
Advanced Quantum Mechanics

Instructor: Kangqiao LIU

Email: kqliu@xhu.edu.cn

Personal page: <https://kangqiaoliu.github.io/>

These lecture notes are accompanied by [interactive exercises and guided reading](#).

© 2026 Kangqiao LIU. Eligible original teaching contributions: CC BY-NC-SA 4.0; third-party rights remain with their holders. See [copyright and licensing](#).

Contents

List of Notation	iii
1 Basic Concepts	1
1.1 Stern–Gerlach Experiment	1
1.2 Linear Vector Spaces	3
1.3 Eigenvalues and Eigenvectors of Operators (Matrices)	6
1.4 Beyond Hermiticity: Loss, Complex Spectra, and Exceptional Points	9
1.5 Measurement	14
1.6 Transformation of Representations	21
1.7 Position and Momentum Representations	22
2 Quantum Dynamics	29
2.1 Time Evolution and the Schrödinger Equation	29
2.2 Schrödinger and Heisenberg Pictures	34
2.3 One-Dimensional Linear Harmonic Oscillator	37
2.4 Coherent States of the Harmonic Oscillator	42
2.5 Squeezed States	44
2.6 Schrödinger Equation in the Coordinate Representation and the WKB Approximation	45
2.7 Feynman Path Integral	48
2.8 Gauge Transformations	52
2.9 Berry Phase	57
2.10 Magnetic Monopoles and Charge Quantization	59
3 Angular Momentum Theory	62
3.1 Rotations and Angular Momentum	62
3.2 Spin Angular Momentum	64
3.3 Ensembles and the Density Operator	68
3.4 Composite Systems and Generalized Quantum Measurements	72
3.5 Eigenvalues and Eigenstates of Angular Momentum	76
3.6 Orbital Angular Momentum	81
3.7 Coupling of Two Angular Momenta	83
3.8 Spin Correlations and Bell’s Inequality	88
3.9 Quantum Teleportation: An Application of Entanglement and Measurement	91
4 Approximation Methods	94
4.1 The Idea of Stationary Perturbation Theory	94
4.2 Nondegenerate Perturbation Theory	95
4.3 Degenerate Perturbation Theory	99
4.4 Variational Method	103
4.5 Strong-Coupling Schrödinger Equation	105

4.6	Time-Dependent Perturbation Theory	107
4.7	Interaction Picture and Perturbation Theory	110
5	Scattering Theory	113
5.1	General Description	113
5.2	Green-Function Method	114
5.3	Lippmann–Schwinger Equation	117
5.4	Born Series	121
5.5	Partial-Wave Method	124
6	Quantum Many-Body Systems	129
6.1	Identical Particles and Permutation Symmetry	129
6.2	Completely Symmetric and Antisymmetric States	131
6.3	Bosons: Fock Space and Creation and Annihilation Operators	132
6.4	Fermions: Antisymmetric States and Operator Algebra	135
7	Relativistic Quantum Mechanics	139
7.1	From the Schrödinger Equation to the Pauli Equation	139
7.2	From Relativistic Energy to the Dirac Equation	140
7.3	The Nonrelativistic Limit of the Dirac Equation	142
7.4	The Relation among the Schrödinger, Pauli, and Dirac Equations	143
7.5	Free Dirac Spinors	143
8	Topological Condensed Matter Physics	145
8.1	From a Periodic Chain to Energy Bands	145
8.2	The Su–Schrieffer–Heeger Chain	146
8.3	Boundary States and Bulk–Boundary Correspondence	149
8.4	Berry Curvature and the Chern Number	150
8.5	Other Topological Phases and Present Perspective	152
8.6	Geometric Phases in a Periodic Lattice	152
8.7	Charge Transport from a Slowly Varying Hamiltonian	153
8.8	Further Worked Checks	154
9	Quantum Neural Networks and Machine Learning	155
9.1	What Is Machine Learning?	155
9.2	Qubits, Quantum Gates, and Readout	156
9.3	A Quantum Neural Network as a Parameterized Circuit	156
9.4	How Quantum Circuits Are Trained	158
9.5	Quantum Kernel Methods	160
9.6	Optimization, Trainability, and Practical Limits	162
9.7	Beyond the Introductory Models	164
9.8	Worked Training and Noise Examples	164
9.9	Learning from Quantum States	165
9.10	Statistical Concentration and a Trainability Check	167
	Bibliography	168

List of Notation

Throughout these notes, *hats* denote abstract operators, *bold symbols* denote spatial vectors, and a dagger denotes the Hermitian adjoint. Matrices in a specified basis may omit hats. The superscript ^{T} denotes ordinary matrix transpose; † includes complex conjugation. The symbol $\mathbf{1}$ always denotes the identity (with a subsystem subscript when needed). Subscripts identify indices, states, or representations as indicated; local parameters are defined where introduced.

Notation	Meaning and global convention
$ \psi\rangle, \langle\psi $	State ket and its dual bra.
$\langle\phi \psi\rangle, \langle\phi \hat{A} \psi\rangle$	Inner product; operator matrix element.
$\hat{A}, A_{mn} = \langle m \hat{A} n\rangle$	Abstract operator; its matrix entries in a basis.
$\hat{A}^\dagger, \mathbf{1}$	Hermitian adjoint; identity operator.
$[A, B], \{A, B\}$	Commutator $AB - BA$; anticommutator $AB + BA$.
$\langle A \rangle_\psi, \langle A \rangle_\rho$	Expectations in a pure state and a mixed state.
$\text{tr}(A)$	Trace, independent of basis.
$\delta_{mn}, \delta^{(3)}(\mathbf{r})$	Kronecker and three-dimensional Dirac deltas.
$i, \hbar, c$	Imaginary unit; reduced Planck constant; light speed.
$\mathbf{r}, \mathbf{p}; \hat{\mathbf{r}}, \hat{\mathbf{p}}$	Position and momentum vectors; corresponding operators.
$\psi(\mathbf{r}, t) = \langle \mathbf{r} \psi \rangle$	Position-space wavefunction (time argument understood).
$\hat{H}, \hat{U}(t, t_0)$	Hamiltonian; unitary time evolution.
$\hat{A}_S, \hat{A}_H, \hat{A}_I$	Schrödinger, Heisenberg and interaction pictures.
$\hat{\rho}, \rho(\mathbf{r}, t)$	Density operator; position-space probability density.
ΔA	Standard deviation of observable A .
$\hat{\mathbf{J}}, \hat{\mathbf{L}}, \hat{\mathbf{S}}$	Total, orbital and spin angular momenta.
$ j, m\rangle, \hat{J}_\pm$	Angular-momentum eigenstate; ladder operators.
$\sigma_i, \hat{\mathbf{s}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}$	Pauli matrices; spin- $\frac{1}{2}$ angular momentum.
$\hat{P}_{ij}, \hat{P}_\pm$	Particle exchange; Bose/Fermi symmetry projectors.
$\hat{a}_i, \hat{a}_i^\dagger$	Annihilation and creation in mode i (index optional).
$\hat{n}_i, \hat{N} = \sum_i \hat{n}_i$	Mode occupation and total number operators.
$ n_1, n_2, \dots\rangle$	Fock occupation-number basis for identical particles.
$\mathcal{T}, \hat{T}$	Time ordering; scattering transition operator.
$f(\Omega), d\sigma/d\Omega$	Scattering amplitude; differential cross section.
$\hat{G}_0^\pm, \hat{G}^\pm$	Free and full resolvents; $+$ is outgoing.
$x^\mu, g_{\mu\nu}$	Spacetime four-coordinate; Minkowski metric.
$\gamma^\mu, \bar{\psi} = \psi^\dagger \gamma^0$	Dirac gamma matrices; Dirac adjoint.

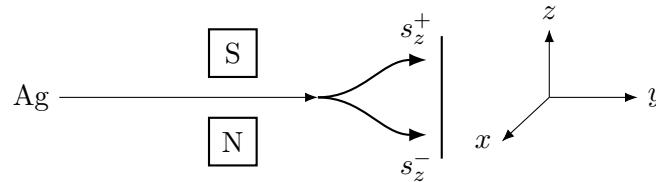
Conventions: σ_i denotes a Pauli matrix, while $\sigma(\Omega)$ may denote a scattering cross section; the index or argument distinguishes them. The relativistic discussion uses $\hbar = c = 1$; factors of c elsewhere follow the units specified in the text.

Chapter 1

Basic Concepts

1.1 Stern–Gerlach Experiment

- The fundamental difference from classical mechanics is readily demonstrated.
- The central issue of quantum mechanics—the measurement problem—is readily demonstrated.
- A two-level system is the most quantum-mechanical system.



1) Results

A heated beam of silver atoms splits into two beams after passing through an inhomogeneous magnetic field.

2) Analysis

- The splitting is caused by interaction with the magnetic field and must therefore be due to the magnetic moment $\mathbf{M}$ of the atoms. The interaction potential is $V = -\mathbf{M} \cdot \mathbf{B}$.
- The magnetic moment is proportional to the angular momentum $\mathbf{J}$: $\mathbf{M} \propto \mathbf{J}$.
- The force on the atoms is

$$\mathbf{F} = -\nabla V = M_z \frac{\partial B_z}{\partial z} \mathbf{e}_z \propto J_z \frac{\partial B_z}{\partial z} \mathbf{e}_z. \quad (1.1.1)$$

The beam separates symmetrically into an upper and a lower component. This indicates that the angular momentum along the direction of the magnetic field (the z direction) has only two components of equal magnitude and opposite directions.

If this angular momentum were produced by the rotation of the atoms themselves, the angular-momentum directions of the thermal atoms would be randomly distributed. After a large number of atoms passed through the magnetic field, their distribution on the screen would be symmetric and continuous, rather than consisting of two separated components. Thus the force does not originate from orbital angular momentum.

A silver atom has 47 electrons. Of these, 46 occupy filled shells with a spherically symmetric distribution and have no net angular momentum. The angular momentum of the silver atom is entirely due to the remaining valence electron. The separated two-component distribution

shows that it is caused by the intrinsic angular momentum of the valence electron, denoted by $\mathbf{s}$. The component s_z has only two values of equal magnitude and opposite direction, s_z^+ and s_z^- .

3) Quantum properties

- Spin angular momentum exists as an intrinsic physical quantity (independent of space and time).
- The values of spin angular momentum are discrete.
- The magnetic field performs a measurement. A magnetic field in the z direction measures the angular momentum in the z direction.

4) Cascaded Stern–Gerlach experiments

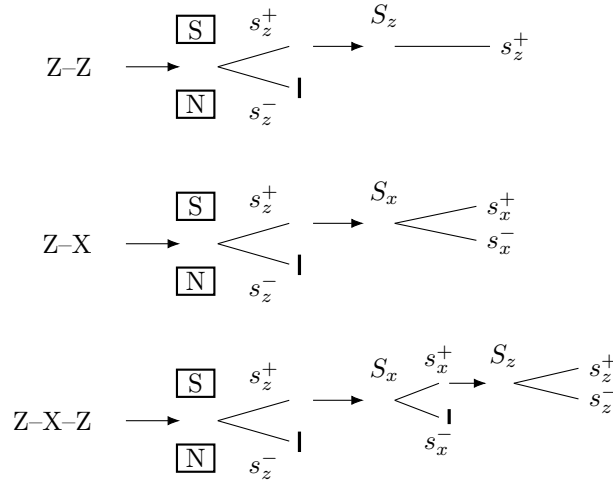


Figure 1.

The incident atomic beam passes successively through two magnetic fields in the z direction, as shown in the upper part of Fig. 1. Before the second magnetic field, s_z has the definite value s_z^+ . Therefore the force on the atoms in that field is definite, and s_z still has the definite value s_z^+ after the second field.

Now let the incident atomic beam pass through two magnetic fields, in the z and x directions, as shown in the middle of Fig. 1. In the second magnetic field the atoms experience a force $\mathbf{F} \propto J_x(\partial B_x/\partial x)\mathbf{e}_x$. The beam splits after the second magnetic field, indicating that s_x has two components s_x^+ and s_x^- before that field, although s_z has the definite value s_z^+ .

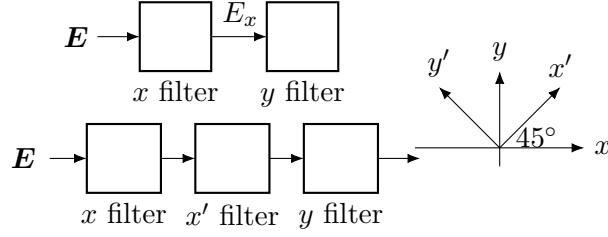
- Quantum property: when s_z has a definite value, s_x does not. s_z and s_x cannot have definite values simultaneously.

Next, let the beam pass successively through three magnetic fields in the z , x , and z directions, as shown in the lower part of Fig. 1. Finally one observes two components s_z^+ and s_z^- . Thus s_z has two components before the third magnetic field, even though s_x has the definite value s_x^+ .

- Quantum property: when s_x has a definite value, s_z likewise does not. s_x and s_z cannot have definite values simultaneously.

Interactive investigation IC-01-03. Open Cascaded Stern–Gerlach. After the three-analyzer discussion, vary the middle analyzer and the selected beam. Compare the joint and conditional detection probabilities, including the case where the final analyzer is aligned with the first.

5) Analogy with classical electromagnetic waves (matter particles and light waves)



An electromagnetic wave propagating along the z direction disappears completely after passing successively through a filter that transmits only the x component (x filter) and a filter that transmits only the y component (y filter):

$$\begin{aligned}
 \mathbf{E}(\mathbf{r}, t) &= E_0(\mathbf{e}_x + \mathbf{e}_y) \cos(kz - \omega t), \\
 \xrightarrow{x \text{ filter}} & E_0 \mathbf{e}_x \cos(kz - \omega t), \\
 \xrightarrow{y \text{ filter}} & 0.
 \end{aligned} \tag{1.1.2}$$

Place an x' filter between the x and y filters, with x' at 45° to both x and y . An electromagnetic wave in the y direction can still be observed at the end:

$$\begin{aligned}
 \mathbf{E}(\mathbf{r}, t) &= E_0(\mathbf{e}_x + \mathbf{e}_y) \cos(kz - \omega t), \\
 \xrightarrow{x \text{ filter}} & E_0 \mathbf{e}_x \cos(kz - \omega t) = \frac{E_0}{\sqrt{2}}(\mathbf{e}_{x'} - \mathbf{e}_{y'}) \cos(kz - \omega t), \\
 \xrightarrow{x' \text{ filter}} & \frac{E_0}{\sqrt{2}} \mathbf{e}_{x'} \cos(kz - \omega t) = \frac{E_0}{2}(\mathbf{e}_x + \mathbf{e}_y) \cos(kz - \omega t), \\
 \xrightarrow{y \text{ filter}} & \frac{E_0}{2} \mathbf{e}_y \cos(kz - \omega t).
 \end{aligned} \tag{1.1.3}$$

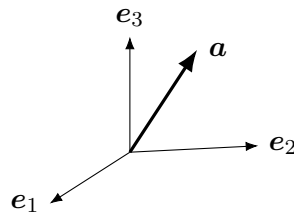
Similarity: s_x, s_y, s_z and $E_{x'}, E_{y'}$ can all be regarded as two-component vectors.

Difference: $\mathbf{s}$ is intrinsic angular momentum, a quantum-mechanical quantity; $\mathbf{E}$ is a spatially dependent mechanical quantity, a classical-mechanical quantity.

1.2 Linear Vector Spaces

From the preceding section, the projection s_n of the electron's spin angular momentum along any direction n can take only two values. We may regard s_n as a two-dimensional vector. To establish the mathematical description of quantum mechanics, we first discuss linear vector spaces.

1) Three-dimensional vector space



Arbitrary vector: $\mathbf{a}$.

Basis vectors: $\mathbf{e}_n$, $n = 1, 2, 3$.

Completeness of the basis:

$$\mathbf{a} = \sum_{n=1}^3 a_n \mathbf{e}_n. \quad (1.2.1)$$

Inner product:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{n,m} a_n b_m \mathbf{e}_n \cdot \mathbf{e}_m. \quad (1.2.2)$$

Squared norm of a vector: $\mathbf{a} \cdot \mathbf{a} \geq 0$.

If the basis vectors are orthonormal, $\mathbf{e}_n \cdot \mathbf{e}_m = \delta_{nm}$, the inner product has the matrix form

$$\mathbf{a} \cdot \mathbf{b} = \sum_n a_n b_n = a^\top b, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad a^\top = (a_1, a_2, a_3), \quad (1.2.3)$$

Here $a^\top$ is the transpose of a column vector; in complex spaces the inner product instead uses the conjugate transpose $a^\dagger = (a^\top)^*$.

Components (matrix elements) of a vector:

$$a_n = \mathbf{e}_n \cdot \mathbf{a}. \quad (1.2.4)$$

The vector $\mathbf{a}$ is its abstract or general form. The matrix a is the representation of $\mathbf{a}$ in a specific coordinate system (representation). Its matrix elements depend on the choice of basis vectors; for example, Cartesian and spherical coordinate systems give different representations.

A linear operator maps vectors to vectors; for example, $\hat{A}\mathbf{a} = \mathbf{b}$, where $\hat{A}$ is generic.

2) Hilbert space

Extend the three-dimensional vector space to a complex vector space of arbitrary dimension:

$$\text{three dimensions} \longrightarrow \text{arbitrary finite, infinite, or continuous dimension}, \quad (1.2.5)$$

$$\text{constant vectors} \longrightarrow \text{complex-function vectors}. \quad (1.2.6)$$

Use Dirac notation (a ket) to denote an arbitrary vector: $|a\rangle$.

For a complex vector, introduce the bra $\langle a|$ to represent its complex-conjugate vector. The bra and ket are not independent; they are conjugate to each other:

$$|a\rangle \longleftrightarrow \langle a|. \quad (1.2.7)$$

A vector can be represented by either a ket $|a\rangle$ or a bra $\langle a|$.

In a complex-function vector space, constants are generally complex, so

$$\alpha |a\rangle \longleftrightarrow \langle a| \alpha^*. \quad (1.2.8)$$

For an operation (operator) $\hat{A}$ in the complex space,

$$\hat{A} |a\rangle \longleftrightarrow \langle a| \hat{A}^\dagger. \quad (1.2.9)$$

$\hat{A}^\dagger$ is the left operator corresponding to the right operator $\hat{A}$, called the Hermitian-conjugate operator. Note that

$$\hat{A}\hat{F} |a\rangle = \hat{A}(\hat{F} |a\rangle) \longleftrightarrow \langle a| \hat{F}^\dagger \hat{A}^\dagger. \quad (1.2.10)$$

We now choose a particular representation, taking an N -dimensional discrete space as an example.

Basis vectors: $|n\rangle$, $n = 1, 2, \dots$

Expansion in a complete basis (concrete representation of a vector):

$$|a\rangle = \sum_n a_n |n\rangle, \quad \langle a| = \sum_n a_n^* \langle n|. \quad (1.2.11)$$

Inner product:

$$\langle a|b\rangle = \sum_{n,m} a_n^* b_m \langle n|m\rangle, \quad (1.2.12)$$

which is a complex number.

Squared norm: $\langle a|a\rangle \geq 0$. (Only when $|a\rangle$ and $\langle a|$ are defined as conjugates can the squared norm be guaranteed nonnegative.)

For a nonzero ket, define a normalized state by

$$|a_{\text{norm}}\rangle = \frac{|a\rangle}{\sqrt{\langle a|a\rangle}}, \quad \langle a_{\text{norm}}|a_{\text{norm}}\rangle = 1. \quad (1.2.13)$$

The subscript *norm* indicates normalization and is unrelated to transpose.

Orthonormality of basis vectors: $\langle n|m\rangle = \delta_{nm}$.

Matrix form of the inner product:

$$\langle a|b\rangle = \sum_n a_n^* b_n = a^\dagger b, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}, \quad (1.2.14)$$

where the Hermitian conjugate (complex-conjugate transpose) of the column matrix a is the row matrix

$$a^\dagger = (a_1^*, a_2^*, \dots, a_N^*). \quad (1.2.15)$$

Evidently,

$$\langle a|b\rangle = (\langle b|a\rangle)^*, \quad \langle a|\hat{A}|b\rangle = (\langle b|\hat{A}^\dagger|a\rangle)^*. \quad (1.2.16)$$

Components (matrix elements) of a vector: multiplying $|a\rangle = \sum_n a_n |n\rangle$ from the left by $\langle m|$ gives

$$a_m = \langle m|a\rangle. \quad (1.2.17)$$

Completeness of the basis:

$$|a\rangle = \sum_n a_n |n\rangle = \sum_n |n\rangle \langle n|a\rangle = \left(\sum_n |n\rangle \langle n| \right) |a\rangle. \quad (1.2.18)$$

Because $|a\rangle$ is arbitrary,

$$\boxed{\sum_n |n\rangle \langle n| = \mathbf{1}}, \quad (1.2.19)$$

which is the completeness relation for the basis.

Representation of an operator (matrix form). Let $|b\rangle = \hat{F}|a\rangle$. Then

$$\langle m|b\rangle = \langle m|\hat{F}|a\rangle = \sum_n \langle m|\hat{F}|n\rangle \langle n|a\rangle, \quad F_{mn} = \langle m|\hat{F}|n\rangle, \quad (1.2.20)$$

and therefore

$$b_m = \sum_n F_{mn} a_n. \quad (1.2.21)$$

In matrix form, $b = Fa$. The matrix F represents the operator $\hat{F}$; it is a square matrix with elements F_{mn} .

Representation of the Hermitian-conjugate operator (left operator). From $\langle b| = \langle a| \hat{F}^\dagger$,

$$\langle b|m\rangle = \langle a|\hat{F}^\dagger|m\rangle = \sum_n \langle a|n\rangle \langle n|\hat{F}^\dagger|m\rangle, \quad (1.2.22)$$

so

$$b_m^* = \sum_n (F^\dagger)_{nm} a_n^*, \quad b_m = \sum_n [(F^\dagger)_{nm}]^* a_n. \quad (1.2.23)$$

Comparison yields

$$[(F^\dagger)_{nm}]^* = F_{mn}, \quad (F^\dagger)_{nm} = F_{mn}^*, \quad F^\dagger = (F^\top)^*. \quad (1.2.24)$$

Thus $F^\dagger$ is the Hermitian-conjugate matrix of F .

Outer product. The outer product is $|a\rangle \langle b|$.

Its representation is a square matrix (a is a column matrix and $b^\dagger$ a row matrix), so the outer product is an operator. In fact,

$$(|a\rangle \langle b|) |c\rangle = |a\rangle (\langle b|c\rangle), \quad (1.2.25)$$

showing that $|a\rangle \langle b|$ maps $|c\rangle$ to another vector parallel to $|a\rangle$. Hence $|a\rangle \langle b|$ is indeed an operator. Its representation is a square matrix with elements

$$(|a\rangle \langle b|)_{mn} = \langle m||a\rangle \langle b|n\rangle = \langle m|a\rangle \langle b|n\rangle = a_m b_n^*. \quad (1.2.26)$$

Interactive investigation IC-01-01. Open Complex amplitudes. After learning inner products and normalization, add two complex amplitudes before taking their squared modulus. Change their relative phase and check how interference differs from adding probabilities.

1.3 Eigenvalues and Eigenvectors of Operators (Matrices)

1) Eigenvalues and eigenstates of a general operator

The eigenvalue equation of an operator is

$$\hat{A} |a\rangle = \lambda |a\rangle, \quad (1.3.1)$$

where λ is called an eigenvalue of $\hat{A}$ and $|a\rangle$ is called an eigenvector of $\hat{A}$.

Use the completeness relation to derive the component equation. Write $|a\rangle = \sum_n a_n |n\rangle$ and multiply $\hat{A} |a\rangle = \lambda |a\rangle$ by $\langle m|$:

$$\langle m|\hat{A}|a\rangle = \sum_n a_n \langle m|\hat{A}|n\rangle = \lambda \sum_n a_n \langle m|n\rangle = \lambda a_m. \quad (1.3.2)$$

Thus $\sum_n A_{mn} a_n = \lambda a_m$ for each m , or,

$$Aa = \lambda a, \quad (A - \lambda \mathbf{1})a = 0. \quad (1.3.3)$$

The condition for a nonzero eigenvector $a \neq 0$ is

$$\det(A - \lambda \mathbf{1}) = 0. \quad (1.3.4)$$

This is the secular equation in an N -dimensional space:

$$\begin{vmatrix} A_{11} - \lambda & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} - \lambda & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} - \lambda \end{vmatrix} = 0. \quad (1.3.5)$$

One thereby obtains N eigenvalues λ_i ($i = 1, 2, \dots, N$). Substituting each into

$$(A - \lambda_i \mathbf{1})a = 0 \quad (1.3.6)$$

gives the corresponding eigenvector a .

Example. Find the eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.3.7)$$

The secular equation is

$$\begin{vmatrix} 1 - \lambda & 0 \\ 0 & -1 - \lambda \end{vmatrix} = 0, \quad \text{i.e.,} \quad \lambda^2 - 1 = 0, \quad (1.3.8)$$

and the eigenvalues are $\lambda_{\pm} = \pm 1$.

Let $a = (a_1, a_2)^T$. For $\lambda = 1$,

$$\begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0. \quad (1.3.9)$$

The normalized eigenvector is $a = (1, 0)^T$. Similarly, taking $\lambda = -1$ gives $a = (0, 1)^T$.

2) Hermitian operators

If $\hat{A}^\dagger = \hat{A}$, or $A^\dagger = A$, the operator $\hat{A}$ is called Hermitian and A is called a Hermitian matrix.

We discuss the following properties of Hermitian operators.

a) Conjugate symmetry.

$$\langle a | \hat{A} | b \rangle = (\langle b | \hat{A}^\dagger | a \rangle)^* = (\langle b | \hat{A} | a \rangle)^*. \quad (1.3.10)$$

In particular,

$$\langle a | \hat{A} | a \rangle = (\langle a | \hat{A} | a \rangle)^*, \quad (1.3.11)$$

which shows that the expectation value $\langle a | \hat{A} | a \rangle$ of a Hermitian operator is real.

For an anti-Hermitian operator, $\hat{A}^\dagger = -\hat{A}$ and

$$\langle a | \hat{A} | a \rangle = -(\langle a | \hat{A} | a \rangle)^*, \quad (1.3.12)$$

so its expectation value is purely imaginary.

b) Reality of eigenvalues and orthogonality. Suppose

$$\hat{A} | i \rangle = \lambda_i | i \rangle, \quad \hat{A} | j \rangle = \lambda_j | j \rangle. \quad (1.3.13)$$

From

$$\langle j | \hat{A} | i \rangle = \lambda_i \langle j | i \rangle \quad (1.3.14)$$

and

$$\langle j | \hat{A} = \langle j | \lambda_j^*, \quad \langle j | \hat{A} | i \rangle = \lambda_j^* \langle j | i \rangle, \quad (1.3.15)$$

we obtain

$$(\lambda_i - \lambda_j^*) \langle j | i \rangle = 0. \quad (1.3.16)$$

For $j = i$ and $\langle i | i \rangle \neq 0$, $\lambda_i = \lambda_i^*$: the eigenvalues of a Hermitian operator are real.

For $j \neq i$, if $\lambda_i - \lambda_j^* \neq 0$, then $\langle j | i \rangle = 0$. Thus eigenstates corresponding to different eigenvalues of a Hermitian operator are orthogonal. Since they can always be normalized, the orthonormality condition is

$$\langle i | j \rangle = \delta_{ij}. \quad (1.3.17)$$

Question: Are eigenstates corresponding to the same eigenvalue orthogonal?

c) Orthogonalization by linear superposition (Schmidt orthogonalization). Suppose several eigenstates correspond to the same eigenvalue, i.e., there is degeneracy. For example, for a g -fold degenerate eigenvalue ($g \geq 2$),

$$\hat{A}|i, j\rangle = \lambda_i |i, j\rangle, \quad j = 1, \dots, g. \quad (1.3.18)$$

Are these g eigenvectors orthogonal?

Redefine g new states:

$$|i, n\rangle = \sum_{j=1}^g C_{nj} |i, j\rangle, \quad n = 1, 2, \dots, g. \quad (1.3.19)$$

Since

$$\hat{A}|i, n\rangle = \sum_{j=1}^g C_{nj} \hat{A}|i, j\rangle = \lambda_i \sum_{j=1}^g C_{nj} |i, j\rangle = \lambda_i |i, n\rangle, \quad (1.3.20)$$

each $|i, n\rangle$ is still an eigenstate of $\hat{A}$ with eigenvalue λ_i . Hence it suffices to show that the $|i, n\rangle$ can be made orthonormal.

Here is an explicit construction of the required vectors. Start with linearly independent eigenvectors $|v_1\rangle, \dots, |v_g\rangle$ in the degenerate eigenspace, and normalize the first:

$$|e_1\rangle = \frac{|v_1\rangle}{\sqrt{\langle v_1|v_1\rangle}}. \quad (1.3.21)$$

Remove the component of the second vector parallel to $|e_1\rangle$:

$$|w_2\rangle = |v_2\rangle - |e_1\rangle \langle e_1|v_2\rangle, \quad \langle e_1|w_2\rangle = \langle e_1|v_2\rangle - \langle e_1|e_1\rangle \langle e_1|v_2\rangle = 0. \quad (1.3.22)$$

The original pair is independent, so $|w_2\rangle \neq 0$. Set $|e_2\rangle = |w_2\rangle / \sqrt{\langle w_2|w_2\rangle}$. For each later vector, repeat:

$$|w_n\rangle = |v_n\rangle - \sum_{k=1}^{n-1} |e_k\rangle \langle e_k|v_n\rangle, \quad |e_n\rangle = \frac{|w_n\rangle}{\sqrt{\langle w_n|w_n\rangle}}. \quad (1.3.23)$$

For $j < n$,

$$\langle e_j|w_n\rangle = \langle e_j|v_n\rangle - \sum_{k < n} \delta_{jk} \langle e_k|v_n\rangle = 0. \quad (1.3.24)$$

Since all $|v_n\rangle$ have the same eigenvalue λ_i , every linear combination $|w_n\rangle$ also satisfies $\hat{A}|w_n\rangle = \lambda_i |w_n\rangle$. This is the Gram–Schmidt construction *inside the eigenspace*. It yields orthonormal eigenvectors, including in the degenerate case:

$$\langle i, m|j, n\rangle = \delta_{ij} \delta_{mn}. \quad (1.3.25)$$

Here distinct eigenvalues supply the factor δ_{ij} , and Gram–Schmidt supplies δ_{mn} .

Conclusion: Eigenstates of a Hermitian operator can be chosen orthonormal, whether the eigenvalues are degenerate or nondegenerate.

d) Completeness. In a finite-dimensional space we can establish completeness by induction. A Hermitian $N \times N$ matrix has an eigenvalue by the fundamental theorem of algebra applied to its characteristic polynomial, and the reality argument above makes that eigenvalue real. Choose a normalized eigenvector $|e_1\rangle$. Its orthogonal complement is invariant: whenever $\langle e_1|v\rangle = 0$,

$$\langle e_1|\hat{A}v\rangle = \langle e_1|\hat{A}|v\rangle = \lambda_1 \langle e_1|v\rangle = 0. \quad (1.3.26)$$

Restrict $\hat{A}$ to this $(N - 1)$ -dimensional invariant subspace. The restriction remains Hermitian. Iterating produces N mutually orthonormal eigenvectors, including any necessary orthogonalization within degenerate subspaces. Every ket can therefore be expanded as

$$|\psi\rangle = \sum_{n=1}^N |e_n\rangle \langle e_n|\psi\rangle = \left(\sum_{n=1}^N |e_n\rangle \langle e_n| \right) |\psi\rangle. \quad (1.3.27)$$

Because $|\psi\rangle$ is arbitrary,

$$\sum_{n=1}^N |e_n\rangle \langle e_n| = \mathbf{1}. \quad (1.3.28)$$

This is the finite-dimensional spectral theorem. Additional domain and continuous-spectrum considerations arise in infinite dimension.

Interactive investigation IC-01-02. Open A Hermitian 2×2 eigenproblem. Now vary the diagonal and off-diagonal entries of a Hermitian matrix. Calculate both eigenvalues and normalized eigenvectors by hand; compare with the display and examine the approach to degeneracy.

1.4 Beyond Hermiticity: Loss, Complex Spectra, and Exceptional Points

A Hermitian Hamiltonian has real eigenvalues and preserves the norm of a state under Schrödinger evolution. A subsystem that exchanges excitations with its environment admits another useful description: an *effective non-Hermitian Hamiltonian*. We now construct one directly from a physical model, write down every matrix element, and follow its eigenvalues and dynamics without assuming diagonalizability. For broader physical context, see Ashida, Gong, and Ueda [1] and the photonics review by El-Ganainy *et al.* [2].

1) What changes when a Hamiltonian is non-Hermitian?

Let the possibly non-Hermitian generator be $\hat{H}_{\text{eff}}$, with

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}_{\text{eff}} |\psi(t)\rangle. \quad (1.4.1)$$

Take the Hermitian adjoint of both sides. The imaginary unit changes sign under complex conjugation, so

$$-i\hbar \frac{d}{dt} \langle\psi(t)| = \langle\psi(t)| \hat{H}_{\text{eff}}^\dagger. \quad (1.4.2)$$

Rearranging these two equations gives

$$\frac{d}{dt} |\psi\rangle = -\frac{i}{\hbar} \hat{H}_{\text{eff}} |\psi\rangle, \quad \frac{d}{dt} \langle\psi| = \frac{i}{\hbar} \langle\psi| \hat{H}_{\text{eff}}^\dagger. \quad (1.4.3)$$

Use the product rule, keeping both contributions:

$$\begin{aligned} \frac{d}{dt} \langle\psi|\psi\rangle &= \left(\frac{d}{dt} \langle\psi| \right) |\psi\rangle + \langle\psi| \left(\frac{d}{dt} |\psi\rangle \right) \\ &= \frac{i}{\hbar} \langle\psi| \hat{H}_{\text{eff}}^\dagger |\psi\rangle - \frac{i}{\hbar} \langle\psi| \hat{H}_{\text{eff}} |\psi\rangle \\ &= \boxed{\frac{i}{\hbar} \langle\psi| \hat{H}_{\text{eff}}^\dagger - \hat{H}_{\text{eff}} |\psi\rangle}. \end{aligned} \quad (1.4.4)$$

For $\hat{H}_{\text{eff}} = \hat{H}_{\text{eff}}^\dagger$, the last line vanishes, recovering unitary norm conservation. For a lossy effective model it will be nonpositive.

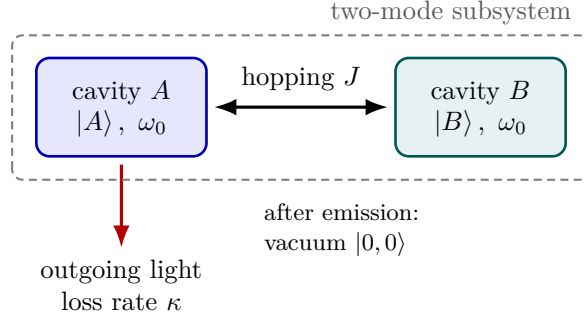


Figure 1.1: Original schematic of the single-photon coupled-cavity model. The external reservoir accepts an excitation from cavity A ; coherent hopping transfers it between the two cavities. The non-Hermitian matrix below describes the *conditional* no-emission sector, rather than unconditional probability loss in the full quantum system.

2) Concrete apparatus: two coupled optical cavities

Consider two optical cavities, each supporting one mode of common frequency ω_0 . A photon can hop coherently between them at a real coupling frequency $J > 0$. Mode A additionally emits into an external reservoir at an *intensity* decay rate $\kappa \geq 0$. We study the single-photon sector with basis $|A\rangle = |1, 0\rangle$ and $|B\rangle = |0, 1\rangle$. We temporarily use units $\hbar = 1$ so that the matrix entries are angular frequencies.

The coherent part of the Hamiltonian is written in ket form as

$$\hat{H}_0 = \omega_0(|A\rangle\langle A| + |B\rangle\langle B|) + J(|A\rangle\langle B| + |B\rangle\langle A|). \quad (1.4.5)$$

To obtain its matrix, insert $\langle A|, \langle B|$ on the left and $|A\rangle, |B\rangle$ on the right. Because $\langle A|A\rangle = \langle B|B\rangle = 1$ and $\langle A|B\rangle = 0$,

$$\begin{aligned} \langle A|\hat{H}_0|A\rangle &= \omega_0, & \langle A|\hat{H}_0|B\rangle &= J, \\ \langle B|\hat{H}_0|A\rangle &= J, & \langle B|\hat{H}_0|B\rangle &= \omega_0. \end{aligned} \quad (1.4.6)$$

Thus $H_0 = \begin{pmatrix} \omega_0 & J \\ J & \omega_0 \end{pmatrix}$ is Hermitian.

To represent the monitored emission channel, introduce $\hat{L} = \sqrt{\kappa}|\text{vac}\rangle\langle A|$ with $|\text{vac}\rangle = |0, 0\rangle$. The standard no-click (no-jump) effective Hamiltonian is

$$\hat{H}_{\text{eff}} = \hat{H}_0 - \frac{i}{2}\hat{L}^\dagger\hat{L}. \quad (1.4.7)$$

Let us perform the multiplication:

$$\begin{aligned} \hat{L}^\dagger\hat{L} &= \kappa|A\rangle\langle\text{vac}|\langle\text{vac}|A\rangle \\ &= \kappa|A\rangle\langle A|, \end{aligned} \quad (1.4.8)$$

since $\langle\text{vac}|\text{vac}\rangle = 1$. Therefore the explicit non-Hermitian matrix is

$$H_{\text{eff}} = \begin{pmatrix} \omega_0 - i\kappa/2 & J \\ J & \omega_0 \end{pmatrix}. \quad (1.4.9)$$

Its AA entry is complex, while the coherent coupling stays real and symmetric. Indeed, $H_{\text{eff}}^\dagger - H_{\text{eff}} = i\kappa|A\rangle\langle A|$.

Check the physical meaning of the imaginary entry. Write the unnormalized conditional state as $|\tilde{\psi}(t)\rangle = a(t)|A\rangle + b(t)|B\rangle$. The Schrödinger equation gives the two *component* equations, before and after division by i :

$$\begin{aligned} i\dot{a} &= (\omega_0 - i\kappa/2)a + Jb, & \dot{a} &= (-i\omega_0 - \kappa/2)a - iJb, \\ i\dot{b} &= Ja + \omega_0 b, & \dot{b} &= -iJa - i\omega_0 b. \end{aligned} \quad (1.4.10)$$

Using $\frac{d}{dt}|a|^2 = a^*\dot{a} + \dot{a}^*a$, we find

$$\frac{d}{dt}|a|^2 = -\kappa|a|^2 + iJ(ab^* - a^*b), \quad (1.4.11)$$

$$\frac{d}{dt}|b|^2 = -iJ(ab^* - a^*b). \quad (1.4.12)$$

The two hopping contributions cancel upon addition:

$$\boxed{\frac{d}{dt}(|a|^2 + |b|^2) = -\kappa|a|^2 \leq 0.} \quad (1.4.13)$$

This verifies Eq. (1.4.4) directly.

For a normalized initial single-photon state, $P_{\text{no jump}}(t) = |a(t)|^2 + |b(t)|^2$ is the probability of detecting *no photon emission* up to time t in this monitored-channel model. Upon conditioning on that event, the surviving state is normalized by dividing $|\tilde{\psi}\rangle$ by $\sqrt{P_{\text{no jump}}}$. An unconditional description includes the vacuum reached after emission and obeys trace-preserving evolution. Thus the decreasing no-jump norm represents a physically identified branch of a larger probability-conserving theory.

3) Eigenvalues: every algebraic step

Solve the right eigenproblem $H_{\text{eff}}\mathbf{v}_{\pm} = \lambda_{\pm}\mathbf{v}_{\pm}$. The characteristic equation reads

$$\begin{aligned} 0 &= \det(H_{\text{eff}} - \lambda\mathbf{1}) \\ &= \begin{vmatrix} \omega_0 - i\kappa/2 - \lambda & J \\ J & \omega_0 - \lambda \end{vmatrix} \\ &= (\omega_0 - i\kappa/2 - \lambda)(\omega_0 - \lambda) - J^2. \end{aligned} \quad (1.4.14)$$

Set $z = \lambda - \omega_0$ and expand:

$$(-z - i\kappa/2)(-z) - J^2 = z^2 + \frac{i\kappa}{2}z - J^2 = 0. \quad (1.4.15)$$

Complete the square:

$$\left(z + \frac{i\kappa}{4}\right)^2 = J^2 - \frac{\kappa^2}{16}. \quad (1.4.16)$$

We obtain the exact eigenvalues

$$\boxed{\lambda_{\pm} = \omega_0 - \frac{i\kappa}{4} \pm \sqrt{J^2 - \left(\frac{\kappa}{4}\right)^2}.} \quad (1.4.17)$$

An eigenamplitude with eigenvalue $\lambda = E - i\Gamma/2$ evolves as $\exp(-iEt - \Gamma t/2)$, so its *intensity* decays as $\exp(-\Gamma t)$. The factor of two between amplitude and intensity decay is essential.

Construct the right eigenvectors explicitly. The top component of $(H_{\text{eff}} - \lambda \mathbf{1})\mathbf{v} = 0$ is

$$(\omega_0 - i\kappa/2 - \lambda)v_A + Jv_B = 0. \quad (1.4.18)$$

For $J > 0$ choose $v_A = 1$ and solve for v_B :

$$\mathbf{v}_{\pm} = \begin{pmatrix} 1 \\ (\lambda_{\pm} - \omega_0 + i\kappa/2)/J \end{pmatrix}. \quad (1.4.19)$$

The second component then vanishes precisely because the secular equation holds. These vectors may be normalized by their Euclidean norms, but eigenvectors of distinct complex eigenvalues need not be orthogonal. For a general non-Hermitian matrix, left eigenvectors form a different set, and biorthogonal expansions replace the Hermitian orthonormal eigenbasis.

4) Exceptional points are more than repeated eigenvalues

Define

$$\Omega = \sqrt{J^2 - \left(\frac{\kappa}{4}\right)^2}. \quad (1.4.20)$$

For $0 \leq \kappa < 4J$, the square root is real: the frequencies $\text{Re } \lambda_{\pm}$ split, while $\text{Im } \lambda_{\pm} = -\kappa/4$ are identical. For $\kappa > 4J$, the square root is imaginary: the two frequencies coincide at ω_0 , while the decay rates separate. At the threshold,

$$\kappa_{\text{EP}} = 4J, \quad \lambda_{\text{EP}} = \omega_0 - iJ. \quad (1.4.21)$$

Insert these values back into the original matrix rather than inferring the eigenspace from the roots alone:

$$H_{\text{eff}} - \lambda_{\text{EP}} \mathbf{1} = J \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix}. \quad (1.4.22)$$

An eigenvector satisfies $-iv_A + v_B = 0$, hence $v_B = iv_A$. With $v_A = 1$,

$$\mathbf{v}_{\text{EP}} = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} -i + i \\ 1 + i^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.4.23)$$

The second row equals i times the first, and the matrix has rank one. Consequently its nullspace has dimension one: a double eigenvalue possesses only one independent right eigenvector. This is an *exceptional point* (EP). For a Hermitian matrix, a twofold degeneracy instead has two independent eigenvectors; eigenvalue coincidence by itself is insufficient to identify an EP.

Connection to PT symmetry, with an important distinction. Subtract a multiple of the identity:

$$H_{\text{eff}} = \left(\omega_0 - \frac{i\kappa}{4}\right) \mathbf{1} + K, \quad K = \begin{pmatrix} -i\kappa/4 & J \\ J & i\kappa/4 \end{pmatrix}. \quad (1.4.24)$$

For $\mathcal{P} = \sigma_x$ (cavity exchange) and $\mathcal{T}$ (complex conjugation), direct multiplication gives $\sigma_x K^* \sigma_x = K$. Thus the shifted matrix K has PT symmetry, and its eigenvalues are $\pm\Omega$. A physically *pumped* balanced gain–loss dimer would instead be

$$H_{\mathcal{PT}} = \begin{pmatrix} \omega_0 + i\gamma & J \\ J & \omega_0 - i\gamma \end{pmatrix}, \quad E_{\pm} = \omega_0 \pm \sqrt{J^2 - \gamma^2}, \quad (1.4.25)$$

with an EP at $|\gamma| = J$. The apparatus of Fig. 1.1 has *only a lossy cavity*; its apparent gain in K is relative to the factored-out common decay. This is an algebraic correspondence, not a claim that cavity B contains an active gain medium.

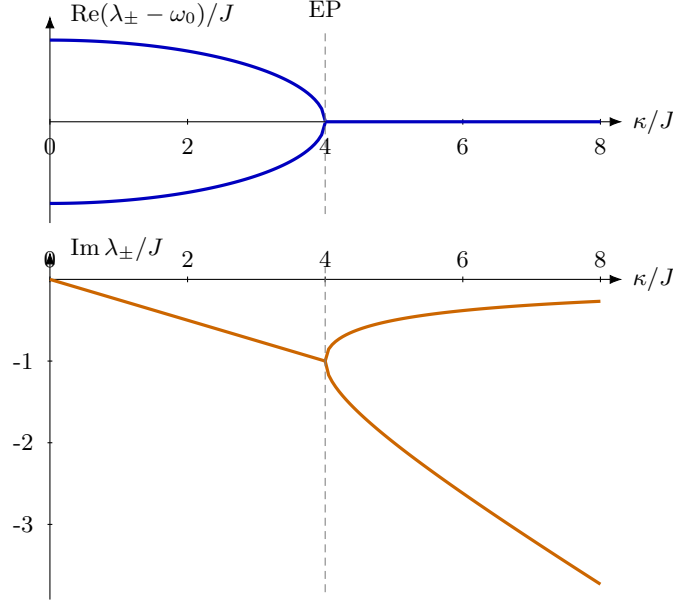


Figure 1.2: Analytic real (upper) and imaginary (lower) parts of the passive coupled-cavity spectrum in Eq. (1.4.17), with $J > 0$. The frequency branches merge at $\kappa/J = 4$, while the two attenuation-related imaginary parts separate beyond that value. Both right eigenvectors also coalesce at the marked EP.

5) Solve the dynamics even at the exceptional point

The eigenvector expansion breaks down at the EP, so we use the matrix exponential directly. From Eq. (1.4.24),

$$\begin{aligned} K^2 &= \begin{pmatrix} -i\kappa/4 & J \\ J & i\kappa/4 \end{pmatrix} \begin{pmatrix} -i\kappa/4 & J \\ J & i\kappa/4 \end{pmatrix} \\ &= \begin{pmatrix} J^2 - \kappa^2/16 & 0 \\ 0 & J^2 - \kappa^2/16 \end{pmatrix} = \Omega^2 \mathbf{1}. \end{aligned} \quad (1.4.26)$$

Hence $K^{2m} = \Omega^{2m} \mathbf{1}$ and $K^{2m+1} = \Omega^{2m} K$. Splitting the exponential series,

$$\begin{aligned} e^{-iKt} &= \sum_{m=0}^{\infty} \frac{(-1)^m (\Omega t)^{2m}}{(2m)!} \mathbf{1} - i \sum_{m=0}^{\infty} \frac{(-1)^m (\Omega t)^{2m+1}}{(2m+1)!} \frac{K}{\Omega} \\ &= \boxed{\mathbf{1} \cos(\Omega t) - i \frac{K}{\Omega} \sin(\Omega t)}. \end{aligned} \quad (1.4.27)$$

Since $\mathbf{1}$ commutes with K , the common diagonal shift in Eq. (1.4.24) factors out:

$$e^{-iH_{\text{eff}}t} = e^{-i\omega_0 t - \kappa t/4} e^{-iKt}. \quad (1.4.28)$$

Prepare a single photon in cavity A , $(a(0), b(0))^T = (1, 0)^T$. Multiplication by the first column of Eq. (1.4.27) gives

$$a(t) = e^{-i\omega_0 t - \kappa t/4} \left[\cos(\Omega t) - \frac{\kappa}{4\Omega} \sin(\Omega t) \right], \quad (1.4.29)$$

$$b(t) = -ie^{-i\omega_0 t - \kappa t/4} \frac{J}{\Omega} \sin(\Omega t). \quad (1.4.30)$$

At $\kappa > 4J$, Ω is imaginary. Use $\sin(ix) = i \sinh x$ and $\cos(ix) = \cosh x$ to obtain real-valued decay envelopes; the formulas themselves remain valid by analytic continuation.

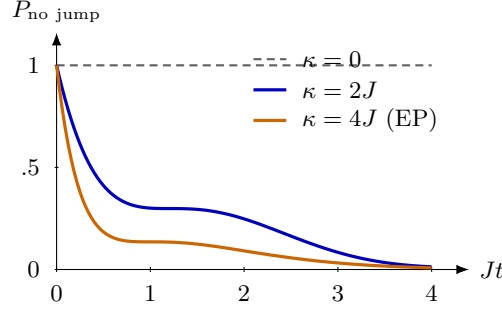


Figure 1.3: Exact no-emission survival probability $|a(t)|^2 + |b(t)|^2$ for a photon initially in cavity A , computed from Eqs. (1.4.29) and (1.4.30). The Hermitian limit conserves norm. For $\kappa = 2J$ the survival decays smoothly; at the EP $\kappa = 4J$, $P_{\text{no jump}} = e^{-2Jt}[(1 - Jt)^2 + (Jt)^2]$. The curves use dimensionless time Jt .

Exactly at the EP, $\Omega = 0$. The limit $\lim_{\Omega \rightarrow 0} \sin(\Omega t)/\Omega = t$ and $\cos 0 = 1$ yield

$$a_{\text{EP}}(t) = e^{-i\omega_0 t - Jt}(1 - Jt), \quad b_{\text{EP}}(t) = -ie^{-i\omega_0 t - Jt}Jt. \quad (1.4.31)$$

The polynomial factor multiplying the exponential is a signature of the non-diagonalizable Jordan structure.

6) Essential consistency checks

1. **Hermitian limit.** For $\kappa = 0$ we have $\Omega = J$, $a(t) = e^{-i\omega_0 t} \cos Jt$, $b(t) = -ie^{-i\omega_0 t} \sin Jt$. The norm is exactly $\cos^2 Jt + \sin^2 Jt = 1$.
2. **Uncoupled lossy cavity.** Set $J = 0$ in the *original* component equations. Then $a(t) = a(0)e^{-i\omega_0 t - \kappa t/2}$ and $b(t) = b(0)e^{-i\omega_0 t}$. An initial photon in A has survival probability $e^{-\kappa t}$: this checks that κ means *intensity* decay rate.
3. **Probability in the larger system.** The norm lost by the no-jump branch is the accumulated probability of emission, so $P_{\text{no jump}}(t) + P_{\text{emission}}(t) = 1$ for the initial one-photon state. There is no loss of probability in the full system plus environment.

Checkpoint. At $\kappa = 4J$, show directly that the characteristic polynomial equals $(\lambda - \omega_0 + iJ)^2$. Compute the nullspace of $H_{\text{eff}} - (\omega_0 - iJ)\mathbf{1}$. Why does the EP provide only one eigenvector? Finally take the $\Omega \rightarrow 0$ limit in Eq. (1.4.27) and verify $e^{-iKt} = \mathbf{1} - iKt$ using $K^2 = 0$.

Independent extension. In the online companion, the [guided non-Hermitian reading question](#) offers an additional matrix example and a chance to check the complex spectrum with symbolic calculation. The hand derivation and norm balance provide the primary physical verification.

1.5 Measurement

We discuss the connection between quantum mechanics and Hilbert space.

1) Measurement of a single physical quantity

In the Stern–Gerlach experiment, s_z has no definite value before the magnetic field (if it did, only one beam of silver atoms would emerge). It has a definite value only after the field: one beam has s_z^+ , and the other has s_z^- . Therefore the values of quantum-mechanical quantities

are closely related to the state of the system. A physical quantity has no definite value in a general state; it has a definite value only in certain particular states.

Postulate of quantum mechanics: Physical quantities of a quantum-mechanical system are represented by Hermitian operators in a linear vector space, and states are represented by vectors. Because the eigenstates of a Hermitian operator are complete, the eigenstates of any physical quantity can form a linear vector space, or a representation.

The eigenvalue equation of $\hat{F}$ is

$$\hat{F}|n\rangle = f_n|n\rangle. \quad (1.5.1)$$

In the F representation, the basis vectors are $|n\rangle$. An arbitrary state is

$$|a\rangle = \sum_n |n\rangle \langle n|a\rangle. \quad (1.5.2)$$

The coefficient $\langle n|a\rangle$ is the concrete representation of $|a\rangle$ in the F representation.

Postulate of quantum mechanics: The physical quantity $\hat{F}$ has a definite value, namely the eigenvalue f_n , only in its eigenstate $|n\rangle$. In an arbitrary state $|a\rangle$, $\hat{F}$ has no definite value, but has a definite average value

$$\langle F \rangle \equiv \langle a|\hat{F}|a\rangle. \quad (1.5.3)$$

(The eigenvalues and expectation values of Hermitian operators are real. This is why physical quantities are represented by Hermitian operators.)

$$\begin{aligned} \langle F \rangle &\equiv \langle a|\hat{F}|a\rangle = \sum_{n,m} \langle a|n\rangle \langle n|\hat{F}|m\rangle \langle m|a\rangle \\ &= \sum_{n,m} \langle a|n\rangle f_m \delta_{nm} \langle m|a\rangle = \sum_n \langle a|n\rangle \langle n|a\rangle f_n = \sum_n |\langle n|a\rangle|^2 f_n. \end{aligned} \quad (1.5.4)$$

Thus the probability that $\hat{F}$ takes the value f_n is $|\langle n|a\rangle|^2$. The quantity $\langle n|a\rangle$ is a probability amplitude, and $|\langle n|a\rangle|^2$ is the probability for state $|a\rangle$ to contain state $|n\rangle$.

In a general state, the value of a physical quantity is uncertain, but its probabilities and expectation value are definite.

In the eigenbasis of $\hat{F}$, the matrix elements are

$$F_{mn} = \langle m|\hat{F}|n\rangle = f_n \langle m|n\rangle = f_n \delta_{mn}. \quad (1.5.5)$$

Thus a physical quantity is represented by a diagonal matrix in its own representation, and the diagonal elements are its eigenvalues.

We now discuss measurement and collapse of the state.

Before the magnetic field in the Stern–Gerlach experiment, s_z of an Ag atom has no definite value; after passing through the field it has a definite value. Applying the magnetic field can be regarded as measuring the spin. Once measured, it has a definite value.

When $\hat{F}$ is measured in an arbitrary state $|a\rangle$, the system collapses into an eigenstate of $\hat{F}$, so $\hat{F}$ has a definite value. Which eigenstate does it collapse into? The probability for state $|a\rangle$ to contain state $|n\rangle$ is $|\langle n|a\rangle|^2$. Therefore the probability of collapse into $|n\rangle$, i.e., of obtaining the measurement value f_n , is $|\langle n|a\rangle|^2$.

Measurement allows the particle nature of the system to become manifest; one may also say that measurement produces a particle.

Question: Measure $\hat{F}$ in state $|a\rangle$ and obtain f_n . If $\hat{F}$ is measured again immediately, what value will be obtained?

First measurement: $|a\rangle \longrightarrow |n\rangle$.

At the second measurement, the system is already in the eigenstate $|n\rangle$ of $\hat{F}$; the result is still f_n . This explains why an Ag atom in the Stern–Gerlach experiment still has spin s_z^+ after passing through two magnetic fields in the z direction.

Interactive investigation IC-01-04. Open Born rule and projectors. After the outcome-probability and repeatability postulates, select different projectors. Check that the probabilities sum to one and that the normalized postmeasurement state predicts the result of an immediate repetition.

