selection rule (§5.3), we recover  $\Delta\ell = \pm 1$ ,  $\Delta m = 0, \pm 1$  (depending on the polarization direction).

## Einstein's $A$ and $B$ coefficients

In 1917, Einstein considered the equilibrium between atoms and a thermal radiation field and introduced three processes:

- **Absorption:** an atom in state  $|i\rangle$  absorbs a photon and transitions to  $|f\rangle$  with  $E_f > E_i$ . The rate is proportional to the energy density  $u(\omega)$  of the radiation field:  $\Gamma_{\text{abs}} = B_{i \rightarrow f} u(\omega_{fi})$ .
- **Stimulated emission:** an atom in  $|f\rangle$  emits a photon *identical* to the incident photon and transitions to  $|i\rangle$ . The rate is  $\Gamma_{\text{stim}} = B_{f \rightarrow i} u(\omega_{fi})$ .
- **Spontaneous emission:** an atom in  $|f\rangle$  emits a photon even in the absence of incident radiation. The rate is  $A_{f \rightarrow i}$  (independent of  $u$ ).

The first two are described by the semiclassical theory (Fermi's golden rule with a classical field). One finds  $B_{i \rightarrow f} = B_{f \rightarrow i}$ , reflecting the symmetry of the interaction. Spontaneous emission requires quantization of the electromagnetic field (quantum electrodynamics), but Einstein's argument gives its rate from the requirement of thermal equilibrium with the Planck distribution:

$$A_{f \rightarrow i} = \frac{\hbar \omega_{fi}^3}{\pi^2 c^3} B_{f \rightarrow i} \quad (503)$$

The  $\omega^3$  dependence explains why spontaneous emission is dominant at optical frequencies but negligible at radio frequencies, a fact of great practical importance in atomic physics and quantum optics.

## Connection with Quantum Information: quantum gates from light-matter interaction

The controlled interaction of qubits with electromagnetic radiation is the experimental foundation of quantum computation and quantum communication. A few key connections:

- **Qubit gates via Rabi oscillations.** As discussed in §6.5, laser pulses of controlled duration and phase implement single-qubit gates (Hadamard, phase gates, arbitrary rotations) through Rabi oscillations. The precision of these gates is determined by the control of the pulse area  $\int_0^T \Omega_R(t) dt$ .
- **Stimulated Raman transitions.** In many physical implementations (trapped ions, neutral atoms), the qubit states are two ground-state hyperfine levels that are not directly connected by a dipole transition. Two-photon stimulated Raman transitions — involving a virtual intermediate state — provide the effective coupling. This is a second-order process in time-dependent perturbation theory.
- **Cavity QED and circuit QED.** When an atom (or artificial atom) is placed inside a high-quality resonator, the electromagnetic field must be treated quantum mechanically. The Jaynes-Cummings model describes the coherent exchange of a single photon between the atom and the cavity mode, leading to vacuum Rabi oscillations. This is the basis of circuit QED implementations of superconducting quantum computers, where microwave photons in coplanar waveguide resonators are used to mediate interactions between superconducting qubits.
- **Decoherence from spontaneous emission.** From the information-theoretic perspective, spontaneous emission is a quantum channel that maps  $|f\rangle \rightarrow |i\rangle$  while emitting a photon into the environment. The emitted photon carries “which-state” information, destroying coherence in the qubit subspace. The amplitude damping channel of quantum information theory is a mathematical model of this process.

## Semiclassical (WKB) approximation

For completeness, we mention the semiclassical (WKB) approximation, which is complementary to perturbation theory. While perturbation theory treats a small additional interaction, the WKB method treats the quantum corrections to classical motion as small. It is most useful in one dimension and for slowly varying potentials.

The WKB ansatz for the wave function is

$$\psi(x) \approx \frac{C}{\sqrt{p(x)}} \exp\left(\pm \frac{i}{\hbar} \int^x p(x') dx'\right) \quad (504)$$

where  $p(x) = \sqrt{2m(E - V(x))}$  is the local classical momentum. This is valid when the potential varies slowly on the scale of the de Broglie wavelength:  $|\frac{d\lambda}{dx}| \ll 1$ , where  $\lambda = 2\pi\hbar/p$ .

The WKB approximation fails at classical turning points ( $p(x) = 0$ ), where the connection formulas bridge the classically allowed and forbidden regions. The Bohr-Sommerfeld quantization condition

$$\oint p(x) dx = \left(n + \frac{1}{2}\right) 2\pi\hbar, \quad n = 0, 1, 2, \dots \quad (505)$$

(where the integral is over one complete classical period) gives approximate energy eigenvalues that become exact for the harmonic oscillator. The WKB method also provides the standard formula for quantum tunneling rates through potential barriers.

A detailed treatment of the WKB approximation can be found in [1], Chapter 5, §5.5, or [14], Volume 1, Complement H<sub>III</sub>.

## 7 Scattering theory

Scattering experiments are one of the primary ways we learn about the microscopic world. From Rutherford’s discovery of the atomic nucleus to the detection of the Higgs boson at the LHC, the basic paradigm

is the same: send a beam of particles toward a target, observe how they deflect, and infer the properties of the interaction responsible for the deflection.

In this chapter we develop the quantum mechanical theory of scattering. The central objects are the *scattering amplitude* and the *cross section*, which encode the probability of deflection as a function of angle and energy. We derive the Lippmann-Schwinger equation, which reformulates the scattering problem as an integral equation; the Born approximation, which provides explicit results for weak potentials; and the method of partial waves, which is exact and particularly powerful for central potentials. We close with the *S*-matrix formalism, which reveals the deep connection between unitarity and the conservation of probability — a connection that admits a natural interpretation in the language of quantum information.

Main references for this chapter are [1], Chapter 7; [17], Chapter 13; and [14], Volume 2, Chapter VIII.

## 7.1 The scattering problem: kinematics and cross section

### Physical setup

Consider a beam of particles with well-defined momentum  $\hbar\mathbf{k}$  incident on a localized potential  $V(\mathbf{x})$  (the “target”). We assume the potential vanishes (or falls off sufficiently fast) outside some finite region. Far from the target, the scattered particles propagate freely and are detected at angles  $(\theta, \phi)$  with respect to the incident direction.

The goal is to compute the probability of scattering into a given direction as a function of the incident energy and the scattering angles. This probability is encoded in the *differential cross section*.

### Differential cross section

The differential cross section is defined operationally. Let  $J_{\text{inc}}$  be the incident flux (number of particles per unit time per unit area) and  $dN/d\Omega$  be the number of particles scattered per unit time into the solid angle  $d\Omega$  around the direction  $(\theta, \phi)$ . Then

$$\frac{d\sigma}{d\Omega} \equiv \frac{1}{J_{\text{inc}}} \frac{dN}{d\Omega} \quad (506)$$

The differential cross section has dimensions of area and depends on  $\theta$ ,  $\phi$ , and the incident energy  $E = \hbar^2 k^2 / 2m$  (we work in the center-of-mass frame with reduced mass  $m$ ).

The *total cross section* is the integral over all solid angles:

$$\sigma_{\text{tot}} = \int d\Omega \frac{d\sigma}{d\Omega} = \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta \frac{d\sigma}{d\Omega} \quad (507)$$

It represents the effective “target area” seen by the incident beam.

### Asymptotic boundary condition

The time-independent Schrödinger equation for the scattering problem is

$$\left( -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right) \psi(\mathbf{x}) = E \psi(\mathbf{x}) \quad (508)$$

with  $E > 0$  (unbound states). The key physical requirement is the *asymptotic boundary condition*: far from the scattering region, the wave function must consist of the incident plane wave plus an outgoing spherical wave:

$$\psi(\mathbf{x}) \xrightarrow{r \rightarrow \infty} e^{i\mathbf{k} \cdot \mathbf{x}} + f(\theta, \phi) \frac{e^{ikr}}{r} \quad (509)$$

where we have chosen the incident wave to propagate along the direction of  $\mathbf{k}$  (conventionally taken as  $\hat{z}$ ). The function  $f(\theta, \phi)$  is the *scattering amplitude*; it encodes all the information about the scattering process.

The outgoing spherical wave  $e^{ikr}/r$  (rather than  $e^{-ikr}/r$ ) is selected by causality: the scattered wave propagates *away* from the target. This choice is enforced by the  $i\epsilon$  prescription in the Green's function, as we shall see below.

### Relation between scattering amplitude and cross section

The probability current density for a wave function  $\psi$  is  $\mathbf{J} = \frac{\hbar}{2mi}(\psi^* \nabla \psi - \psi \nabla \psi^*)$ . For the incident plane wave  $e^{ikz}$ :

$$J_{\text{inc}} = \frac{\hbar k}{m} \equiv v \quad (510)$$

where  $v = \hbar k/m$  is the particle velocity.

For the scattered wave  $f(\theta, \phi) \frac{e^{ikr}}{r}$  at large  $r$ , the radial component of the current is

$$J_{\text{sc}} = \frac{\hbar k}{m} \frac{|f(\theta, \phi)|^2}{r^2} \quad (511)$$

The number of particles crossing a surface element  $r^2 d\Omega$  per unit time is  $J_{\text{sc}} \cdot r^2 d\Omega = v |f|^2 d\Omega$ . Therefore, from (506):

$$\boxed{\frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2} \quad (512)$$

This is a fundamental result: the differential cross section is simply the squared modulus of the scattering amplitude. The entire scattering problem thus reduces to computing  $f(\theta, \phi)$ .

## 7.2 The Lippmann-Schwinger equation

Rather than solving the differential equation (508) with the boundary condition (509) directly, it is more powerful (and more physical) to reformulate the problem as an integral equation. This is the Lippmann-Schwinger approach.

### Operator formulation

Rewrite the Schrödinger equation as

$$(E - \hat{H}_0)|\psi\rangle = \hat{V}|\psi\rangle \quad (513)$$

where  $\hat{H}_0 = \hat{P}^2/2m$  is the free Hamiltonian. If  $(E - \hat{H}_0)$  were invertible, we could write  $|\psi\rangle = (E - \hat{H}_0)^{-1} \hat{V}|\psi\rangle$ . However, for  $E > 0$  the operator  $(E - \hat{H}_0)$  has a zero eigenvalue (the free particle state with energy  $E$ ), so its inverse does not exist. The remedy is to add an infinitesimal imaginary part:

$$|\psi^{(\pm)}\rangle = |\mathbf{k}\rangle + \frac{1}{E - \hat{H}_0 \pm i\epsilon} \hat{V}|\psi^{(\pm)}\rangle \quad (514)$$

where  $|\mathbf{k}\rangle$  is the free-particle eigenstate with  $\hat{H}_0|\mathbf{k}\rangle = E|\mathbf{k}\rangle$  (i.e.,  $E = \hbar^2 k^2/2m$ ), and  $\epsilon \rightarrow 0^+$ . This is the *Lippmann-Schwinger equation*.

The  $\pm i\epsilon$  prescription selects the appropriate boundary condition:  $+i\epsilon$  gives the *retarded* Green's function and outgoing spherical waves (the physical scattering solution), while  $-i\epsilon$  gives the *advanced* Green's function and incoming spherical waves. We focus on the  $+i\epsilon$  case and write  $|\psi^{(+)}\rangle \equiv |\psi\rangle$  for brevity.