2) Spin matrices

From the Stern–Gerlach experiment, the electron has spin $1/2$, and the spin operator $\hat{\mathbf{s}} \cdot \mathbf{e}_n = \hat{s}_n$ along any direction $\mathbf{e}_n$ takes the values $\pm\hbar/2$. Its eigenvalue equation is

$$\hat{s}_n |s_n^\pm\rangle = \pm \frac{\hbar}{2} |s_n^\pm\rangle. \quad (1.5.6)$$

In the representation formed by the eigenstates of $\hat{s}_z$ (the s_z representation), s_z is a diagonal matrix whose diagonal entries are $\pm\hbar/2$:

$$s_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.5.7)$$

Solving

$$s_z \begin{pmatrix} a \\ b \end{pmatrix} = \pm \frac{\hbar}{2} \begin{pmatrix} a \\ b \end{pmatrix} \quad (1.5.8)$$

gives the normalized eigenstates

$$|s_z^+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |s_z^-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.5.9)$$

We now find the matrices for $\hat{s}_x$ and $\hat{s}_y$ in the s_z representation. From the completeness of the eigenstates of $\hat{s}_z$,

$$|s_x^+\rangle = \sum_{i=\pm} |s_z^i\rangle \langle s_z^i | s_x^+\rangle = \langle s_z^+ | s_x^+\rangle |s_z^+\rangle + \langle s_z^- | s_x^+\rangle |s_z^-\rangle. \quad (1.5.10)$$

The cascaded Stern–Gerlach experiment shows that a state with a definite s_x value, after passing through a z -directed magnetic field (a measurement device for s_z), separates into two beams of equal intensity, with equal probabilities for s_z^+ and s_z^- :

$$|\langle s_z^+ | s_x^+\rangle| = |\langle s_z^- | s_x^+\rangle| = \frac{1}{\sqrt{2}}. \quad (1.5.11)$$

Consequently,

$$|s_x^+\rangle = \frac{1}{\sqrt{2}} |s_z^+\rangle + \frac{1}{\sqrt{2}} e^{i\delta_1} |s_z^-\rangle. \quad (1.5.12)$$

Through normalization, the phase factors have been collected into δ_1 .

Using the orthonormality of $|s_x^\pm\rangle$, $\langle s_x^i | s_x^j\rangle = \delta_{ij}$, we obtain

$$|s_x^-\rangle = \frac{1}{\sqrt{2}} |s_z^+\rangle - \frac{1}{\sqrt{2}} e^{i\delta_1} |s_z^-\rangle. \quad (1.5.13)$$

Equivalently,

$$\begin{cases} |s_z^+\rangle = \frac{1}{\sqrt{2}} (|s_x^+\rangle + |s_x^-\rangle), \\ |s_z^-\rangle = \frac{1}{\sqrt{2}} e^{-i\delta_1} (|s_x^+\rangle - |s_x^-\rangle). \end{cases} \quad (1.5.14)$$

Substituting into the matrix elements of $\hat{s}_x$ gives

$$\langle s_z^+ | \hat{s}_x | s_z^+\rangle = 0, \quad \langle s_z^+ | \hat{s}_x | s_z^-\rangle = \frac{\hbar}{2} e^{-i\delta_1}, \quad \langle s_z^- | \hat{s}_x | s_z^+\rangle = \frac{\hbar}{2} e^{i\delta_1}, \quad \langle s_z^- | \hat{s}_x | s_z^-\rangle = 0. \quad (1.5.15)$$

Thus

$$s_x = \frac{\hbar}{2} \begin{pmatrix} 0 & e^{-i\delta_1} \\ e^{i\delta_1} & 0 \end{pmatrix}. \quad (1.5.16)$$

Similarly,

$$|s_y^\pm\rangle = \frac{1}{\sqrt{2}} |s_z^+\rangle \pm \frac{1}{\sqrt{2}} e^{i\delta_2} |s_z^-\rangle, \quad s_y = \frac{\hbar}{2} \begin{pmatrix} 0 & e^{-i\delta_2} \\ e^{i\delta_2} & 0 \end{pmatrix}. \quad (1.5.17)$$

Now consider the values of δ_1, δ_2 . A state with a definite s_x value, after passing through a magnetic field in the y direction (a measurement device for s_y), splits into two equally intense beams corresponding to s_y^+ and s_y^- :

$$|\langle s_y^\pm | s_x^+ \rangle| = \frac{1}{\sqrt{2}}. \quad (1.5.18)$$

That is,

$$\frac{1}{2} |1 \pm e^{i(\delta_1 - \delta_2)}| = \frac{1}{\sqrt{2}}. \quad (1.5.19)$$

Hence

$$\delta_1 - \delta_2 = \pm \frac{\pi}{2}. \quad (1.5.20)$$

Choosing $\delta_1 = 0$ and $\delta_2 = \pi/2$ yields

$$s_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (1.5.21)$$

For s_x, s_y, s_z , one can show that

$$[s_i, s_j] = i\epsilon_{ijk} \hbar s_k, \quad \{s_i, s_j\} = \frac{\hbar^2}{2} \delta_{ij}, \quad (1.5.22)$$

where

$$[A, B] = AB - BA, \quad \{A, B\} = AB + BA. \quad (1.5.23)$$

Define

$$s^2 = \sum_i s_i^2 = \frac{3}{4} \hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (1.5.24)$$

Because s^2 is proportional to the identity matrix, every state is its eigenstate, and it has the definite value $3\hbar^2/4$ in every state. Clearly, s^2 commutes with every physical quantity $\hat{F}$ and may be regarded as a classical quantity.

3) Conditions for simultaneous measurement of several physical quantities

Can two physical quantities $\hat{A}$ and $\hat{B}$ be measured simultaneously?

Theorem. If $\hat{A}$ and $\hat{B}$ have a complete set of common eigenstates, then $\hat{A}$ and $\hat{B}$ commute.

Proof. Let the complete common eigenstates be $|n\rangle$, with

$$\hat{A} |n\rangle = a_n |n\rangle, \quad \hat{B} |n\rangle = b_n |n\rangle. \quad (1.5.25)$$

In the representation formed by $|n\rangle$, an arbitrary state is

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle. \quad (1.5.26)$$

Then

$$\begin{aligned}
 (\hat{A}\hat{B} - \hat{B}\hat{A})|\psi\rangle &= \sum_n (\hat{A}\hat{B} - \hat{B}\hat{A})|n\rangle \langle n|\psi\rangle \\
 &= \sum_n (\hat{A}b_n - \hat{B}a_n)|n\rangle \langle n|\psi\rangle \\
 &= \sum_n (a_nb_n - b_na_n)|n\rangle \langle n|\psi\rangle = 0.
 \end{aligned} \tag{1.5.27}$$

Since $|\psi\rangle$ is arbitrary, $[\hat{A}, \hat{B}] = 0$.

Converse theorem. If $\hat{A}$ and $\hat{B}$ commute, then they have a complete set of common eigenstates.

Proof. Let $|n\rangle$ be an eigenstate of $\hat{A}$:

$$\hat{A}|n\rangle = a_n|n\rangle. \tag{1.5.28}$$

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

$$\hat{A}\hat{B}|n\rangle = \hat{B}\hat{A}|n\rangle = a_n\hat{B}|n\rangle. \tag{1.5.29}$$

Thus $\hat{B}|n\rangle$ is also an eigenstate of $\hat{A}$ with eigenvalue a_n . If $\hat{A}$ is nondegenerate, $\hat{B}|n\rangle$ and $|n\rangle$ are the same state up to a constant:

$$\hat{B}|n\rangle = b_n|n\rangle. \tag{1.5.30}$$

Therefore $|n\rangle$ is also an eigenstate of $\hat{B}$, and $\hat{A}$ and $\hat{B}$ have a complete common eigenbasis.

If $\hat{A}$ is degenerate, one can use the Schmidt method to prove the same result (exercise).

Conclusion. When several physical quantities commute, they can have common eigenstates. In those common eigenstates, all these quantities have definite values, and they can be measured simultaneously.

Note that $[\hat{A}, \hat{B}] = 0$ and $[\hat{B}, \hat{C}] = 0$ mean that A, B can have definite values simultaneously and that B, C can have definite values simultaneously. However, A, C need not have definite values simultaneously.

When $\hat{A}$ and $\hat{B}$ do not commute, they cannot have a complete set of common eigenstates and cannot always be measured simultaneously. They may, however, be simultaneously measurable in a subspace. For example, for orbital angular momentum,

$$[\hat{L}^2, \hat{L}_x] = 0, \quad [\hat{L}^2, \hat{L}_y] = 0, \quad [\hat{L}_x, \hat{L}_y] \neq 0. \tag{1.5.31}$$

Thus L^2, L_x can have definite values simultaneously, and L^2, L_y can have definite values simultaneously, while L_x, L_y cannot generally have definite values simultaneously. In the state with $L = 0$, however, L_x, L_y are simultaneously definite, with $L_x = L_y = 0$.

4) Complete set of physical quantities

When eigenvalues and eigenstates of a physical quantity are degenerate,

$$\hat{F}|i, j\rangle = f_i|i, j\rangle, \quad j = 1, \dots, g, \tag{1.5.32}$$

Schmidt orthogonalization can be used mathematically to obtain orthonormal eigenstates. In physics, however, one generally introduces a complete set of physical quantities to achieve orthonormality.

That several eigenstates share one eigenvalue indicates that this physical quantity is insufficient to characterize the states completely; another physical quantity is needed to distinguish them. Introduce a second commuting observable $\hat{G}$ and require that the eigenstates of $\hat{F}$ also be eigenstates of $\hat{G}$:

$$\hat{F}|i, j\rangle = f_i|i, j\rangle, \quad \hat{G}|i, j\rangle = \gamma_j|i, j\rangle. \tag{1.5.33}$$

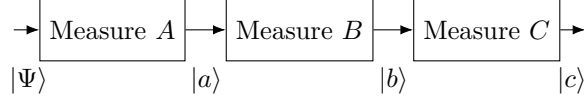
For the set $\{\hat{F}, \hat{G}\}$, if the pair of eigenvalues $\{f_i, \gamma_j\}$ corresponds one-to-one to the eigenstate $|i, j\rangle$, there is no degeneracy, and the orthonormality condition is

$$\langle i, j|i', j'\rangle = \delta_{ii'}\delta_{jj'}. \tag{1.5.34}$$

5) Cascaded measurements

For three physical quantities $\hat{A}, \hat{B}, \hat{C}$, consider the following two measurement procedures.

First procedure:



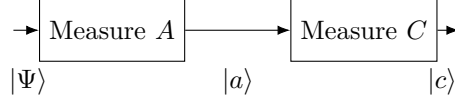
The state sequence is $|\Psi\rangle \rightarrow |a\rangle \rightarrow |b\rangle \rightarrow |c\rangle$; the values of the physical quantities are initially indefinite and then are a, b, c after their respective measurements.

After three measurements, the probability for $\hat{C}$ to take the value c is

$$P(c) = \sum_{a,b} |\langle a|\Psi\rangle|^2 |\langle b|a\rangle|^2 |\langle c|b\rangle|^2. \quad (1.5.35)$$

Here $|\langle a|\Psi\rangle|^2$ is the probability of collapse from $|\Psi\rangle$ to $|a\rangle$; $|\langle b|a\rangle|^2$ is the probability from $|a\rangle$ to $|b\rangle$; and $|\langle c|b\rangle|^2$ is the probability from $|b\rangle$ to $|c\rangle$.

Second procedure:



After two measurements, the probability for $\hat{C}$ to take the value c is

$$P'(c) = \sum_a |\langle a|\Psi\rangle|^2 |\langle c|a\rangle|^2. \quad (1.5.36)$$

Clearly, in general,

$$P(c) \neq P'(c). \quad (1.5.37)$$

However, if $[\hat{A}, \hat{B}] = 0$, $\hat{A}$ and $\hat{B}$ have common eigenstates $|n\rangle$, which replace $|a\rangle$ and $|b\rangle$:

$$P(c) = \sum_n |\langle n|\Psi\rangle|^2 |\langle c|n\rangle|^2 = P'(c). \quad (1.5.38)$$

Or if $[\hat{B}, \hat{C}] = 0$, $\hat{B}$ and $\hat{C}$ have common eigenstates $|m\rangle$, which replace $|b\rangle$ and $|c\rangle$:

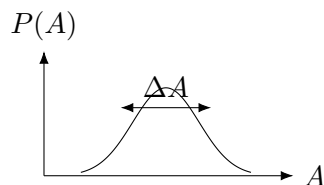
$$P(c) = \sum_a |\langle a|\Psi\rangle|^2 |\langle m|a\rangle|^2 = P'(c). \quad (1.5.39)$$

Conclusion. Since physical quantities have definite values only in their eigenstates, the order of measurements changes the final measurement results.

6) Uncertainty relation

If $[\hat{A}, \hat{B}] \neq 0$, the physical quantities $\hat{A}, \hat{B}$ cannot have definite values simultaneously. How large are their uncertainties?

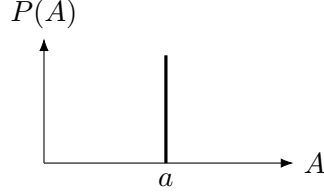
In an arbitrary state $|\psi\rangle$, $\hat{A}$ has only a definite expectation value $\langle \hat{A} \rangle \equiv \langle \psi | \hat{A} | \psi \rangle$ and a distribution $P(A)$.



The uncertainty is represented by the variance:

$$\langle (\Delta \hat{A})^2 \rangle \equiv \langle \psi | (\hat{A} - \langle \hat{A} \rangle)^2 | \psi \rangle \geq 0. \quad (1.5.40)$$

If $|\psi\rangle$ is an eigenstate of $\hat{A}$, the expectation value equals its eigenvalue, the variance is zero, and the distribution is a delta function.



Define the uncertainty between $\hat{A}$ and $\hat{B}$ using the product of variances $\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle$. How can we estimate it?

For arbitrary states $|\alpha\rangle, |\beta\rangle$ and arbitrary complex number λ , the norm of the linear-combination state is nonnegative:

$$\langle (\alpha| + \lambda^* \langle \beta|) (|\alpha\rangle + \lambda |\beta\rangle) \rangle \geq 0. \quad (1.5.41)$$

Choosing

$$\lambda = -\frac{\langle \beta | \alpha \rangle}{\langle \beta | \beta \rangle} \quad (1.5.42)$$

gives the Schwarz inequality,

$$\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq |\langle \alpha | \beta \rangle|^2. \quad (1.5.43)$$

Set

$$|\alpha\rangle = \Delta \hat{A} |\Psi\rangle, \quad |\beta\rangle = \Delta \hat{B} |\Psi\rangle, \quad \Delta \hat{A} = \hat{A} - \langle \hat{A} \rangle, \quad \Delta \hat{B} = \hat{B} - \langle \hat{B} \rangle. \quad (1.5.44)$$

Then

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq |\langle \Delta \hat{A} \Delta \hat{B} \rangle|^2. \quad (1.5.45)$$

Since

$$\begin{aligned} |\langle \Delta \hat{A} \Delta \hat{B} \rangle|^2 &= (\text{Re} \langle \Delta \hat{A} \Delta \hat{B} \rangle)^2 + (\text{Im} \langle \Delta \hat{A} \Delta \hat{B} \rangle)^2 \\ &\geq (\text{Im} \langle \Delta \hat{A} \Delta \hat{B} \rangle)^2 \\ &= \left[\frac{1}{2i} (\langle \Delta \hat{A} \Delta \hat{B} \rangle - \langle \Delta \hat{B} \Delta \hat{A} \rangle) \right]^2, \end{aligned} \quad (1.5.46)$$

we obtain

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \left[\frac{1}{2i} (\langle \Delta \hat{A} \Delta \hat{B} \rangle - \langle \Delta \hat{B} \Delta \hat{A} \rangle) \right]^2. \quad (1.5.47)$$

Equality requires equality in the Schwarz inequality and a vanishing real part of $\langle \Delta \hat{A} \Delta \hat{B} \rangle$.

From

$$\begin{aligned} \langle \Delta \hat{A} \Delta \hat{B} \rangle &= \langle \psi | (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) | \psi \rangle = \langle \hat{A} \hat{B} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle, \\ \langle \Delta \hat{B} \Delta \hat{A} \rangle &= \langle \psi | (\hat{B} - \langle \hat{B} \rangle) (\hat{A} - \langle \hat{A} \rangle) | \psi \rangle = \langle \hat{B} \hat{A} \rangle - \langle \hat{B} \rangle \langle \hat{A} \rangle, \end{aligned} \quad (1.5.48)$$

it follows that

$$\boxed{\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2.} \quad (1.5.49)$$

This shows that the uncertainties of two physical quantities are controlled by their commutation relation and the state $|\psi\rangle$.

1.6 Transformation of Representations

Selecting a suitable representation can simplify a calculation. The transformation between basis vectors of different representations is called a transformation of representation.

Consider two representations:

$$\begin{aligned} F \text{ representation: } & \text{basis } |i\rangle, \quad \hat{F}|i\rangle = f_i|i\rangle, \\ G \text{ representation: } & \text{basis } |m\rangle, \quad \hat{G}|m\rangle = g_m|m\rangle. \end{aligned} \quad (1.6.1)$$

We have

$$|m\rangle = \sum_i |i\rangle \langle i|m\rangle = \sum_i S_{im} |i\rangle. \quad (1.6.2)$$

The transformation matrix S is defined by

$$S_{im} = \langle i|m\rangle. \quad (1.6.3)$$

It represents the basis vectors of the G representation in the F representation; the identity is denoted by $\mathbf{1}$ throughout.

1) Transformation of an arbitrary vector

In the F representation,

$$|a\rangle = \sum_i |i\rangle \langle i|a\rangle. \quad (1.6.4)$$

The coefficient $\langle i|a\rangle$ is the component of $|a\rangle$ along $|i\rangle$, or its representation in the F basis.

In the G representation,

$$|a\rangle = \sum_m |m\rangle \langle m|a\rangle. \quad (1.6.5)$$

The coefficient $\langle m|a\rangle$ is the component along $|m\rangle$, or the representation of $|a\rangle$ in the G basis.

Using completeness,

$$\langle m|a\rangle = \sum_i \langle m|i\rangle \langle i|a\rangle = \sum_i S_{im}^* \langle i|a\rangle = \sum_i S_{mi}^\dagger \langle i|a\rangle. \quad (1.6.6)$$

In matrix form,

$$\boxed{a_G = S^\dagger a_F.} \quad (1.6.7)$$

2) Transformation of an arbitrary operator

In the F representation,

$$(A_F)_{ij} = \langle i|\hat{A}|j\rangle. \quad (1.6.8)$$

In the G representation,

$$\begin{aligned} (A_G)_{mn} &= \langle m|\hat{A}|n\rangle = \sum_{ij} \langle m|i\rangle \langle i|\hat{A}|j\rangle \langle j|n\rangle \\ &= \sum_{ij} S_{im}^* (A_F)_{ij} S_{jn} = \sum_{ij} (S^\dagger)_{mi} (A_F)_{ij} S_{jn}. \end{aligned} \quad (1.6.9)$$

In matrix form,

$$\boxed{A_G = S^\dagger A_F S.} \quad (1.6.10)$$

3) The representation-transformation matrix is unitary

$$(SS^\dagger)_{ij} = \sum_m S_{im}(S^\dagger)_{mj} = \sum_m S_{im}S_{jm}^* = \sum_m \langle i|m\rangle \langle m|j\rangle = \langle i|j\rangle = \delta_{ij}. \quad (1.6.11)$$

Therefore $SS^\dagger = \mathbf{1}$. Similarly, $S^\dagger S = \mathbf{1}$, and hence

$$\boxed{S^\dagger = S^{-1}}. \quad (1.6.12)$$

A representation-transformation matrix is unitary.

Note that S is not necessarily Hermitian: $S^\dagger \neq S$.

4) Transformation of representation does not change eigenvalues

Suppose

$$A_F a_F = \lambda_F a_F. \quad (1.6.13)$$

Then

$$\begin{aligned} A_G a_G &= S^\dagger A_F S S^\dagger a_F = S^\dagger A_F a_F = S^\dagger \lambda_F a_F \\ &= \lambda_F S^\dagger a_F = \lambda_F a_G. \end{aligned} \quad (1.6.14)$$

Thus the eigenvalue in the G representation is still λ_F . A transformation of representation does not change the values of physical quantities.

Example 1: A transformation of representation does not change commutation relations.

If $[A_I, B_I] = C_I$ in the F representation, then in the G representation,

$$\begin{aligned} [A_M, B_M] &= A_M B_M - B_M A_M \\ &= S^\dagger A_I S S^\dagger B_I S - S^\dagger B_I S S^\dagger A_I S \\ &= S^\dagger (A_I B_I - B_I A_I) S = S^\dagger C_I S = C_M. \end{aligned} \quad (1.6.15)$$

Commutation relations are fundamental relations of quantum mechanics and are unchanged by the transformation of representation.

Example 2: A transformation of representation does not change the trace of a matrix:

$$\text{tr}(A_G) = \text{tr}(S^\dagger A_F S) = \text{tr}(S S^\dagger A_F) = \text{tr}(A_F). \quad (1.6.16)$$

Interactive investigation IC-01-05. Open Measurement axes and basis changes. After the unitary-change-of-basis derivation, rotate the spin measurement axes and transform the state components. Verify that the matrix entries change while the probabilities for a fixed physical experiment remain invariant.

1.7 Position and Momentum Representations**1) Continuous spectrum**

Unlike the spin angular momentum $\hat{s}_z$, which takes discrete values, position $\hat{x}$ and momentum $\hat{p}$ take continuous values.

The eigenstates of position or momentum constitute a continuous position or momentum representation.

Eigenvalue equations:

$$\hat{x} |x\rangle = x |x\rangle, \quad \hat{p} |p\rangle = p |p\rangle. \quad (1.7.1)$$

Basis vectors: $|x\rangle, |p\rangle$.

Orthonormality:

$$\langle x|x'\rangle = \delta(x - x'), \quad \langle p|p'\rangle = \delta(p - p'). \quad (1.7.2)$$

Completeness:

$$\int dx |x\rangle \langle x| = \mathbf{1}, \quad \int dp |p\rangle \langle p| = \mathbf{1}. \quad (1.7.3)$$

For an arbitrary state $|\psi\rangle$,

$$|\psi\rangle = \int dx |x\rangle \langle x|\psi\rangle = \int dp |p\rangle \langle p|\psi\rangle. \quad (1.7.4)$$

The functions $\langle x|\psi\rangle$ and $\langle p|\psi\rangle$ are the concrete forms of $|\psi\rangle$ in the position and momentum representations; they are continuous column matrices.

In three-dimensional space,

$$\hat{x} |\mathbf{x}\rangle = \mathbf{x} |\mathbf{x}\rangle, \quad \hat{\mathbf{p}} |\mathbf{p}\rangle = \mathbf{p} |\mathbf{p}\rangle. \quad (1.7.5)$$

Postulate of quantum mechanics: The fundamental commutation relations are

$$[\hat{x}_i, \hat{x}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}, \quad i, j = 1, 2, 3. \quad (1.7.6)$$

The position–momentum uncertainty relation is

$$\langle(\Delta\hat{x})^2\rangle\langle(\Delta\hat{p}_x)^2\rangle \geq \left(\frac{1}{2i}\langle[\hat{x}, \hat{p}_x]\rangle\right)^2 = \frac{\hbar^2}{4}. \quad (1.7.7)$$

2) Spatial translations

Define the spatial-translation operator $\hat{U}_{\text{tr}}$, with the subscript “tr” indicating translation (distinct from the scattering transition operator $\hat{T}$ introduced later), by

$$\hat{U}_{\text{tr}}(d\mathbf{x}) |\mathbf{x}\rangle = |\mathbf{x} + d\mathbf{x}\rangle. \quad (1.7.8)$$

For an arbitrary state $|\psi\rangle$,

$$\begin{aligned} \hat{U}_{\text{tr}}(d\mathbf{x}) |\psi\rangle &= \hat{U}_{\text{tr}}(d\mathbf{x}) \int d^3x |\mathbf{x}\rangle \langle \mathbf{x}|\psi\rangle \\ &= \int d^3x |\mathbf{x} + d\mathbf{x}\rangle \langle \mathbf{x}|\psi\rangle \quad (\text{the operator acts only on vectors}) \\ &= \int d^3x |\mathbf{x}\rangle \langle \mathbf{x} - d\mathbf{x}|\psi\rangle \quad (\text{change of integration variable}). \end{aligned} \quad (1.7.9)$$

For an infinitesimal translation, take

$$\hat{U}_{\text{tr}}(d\mathbf{x}) = \mathbf{1} - i\hat{\mathbf{K}} \cdot d\mathbf{x}, \quad (1.7.10)$$

where the generator $\hat{\mathbf{K}}$ is Hermitian.

A translation is a unitary operator:

$$\hat{U}_{\text{tr}}(d\mathbf{x}) \hat{U}_{\text{tr}}^\dagger(d\mathbf{x}) = (\mathbf{1} - i\hat{\mathbf{K}} \cdot d\mathbf{x})(\mathbf{1} + i\hat{\mathbf{K}} \cdot d\mathbf{x}) = \mathbf{1}, \quad (1.7.11)$$

neglecting second-order infinitesimals.

Composition law for translations:

$$\begin{aligned} \hat{U}_{\text{tr}}(d\mathbf{x}_1) \hat{U}_{\text{tr}}(d\mathbf{x}_2) &= (\mathbf{1} - i\hat{\mathbf{K}} \cdot d\mathbf{x}_1)(\mathbf{1} - i\hat{\mathbf{K}} \cdot d\mathbf{x}_2) \\ &= \mathbf{1} - i\hat{\mathbf{K}} \cdot (d\mathbf{x}_1 + d\mathbf{x}_2) = \hat{U}_{\text{tr}}(d\mathbf{x}_1 + d\mathbf{x}_2). \end{aligned} \quad (1.7.12)$$

A finite translation is obtained by repeated infinitesimal translations.

For an arbitrary state $|\psi\rangle$,

$$\begin{aligned}\hat{\mathbf{x}}\hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x})|\psi\rangle &= \int d^3x \hat{\mathbf{x}}|\mathbf{x} + \mathbf{d}\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle \\ &= \int d^3x (\mathbf{x} + \mathbf{d}\mathbf{x})|\mathbf{x} + \mathbf{d}\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle, \\ \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x})\hat{\mathbf{x}}|\psi\rangle &= \int d^3x \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x})\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle \\ &= \int d^3x \mathbf{x}|\mathbf{x} + \mathbf{d}\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle.\end{aligned}\tag{1.7.13}$$

Hence

$$\begin{aligned}[\hat{\mathbf{x}}, \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x})]|\psi\rangle &= \int d^3x \mathbf{d}\mathbf{x}|\mathbf{x} + \mathbf{d}\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle \\ &= \int d^3x \mathbf{d}\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle = \mathbf{d}\mathbf{x}|\psi\rangle,\end{aligned}\tag{1.7.14}$$

where second-order infinitesimals have been neglected. Thus

$$[\hat{\mathbf{x}}, \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x})] = \mathbf{d}\mathbf{x}, \quad [\hat{\mathbf{x}}, \hat{\mathbf{K}} \cdot \mathbf{d}\mathbf{x}] = \text{id}\mathbf{x}.\tag{1.7.15}$$

If the translation is taken in the x_j direction, $\mathbf{d}\mathbf{x} = \epsilon \mathbf{e}_j$, then

$$[\hat{x}_i, \hat{K}_j] = i\delta_{ij}.\tag{1.7.16}$$

Comparing with the fundamental quantum-mechanical commutation relation,

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij},\tag{1.7.17}$$

we have

$$\hat{\mathbf{p}} = \hbar\hat{\mathbf{K}}.\tag{1.7.18}$$

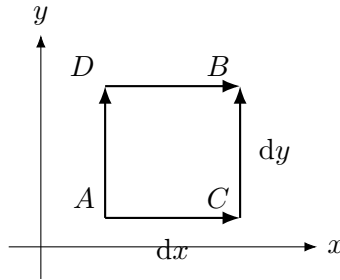
Momentum is therefore the generator of translations, and

$$\boxed{\hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x}) = \mathbf{1} - \frac{i}{\hbar}\hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x}.}\tag{1.7.19}$$

Since $[\hat{p}_x, \hat{p}_y] = 0$,

$$[\hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x}), \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{y})] = \left[\mathbf{1} - \frac{i}{\hbar}\hat{p}_x \mathbf{d}\mathbf{x}, \mathbf{1} - \frac{i}{\hbar}\hat{p}_y \mathbf{d}\mathbf{y}\right] = 0.\tag{1.7.20}$$

This shows that the order of translations is immaterial. Translating $A \rightarrow C$ by $\mathbf{d}\mathbf{x}$ and then $C \rightarrow B$ by $\mathbf{d}\mathbf{y}$ gives the same result as translating $A \rightarrow D$ by $\mathbf{d}\mathbf{y}$ and then $D \rightarrow B$ by $\mathbf{d}\mathbf{x}$.



3) Canonical quantization

Dirac, 1925: quantization of classical mechanical systems.

The Poisson bracket of analytical mechanics is replaced by a commutation relation in quantum mechanics:

$$[A, B]_{\text{P}} = \sum_s \left(\frac{\partial A}{\partial q_s} \frac{\partial B}{\partial p_s} - \frac{\partial A}{\partial p_s} \frac{\partial B}{\partial q_s} \right) \longrightarrow \frac{1}{i\hbar}[\hat{A}, \hat{B}] = \frac{1}{i\hbar}(\hat{A}\hat{B} - \hat{B}\hat{A}).\tag{1.7.21}$$

For example,

$$[x_i, p_j]_{\text{P}} = \delta_{ij} \longrightarrow [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}.\tag{1.7.22}$$

4) Position representation

The matrix elements of $\hat{\mathbf{x}}$ are

$$\langle \mathbf{x}_1 | \hat{\mathbf{x}} | \mathbf{x}_2 \rangle = \mathbf{x}_2 \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{x}_1 \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2), \quad (1.7.23)$$

so the position operator is diagonal.

From

$$\hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x}) |\beta\rangle = \left(\mathbf{1} - \frac{i}{\hbar} \hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x} \right) |\beta\rangle \quad (1.7.24)$$

and

$$\begin{aligned} \hat{U}_{\text{tr}}(\mathbf{d}\mathbf{x}) |\beta\rangle &= \int d^3x |\mathbf{x}\rangle \langle \mathbf{x} - \mathbf{d}\mathbf{x} | \beta \rangle \\ &= \int d^3x |\mathbf{x}\rangle (\langle \mathbf{x} | \beta \rangle - \mathbf{d}\mathbf{x} \cdot \nabla_{\mathbf{x}} \langle \mathbf{x} | \beta \rangle) \\ &= |\beta\rangle - \mathbf{d}\mathbf{x} \cdot \int d^3x |\mathbf{x}\rangle \nabla_{\mathbf{x}} \langle \mathbf{x} | \beta \rangle, \end{aligned} \quad (1.7.25)$$

we obtain

$$\hat{\mathbf{p}} |\beta\rangle = \int d^3x |\mathbf{x}\rangle (-i\hbar \nabla_{\mathbf{x}}) \langle \mathbf{x} | \beta \rangle, \quad (1.7.26)$$

and

$$\langle \alpha | \hat{\mathbf{p}} | \beta \rangle = \int d^3x \langle \alpha | \mathbf{x} \rangle (-i\hbar \nabla_{\mathbf{x}}) \langle \mathbf{x} | \beta \rangle. \quad (1.7.27)$$

Taking $|\alpha\rangle = |\mathbf{x}_1\rangle$ and $|\beta\rangle = |\mathbf{x}_2\rangle$ gives the matrix elements of momentum in the position representation:

$$\langle \mathbf{x}_1 | \hat{\mathbf{p}} | \mathbf{x}_2 \rangle = \int d^3x \langle \mathbf{x}_1 | \mathbf{x} \rangle (-i\hbar \nabla_{\mathbf{x}}) \langle \mathbf{x} | \mathbf{x}_2 \rangle = -i\hbar \nabla_{\mathbf{x}_1} \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2). \quad (1.7.28)$$

Now take $|\alpha\rangle = |\mathbf{x}_1\rangle$ and $|\beta\rangle = |\mathbf{p}\rangle$:

$$\langle \mathbf{x}_1 | \hat{\mathbf{p}} | \mathbf{p} \rangle = \int d^3x \langle \mathbf{x}_1 | \mathbf{x} \rangle (-i\hbar \nabla_{\mathbf{x}}) \langle \mathbf{x} | \mathbf{p} \rangle = -i\hbar \nabla_{\mathbf{x}_1} \langle \mathbf{x}_1 | \mathbf{p} \rangle. \quad (1.7.29)$$

But the momentum eigenvalue equation gives

$$\mathbf{p} \langle \mathbf{x}_1 | \mathbf{p} \rangle = -i\hbar \nabla_{\mathbf{x}_1} \langle \mathbf{x}_1 | \mathbf{p} \rangle. \quad (1.7.30)$$

Solving this first-order differential equation, we find that the momentum eigenstate in the position representation is a plane wave:

$$\langle \mathbf{x} | \mathbf{p} \rangle = N \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}\right). \quad (1.7.31)$$

Using

$$\begin{aligned} \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle &= \int d^3p \langle \mathbf{x}_1 | \mathbf{p} \rangle \langle \mathbf{p} | \mathbf{x}_2 \rangle \\ &= |N|^2 \int d^3p \exp\left[\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{x}_1 - \mathbf{x}_2)\right], \end{aligned} \quad (1.7.32)$$

and

$$\delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2) = |N|^2 (2\pi\hbar)^3 \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2), \quad (1.7.33)$$

we obtain

$$N = \frac{1}{(2\pi\hbar)^{3/2}}, \quad \boxed{\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}\right)}. \quad (1.7.34)$$

5) Momentum representation

The matrix elements of the momentum operator are

$$\langle \mathbf{p}_1 | \hat{\mathbf{p}} | \mathbf{p}_2 \rangle = \mathbf{p}_1 \delta^{(3)}(\mathbf{p}_1 - \mathbf{p}_2), \quad (1.7.35)$$

so the momentum operator is diagonal.

For the position operator,

$$\begin{aligned} \langle \mathbf{p}_1 | \hat{\mathbf{x}} | \mathbf{p}_2 \rangle &= \int d^3x_1 d^3x_2 \langle \mathbf{p}_1 | \mathbf{x}_1 \rangle \langle \mathbf{x}_1 | \hat{\mathbf{x}} | \mathbf{x}_2 \rangle \langle \mathbf{x}_2 | \mathbf{p}_2 \rangle \\ &= \frac{1}{(2\pi\hbar)^3} \int d^3x_1 d^3x_2 \exp\left(-\frac{i}{\hbar} \mathbf{p}_1 \cdot \mathbf{x}_1\right) \mathbf{x}_2 \delta^{(3)}(\mathbf{x}_1 - \mathbf{x}_2) \exp\left(\frac{i}{\hbar} \mathbf{p}_2 \cdot \mathbf{x}_2\right) \\ &= \frac{1}{(2\pi\hbar)^3} \int d^3x_1 \mathbf{x}_1 \exp\left[\frac{i}{\hbar} (\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{x}_1\right] \\ &= i\hbar \nabla_{\mathbf{p}_1} \frac{1}{(2\pi\hbar)^3} \int d^3x_1 \exp\left[\frac{i}{\hbar} (\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{x}_1\right] \\ &= i\hbar \nabla_{\mathbf{p}_1} \delta^{(3)}(\mathbf{p}_1 - \mathbf{p}_2). \end{aligned} \quad (1.7.36)$$

Note that differentiation here is with respect to momentum.

The form of the position eigenstate in the momentum representation is

$$\boxed{\langle \mathbf{p} | \mathbf{x} \rangle = \langle \mathbf{x} | \mathbf{p} \rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{x}\right).} \quad (1.7.37)$$

Example 1: Prove that $\hat{x}$ and $\hat{p}$ are Hermitian operators in the position representation. For position,

$$\langle x_1 | \hat{x} | x_2 \rangle = \delta(x_1 - x_2) x_1, \quad (1.7.38)$$

and

$$\begin{aligned} \langle x_1 | \hat{x}^\dagger | x_2 \rangle &= \langle x_2 | \hat{x} | x_1 \rangle^* = [\delta(x_2 - x_1) x_2]^* = \delta(x_2 - x_1) x_2 \\ &= \delta(x_1 - x_2) x_1 = \langle x_1 | \hat{x} | x_2 \rangle. \end{aligned} \quad (1.7.39)$$

For momentum,

$$\langle x_1 | \hat{p} | x_2 \rangle = -i\hbar \frac{\partial}{\partial x_1} \delta(x_1 - x_2) = -i\hbar \frac{\partial}{\partial(x_1 - x_2)} \delta(x_1 - x_2). \quad (1.7.40)$$

Also,

$$\begin{aligned} \langle x_1 | \hat{p}^\dagger | x_2 \rangle &= \langle x_2 | \hat{p} | x_1 \rangle^* = \left[-i\hbar \frac{\partial}{\partial(x_2 - x_1)} \delta(x_2 - x_1) \right]^* \\ &= i\hbar \frac{\partial}{\partial(x_2 - x_1)} \delta(x_2 - x_1) = -i\hbar \frac{\partial}{\partial(x_1 - x_2)} \delta(x_1 - x_2) \\ &= \langle x_1 | \hat{p} | x_2 \rangle. \end{aligned} \quad (1.7.41)$$

Thus both $\hat{x}$ and $\hat{p}$ are Hermitian.

Example 2: Compute $[\hat{x}, \hat{p}]$ in the position representation.

$$\begin{aligned}
\hat{x}\hat{p} &= \int dx_1 dx_2 \hat{x} |x_1\rangle \langle x_1| \hat{p} |x_2\rangle \langle x_2| \\
&= \int dx_1 dx_2 |x_1\rangle x_1 \left(-i\hbar \frac{\partial}{\partial x_1} \right) \delta(x_1 - x_2) \langle x_2| \\
&= - \int dx_1 dx_2 |x_1\rangle \left(-i\hbar \frac{\partial}{\partial x_1} \right) [x_1 \langle x_1|] \delta(x_1 - x_2) \quad (\text{integration by parts}) \\
&= - \int dx_1 |x_1\rangle \left(-i\hbar \frac{\partial}{\partial x_1} \right) [x_1 \langle x_1|] \\
&= \int dx_1 x_1 |x_1\rangle \left(-i\hbar \frac{\partial}{\partial x_1} \right) \langle x_1| \quad (\text{integration by parts}).
\end{aligned} \tag{1.7.42}$$

Similarly,

$$\hat{p}\hat{x} = -i\hbar + \int dx_1 x_1 |x_1\rangle \left(-i\hbar \frac{\partial}{\partial x_1} \right) \langle x_1|. \tag{1.7.43}$$

Therefore

$$[\hat{x}, \hat{p}] = i\hbar. \tag{1.7.44}$$

This shows that commutation relations do not depend on the representation.

According to the uncertainty relation, position and momentum cannot simultaneously have definite values. For example, in a momentum eigenstate the momentum has a definite value, $\langle p | (\Delta \hat{p})^2 | p \rangle = 0$, but the probability for position to take value x is

$$|\langle x | p \rangle|^2 = \left| \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \right|^2 = \text{const}. \tag{1.7.45}$$

The particle therefore has equal probability of appearing anywhere in $-\infty < x < \infty$; position is completely indefinite, with variance $\langle p | (\Delta \hat{x})^2 | p \rangle = \infty$.

Example 3: Minimum-uncertainty wave packet. For equality to hold in the uncertainty relation

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2, \tag{1.7.46}$$

so that the uncertainty is minimal, the state must satisfy:

1) Equality in the Schwarz inequality:

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle = |\langle \Delta \hat{A} \Delta \hat{B} \rangle|^2. \tag{1.7.47}$$

2) $\text{Re} \langle \Delta \hat{A} \Delta \hat{B} \rangle = 0$.

The solution of condition 1 is

$$\Delta \hat{B} |\psi\rangle = c \Delta \hat{A} |\psi\rangle, \tag{1.7.48}$$

where c is a constant. Condition 2 implies

$$\text{Re} \left(c \langle (\Delta \hat{A})^2 \rangle \right) = 0. \tag{1.7.49}$$

Since $\langle (\Delta \hat{A})^2 \rangle$ is real, $c = ia$, with a real. Hence

$$(\hat{B} - \langle \hat{B} \rangle) |\psi\rangle = ia(\hat{A} - \langle \hat{A} \rangle) |\psi\rangle. \tag{1.7.50}$$

This is the constraint on $|\psi\rangle$ for minimum uncertainty. Choose

$$\hat{A} = \hat{x}, \quad \hat{B} = \hat{p}, \tag{1.7.51}$$

so

$$(\hat{p} - \langle \hat{p} \rangle) |\psi\rangle = ia(\hat{x} - \langle \hat{x} \rangle) |\psi\rangle. \quad (1.7.52)$$

In the position representation,

$$\left(-i\hbar \frac{d}{dx} - \langle \hat{p} \rangle\right) \psi(x) = ia(x - \langle \hat{x} \rangle) \psi(x), \quad \psi(x) = \langle x | \psi \rangle. \quad (1.7.53)$$

The solution is

$$\psi(x) = A \exp \left[-\frac{a(x - \langle \hat{x} \rangle)^2}{2\hbar} \right] \exp \left(\frac{i}{\hbar} \langle \hat{p} \rangle x \right), \quad (1.7.54)$$

a Gaussian wave packet in position space.

Example 4: The ground state of the hydrogen atom. Given the state in the position representation

$$\langle \mathbf{r} | \psi \rangle = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (\text{hydrogen-atom ground state}), \quad (1.7.55)$$

find the expectation value of the momentum.

Since $\langle \mathbf{r} | \psi \rangle$ is not a plane wave, momentum has no definite value. The probability for momentum $\mathbf{p}$ is $|\langle \mathbf{p} | \psi \rangle|^2$, where

$$\begin{aligned} \langle \mathbf{p} | \psi \rangle &= \int d^3r \langle \mathbf{p} | \mathbf{r} \rangle \langle \mathbf{r} | \psi \rangle = \int d^3r \langle \mathbf{r} | \mathbf{p} \rangle^* \langle \mathbf{r} | \psi \rangle \\ &= \int d^3r \frac{1}{(2\pi\hbar)^{3/2}} \exp \left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r} \right) \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \\ &= \frac{(2a_0\hbar)^{3/2} \hbar}{\pi(a_0^2 p^2 + \hbar^2)^2}. \end{aligned} \quad (1.7.56)$$

Since $\langle \mathbf{p} | \psi \rangle$ depends only on the magnitude of momentum, the expectation value of the momentum magnitude is

$$\langle p \rangle = \frac{\int dp p^3 |\langle \mathbf{p} | \psi \rangle|^2}{\int dp p^2 |\langle \mathbf{p} | \psi \rangle|^2}. \quad (1.7.57)$$

Chapter 2

Quantum Dynamics

2.1 Time Evolution and the Schrödinger Equation

This chapter discusses the time evolution of states and physical quantities in quantum mechanics. Unlike other physical quantities, which are represented by operators, time in nonrelativistic quantum mechanics is only a parameter; it remains a classical quantity and is not an operator.

1) Time-evolution (translation) operator

Initial state: $|\alpha, t_0\rangle$.

Evolution from t_0 to t :

$$|\alpha, t\rangle = \hat{U}(t, t_0) |\alpha, t_0\rangle. \quad (2.1.1)$$

Expand $|\alpha, t_0\rangle$ and $|\alpha, t\rangle$ in eigenstates $|a\rangle$ of an operator $\hat{A}$:

$$|\alpha, t_0\rangle = \sum_a |a\rangle \langle a|\alpha, t_0\rangle, \quad |\alpha, t\rangle = \sum_a |a\rangle \langle a|\alpha, t\rangle. \quad (2.1.2)$$

Clearly, the probability of occupying a particular eigenstate is not necessarily conserved:

$$|\langle a|\alpha, t_0\rangle|^2 \neq |\langle a|\alpha, t\rangle|^2. \quad (2.1.3)$$

However, the normalization condition

$$\langle \alpha, t|\alpha, t\rangle = \langle \alpha, t_0|\alpha, t_0\rangle = 1 \quad (2.1.4)$$

implies conservation of the total probability,

$$\sum_a |\langle a|\alpha, t\rangle|^2 = \sum_a |\langle a|\alpha, t_0\rangle|^2 = 1, \quad (2.1.5)$$

and the time-evolution operator is unitary:

$$\hat{U}^\dagger(t, t_0) \hat{U}(t, t_0) = \mathbf{1}. \quad (2.1.6)$$

Require the composition law for time evolution:

$$\hat{U}(t_2, t_0) = \hat{U}(t_2, t_1) \hat{U}(t_1, t_0), \quad t_2 > t_1 > t_0. \quad (2.1.7)$$

An infinitesimal time-evolution operator can be written

$$\hat{U}(t_0 + dt, t_0) = \mathbf{1} - i\hat{\Omega} dt. \quad (2.1.8)$$

Unitarity, $\hat{U}^\dagger \hat{U} = \mathbf{1}$, requires $\hat{\Omega}$ to be Hermitian:

$$\hat{\Omega}^\dagger = \hat{\Omega}. \quad (2.1.9)$$

Since $\hat{U}$ is dimensionless, $\hat{\Omega}$ has dimensions of frequency. By the Planck–Einstein energy–frequency relation $E = \hbar\omega$, we may set

$$\hat{\Omega} = \frac{\hat{H}}{\hbar}, \quad \boxed{\hat{U}(t_0 + dt, t_0) = \mathbf{1} - \frac{i}{\hbar} \hat{H} dt.} \quad (2.1.10)$$

2) Equation of time evolution

Using the composition law,

$$\hat{U}(t + dt, t_0) = \hat{U}(t + dt, t) \hat{U}(t, t_0) = \left(1 - \frac{i}{\hbar} \hat{H} dt\right) \hat{U}(t, t_0). \quad (2.1.11)$$

Hence

$$\frac{\hat{U}(t + dt, t_0) - \hat{U}(t, t_0)}{dt} = -\frac{i}{\hbar} \hat{H} \hat{U}(t, t_0), \quad (2.1.12)$$

so that

$$\boxed{i\hbar \frac{\partial}{\partial t} \hat{U}(t, t_0) = \hat{H} \hat{U}(t, t_0).} \quad (2.1.13)$$

This is the fundamental equation of time evolution in quantum mechanics.

Multiplying the equation on the right by the initial state $|\alpha, t_0\rangle$ gives

$$i\hbar \frac{\partial}{\partial t} \hat{U}(t, t_0) |\alpha, t_0\rangle = \hat{H} \hat{U}(t, t_0) |\alpha, t_0\rangle, \quad (2.1.14)$$

$$\boxed{i\hbar \frac{\partial}{\partial t} |\alpha, t\rangle = \hat{H} |\alpha, t\rangle.} \quad (2.1.15)$$

This is the equation governing the time evolution of a state, the Schrödinger equation.

We now discuss solutions of this fundamental equation. If the Hamiltonian has no explicit time dependence, for example the magnetic interaction $\hat{H} \sim \hat{J}_z B$ in a constant magnetic field directed along z ,

$$\boxed{\hat{U}(t, t_0) = \exp \left[-\frac{i}{\hbar} \hat{H} (t - t_0) \right].} \quad (2.1.16)$$

If the Hamiltonian depends on time but the Hamiltonians at different times commute, for example $\hat{H}(t) \sim \hat{J}_z B(t)$ when the magnitude but not the direction of the magnetic field changes,

$$\hat{U}(t, t_0) = \exp \left[-\frac{i}{\hbar} \int_{t_0}^t \hat{H}(t') dt' \right]. \quad (2.1.17)$$

If the Hamiltonian depends on time and the Hamiltonians at different times do not commute, for example $\hat{H}(t) \sim \mathbf{J} \cdot \mathbf{B}(t)$ when both the magnitude and direction of the field vary,

$$\begin{aligned} \hat{U}(t, t_0) = \mathbf{1} + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar} \right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \\ \times \hat{H}(t_1) \hat{H}(t_2) \cdots \hat{H}(t_n). \end{aligned} \quad (2.1.18)$$

This is called the Dyson series.

Below we mainly consider Hamiltonians that do not depend explicitly on time.

3) Stationary states

Suppose initially the system is in an eigenstate of $\hat{H}$:

$$|\alpha, t_0\rangle = |E\rangle, \quad \hat{H} |E\rangle = E |E\rangle. \quad (2.1.19)$$

Then

$$|\alpha, t\rangle = e^{-i\hat{H}(t-t_0)/\hbar} |\alpha, t_0\rangle = e^{-iE(t-t_0)/\hbar} |\alpha, t_0\rangle. \quad (2.1.20)$$

The state changes only by a phase. The expectation value of any physical quantity $\hat{A}$ is

$$\langle \alpha, t | \hat{A} | \alpha, t \rangle = \langle \alpha, t_0 | e^{iE(t-t_0)/\hbar} \hat{A} e^{-iE(t-t_0)/\hbar} | \alpha, t_0 \rangle = \langle \alpha, t_0 | \hat{A} | \alpha, t_0 \rangle \quad (2.1.21)$$

and does not change with time. Therefore, energy eigenstates are called stationary states.

If the system is initially not in a stationary state, its amplitude for occupying the stationary state $|E\rangle$ at time t is

$$\langle E|\alpha, t\rangle = \langle E|e^{-i\hat{H}(t-t_0)/\hbar}|\alpha, t_0\rangle = e^{-iE(t-t_0)/\hbar}\langle E|\alpha, t_0\rangle, \quad (2.1.22)$$

and its probability satisfies

$$|\langle E|\alpha, t\rangle|^2 = |\langle E|\alpha, t_0\rangle|^2. \quad (2.1.23)$$

The probability of occupying any stationary state is the same at all times as it was initially. Thus, if the system initially occupies a stationary state, it remains in that state; if its initial state is not stationary, it remains nonstationary.

4) Electron-spin precession

Consider the interaction of an electron's spin magnetic moment with an external magnetic field directed along z :

$$\hat{H} = -\left(\frac{e}{mc}\right)\hat{\mathbf{s}} \cdot \mathbf{B} = -\frac{eB}{mc}\hat{s}_z = \omega\hat{s}_z, \quad \omega = -\frac{eB}{mc} > 0. \quad (2.1.24)$$

The time-evolution operator is

$$\hat{U}(t, 0) = e^{-i\hat{H}t/\hbar} = e^{-i\omega\hat{s}_zt/\hbar}. \quad (2.1.25)$$

The operators $\hat{s}_z$ and $\hat{H}$ have common eigenstates $|s_z^\pm\rangle$:

$$\hat{s}_z|s_z^\pm\rangle = \pm\frac{\hbar}{2}|s_z^\pm\rangle, \quad \hat{H}|s_z^\pm\rangle = \pm\frac{\hbar\omega}{2}|s_z^\pm\rangle. \quad (2.1.26)$$

Suppose the system initially occupies the positive eigenstate of $\hat{s}_x$:

$$|\alpha, 0\rangle = |s_x^+\rangle = \frac{1}{\sqrt{2}}|s_z^+\rangle + \frac{1}{\sqrt{2}}|s_z^-\rangle. \quad (2.1.27)$$

At an arbitrary time,

$$|\alpha, t\rangle = \hat{U}(t, 0)|\alpha, 0\rangle. \quad (2.1.28)$$

The probability of finding the system in $|s_x^+\rangle$ is

$$\begin{aligned} |\langle s_x^+|\alpha, t\rangle|^2 &= \left| \left(\frac{\langle s_z^+| + \langle s_z^-|}{\sqrt{2}} \right) e^{-i\omega\hat{s}_zt/\hbar} \left(\frac{|s_z^+\rangle + |s_z^-\rangle}{\sqrt{2}} \right) \right|^2 \\ &= \left| \frac{e^{-i\omega t/2} + e^{i\omega t/2}}{2} \right|^2 = \cos^2 \frac{\omega t}{2}. \end{aligned} \quad (2.1.29)$$

By conservation of probability,

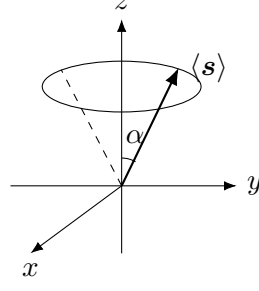
$$|\langle s_x^-|\alpha, t\rangle|^2 = 1 - |\langle s_x^+|\alpha, t\rangle|^2 = \sin^2 \frac{\omega t}{2}. \quad (2.1.30)$$

The expectation value at an arbitrary time is

$$\begin{aligned} \langle s_x \rangle &= \langle \alpha, t|\hat{s}_x|\alpha, t\rangle = \sum_{s_x} s_x |\langle s_x|\alpha, t\rangle|^2 \\ &= \frac{\hbar}{2} \left(\cos^2 \frac{\omega t}{2} - \sin^2 \frac{\omega t}{2} \right) = \frac{\hbar}{2} \cos \omega t. \end{aligned} \quad (2.1.31)$$

Similarly,

$$\langle s_y \rangle = \frac{\hbar}{2} \sin \omega t, \quad \langle s_z \rangle = 0. \quad (2.1.32)$$



The expectation value $\langle \mathbf{s} \rangle$ precesses about the z axis with time; the trajectory is a cone. This is the spin Larmor precession, with ω the Larmor frequency.

Interactive investigation IC-02-01. Open Spin precession. After deriving spin precession, adjust the magnetic field and the initial spin direction. Track the Bloch vector and the spin probabilities, and check the precession frequency and norm conservation.

5) Energy–time uncertainty relation

From

$$\langle (\Delta \hat{x})^2 \rangle \langle (\Delta \hat{p})^2 \rangle \geq \frac{\hbar^2}{4}, \quad (2.1.33)$$

we obtain $\Delta x \Delta p \geq \hbar/2$, where

$$\Delta A = \sqrt{\langle (\Delta \hat{A})^2 \rangle}. \quad (2.1.34)$$

By classical analogy, $x = x(t)$ and $p = p(E, p)$, one may suppose

$$\Delta t \Delta E \geq \frac{\hbar}{2}. \quad (2.1.35)$$

This is the energy–time uncertainty relation. However, unlike x , p , E , $\hat{x}$, $\hat{p}$, and $\hat{H}$ in quantum mechanics, time t remains a classical quantity. What does Δt mean?

What are $\langle (\Delta \hat{H})^2 \rangle$ and $\langle (\Delta t)^2 \rangle$?

For an arbitrary observable $\hat{O}$,

$$\langle O \rangle(t) = \langle \alpha, t | \hat{O} | \alpha, t \rangle. \quad (2.1.36)$$

Its time derivative is

$$\frac{d}{dt} \langle O \rangle = \left(\frac{\partial}{\partial t} \langle \alpha, t | \right) \hat{O} | \alpha, t \rangle + \langle \alpha, t | \frac{\partial \hat{O}}{\partial t} | \alpha, t \rangle + \langle \alpha, t | \hat{O} \left(\frac{\partial}{\partial t} | \alpha, t \rangle \right). \quad (2.1.37)$$

Since

$$i\hbar \frac{\partial}{\partial t} | \alpha, t \rangle = \hat{H} | \alpha, t \rangle, \quad -i\hbar \frac{\partial}{\partial t} \langle \alpha, t | = \langle \alpha, t | \hat{H}, \quad (2.1.38)$$

we have

$$\frac{d}{dt} \langle O \rangle = \frac{1}{i\hbar} \langle \alpha, t | [\hat{O}, \hat{H}] | \alpha, t \rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle. \quad (2.1.39)$$

If $\hat{O}$ has no explicit time dependence,

$$\frac{d}{dt} \langle O \rangle = \frac{1}{i\hbar} \langle [\hat{O}, \hat{H}] \rangle. \quad (2.1.40)$$

Consequently, the uncertainty relation between $\hat{O}$ and $\hat{H}$ gives

$$\langle (\Delta \hat{O})^2 \rangle \langle (\Delta \hat{H})^2 \rangle \geq \left| \frac{1}{2i} \langle [\hat{O}, \hat{H}] \rangle \right|^2 = \frac{\hbar^2}{4} \left(\frac{d\langle O \rangle}{dt} \right)^2. \quad (2.1.41)$$

Therefore

$$\Delta O \Delta E \geq \frac{\hbar}{2} \left| \frac{d\langle O \rangle}{dt} \right|. \quad (2.1.42)$$

If we define

$$\Delta t = \Delta O \left| \frac{d\langle O \rangle}{dt} \right|^{-1}, \quad (2.1.43)$$

then

$$\boxed{\Delta t \Delta E \geq \frac{\hbar}{2}.} \quad (2.1.44)$$

Thus Δt is the time needed for the expectation value of a physical quantity $\hat{O}$ to change by an amount equal to one standard deviation ΔO . Clearly, both Δt and ΔE are averages and uncertainties of physical quantities.

Example. In a stationary state $\Delta E = 0$, and $\langle \hat{O} \rangle$ is time independent:

$$\frac{d}{dt}\langle O \rangle = 0, \quad \Delta t = \Delta O \left| \frac{d\langle O \rangle}{dt} \right|^{-1} \rightarrow \infty. \quad (2.1.45)$$

Worked derivation: from unitarity to the Schrödinger equation

The previous section introduced the infinitesimal time-evolution operator. Let us fill in the intermediate algebra so that the role of Hermiticity is visible, including all terms of first order in dt .