## The free Green's function

Define the free-particle Green's operator

$$\hat{G}_0^{(+)}(E) = \frac{1}{E - \hat{H}_0 + i\epsilon} \quad (515)$$

Its position-space matrix element  $G_0^{(+)}(\mathbf{x}, \mathbf{x}'; E) = \langle \mathbf{x} | \hat{G}_0^{(+)} | \mathbf{x}' \rangle$  is the Green's function satisfying

$$\left( E + \frac{\hbar^2}{2m} \nabla^2 \right) G_0^{(+)}(\mathbf{x}, \mathbf{x}'; E) = \delta^3(\mathbf{x} - \mathbf{x}') \quad (516)$$

with outgoing-wave boundary conditions. The standard calculation proceeds by inserting a complete set of momentum eigenstates  $|\mathbf{p}\rangle$ :

$$G_0^{(+)}(\mathbf{x}, \mathbf{x}'; E) = \int \frac{d^3p}{(2\pi\hbar)^3} \frac{e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')/\hbar}}{E - p^2/2m + i\epsilon} \quad (517)$$

This integral can be evaluated by contour integration. Writing  $E = \hbar^2 k^2/2m$  and  $\mathbf{R} = \mathbf{x} - \mathbf{x}'$ , the angular integration gives a factor  $4\pi \sin(pR/\hbar)/(pR/\hbar)$ , reducing the problem to a one-dimensional integral over  $p$ . The poles of the integrand are at  $p = \pm \hbar k \mp i\epsilon'$  (with  $\epsilon'$  proportional to  $\epsilon$ ). Closing the contour in the upper half-plane and picking up the residue at  $p = +\hbar k + i\epsilon'$ , one obtains

$$G_0^{(+)}(\mathbf{x}, \mathbf{x}'; E) = -\frac{2m}{4\pi\hbar^2} \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} \quad (518)$$

This is an outgoing spherical wave centered at  $\mathbf{x}'$ , as expected from the  $+i\epsilon$  prescription. Had we used  $-i\epsilon$ , we would have obtained  $e^{-ik|\mathbf{x}-\mathbf{x}'|}$  (an incoming wave).

## The Lippmann-Schwinger equation in position space

Projecting (514) onto  $\langle \mathbf{x} |$  and inserting  $\mathbb{1} = \int d^3x' |\mathbf{x}'\rangle \langle \mathbf{x}'|$ :

$$\psi(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} - \frac{2m}{4\pi\hbar^2} \int d^3x' \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} V(\mathbf{x}') \psi(\mathbf{x}') \quad (519)$$

where we used  $\langle \mathbf{x} | \mathbf{k} \rangle = e^{i\mathbf{k}\cdot\mathbf{x}}$  (with the plane-wave normalization convention implicit in  $|\mathbf{k}\rangle$ ). This is an integral equation for the unknown  $\psi(\mathbf{x})$ : the wave function appears on both sides. Solving it exactly is generally as difficult as solving the original differential equation, but the integral form is the natural starting point for systematic approximations.

## Extracting the scattering amplitude

To connect with the asymptotic form (509), consider (519) in the limit  $r = |\mathbf{x}| \rightarrow \infty$  with  $\mathbf{x}'$  confined to the scattering region (where  $V \neq 0$ ). In this limit:

$$|\mathbf{x} - \mathbf{x}'| = r - \hat{r} \cdot \mathbf{x}' + O(1/r) \quad (520)$$

where  $\hat{r} = \mathbf{x}/r$ . Therefore

$$\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} \approx \frac{e^{ikr}}{r} e^{-i\mathbf{k}'\cdot\mathbf{x}'} \quad (521)$$

where  $\mathbf{k}' = k\hat{r}$  is the wave vector of the scattered wave in the direction of observation. Substituting into (519):

$$\psi(\mathbf{x}) \xrightarrow{r \rightarrow \infty} e^{i\mathbf{k}\cdot\mathbf{x}} - \frac{2m}{4\pi\hbar^2} \frac{e^{ikr}}{r} \int d^3x' e^{-i\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \psi(\mathbf{x}') \quad (522)$$

Comparing with (509), we identify the scattering amplitude:

$$f(\theta, \phi) = -\frac{2m}{4\pi\hbar^2} \int d^3x' e^{-i\mathbf{k}' \cdot \mathbf{x}'} V(\mathbf{x}') \psi(\mathbf{x}') \quad (523)$$

or, in operator notation,

$$f(\mathbf{k}' \leftarrow \mathbf{k}) = -\frac{2m}{4\pi\hbar^2} \langle \mathbf{k}' | \hat{V} | \psi^{(+)} \rangle \quad (524)$$

This is an *exact* expression. Note, however, that it is not explicit, since the full scattering state  $|\psi^{(+)}\rangle$  appears on the right-hand side and is itself the unknown. To obtain practical results, we need approximation schemes.

### The $T$ -matrix

It is convenient to define the *transition operator* (or  $T$ -matrix) through the relation

$$\hat{V}|\psi^{(+)}\rangle \equiv \hat{T}(E)|\mathbf{k}\rangle \quad (525)$$

so that

$$f(\mathbf{k}' \leftarrow \mathbf{k}) = -\frac{2m}{4\pi\hbar^2} \langle \mathbf{k}' | \hat{T}(E) | \mathbf{k} \rangle \quad (526)$$

From the Lippmann-Schwinger equation (514), multiplying both sides by  $\hat{V}$ :

$$\hat{T}|\mathbf{k}\rangle = \hat{V}|\mathbf{k}\rangle + \hat{V}\hat{G}_0^{(+)}\hat{T}|\mathbf{k}\rangle \quad (527)$$

Since this holds for arbitrary  $|\mathbf{k}\rangle$ , the  $T$ -matrix satisfies the operator equation

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0^{(+)}\hat{T} \quad (528)$$

This is a reformulation of the Lippmann-Schwinger equation for the  $T$ -matrix. Iterating:

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0^{(+)}\hat{V} + \hat{V}\hat{G}_0^{(+)}\hat{V}\hat{G}_0^{(+)}\hat{V} + \dots \quad (529)$$

This is the *Born series*. Each term describes scattering with an increasing number of interactions with the potential: single scattering, double scattering, and so on. The Born series converges when the potential is sufficiently weak or the energy sufficiently high.