Take $\hat{U}(t + dt, t) = \mathbf{1} + \hat{A}dt + O((dt)^2)$ for some operator $\hat{A}$. Unitarity demands

$$\begin{aligned} \mathbf{1} &= \hat{U}^\dagger \hat{U} \\ &= [\mathbf{1} + \hat{A}^\dagger dt + O((dt)^2)] [\mathbf{1} + \hat{A}dt + O((dt)^2)] \\ &= \mathbf{1} + (\hat{A}^\dagger + \hat{A})dt + O((dt)^2). \end{aligned} \quad (2.1.46)$$

Comparing the coefficients of dt gives $\hat{A}^\dagger = -\hat{A}$. Thus the infinitesimal generator is anti-Hermitian. We can write $\hat{A} = -i\hat{H}/\hbar$, where

$$\hat{H}^\dagger = (i\hbar\hat{A})^\dagger = -i\hbar\hat{A}^\dagger = i\hbar\hat{A} = \hat{H}. \quad (2.1.47)$$

For a time-dependent generator the infinitesimal relation is

$$\hat{U}(t + dt, t) = \mathbf{1} - \frac{i}{\hbar} \hat{H}(t)dt + O((dt)^2). \quad (2.1.48)$$

The composition property gives

$$\begin{aligned} \hat{U}(t + dt, t_0) &= \hat{U}(t + dt, t) \hat{U}(t, t_0) \\ &= \hat{U}(t, t_0) - \frac{i}{\hbar} \hat{H}(t) \hat{U}(t, t_0)dt + O((dt)^2). \end{aligned} \quad (2.1.49)$$

Subtract $\hat{U}(t, t_0)$, divide by dt and take $dt \rightarrow 0$:

$$\boxed{i\hbar \partial_t \hat{U}(t, t_0) = \hat{H}(t) \hat{U}(t, t_0).} \quad (2.1.50)$$

Acting on the fixed initial ket $|\psi(t_0)\rangle$ immediately yields $i\hbar \partial_t |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle$.

What changes if Hamiltonians at different times do not commute? At two successive short intervals Δt , the evolution is

$$\hat{U}(t_2, t_0) \simeq \left[\mathbf{1} - \frac{i}{\hbar} \hat{H}(t_1) \Delta t \right] \left[\mathbf{1} - \frac{i}{\hbar} \hat{H}(t_0) \Delta t \right]. \quad (2.1.51)$$

Reversing their order changes the product by a term proportional to $[\hat{H}(t_1), \hat{H}(t_0)](\Delta t)^2$. For a finite interval these ordering effects accumulate and lead to the time-ordered exponential introduced later in this book. The expression $\exp[-(i/\hbar) \int \hat{H}(t) dt]$ is valid without time ordering only if the relevant Hamiltonians commute. In the special constant two-level example $\hat{H} = \frac{\hbar\Omega}{2} \sigma_x$, $\sigma_x^2 = \mathbf{1}$ gives $\hat{U} = \cos(\Omega t/2) \mathbf{1} - i \sin(\Omega t/2) \sigma_x$. Direct multiplication verifies $\hat{U}^\dagger \hat{U} = \mathbf{1}$, as expected.

Checkpoint. Repeat the first calculation using the *effective* Hamiltonian of Sec. 1.4. Identify precisely which step changes and use Eq. (1.4.13) to interpret the result.

2.2 Schrödinger and Heisenberg Pictures

1) Two kinds of unitary transformations

The spatial-translation operator $\hat{U}(\mathbf{d}\mathbf{x})$ and time-translation operator $\hat{U}(dt)$ discussed earlier both act on the state:

$$|\alpha\rangle \longrightarrow \hat{U}(\mathbf{d}\mathbf{x}) |\alpha\rangle = \left(\mathbf{1} - \frac{i}{\hbar} \hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x} \right) |\alpha\rangle, \quad (2.2.1)$$

$$|\alpha\rangle \longrightarrow \hat{U}(dt) |\alpha\rangle = \left(\mathbf{1} - \frac{i}{\hbar} \hat{H} dt \right) |\alpha\rangle. \quad (2.2.2)$$

The physical quantities remain unchanged; observable differences arise from changes in the state. Under a unitary change of the state,

$$\langle \beta | \hat{A} | \alpha \rangle \longrightarrow \langle \beta | \hat{U}^\dagger \hat{A} \hat{U} | \alpha \rangle. \quad (2.2.3)$$

Equivalently, one may leave the state $|\alpha\rangle$ unchanged and transform the physical quantity:

$$\boxed{\hat{A} \longrightarrow \hat{U}^\dagger \hat{A} \hat{U}, \quad |\alpha\rangle \text{ unchanged.}} \quad (2.2.4)$$

The two methods are completely equivalent: the time dependence may be assigned either to the state or to the observable. For example, under a spatial displacement,

$$\hat{x} \longrightarrow \hat{U}^\dagger(\mathbf{d}\mathbf{x}) \hat{x} \hat{U}(\mathbf{d}\mathbf{x}) = \left(\mathbf{1} + \frac{i}{\hbar} \hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x} \right) \hat{x} \left(\mathbf{1} - \frac{i}{\hbar} \hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x} \right) = \hat{x} + \frac{i}{\hbar} [\hat{\mathbf{p}} \cdot \mathbf{d}\mathbf{x}, \hat{x}] = \hat{x} + \mathbf{d}\mathbf{x}. \quad (2.2.5)$$

2) Schrödinger and Heisenberg pictures of time evolution

In the *Schrödinger picture*, the state evolves as

$$|\alpha, t\rangle_S = \hat{U}(t) |\alpha, 0\rangle_S, \quad \hat{U}(t) = e^{-i\hat{H}t/\hbar}, \quad (2.2.6)$$

while the observable $\hat{A}_S$ remains fixed in time.

In the *Heisenberg picture*, the observable evolves as

$$\hat{A}_H(t) = \hat{U}^\dagger(t) \hat{A}_H(0) \hat{U}(t), \quad (2.2.7)$$

while the state $|\alpha\rangle_H$ remains fixed.

For the same initial conditions,

$$\hat{A}_H(0) = \hat{A}_S, \quad |\alpha, 0\rangle_S = |\alpha\rangle_H. \quad (2.2.8)$$

These are the relations between the two pictures.

Notes.

1. Inner products (including matrix elements, probability amplitudes and expectation values) are independent of picture:

$${}_S \langle \alpha, t | \beta, t \rangle_S = {}_S \langle \alpha, 0 | \hat{U}^\dagger \hat{U} | \beta, 0 \rangle_S = {}_S \langle \alpha, 0 | \beta, 0 \rangle_S = {}_H \langle \alpha | \beta \rangle_H. \quad (2.2.9)$$

2. Since a unitary transformation preserves commutators, if the Schrödinger-picture commutation relation is

$$[\hat{A}_S, \hat{B}_S] = \hat{C}_S, \quad (2.2.10)$$

or equivalently initially $[\hat{A}_H(0), \hat{B}_H(0)] = \hat{C}_H(0)$, then the equal-time relation is

$$[\hat{A}_H(t), \hat{B}_H(t)] = \hat{C}_H(t). \quad (2.2.11)$$

3. The Hamiltonian in the previously introduced time-evolution operator $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$ is the Schrödinger-picture Hamiltonian $\hat{H} = \hat{H}_S$. For a time-independent Hamiltonian,

$$\hat{H}_H(t) = \hat{U}^\dagger(t) \hat{H}_S \hat{U}(t) = e^{i\hat{H}_S t/\hbar} \hat{H}_S e^{-i\hat{H}_S t/\hbar} = \hat{H}_S = \hat{H}_H. \quad (2.2.12)$$

So far we have considered time evolution of states and observables. Do basis vectors also change with time, as general vectors do? Basis vectors are eigenstates of observables; their time dependence is determined by that of the observables.

In the Schrödinger picture,

$$\hat{A}_S |a\rangle_S = a |a\rangle_S. \quad (2.2.13)$$

Since $\hat{A}_S$ is time independent, its eigenbasis $|a\rangle_S$ is also time independent.

In the Heisenberg picture,

$$\hat{A}_H(t) |a, t\rangle_H = a |a, t\rangle_H. \quad (2.2.14)$$

Since $\hat{A}_H$ depends on time, its basis $|a, t\rangle_H$ also depends on time. Using

$$\hat{U}^\dagger(t) \hat{A}_S \hat{U}(t) |a, t\rangle_H = a |a, t\rangle_H, \quad (2.2.15)$$

we obtain

$$\hat{A}_S \hat{U}(t) |a, t\rangle_H = a \hat{U}(t) |a, t\rangle_H. \quad (2.2.16)$$

Comparison with the Schrödinger-picture eigenvalue equation gives

$$\hat{U}(t) |a, t\rangle_H = |a\rangle_S, \quad |a, t\rangle_H = \hat{U}^\dagger(t) |a\rangle_S = \hat{U}^\dagger(t) |a, 0\rangle_H. \quad (2.2.17)$$

In summary:

	Schrödinger picture	Heisenberg picture
State	$ \alpha, t\rangle_S = \hat{U}(t) \alpha, 0\rangle_S$	$ \alpha\rangle_H$ fixed
Observable	$\hat{A}_S$ fixed	$\hat{A}_H(t) = \hat{U}^\dagger \hat{A}_H(0) \hat{U}$
Basis	$ a\rangle_S$ fixed	$ a, t\rangle_H = \hat{U}^\dagger(t) a, 0\rangle_H$

3) Equation of motion in the Heisenberg picture

In classical mechanics, only physical quantities evolve with time and the notion of a state vector is not introduced. The Heisenberg picture of quantum mechanics is therefore easier to compare with classical mechanics.

Differentiate

$$\hat{A}_H(t) = e^{i\hat{H}t/\hbar} \hat{A}_H(0) e^{-i\hat{H}t/\hbar} \quad (2.2.18)$$

with respect to time. We obtain the Heisenberg equation of motion:

$$\boxed{\frac{d\hat{A}_H(t)}{dt} = \frac{1}{i\hbar} [\hat{A}_H(t), \hat{H}].} \quad (2.2.19)$$

It plays a role similar to the Schrödinger-picture equation for the state,

$$i\hbar \frac{\partial}{\partial t} |\alpha, t\rangle_S = \hat{H} |\alpha, t\rangle_S. \quad (2.2.20)$$

Compare the equation of motion of analytical classical mechanics,

$$\frac{dA}{dt} = \{A, H\}_{\text{Poisson}}. \quad (2.2.21)$$

We obtain the route from classical mechanics to quantum mechanics, canonical quantization:

$$\boxed{\{ \ , \ }_{\text{Poisson}} \xrightarrow{\text{quantum mechanics}} \frac{1}{i\hbar} [\ , \].} \quad (2.2.22)$$

4) Ehrenfest theorem

How is $\hat{H}$ constructed? For a quantum system with a classical counterpart, the Hamiltonian $H(\mathbf{x}, \mathbf{p})$ of analytical mechanics is taken to the quantum Hamiltonian $\hat{H}(\hat{\mathbf{x}}, \hat{\mathbf{p}})$.

For a free particle,

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m}. \quad (2.2.23)$$

Heisenberg equations of motion (we henceforth omit the picture index H on observables) give

$$\frac{d\hat{\mathbf{p}}}{dt} = \frac{1}{i\hbar} [\hat{\mathbf{p}}, \hat{H}] = 0, \quad \hat{\mathbf{p}}(t) = \hat{\mathbf{p}}(0), \quad (2.2.24)$$

$$\frac{d\hat{\mathbf{x}}}{dt} = \frac{1}{i\hbar} [\hat{\mathbf{x}}, \hat{H}] = \frac{1}{2im} [\hat{\mathbf{x}}, \hat{\mathbf{p}}^2] = \frac{\hat{\mathbf{p}}(t)}{m} = \frac{\hat{\mathbf{p}}(0)}{m}, \quad \hat{\mathbf{x}}(t) = \hat{\mathbf{x}}(0) + \frac{\hat{\mathbf{p}}(0)}{m} t. \quad (2.2.25)$$

For a general system,

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}}). \quad (2.2.26)$$

The Heisenberg equations are

$$\frac{d\hat{\mathbf{p}}}{dt} = \frac{1}{i\hbar} [\hat{\mathbf{p}}, \hat{H}] = \frac{1}{i\hbar} [\hat{\mathbf{p}}, V(\hat{\mathbf{x}})]. \quad (2.2.27)$$

Expand $V(\hat{\mathbf{x}})$ as a power series in $\hat{\mathbf{x}}$ and use $[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}$. Then

$$\frac{d\hat{\mathbf{p}}}{dt} = -\nabla V(\hat{\mathbf{x}}), \quad (2.2.28)$$

$$\frac{d\hat{\mathbf{x}}}{dt} = \frac{1}{i\hbar} [\hat{\mathbf{x}}, \hat{H}] = \frac{1}{2im} [\hat{\mathbf{x}}, \hat{\mathbf{p}}^2] = \frac{\hat{\mathbf{p}}}{m}, \quad (2.2.29)$$

$$m \frac{d^2 \hat{\mathbf{x}}}{dt^2} = \frac{d\hat{\mathbf{p}}}{dt} = -\nabla V(\hat{\mathbf{x}}). \quad (2.2.30)$$

Taking expectation values on both sides,

$$\boxed{m \frac{d^2 \langle \hat{\mathbf{x}} \rangle}{dt^2} = \frac{d\langle \hat{\mathbf{p}} \rangle}{dt} = -\langle \nabla V(\hat{\mathbf{x}}) \rangle.} \quad (2.2.31)$$

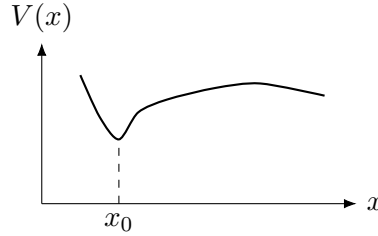
This is the Ehrenfest theorem, corresponding to Newton's law in classical mechanics. Since inner products are picture-independent, the Ehrenfest theorem for expectation values is also picture-independent.

Unless otherwise stated, problems below are generally discussed in the Schrödinger picture.

2.3 One-Dimensional Linear Harmonic Oscillator

For an arbitrary potential, expand it in a Taylor series near a minimum x_0 :

$$V(x) = V(x_0) + V'(x_0)(x - x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 + \cdots. \quad (2.3.1)$$



The constant term $V(x_0)$ can be incorporated into the energy. At the potential minimum $V'(x_0) = 0$. Omitting higher-order terms,

$$V(x) \simeq \frac{1}{2}V''(x_0)(x - x_0)^2. \quad (2.3.2)$$

This approximates a harmonic-oscillator potential; therefore the harmonic oscillator has general significance.

Classically,

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (2.3.3)$$

Quantum mechanically,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 \hat{x}^2, \quad [\hat{x}, \hat{p}] = i\hbar. \quad (2.3.4)$$

Because

$$\hat{H} = \frac{m\omega^2}{2} \left(\hat{x} - \frac{i\hat{p}}{m\omega} \right) \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right) + \frac{1}{2}\hbar\omega, \quad (2.3.5)$$

define new operators

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right), \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i\hat{p}}{m\omega} \right). \quad (2.3.6)$$

Then

$$\boxed{\hat{H} = \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hbar\omega, \quad [\hat{a}, \hat{a}^\dagger] = \mathbf{1}.} \quad (2.3.7)$$

Clearly, $\hat{a}$ is not Hermitian, $\hat{a}^\dagger \neq \hat{a}$. However, $\hat{a}^\dagger \hat{a}$ is Hermitian:

$$(\hat{a}^\dagger \hat{a})^\dagger = \hat{a}^\dagger \hat{a}. \quad (2.3.8)$$

Since $\hat{H}$ differs from $\hat{a}^\dagger \hat{a}$ only by a constant, they have common eigenstates. Let

$$\hat{a}^\dagger \hat{a} |n\rangle = n |n\rangle. \quad (2.3.9)$$

Then

$$\hat{H} |n\rangle = \left(n + \frac{1}{2}\right) \hbar\omega |n\rangle, \quad E_n = \left(n + \frac{1}{2}\right) \hbar\omega. \quad (2.3.10)$$

Questions: What are the eigenvalues n ? What are the coordinate-representation eigenstates $\langle x|n\rangle$?

1) Nonnegativity of the eigenvalues

Let

$$\hat{a} |n\rangle = |b\rangle. \quad (2.3.11)$$

Then

$$\langle n | \hat{a}^\dagger = \langle b |, \quad \langle n | \hat{a}^\dagger \hat{a} |n\rangle = \langle b | b \rangle, \quad n \langle n | n \rangle = \langle b | b \rangle. \quad (2.3.12)$$

Because $\langle b | b \rangle \geq 0$ and $\langle n | n \rangle \geq 0$, we have $n \geq 0$.

2) Raising and lowering operators

From $[\hat{a}, \hat{a}^\dagger] = 1$,

$$\begin{aligned} (\hat{a}^\dagger \hat{a}) \hat{a} |n\rangle &= \hat{a}(\hat{a} \hat{a}^\dagger - 1) |n\rangle = \hat{a}(\hat{a}^\dagger \hat{a} - 1) |n\rangle \\ &= (n - 1) \hat{a} |n\rangle, \end{aligned} \quad (2.3.13)$$

$$\begin{aligned} (\hat{a}^\dagger \hat{a}) \hat{a}^\dagger |n\rangle &= \hat{a}^\dagger(\hat{a} \hat{a}^\dagger) |n\rangle = \hat{a}^\dagger(\hat{a}^\dagger \hat{a} + 1) |n\rangle \\ &= (n + 1) \hat{a}^\dagger |n\rangle. \end{aligned} \quad (2.3.14)$$

If $|n\rangle$ is an eigenstate of $\hat{a}^\dagger \hat{a}$, so are $\hat{a} |n\rangle$ and $\hat{a}^\dagger |n\rangle$, with the following relationships:

Eigenstate	Eigenvalue of $\hat{a}^\dagger \hat{a}$
$(\hat{a}^\dagger)^2 n\rangle$	$n + 2$
$\hat{a}^\dagger n\rangle$	$n + 1$
$ n\rangle$	n
$\hat{a} n\rangle$	$n - 1$
$\hat{a}^2 n\rangle$	$n - 2$

(2.3.15)

Hence $\hat{a}$ is called the lowering (annihilation) operator and $\hat{a}^\dagger$ the raising (creation) operator. Together with $n \geq 0$, the possible eigenvalues have the form

$$n = n_0, n_0 + 1, n_0 + 2, \dots, \quad n_0 \geq 0. \quad (2.3.16)$$

3) Lowest eigenvalue

For the lowest eigenvalue n_0 , necessarily

$$\hat{a} |n_0\rangle = 0. \quad (2.3.17)$$

Otherwise $\hat{a} |n_0\rangle$ would have eigenvalue $n_0 - 1$, contradicting the assumption that n_0 is the smallest eigenvalue. From

$$\hat{a}^\dagger \hat{a} |n_0\rangle = n_0 |n_0\rangle = 0 \quad (2.3.18)$$

we obtain $n_0 = 0$. Therefore

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega, \quad n = 0, 1, 2, \dots$$

(2.3.19)

Notice that so far we have used only the commutation relation, without introducing any specific representation.

4) Matrix elements of creation and annihilation operators

Since $\hat{a}|n\rangle$ has eigenvalue $n-1$, and in the absence of degeneracy (one-dimensional bound states are nondegenerate),

$$\hat{a}|n\rangle = a_n|n-1\rangle. \quad (2.3.20)$$

Similarly,

$$\hat{a}^\dagger|n\rangle = b_n|n+1\rangle. \quad (2.3.21)$$

Because

$$\langle n|\hat{a}^\dagger = a_n^*\langle n-1|, \quad \langle n|\hat{a} = b_n^*\langle n+1|, \quad (2.3.22)$$

we obtain

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle = |a_n|^2\langle n-1|n-1\rangle, \quad n = |a_n|^2, \quad |a_n| = \sqrt{n}, \quad (2.3.23)$$

$$\langle n|\hat{a}\hat{a}^\dagger|n\rangle = |b_n|^2\langle n+1|n+1\rangle, \quad n+1 = |b_n|^2, \quad |b_n| = \sqrt{n+1}. \quad (2.3.24)$$

Choose the phases so that $a_n = \sqrt{n}$ and $b_n = \sqrt{n+1}$. Thus

$$\boxed{\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.} \quad (2.3.25)$$

5) Representation common to $\hat{H}$ and $\hat{a}^\dagger\hat{a}$

$$(a^\dagger a)_{mn} = \langle m|\hat{a}^\dagger\hat{a}|n\rangle = n\langle m|n\rangle = n\delta_{mn}, \quad (2.3.26)$$

$$H_{mn} = \langle m|\hat{H}|n\rangle = \left(n + \frac{1}{2}\right)\hbar\omega\delta_{mn}. \quad (2.3.27)$$

Both are diagonal matrices, as expected in their own representation. Meanwhile,

$$a_{mn} = \langle m|\hat{a}|n\rangle = \sqrt{n}\langle m|n-1\rangle = \sqrt{n}\delta_{m,n-1}, \quad (2.3.28)$$

$$(a^\dagger)_{mn} = \langle m|\hat{a}^\dagger|n\rangle = \sqrt{n+1}\langle m|n+1\rangle = \sqrt{n+1}\delta_{m,n+1}. \quad (2.3.29)$$

Since

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^\dagger + \hat{a}), \quad \hat{p} = i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^\dagger - \hat{a}), \quad (2.3.30)$$

it follows that

$$x_{mn} = \sqrt{\frac{\hbar}{2m\omega}}(\sqrt{n}\delta_{m,n-1} + \sqrt{n+1}\delta_{m,n+1}), \quad (2.3.31)$$

$$p_{mn} = i\sqrt{\frac{m\hbar\omega}{2}}(-\sqrt{n}\delta_{m,n-1} + \sqrt{n+1}\delta_{m,n+1}). \quad (2.3.32)$$

They are nondiagonal matrices. Note that $\langle n|\hat{x}|n\rangle = \langle n|\hat{p}|n\rangle = 0$.

6) Coordinate representation

For the ground state $|0\rangle$,

$$\hat{a}|0\rangle = 0, \quad \left(\hat{x} + \frac{i}{m\omega}\hat{p}\right)|0\rangle = 0. \quad (2.3.33)$$

Multiplying from the left by $\langle x|$ and inserting the completeness relation,

$$\int dx' \langle x|\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right)|x'\rangle\langle x'|0\rangle = 0. \quad (2.3.34)$$

Using

$$\langle x|\hat{x}|x'\rangle = x\delta(x-x'), \quad \langle x|\hat{p}|x'\rangle = -i\hbar \frac{d}{dx}\delta(x-x'), \quad \langle x'|0\rangle = \psi_0(x'), \quad (2.3.35)$$

we obtain

$$\left(x + \frac{\hbar}{m\omega} \frac{d}{dx}\right) \psi_0(x) = 0. \quad (2.3.36)$$

The solution of this first-order differential equation is

$$\psi_0(x) = C \exp\left(-\frac{m\omega x^2}{2\hbar}\right). \quad (2.3.37)$$

With the normalization condition

$$\int_{-\infty}^{\infty} |\psi_0(x)|^2 dx = 1, \quad (2.3.38)$$

we have

$$\boxed{\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right).} \quad (2.3.39)$$

For an excited state,

$$\begin{aligned} \psi_n(x) &= \langle x|n\rangle = \frac{1}{\sqrt{n}} \langle x|\hat{a}^\dagger|n-1\rangle = \frac{1}{\sqrt{n(n-1)}} \langle x|(\hat{a}^\dagger)^2|n-2\rangle \\ &= \frac{1}{\sqrt{n!}} \langle x|(\hat{a}^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{2\hbar}\right)^{n/2} \langle x|\left(\hat{x} - \frac{i}{m\omega}\hat{p}\right)^n|0\rangle. \end{aligned} \quad (2.3.40)$$

Insert n position completeness relations:

$$\begin{aligned} \psi_n(x) &= \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{2\hbar}\right)^{n/2} \int dx_1 \cdots dx_n \left\langle x\left|\hat{x} - \frac{i}{m\omega}\hat{p}\right|x_1\right\rangle \left\langle x_1\left|\hat{x} - \frac{i}{m\omega}\hat{p}\right|x_2\right\rangle \\ &\quad \cdots \left\langle x_{n-1}\left|\hat{x} - \frac{i}{m\omega}\hat{p}\right|x_n\right\rangle \langle x_n|0\rangle. \end{aligned} \quad (2.3.41)$$

Using the position and momentum matrix elements gives

$$\boxed{\psi_n(x) = \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{2\hbar}\right)^{n/2} \left(x - \frac{\hbar}{m\omega} \frac{d}{dx}\right)^n \psi_0(x).} \quad (2.3.42)$$

Thus the one-dimensional harmonic-oscillator problem has been solved completely.

7) Physical meaning of $\hat{a}^\dagger\hat{a}$

The harmonic-oscillator energies are

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega, \quad n = 0, 1, 2, 3, \dots \quad (2.3.43)$$

The ground-state energy $E_0 = \hbar\omega/2 > 0$ is the zero-point energy. An excited state involves quanta of energy $\varepsilon = \hbar\omega$. When the oscillator is in the n -th excited state, n quanta have been excited.

n : quantum number, particle number.

$\hat{N} = \hat{a}^\dagger\hat{a}$ is the particle-number operator; $\hat{a}$ is a particle-annihilation operator and $\hat{a}^\dagger$ a particle-creation operator.

The common representation of $\hat{H}$ and $\hat{N}$ is the particle-number representation.

To summarize the algebraic solution: bound states $\rightarrow$ discrete spectrum $\rightarrow$ identify lowering and raising operators for the discrete quantum numbers. The particle-number representation is the foundation of the second-quantization scheme.

8) Casimir effect

Does the zero-point energy have observable effects? Consider the vacuum energy of the quantized electromagnetic field.

According to Planck's studies of blackbody radiation, the electromagnetic field consists of harmonic oscillations of different frequencies ω_k (energy $\hbar\omega_k$). Here k is the wave number, $\omega_k = ck$, and n is the number of photons of energy $\hbar\omega_k$. The total electromagnetic-field energy is written

$$E = 2 \int \frac{d^3k}{(2\pi)^3} \sum_{n=0}^{\infty} \left(n + \frac{1}{2} \right) \hbar\omega_k. \quad (2.3.44)$$

Boundary conditions change the distribution of frequencies, hence the distribution of wave numbers k .

The vacuum is the unexcited state, $n = 0$, with vacuum energy

$$E = \int \frac{d^3k}{(2\pi)^3} \hbar\omega_k, \quad (2.3.45)$$

which is infinite. In ordinary calculations it is subtracted as the background to excitations. The question is whether the vacuum energy itself has observable effects.

Place two parallel plates in the electromagnetic radiation field, each of area L^2 with L sufficiently large and separation $a \ll L$. Boundary conditions restrict the component of wave number perpendicular to the plates to $k_z = n\pi/a$; in the transverse direction there is no significant restriction because L is large. The vacuum energy between the plates, in volume L^2a , is

$$E(a) = \hbar c \int \frac{d^2k_{\perp}}{(2\pi/L)^2} \sum_{n=0}^{\infty} \sqrt{k_{\perp}^2 + \left(\frac{n\pi}{a} \right)^2}. \quad (2.3.46)$$

Here $d^2k_{\perp}/(2\pi/L)^2$ is the number of transverse modes in the element $d^2k_{\perp}$ near $\mathbf{k}_{\perp}$.

The change in vacuum energy associated with inserting the plates into this volume is

$$\begin{aligned} \Delta E(a) = E(a) - E_0(a) = \hbar c \int \frac{d^2k_{\perp}}{(2\pi/L)^2} & \left[\sum_{n=0}^{\infty} \sqrt{k_{\perp}^2 + \left(\frac{n\pi}{a} \right)^2} \right. \\ & \left. - \int_0^{\infty} dn \sqrt{k_{\perp}^2 + \left(\frac{n\pi}{a} \right)^2} \right]. \end{aligned} \quad (2.3.47)$$

This change produces a force per unit area between the plates,

$$F(a) = -\frac{1}{L^2} \frac{\partial \Delta E(a)}{\partial a}. \quad (2.3.48)$$

Both $E(a)$ and $E_0(a)$ diverge (vacuum divergence), but $\Delta E(a)$ is finite. After calculation (see standard quantum-mechanics textbooks, for example those by Su Ru-Keng and Zhang Yong-De),

$$\boxed{F(a) = -\frac{\hbar c \pi^2}{240a^4}}. \quad (2.3.49)$$

The attractive force caused by the change (fluctuation) in vacuum energy when the plates are inserted, and its dependence on a , were experimentally observed in 1958. This is an observable effect of the vacuum energy.

9) Time evolution of observables in the Heisenberg picture

Omitting the H subscript, the Heisenberg equations of motion give

$$\frac{d\hat{a}}{dt} = \frac{1}{i\hbar}[\hat{a}, \hat{H}] = -i\omega\hat{a}, \quad \frac{d\hat{a}^\dagger}{dt} = \frac{1}{i\hbar}[\hat{a}^\dagger, \hat{H}] = i\omega\hat{a}^\dagger. \quad (2.3.50)$$

The solutions are

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t}, \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}. \quad (2.3.51)$$

Alternatively,

$$\begin{aligned} \hat{x}(t) &= \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}(t) + \hat{a}^\dagger(t)) = \hat{x}(0) \cos \omega t + \frac{\hat{p}(0)}{m\omega} \sin \omega t, \\ \hat{p}(t) &= i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^\dagger(t) - \hat{a}(t)) = \hat{p}(0) \cos \omega t - m\omega\hat{x}(0) \sin \omega t. \end{aligned} \quad (2.3.52)$$

Although $\hat{x}(t)$ and $\hat{p}(t)$ oscillate, their expectation values in an energy eigenstate are

$$\langle n|\hat{x}(t)|n\rangle = \langle n|\hat{x}(0)|n\rangle \cos \omega t + \frac{1}{m\omega} \langle n|\hat{p}(0)|n\rangle \sin \omega t = 0, \quad (2.3.53)$$

$$\langle n|\hat{p}(t)|n\rangle = \langle n|\hat{p}(0)|n\rangle \cos \omega t - m\omega \langle n|\hat{x}(0)|n\rangle \sin \omega t = 0 \quad (2.3.54)$$

and thus show no oscillation.

2.4 Coherent States of the Harmonic Oscillator

In energy eigenstates of the harmonic oscillator, the expectation values of $\hat{x}$ and $\hat{p}$ are zero, and no classical oscillation is observed. However, consider a linear combination of energy eigenstates, for example

$$|a\rangle = c_0|0\rangle + c_1|1\rangle. \quad (2.4.1)$$

In the Heisenberg picture, the expectation value is

$$\begin{aligned} \langle a|\hat{x}(t)|a\rangle &= (c_0^*\langle 0| + c_1^*\langle 1|)\hat{x}(t)(c_0|0\rangle + c_1|1\rangle) \\ &= c_0^*c_1\langle 0|\hat{x}(t)|1\rangle + c_1^*c_0\langle 1|\hat{x}(t)|0\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}(c_0^*c_1e^{-i\omega t} + c_1^*c_0e^{i\omega t}). \end{aligned} \quad (2.4.2)$$

Oscillations appear. Can one find a state whose oscillation is closest to classical oscillation, that is, the oscillation of $\hat{x}(t)$? This leads to the coherent state $|\alpha\rangle$, defined as an eigenstate of the lowering operator:

$$\boxed{\hat{a}|\alpha\rangle = \alpha|\alpha\rangle}. \quad (2.4.3)$$

Note that $\hat{a}$ is not Hermitian; its eigenvalue α is generally complex.

1) Coherent states satisfy the minimum-uncertainty relation

The uncertainty relation is

$$(\Delta x)^2(\Delta p)^2 \geq \frac{\hbar^2}{4}. \quad (2.4.4)$$

In the coherent state $|\alpha\rangle$,

$$\begin{aligned}\langle\alpha|\hat{x}|\alpha\rangle &= \sqrt{\frac{\hbar}{2m\omega}} \langle\alpha|\hat{a}^\dagger + \hat{a}|\alpha\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\alpha^* + \alpha), \\ \langle\alpha|\hat{x}^2|\alpha\rangle &= \frac{\hbar}{2m\omega} [1 + (\alpha + \alpha^*)^2], \\ \langle\alpha|\hat{p}|\alpha\rangle &= -i\sqrt{\frac{\hbar m\omega}{2}}(\alpha - \alpha^*), \\ \langle\alpha|\hat{p}^2|\alpha\rangle &= \frac{\hbar m\omega}{2} [1 - (\alpha - \alpha^*)^2].\end{aligned}\tag{2.4.5}$$

It follows that

$$(\Delta x)^2(\Delta p)^2 = (\langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2)(\langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2) = \frac{\hbar^2}{4}.\tag{2.4.6}$$

Thus coherent states satisfy the minimum-uncertainty relation and are the quantum states closest to classical states.

2) Expansion of a coherent state

Expand $|\alpha\rangle$ in energy eigenstates:

$$|\alpha\rangle = \sum_n c_n |n\rangle.\tag{2.4.7}$$

Since

$$c_n = \langle n|\alpha\rangle = \frac{1}{\sqrt{n!}} \langle 0|\hat{a}^n|\alpha\rangle = \frac{\alpha^n}{\sqrt{n!}} \langle 0|\alpha\rangle = \frac{\alpha^n}{\sqrt{n!}} c_0,\tag{2.4.8}$$

we have

$$|\alpha\rangle = c_0 \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle.\tag{2.4.9}$$

Normalization gives

$$1 = \langle\alpha|\alpha\rangle = \sum_n |c_n|^2 = |c_0|^2 \sum_n \frac{|\alpha|^{2n}}{n!} = |c_0|^2 e^{|\alpha|^2}.\tag{2.4.10}$$

Choose $c_0 = e^{-|\alpha|^2/2}$; then

$$|\alpha\rangle = \sum_{n=0}^{\infty} e^{-|\alpha|^2/2} \frac{\alpha^n}{\sqrt{n!}} |n\rangle.\tag{2.4.11}$$

3) The ground state is a coherent state

For the ground state $|0\rangle$, $\hat{a}|0\rangle = 0$. Thus $|0\rangle$ is an eigenstate of the annihilation operator, with eigenvalue zero.

4) Time evolution of the eigenvalue

In the Heisenberg picture,

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t},\tag{2.4.12}$$

$$\hat{a}(t)|\alpha\rangle = \hat{a}(0)e^{-i\omega t}|\alpha\rangle = \alpha e^{-i\omega t}|\alpha\rangle.\tag{2.4.13}$$

Hence the eigenvalue evolves as

$$\alpha(t) = \alpha e^{-i\omega t}.\tag{2.4.14}$$

Interactive investigation IC-02-03. Open Coherent-state dynamics. After obtaining $\alpha(t)$, prepare an oscillator coherent state and vary its amplitude and phase. Compare the Fock distribution and the moving position density; revisit the experiment after studying squeezed states to see spreading.

2.5 Squeezed States

A coherent state has the minimum uncertainty product and exhibits equal fluctuations of its two *dimensionless* oscillator quadratures. Minimum uncertainty does not require these two fluctuations to be equal. A squeezed state redistributes the fluctuations between conjugate quadratures while maintaining the uncertainty bound.

Define the Hermitian quadratures

$$\hat{X} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \hat{P} = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}}, \quad [\hat{X}, \hat{P}] = i1. \quad (2.5.1)$$

For real r introduce the *single-mode squeezing operator*

$$\hat{S}(r) = \exp\left[\frac{r}{2}(\hat{a}^2 - \hat{a}^{\dagger 2})\right], \quad |0, r\rangle = \hat{S}(r)|0\rangle. \quad (2.5.2)$$

Its exponent is anti-Hermitian; $\hat{S}(r)$ is unitary. The transformation of the lowering operator is the Bogoliubov relation

$$\hat{S}^\dagger(r)\hat{a}\hat{S}(r) = \hat{a} \cosh r - \hat{a}^\dagger \sinh r. \quad (2.5.3)$$

To derive it, set $\hat{a}(r) = \hat{S}^\dagger(r)\hat{a}\hat{S}(r)$. Differentiate with respect to r and use the commutators $[\hat{a}, \hat{a}^{\dagger 2}] = 2\hat{a}^\dagger$ and $[\hat{a}, \hat{a}^2] = 0$. Together with the adjoint equation this gives

$$\frac{d\hat{a}(r)}{dr} = -\hat{a}^\dagger(r), \quad \frac{d\hat{a}^\dagger(r)}{dr} = -\hat{a}(r), \quad (2.5.4)$$

whose solution with $\hat{a}(0) = \hat{a}$ is the displayed transformation.

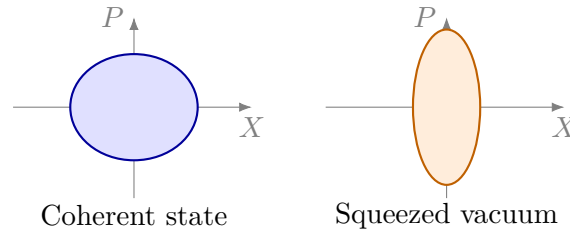
Consequently,

$$\hat{S}^\dagger(r)\hat{X}\hat{S}(r) = e^{-r}\hat{X}, \quad \hat{S}^\dagger(r)\hat{P}\hat{S}(r) = e^r\hat{P}. \quad (2.5.5)$$

The vacuum has $\langle\hat{X}\rangle = \langle\hat{P}\rangle = 0$ and $\langle\hat{X}^2\rangle = \langle\hat{P}^2\rangle = 1/2$. Evaluating the transformed variances in $|0, r\rangle$ gives

$$(\Delta X)^2 = \frac{1}{2}e^{-2r}, \quad (\Delta P)^2 = \frac{1}{2}e^{2r}, \quad \Delta X \Delta P = \frac{1}{2}. \quad (2.5.6)$$

For $r > 0$, the X quadrature is squeezed and the P quadrature is anti-squeezed. Both coherent and squeezed vacuum states saturate the Heisenberg uncertainty relation; they have different distributions of the fluctuations.



The ellipses schematically indicate quantum uncertainties in phase space. The squeezed state retains the minimum uncertainty *area*, but its shape changes. A displaced squeezed state is $\hat{D}(\alpha)\hat{S}(r)|0\rangle$, where $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$; the displacement moves the center without altering these variances.

2.6 Schrödinger Equation in the Coordinate Representation and the WKB Approximation

Consider the abstract state $|\psi, t\rangle$, whose Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = \hat{H} |\psi, t\rangle. \quad (2.6.1)$$

Multiply from the left by the eigenbra $\langle \mathbf{x}|$ of the position operator $\hat{\mathbf{x}}$. Since these eigenstates are time independent in the Schrödinger picture,

$$i\hbar \frac{\partial}{\partial t} \langle \mathbf{x}|\psi, t\rangle = \langle \mathbf{x}|\hat{H}|\psi, t\rangle = \int d^3x' \langle \mathbf{x}|\hat{H}|\mathbf{x}'\rangle \langle \mathbf{x}'|\psi, t\rangle. \quad (2.6.2)$$

Let

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}}). \quad (2.6.3)$$

Since

$$\langle \mathbf{x}|\hat{H}|\mathbf{x}'\rangle = \delta^{(3)}(\mathbf{x} - \mathbf{x}') \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right], \quad (2.6.4)$$

we obtain

$$i\hbar \frac{\partial}{\partial t} \langle \mathbf{x}|\psi, t\rangle = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right] \langle \mathbf{x}|\psi, t\rangle. \quad (2.6.5)$$

Define the wavefunction $\psi(\mathbf{x}, t) = \langle \mathbf{x}|\psi, t\rangle$. This is the Schrödinger equation in the coordinate representation.

1) Time-independent Schrödinger equation

Put

$$\psi(\mathbf{x}, t) = \varphi(\mathbf{x})T(t). \quad (2.6.6)$$

If $\hat{H}$ has no time dependence, inserting this into the Schrödinger equation separates it into two equations. The time-dependent factor is

$$T(t) = e^{-iEt/\hbar}. \quad (2.6.7)$$

The energy E and $\varphi(\mathbf{x})$ are determined by the other equation, the energy eigenvalue equation:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right] \varphi_E(\mathbf{x}) = E\varphi_E(\mathbf{x}). \quad (2.6.8)$$

A stationary state is

$$\psi_E(\mathbf{x}, t) = \varphi_E(\mathbf{x})e^{-iEt/\hbar}. \quad (2.6.9)$$

For bound states the bound-state condition

$$\lim_{|\mathbf{x}| \rightarrow \infty} \varphi_E(\mathbf{x}) = 0 \quad (2.6.10)$$

leads to discrete values of E .

Stationary states are particular solutions when $\hat{H}$ is time independent. The general solution is a linear superposition of stationary states:

$$\psi(\mathbf{x}, t) = \sum_E c_E \psi_E(\mathbf{x}, t). \quad (2.6.11)$$

2) Probability interpretation of the wavefunction

The probability density of finding the particle at a given point is

$$\rho(\mathbf{x}, t) = |\langle \mathbf{x} | \psi(t) \rangle|^2 = |\psi(\mathbf{x}, t)|^2 = \psi^*(\mathbf{x}, t)\psi(\mathbf{x}, t). \quad (2.6.12)$$

When $V(\mathbf{x})$ is real, the Schrödinger equation can be written as the continuity equation,

$$\boxed{\frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{x}, t) = 0}, \quad (2.6.13)$$

$$\mathbf{j}(\mathbf{x}, t) = -\frac{i\hbar}{2m} [\psi^*(\mathbf{x}, t)\nabla\psi(\mathbf{x}, t) - \psi(\mathbf{x}, t)\nabla\psi^*(\mathbf{x}, t)]. \quad (2.6.14)$$

What is the physical meaning of $\mathbf{j}$? Integrate the continuity equation over a finite volume V :

$$\int_V d^3r \frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \int_V d^3r \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0, \quad (2.6.15)$$

$$\frac{d}{dt} \int_V d^3r \rho(\mathbf{r}, t) = - \oint_{\partial V} d\mathbf{S} \cdot \mathbf{j}(\mathbf{r}, t). \quad (2.6.16)$$

Local probability conservation says that the change in the probability inside V equals the negative of the outward probability flux across its boundary S . Thus $\mathbf{j}$ is called the probability-current density.

If we integrate over all space,

$$\frac{d}{dt} \int_{\mathbb{R}^3} d^3r \rho(\mathbf{r}, t) = - \oint_{\infty} d\mathbf{S} \cdot \mathbf{j}(\mathbf{r}, t) = 0. \quad (2.6.17)$$

Therefore $\int d^3r \rho(\mathbf{r}, t)$ is independent of time and is a constant. This implies:

- a) Total probability is conserved; particles are neither created nor annihilated. This describes nonrelativistic quantum mechanics.
- b) The normalization of a wavefunction is independent of time. Schrödinger evolution preserves normalization.
- c) If V is complex, probability is not conserved, and this can effectively describe creation and annihilation of particles.

3) How does a quantum system approach a classical system?

An operator in a general state has no definite value, whereas its expectation value does. Thus expectation values of quantum observables correspond to physical quantities in classical mechanics.

The classical equation of motion is

$$\frac{dO}{dt} = \{O, H\}_{\text{Poisson bracket}}, \quad (2.6.18)$$

while the quantum equation of motion for the expectation value is

$$\frac{d\langle O \rangle}{dt} = \frac{\langle [\hat{O}, \hat{H}] \rangle}{i\hbar}. \quad (2.6.19)$$

In the classical limit, observables commute, and $[\hat{O}, \hat{H}] \rightarrow 0$. To recover the classical equation of motion, one requires

$$\hbar \longrightarrow 0, \quad \frac{\langle [\hat{O}, \hat{H}] \rangle}{i\hbar} \longrightarrow \{O, H\}_{\text{Poisson bracket}}. \quad (2.6.20)$$

4) WKB approximation to the time-independent equation

Only a few Schrödinger equations can be solved exactly. Expanding the Schrödinger equation in powers of $\hbar$ and solving it by successive approximations gives the WKB approximation (Wentzel, Kramers and Brillouin).

Write a stationary-state wavefunction as

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/\hbar} = e^{i\mathbf{s}(\mathbf{r})/\hbar}e^{-iEt/\hbar} = e^{i[\mathbf{s}(\mathbf{r})-Et]/\hbar}. \quad (2.6.21)$$

Insert this into the stationary Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (2.6.22)$$

We find

$$\frac{(\nabla s)^2}{2m} + V(\mathbf{r}) - E - \frac{i\hbar}{2m} \nabla^2 s = 0. \quad (2.6.23)$$

Expand $s(\mathbf{r})$ in $\hbar/i$:

$$s(\mathbf{r}) = s_0(\mathbf{r}) + \frac{\hbar}{i} s_1(\mathbf{r}) + \left(\frac{\hbar}{i} \right)^2 s_2(\mathbf{r}) + \dots. \quad (2.6.24)$$

Equate like powers of $\hbar$:

$$\begin{aligned} \hbar^0 : \quad & (\nabla s_0)^2 + 2m[V(\mathbf{r}) - E] = 0, \\ \hbar^1 : \quad & \nabla s_1 \cdot \nabla s_0 + \frac{1}{2} \nabla^2 s_0 = 0, \\ & \vdots \end{aligned} \quad (2.6.25)$$

For one dimension,

$$s_0(x) = \pm \int^x p(x') dx' + C, \quad s_1(x) = -\ln \sqrt{p(x)} + \tilde{C}, \quad (2.6.26)$$

$$\boxed{p(x) = \sqrt{2m[E - V(x)]}}. \quad (2.6.27)$$

Keeping terms through first order in $\hbar$,

$$\begin{aligned} \psi_{\text{II}}(x) &= \exp \left[\frac{i}{\hbar} \left(s_0(x) + \frac{\hbar}{i} s_1(x) \right) \right] \\ &= \frac{C_1}{\sqrt{p(x)}} \exp \left(\frac{i}{\hbar} \int^x p(x') dx' \right) + \frac{C_2}{\sqrt{p(x)}} \exp \left(-\frac{i}{\hbar} \int^x p(x') dx' \right). \end{aligned} \quad (2.6.28)$$

This solution is valid where $p(x) = \sqrt{2m[E - V(x)]} > 0$, that is, $E > V$. In classical mechanics, $E = T + V$; the condition $E > V$ therefore identifies a classically allowed region.

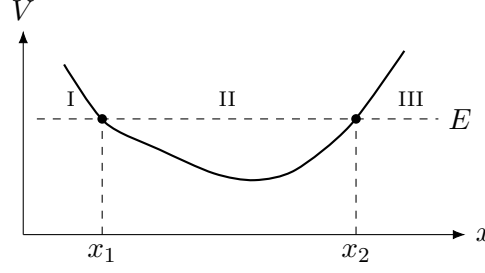
In the classically forbidden region, $E < V$, put

$$p(x) = i\tilde{p}(x) = i\sqrt{2m[V(x) - E]}. \quad (2.6.29)$$

Then

$$\psi_{\text{I,III}}(x) = \frac{C_3}{\sqrt{|p(x)|}} \exp \left(\frac{1}{\hbar} \int^x \tilde{p}(x') dx' \right) + \frac{C_4}{\sqrt{|p(x)|}} \exp \left(-\frac{1}{\hbar} \int^x \tilde{p}(x') dx' \right). \quad (2.6.30)$$

At $E = V$, the momentum $p = 0$ and the solution diverges: WKB is no longer valid. A point x_0 satisfying $E - V(x_0) = 0$ is a classical turning point. How can one solve the equation near x_0 ?



Expand the potential to linear order near x_0 :

$$V(x) \simeq V(x_0) + \left. \frac{dV}{dx} \right|_{x_0} (x - x_0). \quad (2.6.31)$$

Substitution into the stationary Schrödinger equation gives

$$\frac{d^2\varphi}{dx^2} - \frac{2m}{\hbar^2} \left. \frac{dV}{dx} \right|_{x_0} (x - x_0)\varphi = 0. \quad (2.6.32)$$

The exact solutions are Bessel functions of order $\pm \frac{1}{3}$. Consider their asymptotic solutions when $|x - x_0|$ is large and require continuity with the classically allowed and forbidden solutions on either side of x_0 . In this way the constants C_1, C_2, C_3, C_4 can be fixed.

Interactive investigation IC-02-05. Open WKB near a turning point. After the WKB derivation, compare the approximate wavefunction with the exact reference near a turning point. Locate the region where the semiclassical approximation fails and inspect how the error changes with the parameters.

2.7 Feynman Path Integral

1) Propagator in the Schrödinger picture

Start with the time evolution of the state:

$$|\psi, t\rangle = e^{-i\hat{H}(t-t_0)/\hbar} |\psi, t_0\rangle. \quad (2.7.1)$$

In the coordinate representation,

$$\begin{aligned} \langle \mathbf{x}'' | \psi, t \rangle &= \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} |\psi, t_0\rangle \\ &= \int d^3x' \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \langle \mathbf{x}' | \psi, t_0 \rangle. \end{aligned} \quad (2.7.2)$$

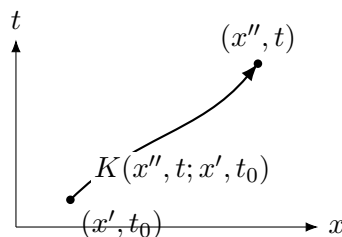
Define the matrix element of the evolution operator,

$$K(\mathbf{x}'', t; \mathbf{x}', t_0) = \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle. \quad (2.7.3)$$

Then

$$\psi(\mathbf{x}'', t) = \int d^3x' K(\mathbf{x}'', t; \mathbf{x}', t_0) \psi(\mathbf{x}', t_0). \quad (2.7.4)$$

The function $K(\mathbf{x}'', t; \mathbf{x}', t_0)$ is the propagator: it propagates the initial wavefunction $\psi(\mathbf{x}', t_0)$ to the final wavefunction $\psi(\mathbf{x}'', t)$. Integration means that every possible position of the particle in the initial wavefunction contributes to its final state.



We now derive the propagator's relationship to the equation of motion. Take $t \geq t_0$ and define a related propagator incorporating the initial condition:

$$K(\mathbf{x}'', t; \mathbf{x}', t_0) = \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \theta(t - t_0). \quad (2.7.5)$$

Differentiate with respect to t :

$$\begin{aligned} i\hbar \frac{\partial K}{\partial t} &= \langle \mathbf{x}'' | \hat{H} e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \theta(t - t_0) \\ &\quad + i\hbar \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \delta(t - t_0) \\ &= \int d^3x''' \langle \mathbf{x}'' | \hat{H} | \mathbf{x}''' \rangle K(\mathbf{x}''', t; \mathbf{x}', t_0) \\ &\quad + i\hbar \delta^{(3)}(\mathbf{x}'' - \mathbf{x}') \delta(t - t_0) \\ &= \hat{H}(\mathbf{x}'') K(\mathbf{x}'', t; \mathbf{x}', t_0) + i\hbar \delta^{(3)}(\mathbf{x}'' - \mathbf{x}') \delta(t - t_0). \end{aligned} \quad (2.7.6)$$

Thus

$$\left(i\hbar \frac{\partial}{\partial t} - \hat{H}(\mathbf{x}'') \right) K(\mathbf{x}'', t; \mathbf{x}', t_0) = i\hbar \delta^{(3)}(\mathbf{x}'' - \mathbf{x}') \delta(t - t_0). \quad (2.7.7)$$

Notice that taking the spatial trace of the propagator yields

$$\begin{aligned} \int d^3x K(\mathbf{x}, t; \mathbf{x}, 0) &= \int d^3x \langle \mathbf{x} | e^{-i\hat{H}t/\hbar} | \mathbf{x} \rangle \\ &= \sum_E \int d^3x \langle \mathbf{x} | E \rangle \langle E | \mathbf{x} \rangle e^{-iEt/\hbar} \\ &= \sum_E e^{-iEt/\hbar}. \end{aligned} \quad (2.7.8)$$

With the imaginary-time substitution $t = -i\hbar\beta$, this becomes the statistical-mechanical partition function $Z = \sum_E e^{-\beta E}$, showing a relation between imaginary-time propagators and statistical mechanics.

Example 1: propagator of a one-dimensional free particle.

$$\begin{aligned} K(x'', t; x', t_0) &= \langle x'' | e^{-i\hat{H}(t-t_0)/\hbar} | x' \rangle \\ &= \int dp \langle x'' | p \rangle e^{-ip^2(t-t_0)/(2m\hbar)} \langle p | x' \rangle \\ &= \int \frac{dp}{2\pi\hbar} \exp \left[\frac{i}{\hbar} \left(p(x'' - x') - \frac{p^2(t-t_0)}{2m} \right) \right] \\ &= \sqrt{\frac{m}{2\pi i\hbar(t-t_0)}} \exp \left[\frac{im(x'' - x')^2}{2\hbar(t-t_0)} \right]. \end{aligned} \quad (2.7.9)$$

Once this propagator is known, the time evolution of an initial plane wave (stationary state), $\psi(x', t_0) = (2\pi)^{-1} \exp [i(px' - p^2 t_0/(2m))/\hbar]$, is

$$\psi(x'', t) = \int dx' K(x'', t; x', t_0) \psi(x', t_0), \quad (2.7.10)$$

which is again a plane wave.

Example 2: propagator of a one-dimensional harmonic oscillator.

$$\begin{aligned} K(x'', t; x', t_0) &= \langle x'' | e^{-i\hat{H}(t-t_0)/\hbar} | x' \rangle \\ &= \sum_n \langle x'' | n \rangle e^{-iE_n(t-t_0)/\hbar} \langle n | x' \rangle \\ &= \sum_n \psi_n(x'') \psi_n^*(x') e^{-iE_n(t-t_0)/\hbar}. \end{aligned} \quad (2.7.11)$$

Here E_n and $\psi_n(x)$ are the oscillator's energies and its coordinate-representation wavefunctions.

2) Transition amplitude in the Heisenberg picture

In the Heisenberg picture,

$$\begin{aligned}
 K(\mathbf{x}'', t; \mathbf{x}', t_0) &= \langle \mathbf{x}'' | e^{-i\hat{H}(t-t_0)/\hbar} | \mathbf{x}' \rangle \\
 &= \left(\langle \mathbf{x}'' | e^{-i\hat{H}t/\hbar} \right) \left(e^{i\hat{H}t_0/\hbar} | \mathbf{x}' \rangle \right) \\
 &= \left(\langle \mathbf{x}'' | \hat{U}(t, 0) \right) \left(\hat{U}^\dagger(t_0, 0) | \mathbf{x}' \rangle \right) = \langle \mathbf{x}'', t | \mathbf{x}', t_0 \rangle.
 \end{aligned} \tag{2.7.12}$$

This is the amplitude for the position eigenstate $|\mathbf{x}', t_0\rangle$ at time t_0 to make a transition to the position eigenstate $|\mathbf{x}'', t\rangle$ at time t . Thus the propagator is also called a transition amplitude.

The completeness relation in the Heisenberg picture must involve basis states at the same time:

$$\int d^3x |\mathbf{x}, t\rangle \langle \mathbf{x}, t| = \mathbf{1}. \tag{2.7.13}$$

It follows that

$$\begin{aligned}
 K(\mathbf{x}_3, t_3; \mathbf{x}_1, t_1) &= \langle \mathbf{x}_3, t_3 | \mathbf{x}_1, t_1 \rangle \\
 &= \int d^3x_2 \langle \mathbf{x}_3, t_3 | \mathbf{x}_2, t_2 \rangle \langle \mathbf{x}_2, t_2 | \mathbf{x}_1, t_1 \rangle \\
 &= \int d^3x_2 K(\mathbf{x}_3, t_3; \mathbf{x}_2, t_2) K(\mathbf{x}_2, t_2; \mathbf{x}_1, t_1), \quad t_3 > t_2 > t_1.
 \end{aligned} \tag{2.7.14}$$

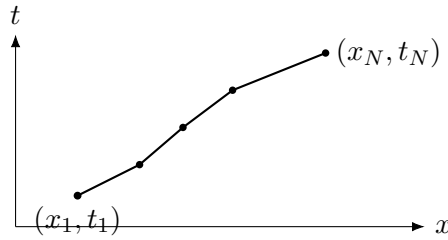
This is the composition law of propagators: propagation from $(\mathbf{x}_1, t_1)$ to $(\mathbf{x}_3, t_3)$ is the sum over propagation first to $(\mathbf{x}_2, t_2)$ and then to $(\mathbf{x}_3, t_3)$.

For simplicity, consider one spatial dimension. Divide the interval from the initial to final time into $N - 1$ equal parts:

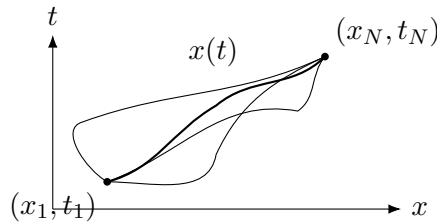
$$dt = \frac{t_N - t_1}{N - 1}. \tag{2.7.15}$$

The full propagator is

$$\begin{aligned}
 K(x_N, t_N; x_1, t_1) &= \int dx_{N-1} dx_{N-2} \cdots dx_2 K(x_N, t_N; x_{N-1}, t_{N-1}) \\
 &\quad \times K(x_{N-1}, t_{N-1}; x_{N-2}, t_{N-2}) \cdots K(x_2, t_2; x_1, t_1).
 \end{aligned} \tag{2.7.16}$$



As $N \rightarrow \infty$, or $dt \rightarrow 0$, this multiple integral becomes a sum over paths.



Unlike quantum mechanics, where every path contributes to evolution in space-time, in classical mechanics a particular initial and final state is connected by a unique path determined by the principle of least action. For example, the unique path of free fall is

$$x = h - \frac{1}{2}gt^2, \quad (2.7.17)$$

where h is the height and g the acceleration of gravity.

3) Path integral

The basic postulate of Feynman's path-integral theory is that the probability amplitude for propagation from (x_1, t_1) to (x_N, t_N) is the sum of contributions from all possible paths $x(t)$:

$$K(x_N, t_N; x_1, t_1) = C \int_{x_1}^{x_N} [dx(t)] e^{iS[x(t)]/\hbar} = C \int_{x_1}^{x_N} [dx(t)] \exp \left[\frac{i}{\hbar} \int_{t_1}^{t_N} L_c(x, \dot{x}) dt \right]. \quad (2.7.18)$$

The contribution of each path is $Ce^{iS[x(t)]/\hbar}$. The action is

$$S[x(t)] = \int_{t_1}^{t_N} L_c(x, \dot{x}) dt, \quad L_c = \frac{1}{2}m\dot{x}^2 - V(x). \quad (2.7.19)$$

To return the path integral to a multiple integral, take $dt = (t_N - t_1)/N$ and obtain

$$K(x_N, t_N; x_1, t_1) = C \lim_{N \rightarrow \infty} \int dx_{N-1} \cdots dx_2 \exp \left[\frac{i}{\hbar} \sum_{n=2}^N L_c(x_n, \dot{x}_n) dt \right]. \quad (2.7.20)$$

If L_c is quadratic in x , the path integral is Gaussian and exactly integrable. For example, for a free particle,

$$L_c = \frac{1}{2}m\dot{x}^2, \quad L_c(x_n, \dot{x}_n)dt = \frac{1}{2}m \left(\frac{x_n - x_{n-1}}{dt} \right)^2 dt = \frac{m(x_n - x_{n-1})^2}{2dt}. \quad (2.7.21)$$

Then

$$\begin{aligned} K(x_N, t_N; x_1, t_1) &= C \lim_{N \rightarrow \infty} \int dx_{N-1} \cdots dx_2 \exp \left[\frac{im}{2\hbar dt} \sum_{n=2}^N (x_n - x_{n-1})^2 \right] \\ &= C \exp \left[\frac{im(x_N - x_1)^2}{2\hbar(t_N - t_1)} \right]. \end{aligned} \quad (2.7.22)$$

The position eigenstates' orthogonality and normalization require

$$\lim_{t_N \rightarrow t_1} K(x_N, t_N; x_1, t_1) = \langle x_N, t_1 | x_1, t_1 \rangle = \delta(x_N - x_1). \quad (2.7.23)$$

Thus

$$K(x_N, t_N; x_1, t_1) = \sqrt{\frac{m}{2\pi i \hbar (t_N - t_1)}} \exp \left[\frac{im(x_N - x_1)^2}{2\hbar(t_N - t_1)} \right]. \quad (2.7.24)$$

This agrees with the calculation in the Schrödinger picture.

As another example, for the harmonic oscillator the classical Lagrangian is

$$L_c = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2. \quad (2.7.25)$$

After evaluating the Gaussian integral,

$$K(x_N, t_N; x_1, t_1) = \sqrt{\frac{m\omega}{2\pi i \hbar \sin[\omega(t_N - t_1)]}} \times \exp \left\{ \frac{im\omega}{2\hbar \sin[\omega(t_N - t_1)]} \left[(x_N^2 + x_1^2) \cos[\omega(t_N - t_1)] - 2x_1 x_N \right] \right\}. \quad (2.7.26)$$

Discussion.

- (1) The phase factor contributed by each path is different. In a classical system, when the action S_{cl} along the classical path is much greater than $\hbar$, the actions on nearby paths differ, and their contributions oscillate rapidly and cancel; the contribution from the stationary-action path remains.
- (2) Only Gaussian path integrals can be evaluated exactly.
- (3) In the Schrödinger picture, the probability amplitude refers to a transition of the state. In Feynman's path integral, the probability amplitude is linked to the particle's motion, namely to paths. Feynman's path-integral theory can be proved equivalent to the Schrödinger equation (see p. 122 of the original source).
- (4) A merit of the path integral is that the difference from classical mechanics is apparent and easily understood. Operators and other q -numbers need not be introduced; all calculations take place using c -numbers. Path integrals are relatively cumbersome in nonrelativistic quantum mechanics, but important in gauge field theory.

Interactive investigation IC-02-06. Open Path phases and stationary action. After writing the path-integral amplitude, vary a family of neighboring paths. Watch their complex phases interfere, identify the stationary-action path, and explain why rapidly changing phases largely cancel.

2.8 Gauge Transformations

1) Quantum coherence caused by a constant potential difference

In classical mechanics, $\mathbf{F} = -\nabla V(\mathbf{x})$, so adding a spatially independent constant to the potential produces no effect.

In quantum mechanics, suppose a particle moves in a potential $V_1(\mathbf{x})$. The evolution from the initial state $|\alpha, t_0\rangle$ is

$$|\alpha, t\rangle_1 = e^{-i\hat{H}(t-t_0)/\hbar} |\alpha, t_0\rangle. \quad (2.8.1)$$

If instead the particle moves in $V_2(\mathbf{x}, t) = V_1(\mathbf{x}) + \Delta V(t)$, the same initial state evolves as

$$\begin{aligned} |\alpha, t\rangle_2 &= e^{-i\hat{H}(t-t_0)/\hbar} \exp \left[-\frac{i}{\hbar} \int_{t_0}^t \Delta V(t') dt' \right] |\alpha, t_0\rangle \\ &= \exp \left[-\frac{i}{\hbar} \int_{t_0}^t \Delta V(t') dt' \right] |\alpha, t\rangle_1. \end{aligned} \quad (2.8.2)$$

The two states differ by a phase factor $e^{i(\varphi_2 - \varphi_1)}$:

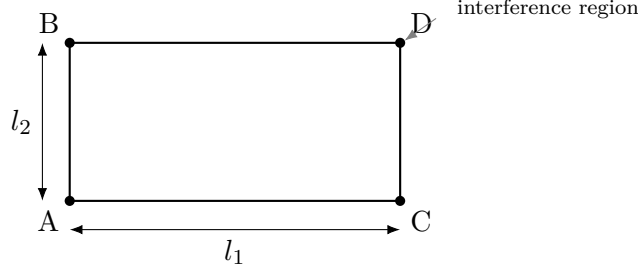
$$\varphi_2 - \varphi_1 = -\frac{1}{\hbar} \int_{t_0}^t \Delta V(t') dt' \xrightarrow{\Delta V \text{ constant}} -\frac{\Delta V}{\hbar} (t - t_0). \quad (2.8.3)$$

This may seem not to lead to any measurable result. However, interference between the two wavefunctions has a term $\cos(\varphi_2 - \varphi_1)$ that depends on the phase difference and is observable.

2) Quantum coherence caused by a gravitational potential difference

Consider the motion of thermal neutrons (uncharged) in a gravitational field.

At point A, the neutron beam is split into two: one travels through A–C–D to D, the other through A–B–D to D. If the frame ABCD is parallel to the ground, the two beams experience the same gravitational potential, and there is no relative gravitational phase at D.



If the frame is rotated upward by an angle δ around AC, the particle in the BD segment has a constant potential difference relative to AC:

$$\Delta V = mgl_2 \sin \delta. \quad (2.8.4)$$

Although the AB and CD segments also have a gravitational potential difference from AC that depends on spatial position, their contributions are the same along both paths and do not affect the relative phase. Let the neutron's travel time from B to D be T . The resulting phase difference at D is

$$\varphi_{ABD} - \varphi_{ACD} = -\frac{m}{\hbar} gl_2 T \sin \delta. \quad (2.8.5)$$

Since $T = l_1/v$, where v is the neutron (wavepacket) velocity, and $mv = p = \hbar/\lambda$ with λ the de Broglie wavelength in the source's convention, we have $T = l_1 m \lambda / \hbar$ and

$$\varphi_{ABD} - \varphi_{ACD} = -\left(\frac{m}{\hbar}\right)^2 gl_1 l_2 \lambda \sin \delta. \quad (2.8.6)$$

As the angle δ changes, the interference term $\cos(\varphi_{ABD} - \varphi_{ACD})$ oscillates. Such oscillations have indeed been observed experimentally (see Fig. 2.6 in the source).

This discussion shows that a spatially uniform potential difference may leave the force unchanged while producing a quantum-interference effect.

3) Gauge transformations in electromagnetism

Scalar and vector potentials:

$$\varphi(\mathbf{x}, t), \quad \mathbf{A}(\mathbf{x}, t). \quad (2.8.7)$$

Fields:

$$\mathbf{E} = -\nabla\varphi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (2.8.8)$$

From the Hamiltonian of classical electrodynamics, the quantum Hamiltonian of a charged particle in an electromagnetic field is

$$\hat{H} = \frac{1}{2m} \left(\hat{\mathbf{p}} - \frac{e}{c} \mathbf{A} \right)^2 + e\varphi, \quad \varphi = \varphi(\hat{\mathbf{x}}, t), \quad \mathbf{A} = \mathbf{A}(\hat{\mathbf{x}}, t). \quad (2.8.9)$$

Using the fundamental commutation relations

$$[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}, \quad (2.8.10)$$

we have

$$\left[\hat{p}_i - \frac{e}{c} A_i, \hat{p}_j - \frac{e}{c} A_j \right] = \frac{i\hbar e}{c} \epsilon_{ijk} B_k. \quad (2.8.11)$$

The Heisenberg equation of motion is

$$\begin{aligned} \frac{d\hat{x}_i}{dt} &= \frac{1}{i\hbar} [\hat{x}_i, \hat{H}] \\ &= \frac{1}{2mi\hbar} \left[\hat{x}_i, \left(\hat{p}_j - \frac{e}{c} A_j \right) \left(\hat{p}_j - \frac{e}{c} A_j \right) \right] \\ &= \frac{1}{m} \left(\hat{p}_i - \frac{e}{c} A_i \right). \end{aligned} \quad (2.8.12)$$

Consequently,

$$\begin{aligned} m \frac{d^2 \hat{x}_i}{dt^2} &= \frac{d}{dt} \left(\hat{p}_i - \frac{e}{c} A_i \right) \\ &= \frac{1}{i\hbar} \left[\hat{p}_i - \frac{e}{c} A_i, \hat{H} \right] - \frac{e}{c} \frac{\partial A_i}{\partial t} \\ &= \frac{e}{2mc} \epsilon_{ijk} \left(\hat{p}_j - \frac{e}{c} A_j \right) B_k - \frac{e}{2mc} \epsilon_{ijk} B_j \left(\hat{p}_k - \frac{e}{c} A_k \right) + eE_i. \end{aligned} \quad (2.8.13)$$

In vector notation,

$$\boxed{m \frac{d^2 \hat{\mathbf{x}}}{dt^2} = e \left[\mathbf{E} + \frac{1}{2c} \left(\frac{d\hat{\mathbf{x}}}{dt} \times \mathbf{B} - \mathbf{B} \times \frac{d\hat{\mathbf{x}}}{dt} \right) \right]}. \quad (2.8.14)$$

This is the Ehrenfest theorem in the Heisenberg picture. Note that the velocity operator $d\hat{\mathbf{x}}/dt$ and $\mathbf{B} = \nabla \times \mathbf{A}$ generally do not commute; operator ordering matters in this expression.

In the coordinate representation of the Schrödinger picture,

$$i\hbar \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = \left[\frac{1}{2m} \left(-i\hbar \nabla - \frac{e}{c} \mathbf{A}(\mathbf{x}, t) \right)^2 + e\varphi(\mathbf{x}, t) \right] \psi(\mathbf{x}, t). \quad (2.8.15)$$

The continuity equation is

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{x}, t) = 0, \quad (2.8.16)$$

where

$$\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2, \quad (2.8.17)$$

$$\mathbf{j}(\mathbf{x}, t) = \frac{\hbar}{m} \text{Im} \left[\psi^*(\mathbf{x}, t) \left(\nabla - \frac{ie}{\hbar c} \mathbf{A}(\mathbf{x}, t) \right) \psi(\mathbf{x}, t) \right]. \quad (2.8.18)$$

Now transform the potentials as

$$\varphi(\mathbf{x}, t) \longrightarrow \varphi(\mathbf{x}, t), \quad \mathbf{A}(\mathbf{x}, t) \longrightarrow \mathbf{A}(\mathbf{x}, t) + \nabla \Lambda(\mathbf{x}), \quad (2.8.19)$$

where $\Lambda(\mathbf{x})$ is an arbitrary scalar function. It is clear that the fields $\mathbf{E}$ and $\mathbf{B}$ remain unchanged. A transformation of potentials leaving the field strengths unchanged is called a *gauge transformation*; distinct functions $\Lambda(\mathbf{x})$ specify different gauges.

Let the state corresponding to the potentials $\varphi, \mathbf{A}$ be $|\psi, t\rangle_1$, and that corresponding to $\varphi, \mathbf{A} + \nabla \Lambda$ be $|\psi, t\rangle_2$. They satisfy respectively

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle_1 = \left[\frac{(\hat{\mathbf{p}} - e\mathbf{A}/c)^2}{2m} + e\varphi \right] |\psi, t\rangle_1, \quad (2.8.20)$$

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle_2 = \left[\frac{[\hat{\mathbf{p}} - e(\mathbf{A} + \nabla \Lambda)/c]^2}{2m} + e\varphi \right] |\psi, t\rangle_2. \quad (2.8.21)$$

Take

$$|\psi, t\rangle_2 = \exp\left(\frac{ie}{\hbar c}\Lambda(\hat{\mathbf{x}})\right) |\psi, t\rangle_1. \quad (2.8.22)$$

Because

$$e^{-ie\Lambda/(\hbar c)} \left[\hat{\mathbf{p}} - \frac{e}{c}(\mathbf{A} + \nabla\Lambda) \right] e^{ie\Lambda/(\hbar c)} = \hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}, \quad (2.8.23)$$

it is easy to show that the two Schrödinger equations are consistent. Thus, under a gauge transformation the states differ by a phase factor:

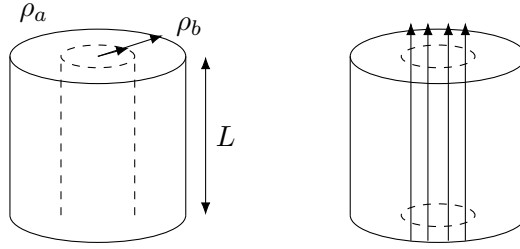
$$|\psi, t\rangle \longrightarrow \exp\left(\frac{ie}{\hbar c}\Lambda(\hat{\mathbf{x}})\right) |\psi, t\rangle. \quad (2.8.24)$$

Does a gauge transformation have observable quantum effects?

4) Aharonov–Bohm (AB) effect

Consider a long hollow cylindrical shell with inner radius ρ_a and outer radius ρ_b . The particle moves in the shell $\rho_a < \rho < \rho_b$; it cannot enter $\rho < \rho_a$ or $\rho > \rho_b$. The wavefunction vanishes at $\rho = \rho_a$ and $\rho = \rho_b$, so this is a bound-state problem.

Now place a long solenoid in the central hollow region, with its magnetic field confined inside the solenoid, ensuring that the field outside vanishes.



Assume

$$\mathbf{B} = \begin{cases} 0, & \rho > \rho_a, \\ B\mathbf{e}_z, & \rho < \rho_a. \end{cases} \quad (2.8.25)$$

In cylindrical coordinates,

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{1}{\rho} \begin{vmatrix} \mathbf{e}_\rho & \rho\mathbf{e}_\phi & \mathbf{e}_z \\ \partial/\partial\rho & \partial/\partial\phi & \partial/\partial z \\ A_\rho & \rho A_\phi & A_z \end{vmatrix}. \quad (2.8.26)$$

One finds

$$\mathbf{A} = \begin{cases} \frac{B\rho_a^2}{2\rho}\mathbf{e}_\phi, & \rho > \rho_a, \\ \frac{B\rho}{2}\mathbf{e}_\phi, & \rho < \rho_a. \end{cases} \quad (2.8.27)$$

Introducing the field into the region $\rho < \rho_a$ changes the vector potential in the particle's accessible region $\rho_a < \rho < \rho_b$ from $\mathbf{A} = 0$ to $\mathbf{A} \neq 0$, while the magnetic field there remains $\mathbf{B} = 0$.

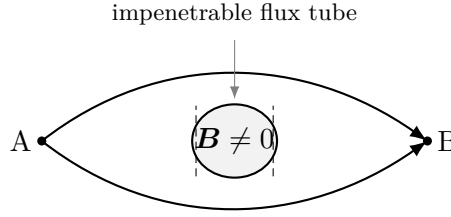
The stationary Schrödinger equation is

$$\frac{1}{2m} \left(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A} \right)^2 \psi = -\frac{\hbar^2}{2m} \left[\mathbf{e}_\rho \frac{\partial}{\partial\rho} + \frac{\mathbf{e}_\phi}{\rho} \left(\frac{\partial}{\partial\phi} - \frac{ieB\rho_a^2}{2\hbar c} \right) + \mathbf{e}_z \frac{\partial}{\partial z} \right]^2 \psi = E\psi, \quad \rho_a < \rho < \rho_b. \quad (2.8.28)$$

Clearly, the magnetic-field term changes the particle's bound-state energies E . Crucially, the magnetic field $\mathbf{B}$ is absent from the particle's accessible region, while $\mathbf{A}$ is present there. In classical mechanics the motion of a charged particle is determined by $\mathbf{E}$ and $\mathbf{B}$, whereas in quantum mechanics a potential without a field strength can still influence the properties of a system. This is a purely quantum effect.

The above is the influence of the AB effect on a bound-state problem.

Now consider the AB effect using the path integral. A charged particle starts at A and passes through two slits in a screen to reach B, where interference occurs. A long solenoid lies between the two slits behind the screen. The magnetic field is inside the solenoid and vanishes outside. For the classical system, the field inside the solenoid cannot affect the interference at B.



For a quantum system, although the field vanishes outside the solenoid, $\mathbf{A} \neq 0$. The action in the path integral is written in the source as

$$\begin{aligned} S[\mathbf{x}(t)] &= \int_{t_A}^{t_B} dt \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A} \right)^2 \\ &\simeq \int_{t_A}^{t_B} dt \left[\frac{m}{2} \left(\frac{d\mathbf{x}}{dt} \right)^2 - \frac{e}{c} \frac{d\mathbf{x}}{dt} \cdot \mathbf{A} \right] \\ &= S_0[\mathbf{x}(t)] - \frac{e}{c} \int_{t_A}^{t_B} dt \frac{d\mathbf{x}}{dt} \cdot \mathbf{A} = S_0[\mathbf{x}(t)] - \frac{e}{c} \int_{\mathbf{x}_A}^{\mathbf{x}_B} \mathbf{A} \cdot d\mathbf{x}. \end{aligned} \quad (2.8.29)$$

The source neglects the term quadratic in $\mathbf{A}$. For any path above the solenoid and any path below it, the two contributions to the amplitude $K(\mathbf{x}_B, t_B; \mathbf{x}_A, t_A)$ have the form

$$\begin{aligned} &\exp \left[-\frac{ie}{\hbar c} \int_{\mathbf{x}_A, \text{above}}^{\mathbf{x}_B} \mathbf{A} \cdot d\mathbf{x} \right] e^{iS_0[\mathbf{x}_{\text{above}}(t)]/\hbar} \\ &+ \exp \left[-\frac{ie}{\hbar c} \int_{\mathbf{x}_A, \text{below}}^{\mathbf{x}_B} \mathbf{A} \cdot d\mathbf{x} \right] e^{iS_0[\mathbf{x}_{\text{below}}(t)]/\hbar}. \end{aligned} \quad (2.8.30)$$

There is a relative phase that affects the interference fringes:

$$\begin{aligned} \frac{e}{\hbar c} \left[\int_{\text{above}} \mathbf{A} \cdot d\mathbf{x} - \int_{\text{below}} \mathbf{A} \cdot d\mathbf{x} \right] &= \frac{e}{\hbar c} \oint \mathbf{A} \cdot d\mathbf{x} \\ &= \frac{e}{\hbar c} \int (\nabla \times \mathbf{A}) \cdot d\mathbf{S} \\ &= \frac{e}{\hbar c} \int \mathbf{B} \cdot d\mathbf{S} = \frac{e}{\hbar c} B \pi R^2. \end{aligned} \quad (2.8.31)$$

This is independent of the precise paths; Stokes' theorem has been used.

The AB effect shows that in quantum mechanics, even when the field strength vanishes, a nonzero potential can influence the interference of particles.

2.9 Berry Phase

The Aharonov–Bohm effect showed that the relative phase of a quantum amplitude can depend on the geometry of a path even where the magnetic field vanishes. A related geometric phase appears when a Hamiltonian depends on external parameters that vary slowly along a path. In this case the path lies in *parameter space*, rather than necessarily in the particle’s physical position space.

1) Instantaneous eigenstates and adiabatic evolution

Let a Hamiltonian depend smoothly on real control parameters $\mathbf{R} = (R_1, \dots, R_d)$:

$$\hat{H}(\mathbf{R}) |n(\mathbf{R})\rangle = E_n(\mathbf{R}) |n(\mathbf{R})\rangle, \quad \langle n(\mathbf{R}) | n(\mathbf{R}) \rangle = 1. \quad (2.9.1)$$

Assume that $\mathbf{R}(t)$ changes sufficiently slowly and that the chosen eigenvalue remains nondegenerate and separated by a gap from the other energy levels along the path. Under the adiabatic approximation a state prepared in $|n(\mathbf{R}(0))\rangle$ follows the corresponding instantaneous eigenspace. Its evolution contains a dynamical phase and a second, geometrical phase:

$$|\psi(t)\rangle \simeq e^{-\frac{i}{\hbar} \int_0^t E_n(\mathbf{R}(t')) dt'} e^{i\gamma_n(t)} |n(\mathbf{R}(t))\rangle. \quad (2.9.2)$$

We determine γ_n from the Schrödinger equation, rather than assuming its form.

Write $|\psi\rangle = e^{-i \int E_n dt / \hbar} e^{i\gamma_n} |n\rangle$ and substitute into

$$i\hbar \frac{d}{dt} |\psi\rangle = \hat{H}(\mathbf{R}(t)) |\psi\rangle. \quad (2.9.3)$$

Differentiation gives the identity within the adiabatic approximation,

$$(E_n - \hbar \dot{\gamma}_n) |n\rangle + i\hbar \frac{d}{dt} |n\rangle = E_n |n\rangle. \quad (2.9.4)$$

Project onto the instantaneous normalized eigenstate to find

$$\dot{\gamma}_n(t) = i \langle n(\mathbf{R}(t)) | \frac{d}{dt} |n(\mathbf{R}(t))\rangle. \quad (2.9.5)$$

The normalization condition implies that $\langle n | \dot{n} \rangle$ is purely imaginary, so $\dot{\gamma}_n$ is real. Apply the chain rule,

$$\frac{d}{dt} |n(\mathbf{R}(t))\rangle = \sum_{\mu} \frac{\partial}{\partial R_{\mu}} |n(\mathbf{R})\rangle \dot{R}_{\mu}, \quad (2.9.6)$$

to rewrite the accumulated phase as an integral along the *parameter path* C :

$$\gamma_n(C) = \int_C \mathcal{A}_n(\mathbf{R}) \cdot d\mathbf{R}, \quad \mathcal{A}_n(\mathbf{R}) = i \langle n(\mathbf{R}) | \nabla_{\mathbf{R}} n(\mathbf{R}) \rangle. \quad (2.9.7)$$

The vector field $\mathcal{A}_n$ is the *Berry connection*. It depends on how the phase of the instantaneous eigenstate has been chosen.

2) Gauge transformations of the eigenstate

The normalized eigenket is physically unchanged by a parameter-dependent phase. Choose a new convention

$$|n'(\mathbf{R})\rangle = e^{-i\beta(\mathbf{R})} |n(\mathbf{R})\rangle, \quad \beta(\mathbf{R}) \in \mathbb{R}. \quad (2.9.8)$$

Substituting it into the Berry connection shows explicitly how the additional derivative acts:

$$\begin{aligned}\mathcal{A}'_n &= i \langle n | e^{i\beta} \nabla_R (e^{-i\beta} |n\rangle) \\ &= i \langle n | \nabla_R n \rangle + \nabla_R \beta \\ &= \boxed{\mathcal{A}_n + \nabla_R \beta}.\end{aligned}\tag{2.9.9}$$

Hence an *open-path* integral transforms as

$$\gamma'_n(C) = \gamma_n(C) + \beta(\mathbf{R}_f) - \beta(\mathbf{R}_i).\tag{2.9.10}$$

For a closed path C , with a single-valued choice of phase along the loop, the two endpoint contributions cancel modulo 2π . Thus the Berry phase around a closed cycle is a gauge-invariant quantity:

$$\boxed{e^{i\gamma_n(C)} \text{ is gauge invariant for a closed } C.}\tag{2.9.11}$$

The distinction matters: neither the connection at one point nor the Berry phase assigned to an open segment is independently gauge invariant. A physically measured phase comes from interference, from a cyclic state evolution, or from an appropriate endpoint comparison.

3) Berry curvature and geometry of parameter space

For three control parameters define the Berry curvature by

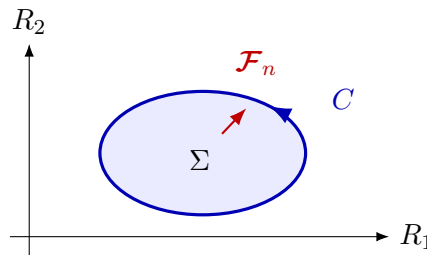
$$\boxed{\mathcal{F}_n(\mathbf{R}) = \nabla_R \times \mathcal{A}_n(\mathbf{R}).}\tag{2.9.12}$$

It is unaffected by the gauge change because $\nabla_R \times \nabla_R \beta = 0$. For a surface Σ whose boundary is C , Stokes' theorem gives

$$\boxed{\gamma_n(C) = \oint_C \mathcal{A}_n \cdot d\mathbf{R} = \int_{\Sigma} \mathcal{F}_n \cdot d\mathbf{S} \pmod{2\pi},}\tag{2.9.13}$$

provided the eigenstate can be chosen smoothly over the surface; singular gauges or degeneracies require appropriate patches. In more than three parameters use the antisymmetric tensor

$$\mathcal{F}_{n,\mu\nu} = \partial_{\mu} \mathcal{A}_{n,\nu} - \partial_{\nu} \mathcal{A}_{n,\mu} = i (\langle \partial_{\mu} n | \partial_{\nu} n \rangle - \langle \partial_{\nu} n | \partial_{\mu} n \rangle).\tag{2.9.14}$$



The analogy with the AB effect is now precise: the AB phase involves the electromagnetic vector potential integrated over a path in physical space; the Berry phase involves a connection generated by the *parameter dependence of eigenstates*. Their common mathematical structure is a line integral of a connection, its gauge transformation by a gradient, and a gauge-invariant curvature or closed-loop phase. The Berry curvature is a property of the quantum eigenstate bundle, not an electromagnetic field in ordinary space. This construction will reappear when Bloch-band eigenstates depend on crystal momentum.

Interactive investigation IC-02-08. Open Berry phase on the Bloch sphere. After separating dynamical and geometric phases, vary a closed adiabatic path on the Bloch sphere. Estimate its enclosed solid angle and compare the observed geometric phase with the predicted half-solid-angle relation.

2.10 Magnetic Monopoles and Charge Quantization

In classical electromagnetic theory, static electric and magnetic fields are asymmetric. The static electric field has sources but no curl:

$$\nabla \cdot \mathbf{E} = 4\pi\rho_e, \quad \nabla \times \mathbf{E} = 0. \quad (2.10.1)$$

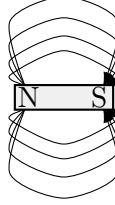
Electric field-line patterns (single-charge configuration):



The static magnetic field has no sources and does have curl:

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J}. \quad (2.10.2)$$

The magnetic field-line pattern has a dipole structure:



Question: Can there be an isolated magnetic charge, that is, a magnetic monopole?

Dirac string. If there is a magnetic charge e_m at the origin, the static field it produces satisfies

$$\nabla \cdot \mathbf{B} = 4\pi e_m \delta^{(3)}(\mathbf{r}). \quad (2.10.3)$$

Its solution is

$$\boxed{\mathbf{B}(\mathbf{r}) = \frac{e_m}{r^2} \mathbf{e}_r.} \quad (2.10.4)$$

Can a vector potential $\mathbf{A}(\mathbf{r})$ be found such that the original relation $\mathbf{B} = \nabla \times \mathbf{A}$ still holds everywhere? The answer is no.

Proof: If $\mathbf{B} = \nabla \times \mathbf{A}$, then necessarily

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0, \quad (2.10.5)$$

contradicting the monopole source equation.

Likewise, by Gauss's theorem, for a closed surface surrounding the monopole,

$$\oint \mathbf{B} \cdot d\mathbf{S} = 4\pi e_m. \quad (2.10.6)$$

But if $\mathbf{B} = \nabla \times \mathbf{A}$, then

$$\oint \mathbf{B} \cdot d\mathbf{S} = \oint (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \int \nabla \cdot (\nabla \times \mathbf{A}) dV = 0, \quad (2.10.7)$$

contradicting Gauss's theorem. Thus no single globally regular vector potential $\mathbf{A}(\mathbf{r})$ can produce $\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}$.

For example, in spherical coordinates take

$$\mathbf{A} = e_m \frac{1 - \cos \theta}{r \sin \theta} \mathbf{e}_\varphi. \quad (2.10.8)$$

For $\theta < \pi - \varepsilon$ (ε infinitesimal), using

$$\begin{aligned} \nabla \times \mathbf{A} &= \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} (A_\varphi \sin \theta) - \frac{\partial A_\theta}{\partial \varphi} \right] \mathbf{e}_r \\ &+ \frac{1}{r} \left[\frac{1}{\sin \theta} \frac{\partial A_r}{\partial \varphi} - \frac{\partial}{\partial r} (r A_\varphi) \right] \mathbf{e}_\theta \\ &+ \frac{1}{r} \left[\frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial \theta} \right] \mathbf{e}_\varphi = \frac{e_m}{r^2} \mathbf{e}_r = \mathbf{B}. \end{aligned} \quad (2.10.9)$$

However, $\mathbf{A}$ diverges at $\theta = \pi$. The negative z axis therefore contains a singular string, called the Dirac string. Its singular contribution means that the distributional curl of this globally extended singular potential does not equal the monopole field alone.

Quantization of electric charge. Although a single vector potential satisfying $\mathbf{B} = \nabla \times \mathbf{A}$ cannot cover all space, such potentials can be found in separate regions. For example, take

$$\boxed{\begin{aligned} \mathbf{A}_I &= e_m \frac{1 - \cos \theta}{r \sin \theta} \mathbf{e}_\varphi, & \theta < \pi - \varepsilon, \\ \mathbf{A}_{II} &= -e_m \frac{1 + \cos \theta}{r \sin \theta} \mathbf{e}_\varphi, & \theta > \varepsilon. \end{aligned}} \quad (2.10.10)$$

$\mathbf{A}_I$ diverges at $\theta = \pi$, while $\mathbf{A}_{II}$ diverges at $\theta = 0$. With this choice of potentials,

$$\nabla \times \mathbf{A} = \frac{e_m}{r^2} \mathbf{e}_r = \mathbf{B} \quad (2.10.11)$$

in each nonsingular region.

In their overlap, since both potentials produce the same magnetic field, they are related by a gauge transformation Λ :

$$\mathbf{A}_{II} - \mathbf{A}_I = -\frac{2e_m}{r \sin \theta} \mathbf{e}_\varphi = \nabla \Lambda. \quad (2.10.12)$$

In spherical coordinates,

$$\nabla \Lambda = \frac{\partial \Lambda}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial \Lambda}{\partial \theta} \mathbf{e}_\theta + \frac{1}{r \sin \theta} \frac{\partial \Lambda}{\partial \varphi} \mathbf{e}_\varphi. \quad (2.10.13)$$

Hence

$$\boxed{\Lambda(\mathbf{r}) = -2e_m \varphi.} \quad (2.10.14)$$

Consider a particle of electric charge e moving in the monopole field. The wavefunctions ψ_I and ψ_{II} corresponding to $\mathbf{A}_I$ and $\mathbf{A}_{II}$ satisfy in the overlapping region

$$\psi_{II} = e^{ie\Lambda/(hc)} \psi_I = e^{-2iee_m\varphi/(hc)} \psi_I. \quad (2.10.15)$$

By single-valuedness of the wavefunction, ψ_{II} must be single valued in the overlap. For example, at $\theta = \pi/2$,

$$\psi_{II}(r, \theta = \pi/2, \varphi = 0) = \psi_{II}(r, \theta = \pi/2, \varphi = 2\pi). \quad (2.10.16)$$

Consequently,

$$\frac{2ee_m}{\hbar c} = \pm n, \quad n = 0, \pm 1, \pm 2, \dots \quad (2.10.17)$$

This is electric-charge quantization:

$$e = \pm \frac{\hbar c}{2e_m} n. \quad (2.10.18)$$

There is a smallest unit of electric charge.

Quantum mechanics does not require the existence of magnetic monopoles. If a magnetic monopole exists, however, electric charge must be quantized and there is a smallest unit of electric charge.

Chapter 3

Angular Momentum Theory

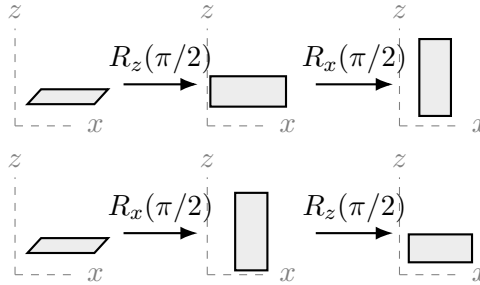
3.1 Rotations and Angular Momentum

Rotations and angular momentum have a unique position in quantum mechanics.

1) Classical rotations

Two rotations about the same axis commute. For instance, rotating first through 30° about the Z axis and then through 60° is exactly equivalent to performing the 60° rotation first and the 30° rotation second.

However, two rotations about different axes do not commute. As shown in the figure, a counterclockwise rotation through 90° about the Z axis followed by a 90° rotation about the X axis gives a result different from performing the X rotation first and the Z rotation second.



The upper and lower sequences show the different final orientations produced by reversing the order of the two rotations.

A three-dimensional vector $\mathbf{V} = V_x \mathbf{e}_x + V_y \mathbf{e}_y + V_z \mathbf{e}_z$ becomes $\mathbf{V}' = V'_x \mathbf{e}_x + V'_y \mathbf{e}_y + V'_z \mathbf{e}_z$ after a rotation $R(\varphi)$. In matrix form,

$$\mathbf{V}' = R(\varphi)\mathbf{V}. \quad (3.1.1)$$

Here $R(\varphi)$ is a 3×3 matrix. Obviously, rotation does not change the length of a vector:

$$|\mathbf{V}'| = |\mathbf{V}|. \quad (3.1.2)$$

The rotation matrix through an angle φ about the Z axis is easily obtained:

$$R_z(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.1.3)$$

For an infinitesimal rotation ϵ ,

$$R_z(\epsilon) = \begin{pmatrix} 1 - \epsilon^2/2 & -\epsilon & 0 \\ \epsilon & 1 - \epsilon^2/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (3.1.4)$$

where terms higher than the second order in ϵ have been omitted. Similarly,

$$R_x(\epsilon) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \epsilon^2/2 & -\epsilon \\ 0 & \epsilon & 1 - \epsilon^2/2 \end{pmatrix}, \quad R_y(\epsilon) = \begin{pmatrix} 1 - \epsilon^2/2 & 0 & \epsilon \\ 0 & 1 & 0 \\ -\epsilon & 0 & 1 - \epsilon^2/2 \end{pmatrix}. \quad (3.1.5)$$

Clearly,

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = \begin{pmatrix} 0 & -\epsilon^2 & 0 \\ \epsilon^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_z(\epsilon^2) - \mathbf{1}. \quad (3.1.6)$$

This expression shows that rotations about X and Y do not commute. In general,

$$R_i(\epsilon)R_j(\epsilon) - R_j(\epsilon)R_i(\epsilon) = \epsilon_{ijk}[R_k(\epsilon^2) - \mathbf{1}]. \quad (3.1.7)$$

2) Quantum rotations

Let the quantum rotation operator corresponding to the classical rotation matrix $R(\varphi)$ be $\hat{D}(\varphi)$:

$$|\alpha\rangle_\varphi = \hat{D}(\varphi) |\alpha\rangle. \quad (3.1.8)$$

Its representation is a matrix $D(\varphi)$.

First discuss the general properties of the rotation operator. The generator of an infinitesimal spatial translation is the momentum operator:

$$\hat{U}_{\text{tr}}(d\mathbf{x}) = \mathbf{1} - \frac{i}{\hbar} \hat{\mathbf{p}} \cdot d\mathbf{x} \simeq \exp\left(-\frac{i}{\hbar} \hat{\mathbf{p}} \cdot d\mathbf{x}\right). \quad (3.1.9)$$

Spatial-translation invariance implies momentum conservation. For an infinitesimal time evolution, the generator is the Hamiltonian:

$$\hat{U}(t + dt, t) = \mathbf{1} - \frac{i}{\hbar} \hat{H} dt \simeq \exp\left(-\frac{i}{\hbar} \hat{H} dt\right). \quad (3.1.10)$$

Time-translation invariance implies energy conservation. For a rotation, classical mechanics tells us that its generator is the corresponding conserved mechanical quantity: angular momentum. Let $\hat{\mathbf{J}}$ be the generator of the quantum rotation operator $\hat{D}(\varphi)$. For an infinitesimal rotation $d\varphi$ about the direction $\mathbf{e}_n$,

$$\hat{D}(d\varphi) = \mathbf{1} - \frac{i}{\hbar} \hat{\mathbf{J}} \cdot \mathbf{e}_n d\varphi \simeq \exp\left(-\frac{i}{\hbar} \hat{\mathbf{J}} \cdot \mathbf{e}_n d\varphi\right). \quad (3.1.11)$$

Because rotations about one axis commute, a finite rotation can be regarded as a sequence of infinitesimal rotations. Thus

$$\boxed{\hat{D}(\varphi) = \exp\left(-\frac{i}{\hbar} \hat{\mathbf{J}} \cdot \mathbf{e}_n \varphi\right)}. \quad (3.1.12)$$

Assume that $\hat{D}$ in quantum mechanics and R in classical mechanics satisfy the same commutation relations:

$$[\hat{D}_i(\epsilon), \hat{D}_j(\epsilon)] = \epsilon_{ijk} [\hat{D}_k(\epsilon^2) - \mathbf{1}]. \quad (3.1.13)$$

Retaining terms through ϵ^2 ,

$$\left[\mathbf{1} - \frac{i}{\hbar} \hat{J}_i \epsilon - \frac{\hat{J}_i^2 \epsilon^2}{2\hbar^2}, \mathbf{1} - \frac{i}{\hbar} \hat{J}_j \epsilon - \frac{\hat{J}_j^2 \epsilon^2}{2\hbar^2} \right] = \epsilon_{ijk} \left(\mathbf{1} - \frac{i}{\hbar} \hat{J}_k \epsilon^2 - \mathbf{1} \right). \quad (3.1.14)$$

Hence

$$\boxed{[\hat{J}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{J}_k}, \quad \hat{\mathbf{J}} \times \hat{\mathbf{J}} = i\hbar \hat{\mathbf{J}}. \quad (3.1.15)$$

This is the fundamental angular-momentum commutation relation in quantum mechanics.

3) Expectation values of angular momentum

Consider a rotation by φ about Z . The state changes as

$$|\alpha\rangle \longrightarrow |\alpha\rangle_\varphi = \hat{D}_z(\varphi) |\alpha\rangle = e^{-i\hat{J}_z \varphi / \hbar} |\alpha\rangle. \quad (3.1.16)$$

The expectation value of $\hat{J}_x$ changes from

$$\langle \hat{J}_x \rangle_0 = \langle \alpha | \hat{J}_x | \alpha \rangle \quad (3.1.17)$$

to

$$\langle \hat{J}_x \rangle_\varphi = \langle \alpha_\varphi | \hat{J}_x | \alpha_\varphi \rangle = \langle \alpha | e^{i\hat{J}_z \varphi / \hbar} \hat{J}_x e^{-i\hat{J}_z \varphi / \hbar} | \alpha \rangle. \quad (3.1.18)$$

Expand $e^{\pm i\hat{J}_z \varphi / \hbar}$ in a power series and use

$$[\hat{J}_z, \hat{J}_x] = i\hbar \hat{J}_y. \quad (3.1.19)$$

We obtain

$$e^{i\hat{J}_z \varphi / \hbar} \hat{J}_x e^{-i\hat{J}_z \varphi / \hbar} = \hat{J}_x \cos \varphi - \hat{J}_y \sin \varphi. \quad (3.1.20)$$

Therefore

$$\langle \hat{J}_x \rangle_\varphi = \langle \hat{J}_x \rangle_0 \cos \varphi - \langle \hat{J}_y \rangle_0 \sin \varphi. \quad (3.1.21)$$

Similarly,

$$\langle \hat{J}_y \rangle_\varphi = \langle \hat{J}_y \rangle_0 \cos \varphi + \langle \hat{J}_x \rangle_0 \sin \varphi, \quad \langle \hat{J}_z \rangle_\varphi = \langle \hat{J}_z \rangle_0. \quad (3.1.22)$$

If $\langle \hat{J}_x \rangle$, $\langle \hat{J}_y \rangle$ and $\langle \hat{J}_z \rangle$ are regarded as three components of $\langle \hat{\mathbf{J}} \rangle$, then

$$\boxed{\langle \hat{\mathbf{J}} \rangle_\varphi = R_z(\varphi) \langle \hat{\mathbf{J}} \rangle_0}. \quad (3.1.23)$$

This obeys the same transformation law as a vector mechanical quantity in classical mechanics.

3.2 Spin Angular Momentum

1) Spin precession (relation between the rotation operator and the time-evolution operator)

We have previously discussed the time evolution of spin precession about the Z axis. In the particle's own coordinate frame there is no kinetic energy, only the interaction of its spin with the external magnetic field:

$$\hat{H} = \omega \hat{s}_z, \quad \omega = -\frac{eB}{mc}, \quad \hat{U}(t, 0) = e^{-i\hat{H}t/\hbar} = e^{-i\omega \hat{s}_z t / \hbar}. \quad (3.2.1)$$

Suppose the initial state ($t = 0$) is the eigenstate $|s_x^+\rangle$ of $\hat{s}_x$. Then at any time t the expectation values are

$$\langle \hat{s}_x \rangle_t = \frac{\hbar}{2} \cos \omega t, \quad \langle \hat{s}_y \rangle_t = \frac{\hbar}{2} \sin \omega t, \quad \langle \hat{s}_z \rangle_t = 0. \quad (3.2.2)$$

Compare $\hat{U}(t, 0) = e^{-i\omega\hat{s}_z t/\hbar}$ with the rotation operator about Z , $\hat{D}_z(\varphi) = e^{-i\hat{s}_z\varphi/\hbar}$. If $\varphi = \omega t$, the angle φ corresponds one-to-one with t . From the relation between angular-momentum expectation values before and after a rotation,

$$\begin{cases} \langle \hat{s}_x \rangle_\varphi = \langle \hat{s}_x \rangle_0 \cos \varphi - \langle \hat{s}_y \rangle_0 \sin \varphi, \\ \langle \hat{s}_y \rangle_\varphi = \langle \hat{s}_y \rangle_0 \cos \varphi + \langle \hat{s}_x \rangle_0 \sin \varphi, \\ \langle \hat{s}_z \rangle_\varphi = \langle \hat{s}_z \rangle_0, \end{cases} \quad (3.2.3)$$

we obtain for any initial state

$$\begin{cases} \langle \hat{s}_x \rangle_t = \langle \hat{s}_x \rangle_0 \cos \omega t - \langle \hat{s}_y \rangle_0 \sin \omega t, \\ \langle \hat{s}_y \rangle_t = \langle \hat{s}_y \rangle_0 \cos \omega t + \langle \hat{s}_x \rangle_0 \sin \omega t, \\ \langle \hat{s}_z \rangle_t = \langle \hat{s}_z \rangle_0. \end{cases} \quad (3.2.4)$$

These expressions are the expectation values of spin angular momentum at time t for an arbitrary initial state.

For $t = 2\pi/\omega$,

$$\langle \hat{s}_x \rangle_t = \langle \hat{s}_x \rangle_0, \quad \langle \hat{s}_y \rangle_t = \langle \hat{s}_y \rangle_0, \quad \langle \hat{s}_z \rangle_t = \langle \hat{s}_z \rangle_0. \quad (3.2.5)$$

Thus $2\pi/\omega$ is called the spin-precession period.

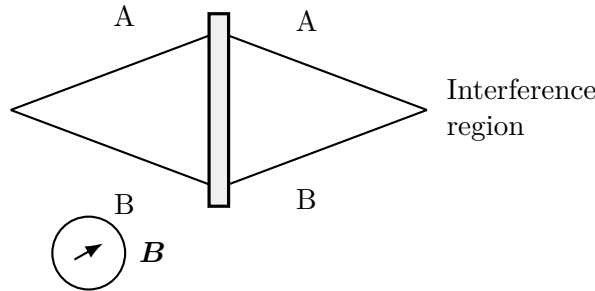
Expand an arbitrary initial state in the eigenstates $|s_z^\pm\rangle$ of $\hat{s}_z$. Under a rotation about Z ,

$$\begin{aligned} |\alpha\rangle &= |s_z^+\rangle \langle s_z^+|\alpha\rangle + |s_z^-\rangle \langle s_z^-|\alpha\rangle \\ &\rightarrow |\alpha\rangle_\varphi = e^{-i\hat{s}_z\varphi/\hbar} |\alpha\rangle = e^{-i\varphi/2} |s_z^+\rangle \langle s_z^+|\alpha\rangle + e^{i\varphi/2} |s_z^-\rangle \langle s_z^-|\alpha\rangle. \end{aligned} \quad (3.2.6)$$

When $\varphi = 2\pi$, or $t = 2\pi/\omega$, $|\alpha\rangle_\varphi = -|\alpha\rangle$; the state does not return to itself! Only after $\varphi = 4\pi$ does it return. Unlike the period of the expectation values, $2\pi/\omega$, the period of the state is $4\pi/\omega$.

2) Neutron interference experiment (experimental verification of spin precession)

The neutron has spin 1/2. A neutron beam with almost fixed energy is split into two beams, which travel along paths A and B to an interference region. Along path B a region containing a nonzero magnetic field is inserted, as illustrated.



The magnetic moment of the neutron is

$$\mu_n = \frac{g_n e \hbar}{2m_n c}, \quad g_n = -1.91, \quad (3.2.7)$$

and the spin-precession frequency is

$$\omega = \frac{g_n e B}{m_n c}. \quad (3.2.8)$$

(Note: for an electron, $g_e = -1$, and the precession frequency is $\omega = g_e e B / (m_e c)$ in the source's convention.) From the time evolution

$$\hat{U}(t, 0) = e^{-i\hat{H}t/\hbar} = e^{-i\omega\hat{s}_z t/\hbar}, \quad (3.2.9)$$

the state of the neutron following path B acquires an additional phase factor $e^{\pm i\omega T/2}$ after passing through the field region. Here T is the time the neutron takes to traverse that region:

$$T = l/v_n = lm_n/p_n = lm_n\lambda/\hbar. \quad (3.2.10)$$

The interference consequently has a term

$$\cos\left(\frac{\omega T}{2} + \delta\right), \quad (3.2.11)$$

where δ is the phase difference between paths A and B in the absence of the magnetic field.

Changing the external field B changes the Larmor frequency and produces a periodic change in the interference intensity. The frequency change corresponding to one full period obeys

$$\frac{\Delta\omega T}{2} = 2\pi, \quad (3.2.12)$$

corresponding to

$$\boxed{\Delta B = \frac{4\pi m_n c}{g_n e T}.} \quad (3.2.13)$$

Such changes in interference intensity caused by varying the magnetic field have been confirmed experimentally.

3) Pauli matrices

Introduce the Pauli operator $\hat{\sigma}$:

$$\hat{s} = \frac{\hbar}{2} \hat{\sigma}. \quad (3.2.14)$$

The eigenvalues of $\hat{\sigma}_i$ are ± 1 ; $\hat{\sigma}_i^2 = \mathbf{1}$, and $\hat{\sigma}^2 = 3\mathbf{1}$. Using the commutation relations

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk}\hat{\sigma}_k, \quad (3.2.15)$$

we have, for example,

$$\begin{aligned} \hat{\sigma}_x \hat{\sigma}_y + \hat{\sigma}_y \hat{\sigma}_x &= \frac{1}{2i}(\hat{\sigma}_y \hat{\sigma}_z - \hat{\sigma}_z \hat{\sigma}_y)\hat{\sigma}_y + \frac{1}{2i}\hat{\sigma}_y(\hat{\sigma}_y \hat{\sigma}_z - \hat{\sigma}_z \hat{\sigma}_y) \\ &= \frac{1}{2i}(-\hat{\sigma}_z \hat{\sigma}_y^2 + \hat{\sigma}_y^2 \hat{\sigma}_z) = \frac{1}{2i}(-\hat{\sigma}_z + \hat{\sigma}_z) = 0. \end{aligned} \quad (3.2.16)$$

Hence the anticommutation relations are

$$\boxed{\{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{ij}\mathbf{1}.} \quad (3.2.17)$$

In the $\hat{\sigma}_z$ representation,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.2.18)$$

The eigenstates of $\hat{\sigma}_z$ are

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (3.2.19)$$

By completeness, any spin state has the form $\chi = c_+\chi_+ + c_-\chi_-$.

Consider rotations of a spin state. In the s_z representation,

$$D_n(\varphi) = e^{-i\hat{s}\cdot\mathbf{e}_n\varphi/\hbar} = e^{-i\boldsymbol{\sigma}\cdot\mathbf{e}_n\varphi/2}. \quad (3.2.20)$$

It is straightforward to prove for arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$ that

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \quad (3.2.21)$$

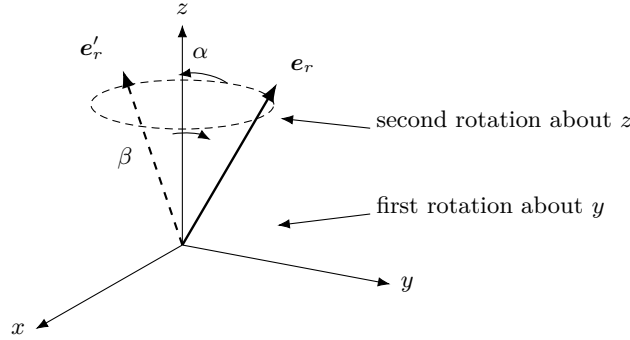
In particular, $(\boldsymbol{\sigma} \cdot \mathbf{a})^2 = |\mathbf{a}|^2$, and

$$(\boldsymbol{\sigma} \cdot \mathbf{e}_n)^r = \begin{cases} 1, & r \text{ even,} \\ \boldsymbol{\sigma} \cdot \mathbf{e}_n, & r \text{ odd.} \end{cases} \quad (3.2.22)$$

Therefore

$$D_n(\varphi) = \cos \frac{\varphi}{2} \mathbf{1} - i\boldsymbol{\sigma} \cdot \mathbf{e}_n \sin \frac{\varphi}{2}. \quad (3.2.23)$$

Example: any spin state can be generated by rotating an eigenstate of s_z (or σ_z). Suppose that the orientation of an arbitrary spin state χ is $\mathbf{e}_n$, whose polar and azimuthal angles are β and α . Assume it has eigenvalue $+1$:



$$(\boldsymbol{\sigma} \cdot \mathbf{e}_n)\chi = \chi. \quad (3.2.24)$$

It may be obtained by rotating the eigenstate $\sigma_z\chi_+ = \chi_+$, $\chi_+ = (1, 0)^T$, through an angle β about Y and then through an angle α about Z :

$$\begin{aligned} \chi &= D_z(\alpha)D_y(\beta)\chi_+ = e^{-i\sigma_z\alpha/2}e^{-i\sigma_y\beta/2}\chi_+ \\ &= \left(\cos \frac{\alpha}{2} - i\sigma_z \sin \frac{\alpha}{2} \right) \left(\cos \frac{\beta}{2} - i\sigma_y \sin \frac{\beta}{2} \right) \chi_+ \\ &= \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \begin{pmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\beta/2)e^{-i\alpha/2} \\ \sin(\beta/2)e^{i\alpha/2} \end{pmatrix} = \cos \frac{\beta}{2}e^{-i\alpha/2}\chi_+ + \sin \frac{\beta}{2}e^{i\alpha/2}\chi_-. \end{aligned} \quad (3.2.25)$$

4) Euler rotations

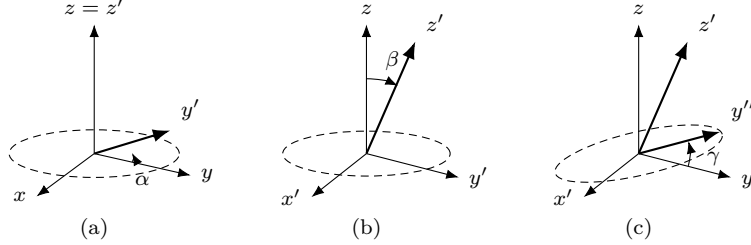
We now discuss how an arbitrary rotation can be expressed as a product of rotations about the X , Y and Z axes.

Let (X, Y, Z) be fixed coordinates and (X', Y', Z') a frame that moves with the vector. In analytical mechanics, an arbitrary rigid-body rotation can be described by three Euler angles. Initially the two frames coincide. First rotate counterclockwise about Z by α , as shown in (a); Y' then no longer coincides with Y . Next rotate about Y' by β , as in (b), after which Z' no

longer coincides with Z . Finally rotate about Z' by γ , as in (c); the Y' axis changes again, becoming Y'' . Thus an arbitrary rigid-body rotation can be written as

$$R(\alpha, \beta, \gamma) = R_{z'}(\gamma)R_{y'}(\beta)R_z(\alpha). \quad (3.2.26)$$

The angles α, β, γ are the Euler angles.



Two of these three Euler rotations are about moving axes, y' and z' , which is inconvenient for calculations: rotations generated by angular momentum are defined with respect to the fixed frame. We must therefore express rotations about moving axes as rotations about fixed axes. The preceding three rotations are equivalent to three rotations about the fixed axes:

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma). \quad (3.2.27)$$

The corresponding quantum rotation operator factorizes as

$$\hat{D}(\alpha, \beta, \gamma) = \hat{D}_z(\alpha)\hat{D}_y(\beta)\hat{D}_z(\gamma) = e^{-i\hat{J}_z\alpha/\hbar}e^{-i\hat{J}_y\beta/\hbar}e^{-i\hat{J}_z\gamma/\hbar}. \quad (3.2.28)$$

For spin 1/2,

$$D(\alpha, \beta, \gamma) = e^{-i\sigma_z\alpha/2}e^{-i\sigma_y\beta/2}e^{-i\sigma_z\gamma/2}. \quad (3.2.29)$$

In the s_z representation,

$$\begin{aligned} D(\alpha, \beta, \gamma) &= \left(\cos \frac{\alpha}{2} - i\sigma_z \sin \frac{\alpha}{2} \right) \left(\cos \frac{\beta}{2} - i\sigma_y \sin \frac{\beta}{2} \right) \left(\cos \frac{\gamma}{2} - i\sigma_z \sin \frac{\gamma}{2} \right) \\ &= \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \begin{pmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{pmatrix} \begin{pmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{pmatrix} \\ &= \begin{pmatrix} e^{-i(\alpha+\gamma)/2} \cos(\beta/2) & -e^{-i(\alpha-\gamma)/2} \sin(\beta/2) \\ e^{i(\alpha-\gamma)/2} \sin(\beta/2) & e^{i(\alpha+\gamma)/2} \cos(\beta/2) \end{pmatrix}. \end{aligned} \quad (3.2.30)$$

3.3 Ensembles and the Density Operator

We consider the spin properties of a many-particle system.