## 7.3 The Born approximation

### First Born approximation

The simplest approximation is to truncate the Born series (529) at the first term:  $\hat{T} \approx \hat{V}$ , or equivalently, to replace the exact wave function  $\psi(\mathbf{x}')$  in (523) by the incident plane wave  $e^{i\mathbf{k} \cdot \mathbf{x}'}$ . This gives the *first Born approximation*:

$$f^{(1)}(\theta, \phi) = -\frac{2m}{4\pi\hbar^2} \int d^3x' e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}'} V(\mathbf{x}') = -\frac{2m}{4\pi\hbar^2} \tilde{V}(\mathbf{q}) \quad (530)$$

where  $\mathbf{q} = \mathbf{k} - \mathbf{k}'$  is the *momentum transfer* and

$$\tilde{V}(\mathbf{q}) = \int d^3x e^{i\mathbf{q} \cdot \mathbf{x}} V(\mathbf{x}) \quad (531)$$

is the Fourier transform of the potential. The Born approximation thus gives a beautifully simple result: the scattering amplitude is (up to a constant factor) the Fourier transform of the potential evaluated at the momentum transfer.

For elastic scattering,  $|\mathbf{k}| = |\mathbf{k}'| = k$ , and the magnitude of the momentum transfer is

$$q = |\mathbf{q}| = 2k \sin \frac{\theta}{2} \quad (532)$$

where  $\theta$  is the scattering angle. This follows from the geometry:  $\mathbf{k}$  and  $\mathbf{k}'$  have the same length  $k$ , and  $\mathbf{q}$  is the base of the isosceles triangle they form.

### Born approximation for central potentials

If  $V(\mathbf{x}) = V(r)$  depends only on the distance from the center, the Fourier transform can be performed using the angular part explicitly. Setting  $\hat{q}$  along the polar axis:

$$\tilde{V}(q) = \int_0^\infty dr r^2 \int_0^\pi d\vartheta \sin \vartheta \int_0^{2\pi} d\varphi e^{iqr \cos \vartheta} V(r) \quad (533)$$

The angular integration gives  $\int_0^\pi d\vartheta \sin \vartheta e^{iqr \cos \vartheta} = \frac{2 \sin(qr)}{qr}$ , so that

$$\tilde{V}(q) = \frac{4\pi}{q} \int_0^\infty dr r \sin(qr) V(r) \quad (534)$$

and therefore

$$f^{(1)}(\theta) = -\frac{2m}{\hbar^2 q} \int_0^\infty dr r \sin(qr) V(r) \quad (535)$$

Note that  $f$  depends only on  $\theta$  (not on  $\phi$ ) for a central potential, as expected from azimuthal symmetry.

### Example: Yukawa potential

The Yukawa potential

$$V(r) = V_0 \frac{e^{-\mu r}}{r} \quad (536)$$

(with  $\mu > 0$ ) describes a screened Coulomb interaction with range  $\mu^{-1}$ . It was originally introduced by Yukawa (1935) to model the nuclear force mediated by pion exchange, with  $\mu = m_\pi c/\hbar$ .

Substituting into (535):

$$f^{(1)}(\theta) = -\frac{2mV_0}{\hbar^2 q} \int_0^\infty dr \sin(qr) e^{-\mu r} \quad (537)$$

The integral is elementary:  $\int_0^\infty dr \sin(qr) e^{-\mu r} = \frac{q}{\mu^2 + q^2}$ . Therefore

$$f^{(1)}(\theta) = -\frac{2mV_0}{\hbar^2(\mu^2 + q^2)} = -\frac{2mV_0}{\hbar^2(\mu^2 + 4k^2 \sin^2 \frac{\theta}{2})} \quad (538)$$

and the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{4m^2 V_0^2}{\hbar^4 (\mu^2 + 4k^2 \sin^2 \frac{\theta}{2})^2} \quad (539)$$

### Example: Coulomb potential (Rutherford scattering)

The Coulomb potential between two charges  $Z_1e$  and  $Z_2e$  is obtained from the Yukawa result in the limit  $\mu \rightarrow 0$ , with  $V_0 = Z_1Z_2e^2/(4\pi\epsilon_0)$ :

$$f_{\text{Coul}}^{(1)}(\theta) = -\frac{2mZ_1Z_2e^2}{4\pi\epsilon_0\hbar^2} \frac{1}{4k^2 \sin^2 \frac{\theta}{2}} = -\frac{a}{2k \sin^2 \frac{\theta}{2}} \quad (540)$$

where  $a = Z_1Z_2e^2m/(4\pi\epsilon_0\hbar^2k)$  is a dimensionless parameter related to the ratio of the Coulomb energy to the kinetic energy. The differential cross section is

$$\boxed{\left(\frac{d\sigma}{d\Omega}\right)_{\text{Ruth}} = \left(\frac{Z_1Z_2e^2}{16\pi\epsilon_0E}\right)^2 \frac{1}{\sin^4 \frac{\theta}{2}}} \quad (541)$$

where  $E = \hbar^2k^2/2m$  is the center-of-mass kinetic energy.