1) Pure ensembles and mixed ensembles

Identical physical systems form an ensemble; for example, an ensemble may consist of particles with identical spins. The spin state of a spin-1/2 particle, with azimuthal and polar angles α, β , is

$$\chi(\alpha, \beta) = c_+(\alpha, \beta)\chi_+ + c_-(\alpha, \beta)\chi_- = \cos \frac{\beta}{2} e^{-i\alpha/2} \chi_+ + \sin \frac{\beta}{2} e^{i\alpha/2} \chi_-, \quad (3.3.1)$$

where χ_+, χ_- are eigenstates of $\hat{s}_z$ and

$$\frac{c_+}{c_-} = \frac{\cos(\beta/2)}{e^{i\alpha} \sin(\beta/2)}. \quad (3.3.2)$$

If all particles are in the same state, e.g., their spins point in the same direction, the ensemble is called a *pure ensemble* or a polarized system. If particles are in different states, e.g., their spins do not all point in one direction, the ensemble is called a *mixed ensemble*. For example, if 70% of particles have spin up and 30% spin down, the system is partially polarized and forms a partially mixed ensemble. In a system with completely random spin directions, the probabilities of spin up and spin down are each 50%. Their overall contributions cancel and there is no net spin polarization.

2) Ensemble averages and the density operator

For a fixed state $|\alpha\rangle$, the average of a mechanical quantity $\hat{A}$ is

$$\langle A \rangle_\alpha = \langle \alpha | \hat{A} | \alpha \rangle. \quad (3.3.3)$$

This is a *quantum average*. Insert any complete basis B :

$$\langle A \rangle_\alpha = \sum_{b,b'} \langle \alpha | b' \rangle \langle b' | \hat{A} | b \rangle \langle b | \alpha \rangle. \quad (3.3.4)$$

The ensemble average is

$$\langle A \rangle_\rho = \sum_\alpha w_\alpha \langle A \rangle_\alpha, \quad (3.3.5)$$

where w_α is the probability that the system occupies state $|\alpha\rangle$. It satisfies the normalization condition

$$\sum_\alpha w_\alpha = 1. \quad (3.3.6)$$

This is a *statistical average*. Since

$$\begin{aligned} \langle A \rangle_\rho &= \sum_{\alpha,b,b'} w_\alpha \langle \alpha | b' \rangle \langle b' | \hat{A} | b \rangle \langle b | \alpha \rangle \\ &= \sum_{b,b'} \left(\sum_\alpha w_\alpha \langle b | \alpha \rangle \langle \alpha | b' \rangle \right) \langle b' | \hat{A} | b \rangle, \end{aligned} \quad (3.3.7)$$

we define the density operator by

$$\hat{\rho} = \sum_\alpha w_\alpha |\alpha\rangle \langle \alpha|. \quad (3.3.8)$$

Its matrix elements in the B representation are

$$\rho_{bb'} = \langle b | \hat{\rho} | b' \rangle = \sum_\alpha w_\alpha \langle b | \alpha \rangle \langle \alpha | b' \rangle. \quad (3.3.9)$$

Therefore

$$\langle A \rangle_\rho = \sum_{b,b'} \langle b | \hat{\rho} | b' \rangle \langle b' | \hat{A} | b \rangle = \sum_b \langle b | \hat{\rho} \hat{A} | b \rangle \equiv \text{tr}(\hat{\rho} \hat{A}). \quad (3.3.10)$$

This fundamental equation of quantum statistical mechanics includes both quantum and statistical averaging. The trace is unchanged by a change of representation, so the formula is independent of the chosen representation.

In a continuous representation, e.g., the position basis, the matrix elements are

$$\rho_{\mathbf{x}\mathbf{x}'} = \langle \mathbf{x} | \hat{\rho} | \mathbf{x}' \rangle = \sum_\alpha w_\alpha \langle \mathbf{x} | \alpha \rangle \langle \alpha | \mathbf{x}' \rangle = \sum_\alpha w_\alpha \psi_\alpha(\mathbf{x}) \psi_\alpha^*(\mathbf{x}'). \quad (3.3.11)$$

The ensemble average becomes

$$\langle A \rangle_\rho = \text{tr}(\hat{\rho}\hat{A}) = \int d^3x \langle \mathbf{x} | \hat{\rho}\hat{A} | \mathbf{x} \rangle. \quad (3.3.12)$$

The density matrix satisfies the normalization condition

$$\begin{aligned} \text{tr } \hat{\rho} &= \sum_{\alpha, b} w_\alpha \langle b | \alpha \rangle \langle \alpha | b \rangle \\ &= \sum_{\alpha} w_\alpha \langle \alpha | \left(\sum_b | b \rangle \langle b | \right) | \alpha \rangle \quad (\text{completeness}) \\ &= \sum_{\alpha} w_\alpha \langle \alpha | \alpha \rangle \quad (\text{quantum-state normalization}) \\ &= \sum_{\alpha} w_\alpha = 1 \quad (\text{statistical normalization}). \end{aligned} \quad (3.3.13)$$

Suppose the density operator has an eigenstate $|\theta\rangle$:

$$\hat{\rho} |\theta\rangle = \theta |\theta\rangle, \quad \hat{\rho}^2 |\theta\rangle = \theta^2 |\theta\rangle. \quad (3.3.14)$$

For a pure ensemble all systems are in the same state $|n\rangle$:

$$w_\alpha = \begin{cases} 1, & \alpha = n, \\ 0, & \alpha \neq n. \end{cases} \quad (3.3.15)$$

Then

$$\hat{\rho} = |n\rangle \langle n|, \quad \hat{\rho}^2 = |n\rangle \langle n| n \rangle \langle n| = |n\rangle \langle n| = \hat{\rho}, \quad (3.3.16)$$

so $\theta^2 = \theta$ and its eigenvalues are $\theta = 0, 1$. If $|n\rangle$ is a basis state of the chosen representation, the density matrix is diagonal:

$$\rho_{bb'} = \langle b | n \rangle \langle n | b' \rangle = \delta_{bn} \delta_{b'n}, \quad \rho = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}. \quad (3.3.17)$$

Only one matrix entry is 1; all others are 0.

For a completely mixed ensemble, all possible states occur with equal probabilities $w_\alpha = w_0 = 1/N$, where N is the number of states. Then

$$\rho_{bb'} = \sum_{\alpha} w_\alpha \langle b | \alpha \rangle \langle \alpha | b' \rangle = \frac{1}{N} \langle b | b' \rangle = \frac{1}{N} \delta_{bb'}, \quad (3.3.18)$$

and the density matrix is proportional to the identity:

$$\rho = \frac{1}{N} \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}. \quad (3.3.19)$$

Example 1: pure ensemble. Suppose that the completely polarized state is $|n\rangle = |+\rangle$, with $\hat{s}_z |\pm\rangle = \pm(\hbar/2) |\pm\rangle$. Then

$$\hat{\rho} = |+\rangle \langle +|. \quad (3.3.20)$$

In the s_z basis ρ is a 2×2 matrix:

$$\rho_{++} = \langle +|+\rangle \langle +|+\rangle = 1, \quad \rho_{+-} = \langle +|+\rangle \langle +|-\rangle = 0, \quad (3.3.21)$$

$$\rho_{-+} = \langle -|+\rangle \langle +|+\rangle = 0, \quad \rho_{--} = \langle -|+\rangle \langle +|-\rangle = 0, \quad \rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.3.22)$$

If the completely polarized state is

$$|n\rangle = |s_x^+\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle, \quad \hat{s}_x |s_x^+\rangle = \frac{\hbar}{2} |s_x^+\rangle, \quad (3.3.23)$$

then $\hat{\rho} = |s_x^+\rangle \langle s_x^+|$. In the s_z representation,

$$\rho_{++} = \langle +|s_x^+\rangle \langle s_x^+|+\rangle = \frac{1}{2}, \quad \rho_{+-} = \langle +|s_x^+\rangle \langle s_x^+|-\rangle = \frac{1}{2}, \quad (3.3.24)$$

$$\rho_{-+} = \langle -|s_x^+\rangle \langle s_x^+|+\rangle = \frac{1}{2}, \quad \rho_{--} = \langle -|s_x^+\rangle \langle s_x^+|-\rangle = \frac{1}{2}, \quad \rho = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (3.3.25)$$

Example 2: density matrix for a completely unpolarized system. With $N = 2$,

$$\hat{\rho} = \frac{1}{2} |+\rangle \langle +| + \frac{1}{2} |-\rangle \langle -|. \quad (3.3.26)$$

In the s_z representation,

$$\rho_{++} = \frac{1}{2}, \quad \rho_{+-} = 0, \quad \rho_{-+} = 0, \quad \rho_{--} = \frac{1}{2}, \quad \rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.3.27)$$

The density matrix of a completely mixed ensemble is proportional to the identity. The expectation values in a completely unpolarized ensemble obviously cancel:

$$\langle s_i \rangle_\rho = \text{tr}(\hat{\rho} \hat{s}_i) = \frac{1}{2} \text{tr}(\hat{s}_i). \quad (3.3.28)$$

Inserting the spin matrices s_x, s_y, s_z gives $\langle s_i \rangle_\rho = 0$.

Example 3: density matrix for a partially polarized system. Let

$$\hat{\rho} = \frac{3}{4} |+\rangle \langle +| + \frac{1}{4} |s_x^+\rangle \langle s_x^+|. \quad (3.3.29)$$

In the s_z representation,

$$\rho_{++} = \frac{3}{4} + \frac{1}{4} \frac{1}{2} = \frac{7}{8}, \quad \rho_{+-} = \rho_{-+} = \rho_{--} = \frac{1}{8}, \quad (3.3.30)$$

so that

$$\rho = \frac{1}{8} \begin{pmatrix} 7 & 1 \\ 1 & 1 \end{pmatrix}, \quad \langle s_x \rangle_\rho = \frac{\hbar}{8}, \quad \langle s_y \rangle_\rho = 0, \quad \langle s_z \rangle_\rho = \frac{3\hbar}{8}. \quad (3.3.31)$$

Interactive investigation IC-03-03. Open Pure and mixed density operators. After comparing ensembles, use different Bloch-ball decompositions to prepare the same density matrix. Check trace normalization, purity $\text{tr}(\rho^2)$, and which ensemble details a measurement can actually distinguish.

3) Evolution of the ensemble

How does the density operator evolve in time? If the statistical properties of the ensemble do not vary in time, its probability distribution w_α is constant and

$$\hat{\rho}(t) = \sum_{\alpha} w_{\alpha} |\alpha, t\rangle \langle \alpha, t|. \quad (3.3.32)$$

By the state evolution equations

$$i\hbar \frac{\partial}{\partial t} |\alpha, t\rangle = \hat{H} |\alpha, t\rangle, \quad -i\hbar \frac{\partial}{\partial t} \langle \alpha, t| = \langle \alpha, t| \hat{H}, \quad (3.3.33)$$

we obtain

$$\begin{aligned} i\hbar \frac{\partial \hat{\rho}}{\partial t} &= \sum_{\alpha} w_{\alpha} \left(\hat{H} |\alpha, t\rangle \langle \alpha, t| - |\alpha, t\rangle \langle \alpha, t| \hat{H} \right) \\ &= -[\hat{\rho}, \hat{H}]. \end{aligned} \quad (3.3.34)$$

This is the density operator's time-evolution equation. It resembles the equation of motion of mechanical quantities in the Heisenberg picture, differing by a minus sign; however $\hat{\rho}$ is not an observable operator but an operator constructed from states in the Schrödinger picture.

3.4 Composite Systems and Generalized Quantum Measurements

The preceding discussion described pure and mixed ensembles by a density operator $\hat{\rho}$. We now ask how a measurement can be modeled as a physical interaction. The essential observation is that the system of interest and the measuring device together form a composite quantum system. Unitary evolution of that larger system, followed by a readout of the probe, generates measurement probabilities and conditional state changes of the measured subsystem.

1) Composite systems and reduced states

Let two subsystems A, B have Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$. The composite Hilbert space is

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (3.4.1)$$

For bases $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$, an arbitrary pure composite state has the form

$$|\Psi_{AB}\rangle = \sum_{ij} c_{ij} |i\rangle_A \otimes |j\rangle_B, \quad \sum_{ij} |c_{ij}|^2 = 1. \quad (3.4.2)$$

Interactive investigation IC-01-08. Open Tensor-product indices. Once the composite basis $ket_{i_A} otimes ket_{j_B}$ is defined, construct two-qubit product states and their Kronecker-product matrices. Track index ordering and confirm that the combined Hilbert space has dimension four.

If $|\Psi_{AB}\rangle$ cannot be factored as a product $|\psi_A\rangle \otimes |\psi_B\rangle$, the two subsystems are entangled. The density operator of the full system is $\hat{\rho}_{AB} = |\Psi_{AB}\rangle \langle \Psi_{AB}|$; the state of subsystem A is obtained by *partial trace*,

$$\hat{\rho}_A = \text{tr}_B \hat{\rho}_{AB} = \sum_j {}_B \langle j| \hat{\rho}_{AB} |j\rangle_B. \quad (3.4.3)$$

Expectation values of every observable acting only on A can be computed equivalently as

$$\text{tr}_{AB} [\hat{\rho}_{AB} (\hat{A} \otimes \mathbf{1}_B)] = \text{tr}_A (\hat{\rho}_A \hat{A}). \quad (3.4.4)$$

For example, for the entangled Bell state

$$|\Phi^+\rangle_{AB} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad (3.4.5)$$

the reduced state is

$$\hat{\rho}_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}\mathbf{1}_A. \quad (3.4.6)$$

The composite state is pure even though its subsystem is maximally mixed. The purity $\text{tr} \hat{\rho}^2$ equals one for a pure state and is smaller for a mixed state; it is preserved under unitary evolution of the closed composite system but need not be preserved under a partial trace.

Interactive investigation IC-03-04. Open Entanglement and partial trace. After tracing out one half of a Bell state, vary the amplitudes of a two-qubit pure state. Calculate the reduced density matrices and compare their eigenvalues, purity, and entropy with the full pure state.

2) Projective measurement and conditional states

A projective measurement of an observable $\hat{A} = \sum_j a_j \hat{Q}_j$ uses mutually orthogonal projectors $\hat{Q}_j \hat{Q}_k = \delta_{jk} \hat{Q}_j$ and $\sum_j \hat{Q}_j = \mathbf{1}$. As established earlier, outcome j has probability

$$p_j = \text{tr}(\hat{Q}_j \hat{\rho}), \quad (3.4.7)$$

and the selectively updated state for a Lüders measurement is

$$\hat{\rho}_j = \frac{\hat{Q}_j \hat{\rho} \hat{Q}_j}{p_j} \quad (p_j > 0). \quad (3.4.8)$$

If the outcome is ignored, the state is $\sum_j \hat{Q}_j \hat{\rho} \hat{Q}_j$. General laboratory measurements need not be described by orthogonal projectors on the system alone; this is the motivation for the system–probe construction.

3) A measurement as a system–probe interaction

Denote the measured system by S and the probe or meter by M . Prepare an initially uncorrelated state

$$|\Psi_{SM}(0)\rangle = |\psi\rangle_S \otimes |\theta\rangle_M. \quad (3.4.9)$$

For an interaction of duration τ , the combined evolution is unitary:

$$|\Psi_{SM}(\tau)\rangle = \hat{U}_{SM} |\psi\rangle_S |\theta\rangle_M, \quad \hat{U}_{SM} = \mathcal{T} \exp\left(-\frac{i}{\hbar} \int_0^\tau \hat{H}_{SM}(t) dt\right). \quad (3.4.10)$$

When the interaction is sufficiently short and strong, the free subsystem Hamiltonians can be neglected relative to the interaction Hamiltonian during this interval. A general interaction correlates the system and probe. We then make a projective measurement on the *probe* in an orthonormal basis $\{|r_j\rangle_M\}$.

Define the induced operator on the system Hilbert space by

$$\hat{M}_j = {}_M \langle r_j | \hat{U}_{SM} | \theta \rangle_M. \quad (3.4.11)$$

Projecting the composite state onto $|r_j\rangle_M$ gives the unnormalized system state $\hat{M}_j |\psi\rangle_S$. Its squared norm is the probability of outcome j :

$$p_j = \langle \psi | \hat{M}_j^\dagger \hat{M}_j | \psi \rangle, \quad |\psi_j\rangle = \frac{\hat{M}_j |\psi\rangle}{\sqrt{p_j}}. \quad (3.4.12)$$

Completeness of the probe readout basis and unitarity of $\hat{U}_{SM}$ imply

$$\sum_j \hat{M}_j^\dagger \hat{M}_j = {}_M \langle \theta | \hat{U}_{SM}^\dagger (\mathbf{1}_S \otimes \sum_j |r_j\rangle_M \langle r_j|) \hat{U}_{SM} | \theta \rangle_M = \mathbf{1}_S. \quad (3.4.13)$$

The system's state generally changes even if no direct projective measurement was performed on it. This is *measurement backaction*, illustrated in Fig. 3.1. For an ideal rank-one probe readout conditioned on j , the post-readout pure composite state is the product $|\psi_j\rangle_S \otimes |r_j\rangle_M$, whereas the pre-readout state may have been entangled.

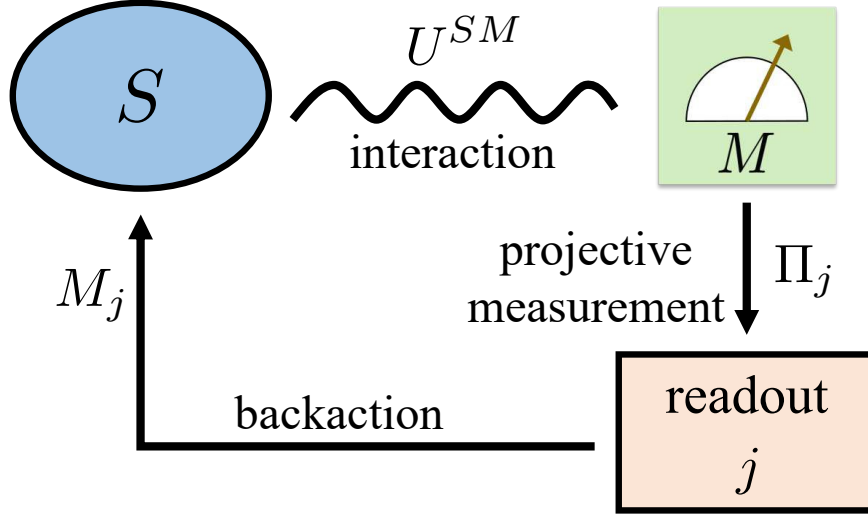


Figure 3.1: Indirect measurement as an interaction between a system S and a probe M , followed by a projective meter readout. The observed result j determines the induced measurement operator and the backaction on S . Reproduced from the instructor's doctoral dissertation, Fig. 2.1.

4) POVMs and quantum instruments

The positive operators

$$\boxed{\hat{E}_j = \hat{M}_j^\dagger \hat{M}_j, \quad \hat{E}_j \geq 0, \quad \sum_j \hat{E}_j = \mathbf{1}_S} \quad (3.4.14)$$

are called *effects*, and their complete collection forms a positive operator-valued measure (POVM). They determine outcome probabilities, $p_j = \text{tr}(\hat{\rho}_S \hat{E}_j)$. A projective measurement is the special case $\hat{M}_j = \hat{E}_j = \hat{Q}_j$.

For a mixed probe state $\hat{\rho}_M = \sum_k q_k |\theta_k\rangle \langle \theta_k|$, the same recorded result j may arise from several unobserved probe components. Define the Kraus operators

$$\hat{M}_{jk} = \sqrt{q_k} {}_M \langle r_j | \hat{U}_{SM} | \theta_k \rangle_M. \quad (3.4.15)$$

Now

$$\boxed{\begin{aligned} \hat{E}_j &= \sum_k \hat{M}_{jk}^\dagger \hat{M}_{jk}, \\ p_j &= \text{tr}(\hat{\rho}_S \hat{E}_j), \\ \hat{\rho}_{S|j} &= \frac{\sum_k \hat{M}_{jk} \hat{\rho}_S \hat{M}_{jk}^\dagger}{p_j}. \end{aligned}} \quad (3.4.16)$$

The *nonselective* state after the measurement is

$$\hat{\rho}'_S = \sum_{jk} \hat{M}_{jk} \hat{\rho}_S \hat{M}_{jk}^\dagger, \quad \text{tr } \hat{\rho}'_S = 1. \quad (3.4.17)$$

The individual outcome maps $\mathcal{I}_j(\hat{\rho}) = \sum_k \hat{M}_{jk} \hat{\rho} \hat{M}_{jk}^\dagger$ comprise a quantum *instrument*; their sum is completely positive and trace-preserving. An effect $\hat{E}_j$ specifies a probability but *not, by itself*, the conditional state: different instruments can have identical POVM effects.

Interactive investigation IC-03-05. Open POVMs and quantum instruments. After deriving effects and Kraus maps, compare distinct measurement instruments with the same POVM effects. Verify equal outcome probabilities and examine how their conditional and nonselective postmeasurement states differ.

5) A qubit interacting with a qubit meter

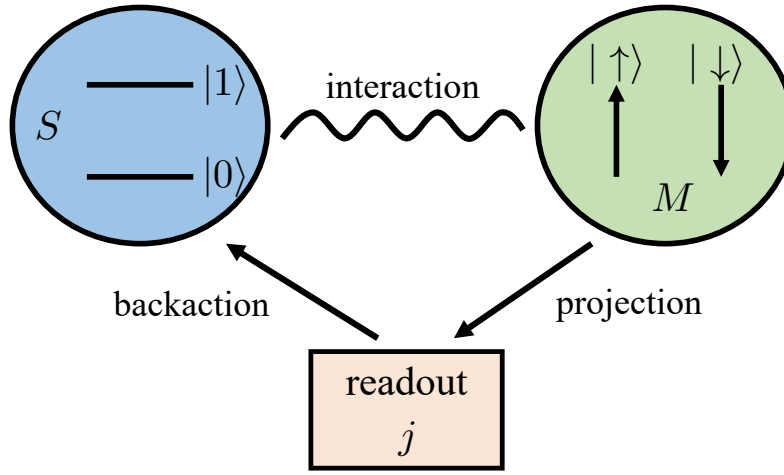


Figure 3.2: A concrete two-level system S and qubit probe M . Interaction correlates their states; projective probe readout produces information about S and a conditional state change. Reproduced from the instructor's doctoral dissertation, Fig. 2.2.

Figure 3.2 illustrates the two-level system and probe used below. As an explicit example, prepare a system qubit and a two-level meter as

$$|\psi\rangle_S = a_0 |0\rangle + a_1 |1\rangle, \quad |\theta\rangle_M = \frac{|\downarrow\rangle + |\uparrow\rangle}{\sqrt{2}}, \quad (3.4.18)$$

with $|a_0|^2 + |a_1|^2 = 1$. To allow the meter preparation to vary, first write $|\theta\rangle_M = b_0 |\downarrow\rangle + b_1 |\uparrow\rangle$ with $|b_0|^2 + |b_1|^2 = 1$. We will subsequently choose $b_0 = b_1 = 1/\sqrt{2}$ to obtain a particularly transparent measurement. The handout uses the interaction Hamiltonian

$$\hat{H}_{\text{int}} = ig(|1\rangle\langle 1| - |0\rangle\langle 0|) \otimes (|\uparrow\rangle\langle \downarrow| - |\downarrow\rangle\langle \uparrow|), \quad (3.4.19)$$

which is Hermitian because the operator in the second factor is anti-Hermitian. Introduce $\sigma_y^{(M)} = i(|\uparrow\rangle\langle \downarrow| - |\downarrow\rangle\langle \uparrow|)$ and $\theta = g\tau/\hbar$. Then

$$\hat{U}_{SM} = e^{-i\theta(|1\rangle\langle 1| - |0\rangle\langle 0|) \otimes \sigma_y^{(M)}}. \quad (3.4.20)$$

For a general meter preparation b_0, b_1 , direct evaluation of this unitary and projection of the meter onto $|\downarrow\rangle$ or $|\uparrow\rangle$ yield

$$\begin{aligned}\hat{M}_\downarrow &= (b_0 \cos \theta + b_1 \sin \theta) |0\rangle \langle 0| + (b_0 \cos \theta - b_1 \sin \theta) |1\rangle \langle 1|, \\ \hat{M}_\uparrow &= (b_1 \cos \theta - b_0 \sin \theta) |0\rangle \langle 0| + (b_1 \cos \theta + b_0 \sin \theta) |1\rangle \langle 1|.\end{aligned}\quad (3.4.21)$$

These two operators satisfy the completeness relation for any normalized b_0, b_1 . Specialize now to $b_0 = b_1 = 1/\sqrt{2}$. For $|m_\pm\rangle = (|\downarrow\rangle \pm |\uparrow\rangle)/\sqrt{2}$, the state after the interaction is

$$\begin{aligned}|\Psi_{SM}(\tau)\rangle &= a_0 |0\rangle (\cos \theta |m_+\rangle + \sin \theta |m_-\rangle) \\ &\quad + a_1 |1\rangle (\cos \theta |m_+\rangle - \sin \theta |m_-\rangle).\end{aligned}\quad (3.4.22)$$

A measurement of the meter in the $\{|\downarrow\rangle, |\uparrow\rangle\}$ basis induces the system operators

$$\begin{aligned}\hat{M}_\downarrow &= \frac{1}{\sqrt{2}} [(\cos \theta + \sin \theta) |0\rangle \langle 0| + (\cos \theta - \sin \theta) |1\rangle \langle 1|], \\ \hat{M}_\uparrow &= \frac{1}{\sqrt{2}} [(\cos \theta - \sin \theta) |0\rangle \langle 0| + (\cos \theta + \sin \theta) |1\rangle \langle 1|].\end{aligned}\quad (3.4.23)$$

Set $\eta = \sin(2\theta)$. The two POVM effects are

$$\hat{E}_\downarrow = \frac{1}{2}[(1 + \eta) |0\rangle \langle 0| + (1 - \eta) |1\rangle \langle 1|], \quad \hat{E}_\uparrow = \mathbf{1} - \hat{E}_\downarrow. \quad (3.4.24)$$

They are positive and sum to the identity for $|\eta| \leq 1$. When $\theta = \pi/4$, the measurement is projective:

$$\hat{M}_\downarrow = |0\rangle \langle 0|, \quad \hat{M}_\uparrow = |1\rangle \langle 1|, \quad |\Psi_{SM}\rangle = a_0 |0\rangle |\downarrow\rangle + a_1 |1\rangle |\uparrow\rangle. \quad (3.4.25)$$

The outcomes identify the computational-basis state, with probabilities $|a_0|^2$ and $|a_1|^2$. At $\theta = 3\pi/4$ the two projector outcomes are exchanged (up to an irrelevant phase): the meter result $\downarrow$ selects system state $|1\rangle$, and $\uparrow$ selects $|0\rangle$. When $\theta = 0$, both effects are $\mathbf{1}/2$: the result contains no information about the system and the conditional state is unchanged. Intermediate $0 < \theta < \pi/4$ implements a tunable *unsharp* measurement, whose backaction and information gain vary continuously with the interaction strength and time. The result follows entirely from a unitary system–probe interaction and an ordinary projective readout of the meter.

3.5 Eigenvalues and Eigenstates of Angular Momentum

We study the eigenvalues and eigenstates of a general angular momentum. The general definition of angular momentum $\hat{\mathbf{J}}$ is its commutation relation,

$$[\hat{J}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{J}_k. \quad (3.5.1)$$

A mechanical quantity satisfying this relation is called an angular momentum. Orbital angular momentum $\hat{\mathbf{L}}$ and spin angular momentum are special cases. We employ an algebraic method, analogous to that used for the harmonic oscillator, to find the eigenvalues of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$.

1) Definition of angular momentum squared

Define the operator describing the magnitude of angular momentum:

$$\hat{\mathbf{J}}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \hat{J}_i \hat{J}_i. \quad (3.5.2)$$

It commutes with angular momentum in any direction:

$$\begin{aligned}[\hat{\mathbf{J}}^2, \hat{J}_j] &= \hat{J}_i [\hat{J}_i, \hat{J}_j] + [\hat{J}_i, \hat{J}_j] \hat{J}_i \\ &= i\hbar \epsilon_{ijk} \hat{J}_i \hat{J}_k + i\hbar \epsilon_{ijk} \hat{J}_k \hat{J}_i \\ &= i\hbar \epsilon_{ijk} \hat{J}_i \hat{J}_k + i\hbar \epsilon_{kji} \hat{J}_i \hat{J}_k = 0.\end{aligned}\quad (3.5.3)$$

Here we used the antisymmetry of the Levi-Civita tensor, $\epsilon_{ijk} = -\epsilon_{kji}$.

Because $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ commute, they possess common eigenstates. Since $\hat{\mathbf{J}}^2$ is not determined by $\hat{J}_z$ alone, at least two quantum numbers are required. Denote a common eigenstate by $|\lambda, m\rangle$; quantum numbers are conventionally dimensionless. Taking the dimensions of the operators into account,

$$\hat{\mathbf{J}}^2 |\lambda, m\rangle = \lambda \hbar^2 |\lambda, m\rangle, \quad \hat{J}_z |\lambda, m\rangle = m \hbar |\lambda, m\rangle. \quad (3.5.4)$$

The question is: $\lambda = ?$, $m = ?$

2) Construct a new set of operators

Replace $\hat{J}_x, \hat{J}_y, \hat{J}_z$ by

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y, \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y = \hat{J}_+^\dagger, \quad \hat{J}_z. \quad (3.5.5)$$

It follows that

$$\hat{J}_+ \hat{J}_- = \hat{\mathbf{J}}^2 - \hat{J}_z^2 + \hbar \hat{J}_z, \quad (3.5.6)$$

$$\hat{J}_- \hat{J}_+ = \hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar \hat{J}_z, \quad (3.5.7)$$

$$\hat{\mathbf{J}}^2 = \frac{1}{2}(\hat{J}_+ \hat{J}_- + \hat{J}_- \hat{J}_+) + \hat{J}_z^2. \quad (3.5.8)$$

The new commutation relations are

$$[\hat{J}_+, \hat{J}_-] = 2\hbar \hat{J}_z, \quad [\hat{J}_-, \hat{J}_z] = \hbar \hat{J}_-, \quad [\hat{J}_+, \hat{J}_z] = -\hbar \hat{J}_+, \quad (3.5.9)$$

$$[\hat{\mathbf{J}}^2, \hat{J}_+] = [\hat{\mathbf{J}}^2, \hat{J}_-] = [\hat{\mathbf{J}}^2, \hat{J}_z] = 0. \quad (3.5.10)$$

3) Positivity bound

Take an expectation value in $|\lambda, m\rangle$:

$$\begin{aligned} \langle \lambda, m | \hat{\mathbf{J}}^2 - \hat{J}_z^2 | \lambda, m \rangle &= \frac{1}{2} \langle \lambda, m | \hat{J}_+ \hat{J}_- + \hat{J}_- \hat{J}_+ | \lambda, m \rangle, \\ (\lambda - m^2) \hbar^2 \langle \lambda, m | \lambda, m \rangle &= \frac{1}{2} \langle \lambda, m | \hat{J}_+ \hat{J}_- | \lambda, m \rangle + \frac{1}{2} \langle \lambda, m | \hat{J}_- \hat{J}_+ | \lambda, m \rangle \geq 0. \end{aligned} \quad (3.5.11)$$

Thus $\lambda \geq m^2$.

4) Raising and lowering operators

Using their commutators with $\hat{\mathbf{J}}^2$ and $\hat{J}_z$,

$$\begin{cases} \hat{\mathbf{J}}^2(\hat{J}_+ |\lambda, m\rangle) = \lambda \hbar^2 (\hat{J}_+ |\lambda, m\rangle), \\ \hat{J}_z(\hat{J}_+ |\lambda, m\rangle) = (\hbar \hat{J}_+ + \hat{J}_+ \hat{J}_z) |\lambda, m\rangle = (m+1) \hbar (\hat{J}_+ |\lambda, m\rangle), \end{cases} \quad (3.5.12)$$

and

$$\begin{cases} \hat{\mathbf{J}}^2(\hat{J}_- |\lambda, m\rangle) = \lambda \hbar^2 (\hat{J}_- |\lambda, m\rangle), \\ \hat{J}_z(\hat{J}_- |\lambda, m\rangle) = (-\hbar \hat{J}_- + \hat{J}_- \hat{J}_z) |\lambda, m\rangle = (m-1) \hbar (\hat{J}_- |\lambda, m\rangle). \end{cases} \quad (3.5.13)$$

Thus if $|\lambda, m\rangle$ is a common eigenstate of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$, so are $\hat{J}_+ |\lambda, m\rangle$ and $\hat{J}_- |\lambda, m\rangle$, when nonzero. Their eigenvalues form the sequence

Common eigenstate	$\hat{J}_z$ eigenvalue	$\hat{\mathbf{J}}^2$ eigenvalue
$\vdots$	$\vdots$	$\vdots$
$(\hat{J}_+)^2 \lambda, m\rangle$	$(m+2)\hbar$	$\lambda\hbar^2$
$\hat{J}_+ \lambda, m\rangle$	$(m+1)\hbar$	$\lambda\hbar^2$
$ \lambda, m\rangle$	$m\hbar$	$\lambda\hbar^2$
$\hat{J}_- \lambda, m\rangle$	$(m-1)\hbar$	$\lambda\hbar^2$
$(\hat{J}_-)^2 \lambda, m\rangle$	$(m-2)\hbar$	$\lambda\hbar^2$
$\vdots$	$\vdots$	$\vdots$

(3.5.14)

Therefore $\hat{J}_-$ is called a lowering operator and $\hat{J}_+$ a raising operator. Since the bound $\lambda \geq m^2$ limits m , there must be largest and smallest values of m , denoted j and j' . The values of m form the sequence $j', j'+1, \dots, j-1, j$, and $\lambda \geq j^2, \lambda \geq j'^2$.

5) Determine the upper and lower bounds

For the largest value j ,

$$\hat{J}_+ |\lambda, j\rangle = 0, \quad (3.5.15)$$

otherwise $j\hbar$ would not be the greatest eigenvalue of $\hat{J}_z$. Using

$$0 = \hat{J}_- \hat{J}_+ |\lambda, j\rangle = (\hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar \hat{J}_z) |\lambda, j\rangle = (\lambda - j^2 - j)\hbar^2 |\lambda, j\rangle, \quad (3.5.16)$$

we obtain $\lambda = j(j+1)$.

For the smallest value j' ,

$$\hat{J}_- |\lambda, j'\rangle = 0, \quad (3.5.17)$$

otherwise $j'\hbar$ would not be the smallest eigenvalue. Similarly,

$$0 = \hat{J}_+ \hat{J}_- |\lambda, j'\rangle = (\lambda - j'^2 + j')\hbar^2 |\lambda, j'\rangle, \quad (3.5.18)$$

so $\lambda = j'(j'-1)$. Since $j(j+1) = j'(j'-1)$,

$$j' = \begin{cases} j+1, \\ -j. \end{cases} \quad (3.5.19)$$

The choice $j' = j+1 > j$ contradicts the fact that j is the maximum and j' the minimum. Thus $j' = -j$, and

$$\hat{\mathbf{J}}^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle, \quad \hat{J}_z |j, m\rangle = m\hbar |j, m\rangle, \quad (3.5.20)$$

where $m = -j, -j+1, \dots, j-1, j$ comprises $2j+1$ values. We relabel $|\lambda, m\rangle$ as $|j, m\rangle$. The remaining question is: which values of j are possible?

6) Allowed angular-momentum quantum numbers

Applying $\hat{J}_-$ to $|j, j\rangle$ exactly $2j$ times produces $|j, -j\rangle$; applying $\hat{J}_+$ to $|j, -j\rangle$ exactly $2j$ times produces $|j, j\rangle$. Therefore $2j$ is zero or a positive integer, so j may be zero, a positive integer, or a positive half-integer.

In summary,

$$\begin{aligned} \hat{\mathbf{J}}^2 |j, m\rangle &= j(j+1)\hbar^2 |j, m\rangle, & j &= 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \\ \hat{J}_z |j, m\rangle &= m\hbar |j, m\rangle, & m &= -j, -j+1, \dots, j \quad (2j+1 \text{ values}). \end{aligned}$$

(3.5.21)

For j zero or a positive integer, $\hat{\mathbf{J}}$ can be explained by orbital angular momentum $\hat{\mathbf{L}}$. What is the physical meaning of $\hat{\mathbf{J}}$ when j is a positive half-integer? The angular-momentum commutation relation itself establishes the possibility of a new type of angular momentum: spin angular momentum.

7) Matrix elements of raising and lowering operators

Because the eigenvalue pair (j, m) labels a unique eigenstate in an irreducible fixed- j multiplet,

$$\hat{J}_+ |j, m\rangle = a_{jm} |j, m+1\rangle, \quad \hat{J}_- |j, m\rangle = b_{jm} |j, m-1\rangle. \quad (3.5.22)$$

Taking Hermitian conjugates,

$$\langle j, m | \hat{J}_- = a_{jm}^* \langle j, m+1 |, \quad \langle j, m | \hat{J}_+ = b_{jm}^* \langle j, m-1 |. \quad (3.5.23)$$

Consequently,

$$\begin{aligned} |a_{jm}|^2 &= \langle j, m | \hat{J}_- \hat{J}_+ |j, m\rangle = \langle j, m | \hat{\mathbf{J}}^2 - \hat{J}_z^2 - \hbar \hat{J}_z |j, m\rangle \\ &= [j(j+1) - m^2 - m] \hbar^2. \end{aligned} \quad (3.5.24)$$

Choose the phase convention

$$a_{jm} = \hbar \sqrt{j(j+1) - m(m+1)} = \hbar \sqrt{(j-m)(j+m+1)}. \quad (3.5.25)$$

Similarly,

$$b_{jm} = \hbar \sqrt{(j+m)(j-m+1)}. \quad (3.5.26)$$

Thus

$$\begin{aligned} \hat{J}_+ |j, m\rangle &= \hbar \sqrt{(j-m)(j+m+1)} |j, m+1\rangle, \\ \hat{J}_- |j, m\rangle &= \hbar \sqrt{(j+m)(j-m+1)} |j, m-1\rangle. \end{aligned} \quad (3.5.27)$$

Because

$$\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-), \quad \hat{J}_y = \frac{1}{2i}(\hat{J}_+ - \hat{J}_-), \quad (3.5.28)$$

we obtain

$$\begin{aligned} \hat{J}_x |j, m\rangle &= \frac{\hbar}{2} \sqrt{(j-m)(j+m+1)} |j, m+1\rangle + \frac{\hbar}{2} \sqrt{(j+m)(j-m+1)} |j, m-1\rangle, \\ \hat{J}_y |j, m\rangle &= \frac{\hbar}{2i} \sqrt{(j-m)(j+m+1)} |j, m+1\rangle - \frac{\hbar}{2i} \sqrt{(j+m)(j-m+1)} |j, m-1\rangle. \end{aligned} \quad (3.5.29)$$

8) Common representation of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$

The basis consists of $|j, m\rangle$. For fixed j , the eigenvalue of $\hat{\mathbf{J}}^2$ is fixed, and the eigenstates of $\hat{J}_z$ span a subspace of dimension $D = 2j + 1$. The matrix elements are

$$\langle j, n | \hat{\mathbf{J}}^2 |j, m\rangle = j(j+1) \hbar^2 \delta_{nm}, \quad \langle j, n | \hat{J}_z |j, m\rangle = m \hbar \delta_{nm}, \quad (3.5.30)$$

$$\begin{aligned} \langle j, n | \hat{J}_x |j, m\rangle &= \frac{\hbar}{2} \sqrt{(j-m)(j+m+1)} \delta_{n, m+1} + \frac{\hbar}{2} \sqrt{(j+m)(j-m+1)} \delta_{n, m-1}, \\ \langle j, n | \hat{J}_y |j, m\rangle &= \frac{\hbar}{2i} \sqrt{(j-m)(j+m+1)} \delta_{n, m+1} - \frac{\hbar}{2i} \sqrt{(j+m)(j-m+1)} \delta_{n, m-1}. \end{aligned} \quad (3.5.31)$$

These are the matrix representations of $\hat{J}_x, \hat{J}_y, \hat{J}_z$ in the common eigenbasis.

Example 1: $j = 1/2$ (electron spin). In this case

$$\hat{\mathbf{J}}^2 = \frac{3}{4} \hbar^2, \quad m = \pm \frac{1}{2}, \quad D = 2. \quad (3.5.32)$$

Choose the first basis state $|1/2, 1/2\rangle$ and the second $|1/2, -1/2\rangle$. Then

$$J_z = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} = \frac{\hbar}{2}\sigma_z, \quad J_x = \begin{pmatrix} 0 & \hbar/2 \\ \hbar/2 & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_x, \quad (3.5.33)$$

$$J_y = \begin{pmatrix} 0 & -i\hbar/2 \\ i\hbar/2 & 0 \end{pmatrix} = \frac{\hbar}{2}\sigma_y. \quad (3.5.34)$$

If an eigenstate of $\hat{J}_z$ has components $(a, b)^\top$, its eigenvalue equation is

$$\begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = m\hbar \begin{pmatrix} a \\ b \end{pmatrix}. \quad (3.5.35)$$

After normalization, the eigenstate is $(1, 0)^\top$ for $m = 1/2$ and $(0, 1)^\top$ for $m = -1/2$.

Example 2: $j = 1$ (photon spin). Here $\hat{\mathbf{J}}^2 = 2\hbar^2$, $m = -1, 0, 1$, and $D = 3$. Choose the ordered basis $|1, 1\rangle, |1, 0\rangle, |1, -1\rangle$:

$$J_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (3.5.36)$$

The eigenstates of J_z have components

$$m = 1 : \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad m = 0 : \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad m = -1 : \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.5.37)$$

9) Matrix representation of rotation operators

The rotation operator is

$$\hat{D}_n(\varphi) = \exp\left(-\frac{i}{\hbar} \hat{\mathbf{J}} \cdot \mathbf{e}_n \varphi\right). \quad (3.5.38)$$

Since

$$[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}] = 0, \quad [\hat{\mathbf{J}}^2, \hat{D}_n(\varphi)] = 0, \quad (3.5.39)$$

$\hat{D}_n(\varphi)$ and $\hat{\mathbf{J}}^2$ have common eigenstates. In the $\hat{\mathbf{J}}^2$ representation, the rotation matrix is block diagonal; in the common basis of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$,

$$D_{mm'}^{jj'} = \langle j, m | \hat{D}_n(\varphi) | j', m' \rangle = 0 \quad (j \neq j'). \quad (3.5.40)$$

But $[\hat{J}_z, \hat{D}_n(\varphi)] \neq 0$ in general, so within a fixed- j subspace D has off-diagonal entries. Since any rotation can be written in terms of Euler angles α, β, γ , the matrix elements in a fixed- j subspace obey

$$\begin{aligned} D_{mm'}^j &= \langle j, m | \hat{D}_n(\varphi) | j, m' \rangle = \langle j, m | e^{-i\hat{\mathbf{J}} \cdot \mathbf{e}_n \varphi / \hbar} | j, m' \rangle \\ &= \langle j, m | e^{-i\hat{J}_z \alpha / \hbar} e^{-i\hat{J}_y \beta / \hbar} e^{-i\hat{J}_z \gamma / \hbar} | j, m' \rangle \\ &= e^{-i(m\alpha + m'\gamma)} \langle j, m | e^{-i\hat{J}_y \beta / \hbar} | j, m' \rangle \\ &= e^{-i(m\alpha + m'\gamma)} d_{mm'}^j(\beta). \end{aligned} \quad (3.5.41)$$

Thus D is block diagonal in j , with nonzero matrix elements generally occurring only within the same j block.

For $j = 1/2$, we previously found

$$d^{1/2}(\beta) = \begin{pmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{pmatrix}. \quad (3.5.42)$$

For $j = 1$, in the common representation,

$$J_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \quad (3.5.43)$$

One can verify

$$e^{-iJ_y\beta/\hbar} = \mathbf{1} - \frac{i}{\hbar} \sin \beta J_y + \left(\frac{i}{\hbar}\right)^2 (1 - \cos \beta) J_y^2, \quad (3.5.44)$$

which yields

$$d^1(\beta) = \frac{1}{2} \begin{pmatrix} 1 + \cos \beta & -\sqrt{2} \sin \beta & 1 - \cos \beta \\ \sqrt{2} \sin \beta & 2 \cos \beta & -\sqrt{2} \sin \beta \\ 1 - \cos \beta & \sqrt{2} \sin \beta & 1 + \cos \beta \end{pmatrix}. \quad (3.5.45)$$

3.6 Orbital Angular Momentum

1) Eigenvalues and eigenstates

The classical orbital angular momentum and its quantum operator are

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \longrightarrow \hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}, \quad \hat{L}_i = \epsilon_{ijk} \hat{x}_j \hat{p}_k \quad (i, j, k = 1, 2, 3). \quad (3.6.1)$$

From the fundamental commutation relations

$$[\hat{x}_i, \hat{x}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}, \quad (3.6.2)$$

one can show that

$$[\hat{L}_i, \hat{L}_j] = i\hbar \epsilon_{ijk} \hat{L}_k. \quad (3.6.3)$$

Thus the orbital angular-momentum operators satisfy the general definition of angular momentum.

Because this is orbital angular momentum, we calculate eigenvalues and eigenstates in the coordinate representation:

$$\hat{\mathbf{r}} = \mathbf{r}, \quad \hat{\mathbf{p}} = -i\hbar \nabla, \quad (3.6.4)$$

$$\hat{\mathbf{L}} = -i\hbar \mathbf{r} \times \nabla, \quad \hat{L}_i = -i\hbar \epsilon_{ijk} x_j \frac{\partial}{\partial x_k}. \quad (3.6.5)$$

It is convenient to consider rotations in spherical coordinates, $(x, y, z) \rightarrow (r, \theta, \varphi)$. The operators are

$$\begin{aligned} \hat{L}_x &= i\hbar \left(\sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right), \\ \hat{L}_y &= -i\hbar \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right), \end{aligned} \quad (3.6.6)$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi},$$

$$\hat{\mathbf{L}}^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right]. \quad (3.6.7)$$

This shows that $\hat{\mathbf{L}}$ depends only on the angles θ and φ .

First consider the eigenvalue equation of $\hat{L}_z$. As $\hat{L}_z$ does not depend on θ ,

$$\hat{L}_z \Phi(\varphi) = L_z \Phi(\varphi), \quad -i\hbar \frac{d\Phi}{d\varphi} = L_z \Phi(\varphi). \quad (3.6.8)$$

The solution is

$$\Phi(\varphi) = A \exp(iL_z \varphi / \hbar). \quad (3.6.9)$$

The single-valuedness condition

$$\Phi(\varphi) = \Phi(\varphi + 2\pi) \quad (3.6.10)$$

implies $L_z = m\hbar$, where $m = 0, \pm 1, \dots, \pm\infty$, and $\Phi(\varphi) = Ae^{im\varphi}$. Normalize:

$$\int_0^{2\pi} |\Phi(\varphi)|^2 d\varphi = 1 \implies A = \frac{1}{\sqrt{2\pi}}. \quad (3.6.11)$$

Thus

$$\boxed{\Phi(\varphi) = \frac{e^{im\varphi}}{\sqrt{2\pi}}.} \quad (3.6.12)$$

Next consider the eigenvalue equation of $\hat{\mathbf{L}}^2$:

$$\hat{\mathbf{L}}^2 Y(\theta, \varphi) = L^2 Y(\theta, \varphi). \quad (3.6.13)$$

This is

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \frac{L^2}{\hbar^2} \right] Y_{lm}(\theta, \varphi) = 0. \quad (3.6.14)$$

Using separation of variables and the single-valuedness and finiteness conditions, one finds

$$L^2 = l(l+1)\hbar^2, \quad l = 0, 1, 2, \dots \quad (3.6.15)$$

The eigenstates are spherical harmonics

$$Y_{lm}(\theta, \varphi) = N_{lm} P_l^m(\cos \theta) e^{im\varphi}, \quad (3.6.16)$$

where $P_l^m(x)$ is the associated Legendre polynomial. The magnetic quantum number is

$$m = 0, \pm 1, \pm 2, \dots, \pm l, \quad (3.6.17)$$

with $2l+1$ values. The eigenvalue L^2 depends only on l whereas the eigenstates also depend on m , so the degeneracy is $g = 2l+1$.

The normalization condition is

$$\int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta |Y_{lm}(\theta, \varphi)|^2 = 1, \quad (3.6.18)$$

which gives the normalization factor

$$N_{lm} = \sqrt{\frac{(l-m)!(2l+1)}{(l+m)!4\pi}}. \quad (3.6.19)$$

Because the φ -dependent factor in Y_{lm} is an eigenstate of $\hat{L}_z$, the spherical harmonics are common eigenstates of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$:

$$\boxed{\begin{aligned} \hat{\mathbf{L}}^2 Y_{lm}(\theta, \varphi) &= l(l+1)\hbar^2 Y_{lm}(\theta, \varphi), \\ \hat{L}_z Y_{lm}(\theta, \varphi) &= m\hbar Y_{lm}(\theta, \varphi), \end{aligned}} \quad (3.6.20)$$

where $l = 0, 1, 2, \dots$ and $m = 0, \pm 1, \dots, \pm l$. The component of angular momentum is constrained by its total magnitude, so m cannot extend to $\pm\infty$ for fixed l .

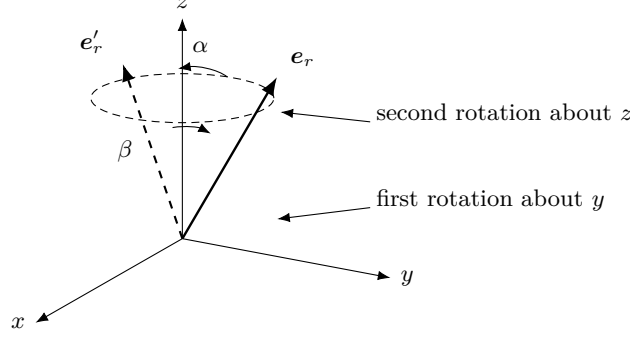
This shows that orbital angular momentum realizes the general angular-momentum algebra in coordinate space. It is only one form of angular momentum: l is restricted to zero and positive integers and cannot be a positive half-integer.

2) Spatial rotation operator

The orbital-angular-momentum eigenstate is the projection of $|l, m\rangle$ onto a radial position basis state $|\mathbf{e}_r(\theta, \varphi)\rangle$:

$$Y_{lm}(\theta, \varphi) = \langle \mathbf{e}_r(\theta, \varphi) | l, m \rangle. \quad (3.6.21)$$

The unit vector $\mathbf{e}_r(\theta, \varphi)$ can be obtained from $\mathbf{e}_z(0, \varphi)$ by two rotations, first by θ about the y axis and then by φ about the z axis:



$$|\mathbf{e}_r(\theta, \varphi)\rangle = \hat{D}_z(\varphi) \hat{D}_y(\theta) |\mathbf{e}_z(0, \varphi)\rangle. \quad (3.6.22)$$

Insert the complete basis:

$$|\mathbf{e}_r(\theta, \varphi)\rangle = \sum_{m'} \hat{D}_z(\varphi) \hat{D}_y(\theta) |l, m'\rangle \langle l, m' | \mathbf{e}_z(0, \varphi)\rangle. \quad (3.6.23)$$

Hence

$$\begin{aligned} Y_{lm}^*(\theta, \varphi) &= \langle l, m | \mathbf{e}_r(\theta, \varphi) \rangle \\ &= \sum_{m'} \langle l, m | \hat{D}_z(\varphi) \hat{D}_y(\theta) | l, m' \rangle \langle l, m' | \mathbf{e}_z(0, \varphi) \rangle. \end{aligned} \quad (3.6.24)$$

Since

$$\langle l, m' | \mathbf{e}_z(0, \varphi) \rangle = Y_{lm'}^*(0, \varphi) = \sqrt{\frac{2l+1}{4\pi}} \delta_{m'0}, \quad (3.6.25)$$

and

$$\langle l, m | \hat{D}_z(\varphi) = \langle l, m | e^{-i\hat{L}_z\varphi/\hbar} = e^{-im\varphi} \langle l, m |, \quad (3.6.26)$$

we obtain

$$\boxed{Y_{lm}^*(\theta, \varphi) = e^{-im\varphi} \sqrt{\frac{2l+1}{4\pi}} d_{m0}^l(\theta), \quad d_{mm'}^l(\theta) = \langle l, m | \hat{D}_y(\theta) | l, m' \rangle.} \quad (3.6.27)$$

A spherical harmonic in any direction can thus be obtained by rotating a spherical harmonic along the Z direction.

3.7 Coupling of Two Angular Momenta

Consider the addition of two arbitrary independent angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$:

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2. \quad (3.7.1)$$

They satisfy

$$[\hat{J}_{1i}, \hat{J}_{1j}] = i\hbar\epsilon_{ijk}\hat{J}_{1k}, \quad [\hat{J}_{2i}, \hat{J}_{2j}] = i\hbar\epsilon_{ijk}\hat{J}_{2k}, \quad [\hat{J}_{1i}, \hat{J}_{2j}] = 0. \quad (3.7.2)$$

The respective eigenvalues are

$$J_1^2 = j_1(j_1 + 1)\hbar^2, \quad J_{1z} = m_1\hbar, \quad m_1 = -j_1, \dots, j_1, \quad (3.7.3)$$

$$J_2^2 = j_2(j_2 + 1)\hbar^2, \quad J_{2z} = m_2\hbar, \quad m_2 = -j_2, \dots, j_2. \quad (3.7.4)$$

It is easy to prove

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k. \quad (3.7.5)$$

Thus $\hat{\mathbf{J}}$ is again an angular-momentum operator, called the total angular momentum, with eigenvalues

$$J^2 = j(j + 1)\hbar^2, \quad J_z = m\hbar, \quad m = -j, \dots, j. \quad (3.7.6)$$

The problem is to find the relation between $|j, m\rangle$ and $|j_1, m_1\rangle |j_2, m_2\rangle$.

1) Two representations

Uncoupled representation. Since the two angular momenta commute with one another, the four operators $\hat{\mathbf{J}}_1^2$, $\hat{J}_{1z}$, $\hat{\mathbf{J}}_2^2$, $\hat{J}_{2z}$ have common eigenvectors $|j_1 m_1 j_2 m_2\rangle$, which form the uncoupled representation:

$$\begin{aligned} \hat{\mathbf{J}}_1^2 |j_1 m_1 j_2 m_2\rangle &= j_1(j_1 + 1)\hbar^2 |j_1 m_1 j_2 m_2\rangle, \\ \hat{J}_{1z} |j_1 m_1 j_2 m_2\rangle &= m_1\hbar |j_1 m_1 j_2 m_2\rangle, \\ \hat{\mathbf{J}}_2^2 |j_1 m_1 j_2 m_2\rangle &= j_2(j_2 + 1)\hbar^2 |j_1 m_1 j_2 m_2\rangle, \\ \hat{J}_{2z} |j_1 m_1 j_2 m_2\rangle &= m_2\hbar |j_1 m_1 j_2 m_2\rangle. \end{aligned} \quad (3.7.7)$$

But $[\hat{\mathbf{J}}^2, \hat{J}_{1z}] \neq 0$ and $[\hat{\mathbf{J}}^2, \hat{J}_{2z}] \neq 0$. Consequently, $\hat{\mathbf{J}}^2$ is not diagonal in this basis, making its eigenvalues inconvenient to obtain directly.

Coupled representation. Since

$$[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_1^2] = [\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_2^2] = [\hat{\mathbf{J}}^2, \hat{J}_z] = 0, \quad (3.7.8)$$

and the remaining relevant pairs also commute, the four operators $\hat{\mathbf{J}}_1^2$, $\hat{\mathbf{J}}_2^2$, $\hat{\mathbf{J}}^2$, and $\hat{J}_z$ have common eigenvectors $|j_1 j_2 j m\rangle$, which form the coupled representation:

$$\begin{aligned} \hat{\mathbf{J}}_1^2 |j_1 j_2 j m\rangle &= j_1(j_1 + 1)\hbar^2 |j_1 j_2 j m\rangle, \\ \hat{\mathbf{J}}_2^2 |j_1 j_2 j m\rangle &= j_2(j_2 + 1)\hbar^2 |j_1 j_2 j m\rangle, \\ \hat{\mathbf{J}}^2 |j_1 j_2 j m\rangle &= j(j + 1)\hbar^2 |j_1 j_2 j m\rangle, \\ \hat{J}_z |j_1 j_2 j m\rangle &= m\hbar |j_1 j_2 j m\rangle. \end{aligned} \quad (3.7.9)$$

Because the eigenvalues and eigenvectors of the four uncoupled operators are known, obtaining those of total angular momentum requires connecting the two representations by a change of basis.

From completeness of the uncoupled basis,

$$\sum_{m_1, m_2} |j_1 m_1 j_2 m_2\rangle \langle j_1 m_1 j_2 m_2| = \mathbf{1}, \quad (3.7.10)$$

the transformation is

$$|j_1 j_2 j m\rangle = \sum_{m_1, m_2} |j_1 m_1 j_2 m_2\rangle \langle j_1 m_1 j_2 m_2 | j_1 j_2 j m\rangle. \quad (3.7.11)$$

The overlaps $\langle j_1 m_1 j_2 m_2 | j_1 j_2 j m\rangle$ are the Clebsch–Gordan (CG) coefficients.

2) Eigenvalues (j, m) of the coupled representation

Apply $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$ to the expansion:

$$\begin{aligned} m\hbar |j_1 j_2 j m\rangle &= \sum_{m_1, m_2} \langle j_1 m_1 j_2 m_2 | j_1 j_2 j m \rangle (\hat{J}_{1z} + \hat{J}_{2z}) |j_1 m_1 j_2 m_2\rangle \\ &= \sum_{m_1, m_2} (m_1 + m_2)\hbar \langle j_1 m_1 j_2 m_2 | j_1 j_2 j m \rangle |j_1 m_1 j_2 m_2\rangle. \end{aligned} \quad (3.7.12)$$

Thus

$$\sum_{m_1, m_2} (m - m_1 - m_2)\hbar \langle j_1 m_1 j_2 m_2 | j_1 j_2 j m \rangle |j_1 m_1 j_2 m_2\rangle = 0. \quad (3.7.13)$$

The uncoupled basis vectors are linearly independent. Therefore for each term either the CG coefficient is zero or $m = m_1 + m_2$. We seek nonzero CG coefficients, and hence

$$\boxed{m = m_1 + m_2.} \quad (3.7.14)$$

Next consider the possible values of j . Let $j_{\min} \leq j \leq j_{\max}$. The maximum is

$$j_{\max} = m_{\max} = (m_1)_{\max} + (m_2)_{\max} = j_1 + j_2. \quad (3.7.15)$$

The dimension of the uncoupled representation is

$$D = (2j_1 + 1)(2j_2 + 1), \quad (3.7.16)$$

while the dimension of the coupled representation is

$$D = \sum_{j=j_{\min}}^{j_{\max}} (2j + 1) = (j_{\max} + 1)^2 - j_{\min}^2. \quad (3.7.17)$$

Equating the dimensions,

$$(2j_1 + 1)(2j_2 + 1) = (j_1 + j_2 + 1)^2 - j_{\min}^2, \quad (3.7.18)$$

gives $j_{\min}^2 = (j_1 - j_2)^2$ and $j_{\min} = |j_1 - j_2|$. Thus for fixed (j_1, j_2) ,

$$\boxed{\begin{aligned} J^2 &= j(j+1)\hbar^2, & j &= |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2, \\ J_z &= m\hbar, & m &= m_1 + m_2. \end{aligned}} \quad (3.7.19)$$

3) Eigenstates $|j_1 j_2 j m\rangle$ of the coupled representation

The central task is to calculate the CG coefficients. We do not undertake a general treatment; standard tables are available.

Example 1: spin and orbital angular momentum. Let $\hat{\mathbf{J}}_1 = \hat{\mathbf{L}}$, $\hat{\mathbf{J}}_2 = \hat{\mathbf{S}}$, and $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. The coupled basis is $|l, \frac{1}{2}; j, m\rangle$ and the uncoupled basis is $|l, m_l; \frac{1}{2}, m_s\rangle$. The transformation reads

$$|l, \frac{1}{2}; j, m\rangle = \sum_{m_l, m_s} C_{m_l m_s} |l, m_l; \frac{1}{2}, m_s\rangle = \sum_{m_l} [A_{m_l} |l, m_l; \frac{1}{2}, \frac{1}{2}\rangle + B_{m_l} |l, m_l; \frac{1}{2}, -\frac{1}{2}\rangle]. \quad (3.7.20)$$

Because $m = m_l + m_s$,

$$m_l = m - m_s = \begin{cases} m - \frac{1}{2}, & m_s = \frac{1}{2}, \\ m + \frac{1}{2}, & m_s = -\frac{1}{2}, \end{cases} \quad (3.7.21)$$

there are only two terms:

$$\left| l, \frac{1}{2}; j, m \right\rangle = A \left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + B \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle. \quad (3.7.22)$$

We determine the coefficients A, B .

In the spin subspace of the uncoupled representation,

$$\begin{aligned} \hat{\mathbf{J}}^2 &= \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \\ &= \begin{pmatrix} \hat{\mathbf{L}}^2 + \frac{3}{4}\hbar^2 + \hbar\hat{L}_z & \hbar\hat{L}_- \\ \hbar\hat{L}_+ & \hat{\mathbf{L}}^2 + \frac{3}{4}\hbar^2 - \hbar\hat{L}_z \end{pmatrix}. \end{aligned} \quad (3.7.23)$$

Here $\hat{L}_{\pm} = \hat{L}_x \pm i\hat{L}_y$ are the orbital raising and lowering operators. The uncoupled states are represented as

$$\left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle = \begin{pmatrix} |l, m - \frac{1}{2}\rangle \\ 0 \end{pmatrix}, \quad \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle = \begin{pmatrix} 0 \\ |l, m + \frac{1}{2}\rangle \end{pmatrix}. \quad (3.7.24)$$

Insert the expansion into the eigenvalue equation

$$\hat{\mathbf{J}}^2 \left| l, \frac{1}{2}; j, m \right\rangle = j(j+1)\hbar^2 \left| l, \frac{1}{2}; j, m \right\rangle. \quad (3.7.25)$$

It becomes

$$j(j+1)\hbar^2 (A \left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + B \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle) = (\hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) \times (A \left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + B \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle). \quad (3.7.26)$$

In the spin subspace this gives

$$\begin{cases} A \left(\hat{\mathbf{L}}^2 + \frac{3}{4}\hbar^2 + \hbar\hat{L}_z - j(j+1)\hbar^2 \right) \left| l, m - \frac{1}{2} \right\rangle + B\hbar\hat{L}_- \left| l, m + \frac{1}{2} \right\rangle = 0, \\ A\hbar\hat{L}_+ \left| l, m - \frac{1}{2} \right\rangle + B \left(\hat{\mathbf{L}}^2 + \frac{3}{4}\hbar^2 - \hbar\hat{L}_z - j(j+1)\hbar^2 \right) \left| l, m + \frac{1}{2} \right\rangle = 0. \end{cases} \quad (3.7.27)$$

Using the orbital-angular-momentum eigenvalues and ladder-operator coefficients yields

$$\begin{cases} \left[l(l+1) + \frac{3}{4} + \left(m - \frac{1}{2} \right) - j(j+1) \right] A + \sqrt{\left(l - m + \frac{1}{2} \right) \left(l + m + \frac{1}{2} \right)} B = 0, \\ \sqrt{\left(l + m + \frac{1}{2} \right) \left(l - m + \frac{1}{2} \right)} A + \left[l(l+1) + \frac{3}{4} - \left(m + \frac{1}{2} \right) - j(j+1) \right] B = 0. \end{cases} \quad (3.7.28)$$

These equations determine A, B . The allowed total angular momenta are $j = l - \frac{1}{2}$ and $j = l + \frac{1}{2}$ (where permitted).

For $j = l + \frac{1}{2}$, normalization yields

$$\left| l, \frac{1}{2}; j, m \right\rangle = \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} \left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle. \quad (3.7.29)$$

For $j = l - \frac{1}{2}$,

$$\left| l, \frac{1}{2}; j, m \right\rangle = -\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} \left| l, m - \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} \left| l, m + \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle. \quad (3.7.30)$$

Example 2: coupling two electron spins. Let

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_2, \quad s_1 = s_2 = \frac{1}{2}, \quad m_1, m_2 = \pm \frac{1}{2}, \quad s = 0, 1. \quad (3.7.31)$$

The uncoupled basis states are

$$\begin{aligned} & \left| \frac{1}{2}, \frac{1}{2} \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \quad \left| \frac{1}{2}, \frac{1}{2} \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle, \\ & \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \quad \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle. \end{aligned} \quad (3.7.32)$$

The coupled basis is $\left| \frac{1}{2}, \frac{1}{2}; s, m \right\rangle$ with $(s, m) = (1, 1), (1, 0), (1, -1), (0, 0)$. As in Example 1, coupled states can be expressed in the uncoupled basis.

For $s = 1$, the *spin triplet* is

$$\begin{aligned} \left| \frac{1}{2}, \frac{1}{2}; 1, 1 \right\rangle &= \left| \frac{1}{2}, \frac{1}{2} \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle, \\ \left| \frac{1}{2}, \frac{1}{2}; 1, 0 \right\rangle &= \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right), \\ \left| \frac{1}{2}, \frac{1}{2}; 1, -1 \right\rangle &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle. \end{aligned} \quad (3.7.33)$$

For $s = 0$, the *spin singlet* is

$$\left| \frac{1}{2}, \frac{1}{2}; 0, 0 \right\rangle = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right). \quad (3.7.34)$$

Worked example: reconstruct every coefficient for two spin- $\frac{1}{2}$ particles

The coupled and uncoupled bases introduced above are two representations of the same four-dimensional Hilbert space. To obtain the Clebsch–Gordan coefficients concretely, write the product kets $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$. Their J_z values are respectively $\hbar$, 0, 0 and $-\hbar$. The state with maximal m is uniquely

$$|1, 1\rangle = |\uparrow\uparrow\rangle. \quad (3.7.35)$$

Now apply the total lowering operator $\hat{J}_- = \hat{J}_{1-} + \hat{J}_{2-}$. For spin one, $\hat{J}_- |1, 1\rangle = \hbar\sqrt{2} |1, 0\rangle$. On the product ket, use $\hat{J}_{1-} |\uparrow\uparrow\rangle = \hbar |\downarrow\uparrow\rangle$ and $\hat{J}_{2-} |\uparrow\uparrow\rangle = \hbar |\uparrow\downarrow\rangle$:

$$\begin{aligned} \hbar\sqrt{2} |1, 0\rangle &= (\hat{J}_{1-} + \hat{J}_{2-}) |\uparrow\uparrow\rangle \\ &= \hbar |\downarrow\uparrow\rangle + \hbar |\uparrow\downarrow\rangle. \end{aligned} \quad (3.7.36)$$

Divide by $\hbar\sqrt{2}$:

$$|1, 0\rangle = \frac{1}{\sqrt{2}} (|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle). \quad (3.7.37)$$

Applying $\hat{J}_-$ once more,

$$\begin{aligned} \hbar\sqrt{2} |1, -1\rangle &= (\hat{J}_{1-} + \hat{J}_{2-}) |1, 0\rangle \\ &= \frac{\hbar}{\sqrt{2}} (|\downarrow\downarrow\rangle + |\downarrow\downarrow\rangle) \\ &= \hbar\sqrt{2} |\downarrow\downarrow\rangle, \end{aligned} \quad (3.7.38)$$

which yields $|1, -1\rangle = |\downarrow\downarrow\rangle$.

What is the fourth orthonormal state? It must have $m = 0$ and be orthogonal to $|1, 0\rangle$. Write $|\chi\rangle = a |\uparrow\downarrow\rangle + b |\downarrow\uparrow\rangle$. Orthogonality requires $(a + b)/\sqrt{2} = 0$, so $b = -a$. Normalization requires $|a|^2 + |b|^2 = 1$, hence $|a| = 1/\sqrt{2}$. Fixing the overall phase,

$$|0, 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \quad (3.7.39)$$

The three $j = 1$ triplets and one $j = 0$ singlet account for all four independent product states: $2 \otimes 2 = 3 \oplus 1$.

Check by applying total $\hat{\mathbf{J}}^2$. Since $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$,

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + 2\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}_2, \quad \hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}_2 = \hat{J}_{1z}\hat{J}_{2z} + \frac{1}{2}(\hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+}). \quad (3.7.40)$$

On the $m = 0$ basis ($|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle$), the operators become

$$\hat{J}_{1z}\hat{J}_{2z} = -\frac{\hbar^2}{4}\mathbf{1}, \quad \frac{1}{2}(\hat{J}_{1+}\hat{J}_{2-} + \hat{J}_{1-}\hat{J}_{2+}) = \frac{\hbar^2}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.7.41)$$

As $\hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 = \frac{3\hbar^2}{2}\mathbf{1}$, we obtain

$$J^2 = \hbar^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (3.7.42)$$

Multiplying by $(1, 1)^T/\sqrt{2}$ gives $2\hbar^2$ (triplet $j = 1$), while multiplying by $(1, -1)^T/\sqrt{2}$ gives 0 (singlet $j = 0$). The two calculations agree.

Self-check. Apply $\hat{J}_+$ to the singlet explicitly. Every resulting term cancels: $\hat{J}_+|0, 0\rangle = 0$. The coupled basis is therefore a direct consequence of angular-momentum algebra and normalization, without memorizing a table of coefficients.

Interactive investigation IC-03-02. Open Adding two angular momenta. After deriving the singlet and triplet states, convert between the product and coupled spin bases. Check the Clebsch–Gordan coefficients and verify normalization and total-spin eigenvalues.

3.8 Spin Correlations and Bell's Inequality

1) Correlations in spin measurements

The spin singlet of two electrons may be expanded in the eigenstates of $\hat{s}_{1z}$ and $\hat{s}_{2z}$:

$$|\tfrac{1}{2}, \tfrac{1}{2}; 0, 0\rangle = \frac{1}{\sqrt{2}} \left(|s_{1z}^+\rangle |s_{2z}^-\rangle - |s_{1z}^-\rangle |s_{2z}^+\rangle \right). \quad (3.8.1)$$

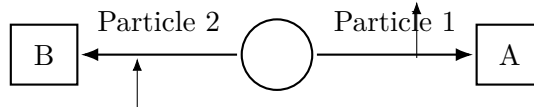
Because

$$|s_z^\pm\rangle = \frac{1}{\sqrt{2}} \left(|s_x^+\rangle \pm |s_x^-\rangle \right), \quad (3.8.2)$$

the spin singlet can also be expanded in the eigenstates of $\hat{s}_{1x}, \hat{s}_{2x}$:

$$|\tfrac{1}{2}, \tfrac{1}{2}; 0, 0\rangle = \frac{1}{\sqrt{2}} \left(|s_{1x}^-\rangle |s_{2x}^+\rangle - |s_{1x}^+\rangle |s_{2x}^-\rangle \right). \quad (3.8.3)$$

As illustrated below, let two electrons in the singlet state move from a central source to detectors A and B on opposite sides. Detector A first measures the spin of electron 1; immediately afterwards detector B measures the spin of electron 2.



If A measures s_z for electron 1 and B measures s_x for electron 2, then the probabilities for electron 2 to have s_x^+ or s_x^- are each 50%: there is no correlation between these measurement outcomes.

If A measures s_x for electron 1 and B also measures s_x for electron 2, the spin of electron 2 is determined by the result at A. When electron 1 is found to have s_x^+ , electron 2 has s_x^- ; when electron 1 has s_x^- , electron 2 has s_x^+ . The two measurements are correlated.

If A does not measure electron 1, while B measures s_x for electron 2, then electron 2 has s_x^+ and s_x^- with equal probability, 50% each. No correlation with an A measurement is observed.

2) Einstein locality and Bell's inequality

The spin-measurement correlations discussed above suggest that what is measured for particle 2 at B is related to the physical quantity previously measured for particle 1 at A, despite a large separation between A and B.

Many physicists found this long-range correlation puzzling. Einstein and collaborators proposed the principle of locality: the behavior of system 2 should not depend on the state of distant system 1 (the EPR paradox). To resolve the puzzle, many proposed hidden-variable theories: quantum mechanics results from a theory containing additional parameters that are not observed (hidden variables). We will not discuss a particular hidden-variable theory, but consider how experiments can distinguish the implications of Einstein locality from quantum-mechanical predictions.

Consider two particles. Assume a particle is of a definite type such that measuring s_z gives s_z^+ and measuring s_x gives s_x^+ , irrespective of the other distant particle's state, as required by the locality assumption in this model. Denote this type by (s_z^+, s_x^+) . Of course s_z and s_x are still not measured simultaneously. To maintain the total spin singlet, its partner must have the opposite type (s_z^-, s_x^-) . Continuing in this way, the possible pairs are

Particle 1	Particle 2	
(s_z^+, s_x^+)	(s_z^-, s_x^-)	(3.8.4)
(s_z^+, s_x^-)	(s_z^-, s_x^+)	
(s_z^-, s_x^+)	(s_z^+, s_x^-)	
(s_z^-, s_x^-)	(s_z^+, s_x^+)	

This reproduces the spin-measurement correlations described above. For example, measuring s_z for particle 1 at A and s_x for particle 2 at B allows both s_x^+ and s_x^- ; measuring s_x for both particles gives opposite results. Thus two measurement directions do not reveal a difference between these predictions.

Continue to consider two particles, but now measure spin along three directions $\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c$, i.e.

$$s_a = \mathbf{s} \cdot \mathbf{e}_a, \quad s_b = \mathbf{s} \cdot \mathbf{e}_b, \quad s_c = \mathbf{s} \cdot \mathbf{e}_c. \quad (3.8.5)$$

The directions need not be mutually perpendicular. A particle of type (s_a^+, s_b^+, s_c^+) means it yields s_a^+ if s_a is measured, s_b^+ if s_b is measured, and s_c^+ if s_c is measured (not all simultaneously). Requiring the two-particle system to be a spin singlet, the locality model gives the following types for particles 1 and 2 measured at A and B:

Particle 1	Particle 2	Count	
(s_a^+, s_b^+, s_c^+)	(s_a^-, s_b^-, s_c^-)	N_1	(3.8.6)
(s_a^+, s_b^+, s_c^-)	(s_a^-, s_b^-, s_c^+)	N_2	
(s_a^+, s_b^-, s_c^+)	(s_a^-, s_b^+, s_c^-)	N_3	
(s_a^+, s_b^-, s_c^-)	(s_a^-, s_b^+, s_c^+)	N_4	
(s_a^-, s_b^+, s_c^+)	(s_a^+, s_b^-, s_c^-)	N_5	
(s_a^-, s_b^+, s_c^-)	(s_a^+, s_b^-, s_c^+)	N_6	
(s_a^-, s_b^-, s_c^+)	(s_a^+, s_b^+, s_c^-)	N_7	
(s_a^-, s_b^-, s_c^-)	(s_a^+, s_b^+, s_c^+)	N_8	

The probability that particle 1 at A is found in s_a^+ and particle 2 at B in s_b^+ is

$$P(s_a^+, s_b^+) = \frac{N_3 + N_4}{\sum_{i=1}^8 N_i}. \quad (3.8.7)$$

Similarly,

$$P(s_a^+, s_c^+) = \frac{N_2 + N_4}{\sum_{i=1}^8 N_i}, \quad P(s_c^+, s_b^+) = \frac{N_3 + N_7}{\sum_{i=1}^8 N_i}. \quad (3.8.8)$$

Since

$$N_3 + N_4 \leq (N_3 + N_7) + (N_2 + N_4), \quad (3.8.9)$$

we have

$$\boxed{P(s_a^+, s_b^+) \leq P(s_a^+, s_c^+) + P(s_c^+, s_b^+).} \quad (3.8.10)$$

This is Bell's inequality in the form used here; it expresses a consequence of Einstein locality under the stated assignments.

3) Quantum mechanics and Bell's inequality

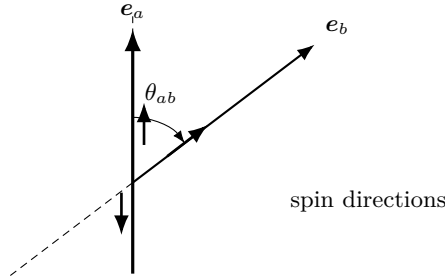
The preceding inequality is based on Einstein locality. Is it satisfied by quantum mechanics? We calculate the three probabilities in the inequality using quantum mechanics.

When particle 1 is measured along direction a at A and found in s_a^+ , we wish to determine the probability that particle 2 is measured along b at B and found in s_b^+ . For the spin singlet, particle 2 is in s_a^- . Using the rotation of spin states discussed previously, $|s_a^- \rangle$ can be expressed in the eigenbasis of $\hat{s}_b$. Let θ_{ab} be the angle between e_b and e_a , with relative azimuth α_{ab} :

$$|s_a^- \rangle = \cos \frac{\pi - \theta_{ab}}{2} e^{-i\alpha_{ab}/2} |s_b^+ \rangle + \sin \frac{\pi - \theta_{ab}}{2} e^{i\alpha_{ab}/2} |s_b^- \rangle. \quad (3.8.11)$$

Thus, conditional on s_a^- for particle 2, the probability of s_b^+ is

$$\cos^2 \frac{\pi - \theta_{ab}}{2} = \sin^2 \frac{\theta_{ab}}{2}. \quad (3.8.12)$$



Accounting for the probability 1/2 that particle 1 is in s_a^+ ,

$$\boxed{P(s_a^+, s_b^+) = \frac{1}{2} \sin^2 \frac{\theta_{ab}}{2}.} \quad (3.8.13)$$

Likewise,

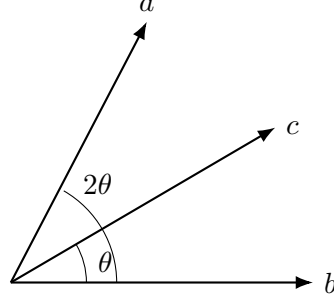
$$P(s_a^+, s_c^+) = \frac{1}{2} \sin^2 \frac{\theta_{ac}}{2}, \quad P(s_c^+, s_b^+) = \frac{1}{2} \sin^2 \frac{\theta_{cb}}{2}. \quad (3.8.14)$$

Bell's inequality would therefore require, for any three directions e_a, e_b, e_c ,

$$\boxed{\sin^2 \frac{\theta_{ab}}{2} \leq \sin^2 \frac{\theta_{ac}}{2} + \sin^2 \frac{\theta_{cb}}{2}.} \quad (3.8.15)$$

However, this inequality is not always satisfied. Let the three directions lie in one plane and let

$$\theta_{ab} = 2\theta, \quad \theta_{ac} = \theta_{cb} = \theta. \quad (3.8.16)$$



For $0 < \theta < \pi/2$, Bell's inequality reads

$$\sin^2 \theta \leq 2 \sin^2 \frac{\theta}{2}. \quad (3.8.17)$$

This fails, for example, at $\theta = \pi/4$: it would demand the impossible numerical relation $0.5 \leq 0.292$. Thus the quantum-mechanical calculation and Bell's inequality are inconsistent.

4) Experimental results

Many precise experiments have shown violations of Bell's inequality, with results agreeing with quantum-mechanical predictions within experimental errors.

Nevertheless, the interpretation of quantum mechanics, especially measurement, remains a question that has not been completely resolved.

Interactive investigation IC-03-06. Open Bell correlations and CHSH. After deriving the CHSH bound, vary the four analyzer directions. Calculate the correlators and the CHSH combination, then compare the classical bound with the quantum prediction and its maximum.

3.9 Quantum Teleportation: An Application of Entanglement and Measurement

The Bell correlations developed above and the general measurement formalism in this chapter can be combined into a communication protocol. *Quantum teleportation* transfers an unknown qubit state using one previously shared entangled pair and two classical bits. The protocol moves the quantum *state*, not the physical particle carrying it.

1) Bell states and an unknown input qubit

Use the computational basis $|0\rangle, |1\rangle$ and the Pauli operators

$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|, \quad \sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|. \quad (3.9.1)$$

The four Bell states are

$$\begin{aligned} |\Phi^\pm\rangle &= \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}, \\ |\Psi^\pm\rangle &= \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}. \end{aligned} \quad (3.9.2)$$

They form an orthonormal basis of two qubits. One may obtain them from computational-basis states with a Hadamard gate on the first qubit and a controlled-NOT (CNOT) gate from the first qubit to the second: a Hadamard creates a superposition; the CNOT correlates the target with its control. More precisely,

$$|\beta_{xy}\rangle = \text{CNOT}(\hat{U}_H \otimes \mathbf{1}) |x, y\rangle = \frac{|0, y\rangle + (-1)^x |1, 1-y\rangle}{\sqrt{2}}, \quad x, y \in \{0, 1\}. \quad (3.9.3)$$

This convention gives $|\beta_{00}\rangle = |\Phi^+\rangle$, $|\beta_{01}\rangle = |\Psi^+\rangle$, $|\beta_{10}\rangle = |\Phi^-\rangle$, and $|\beta_{11}\rangle = |\Psi^-\rangle$.

Let Alice hold an unknown normalized state

$$|\psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (3.9.4)$$

Alice and Bob have previously shared the entangled state $|\Phi^+\rangle_{23}$, with qubit 2 at Alice's location and qubit 3 at Bob's location. Their initial three-qubit state is

$$|\Psi_0\rangle = |\psi\rangle_1 \otimes |\Phi^+\rangle_{23} = \frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|100\rangle + |111\rangle)]. \quad (3.9.5)$$

The ordering of the three tensor factors is always 1, 2, 3.

2) The two-gate derivation

Alice first applies a CNOT with qubit 1 as the control and qubit 2 as the target. The target flips only when the control is $|1\rangle$:

$$|\Psi_1\rangle = \frac{1}{\sqrt{2}}[\alpha(|000\rangle + |011\rangle) + \beta(|110\rangle + |101\rangle)]. \quad (3.9.6)$$

Next Alice applies the Hadamard gate $\hat{U}_H$ to qubit 1, with

$$\hat{U}_H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad \hat{U}_H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (3.9.7)$$

Grouping the resulting three-qubit amplitudes according to Alice's two-qubit state gives

$$\boxed{|\Psi_2\rangle = \frac{1}{2}\{ |00\rangle_{12}(\alpha|0\rangle + \beta|1\rangle)_3 + |01\rangle_{12}(\alpha|1\rangle + \beta|0\rangle)_3 + |10\rangle_{12}(\alpha|0\rangle - \beta|1\rangle)_3 + |11\rangle_{12}(\alpha|1\rangle - \beta|0\rangle)_3 \}. \quad (3.9.8)}$$

The expression contains the entire protocol. Alice measures both qubits in the computational basis. Each two-bit result occurs with probability $1/4$, independent of the unknown amplitudes α, β . Conditioned on this outcome, Bob's qubit is one of four states related to $|\psi\rangle$ by Pauli operations.

3) Classical communication and the correction

Write Alice's two measured bits as $m_1 m_2$. Bob applies the appropriate correction:

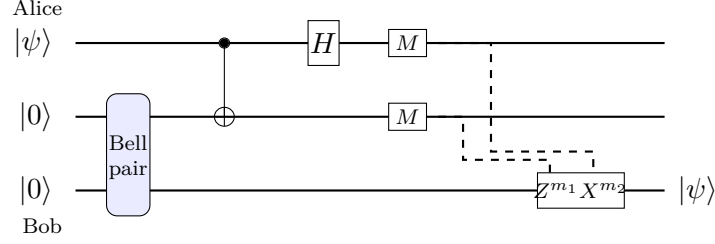
Alice's result	Bob's conditional state	Correction
00	$ \psi\rangle$	1
01	$\sigma_x \psi\rangle$	σ_x
10	$\sigma_z \psi\rangle$	σ_z
11	$\sigma_x \sigma_z \psi\rangle$	$\sigma_z \sigma_x$

(3.9.9)

The corrections are summarized, up to irrelevant global phases, by

$$\boxed{\hat{U}_{\text{corr}} = \sigma_z^{m_1} \sigma_x^{m_2}}. \quad (3.9.10)$$

Because the Pauli operators square to the identity, after Bob receives Alice's two classical bits his final state is $|\psi\rangle$. The prior entangled resource is consumed by the measurement.



In the circuit the pair shared before the protocol is represented by the initialization of qubits 2 and 3; the dashed lines indicate the two classical outcomes sent to Bob. The CNOT control is on qubit 1 and its target is qubit 2.

4) Causality and the no-cloning principle

Without Alice's classical message, Bob must average over her possible outcomes. His density operator is

$$\hat{\rho}_3 = \frac{1}{4}(\hat{\rho} + \sigma_x \hat{\rho} \sigma_x + \sigma_z \hat{\rho} \sigma_z + \sigma_x \sigma_z \hat{\rho} \sigma_z \sigma_x) = \frac{1}{2} \mathbf{1}, \quad (3.9.11)$$

where $\hat{\rho} = |\psi\rangle \langle \psi|$. Bob's local state is independent of α, β before he receives the classical bits, so the protocol cannot transmit information faster than light.

Teleportation also respects the no-cloning theorem. Alice measures and thereby changes her original input system; after the full protocol the unknown state is available at Bob's site, with no additional intact copy left at Alice's site. One shared Bell pair plus two classical bits is therefore sufficient to transfer a single arbitrary qubit state, but the entanglement is a consumed resource and classical communication remains necessary.

Chapter 4

Approximation Methods

4.1 The Idea of Stationary Perturbation Theory

The Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = \hat{H} |\psi, t\rangle. \quad (4.1.1)$$

Only a very small number of physical problems can be solved exactly. For most actual problems, the construction of $\hat{H}$ is so complicated that an exact solution is impossible; approximate methods must be used. Different problems call for different methods of approximation.

When $\hat{H}$ does not depend on time, separation of variables gives a stationary state

$$|\psi_n, t\rangle = |n\rangle e^{-iE_n t/\hbar}, \quad (4.1.2)$$

where the energy E_n and the energy eigenstate $|n\rangle$ satisfy the stationary Schrödinger equation

$$\hat{H} |n\rangle = E_n |n\rangle. \quad (4.1.3)$$

Write

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}. \quad (4.1.4)$$

The following are required:

1. $\hat{H}^{(0)}$ contains the main part of $\hat{H}$, so $\hat{H}^{(1)}$ is small. Since both $\hat{H}^{(0)}$ and $\hat{H}^{(1)}$ are operators, their relative magnitudes can be understood from the classical correspondence principle. Strictly speaking, one compares the matrix elements of the two operators, as discussed below.
2. The stationary equation for $\hat{H}^{(0)}$,

$$\hat{H}^{(0)} |n\rangle^{(0)} = E_n^{(0)} |n\rangle^{(0)}, \quad (4.1.5)$$

can be solved exactly.

Expand E_n and $|n\rangle$:

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots, \quad |n\rangle = |n\rangle^{(0)} + |n\rangle^{(1)} + |n\rangle^{(2)} + \cdots. \quad (4.1.6)$$

Insert these into the stationary Schrödinger equation:

$$\begin{aligned} (\hat{H}^{(0)} + \hat{H}^{(1)})(|n\rangle^{(0)} + |n\rangle^{(1)} + |n\rangle^{(2)} + \cdots) \\ = (E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots)(|n\rangle^{(0)} + |n\rangle^{(1)} + |n\rangle^{(2)} + \cdots). \end{aligned} \quad (4.1.7)$$

The equation is then solved successively, order by order.

Zeroth-order approximation:

$$\hat{H}^{(0)} |n\rangle^{(0)} = E_n^{(0)} |n\rangle^{(0)} \longrightarrow E_n^{(0)}, |n\rangle^{(0)}. \quad (4.1.8)$$

First-order approximation:

$$\hat{H}^{(0)} |n\rangle^{(1)} + \hat{H}^{(1)} |n\rangle^{(0)} = E_n^{(0)} |n\rangle^{(1)} + E_n^{(1)} |n\rangle^{(0)}, \quad (4.1.9)$$

or

$$(\hat{H}^{(0)} - E_n^{(0)}) |n\rangle^{(1)} = -(\hat{H}^{(1)} - E_n^{(1)}) |n\rangle^{(0)} \longrightarrow E_n^{(1)}, |n\rangle^{(1)}. \quad (4.1.10)$$

Second-order approximation:

$$(\hat{H}^{(0)} - E_n^{(0)}) |n\rangle^{(2)} = -(\hat{H}^{(1)} - E_n^{(1)}) |n\rangle^{(1)} + E_n^{(2)} |n\rangle^{(0)} \longrightarrow E_n^{(2)}, |n\rangle^{(2)}. \quad (4.1.11)$$

And so forth. In general, only the first nonzero correction is calculated.

Perturbation theory is based on the eigenvalues $E_n^{(0)}$ and eigenstates $|n\rangle^{(0)}$ of $\hat{H}^{(0)}$. In obtaining higher-order corrections, it is necessary to consider whether $E_n^{(0)}$ is degenerate. We first consider the nondegenerate case.

4.2 Nondegenerate Perturbation Theory

1) First-order correction

Expand the first-order equation in the $\hat{H}^{(0)}$ representation:

$$|n\rangle^{(1)} = \sum_i |i\rangle^{(0)} {}^{(0)}\langle i|n^{(1)}\rangle, \quad (4.2.1)$$

$$\sum_i (\hat{H}^{(0)} - E_n^{(0)}) |i\rangle^{(0)} {}^{(0)}\langle i|n^{(1)}\rangle = -(\hat{H}^{(1)} - E_n^{(1)}) |n\rangle^{(0)}. \quad (4.2.2)$$

Multiply on the left by ${}^{(0)}\langle m|$ to obtain

$$\begin{aligned} \sum_i (E_i^{(0)} - E_n^{(0)}) \delta_{mi} {}^{(0)}\langle i|n^{(1)}\rangle &= -{}^{(0)}\langle m|\hat{H}^{(1)}|n^{(0)}\rangle + E_n^{(1)} \delta_{mn} \\ &= -H_{mn}^{(1)} + E_n^{(1)} \delta_{mn}. \end{aligned} \quad (4.2.3)$$

Here $H_{mn}^{(1)} = {}^{(0)}\langle m|\hat{H}^{(1)}|n^{(0)}\rangle$ is the matrix element of $\hat{H}^{(1)}$ in the $\hat{H}^{(0)}$ representation. The orthonormality of the eigenstates of $\hat{H}^{(0)}$ has already been used. Thus

$$(E_m^{(0)} - E_n^{(0)}) {}^{(0)}\langle m|n^{(1)}\rangle = -H_{mn}^{(1)} + E_n^{(1)} \delta_{mn}. \quad (4.2.4)$$

For $m = n$,

$$E_n^{(1)} = H_{nn}^{(1)}. \quad (4.2.5)$$

For $m \neq n$,

$${}^{(0)}\langle m|n^{(1)}\rangle = \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}}. \quad (4.2.6)$$

To obtain the full expansion

$$|n\rangle^{(1)} = \sum_m |m\rangle^{(0)} {}^{(0)}\langle m|n^{(1)}\rangle, \quad (4.2.7)$$

we must also determine ${}^{(0)}\langle n|n^{(1)}\rangle$.

Normalize to first order, neglecting terms of second and higher order:

$$\begin{aligned} 1 = \langle n|n \rangle &= {}^{(0)}\langle n| + {}^{(1)}\langle n| \rangle (|n\rangle^{(0)} + |n\rangle^{(1)}) \\ &\simeq \langle n^{(0)}|n^{(0)} \rangle + \langle n^{(1)}|n^{(0)} \rangle + \langle n^{(0)}|n^{(1)} \rangle. \end{aligned} \quad (4.2.8)$$

Since $\langle n^{(0)}|n^{(0)} \rangle = 1$,

$$\langle n^{(0)}|n^{(1)} \rangle + \langle n^{(1)}|n^{(0)} \rangle = 0, \quad (4.2.9)$$

$$\langle n^{(0)}|n^{(1)} \rangle + (\langle n^{(0)}|n^{(1)} \rangle)^* = 0. \quad (4.2.10)$$

The real part of $\langle n^{(0)}|n^{(1)} \rangle$ must vanish; it has only an imaginary part. Write

$$\langle n^{(0)}|n^{(1)} \rangle = i\alpha. \quad (4.2.11)$$

Therefore

$$\begin{aligned} |n\rangle &= |n\rangle^{(0)} + \sum_m |m\rangle^{(0)} \langle m^{(0)}|n^{(1)} \rangle \\ &= |n\rangle^{(0)} + i\alpha |n\rangle^{(0)} + \sum_{m \neq n} \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)} \\ &\simeq e^{i\alpha} |n\rangle^{(0)} + \sum_{m \neq n} \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)} \\ &\simeq e^{i\alpha} \left(|n\rangle^{(0)} + \sum_{m \neq n} \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)} \right) \quad (\text{to first order}). \end{aligned} \quad (4.2.12)$$

The contribution $\langle n^{(0)}|n^{(1)} \rangle = i\alpha$ to the first-order correction merely supplies a phase factor and does not affect the probability distribution. One may choose $\alpha = 0$.

The solution of the stationary Schrödinger equation to first order is thus

$$\boxed{\begin{aligned} E_n &= E_n^{(0)} + H_{nn}^{(1)}, \\ |n\rangle &= |n\rangle^{(0)} + \sum_{m \neq n} \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)}. \end{aligned}} \quad (4.2.13)$$

2) Second-order correction

Expand the second-order equation in the $\hat{H}^{(0)}$ representation:

$$|n\rangle^{(2)} = \sum_i |i\rangle^{(0)} \langle i^{(0)}|n^{(2)} \rangle, \quad (4.2.14)$$

$$\begin{aligned} \sum_i (\hat{H}^{(0)} - E_n^{(0)}) |i\rangle^{(0)} \langle i^{(0)}|n^{(2)} \rangle \\ = - \sum_{i \neq n} (\hat{H}^{(1)} - E_n^{(1)}) |i\rangle^{(0)} \langle i^{(0)}|n^{(1)} \rangle + E_n^{(2)} |n\rangle^{(0)}. \end{aligned} \quad (4.2.15)$$

Multiplying by ${}^{(0)}\langle m|$ gives

$$\begin{aligned} \sum_i (E_i^{(0)} - E_n^{(0)}) \delta_{mi} \langle i^{(0)}|n^{(2)} \rangle \\ = - \sum_{i \neq n} (H_{mi}^{(1)} - E_n^{(1)} \delta_{mi}) \langle i^{(0)}|n^{(1)} \rangle + E_n^{(2)} \delta_{mn}, \\ (E_m^{(0)} - E_n^{(0)}) \langle m^{(0)}|n^{(2)} \rangle = - \sum_{i \neq n} H_{mi}^{(1)} \langle i^{(0)}|n^{(1)} \rangle + E_n^{(1)} \langle m^{(0)}|n^{(1)} \rangle + E_n^{(2)} \delta_{mn}. \end{aligned} \quad (4.2.16)$$

For $m = n$, taking $\langle n^{(0)} | n^{(1)} \rangle = i\alpha = 0$,

$$\begin{aligned} E_n^{(2)} &= \sum_{i \neq n} H_{ni}^{(1)} \langle i^{(0)} | n^{(1)} \rangle = \sum_{i \neq n} \frac{H_{ni}^{(1)} H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} \\ &= \sum_{i \neq n} \frac{|H_{in}^{(1)}|^2}{E_n^{(0)} - E_i^{(0)}}. \end{aligned} \quad (4.2.17)$$

For $m \neq n$, the coefficient $\langle m^{(0)} | n^{(2)} \rangle$ can also be found (see the textbook and calculate it yourself).

The approximate solution of the stationary Schrödinger equation through second order is

$$\begin{aligned} E_n &= E_n^{(0)} + H_{nn}^{(1)} + \sum_{i \neq n} \frac{|H_{in}^{(1)}|^2}{E_n^{(0)} - E_i^{(0)}}, \\ |n\rangle &= |n\rangle^{(0)} + \sum_{i \neq n} \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} |i\rangle^{(0)} + |n\rangle^{(2)}. \end{aligned} \quad (4.2.18)$$

3) Discussion of convergence

The condition for perturbation theory to be applicable is convergence, namely,

$$\left| \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} \right| \ll 1, \quad \text{i.e.,} \quad |H_{in}^{(1)}| \ll |E_n^{(0)} - E_i^{(0)}|. \quad (4.2.19)$$

This is the meaning of $\hat{H}^{(1)}$ being much smaller than $\hat{H}^{(0)}$: in the $H^{(0)}$ representation, its matrix elements are much smaller than the differences between the corresponding energy levels.

It follows that:

1. The above method applies only when $\hat{H}^{(0)}$ has a discrete spectrum. For a continuous spectrum, $|E_n^{(0)} - E_i^{(0)}| \rightarrow 0$ and the above inequality cannot hold.
2. It cannot be used when $\hat{H}^{(0)}$ has degeneracies. In that case, $|E_n^{(0)} - E_i^{(0)}| = 0$ and the inequality cannot hold either.

Example: A charged harmonic oscillator in an external electric field. Let the electric charge be q and the electric-field strength be $\mathcal{E}$. The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2 - q \mathcal{E} x. \quad (4.2.20)$$

For a weak electric field $\mathcal{E}$, take

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}, \quad \hat{H}^{(0)} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m \omega^2 x^2, \quad \hat{H}^{(1)} = -q \mathcal{E} x. \quad (4.2.21)$$

Zeroth-order approximation:

$$\hat{H}^{(0)} |n\rangle^{(0)} = E_n^{(0)} |n\rangle^{(0)}, \quad E_n^{(0)} = \left(n + \frac{1}{2}\right) \hbar \omega, \quad \langle x^{(0)} | n^{(0)} \rangle = \psi_n^{(0)}(x). \quad (4.2.22)$$

The harmonic-oscillator eigenfunctions are Hermite-polynomial functions, with discrete, non-degenerate energies and separated variables.

First-order correction:

$$E_n^{(1)} = H_{nn}^{(1)} = -q \mathcal{E} x_{nn}, \quad (4.2.23)$$

$$x_{nm} = {}^{(0)}\langle n|x|m^{(0)}\rangle = \sqrt{\frac{\hbar}{m\omega}} \left(\sqrt{\frac{n+1}{2}} \delta_{n+1,m} + \sqrt{\frac{n}{2}} \delta_{n-1,m} \right), \quad x_{nn} = 0, \quad E_n^{(1)} = 0. \quad (4.2.24)$$

For the first-order change of the state,

$$\begin{aligned} |n\rangle^{(1)} &= \sum_{m \neq n} \frac{H_{mn}^{(1)}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)} = -q\mathcal{E} \sum_{m \neq n} \frac{x_{mn}}{E_n^{(0)} - E_m^{(0)}} |m\rangle^{(0)} \\ &= -q\mathcal{E} \sqrt{\frac{\hbar}{2m\omega}} \left(\frac{\sqrt{n+1}}{E_n^{(0)} - E_{n+1}^{(0)}} |n+1\rangle^{(0)} + \frac{\sqrt{n}}{E_n^{(0)} - E_{n-1}^{(0)}} |n-1\rangle^{(0)} \right) \\ &= \frac{q\mathcal{E}}{\sqrt{2m\hbar\omega^3}} \left(\sqrt{n+1} |n+1\rangle^{(0)} - \sqrt{n} |n-1\rangle^{(0)} \right). \end{aligned} \quad (4.2.25)$$

In the coordinate representation,

$$\psi_n^{(1)}(x) = \langle x|n^{(1)}\rangle = \frac{q\mathcal{E}}{\sqrt{2m\hbar\omega^3}} \left[\sqrt{n+1} \psi_{n+1}^{(0)}(x) - \sqrt{n} \psi_{n-1}^{(0)}(x) \right]. \quad (4.2.26)$$

Second-order correction:

$$\begin{aligned} E_n^{(2)} &= \sum_{m \neq n} \frac{|H_{mn}^{(1)}|^2}{E_n^{(0)} - E_m^{(0)}} \\ &= \frac{q^2 \mathcal{E}^2 \hbar}{m\omega} \sum_{m \neq n} \frac{|\sqrt{(n+1)/2} \delta_{m,n+1} + \sqrt{n/2} \delta_{m,n-1}|^2}{E_n^{(0)} - E_m^{(0)}} \\ &= -\frac{q^2 \mathcal{E}^2}{2m\omega^2}. \end{aligned} \quad (4.2.27)$$

Thus, keeping the first two nonzero corrections,

$$\boxed{\begin{aligned} E_n &= \left(n + \frac{1}{2}\right) \hbar\omega - \frac{q^2 \mathcal{E}^2}{2m\omega^2}, \\ \psi_n(x) &\simeq \psi_n^{(0)}(x) + \frac{q\mathcal{E}}{\sqrt{2m\hbar\omega^3}} \left(\sqrt{n+1} \psi_{n+1}^{(0)}(x) - \sqrt{n} \psi_{n-1}^{(0)}(x) \right). \end{aligned}} \quad (4.2.28)$$

Can an exact solution be found? Start from

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2 - q\mathcal{E}x \right) \psi(x) = E\psi(x). \quad (4.2.29)$$

Complete the square:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 \left(x - \frac{q\mathcal{E}}{m\omega^2} \right)^2 \right] \psi(x) = \left(E + \frac{q^2 \mathcal{E}^2}{2m\omega^2} \right) \psi(x). \quad (4.2.30)$$

Put $\xi = x - q\mathcal{E}/(m\omega^2)$; then

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{d\xi^2} + \frac{1}{2} m\omega^2 \xi^2 \right) \psi(\xi) = \left(E + \frac{q^2 \mathcal{E}^2}{2m\omega^2} \right) \psi(\xi). \quad (4.2.31)$$

The solution is

$$\boxed{\begin{aligned} E_n &= \left(n + \frac{1}{2}\right) \hbar\omega - \frac{q^2 \mathcal{E}^2}{2m\omega^2}, \\ \psi_n(\xi) &= \psi_n^{(0)} \left(x - \frac{q\mathcal{E}}{m\omega^2} \right), \end{aligned}} \quad (4.2.32)$$

where $\psi_n^{(0)}$ is a Hermite-polynomial function.

The energy computed through second order is already exact.

4.3 Degenerate Perturbation Theory

Suppose

$$\hat{H}^{(0)} |n, i\rangle^{(0)} = E_n^{(0)} |n, i\rangle^{(0)}, \quad i = 1, 2, \dots, a. \quad (4.3.1)$$

The zeroth-order energy has a -fold degeneracy. Schmidt orthogonalization or the introduction of a complete set of mechanical quantities can guarantee mutual orthonormality:

$${}^{(0)}\langle n, i | n, j \rangle^{(0)} = \delta_{ij}. \quad (4.3.2)$$

Question: Which one of the states $|n, i\rangle^{(0)}$ should be chosen as the wavefunction for $E_n^{(0)}$?

Take the general form

$$|n\rangle^{(0)} = \sum_{i=1}^a c_i |n, i\rangle^{(0)}. \quad (4.3.3)$$

It is still an eigenstate of $\hat{H}^{(0)}$ belonging to $E_n^{(0)}$.

Question: How should the c_i be chosen?

Substitute $|n\rangle^{(0)}$ into the first-order approximation equation:

$$(\hat{H}^{(0)} - E_n^{(0)}) |n\rangle^{(1)} = -(\hat{H}^{(1)} - E_n^{(1)}) |n\rangle^{(0)}, \quad (4.3.4)$$

so that

$$(\hat{H}^{(0)} - E_n^{(0)}) |n\rangle^{(1)} = - \sum_{i=1}^a c_i (\hat{H}^{(1)} - E_n^{(1)}) |n, i\rangle^{(0)}. \quad (4.3.5)$$

Multiply on the left by ${}^{(0)}\langle n, j |$:

$$0 = \sum_{i=1}^a c_i (H_{ji}^{(1)} - E_n^{(1)} \delta_{ji}), \quad H_{ji}^{(1)} = {}^{(0)}\langle n, j | \hat{H}^{(1)} | n, i \rangle^{(0)}. \quad (4.3.6)$$

This is a homogeneous system of equations for the coefficients c_i (a equations, $j = 1, \dots, a$, determining a coefficients, $i = 1, \dots, a$). For a nonzero solution, its determinant must vanish:

$$\begin{vmatrix} H_{11}^{(1)} - E_n^{(1)} & H_{12}^{(1)} & \cdots & H_{1a}^{(1)} \\ H_{21}^{(1)} & H_{22}^{(1)} - E_n^{(1)} & \cdots & H_{2a}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ H_{a1}^{(1)} & H_{a2}^{(1)} & \cdots & H_{aa}^{(1)} - E_n^{(1)} \end{vmatrix} = 0. \quad (4.3.7)$$

The secular equation therefore determines a roots $E_{ni}^{(1)}$ of the first-order energy correction:

$$E_n^{(0)} \longrightarrow E_{ni} = E_n^{(0)} + E_{ni}^{(1)}, \quad i = 1, \dots, a. \quad (4.3.8)$$

If the a roots of the secular equation are distinct, the degeneracy is completely removed; if roots repeat, it is only partially removed.

Note. If $\hat{H}^{(1)}$ is diagonal in the degenerate subspace, $E_{ni}^{(1)} = H_{ii}^{(1)}$.

Substituting each $E_{ni}^{(1)}$ back into the homogeneous equations gives a set of coefficients $\{c_{i1}, c_{i2}, \dots, c_{ia}\}$, and hence the zeroth-order wavefunction of $\hat{H}$ corresponding to E_{ni} :

$$E_{ni} = E_n^{(0)} + E_{ni}^{(1)}, \quad |n_i\rangle^{(0)} = \sum_{j=1}^a c_{ij} |n, j\rangle^{(0)}. \quad (4.3.9)$$

Higher-order corrections can be found in the references.

Example 1: A doubly degenerate system. Let $E_n^{(0)}$ have basis states $|n, 1\rangle^{(0)}$ and $|n, 2\rangle^{(0)}$. The secular equation in this two-dimensional subspace is

$$\begin{vmatrix} H_{11}^{(1)} - E_n^{(1)} & H_{12}^{(1)} \\ H_{21}^{(1)} & H_{22}^{(1)} - E_n^{(1)} \end{vmatrix} = 0, \quad H_{ij}^{(1)} = {}^{(0)}\langle n, i | \hat{H}^{(1)} | n, j \rangle^{(0)}, \quad i, j = 1, 2. \quad (4.3.10)$$

Thus

$$(E_n^{(1)})^2 - (H_{11}^{(1)} + H_{22}^{(1)})E_n^{(1)} + H_{11}^{(1)}H_{22}^{(1)} - |H_{12}^{(1)}|^2 = 0, \quad (4.3.11)$$

$$E_n^{(1)} = \frac{1}{2} \left[H_{11}^{(1)} + H_{22}^{(1)} \pm \sqrt{(H_{11}^{(1)} - H_{22}^{(1)})^2 + 4|H_{12}^{(1)}|^2} \right]. \quad (4.3.12)$$

If $H_{12}^{(1)} = 0$, then

$$E_{n1}^{(1)} = H_{11}^{(1)}, \quad E_{n2}^{(1)} = H_{22}^{(1)}, \quad E_{n1} = E_n^{(0)} + H_{11}^{(1)}, \quad E_{n2} = E_n^{(0)} + H_{22}^{(1)}. \quad (4.3.13)$$

Example 2: The hydrogen atom in an external electric field (Stark effect). Choose the z axis along the external electric field $\mathcal{E}$:

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{r} + e\mathcal{E} \cdot \mathbf{r} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{r} + e\mathcal{E}r \cos \theta. \quad (4.3.14)$$

For a weak field $\mathcal{E}$,

$$\hat{H}^{(0)} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{r}, \quad \hat{H}^{(1)} = e\mathcal{E}r \cos \theta. \quad (4.3.15)$$

Zeroth-order approximation:

$$E_n^{(0)} = -\frac{\mu e^4}{2n^2 \hbar^2}, \quad \psi_{nlm}(\mathbf{r}), \quad l = 0, \dots, n-1, \quad m = -l, \dots, l, \quad (4.3.16)$$

with degeneracy

$$g = \sum_{l=0}^{n-1} (2l+1) = n^2. \quad (4.3.17)$$

For $n = 2$, the zeroth-order energy $E_2^{(0)}$ is fourfold degenerate, with states

$$\phi_1 = \psi_{200}, \quad \phi_2 = \psi_{210}, \quad \phi_3 = \psi_{211}, \quad \phi_4 = \psi_{21,-1}. \quad (4.3.18)$$

To find the first-order splitting, calculate the matrix elements

$$H_{ij}^{(1)} = \int d^3r \phi_i^*(\mathbf{r}) \hat{H}^{(1)} \phi_j(\mathbf{r}), \quad i, j = 1, 2, 3, 4. \quad (4.3.19)$$

The matrix element is

$$H_{12}^{(1)} = H_{21}^{(1)} = -3e\mathcal{E}a_0, \quad (4.3.20)$$

with all other entries zero. The secular equation in the four-dimensional subspace is

$$\begin{vmatrix} -E_2^{(1)} & -3e\mathcal{E}a_0 & 0 & 0 \\ -3e\mathcal{E}a_0 & -E_2^{(1)} & 0 & 0 \\ 0 & 0 & -E_2^{(1)} & 0 \\ 0 & 0 & 0 & -E_2^{(1)} \end{vmatrix} = 0. \quad (4.3.21)$$

Hence

$$E_{21}^{(1)} = 3e\mathcal{E}a_0, \quad E_{22}^{(1)} = -3e\mathcal{E}a_0, \quad E_{23}^{(1)} = E_{24}^{(1)} = 0. \quad (4.3.22)$$

The original fourfold-degenerate level splits into one level shifted upward by $3e\mathcal{E}a_0$, a twofold-degenerate level with zero shift, and one level shifted downward by $3e\mathcal{E}a_0$. The degenerate perturbation is partially removed.

Example 3: Relativistic correction for the hydrogen atom. The nonrelativistic kinetic energy is

$$T = \frac{p^2}{2m}. \quad (4.3.23)$$

The relativistic kinetic energy is

$$\begin{aligned} T &= \sqrt{p^2c^2 + m^2c^4} - mc^2 = mc^2 \sqrt{1 + \frac{p^2}{m^2c^2}} - mc^2 \\ &= \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \end{aligned} \quad (4.3.24)$$

Keeping only the p^4 correction gives the relativistic Hamiltonian

$$\hat{H} = \sqrt{\hat{p}^2c^2 + m^2c^4} - mc^2 - \frac{e^2}{r} \simeq \frac{\hat{p}^2}{2m} - \frac{e^2}{r} - \frac{\hat{p}^4}{8m^3c^2}. \quad (4.3.25)$$

Take

$$\hat{H}^{(0)} = \frac{\hat{p}^2}{2m} - \frac{e^2}{r}, \quad \hat{H}^{(1)} = -\frac{\hat{p}^4}{8m^3c^2}. \quad (4.3.26)$$

The zeroth-order energy is $E_n^{(0)}$ and the n^2 eigenstates are $|nlm\rangle$, with wavefunctions

$$\langle \mathbf{r} | nlm \rangle = \psi_{nlm}(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi). \quad (4.3.27)$$

In the n^2 -dimensional subspace,

$$H_{nlm, n'l'm'}^{(1)} = -\frac{1}{8m^3c^2} \langle nlm | \hat{p}^4 | n'l'm' \rangle. \quad (4.3.28)$$

The zeroth-order Schrödinger equation is

$$\left(\frac{\hat{p}^2}{2m} + V(r) \right) |nlm\rangle = E_n^{(0)} |nlm\rangle, \quad \hat{p}^2 |nlm\rangle = 2m(E_n^{(0)} - V(r)) |nlm\rangle. \quad (4.3.29)$$

Using this relation, the matrix elements become

$$\begin{aligned} H_{nlm, n'l'm'}^{(1)} &= -\frac{1}{2mc^2} \langle nlm | (E_n^{(0)} - V(r))^2 | n'l'm' \rangle \\ &= -\frac{1}{2mc^2} \left[(E_n^{(0)})^2 \delta_{nn'} \delta_{ll'} \delta_{mm'} - 2E_n^{(0)} \langle nlm | V(r) | n'l'm' \rangle + \langle nlm | V^2(r) | n'l'm' \rangle \right]. \end{aligned} \quad (4.3.30)$$

Since $V(r)$ depends only on r and not on the angles,

$$\langle nlm | V(r) | n'l'm' \rangle = \int d^3r R_{nl}(r)Y_{lm}^*(\theta, \varphi) \left(-\frac{e^2}{r} \right) R_{n'l'}(r)Y_{l'm'}(\theta, \varphi) = -\frac{e^2}{n^2a_0} \delta_{ll'} \delta_{mm'} \quad (4.3.31)$$

within a fixed principal-energy subspace. Similarly,

$$\langle nlm | V^2(r) | n'l'm' \rangle = \frac{e^4}{(l+1/2)n^3a_0^2} \delta_{ll'} \delta_{mm'}. \quad (4.3.32)$$

Therefore the first-order matrix is diagonal:

$$H_{nlm, nl'm'}^{(1)} = H_{nl}^{(1)} \delta_{ll'} \delta_{mm'}, \quad (4.3.33)$$

$$H_{nl}^{(1)} = -\frac{1}{2mc^2} \left[(E_n^{(0)})^2 + \frac{2e^2}{n^2 a_0} E_n^{(0)} + \frac{e^4}{(l+1/2)n^3 a_0^2} \right]. \quad (4.3.34)$$

This shows that the eigenstates of $\hat{H}^{(0)}$ are also eigenstates of $\hat{H}^{(1)}$ in this subspace, and the secular equation reduces to $E_{nl}^{(1)} = H_{nl}^{(1)}$:

$$E_n^{(0)} \quad (n^2\text{-fold degenerate}) \longrightarrow E_{nl} = E_n^{(0)} + H_{nl}^{(1)} \quad ((2l+1)\text{-fold degenerate}). \quad (4.3.35)$$

Example 4: The fine structure of the spectrum. Consider the spin-orbit coupling of the electron:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}, \quad \hat{H}^{(0)} = -\frac{\hbar^2}{2\mu} \nabla^2 + V(r), \quad (4.3.36)$$

$$\hat{H}^{(1)} = \frac{1}{2\mu^2 r} \frac{dV(r)}{dr} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \equiv \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}. \quad (4.3.37)$$

Zeroth-order states: the system is in an energy eigenstate $E_{nl}^{(0)}$. For a given l , the energy level has $2(2l+1)$ degeneracy. For this level the basis may be chosen as

$$\{|n, l, m, s_z\rangle\}, \quad |n, l, j, m_j\rangle, \quad (4.3.38)$$

using either the $\{\hat{H}^{(0)}, \hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$ representation or the $\{\hat{H}^{(0)}, \hat{\mathbf{J}}^2, \hat{\mathbf{L}}^2, \hat{J}_z\}$ representation. The latter describes an eigenstate $|nljm_j\rangle$.