This is the *Rutherford formula*, originally derived by Rutherford (1911) using classical mechanics. It is a remarkable coincidence of quantum mechanics that the first Born approximation reproduces the exact classical result for the Coulomb potential. The exact quantum mechanical calculation (using Coulomb wave functions) gives the same angular dependence, differing only by an overall phase in the scattering amplitude — which does not affect the cross section.

The  $1/\sin^4(\theta/2)$  dependence implies strong forward scattering ( $\theta \rightarrow 0$ ) and a divergent total cross section. Both features reflect the infinite range of the Coulomb potential: even very distant particles are deflected by a small angle. For the screened Yukawa potential (539), the forward divergence is regularized by the screening parameter  $\mu$ .

### Second Born approximation

The second term in the Born series (529) gives the second Born approximation:

$$f^{(2)}(\mathbf{k}' \leftarrow \mathbf{k}) = -\frac{2m}{4\pi\hbar^2} \langle \mathbf{k}' | \hat{V} \hat{G}_0^{(+)} \hat{V} | \mathbf{k} \rangle \quad (542)$$

Inserting a complete set of momentum eigenstates:

$$f^{(2)}(\mathbf{k}' \leftarrow \mathbf{k}) = -\frac{2m}{4\pi\hbar^2} \int \frac{d^3p}{(2\pi)^3} \frac{\tilde{V}(\mathbf{k}' - \mathbf{p}) \tilde{V}(\mathbf{p} - \mathbf{k})}{E - \hbar^2p^2/2m + i\epsilon} \quad (543)$$

The physical interpretation is clear: the particle first scatters from initial momentum  $\hbar\mathbf{k}$  to an intermediate momentum  $\hbar\mathbf{p}$  (acquiring a factor  $\tilde{V}(\mathbf{p} - \mathbf{k})$ ), propagates freely with the Green's function  $(E - \hbar^2p^2/2m + i\epsilon)^{-1}$ , and then scatters again from  $\hbar\mathbf{p}$  to the final momentum  $\hbar\mathbf{k}'$  (acquiring a factor  $\tilde{V}(\mathbf{k}' - \mathbf{p})$ ). The integral over  $\mathbf{p}$  sums over all intermediate momenta.

This “multiple scattering” interpretation extends to all higher orders in the Born series and is the precursor of the Feynman diagrammatic approach in quantum field theory.

### Validity of the Born approximation

The Born approximation is reliable when the scattered wave is much weaker than the incident wave. A sufficient condition can be estimated by demanding that  $|f^{(1)}|$  be small compared to the “natural” length scale  $1/k$  (i.e., the de Broglie wavelength divided by  $2\pi$ ). For a central potential of characteristic strength  $V_0$  and range  $R$ , this gives two regimes:

**High-energy limit** ( $kR \gg 1$ ):

$$\frac{2m|V_0|R}{\hbar^2 k} \ll 1 \iff |V_0| \ll E \quad (544)$$

The potential energy must be small compared to the kinetic energy. At high energies, even strong potentials can be treated perturbatively.

**Low-energy limit** ( $kR \ll 1$ ):

$$\frac{2m|V_0|R^2}{\hbar^2} \ll 1 \quad (545)$$

This is a condition on the potential alone, independent of energy. It says that the potential must be too weak to support a bound state (or, more precisely, that the dimensionless coupling  $2mV_0R^2/\hbar^2$  must be small).

## 7.4 Partial wave analysis

The Born approximation is perturbative in the potential. For strong potentials at low energies, an alternative approach is needed. The method of *partial waves* is exact and exploits the rotational symmetry of central potentials to decompose the scattering problem into independent one-dimensional problems, one for each angular momentum channel.

### Partial wave expansion of the scattering amplitude

For a central potential  $V(r)$ , the scattering amplitude depends only on the scattering angle  $\theta$  (azimuthal symmetry). It can be expanded in Legendre polynomials:

$$f(\theta) = \sum_{\ell=0}^{\infty} (2\ell+1) f_{\ell}(k) P_{\ell}(\cos \theta) \quad (546)$$

where  $f_{\ell}(k)$  is the *partial wave amplitude* for orbital angular momentum  $\ell$ .

The incident plane wave  $e^{ikz}$  has a well-known expansion in terms of spherical Bessel functions  $j_{\ell}(kr)$  and Legendre polynomials (the Rayleigh formula):

$$e^{ikz} = e^{ikr \cos \theta} = \sum_{\ell=0}^{\infty} i^{\ell} (2\ell+1) j_{\ell}(kr) P_{\ell}(\cos \theta) \quad (547)$$

where  $j_{\ell}(x)$  is the spherical Bessel function of the first kind. At large  $r$ :

$$j_{\ell}(kr) \xrightarrow{kr \gg 1} \frac{\sin(kr - \ell\pi/2)}{kr} = \frac{1}{2ikr} (e^{i(kr - \ell\pi/2)} - e^{-i(kr - \ell\pi/2)}) \quad (548)$$

This can be interpreted as a superposition of an outgoing wave  $e^{ikr}/r$  and an incoming wave  $e^{-ikr}/r$  with equal amplitudes.

### Phase shifts

For the full scattering solution, each partial wave is modified by the potential. Far from the scattering region, the radial wave function for the  $\ell$ -th partial wave must take the general form

$$R_{\ell}(r) \xrightarrow{r \rightarrow \infty} \frac{A_{\ell}}{r} \sin \left( kr - \frac{\ell\pi}{2} + \delta_{\ell} \right) \quad (549)$$

where  $\delta_\ell(k)$  is the *phase shift*. Compared to the free-particle solution (548), the effect of the potential is to shift the phase of the outgoing wave relative to the incoming wave by an amount  $2\delta_\ell$ .