The zeroth-order energy is degenerate, so find the matrix elements of $\hat{H}^{(1)}$. Which representation should be used?

In the uncoupled representation, $[\hat{H}^{(1)}, \hat{L}_z] \neq 0$ and $[\hat{H}^{(1)}, \hat{S}_z] \neq 0$, so $|nlms_z\rangle$ is not an eigenstate of $\hat{H}^{(1)}$. Its matrix elements need not be diagonal and the secular equation is more complicated.

In the coupled representation, $[\hat{H}^{(1)}, \hat{\mathbf{J}}^2] = 0$ and $[\hat{H}^{(1)}, \hat{J}_z] = 0$. Thus $|nljm_j\rangle$ is an eigenstate of $\hat{H}^{(1)}$, whose matrix is diagonal. The first-order energy correction is

$$\begin{aligned} E_{nlj}^{(1)} &= \langle nljm_j | \hat{H}^{(1)} | nljm_j \rangle \\ &= \langle nljm_j | \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} | nljm_j \rangle \\ &= \langle nljm_j | \xi(r) \frac{1}{2} (\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2) | nljm_j \rangle \\ &= \frac{\hbar^2}{2} [j(j+1) - l(l+1) - s(s+1)] \int_0^\infty R_{nl}^2(r) \xi(r) r^2 dr. \end{aligned} \quad (4.3.39)$$

The level changes as

$$E_{nl}^{(0)} \quad (2(2l+1)\text{-fold degenerate}) \longrightarrow E_{nlj} = E_{nl}^{(0)} + E_{nlj}^{(1)} \quad ((2j+1)\text{-fold degenerate}). \quad (4.3.40)$$

For example, when $n = 2$,

$$E_2^{(0)} \quad (8\text{-fold degenerate}) \longrightarrow \begin{cases} E_2^{(0)} + E_{2,1,3/2}^{(1)} & (4\text{-fold}), \\ E_2^{(0)} + E_{2,1,1/2}^{(1)} & (2\text{-fold}), \\ E_2^{(0)} + E_{2,0,1/2}^{(1)} & (2\text{-fold}). \end{cases} \quad (4.3.41)$$

Note: when $l = 0$, $j = s$ and $E_{2,0,1/2}^{(1)} = 0$.

Interactive investigation IC-04-01. Open Avoided crossings. After studying degenerate perturbation theory, tune the detuning and coupling in an exactly solvable two-level Hamiltonian. Compare exact eigenvalues and state mixing with perturbative estimates near the avoided crossing.

4.4 Variational Method

The variational method is used to solve the stationary equation

$$\hat{H} |n\rangle = E_n |n\rangle. \quad (4.4.1)$$

For an arbitrary state,

$$|\psi\rangle = \sum_n c_n |n\rangle, \quad (4.4.2)$$

the energy expectation value is

$$\begin{aligned} \langle E \rangle &= \langle \psi | \hat{H} | \psi \rangle = \sum_{n,m} c_n^* c_m \langle n | \hat{H} | m \rangle \\ &= \sum_{n,m} c_n^* c_m E_n \delta_{nm} = \sum_n |c_n|^2 E_n \geq E_0. \end{aligned} \quad (4.4.3)$$

This is the basis of the variational method: take different trial functions $|\psi\rangle$, calculate $\langle E \rangle = \langle \psi | \hat{H} | \psi \rangle$, and choose the lowest $\langle E \rangle$ to approximate the ground-state energy E_0 .

Method. From the physical properties of the system, guess a trial wavefunction $|\psi(\lambda)\rangle$ containing a parameter λ . Calculate

$$\langle E \rangle(\lambda) = \langle \psi(\lambda) | \hat{H} | \psi(\lambda) \rangle. \quad (4.4.4)$$

Solve

$$\frac{d\langle E \rangle(\lambda)}{d\lambda} = 0, \quad \frac{d^2\langle E \rangle(\lambda)}{d\lambda^2} > 0 \quad \longrightarrow \quad \lambda_0, \quad (4.4.5)$$

which gives the minimum $\langle E \rangle(\lambda_0)$. The ground-state energy is approximately

$$E_0 \simeq \langle E \rangle(\lambda_0). \quad (4.4.6)$$

Example: The helium atom.

$$\begin{aligned} \hat{H} &= -\frac{\hbar^2}{2\mu} \nabla_1^2 - \frac{\hbar^2}{2\mu} \nabla_2^2 - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &= \hat{H}^{(0)} + \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \end{aligned} \quad (4.4.7)$$

Here the atomic nucleus is approximately considered fixed. In the ground state of $\hat{H}^{(0)}$,

$$E_0^{(0)} = -\frac{4e^2}{a_0}, \quad (4.4.8)$$

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2), \quad \psi_{100}(\mathbf{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0}, \quad Z = 2. \quad (4.4.9)$$

Taking the interaction between the two electrons into account, the screening by one electron reduces the effective nuclear charge felt by the other to $Z < 2$. Treat Z as a parameter, and use $\psi_0(\mathbf{r}_1, \mathbf{r}_2, Z)$ as the trial wavefunction for $\hat{H}$. Since $\hat{H}$ itself is independent of Z ,

$$\begin{aligned} \langle E \rangle(Z) &= \int d^3r_1 d^3r_2 \psi_0^*(\mathbf{r}_1, \mathbf{r}_2, Z) \hat{H} \psi_0(\mathbf{r}_1, \mathbf{r}_2, Z) \\ &= \left(2Z^2 - \frac{27}{4}Z \right) \frac{e^2}{2a_0}. \end{aligned} \quad (4.4.10)$$

From

$$\frac{d\langle E \rangle(Z)}{dZ} = 0, \quad \frac{d^2\langle E \rangle(Z)}{dZ^2} > 0, \quad (4.4.11)$$

one obtains

$$Z_0 = \frac{27}{16} < 2. \quad (4.4.12)$$

The ground-state energy is

$$E_0 \simeq \langle E \rangle(Z_0) = -2.85 \frac{e^2}{a_0}, \quad (4.4.13)$$

with corresponding trial wavefunction $\psi_0(\mathbf{r}_1, \mathbf{r}_2, Z_0)$.

Perturbation theory can also be used to obtain the correction to the ground-state energy. Take

$$\hat{H}^{(1)} = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (4.4.14)$$

The ground state is nondegenerate; nondegenerate perturbation theory gives

$$E_0^{(1)} = \int d^3r_1 d^3r_2 \psi_0^*(\mathbf{r}_1, \mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_0(\mathbf{r}_1, \mathbf{r}_2) = \frac{5e^2}{4a_0}. \quad (4.4.15)$$

Therefore

$$E_0^{(0)} + E_0^{(1)} = -\frac{4e^2}{a_0} + \frac{5e^2}{4a_0} = -2.75 \frac{e^2}{a_0}. \quad (4.4.16)$$

The experimental ground-state energy of helium is $-2.904e^2/a_0$. Thus the variational result is closer to the experimental value and is better.

Worked calculation: Gaussian upper bound for an anharmonic oscillator

The helium calculation above illustrates screening, but its Coulomb integrals can obscure the variational principle. A one-dimensional example lets us calculate every integral and every derivative explicitly before handling more elaborate atoms.

Choose dimensionless units $\hbar = m = \omega = 1$ and

$$\hat{H} = -\frac{1}{2} \frac{d^2}{dx^2} + \frac{x^2}{2} + \lambda x^4, \quad \lambda \geq 0. \quad (4.4.17)$$

Take a normalized Gaussian trial state parametrized by its positive width σ ,

$$\psi_\sigma(x) = \frac{1}{\pi^{1/4} \sqrt{\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad |\psi_\sigma(x)|^2 = \frac{e^{-x^2/\sigma^2}}{\sqrt{\pi} \sigma}. \quad (4.4.18)$$

Check normalization by setting $u = x/\sigma$:

$$\int_{-\infty}^{\infty} |\psi_\sigma|^2 dx = \frac{1}{\sqrt{\pi} \sigma} \int_{-\infty}^{\infty} e^{-x^2/\sigma^2} dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2} du = 1. \quad (4.4.19)$$

The Gaussian integral is $\int_{\mathbb{R}} e^{-u^2} du = \sqrt{\pi}$. Differentiate it with respect to a width parameter to obtain

$$\langle x^2 \rangle = \frac{1}{\sqrt{\pi} \sigma} \int_{\mathbb{R}} x^2 e^{-x^2/\sigma^2} dx = \frac{\sigma^2}{2}, \quad \langle x^4 \rangle = \frac{3\sigma^4}{4}. \quad (4.4.20)$$

For the kinetic term, first differentiate $\psi'_\sigma = -x\psi_\sigma/\sigma^2$. Integration by parts, with the boundary term vanishing, gives

$$\begin{aligned} \left\langle -\frac{1}{2} \frac{d^2}{dx^2} \right\rangle &= \frac{1}{2} \int_{\mathbb{R}} |\psi'_\sigma(x)|^2 dx \\ &= \frac{1}{2\sigma^4} \int_{\mathbb{R}} x^2 |\psi_\sigma(x)|^2 dx = \frac{1}{4\sigma^2}. \end{aligned} \quad (4.4.21)$$

Combining these separate contributions,

$$E_{\text{trial}}(\sigma) = \frac{1}{4\sigma^2} + \frac{\sigma^2}{4} + \frac{3\lambda\sigma^4}{4}. \quad (4.4.22)$$

To find the optimum, calculate

$$\frac{dE_{\text{trial}}}{d\sigma} = -\frac{1}{2\sigma^3} + \frac{\sigma}{2} + 3\lambda\sigma^3. \quad (4.4.23)$$

The stationarity condition can be multiplied by $2\sigma^3$ (positive because $\sigma > 0$):

$$\sigma^4 + 6\lambda\sigma^6 = 1. \quad (4.4.24)$$

For $\lambda \geq 0$, the left side is strictly increasing in $\sigma > 0$, ranging from zero to infinity. Thus it has exactly one positive solution. Furthermore,

$$\frac{d^2E_{\text{trial}}}{d\sigma^2} = \frac{3}{2\sigma^4} + \frac{1}{2} + 9\lambda\sigma^2 > 0. \quad (4.4.25)$$

The stationary point is therefore the unique global minimum. The variational bound is $E_0 \leq E_{\text{trial}}(\sigma_*)$. At $\lambda = 0$, the equation gives $\sigma_* = 1$ and $E_{\text{trial}} = 1/2$, recovering the exact harmonic-oscillator ground energy.

Why the inequality always points upward. Expand any *normalized* trial ket in exact normalized energy eigenstates: $|\psi\rangle = \sum_n c_n |n\rangle$, so $\sum_n |c_n|^2 = 1$. Then

$$\langle H \rangle - E_0 = \sum_n |c_n|^2 (E_n - E_0) \geq 0, \quad (4.4.26)$$

because each term is nonnegative. If the trial ket were not normalized, the correct variational object would be $\langle \psi | \hat{H} | \psi \rangle / \langle \psi | \psi \rangle$.

Interpretation and experiment. Increasing λ penalizes large $\langle x^4 \rangle$, so the optimal wavepacket becomes narrower. Its kinetic energy then increases, and the width balances these effects. The after-class [variational energy interactive investigation](#) plots the same formula against an independent finite-basis diagonalization, giving two variational upper bounds that can be compared for convergence.

Interactive investigation IC-04-02. Open Variational energy landscape. After minimizing the Gaussian trial energy, vary its width and potential parameters in the experiment. Locate the numerical minimum and verify that every normalized trial-state energy remains an upper bound to the exact ground energy.

4.5 Strong-Coupling Schrödinger Equation

In 1999, T. D. Lee and collaborators proposed a method for the strongly coupled Schrödinger equation. We give a brief introduction using the Yukawa potential as an example.

The Yukawa potential is

$$V(r) = -g^2 \frac{e^{-\alpha r}}{r}, \quad g > 1 \quad (\text{strong coupling}). \quad (4.5.1)$$

The energy eigenvalue equation is

$$\left(-\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (4.5.2)$$

An exact analytic solution is unavailable. Set $\hbar = 1$ to simplify the calculation. The ground-state wavefunction does not depend on the angular coordinates. Let

$$\psi(r) = e^{-S(r)}. \quad (4.5.3)$$

Expand the eigenvalue E and the exponent function $S(r)$ in powers of $1/g^2$:

$$E = g^4 E_0 + g^2 E_1 + E_2 + \cdots, \quad S = g^2 S_0 + S_1 + g^{-2} S_2 + \cdots. \quad (4.5.4)$$

The purpose of this expansion is to eliminate V from the equations for E_0 and S_0 . Substitute it into the stationary Schrödinger equation, and compare the coefficients of equal powers of g :

$$\begin{aligned} \text{Zeroth-order equation: } (\nabla S_0)^2 &= -2mE_0, \\ \text{First-order equation: } \nabla S_0 \cdot \nabla S_1 &= \frac{1}{2} \nabla^2 S_0 - m \left(\frac{e^{-\alpha r}}{r} + E_1 \right), \\ \text{Second-order equation: } \nabla S_0 \cdot \nabla S_2 &= -\frac{1}{2} (\nabla S_1)^2 + \frac{1}{2} \nabla^2 S_1 - mE_2, \\ &\vdots \end{aligned} \quad (4.5.5)$$

The solution of the zeroth-order equation is

$$S_0(r) = \sqrt{-2mE_0} r. \quad (4.5.6)$$

The positive square root is chosen in view of the boundary condition that the wavefunction $\psi(r)$ remain finite at infinity. Substituting $S_0(r)$ into the first-order equation gives

$$\sqrt{-2mE_0} \frac{dS_1}{dr} = \frac{1}{r} (\sqrt{-2mE_0} - me^{-\alpha r}) - mE_1. \quad (4.5.7)$$

The wavefunction ψ and its first derivative must be continuous. Thus $S_0, S_1, S_2, \dots$ and their first derivatives must be continuous, requiring dS_1/dr to be nonsingular at $r = 0$ (since S_0 and its first derivative are continuous). Therefore

$$\lim_{r \rightarrow 0} \frac{\sqrt{-2mE_0} - me^{-\alpha r}}{r} \quad \text{must be finite.} \quad (4.5.8)$$

It follows that

$$E_0 = -\frac{m}{2}, \quad S_0(r) = mr, \quad (4.5.9)$$

$$S_1(r) = \int_0^r dr' \left[\frac{1 - e^{-\alpha r'}}{r'} - E_1 \right]. \quad (4.5.10)$$

Substitute into the second-order equation and require dS_2/dr to be nonsingular at $r = 0$. This gives

$$E_1 = \alpha. \quad (4.5.11)$$

The successive solution gives the ground-state energy

$$E = -\frac{m}{2} g^4 + \alpha g^2 - \frac{3\alpha^2}{4m} + \frac{\alpha^2}{2m^2} g^{-2} + \cdots. \quad (4.5.12)$$

If $\alpha = 0$, the Yukawa potential reduces to the Coulomb potential. In this case,

$$\boxed{\begin{aligned} E &= -\frac{m}{2} g^4, \\ \psi(r) &= e^{-g^2 S_0(r)} = e^{-mg^2 r}. \end{aligned}} \quad (4.5.13)$$

This is exactly the same as the exact ground-state solution of the hydrogen-atom problem.

4.6 Time-Dependent Perturbation Theory

The Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle. \quad (4.6.1)$$

The Hamiltonian can be written as

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{H}^{(1)}(t), \quad \hat{H}^{(0)} |m\rangle = E_m |m\rangle. \quad (4.6.2)$$

Question. When $t \leq 0$, the Hamiltonian is $\hat{H} = \hat{H}^{(0)}$, independent of time, and the system may be in a stationary state. Assume that at $t = 0$ it is in the stationary state $|n\rangle$. For $t > 0$, $\hat{H}(t)$ depends on time; the system is no longer stably in the stationary state $|n\rangle$, but has certain probabilities of being in other stationary states. At time t the state is

$$|\psi(t)\rangle = \sum_m a_m(t) |m\rangle e^{-iE_m t/\hbar}. \quad (4.6.3)$$

The probability of the stationary state $|m\rangle$ is $|a_m(t)|^2$. For a time-dependent perturbation, the quantity to be found is the probability that, after a macroscopic time ($t \rightarrow \infty$), the system has made a transition from the initial stationary state $|n\rangle$ to another stationary state $|m\rangle$:

$$\lim_{t \rightarrow \infty} W_{n \rightarrow m}(t) = \lim_{t \rightarrow \infty} |a_m(t)|^2. \quad (4.6.4)$$

1) Transition equations

Substitute the expansion of $|\psi(t)\rangle$ into the Schrödinger equation:

$$\begin{aligned} i\hbar \sum_m \dot{a}_m |m\rangle e^{-iE_m t/\hbar} + \sum_m a_m E_m |m\rangle e^{-iE_m t/\hbar} \\ = \sum_m a_m \hat{H}^{(0)} |m\rangle e^{-iE_m t/\hbar} + \sum_m a_m \hat{H}^{(1)}(t) |m\rangle e^{-iE_m t/\hbar}. \end{aligned} \quad (4.6.5)$$

The second term on the left cancels the first term on the right, giving

$$i\hbar \sum_m \dot{a}_m |m\rangle e^{-iE_m t/\hbar} = \sum_m a_m \hat{H}^{(1)}(t) |m\rangle e^{-iE_m t/\hbar}. \quad (4.6.6)$$

Multiply on the left by $\langle k|$:

$$i\hbar \dot{a}_k e^{-iE_k t/\hbar} = \sum_m a_m H_{km}^{(1)}(t) e^{-iE_m t/\hbar}, \quad H_{km}^{(1)}(t) = \langle k | \hat{H}^{(1)}(t) | m \rangle. \quad (4.6.7)$$

Here $H_{km}^{(1)}(t)$ is the matrix element of $\hat{H}^{(1)}(t)$ in the $\hat{H}^{(0)}$ representation. Thus

$$\boxed{\begin{aligned} i\hbar \dot{a}_k &= \sum_m a_m H_{km}^{(1)}(t) e^{i\omega_{km} t}, \\ a_k(0) &= \delta_{kn}, \quad \omega_{km} = \frac{E_k - E_m}{\hbar}. \end{aligned}} \quad (4.6.8)$$

2) Perturbative solution

Write

$$a_k(t) = a_k^{(0)}(t) + a_k^{(1)}(t) + \cdots. \quad (4.6.9)$$

The transition equation becomes

$$i\hbar (\dot{a}_k^{(0)} + \dot{a}_k^{(1)} + \cdots) = \sum_m (a_m^{(0)} + a_m^{(1)} + \cdots) H_{km}^{(1)}(t) e^{i\omega_{km} t}. \quad (4.6.10)$$

The zeroth-order equation is

$$i\hbar\dot{a}_k^{(0)}(t) = 0, \quad a_k^{(0)}(0) = \delta_{kn}, \quad (4.6.11)$$

with solution

$$a_k^{(0)}(t) = \delta_{kn}. \quad (4.6.12)$$

The first-order equation is

$$i\hbar\dot{a}_k^{(1)}(t) = \sum_m a_m^{(0)}(t) H_{km}^{(1)}(t) e^{i\omega_{km}t} = H_{kn}^{(1)}(t) e^{i\omega_{kn}t}. \quad (4.6.13)$$

Its solution is

$$a_k^{(1)}(t) = \frac{1}{i\hbar} \int_0^t H_{kn}^{(1)}(t') e^{i\omega_{kn}t'} dt'. \quad (4.6.14)$$

For a transition from $|n\rangle$ to $|k\rangle$ ($k \neq n$), $a_k(t) \simeq a_k^{(1)}(t)$, so the transition probability is

$$W_{n \rightarrow k}(t) \simeq |a_k^{(1)}(t)|^2. \quad (4.6.15)$$

3) Periodic perturbation

Take

$$\hat{H}^{(1)}(t) = \hat{A} \cos \omega t = \hat{F}(e^{i\omega t} + e^{-i\omega t}), \quad (4.6.16)$$

where $\hat{A}$, $\hat{F}$ are independent of t . The first-order transition amplitude is

$$\begin{aligned} a_m^{(1)}(t) &= \frac{1}{i\hbar} \int_0^t H_{mn}^{(1)}(t') e^{i\omega_{mn}t'} dt' \\ &= \frac{F_{mn}}{i\hbar} \int_0^t [e^{i(\omega_{mn}+\omega)t'} + e^{i(\omega_{mn}-\omega)t'}] dt', \end{aligned} \quad (4.6.17)$$

where ω_{mn} is the transition frequency and ω is the driving frequency. Integrating gives

$$a_m^{(1)}(t) = -\frac{F_{mn}}{\hbar} \left[\frac{e^{i(\omega_{mn}+\omega)t} - 1}{\omega_{mn} + \omega} + \frac{e^{i(\omega_{mn}-\omega)t} - 1}{\omega_{mn} - \omega} \right]. \quad (4.6.18)$$

When $\omega \neq \pm\omega_{mn}$, $a_m^{(1)}(t)$ oscillates with t and does not increase monotonically with time. When $\omega \rightarrow \pm\omega_{mn}$, the nearly resonant term dominates:

$$a_m^{(1)}(t) \simeq -\frac{F_{mn}}{\hbar} \frac{e^{i(\omega_{mn} \mp \omega)t} - 1}{\omega_{mn} \mp \omega}. \quad (4.6.19)$$

The transition probability is

$$W_{n \rightarrow m}(t) = |a_m^{(1)}(t)|^2 \simeq \frac{4|F_{mn}|^2}{\hbar^2} \frac{\sin^2[(\omega_{mn} \mp \omega)t/2]}{(\omega_{mn} \mp \omega)^2}. \quad (4.6.20)$$

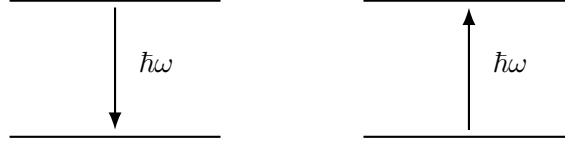
Since

$$\lim_{t \rightarrow \infty} \frac{\sin^2(xt)}{\pi t x^2} = \delta(x), \quad (4.6.21)$$

it follows that

$$\lim_{t \rightarrow \infty} W_{n \rightarrow m}(t) \simeq \frac{2\pi t |F_{mn}|^2}{\hbar^2} \delta(\omega_{mn} \mp \omega) = \frac{2\pi t |F_{mn}|^2}{\hbar} \delta(E_m - E_n \mp \hbar\omega). \quad (4.6.22)$$

It grows linearly with t . The δ function expresses energy conservation, illustrated schematically below:



The transition rate is

$$\frac{dW_{n \rightarrow m}(t)}{dt} = \frac{2\pi}{\hbar^2} |F_{mn}|^2 \delta(\omega_{mn} \mp \omega). \quad (4.6.23)$$

A real perturbation is not a strictly monochromatic field but has a frequency range $\Delta\omega$. The total transition rate is

$$\begin{aligned} \frac{dW_{n \rightarrow m}(t)}{dt} &= \frac{2\pi}{\hbar^2} |F_{mn}|^2 \int_{\Delta\omega} \delta(\omega_{mn} \mp \omega) d\omega \\ &= \begin{cases} \frac{2\pi}{\hbar^2} |F_{mn}|^2, & \Delta\omega \text{ contains } |\omega_{mn}|, \\ 0, & \Delta\omega \text{ does not contain } |\omega_{mn}|. \end{cases} \end{aligned} \quad (4.6.24)$$

Example. At $t = 0$ a hydrogen atom is in its ground state $|100\rangle$. A perturbation

$$\hat{H}^{(1)}(t) = -e\mathcal{E}_0 z \delta(t) \quad (4.6.25)$$

is applied. Find the probability for transitions to all excited states.

Solution. The first-order amplitude is

$$\begin{aligned} a_m^{(1)}(t) &= \frac{1}{i\hbar} \int_0^t H_{mn}^{(1)}(t') e^{i\omega_{mn}t'} dt' \\ &= -\frac{e\mathcal{E}_0 z_{mn}}{i\hbar} \int_0^t e^{i\omega_{mn}t'} \delta(t') dt' \\ &= -\frac{e\mathcal{E}_0 z_{mn}}{i\hbar}. \end{aligned} \quad (4.6.26)$$

Here $|n\rangle = |100\rangle$ is the ground state and $|m\rangle = |nlm\rangle$ ($n \neq 1$) are excited states. The probability of a transition from the ground state to any excited state is

$$W = \sum'_{nlm} \left(\frac{e\mathcal{E}_0}{\hbar} \right)^2 |\langle nlm|z|100\rangle|^2, \quad (4.6.27)$$

where the prime means that the ground state is excluded from the sum. Since

$$z = r \cos \theta, \quad (4.6.28)$$

the corresponding matrix elements are

$$\langle nlm|z|n'l'm'\rangle = \langle nl|r|n'l'\rangle \langle lm|\cos\theta|l'm'\rangle. \quad (4.6.29)$$

The action of $\cos\theta$ is

$$\cos\theta |lm\rangle = \sqrt{\frac{(l+1)^2 - m^2}{(2l+1)(2l+3)}} |l+1, m\rangle + \sqrt{\frac{l^2 - m^2}{(2l-1)(2l+1)}} |l-1, m\rangle. \quad (4.6.30)$$

Therefore, the conditions for $\langle lm|\cos\theta|l'm'\rangle$ to be nonzero (the selection rules for transitions) are

$$\Delta l = l - l' = \pm 1. \quad (4.6.31)$$

Thus $\langle 100|z|100\rangle = 0$, and

$$\begin{aligned} W &= \sum_{nlm} \left(\frac{e\mathcal{E}_0}{\hbar} \right)^2 |\langle nlm|z|100\rangle|^2 \\ &= \left(\frac{e\mathcal{E}_0}{\hbar} \right)^2 \sum_{nlm} \langle 100|z|nlm\rangle \langle nlm|z|100\rangle \\ &= \left(\frac{e\mathcal{E}_0}{\hbar} \right)^2 \langle 100|z^2|100\rangle. \end{aligned} \quad (4.6.32)$$

Since

$$\langle 100|z^2|100\rangle = a_0^2, \quad (4.6.33)$$

where a_0 is the Bohr radius, the result is

$$W = \left(\frac{e\mathcal{E}_0 a_0}{\hbar} \right)^2. \quad (4.6.34)$$

The probability for the hydrogen atom to remain in the ground state is therefore

$$1 - W = 1 - \left(\frac{e\mathcal{E}_0 a_0}{\hbar} \right)^2. \quad (4.6.35)$$

Interactive investigation IC-04-03. Open Driven two-level resonance. After the periodic-drive transition calculation, scan the drive frequency and duration. Examine resonant versus off-resonant transition probabilities and the finite-time linewidth; compare weak-driving predictions with the exact two-level model.

4.7 Interaction Picture and Perturbation Theory

1) Interaction picture

First review the Schrödinger picture and the Heisenberg picture.

Schrödinger picture: The state depends on time t :

$$|\alpha, t\rangle_S = \hat{U}(t) |\alpha, 0\rangle_S, \quad \hat{U}(t) = e^{-i\hat{H}t/\hbar}. \quad (4.7.1)$$

A mechanical quantity does not depend explicitly on time: $\hat{A}_S$.

The dynamical equation is

$$i\hbar \frac{\partial}{\partial t} |\alpha, t\rangle_S = \hat{H} |\alpha, t\rangle_S. \quad (4.7.2)$$

Heisenberg picture: The state does not depend on time: $|\alpha\rangle_H$.

The mechanical quantities depend on time:

$$\hat{A}_H(t) = \hat{U}^\dagger(t) \hat{A}_H(0) \hat{U}(t). \quad (4.7.3)$$

The dynamical equation is

$$\frac{d\hat{A}_H}{dt} = \frac{1}{i\hbar} [\hat{A}_H, \hat{H}]. \quad (4.7.4)$$

The Hamiltonian $\hat{H}$ is the same in both pictures.

Now introduce the interaction picture. In the Schrödinger picture, separate the Hamiltonian into two parts:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}. \quad (4.7.5)$$

Define the relation between states in the interaction and Schrödinger pictures by

$$|\alpha, t\rangle_I = e^{i\hat{H}^{(0)}t/\hbar} |\alpha, t\rangle_S. \quad (4.7.6)$$

The expectation value of a mechanical quantity $\hat{O}$ is independent of the picture:

$$\begin{aligned} {}_I\langle\alpha, t|\hat{O}_I|\alpha, t\rangle_I &= {}_S\langle\alpha, t|\hat{O}_S|\alpha, t\rangle_S \\ &= {}_I\langle\alpha, t|e^{i\hat{H}^{(0)}t/\hbar}\hat{O}_Se^{-i\hat{H}^{(0)}t/\hbar}|\alpha, t\rangle_I. \end{aligned} \quad (4.7.7)$$

Hence mechanical quantities are related by

$$\boxed{\hat{O}_I = e^{i\hat{H}^{(0)}t/\hbar}\hat{O}_Se^{-i\hat{H}^{(0)}t/\hbar}.} \quad (4.7.8)$$

Note: $\hat{H}^{(0)}$ is the same in the Schrödinger and interaction pictures, but $\hat{H}^{(1)}$ is not.

In the interaction picture, both states and mechanical quantities depend on time. The equation of motion for states is

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\alpha, t\rangle_I &= -\hat{H}^{(0)}e^{i\hat{H}^{(0)}t/\hbar}|\alpha, t\rangle_S + e^{i\hat{H}^{(0)}t/\hbar}i\hbar\frac{\partial}{\partial t}|\alpha, t\rangle_S \\ &= -e^{i\hat{H}^{(0)}t/\hbar}\hat{H}^{(0)}|\alpha, t\rangle_S + e^{i\hat{H}^{(0)}t/\hbar}\hat{H}|\alpha, t\rangle_S \\ &= e^{i\hat{H}^{(0)}t/\hbar}\hat{H}_S^{(1)}|\alpha, t\rangle_S \\ &= e^{i\hat{H}^{(0)}t/\hbar}\hat{H}_S^{(1)}e^{-i\hat{H}^{(0)}t/\hbar}e^{i\hat{H}^{(0)}t/\hbar}|\alpha, t\rangle_S \\ &= \hat{H}_I^{(1)}|\alpha, t\rangle_I. \end{aligned} \quad (4.7.9)$$

The equation of motion for an operator in the interaction picture is

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{O}_I &= -\hat{H}^{(0)}\hat{O}_I + \hat{O}_I\hat{H}^{(0)} \\ &= [\hat{O}_I, \hat{H}^{(0)}]. \end{aligned} \quad (4.7.10)$$

Although both states and mechanical quantities depend on time in the interaction picture (only states depend on time in the Schrödinger picture, and only mechanical quantities depend on time in the Heisenberg picture), the time evolution of the states depends only on $\hat{H}^{(1)}$, while the time evolution of the mechanical quantities depends only on $\hat{H}^{(0)}$. This is very useful in perturbative expansions.

2) Perturbative expansion

The time evolution of a state in the interaction picture is

$$|\alpha, t\rangle_I = \hat{U}(t, t_0)|\alpha, t_0\rangle_I. \quad (4.7.11)$$

The equation of motion for the interaction-picture state gives

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t, t_0) = \hat{H}_I^{(1)}(t)\hat{U}(t, t_0), \quad (4.7.12)$$

subject to

$$\hat{U}(t_0, t_0) = \mathbf{1}. \quad (4.7.13)$$

Its integral solution is

$$\hat{U}(t, t_0) = \mathbf{1} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}_I^{(1)}(t')\hat{U}(t', t_0). \quad (4.7.14)$$

By iteration,

$$\begin{aligned}\hat{U}(t, t_0) &= \mathbf{1} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}_I^{(1)}(t') \\ &\quad + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') + \dots\end{aligned}\quad (4.7.15)$$

For the third term on the right,

$$\begin{aligned}&\int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') \\ &= \frac{1}{2!} \left[\int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') + \int_{t_0}^t dt'' \int_{t_0}^{t''} dt' \hat{H}_I^{(1)}(t'') \hat{H}_I^{(1)}(t') \right] \\ &= \frac{1}{2!} \int_{t_0}^t dt' \int_{t_0}^t dt'' \mathcal{T} \left(\hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') \right).\end{aligned}\quad (4.7.16)$$

The second term on the right of the first equality is obtained by interchanging the integration variables. The time-ordered product $\mathcal{T}$ is defined by

$$\begin{aligned}\mathcal{T} \left(\hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') \right) \\ = \hat{H}_I^{(1)}(t') \hat{H}_I^{(1)}(t'') \theta(t' - t'') + \hat{H}_I^{(1)}(t'') \hat{H}_I^{(1)}(t') \theta(t'' - t').\end{aligned}\quad (4.7.17)$$

The operator corresponding to the later time always stands to the left.

Thus the solution for the time-evolution operator can be written as

$$\boxed{\begin{aligned}\hat{U}(t, t_0) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n \\ &\quad \times \mathcal{T} \left(\hat{H}_I^{(1)}(t_1) \hat{H}_I^{(1)}(t_2) \cdots \hat{H}_I^{(1)}(t_n) \right).\end{aligned}}\quad (4.7.18)$$

This is the perturbative expansion of the evolution operator for states in the interaction picture, and it is the foundation of perturbation theory in quantum mechanics and quantum field theory.

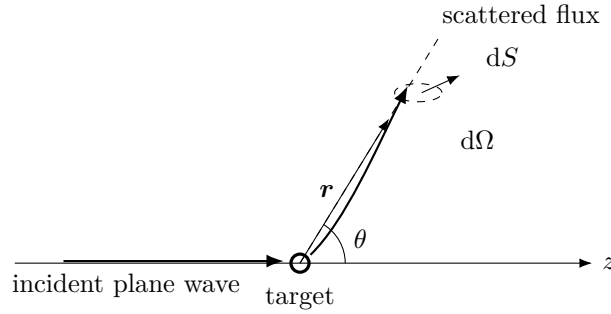
Chapter 5

Scattering Theory

5.1 General Description

The perturbation method discussed above is mainly suitable for discrete spectra, that is, for the case of separated energy levels. How should one solve continuous-spectrum problems? A typical physical problem involving a continuous spectrum is scattering.

1) Differential scattering cross section



An incident plane wave is changed into a spherical wave after scattering. The problem is to find the probability per unit time that an incident particle is scattered into a unit solid angle in the direction Ω , denoted by $\sigma(\theta, \varphi)$. The quantity $\sigma(\theta, \varphi)$ depends on the nature of the target and of the interaction. By measuring $\sigma(\theta, \varphi)$, we may learn about the internal structure of the target or discover new particles.

Scattering is divided into elastic and inelastic scattering:

- **Elastic scattering:** The energy of the particle does not change before and after scattering; only its direction of motion changes.
- **Inelastic scattering:** After scattering, the properties of the particle change; new particles may even be created.

We consider only elastic scattering, in the rest frame of the target particle. Let the incident plane wave be

$$\psi_1(\mathbf{r}) = Ae^{ikz}. \quad (5.1.1)$$

The probability-current density of the incident particle (probability crossing unit area per unit time) is

$$J_z = \frac{i\hbar}{2\mu} \left(\psi_1 \frac{\partial \psi_1^*}{\partial z} - \psi_1^* \frac{\partial \psi_1}{\partial z} \right) = \frac{\hbar k}{\mu} |A|^2. \quad (5.1.2)$$

The probability per unit time that these incident particles are scattered into the direction $d\Omega$ is

$$dN = J_z \sigma(\theta, \varphi) d\Omega. \quad (5.1.3)$$

Analyze the dimension of σ : since $[dN] = T^{-1}$, $[J_z] = L^{-2}T^{-1}$ and $[d\Omega] = 1$, it follows that

$$[\sigma] = L^2. \quad (5.1.4)$$

It has the dimension of area and therefore $\sigma(\theta, \varphi)$ is called the *differential scattering cross section*.

2) General method for calculating $\sigma(\theta, \varphi)$

As $r \rightarrow \infty$, $V(\mathbf{r}) \rightarrow 0$, and $\psi(\mathbf{r})$ consists of two parts: the incident plane wave propagating in the z direction without being affected by the interaction, and an outgoing scattered spherical wave propagating in the $\mathbf{r}$ direction,

$$\psi(\mathbf{r}) \xrightarrow{r \rightarrow \infty} \psi_1(\mathbf{r}) + \psi_2(\mathbf{r}) = Ae^{ikz} + Af(\theta, \varphi) \frac{e^{ikr}}{r}. \quad (5.1.5)$$

Note: (1) The spherical wave depends on θ and φ ; $f(\theta, \varphi)$ can be chosen so that ψ_2 and ψ_1 have the same constant A . The specific form of $f(\theta, \varphi)$ is determined by the interaction. (2) For elastic scattering, the energy and the magnitude of k remain unchanged, while its direction changes.

The probability-current density of the scattered particles is

$$J_r = \frac{i\hbar}{2\mu} \left(\psi_2 \frac{\partial \psi_2^*}{\partial r} - \psi_2^* \frac{\partial \psi_2}{\partial r} \right) = \frac{\hbar k}{\mu} |A|^2 \frac{|f(\theta, \varphi)|^2}{r^2} = J_z \frac{|f(\theta, \varphi)|^2}{r^2}. \quad (5.1.6)$$

The probability per unit time scattered into $d\Omega$ is

$$dN = J_r dS = J_r r^2 d\Omega = J_z |f(\theta, \varphi)|^2 d\Omega. \quad (5.1.7)$$

Comparing with $dN = J_z \sigma(\theta, \varphi) d\Omega$, we obtain

$$\boxed{\sigma(\theta, \varphi) = |f(\theta, \varphi)|^2}. \quad (5.1.8)$$

Here $f(\theta, \varphi)$ is called the *scattering amplitude*.

Conclusion: The general method for calculating the differential cross section $\sigma(\theta, \varphi)$ is:

1. Solve the specific Schrödinger equation to obtain $\psi(\mathbf{r})$.
2. Compare its asymptotic form as $r \rightarrow \infty$ with the standard asymptotic solution

$$\psi(\mathbf{r}) \xrightarrow{r \rightarrow \infty} Ae^{ikz} + Af(\theta, \varphi) \frac{e^{ikr}}{r} \quad (5.1.9)$$

to find $f(\theta, \varphi)$.

3. Obtain $\sigma(\theta, \varphi) = |f(\theta, \varphi)|^2$.

5.2 Green-Function Method

1) Integral equation

The stationary Schrödinger equation is

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (5.2.1)$$

In elastic scattering the energy is conserved. E is the known incident energy, and only $\psi(\mathbf{r})$ needs to be determined. (In a bound-state problem, the bound-state condition is used to determine the allowed values of E .) Set

$$k^2 = \frac{2mE}{\hbar^2}, \quad U(\mathbf{r}) = \frac{2m}{\hbar^2}V(\mathbf{r}). \quad (5.2.2)$$

Then

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = U(\mathbf{r})\psi(\mathbf{r}), \quad (5.2.3)$$

which has a continuous source (potential) $U(\mathbf{r})$.

The idea of the Green-function method is to solve first the stationary equation corresponding to a point source $\delta^{(3)}(\mathbf{r} - \mathbf{r}')$,

$$(\nabla_{\mathbf{r}}^2 + k^2)G(\mathbf{r}, \mathbf{r}') = \delta^{(3)}(\mathbf{r} - \mathbf{r}'), \quad (5.2.4)$$

thus obtaining the Green function $G(\mathbf{r}, \mathbf{r}')$, and then solving the stationary equation

$$\begin{aligned} (\nabla^2 + k^2)\psi(\mathbf{r}) &= U(\mathbf{r})\psi(\mathbf{r}) \\ &= \int d^3r' U(\mathbf{r}')\psi(\mathbf{r}')\delta^{(3)}(\mathbf{r} - \mathbf{r}') \\ &= \int d^3r' U(\mathbf{r}')\psi(\mathbf{r}')(\nabla_{\mathbf{r}}^2 + k^2)G(\mathbf{r}, \mathbf{r}'). \end{aligned} \quad (5.2.5)$$

Therefore one solution of the stationary equation is

$$\psi(\mathbf{r}) = \int d^3r' U(\mathbf{r}')\psi(\mathbf{r}')G(\mathbf{r}, \mathbf{r}'). \quad (5.2.6)$$

Taking into account the solution to the sourceless equation

$$(\nabla^2 + k^2)\psi^{(0)}(\mathbf{r}) = 0, \quad \psi^{(0)}(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}}, \quad (5.2.7)$$

the general solution of the stationary equation is

$$\boxed{\psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} + \int d^3r' U(\mathbf{r}')\psi(\mathbf{r}')G(\mathbf{r}, \mathbf{r}')} \quad (5.2.8)$$

This satisfies the requirement

$$\lim_{V \rightarrow 0} \psi(\mathbf{r}) = \psi^{(0)}(\mathbf{r}). \quad (5.2.9)$$

It is an integral equation for $\psi(\mathbf{r})$, completely equivalent to the differential form of the stationary equation.

The integral equation has advantages: (1) The boundary condition is already included in the equation (the incident plane wave $Ae^{i\mathbf{k}\cdot\mathbf{r}}$); (2) an iterative method may be used to solve it.

2) Green functions

Perform a Fourier transformation,

$$\begin{aligned} G(\mathbf{r}, \mathbf{r}') &= \frac{1}{(2\pi)^3} \int d^3k' g(\mathbf{k}', \mathbf{r}') e^{i\mathbf{k}'\cdot\mathbf{r}}, \\ \delta^{(3)}(\mathbf{r} - \mathbf{r}') &= \frac{1}{(2\pi)^3} \int d^3k' e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}. \end{aligned} \quad (5.2.10)$$

Substitution into the equation satisfied by $G(\mathbf{r}, \mathbf{r}')$ gives

$$g(\mathbf{k}', \mathbf{r}') = \frac{e^{-i\mathbf{k}'\cdot\mathbf{r}'}}{k'^2 - k'^2}. \quad (5.2.11)$$

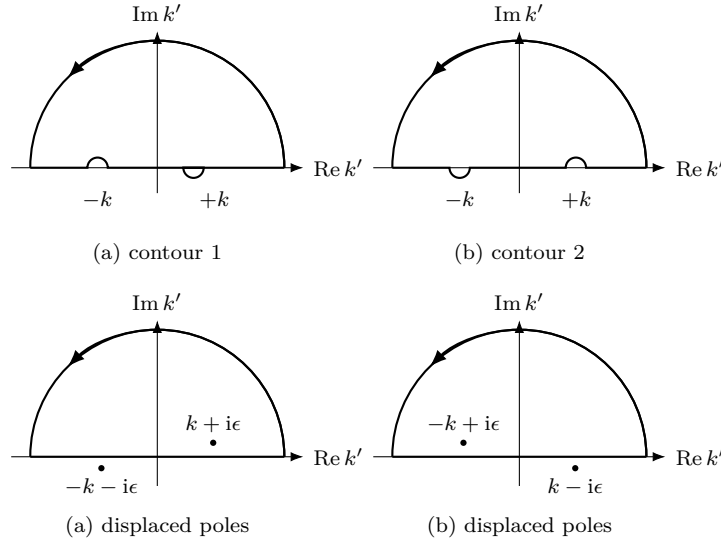
Hence

$$G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r} - \mathbf{r}') = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d^3k' \frac{e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}')}}{k'^2 - k^2}. \quad (5.2.12)$$

Take $\mathbf{r} - \mathbf{r}'$ as the polar axis. After angular integration, with $R = |\mathbf{r} - \mathbf{r}'|$,

$$\begin{aligned} G(\mathbf{r} - \mathbf{r}') &= \frac{1}{4\pi^2 R} \int_{-\infty}^{\infty} dk' \frac{k' \sin(k'R)}{k'^2 - k^2} \\ &= \frac{1}{4\pi^2 i R} \int_{-\infty}^{\infty} dk' \frac{k' e^{ik'R}}{k'^2 - k^2} \\ &= -\frac{1}{8\pi^2 i R} \int_{-\infty}^{\infty} dk' \left(\frac{e^{ik'R}}{k' - k} + \frac{e^{ik'R}}{k' + k} \right). \end{aligned} \quad (5.2.13)$$

The part of the integrand involving $\cos(k'R)$ is an odd function of k' and integrates to zero. The integral has two poles, $k' = \pm k$, and is treated by complex-variable methods.



In the complex k' plane, consider contour 1 closed in the upper half-plane (left). The first pole contributes and the second does not. On the semicircle,

$$e^{ik'R} = e^{i \operatorname{Re} k' R} e^{-\operatorname{Im} k' R}, \quad (5.2.14)$$

since $\operatorname{Im} k' > 0$ in the upper half-plane, this tends to zero as $R \rightarrow \infty$. The residue theorem gives

$$G^+(\mathbf{r} - \mathbf{r}') = -\frac{1}{4\pi} \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}, \quad (5.2.15)$$

called the *retarded Green function*. For contour 2 in the upper half-plane (right), the first pole does not contribute and the second does:

$$G^-(\mathbf{r} - \mathbf{r}') = -\frac{1}{4\pi} \frac{e^{-ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}, \quad (5.2.16)$$

called the *advanced Green function*.

A contour containing both poles gives what is called the causal Green function. The contours above each make small semicircular indentations around the poles. Equivalently, one may displace the two poles slightly upward or downward and use a straight integration path (bottom diagrams in the source figure).

Retarded, advanced, and causal Green functions are all solutions of

$$(\nabla_{\mathbf{r}}^2 + k^2)G(\mathbf{r}, \mathbf{r}') = \delta^{(3)}(\mathbf{r} - \mathbf{r}'). \quad (5.2.17)$$

Substituting the retarded and advanced Green functions into the integral equation for $\psi(\mathbf{r})$, for example, gives

$$\psi^\pm(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi} \int d^3r' U(\mathbf{r}') \psi^\pm(\mathbf{r}') \frac{e^{\pm ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}. \quad (5.2.18)$$

5.3 Lippmann–Schwinger Equation

1) Lippmann–Schwinger equation in Hilbert space

Let

$$\hat{H} = \hat{H}_0 + \hat{V}, \quad \hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m}. \quad (5.3.1)$$

The incident plane wave is $|\varphi\rangle$, with

$$\hat{H}_0 |\varphi\rangle = E |\varphi\rangle, \quad E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}. \quad (5.3.2)$$

The scattered wave is $|\psi\rangle$. In elastic scattering the energy is unchanged:

$$\hat{H} |\psi\rangle = (\hat{H}_0 + \hat{V}) |\psi\rangle = E |\psi\rangle, \quad (E - \hat{H}_0) |\psi\rangle = \hat{V} |\psi\rangle. \quad (5.3.3)$$

The formal solution satisfying the initial condition $\lim_{V \rightarrow 0} |\psi\rangle = |\varphi\rangle$ is

$$|\psi\rangle = |\varphi\rangle + \frac{1}{E - \hat{H}_0} \hat{V} |\psi\rangle. \quad (5.3.4)$$

However, the operator $(E - \hat{H}_0)^{-1}$ diverges when acting on eigenstates of $\hat{H}_0$. To avoid the divergence, extend the energy into the complex plane and write the formal solution as

$$|\psi^\pm\rangle = |\varphi\rangle + \frac{1}{E - \hat{H}_0 \pm i\epsilon} \hat{V} |\psi^\pm\rangle, \quad \epsilon \rightarrow 0^+. \quad (5.3.5)$$

This is the Lippmann–Schwinger equation in Hilbert space.

Enter the coordinate representation:

$$\langle \mathbf{r} | \psi^\pm \rangle = \langle \mathbf{r} | \varphi \rangle + \int d^3r' \left\langle \mathbf{r} \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{r}' \right\rangle \langle \mathbf{r}' | \hat{V} | \psi^\pm \rangle. \quad (5.3.6)$$

Because

$$\begin{aligned} \left\langle \mathbf{r} \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{r}' \right\rangle &= \int d^3p' d^3p'' \langle \mathbf{r} | \mathbf{p}' \rangle \left\langle \mathbf{p}' \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{p}'' \right\rangle \langle \mathbf{p}'' | \mathbf{r}' \rangle, \\ \left\langle \mathbf{p}' \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{p}'' \right\rangle &= \frac{\delta^{(3)}(\mathbf{p}' - \mathbf{p}'')}{E - E' \pm i\epsilon}, \\ \langle \mathbf{r} | \mathbf{p}' \rangle &= \frac{e^{i\mathbf{p}'\cdot\mathbf{r}/\hbar}}{(2\pi\hbar)^{3/2}}, \quad \langle \mathbf{p}'' | \mathbf{r}' \rangle = \frac{e^{-i\mathbf{p}''\cdot\mathbf{r}'/\hbar}}{(2\pi\hbar)^{3/2}}, \end{aligned} \quad (5.3.7)$$

we have

$$\begin{aligned} \left\langle \mathbf{r} \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{r}' \right\rangle &= \int \frac{d^3p'}{(2\pi\hbar)^3} \frac{e^{i\mathbf{p}'\cdot(\mathbf{r}-\mathbf{r}')/\hbar}}{E - E' \pm i\epsilon} \\ &= \int \frac{d^3k'}{(2\pi)^3} \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{E - E' \pm i\epsilon}. \end{aligned} \quad (5.3.8)$$

By the same contour deformation used to define the Green function, moving a pole slightly upward or downward makes the integration path straight. With $E = \hbar^2 k^2 / (2m)$,

$$G^\pm(\mathbf{r} - \mathbf{r}') = \int \frac{d^3 k'}{(2\pi)^3} \frac{e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}')}}{k^2 - k'^2 \pm i0},$$

$$\left\langle \mathbf{r} \left| \frac{1}{E - \hat{H}_0 \pm i\epsilon} \right| \mathbf{r}' \right\rangle = \frac{2m}{\hbar^2} G^\pm(\mathbf{r} - \mathbf{r}'). \quad (5.3.9)$$

Hence the Lippmann–Schwinger equation in the coordinate representation becomes

$$\langle \mathbf{r} | \psi^\pm \rangle = \langle \mathbf{r} | \varphi \rangle + \frac{2m}{\hbar^2} \int d^3 r' G^\pm(\mathbf{r} - \mathbf{r}') \langle \mathbf{r}' | \hat{V} | \psi^\pm \rangle. \quad (5.3.10)$$

Since

$$\begin{aligned} \langle \mathbf{r}' | \hat{V} | \psi^\pm \rangle &= \int d^3 r'' \langle \mathbf{r}' | \hat{V} | \mathbf{r}'' \rangle \langle \mathbf{r}'' | \psi^\pm \rangle \\ &= \int d^3 r'' V(\mathbf{r}') \delta^{(3)}(\mathbf{r}' - \mathbf{r}'') \langle \mathbf{r}'' | \psi^\pm \rangle = V(\mathbf{r}') \langle \mathbf{r}' | \psi^\pm \rangle, \end{aligned} \quad (5.3.11)$$

we obtain

$$\boxed{\langle \mathbf{r} | \psi^\pm \rangle = \langle \mathbf{r} | \varphi \rangle + \int d^3 r' U(\mathbf{r}') G^\pm(\mathbf{r} - \mathbf{r}') \langle \mathbf{r}' | \psi^\pm \rangle.} \quad (5.3.12)$$

This is precisely the integral equation corresponding to the coordinate-space Schrödinger equation obtained in the preceding section by the Green-function method. It is the coordinate representation of the Lippmann–Schwinger equation.

2) Dyson equation

In the general Lippmann–Schwinger equation,

$$|\psi^\pm\rangle = |\varphi\rangle + \frac{1}{E - \hat{H}_0 \pm i\epsilon} \hat{V} |\psi^\pm\rangle, \quad (5.3.13)$$

let

$$\hat{G}_0^\pm = \frac{1}{E - \hat{H}_0 \pm i\epsilon}. \quad (5.3.14)$$

The coordinate representation $\langle \mathbf{r} | \hat{G}_0^\pm | \mathbf{r}' \rangle$ is a Green function, and thus $\hat{G}_0^\pm$ is called the *free Green operator*, related to $\hat{H}_0$.

The Lippmann–Schwinger equation can also be rewritten as

$$\begin{aligned} (E - \hat{H}_0 \pm i\epsilon) |\psi^\pm\rangle &= (E - \hat{H}_0 \pm i\epsilon) |\varphi\rangle + \hat{V} |\psi^\pm\rangle, \\ (E - \hat{H} \pm i\epsilon) |\psi^\pm\rangle &= (E - \hat{H} \pm i\epsilon) |\varphi\rangle + \hat{V} |\varphi\rangle, \\ |\psi^\pm\rangle &= |\varphi\rangle + \frac{1}{E - \hat{H} \pm i\epsilon} \hat{V} |\varphi\rangle. \end{aligned} \quad (5.3.15)$$

Note that $|\varphi\rangle$ is an eigenstate of $\hat{H}_0$, not of $\hat{H}$. Define

$$\hat{G}^\pm = \frac{1}{E - \hat{H} \pm i\epsilon} \quad (5.3.16)$$

as the *full Green operator*, related to $\hat{H}$.

From

$$\begin{aligned} (E - \hat{H} \pm i\epsilon) \hat{G}^\pm &= \mathbf{1}, \\ (E - \hat{H}_0 \pm i\epsilon) \hat{G}^\pm - \hat{V} \hat{G}^\pm &= \mathbf{1}, \end{aligned} \quad (5.3.17)$$

we have

$$\hat{G}^{\pm} = \hat{G}_0^{\pm} + \hat{G}_0^{\pm} \hat{V} \hat{G}^{\pm}. \quad (5.3.18)$$

This is the *Dyson equation*, which describes the relation between the free and full Green operators. It has important applications in quantum statistical mechanics and quantum field theory.

3) Scattering amplitude

Consider the state $|\psi^+\rangle$ associated with the retarded Green function G^+ . The coordinate-space Lippmann–Schwinger equation gives

$$\psi^+(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi} \int d^3r' U(\mathbf{r}') \psi^+(\mathbf{r}') \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}. \quad (5.3.19)$$

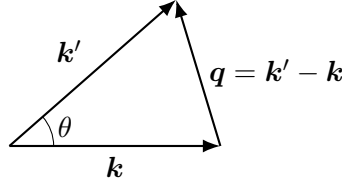
Because the effective region of the potential $V(\mathbf{r}')$ is finite, when $r \rightarrow \infty$,

$$\begin{aligned} |\mathbf{r}'| &\ll r, \\ |\mathbf{r}-\mathbf{r}'| &= \sqrt{(\mathbf{r}-\mathbf{r}')^2} = \sqrt{r^2 + r'^2 - 2\mathbf{r}\cdot\mathbf{r}'} \\ &\simeq r \left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}'}{r^2}\right)^{1/2} \simeq r - \frac{\mathbf{r}\cdot\mathbf{r}'}{r}. \end{aligned} \quad (5.3.20)$$

Consequently,

$$\psi^+(\mathbf{r}) \xrightarrow{r \rightarrow \infty} Ae^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi} \frac{2m}{\hbar^2} \frac{e^{ikr}}{r} \int d^3r' e^{-i\mathbf{k}'\cdot\mathbf{r}'} V(\mathbf{r}') \psi^+(\mathbf{r}'). \quad (5.3.21)$$

In the phase, $|\mathbf{r}-\mathbf{r}'|$ is retained to first order; in the denominator only the zeroth-order term is kept. The scattered wavevector is $\mathbf{k}' = k\mathbf{r}/r$, with $|\mathbf{k}'| = k$.