The scattering amplitude is, by definition, the coefficient of the outgoing spherical wave after subtracting the incident wave. Comparing the asymptotic expansions, one finds:

$$f_\ell(k) = \frac{e^{2i\delta_\ell} - 1}{2ik} = \frac{e^{i\delta_\ell} \sin \delta_\ell}{k} \quad (550)$$

so that

$$f(\theta) = \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1) e^{i\delta_\ell} \sin \delta_\ell P_\ell(\cos \theta) \quad (551)$$

The phase shift  $\delta_\ell$  contains all the physical information about scattering in the  $\ell$ -th partial wave. Computing  $\delta_\ell(k)$  requires solving the radial Schrödinger equation for the given potential and matching the solution to the asymptotic form (549).

### Partial wave cross sections

The differential cross section is obtained from (512) and (551). The total cross section involves the integral  $\int d\Omega |f(\theta)|^2$ . Using the orthogonality of Legendre polynomials,  $\int_{-1}^1 d(\cos \theta) P_\ell(\cos \theta) P_{\ell'}(\cos \theta) = \frac{2}{2\ell+1} \delta_{\ell\ell'}$ , the cross terms vanish and we obtain

$$\sigma_{\text{tot}} = \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} (2\ell + 1) \sin^2 \delta_\ell \quad (552)$$

Each partial wave contributes independently to the total cross section. The maximum contribution of a single partial wave occurs when  $\delta_\ell = \pi/2 + n\pi$  (with  $n$  integer), giving  $\sigma_\ell = \frac{4\pi(2\ell+1)}{k^2}$ . This is the *unitarity bound* for the  $\ell$ -th partial wave.

### Low-energy scattering and the *s*-wave

At low energies ( $kR \ll 1$ , where  $R$  is the range of the potential), the centrifugal barrier  $\ell(\ell+1)\hbar^2/(2mr^2)$  suppresses the contribution of higher partial waves: classically, a particle with impact parameter  $b > R$  has angular momentum  $\ell \sim kb$  and misses the target. More precisely, for  $\ell > kR$  the particle cannot penetrate the centrifugal barrier and the phase shift is exponentially small. Therefore, at low energies, only the  $\ell = 0$  (*s*-wave) partial wave contributes significantly:

$$f(\theta) \approx f_0 = \frac{e^{i\delta_0} \sin \delta_0}{k} \xrightarrow{k \rightarrow 0} -a_s \quad (553)$$

where  $a_s = -\lim_{k \rightarrow 0} \frac{\delta_0(k)}{k}$  is the *scattering length*. In this regime:

$$\sigma_{\text{tot}} \approx 4\pi a_s^2 \quad (554)$$

The cross section is isotropic (no angular dependence) and determined by a single parameter  $a_s$ . The scattering length encodes whether the potential is attractive ( $a_s > 0$  for a potential that almost supports a bound state) or repulsive ( $a_s < 0$ ). In atomic and nuclear physics,  $a_s$  is one of the most fundamental characterizations of a short-range interaction.

### Example: hard sphere scattering

For a hard sphere of radius  $R$  (i.e.,  $V(r) = \infty$  for  $r < R$  and  $V(r) = 0$  for  $r > R$ ), the boundary condition is  $R_\ell(R) = 0$  for all  $\ell$ . The radial solution for  $r > R$  is a linear combination of  $j_\ell(kr)$  and  $n_\ell(kr)$  (spherical Neumann function). Imposing the boundary condition at  $r = R$ :

$$\tan \delta_\ell = \frac{j_\ell(kR)}{n_\ell(kR)} \quad (555)$$

In the low-energy limit ( $kR \ll 1$ ), using  $j_\ell(x) \sim \frac{x^\ell}{(2\ell+1)!!}$  and  $n_\ell(x) \sim -\frac{(2\ell-1)!!}{x^{\ell+1}}$ :

$$\tan \delta_\ell \approx -\frac{(kR)^{2\ell+1}}{(2\ell+1)[(2\ell-1)!!]^2} \quad (556)$$

This confirms that higher partial waves are suppressed as  $(kR)^{2\ell+1}$ . For the  $s$ -wave:

$$\delta_0 \approx -kR \implies a_s = R \quad (557)$$

and  $\sigma_{\text{tot}} \approx 4\pi R^2$ . This is four times the classical geometric cross section  $\pi R^2$  — a purely wave-mechanical effect due to diffraction around the sphere.

## 7.5 The optical theorem

There is a deep relation between the total cross section and the forward scattering amplitude, known as the *optical theorem*. It is a direct consequence of probability conservation (unitarity).

### Statement and proof

**Theorem (Optical theorem).**

$$\boxed{\sigma_{\text{tot}} = \frac{4\pi}{k} \text{Im } f(\theta = 0)} \quad (558)$$

*Proof.* From the partial wave expansion (551), the forward scattering amplitude is

$$f(0) = \frac{1}{k} \sum_{\ell} (2\ell+1) e^{i\delta_\ell} \sin \delta_\ell \quad (559)$$

since  $P_\ell(1) = 1$  for all  $\ell$ . Taking the imaginary part:

$$\text{Im } f(0) = \frac{1}{k} \sum_{\ell} (2\ell+1) \sin^2 \delta_\ell \quad (560)$$

Comparing with (552):

$$\sigma_{\text{tot}} = \frac{4\pi}{k^2} \sum_{\ell} (2\ell+1) \sin^2 \delta_\ell = \frac{4\pi}{k} \text{Im } f(0) \quad (561)$$

□

The optical theorem has a compelling physical interpretation. The total cross section measures the total flux removed from the incident beam by scattering. This flux must be extracted from the forward direction — it manifests as destructive interference between the incident wave and the forward-scattered wave. The imaginary part of  $f(0)$  quantifies precisely this interference, and the optical theorem states that it accounts for all the scattered flux, regardless of the angular distribution.

The optical theorem is not restricted to potential scattering; it holds in any theory where the  $S$ -matrix is unitary, including relativistic quantum field theory. It is one of the most general and powerful constraints in scattering theory.

## 7.6 The $S$ -matrix

We now develop the  $S$ -matrix formalism, which provides the most fundamental and elegant description of scattering. The  $S$ -matrix unifies the concepts developed above and reveals the role of unitarity as the central structural principle.

### Definition

The  $S$ -matrix (scattering matrix) is the operator that relates the asymptotic “in” state of the system (the state it was prepared in, long before the interaction) to the asymptotic “out” state (the state it is detected in, long after the interaction):

$$|\text{out}\rangle = \hat{S} |\text{in}\rangle \quad (562)$$

For our scattering problem, the “in” state is the incident plane wave  $|\mathbf{k}\rangle$  and the “out” state is the full scattered state  $|\psi^{(+)}\rangle$  in the asymptotic future. The matrix element

$$S_{\mathbf{k}'\mathbf{k}} = \langle \mathbf{k}' | \hat{S} | \mathbf{k} \rangle \quad (563)$$

gives the amplitude for the transition  $\mathbf{k} \rightarrow \mathbf{k}'$ .

In the absence of scattering ( $V = 0$ ), the  $S$ -matrix is the identity. We separate this trivial part by writing

$$\hat{S} = \mathbb{1} + i\hat{R} \quad (564)$$

where  $\hat{R}$  encodes the non-trivial scattering. The matrix element of  $\hat{R}$  is related to the scattering amplitude by

$$\langle \mathbf{k}' | \hat{R} | \mathbf{k} \rangle = -\frac{2\pi i}{k} f(\mathbf{k}' \leftarrow \mathbf{k}) \delta(E_{k'} - E_k) \quad (565)$$

where the energy-conserving delta function reflects the fact that elastic scattering preserves energy.

### Unitarity

The  $S$ -matrix must be *unitary*:

$$\hat{S}^\dagger \hat{S} = \hat{S} \hat{S}^\dagger = \mathbb{1} \quad (566)$$

This is a direct consequence of probability conservation: the total probability of scattering into *some* final state must equal the total initial probability, which is 1. In the language of quantum mechanics, unitary time evolution preserves the norm of states.

In terms of the partial wave decomposition, the  $S$ -matrix is diagonal in the angular momentum basis:

$$S_\ell = e^{2i\delta_\ell} \quad (567)$$

The unitarity condition  $|S_\ell|^2 = 1$  is automatically satisfied, since  $|e^{2i\delta_\ell}| = 1$  for real phase shifts. The partial wave amplitude (550) can be rewritten as

$$f_\ell = \frac{S_\ell - 1}{2ik} \quad (568)$$

The constraint  $|S_\ell| = 1$  is the origin of the unitarity bound  $\text{Im } f_\ell \leq 1/k$ , and the optical theorem (558) is the “off-diagonal” expression of this same unitarity constraint.

## Unitarity with absorption

In some physical situations, the scattering potential can absorb particles (for instance, nuclear reactions that convert the incident particle into different products). In this case, the outgoing flux in the elastic channel is less than the incoming flux, and the partial wave  $S$ -matrix element satisfies  $|S_\ell| \leq 1$ . We can parametrize this as

$$S_\ell = \eta_\ell e^{2i\delta_\ell}, \quad 0 \leq \eta_\ell \leq 1 \quad (569)$$

where  $\eta_\ell$  is the *inelasticity parameter*. The elastic cross section is

$$\sigma_{\text{el}} = \frac{\pi}{k^2} \sum_{\ell} (2\ell + 1) |1 - \eta_\ell e^{2i\delta_\ell}|^2 \quad (570)$$

and the *reaction* (absorption) cross section is

$$\sigma_{\text{reac}} = \frac{\pi}{k^2} \sum_{\ell} (2\ell + 1) (1 - \eta_\ell^2) \quad (571)$$

The total cross section  $\sigma_{\text{tot}} = \sigma_{\text{el}} + \sigma_{\text{reac}}$  still satisfies the optical theorem (558).

Maximum absorption occurs when  $\eta_\ell = 0$  (“black disk”), giving  $\sigma_{\text{el}} = \sigma_{\text{reac}} = \frac{\pi(2\ell+1)}{k^2}$ . The fact that the elastic cross section is nonzero even for total absorption is a wave-mechanical effect: the removal of the  $\ell$ -th partial wave from the forward direction produces a “shadow” that is itself a scattering phenomenon.

## 7.7 Connection with quantum information

We close this chapter by highlighting the deep conceptual connections between scattering theory and quantum information.

### Unitarity as conservation of information

The unitarity of the  $S$ -matrix,  $\hat{S}^\dagger \hat{S} = \mathbb{1}$ , is the statement that scattering is a *reversible* transformation: no information is lost in the scattering process. Given the final state, the initial state can in principle be reconstructed (by applying  $\hat{S}^\dagger$ ). In the language of Chapter 2, the  $S$ -matrix implements a unitary quantum channel — the most special type of channel, with zero entropy production.

This principle has profound implications beyond non-relativistic scattering:

- In quantum field theory, the unitarity of the  $S$ -matrix constrains the allowed interactions and leads to the optical theorem, dispersion relations, and the Froissart bound on high-energy cross sections.
- In quantum gravity, the question of whether black hole evaporation preserves unitarity (the *black hole information paradox*, raised by Hawking in 1975) has been one of the central problems in theoretical physics for half a century. The AdS/CFT correspondence provides strong evidence that unitarity is indeed preserved.

### The $S$ -matrix as a quantum gate

In the partial wave basis, the  $S$ -matrix for each  $\ell$  is a phase gate:

$$S_\ell = e^{2i\delta_\ell} \quad (572)$$

This is precisely a single-qubit phase gate  $R_z(2\delta_\ell)$  acting on the “partial wave” degree of freedom. The full  $S$ -matrix, acting on all partial waves simultaneously, can be viewed as a diagonal unitary matrix in the angular momentum basis — a collection of independent phase gates.

In this perspective, a scattering experiment is a quantum computation: the initial state (a plane wave, which is a superposition of angular momentum eigenstates by the Rayleigh expansion (547)) is processed by the  $S$ -matrix “circuit” (which applies a different phase to each angular momentum component), and the output state is measured (the angular distribution of scattered particles is the outcome probability distribution).

## Scattering and quantum channels

When absorption is present ( $\eta_\ell < 1$ ), the elastic scattering channel alone does not preserve probability: the  $S$ -matrix (569) is no longer unitary. From the quantum information perspective, this means that if we trace over the reaction products (the “environment”), the elastic scattering is described by a non-unitary quantum channel  $\mathcal{E}(\rho) = \sum_n M_n \rho M_n^\dagger$ , where the Kraus operators encode both the elastic phase shifts and the absorption. The full process (elastic + reaction channels together) remains unitary — information is not lost, but rather transferred to the reaction products.

This is completely analogous to the discussion of open quantum systems in Chapter 2: the closed system (all scattering channels) evolves unitarily, but the reduced description of a subsystem (the elastic channel) involves a quantum channel with possible entropy increase.

## Scattering and entanglement

The scattering of two distinguishable particles can generate entanglement between them. Consider two particles, initially in a product state  $|\mathbf{k}_1\rangle \otimes |\mathbf{k}_2\rangle$ , that interact through a potential and scatter. The final state, obtained by applying the two-body  $S$ -matrix, is generally an entangled superposition of different momenta:

$$\hat{S} |\mathbf{k}_1, \mathbf{k}_2\rangle = \sum_{\mathbf{k}'_1, \mathbf{k}'_2} S_{\mathbf{k}'_1 \mathbf{k}'_2; \mathbf{k}_1 \mathbf{k}_2} |\mathbf{k}'_1, \mathbf{k}'_2\rangle \quad (573)$$

The amount of entanglement generated depends on the scattering amplitude and can be quantified by the von Neumann entropy of the reduced density matrix of one particle. Scattering-generated entanglement has been studied in the context of quantum communication protocols and as a probe of the entanglement structure of quantum field theories.

These connections illustrate a broader theme: scattering theory and quantum information theory are two faces of the same underlying structure — the unitary evolution of quantum states. The tools developed in each field illuminate the other, and their intersection is an active area of contemporary research.

## References

- [1] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics, 3rd edition*. Cambridge University Press, 2020.
- [2] John Preskill. *Lecture Notes on Quantum Information*. URL: <http://theory.caltech.edu/~preskill/ph219/>.
- [3] John Von Neumann. *Mathematical foundations of quantum mechanics: New edition*. Vol. 53. Princeton university press, 2018.
- [4] Rudolf Haag. *Local quantum physics: Fields, particles, algebras*. Springer Science & Business Media, 2012.
- [5] Paul Adrien Maurice Dirac. *The principles of quantum mechanics*. 27. Oxford university press, 1981.
- [6] Xiao-Gang Wen. *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons*. OUP Oxford, 2004. ISBN: 0191523968, 9780191523960.

- [7] Michael V. Berry. “Quantal phase factors accompanying adiabatic changes”. In: *Proc. Roy. Soc. Lond. A* 392 (1984), pp. 45–57. DOI: 10.1098/rspa.1984.0023.
- [8] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. ISBN: 9780521635035.
- [9] Steven Weinberg. *The Quantum Theory of Fields. Volume 1: Foundations*. Cambridge University Press, 1995. ISBN: 9780521670531.
- [10] Valentine Bargmann. “Note on Wigner’s Theorem on Symmetry Operations”. In: *Journal of Mathematical Physics* 5.7 (1964), pp. 862–868. DOI: 10.1063/1.1704188.
- [11] H. Rauch, A. Zeilinger, G. Badurek, A. Wilfing, W. Bauspiess, and U. Bonse. “Verification of Coherent Spinor Rotation of Fermions”. In: *Physics Letters A* 54.6 (1975), pp. 425–427. DOI: 10.1016/0375-9601(75)90798-7.
- [12] Asher Peres. “Separability Criterion for Density Matrices”. In: *Physical Review Letters* 77.8 (1996), pp. 1413–1415. DOI: 10.1103/PhysRevLett.77.1413.
- [13] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. “Separability of Mixed States: Necessary and Sufficient Conditions”. In: *Physics Letters A* 223.1–2 (1996), pp. 1–8. DOI: 10.1016/S0375-9601(96)00706-2.
- [14] Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë. *Quantum Mechanics, Volume 1: Basic Concepts, Tools, and Applications*. 2nd ed. Wiley-VCH, 2020.
- [15] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. 3rd ed. Cambridge University Press, 2018. ISBN: 9781107189638.
- [16] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. “A variational eigenvalue solver on a photonic quantum processor”. In: *Nature Communications* 5 (2014), p. 4213. DOI: 10.1038/ncomms5213.
- [17] Leslie E. Ballentine. *Quantum Mechanics: A Modern Development*. 2nd ed. World Scientific, 2014. DOI: 10.1142/9038.