Comparing with the standard asymptotic solution

$$\psi(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} + Af(\theta, \varphi) \frac{e^{ikr}}{r}, \quad (5.3.22)$$

we obtain the scattering amplitude

$$\begin{aligned} f(\theta, \varphi) &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \int d^3r' \frac{e^{-i\mathbf{k}'\cdot\mathbf{r}'}}{(2\pi)^{3/2}} V(\mathbf{r}') \psi^+(\mathbf{r}') \\ &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \int d^3r' \langle \mathbf{k}' | \mathbf{r}' \rangle V(\mathbf{r}') \langle \mathbf{r}' | \psi^+ \rangle \\ &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \langle \mathbf{k}' | \hat{V} | \psi^+ \rangle \equiv f(\mathbf{k}, \mathbf{k}'). \end{aligned} \quad (5.3.23)$$

Here we have taken

$$\langle \mathbf{r} | \varphi \rangle = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(2\pi)^{3/2}}. \quad (5.3.24)$$

Define the transition operator $\hat{T}$ by

$$\hat{V} |\psi^+\rangle \equiv \hat{T} |\varphi\rangle. \quad (5.3.25)$$

Then the scattering amplitude is

$$f(\mathbf{k}, \mathbf{k}') = -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \langle \mathbf{k}' | \hat{T} | \mathbf{k} \rangle. \quad (5.3.26)$$

The physical meaning of the transition matrix element $\langle \mathbf{k}' | \hat{T} | \mathbf{k} \rangle$ is clear: it is the transition amplitude from the plane wave $|\mathbf{k}\rangle$ to the free state $|\mathbf{k}'\rangle$.

If one considers the advanced Green function G^- , the resulting $\psi^-(\mathbf{r})$ is a superposition of a plane wave and an inward-propagating spherical wave, rather than a solution of the outgoing scattering problem. Advanced and causal Green functions are relevant to time evolution in dynamical systems.

4) Optical theorem

The forward scattering amplitude is

$$f(0, \varphi) = f(\mathbf{k}, \mathbf{k}) = -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \langle \mathbf{k} | \hat{V} | \psi^+ \rangle. \quad (5.3.27)$$

From the Lippmann–Schwinger equation (with $|\varphi\rangle = |\mathbf{k}\rangle$),

$$\begin{aligned} |\psi^+\rangle &= |\varphi\rangle + \frac{1}{E - \hat{H}_0 + i\epsilon} \hat{V} |\psi^+\rangle, \\ \langle \mathbf{k} | &= \langle \psi^+ | - \langle \psi^+ | \hat{V} \frac{1}{E - \hat{H}_0 - i\epsilon}, \end{aligned} \quad (5.3.28)$$

so that

$$\begin{aligned} \langle \mathbf{k} | \hat{V} | \psi^+ \rangle &= \left(\langle \psi^+ | - \langle \psi^+ | \hat{V} \frac{1}{E - \hat{H}_0 - i\epsilon} \right) \hat{V} | \psi^+ \rangle \\ &= \langle \psi^+ | \hat{V} | \psi^+ \rangle - \langle \psi^+ | \hat{V} \frac{1}{E - \hat{H}_0 - i\epsilon} \hat{V} | \psi^+ \rangle. \end{aligned} \quad (5.3.29)$$

Using

$$\frac{1}{E - \hat{H}_0 - i\epsilon} = \text{Pr} \frac{1}{E - \hat{H}_0} + i\pi \delta(E - \hat{H}_0), \quad (5.3.30)$$

and noting that $\hat{V}$ and $\hat{V}(\text{Pr}(E - \hat{H}_0)^{-1})\hat{V}$ are Hermitian operators, whose expectation values in $|\psi^+\rangle$ are real, we find

$$\begin{aligned} \text{Im} \langle \mathbf{k} | \hat{V} | \psi^+ \rangle &= -\pi \langle \psi^+ | \hat{V} \delta(E - \hat{H}_0) \hat{V} | \psi^+ \rangle \\ &= -\pi \int d^3k' \langle \psi^+ | \hat{V} | \mathbf{k}' \rangle \langle \mathbf{k}' | \hat{V} | \psi^+ \rangle \delta(E - \hbar^2 k'^2 / 2m) \\ &= -\pi \int d^3k' \left| \langle \mathbf{k}' | \hat{V} | \psi^+ \rangle \right|^2 \delta\left(E - \frac{\hbar^2 k'^2}{2m}\right) \\ &= -\frac{\pi m k}{\hbar^2} \int d\Omega' \left| \langle \mathbf{k}' | \hat{V} | \psi^+ \rangle \right|^2. \end{aligned} \quad (5.3.31)$$

Because

$$\sigma(\mathbf{k}', \mathbf{k}) = |f(\mathbf{k}', \mathbf{k})|^2 = \frac{(2\pi)^6}{16\pi^2} \frac{4m^2}{\hbar^4} \left| \langle \mathbf{k}' | \hat{V} | \psi^+ \rangle \right|^2, \quad \sigma_{\text{tot}} = \int d\Omega \sigma(\theta, \varphi), \quad (5.3.32)$$

we obtain

$$\text{Im} \langle \mathbf{k} | \hat{V} | \psi^+ \rangle = -\frac{\hbar^2 k}{16\pi^3 m} \int d\Omega' \sigma(\theta, \varphi) = -\frac{\hbar^2 k}{16\pi^3 m} \sigma_{\text{tot}}. \quad (5.3.33)$$

Therefore

$$\begin{aligned} \text{Im} f(\theta = 0) &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \text{Im} \langle \mathbf{k} | \hat{V} | \psi^+ \rangle \\ &= \frac{k}{4\pi} \sigma_{\text{tot}}. \end{aligned} \quad (5.3.34)$$

This relation between the forward scattering amplitude and the total scattering cross section is called the *optical theorem*:

$$\boxed{\text{Im} f(0) = \frac{k}{4\pi} \sigma_{\text{tot}}}. \quad (5.3.35)$$

5.4 Born Series

1) First Born approximation

For the integral equation and scattering amplitude,

$$\psi^+(\mathbf{r}) \xrightarrow{r \rightarrow \infty} A e^{i\mathbf{k} \cdot \mathbf{r}} - \frac{1}{4\pi} \frac{2m}{\hbar^2} \frac{e^{ikr}}{r} \int d^3 r' e^{-i\mathbf{k}' \cdot \mathbf{r}'} V(\mathbf{r}') \psi^+(\mathbf{r}'), \quad (5.4.1)$$

$$f(\theta, \varphi) = -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \int d^3 r' \frac{e^{-i\mathbf{k}' \cdot \mathbf{r}'}}{(2\pi)^{3/2}} V(\mathbf{r}') \psi^+(\mathbf{r}'), \quad (5.4.2)$$

if the potential $V(\mathbf{r})$ is weak (incident energy $E \gg V$), one may solve iteratively. The zeroth-order approximation is

$$\psi^{(0)}(\mathbf{r}) = \frac{e^{i\mathbf{k} \cdot \mathbf{r}}}{(2\pi)^{3/2}}. \quad (5.4.3)$$

The first Born approximation then gives

$$\begin{aligned} f(\theta, \varphi) &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \int d^3 r' \frac{e^{-i\mathbf{k}' \cdot \mathbf{r}'}}{(2\pi)^{3/2}} V(\mathbf{r}') \frac{e^{i\mathbf{k} \cdot \mathbf{r}'}}{(2\pi)^{3/2}} \\ &= -\frac{1}{4\pi} \int d^3 r' e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}'} U(\mathbf{r}'). \end{aligned} \quad (5.4.4)$$

For low-energy scattering (still satisfying $E \gg V$),

$$\begin{aligned} q &= |\mathbf{k} - \mathbf{k}'| = \sqrt{k^2 + k'^2 - 2kk' \cos \theta} \\ &= \sqrt{2k^2(1 - \cos \theta)} = \sqrt{4k^2 \sin^2 \frac{\theta}{2}} = 2k \sin \frac{\theta}{2}. \end{aligned} \quad (5.4.5)$$

If q is small, the phase may be regarded as constant over the effective region of the interaction, and contributes no change to the cross section. Hence

$$\boxed{f(\theta, \varphi) = -\frac{1}{4\pi} \int d^3 r' U(\mathbf{r}').} \quad (5.4.6)$$

Example. Consider low-energy scattering by

$$V(r) = \begin{cases} V_0, & r \leq a, \\ 0, & r > a. \end{cases} \quad (5.4.7)$$

Worked example: outgoing Green function, Born series, and exact scattering

The formal integral equation becomes much more concrete in one dimension. Work in units $\hbar^2/(2m) = 1$, choose positive energy $E = k^2$ with $k > 0$, and consider

$$\left(-\frac{d^2}{dx^2} + g\delta(x)\right)\psi(x) = k^2\psi(x). \quad (5.4.16)$$

The incident wave comes from the left. Outside $x = 0$ the stationary equation is free, so the most general outgoing scattering ansatz is

$$\psi(x) = \begin{cases} e^{ikx} + re^{-ikx}, & x < 0, \\ te^{ikx}, & x > 0. \end{cases} \quad (5.4.17)$$

Continuity of the wavefunction gives $\psi(0^-) = \psi(0^+)$, hence $1 + r = t$. Next integrate the differential equation from $-\epsilon$ to $+\epsilon$:

$$-\psi'(\epsilon) + \psi'(-\epsilon) + g\psi(0) = k^2 \int_{-\epsilon}^{\epsilon} \psi(x)dx. \quad (5.4.18)$$

The right side tends to zero as $\epsilon \rightarrow 0$; the delta function contributes $g\psi(0)$. Therefore

$$\psi'(0^+) - \psi'(0^-) = g\psi(0). \quad (5.4.19)$$

The two derivatives obtained from Eq. (5.4.17) are $\psi'(0^+) = ikt$ and $\psi'(0^-) = ik(1 - r)$. Use $t = 1 + r$ to find

$$ik[(1 + r) - (1 - r)] = g(1 + r), \quad 2ikr = g(1 + r). \quad (5.4.20)$$

Solving the last equation step by step,

$$(2ik - g)r = g, \quad r = \frac{g}{2ik - g} = \boxed{\frac{-iu}{1 + iu}}, \quad t = 1 + r = \boxed{\frac{1}{1 + iu}}, \quad u = \frac{g}{2k}. \quad (5.4.21)$$

The outgoing current carried by each plane wave is proportional to its signed wavenumber. Incoming and outgoing wavenumber magnitudes are identical, so reflection and transmission probabilities are simply squared amplitudes:

$$R = |r|^2 = \frac{u^2}{1 + u^2}, \quad T = |t|^2 = \frac{1}{1 + u^2}, \quad \boxed{R + T = 1}. \quad (5.4.22)$$

This current-conservation identity is an independent check.

Recover the same result from the integral equation. The one-dimensional outgoing free Green function for $k^2 + \partial_x^2$ is

$$G_0^+(x, x'; k) = -\frac{i}{2k} e^{ik|x-x'|}. \quad (5.4.23)$$

Its first derivative jumps by 1 in the equation $(k^2 + \partial_x^2)G_0^+ = \delta(x - x')$, which can be checked on either side of $x = x'$. The Lippmann–Schwinger equation becomes

$$\psi(x) = e^{ikx} + g G_0^+(x, 0; k)\psi(0). \quad (5.4.24)$$

Setting $x = 0$ gives an *algebraic* equation:

$$\psi(0) = 1 - \frac{ig}{2k}\psi(0) \implies \psi(0) = \frac{1}{1 + iu}. \quad (5.4.25)$$

For $x < 0$, the scattered term is $-iu\psi(0)e^{-ikx}$; therefore $r = -iu/(1 + iu)$, exactly as in Eq. (5.4.21). For $x > 0$ the total amplitude is $1 - iu\psi(0) = 1/(1 + iu) = t$.

Show precisely where the Born series converges. At the origin, the repeated interaction gives a geometric series

$$\psi(0) = \frac{1}{1 + iu} = \sum_{n=0}^{\infty} (-iu)^n, \quad |u| < 1. \quad (5.4.26)$$

The order- N reflection amplitude is

$$r_N = -iu \sum_{n=0}^{N-1} (-iu)^n. \quad (5.4.27)$$

Subtract it from the exact geometric sum:

$$r - r_N = r(-iu)^N, \quad |r - r_N| = |r| |u|^N. \quad (5.4.28)$$

For $|u| > 1$, successive Born terms grow, while the exact amplitude remains bounded and satisfies $R+T=1$. The radius of convergence here is a property of this specified delta-potential series; other potentials require their own analysis. The complete [interactive scattering problem](#) illustrates both behaviors.

Interactive investigation IC-05-03. Open Lippmann–Schwinger and Born series. After comparing the exact delta-scatterer amplitude with its Born expansion, tune the interaction strength. Inspect the first-order error and higher-order corrections, and determine where the geometric series ceases to converge.

5.5 Partial-Wave Method

Find the scattering amplitude in a central potential field.

1) Idea

The incident plane wave Ae^{ikz} is a common eigenstate of the free-particle observables $\{\hat{H}, \hat{\mathbf{p}}, \hat{L}_z\}$, with eigenvalues

$$E = \frac{\hbar^2 k^2}{2\mu}, \quad p_x = p_y = 0, \quad p_z = \hbar k, \quad L_z = 0. \quad (5.5.1)$$

For a central potential, $[\hat{\mathbf{p}}, \hat{H}] \neq 0$, so momentum $\mathbf{p}$ is not conserved, while $[\hat{\mathbf{L}}^2, \hat{H}] = 0$: orbital angular momentum is conserved. A commuting set of conserved observables is $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z\}$. Thus a plane wave is not an eigenstate of the conserved-observable set for a central field, and it is inconvenient for describing directly the relation between incoming and outgoing states.

The partial-wave method: Under conservation of energy (unchanged k) and conservation of $\hat{L}_z$ (unchanged $m=0$), expand the incident plane wave in eigenstates of $\hat{\mathbf{L}}^2$, $P_l(\cos\theta)$ (a partial-wave expansion). The scattering of a plane wave is thereby converted into the sum of scatterings of eigenstates (partial waves) of the conserved-observable set.

2) Partial-wave expansion of the plane wave

$$e^{ikz} = e^{ikr \cos\theta} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos\theta). \quad (5.5.2)$$

For the spherical Bessel function of order l ,

$$j_l(kr) \xrightarrow{r \rightarrow \infty} \frac{1}{kr} \sin\left(kr - \frac{l\pi}{2}\right). \quad (5.5.3)$$

Hence

$$e^{ikz} \xrightarrow{r \rightarrow \infty} \sum_{l=0}^{\infty} (2l+1) i^l \frac{1}{kr} \sin \left(kr - \frac{l\pi}{2} \right) P_l(\cos \theta). \quad (5.5.4)$$

The standard asymptotic solution is

$$\begin{aligned} \psi(\mathbf{r}) &\xrightarrow{r \rightarrow \infty} A e^{ikz} + A f(\theta) \frac{e^{ikr}}{r} \\ &= A \sum_{l=0}^{\infty} (2l+1) i^l \frac{1}{kr} \sin \left(kr - \frac{l\pi}{2} \right) P_l(\cos \theta) + A f(\theta) \frac{e^{ikr}}{r}. \end{aligned} \quad (5.5.5)$$

3) Partial-wave expansion of the state in a central field

A particular solution of the stationary Schrödinger equation for a central field is

$$\psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) Y_{lm}(\theta, \varphi). \quad (5.5.6)$$

The general solution is

$$\psi(r, \theta, \varphi) = \sum_{n,l,m} B_{nlm} R_{nl}(r) Y_{lm}(\theta, \varphi). \quad (5.5.7)$$

The general solution satisfying conservation of energy (k unchanged) and $\hat{L}_z$ ($m = 0$ unchanged) is

$$\psi(r, \theta) = \sum_{l=0}^{\infty} B_l R_l(r) P_l(\cos \theta). \quad (5.5.8)$$

Here $R_l(r)$ is determined by the radial equation

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_l}{dr} \right) + \left(k^2 - \frac{2\mu}{\hbar^2} V(r) - \frac{l(l+1)}{r^2} \right) R_l = 0. \quad (5.5.9)$$

As $r \rightarrow \infty$, $V(r) \rightarrow 0$, and

$$\frac{d^2(rR_l)}{dr^2} + k^2(rR_l) = 0. \quad (5.5.10)$$

Thus

$$R_l(r \rightarrow \infty) = \frac{C_l}{kr} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right). \quad (5.5.11)$$

The two undetermined constants C_l and δ_l depend on the particular central potential $V(r)$; $-\pi/2$ is introduced for convenience in calculation. The asymptotic solution is therefore

$$\psi(r, \theta) \xrightarrow{r \rightarrow \infty} A \sum_{l=0}^{\infty} \frac{A_l}{kr} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right) P_l(\cos \theta), \quad AA_l = B_l C_l. \quad (5.5.12)$$

Comparing with the standard asymptotic solution gives

$$\begin{aligned} &\sum_{l=0}^{\infty} (2l+1) i^l \frac{1}{kr} \sin \left(kr - \frac{l\pi}{2} \right) P_l(\cos \theta) + f(\theta) \frac{e^{ikr}}{r} \\ &= \sum_{l=0}^{\infty} \frac{A_l}{kr} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right) P_l(\cos \theta). \end{aligned} \quad (5.5.13)$$

Use

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}. \quad (5.5.14)$$

Then

$$0 = \frac{e^{ikr}}{2ikr} \left[2ikf(\theta) + \sum_{l=0}^{\infty} (2l+1) i^l e^{-il\pi/2} P_l(\cos \theta) - \sum_{l=0}^{\infty} A_l e^{i(\delta_l - l\pi/2)} P_l(\cos \theta) \right] \\ - \frac{e^{-ikr}}{2ikr} \left[\sum_{l=0}^{\infty} (2l+1) i^l e^{il\pi/2} P_l(\cos \theta) - \sum_{l=0}^{\infty} A_l e^{-i(\delta_l - l\pi/2)} P_l(\cos \theta) \right]. \quad (5.5.15)$$

For this to hold for arbitrary r , the coefficients of e^{ikr} and e^{-ikr} must vanish separately. We find

$$A_l = (2l+1) i^l e^{i\delta_l}, \quad (5.5.16)$$

and

$$f(\theta) = \sum_{l=0}^{\infty} \frac{2l+1}{k} P_l(\cos \theta) e^{i\delta_l} \sin \delta_l. \quad (5.5.17)$$

Define the partial-wave scattering amplitude

$$f_l(\theta) = \frac{1}{k} (2l+1) P_l(\cos \theta) e^{i\delta_l} \sin \delta_l. \quad (5.5.18)$$

The total scattering amplitude is the sum of the partial-wave amplitudes:

$$f(\theta) = \sum_{l=0}^{\infty} f_l(\theta). \quad (5.5.19)$$

The differential scattering cross section is

$$\sigma(\theta) = \left| \sum_{l=0}^{\infty} f_l(\theta) \right|^2. \quad (5.5.20)$$

The partial-wave scattering amplitudes depend only on δ_l . The key is to find the phase shifts δ_l :

1. Substitute $V(r)$ into the radial equation and obtain $R_l(r)$.
2. Compare $\lim_{r \rightarrow \infty} R_l(r)$ with $C_l(kr)^{-1} \sin(kr - l\pi/2 + \delta_l)$ to find δ_l .

The physical meaning of δ_l is as follows. The partial wave before scattering is

$$\frac{1}{kr} \sin \left(kr - \frac{l\pi}{2} \right), \quad (5.5.21)$$

whereas the partial wave after scattering is

$$\frac{1}{kr} \sin \left(kr - \frac{l\pi}{2} + \delta_l \right). \quad (5.5.22)$$



Thus scattering involves the same partial wave before and after, with only an overall phase shift δ_l .

4) Convergence

A semiclassical estimate of the convergence of the sum gives $L \sim pr$. Suppose the potential $V(r)$ has effective radius a . Scattering occurs only for $L < pa$, so that partial waves satisfying approximately $\sqrt{l(l+1)} < ka$ contribute. When the incident energy $E(k)$ is small, one need retain only low-order partial waves, such as the s wave.

Conclusion: The partial-wave method is suitable for low-energy scattering by a central field.

Example. Scattering of a low-energy particle by a spherically symmetric attractive potential well:

$$V(r) = \begin{cases} -V_0, & r \leq a, \\ 0, & r > a. \end{cases} \quad (5.5.23)$$

The radial equation is

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_l}{dr} \right) + \left(k^2 - \frac{2\mu}{\hbar^2} V(r) - \frac{l(l+1)}{r^2} \right) R_l = 0. \quad (5.5.24)$$

Consider only the s partial wave ($l = 0$). Then

$$\begin{cases} \frac{d^2(rR_0)}{dr^2} + k'^2(rR_0) = 0, & r \leq a, \\ \frac{d^2(rR_0)}{dr^2} + k^2(rR_0) = 0, & r > a, \end{cases} \quad k'^2 = k^2 + \frac{2\mu}{\hbar^2} V_0. \quad (5.5.25)$$

The solution is

$$R_0(r) = \begin{cases} \frac{C'_0}{k'r} \sin(k'r + B'_0), & r \leq a, \\ \frac{C_0}{kr} \sin(kr + B_0), & r > a. \end{cases} \quad (5.5.26)$$

Requiring $R_0(r)$ to remain finite at $r = 0$ gives $B'_0 = 0$. Furthermore, continuity of $R_0(r)$ and dR_0/dr at $r = a$ gives

$$\begin{aligned} B_0 &= \arctan \left(\frac{k}{k'} \tan(k'a) \right) - ka, \\ C'_0 &= C_0 \frac{k'}{k} \frac{\sin(ka + B_0)}{\sin(k'a)}. \end{aligned} \quad (5.5.27)$$

Compare

$$\lim_{r \rightarrow \infty} R_0(r) = \frac{C_0}{kr} \sin(kr + B_0) \quad (5.5.28)$$

with the standard asymptotic form for a central potential,

$$\lim_{r \rightarrow \infty} R_0(r) = \frac{C_0}{kr} \sin(kr + \delta_0). \quad (5.5.29)$$

The s -wave phase shift is

$$\boxed{\delta_0 = B_0}. \quad (5.5.30)$$

Since $P_0(\cos \theta) = 1$,

$$f_0(\theta) = \frac{1}{k} e^{i\delta_0} \sin \delta_0, \quad (5.5.31)$$

$$\sigma(\theta) = |f(\theta)|^2 \simeq |f_0(\theta)|^2 = \frac{\sin^2 \delta_0}{k^2}. \quad (5.5.32)$$

At low energy, $k \rightarrow 0$,

$$\arctan \left(\frac{k}{k'} \tan k'a \right) \longrightarrow \frac{k}{k'} \tan k'a, \quad k'^2 \longrightarrow \frac{2\mu V_0}{\hbar^2}. \quad (5.5.33)$$

Therefore

$$\delta_0 \simeq \frac{k}{k'} \tan k'a - ka = ka \left(\frac{\tan k'a}{k'a} - 1 \right), \quad (5.5.34)$$

and

$$\sigma(\theta) \simeq \frac{\sin^2 \delta_0}{k^2} \simeq \frac{\delta_0^2}{k^2} \simeq a^2 \left(\frac{\tan k'a}{k'a} - 1 \right)^2. \quad (5.5.35)$$

Since only the s partial wave has been considered, σ is independent of θ . The total scattering cross section is

$$\sigma_{\text{tot}} = \int \sigma(\theta) d\Omega = 4\pi a^2 \left(\frac{\tan k'a}{k'a} - 1 \right)^2. \quad (5.5.36)$$

Interactive investigation IC-05-05. Open Partial-wave interference. After computing the partial-wave cross section, vary individual phase shifts and inspect the total angular pattern. Confirm the channel contributions to the integrated cross section and the optical-theorem consistency check.

Chapter 6

Quantum Many-Body Systems

6.1 Identical Particles and Permutation Symmetry

In a system of identical particles, all particles have the same intrinsic properties: mass, charge, spin, and any other labels that identify the species. The spin–statistics connection is established in relativistic quantum field theory: integer-spin particles obey Bose–Einstein statistics, while half-integer-spin particles obey Fermi–Dirac statistics. Our immediate purpose is to understand how these two classes arise in the state space of a many-particle system. The analysis begins with permutations of particle labels.

1) States, observables, and permutations

Consider N identical particles with particle labels $1, \dots, N$. Their Hamiltonian is symmetric under exchanging the particle labels, including the associated spatial and spin coordinates. A many-body wavefunction is written

$$\Psi(1, 2, \dots, N), \quad (6.1.1)$$

where each argument collectively denotes the coordinates of one particle. Define the exchange operator $\hat{P}_{ij}$ by

$$\hat{P}_{ij}\Psi(1, \dots, i, \dots, j, \dots, N) = \Psi(1, \dots, j, \dots, i, \dots, N). \quad (6.1.2)$$

Exchanging two particles twice restores the original state, hence

$$\hat{P}_{ij}^2 = \mathbf{1}, \quad \hat{P}_{ij}^{-1} = \hat{P}_{ij}. \quad (6.1.3)$$

The Hamiltonian commutes with every permutation operator:

$$\boxed{[\hat{H}, \hat{P}] = 0, \quad P \in S_N.} \quad (6.1.4)$$

Here S_N is the symmetric group consisting of the $N!$ permutations of N objects. Any permutation is a product of elementary transpositions; its parity is even or odd according to the parity of the number of transpositions in such a decomposition. Different decompositions have the same parity.

On the many-particle Hilbert space each permutation is represented by a unitary operator. To see this, let the inner product be the integral over all particle coordinates. Relabeling dummy integration variables gives

$$\begin{aligned} \langle \Psi | \hat{P} \Phi \rangle &= \int d1 \cdots dN \Psi^*(1, \dots, N) (\hat{P}\Phi)(1, \dots, N) \\ &= \langle \hat{P}^{-1} \Psi | \Phi \rangle, \end{aligned} \quad (6.1.5)$$

so $\hat{P}^\dagger = \hat{P}^{-1}$. In particular, $\hat{P}_{ij}$ is Hermitian as well as unitary, and its eigenvalues are $+1$ or -1 .

An observable that treats identical particles identically is invariant under relabeling, and therefore

$$\boxed{\hat{P}^{-1}\hat{A}\hat{P} = \hat{A}, \quad [\hat{A}, \hat{P}] = 0 \quad (P \in S_N).} \quad (6.1.6)$$

The converse follows immediately by multiplying the commutator by $\hat{P}^{-1}$. Invariance of the Hamiltonian means that permutation symmetry is preserved under time evolution: for $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$,

$$\hat{P}\hat{U}(t)|\Psi\rangle = \hat{U}(t)\hat{P}|\Psi\rangle. \quad (6.1.7)$$

Therefore, if a state is symmetric or antisymmetric at the initial time, its symmetry character remains unchanged.

2) Examples: two and three particles

For two particles, $S_2 = \{1, P_{12}\}$. Given an arbitrary wavefunction $\Psi(1, 2)$, its symmetric and antisymmetric combinations are

$$\Psi_\pm(1, 2) \propto \Psi(1, 2) \pm \Psi(2, 1), \quad (6.1.8)$$

which satisfy $\hat{P}_{12}\Psi_\pm = \pm\Psi_\pm$ whenever the combinations are nonzero. The normalization depends on the overlap of the two terms; it is derived explicitly in the next section.

The three-particle example reveals an important additional possibility. The six elements of S_3 are the identity, three exchanges, and two cyclic permutations. For instance,

$$\hat{P}_{12}\hat{P}_{23}\hat{P}_{12} = \hat{P}_{13}, \quad \hat{P}_{12}\hat{P}_{23} \neq \hat{P}_{23}\hat{P}_{12}. \quad (6.1.9)$$

These operators need not commute. Hence a basis of simultaneous eigenstates of *all* exchange operators generally does not exist in the full distinguishable-particle tensor-product space.

Take three orthonormal one-particle orbitals $|a\rangle, |b\rangle, |c\rangle$. Acting with all elements of S_3 on $|abc\rangle := |a\rangle \otimes |b\rangle \otimes |c\rangle$ produces six orthonormal ordered products. This six-dimensional permutation space contains an invariant symmetric combination

$$|\Psi_s\rangle = \frac{1}{\sqrt{6}}(|abc\rangle + |acb\rangle + |bac\rangle + |bca\rangle + |cab\rangle + |cba\rangle), \quad (6.1.10)$$

and an invariant antisymmetric combination

$$|\Psi_a\rangle = \frac{1}{\sqrt{6}}(|abc\rangle - |acb\rangle - |bac\rangle + |bca\rangle + |cab\rangle - |cba\rangle). \quad (6.1.11)$$

They transform as $\hat{P}|\Psi_s\rangle = |\Psi_s\rangle$ and $\hat{P}|\Psi_a\rangle = \text{sgn}(P)|\Psi_a\rangle$.

There are also states in this six-dimensional space that transform into *linear combinations of one another* under permutations. These form two copies of the two-dimensional standard irreducible representation of S_3 . Two orthonormal vectors in the remaining four-dimensional complement are

$$\begin{aligned} |u_1\rangle &= \frac{1}{\sqrt{6}}(2|abc\rangle - |bca\rangle - |cab\rangle), \\ |u_2\rangle &= \frac{1}{\sqrt{2}}(|bca\rangle - |cab\rangle). \end{aligned} \quad (6.1.12)$$

Permutation operators mix vectors in this four-dimensional complement. By choosing appropriate two-dimensional invariant subspaces, each permutation is represented by a 2×2 matrix

rather than a single overall sign. More abstractly, the six-dimensional regular representation decomposes as

$$\mathbf{6} \simeq \mathbf{1}_s \oplus \mathbf{1}_a \oplus \mathbf{2} \oplus \mathbf{2}, \quad 1^2 + 1^2 + 2^2 = 6. \quad (6.1.13)$$

This example explains why permutation-group symmetry alone permits mathematical representations beyond the completely symmetric and completely antisymmetric ones. In the usual nonrelativistic treatment of ordinary identical bosons and fermions, we select the corresponding one-dimensional sectors as the physical exchange-symmetry postulate. Its relation to particle spin is supplied by the spin–statistics theorem under its relativistic assumptions.

3) Permutation-invariant subspaces

A subspace $\mathcal{V}$ is invariant under a representation of S_N if $\hat{P}\mathcal{V} \subseteq \mathcal{V}$ for every $P \in S_N$. Symmetric and antisymmetric states constitute two especially simple invariant subspaces. The physical action of a permutation never changes measurable expectation values:

$$\langle \hat{P}\Psi | \hat{A} | \hat{P}\Psi \rangle = \langle \Psi | \hat{P}^\dagger \hat{A} \hat{P} | \Psi \rangle = \langle \Psi | \hat{A} | \Psi \rangle. \quad (6.1.14)$$

This is the operational meaning of particle indistinguishability. The next section constructs normalized many-particle states in the two physical exchange-symmetry sectors.

Interactive investigation IC-06-01. Open Particle exchange. After defining particle permutations, exchange two identical particles in symmetric and antisymmetric states. Compare eigenvalues of the exchange operator and check which observables are invariant.

6.2 Completely Symmetric and Antisymmetric States

Let $\{|i\rangle\}$ be an orthonormal one-particle basis. An ordered basis vector in the N -particle tensor-product space is

$$|i_1, \dots, i_N\rangle := |i_1\rangle_1 \otimes \cdots \otimes |i_N\rangle_N. \quad (6.2.1)$$

Write the sum of the permutation operators as

$$\hat{\mathcal{S}}_\pm = \frac{1}{\sqrt{N!}} \sum_{P \in S_N} [\text{sgn}(P)]_\pm \hat{P}, \quad (6.2.2)$$

where $[\text{sgn}(P)]_+ = 1$ and $[\text{sgn}(P)]_- = \text{sgn}(P)$. We deliberately use $\mathcal{S}$ for the unnormalized sum with the square-root prefactor; normalized projectors are introduced below. Changing variables $Q = PP'$ in the group sum yields

$$\hat{P}' \hat{\mathcal{S}}_+ = \hat{\mathcal{S}}_+ \hat{P}' = \hat{\mathcal{S}}_+, \quad \hat{P}' \hat{\mathcal{S}}_- = \hat{\mathcal{S}}_- \hat{P}' = \text{sgn}(P') \hat{\mathcal{S}}_-. \quad (6.2.3)$$

Consequently $\hat{\mathcal{S}}_+ |i_1, \dots, i_N\rangle$ is symmetric and $\hat{\mathcal{S}}_- |i_1, \dots, i_N\rangle$ is antisymmetric. They are also orthogonal for $N > 1$: an exchange P_{12} preserves the first state and reverses the sign of the second, so their overlap equals its own negative.

1) The symmetrization and antisymmetrization projectors

Introduce the actual orthogonal projectors

$$\hat{\mathcal{P}}_\pm = \frac{1}{N!} \sum_{P \in S_N} [\text{sgn}(P)]_\pm \hat{P}. \quad (6.2.4)$$

Their defining properties follow directly from the group multiplication law:

$$\hat{\mathcal{P}}_\pm^\dagger = \hat{\mathcal{P}}_\pm, \quad \hat{\mathcal{P}}_\pm^2 = \hat{\mathcal{P}}_\pm, \quad \hat{\mathcal{P}}_+ \hat{\mathcal{P}}_- = 0 \quad (N > 1). \quad (6.2.5)$$

For example, in the square of the plus projector, for each fixed $R \in S_N$ there are $N!$ pairs (P, Q) with $PQ = R$; therefore

$$\hat{\mathcal{P}}_+^2 = \frac{1}{(N!)^2} \sum_{P, Q} \hat{P} \hat{Q} = \frac{1}{N!} \sum_R \hat{R} = \hat{\mathcal{P}}_+. \quad (6.2.6)$$

For the minus projector the factor $\text{sgn}(P) \text{sgn}(Q) = \text{sgn}(PQ)$ yields the identical argument. Every completely symmetric state can be written as a linear combination of projected ordered basis states, since the identity resolution can be inserted before projection.

2) Normalization and repeated orbital occupation

Suppose the ordered list $(i_1, \dots, i_N)$ contains the orbital i precisely n_i times, with $\sum_i n_i = N$. For bosons, there are $N! / \prod_i n_i!$ distinct permutations of that list, and every distinct ordered product occurs $\prod_i n_i!$ times in the full permutation sum. The norm is therefore

$$\left\| \sum_{P \in S_N} \hat{P} |i_1, \dots, i_N\rangle \right\|^2 = N! \prod_i n_i!. \quad (6.2.7)$$

Thus the normalized Bose state is

$$|\{n_i\}\rangle_B = \frac{1}{\sqrt{N! \prod_i n_i!}} \sum_{P \in S_N} \hat{P} |i_1, \dots, i_N\rangle. \quad (6.2.8)$$

In terms of the handout's $\hat{\mathcal{S}}_+$, this is $|\{n_i\}\rangle_B = \hat{\mathcal{S}}_+ |i_1, \dots, i_N\rangle / \sqrt{\prod_i n_i!}$. This factorial is essential whenever an orbital is occupied more than once.

For fermions, if $i_a = i_b$ for $a \neq b$, the exchanged ordered product is unchanged but the antisymmetrization sign is reversed, giving zero. Therefore all occupied orbitals must be distinct. For distinct $i_1, \dots, i_N$, every ordered product in the permutation sum is orthogonal to every other, and

$$|i_1, \dots, i_N\rangle_F = \frac{1}{\sqrt{N!}} \sum_{P \in S_N} \text{sgn}(P) \hat{P} |i_1, \dots, i_N\rangle. \quad (6.2.9)$$

In coordinate representation the same expression is the normalized Slater determinant,

$$\Psi_F(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \varphi_{i_1}(x_1) & \cdots & \varphi_{i_N}(x_1) \\ \vdots & \ddots & \vdots \\ \varphi_{i_1}(x_N) & \cdots & \varphi_{i_N}(x_N) \end{pmatrix}. \quad (6.2.10)$$

The variable x may include spin as well as position. Both constructions provide complete orthonormal bases of their respective symmetric and antisymmetric N -particle sectors. The next chapter assembles sectors with different N into Fock space and derives operator rules directly from these states.

6.3 Bosons: Fock Space and Creation and Annihilation Operators

A completely symmetric N -particle state can be labeled by the *occupation numbers* n_i of one-particle orbitals $|i\rangle$. The actual particle labels have disappeared from this description. Define

$$|n_1, n_2, \dots\rangle_B, \quad n_i = 0, 1, 2, \dots, \quad \sum_i n_i = N, \quad (6.3.1)$$

by the normalized symmetrization in the preceding chapter. Because different occupation lists are orthogonal,

$${}_B \langle n_1, n_2, \dots | n'_1, n'_2, \dots \rangle_B = \prod_i \delta_{n_i n'_i}. \quad (6.3.2)$$

For a fixed N these states span the symmetric N -particle Hilbert space $\mathcal{H}_N^{(B)}$. Combining the spaces for every particle number leads to the Bose Fock space,

$$\mathcal{F}_B = \bigoplus_{N=0}^{\infty} \mathcal{H}_N^{(B)}. \quad (6.3.3)$$

The $N = 0$ subspace contains a distinguished normalized state, the *vacuum*, written $|0\rangle = |0, 0, \dots\rangle$. A state with several occupied orbitals belongs to one fixed- N sector; a general Fock-space state may be a superposition of sectors.

1) Creation and annihilation operators

The occupation-number basis makes the effect of adding or removing one particle in orbital i explicit. Define

$$\begin{aligned} \hat{a}_i^\dagger |\dots, n_i, \dots\rangle_B &= \sqrt{n_i + 1} |\dots, n_i + 1, \dots\rangle_B, \\ \hat{a}_i |\dots, n_i, \dots\rangle_B &= \sqrt{n_i} |\dots, n_i - 1, \dots\rangle_B. \end{aligned} \quad (6.3.4)$$

The second formula gives zero when $n_i = 0$. The square-root factors follow from the normalized symmetric states, including the factorial $n_i!$. In particular,

$$\hat{a}_i |0\rangle = 0, \quad |i\rangle = \hat{a}_i^\dagger |0\rangle, \quad |n_1, n_2, \dots\rangle_B = \prod_i \frac{(\hat{a}_i^\dagger)^{n_i}}{\sqrt{n_i!}} |0\rangle. \quad (6.3.5)$$

The order of factors is immaterial for bosons. To check the normalization of the last expression, use the norm relation

$$\|\hat{a}_i^\dagger |n_i\rangle\|^2 = n_i + 1 \implies \|(\hat{a}_i^\dagger)^{n_i} |0\rangle\|^2 = n_i!. \quad (6.3.6)$$

Thus the operator algebra reproduces the explicit permutation sum without referring to labels of individual particles.

2) Derivation of Bose commutation relations

For different modes $i \neq j$, creation of particles in modes i, j can be carried out in either order. Applying the preceding formulas twice yields

$$\hat{a}_i \hat{a}_j |\{n_k\}\rangle = \sqrt{n_i n_j} |\dots, n_i - 1, \dots, n_j - 1, \dots\rangle = \hat{a}_j \hat{a}_i |\{n_k\}\rangle. \quad (6.3.7)$$

The same argument applies to two creation operators and to $\hat{a}_i \hat{a}_j^\dagger$ for $i \neq j$. For one fixed mode,

$$\begin{aligned} \hat{a}_i \hat{a}_i^\dagger |\{n_k\}\rangle &= (n_i + 1) |\{n_k\}\rangle, \\ \hat{a}_i^\dagger \hat{a}_i |\{n_k\}\rangle &= n_i |\{n_k\}\rangle. \end{aligned} \quad (6.3.8)$$

Their difference is the identity on every occupation-number basis state. We have therefore established

$$[\hat{a}_i, \hat{a}_j] = 0, \quad [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \quad [\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij} \mathbf{1}. \quad (6.3.9)$$

Here $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$. The positive sign in the last relation is characteristic of bosonic statistics.

3) Particle-number operators and independent particles

Define the occupation-number operator for mode i by

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i, \quad \hat{n}_i |\{n_k\}\rangle = n_i |\{n_k\}\rangle. \quad (6.3.10)$$

Summing over all modes gives the total particle-number operator,

$$\boxed{\hat{N} = \sum_i \hat{n}_i = \sum_i \hat{a}_i^\dagger \hat{a}_i.} \quad (6.3.11)$$

For a noninteracting system with one-particle energies ϵ_i , the many-particle Hamiltonian acts on an occupation state as

$$\hat{H}_0 |\{n_i\}\rangle = \left(\sum_i n_i \epsilon_i \right) |\{n_i\}\rangle, \quad (6.3.12)$$

which is reproduced exactly by

$$\boxed{\hat{H}_0 = \sum_i \epsilon_i \hat{a}_i^\dagger \hat{a}_i.} \quad (6.3.13)$$

This formula is a first example of a sum over individually identical one-particle operators represented entirely in terms of occupation modes.

4) General one-particle operator

Consider a one-particle observable $\hat{t}$ and the additive many-particle operator

$$\hat{\mathcal{O}} = \sum_{\alpha=1}^N \hat{t}_\alpha, \quad t_{ij} = \langle i | \hat{t} | j \rangle. \quad (6.3.14)$$

The completeness of the one-particle basis gives

$$\hat{t} = \sum_{ij} t_{ij} |i\rangle \langle j|, \quad \hat{\mathcal{O}} = \sum_{ij} t_{ij} \left(\sum_{\alpha=1}^N |i\rangle_\alpha \langle j|_\alpha \right). \quad (6.3.15)$$

The operator in parentheses removes one particle from $|j\rangle$ and returns one in $|i\rangle$. On a bosonic occupation state with $i \neq j$, there are n_j possible removals and a corresponding symmetrization factor, which together produce

$$\begin{aligned} \left(\sum_{\alpha} |i\rangle_{\alpha} \langle j|_{\alpha} \right) |\dots, n_i, \dots, n_j, \dots\rangle_B &= \sqrt{n_j(n_i+1)} |\dots, n_i+1, \dots, n_j-1, \dots\rangle_B \\ &= \hat{a}_i^\dagger \hat{a}_j |\{n_k\}\rangle_B. \end{aligned} \quad (6.3.16)$$

For $i = j$ it simply counts n_i particles. Since the actions agree on a complete Fock basis,

$$\boxed{\hat{\mathcal{O}} = \sum_{ij} t_{ij} \hat{a}_i^\dagger \hat{a}_j.} \quad (6.3.17)$$

The noninteracting Hamiltonian above follows by setting $t_{ij} = \epsilon_i \delta_{ij}$. Notice the operator order: first the particle is removed from j , and then one is created in i .

5) Two-particle interactions

For a symmetric pair potential $V(x_1, x_2) = V(x_2, x_1)$, the first-quantized operator counts each unordered pair once:

$$\hat{V} = \sum_{\alpha < \beta} V(\hat{x}_\alpha, \hat{x}_\beta) = \frac{1}{2} \sum_{\alpha \neq \beta} V(\hat{x}_\alpha, \hat{x}_\beta). \quad (6.3.18)$$

Introduce the *unsymmetrized* orbital matrix elements

$$V_{ij;kl} = \int dx dy \varphi_i^*(x) \varphi_j^*(y) V(x, y) \varphi_k(x) \varphi_l(y). \quad (6.3.19)$$

They obey $V_{ij;kl} = V_{ji;lk}$ for a symmetric interaction. Acting on an occupation state with the ordered product $\hat{a}_l \hat{a}_k$ removes the two incoming particles, and $\hat{a}_i^\dagger \hat{a}_j^\dagger$ inserts the two outgoing particles. Summing over the orbital indices exactly reconstructs the matrix elements of $\sum_{\alpha \neq \beta} V_{\alpha\beta}$; the factor $1/2$ removes double counting. Thus

$$\hat{V} = \frac{1}{2} \sum_{ijkl} V_{ij;kl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k. \quad (6.3.20)$$

The ordering of k, l on the right is deliberate and will be retained for fermions. This formula uses unsymmetrized matrix elements; antisymmetrizing those matrix elements would change the appropriate prefactor and indexing conventions.

Interactive investigation IC-06-03. Open Bosonic Fock occupations. After constructing the bosonic ladder operators, apply creation and annihilation to different number states. Check the *sqrtn* + 1 and *sqrtn* factors by evaluating state norms and commutators.

6.4 Fermions: Antisymmetric States and Operator Algebra

Fermionic many-particle states are completely antisymmetric under exchanging any two particles. From the preceding chapter, the normalized wavefunction made of N distinct single-particle orbitals is a Slater determinant:

$$\Psi_F(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det[\varphi_{i_a}(x_b)]_{a,b=1}^N. \quad (6.4.1)$$

If any two orbitals coincide, two rows of the determinant are identical and the wavefunction vanishes. This is the Pauli exclusion principle: the occupation number of each one-particle state is either zero or one.

1) Fermionic Fock space and ordered occupation states

Choose a fixed ordering $1 < 2 < \dots$ of the one-particle orbitals. The antisymmetric states with different occupation lists form an orthonormal basis,

$$|n_1, n_2, \dots\rangle_F, \quad n_i \in \{0, 1\}, \quad N = \sum_i n_i. \quad (6.4.2)$$

The full fermionic Fock space is the direct sum

$$\mathcal{F}_F = \bigoplus_{N=0}^{\infty} \mathcal{H}_N^{(F)}, \quad (6.4.3)$$

with vacuum $|0\rangle$. The fixed ordering is essential: exchanging the order of fermion creation operators changes the sign of the state.

Define creation operators by the ordered convention

$$|n_1, n_2, \dots\rangle_F = \prod_{i=1}^{\infty} (\hat{a}_i^\dagger)^{n_i} |0\rangle, \quad (6.4.4)$$

where the product is taken in increasing order of i , from left to right. Since adding a particle to an occupied orbital would produce a repeated column in the Slater determinant,

$$(\hat{a}_i^\dagger)^2 = 0. \quad (6.4.5)$$

Exchanging two distinct orbitals reverses the Slater determinant sign, hence

$$\hat{a}_i^\dagger \hat{a}_j^\dagger = -\hat{a}_j^\dagger \hat{a}_i^\dagger \quad (i \neq j). \quad (6.4.6)$$

2) Fermionic creation and annihilation operators

To put a newly created operator into the fixed orbital order, it must pass every occupied lower-index mode. Define the parity count

$$\nu_i = \sum_{j < i} n_j. \quad (6.4.7)$$

The creation and annihilation actions are then

$$\begin{aligned} \hat{a}_i^\dagger |\dots, n_i, \dots\rangle_F &= (1 - n_i)(-1)^{\nu_i} |\dots, n_i + 1, \dots\rangle_F, \\ \hat{a}_i |\dots, n_i, \dots\rangle_F &= n_i(-1)^{\nu_i} |\dots, n_i - 1, \dots\rangle_F. \end{aligned} \quad (6.4.8)$$

Thus creation at an already occupied mode and annihilation at an empty mode both give zero. For the vacuum, $\hat{a}_i |0\rangle = 0$ as in the bosonic case. The sign is the number of already occupied orbitals passed when inserting or removing a fermion.

3) Derivation of anticommutation relations

Apply the operators in the two possible orders to an arbitrary occupation state. For the same mode,

$$\begin{aligned} \hat{a}_i \hat{a}_i^\dagger |\{n_k\}\rangle_F &= (1 - n_i) |\{n_k\}\rangle_F, \\ \hat{a}_i^\dagger \hat{a}_i |\{n_k\}\rangle_F &= n_i |\{n_k\}\rangle_F. \end{aligned} \quad (6.4.9)$$

Their sum is the identity. For distinct $i < j$, exchanging the order of the two operations changes the number of occupied orbitals through which one operator passes by one. Therefore

$$\hat{a}_i \hat{a}_j^\dagger |\{n_k\}\rangle_F = -\hat{a}_j^\dagger \hat{a}_i |\{n_k\}\rangle_F, \quad i \neq j, \quad (6.4.10)$$

and analogously for two creation or two annihilation operators. Define the anticommutator $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. We have

$$\{\hat{a}_i, \hat{a}_j\} = 0, \quad \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0, \quad \{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij} \mathbf{1}. \quad (6.4.11)$$

In particular, the nilpotence of the creation operator follows again from $\{\hat{a}_i^\dagger, \hat{a}_i^\dagger\} = 0$. The occupation-number operator has the same formal expression as for bosons,

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i, \quad \hat{N} = \sum_i \hat{a}_i^\dagger \hat{a}_i, \quad (6.4.12)$$

but its eigenvalues are restricted to 0, 1.

4) One-particle operators and the importance of order

Take the additive operator $\hat{O} = \sum_{\alpha=1}^N \hat{t}_{\alpha}$, and let $t_{ij} = \langle i | \hat{t} | j \rangle$. As for bosons, each term $|i\rangle_{\alpha} \langle j|_{\alpha}$ removes one occupied orbital j and replaces it by i . In the ordered Slater determinant, the signs from removing and inserting exactly match those in $\hat{a}_i^{\dagger} \hat{a}_j$. Thus on every antisymmetric basis state,

$$\sum_{\alpha=1}^N |i\rangle_{\alpha} \langle j|_{\alpha} = \hat{a}_i^{\dagger} \hat{a}_j, \quad (6.4.13)$$

and hence

$$\hat{O} = \sum_{ij} t_{ij} \hat{a}_i^{\dagger} \hat{a}_j. \quad (6.4.14)$$

No special parity factor is inserted outside the operators: the fermionic sign is already contained in the creation and annihilation algebra.

For a symmetric two-particle interaction, apply the same reasoning with two incoming and two outgoing orbitals. Use the unsymmetrized matrix elements $V_{ij;kl}$ defined in the boson section. With the explicit operator order

$$\hat{a}_i^{\dagger} \hat{a}_j^{\dagger} \hat{a}_l \hat{a}_k, \quad (6.4.15)$$

particles are first annihilated from k and l , and then created in j and i , with all signs supplied by anticommutation. One obtains

$$\hat{V} = \frac{1}{2} \sum_{ijkl} V_{ij;kl} \hat{a}_i^{\dagger} \hat{a}_j^{\dagger} \hat{a}_l \hat{a}_k. \quad (6.4.16)$$

The factor $1/2$ counts each *unordered particle pair* once. It is independent of Bose or Fermi statistics when $V_{ij;kl}$ is defined by unsymmetrized orbital integrals.

5) Unified form for bosons and fermions

Both particle species have the same *normal-ordered* form of a general one- plus two-body Hamiltonian:

$$\begin{aligned} \hat{H} = & \sum_{ij} h_{ij} \hat{a}_i^{\dagger} \hat{a}_j \\ & + \frac{1}{2} \sum_{ijkl} V_{ij;kl} \hat{a}_i^{\dagger} \hat{a}_j^{\dagger} \hat{a}_l \hat{a}_k, \end{aligned} \quad (6.4.17)$$

where $h_{ij} = \langle i | \hat{p}^2/(2m) + \hat{U} | j \rangle$. The entire difference between the statistics resides in the canonical commutator or anticommutator relations and the allowed occupation numbers:

	Bosons	Fermions
n_i	$0, 1, 2, \dots$	$0, 1$
$\hat{a}_i \hat{a}_j^{\dagger} \mp \hat{a}_j^{\dagger} \hat{a}_i$	δ_{ij}	δ_{ij}

(6.4.18)

Here the minus sign applies to bosons and the plus sign to fermions. For a one-body potential U , its matrix elements are included in h_{ij} ; interactions between two particles are included in $V_{ij;kl}$. This is the general result toward which the lecture derivations have been building.

Worked calculation: tracking fermionic signs in three orbitals

The antisymmetric Fock algebra is sometimes difficult because operators act *from right to left* while the canonical occupation convention orders creation operators from left to right. We calculate a three-mode example explicitly.

Choose orbitals $1 < 2 < 3$, vacuum $|0\rangle$, and

$$|1, 0, 1\rangle = \hat{a}_1^\dagger \hat{a}_3^\dagger |0\rangle. \quad (6.4.19)$$

Act with $\hat{a}_2^\dagger$. It must cross the occupied lower-index mode 1 to restore canonical order:

$$\begin{aligned} \hat{a}_2^\dagger |1, 0, 1\rangle &= \hat{a}_2^\dagger \hat{a}_1^\dagger \hat{a}_3^\dagger |0\rangle \\ &= -\hat{a}_1^\dagger \hat{a}_2^\dagger \hat{a}_3^\dagger |0\rangle = \boxed{-|1, 1, 1\rangle}. \end{aligned} \quad (6.4.20)$$

The minus sign is exactly $(-1)^{\nu_2}$, since one lower-index orbital is occupied. Acting instead with $\hat{a}_1$ on $|1, 0, 1\rangle$,

$$\hat{a}_1 \hat{a}_1^\dagger \hat{a}_3^\dagger |0\rangle = (1 - \hat{a}_1^\dagger \hat{a}_1) \hat{a}_3^\dagger |0\rangle = |0, 0, 1\rangle, \quad (6.4.21)$$

because $\hat{a}_1 |0\rangle = 0$. This result has positive sign: mode 1 has no occupied lower-index orbitals.

A complete one-body transition. Evaluate $\hat{a}_2^\dagger \hat{a}_3$ on $|1, 0, 1\rangle$ by applying the rightmost operator first. Since $\hat{a}_3$ must pass the particle in mode 1,

$$\hat{a}_3 |1, 0, 1\rangle = \hat{a}_3 \hat{a}_1^\dagger \hat{a}_3^\dagger |0\rangle = -\hat{a}_1^\dagger (1 - \hat{a}_3^\dagger \hat{a}_3) |0\rangle = -|1, 0, 0\rangle. \quad (6.4.22)$$

Now create in mode 2:

$$\hat{a}_2^\dagger \hat{a}_3 |1, 0, 1\rangle = -\hat{a}_2^\dagger \hat{a}_1^\dagger |0\rangle = +\hat{a}_1^\dagger \hat{a}_2^\dagger |0\rangle = \boxed{+|1, 1, 0\rangle}. \quad (6.4.23)$$

Two minus signs cancel. This small calculation is the concrete meaning of the matrix element $t_{23} \hat{a}_2^\dagger \hat{a}_3$ in a one-body Hamiltonian.

Where does the two-body factor of one half come from? Start with the first-quantized sum over unordered pairs $\hat{V} = \sum_{\alpha < \beta} v(\mathbf{r}_\alpha, \mathbf{r}_\beta)$. For a symmetric interaction $v(\mathbf{r}_\alpha, \mathbf{r}_\beta) = v(\mathbf{r}_\beta, \mathbf{r}_\alpha)$,

$$\sum_{\alpha < \beta} v_{\alpha\beta} = \frac{1}{2} \sum_{\alpha \neq \beta} v_{\alpha\beta}. \quad (6.4.24)$$

The right-hand sum counts each pair twice. After expanding the one-particle coordinate basis and using creation and annihilation operators, this becomes

$$\hat{V} = \frac{1}{2} \sum_{ijkl} V_{ij;kl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k. \quad (6.4.25)$$

The prefactor follows from *pair counting*; anticommutation then supplies exchange signs. If antisymmetrized two-body matrix elements are used instead of the unsymmetrized $V_{ij;kl}$ defined above, the numerical prefactor must be reconsidered to avoid double counting.

Self-check. Calculate $\hat{a}_3^\dagger \hat{a}_2 |1, 1, 0\rangle$ and confirm that the outcome is $+|1, 0, 1\rangle$, the Hermitian-adjoint transition of the preceding example.

Interactive investigation IC-06-04. Open Fermionic ordering signs. After the three-orbital sign calculation, change the occupation pattern and the order of fermionic operators. Count occupied lower-index modes and compare predicted parity signs with the resulting Fock states.

Chapter 7

Relativistic Quantum Mechanics

7.1 From the Schrödinger Equation to the Pauli Equation

The nonrelativistic Schrödinger equation describes a scalar wavefunction and, for a free particle, uses the Hamiltonian $\hat{H} = \hat{\mathbf{p}}^2/(2m)$. Electrons also possess an intrinsic spin and a corresponding magnetic moment. To incorporate spin into the wave equation, we first construct a two-component spinor theory. We shall then derive the Pauli Hamiltonian directly and show that it emerges again as the low-energy limit of the relativistic Dirac equation.

1) A two-component wavefunction

For spin 1/2 write

$$\psi(\mathbf{r}, t) = \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix}, \quad \hat{\mathbf{s}} = \frac{\hbar}{2} \boldsymbol{\sigma}, \quad (7.1.1)$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.1.2)$$

They satisfy $\sigma_i \sigma_j = \delta_{ij} \mathbf{1}_2 + i \sum_k \epsilon_{ijk} \sigma_k$. Thus, for vectors of mutually commuting scalars, $(\boldsymbol{\sigma} \cdot \mathbf{a})^2 = \mathbf{a}^2 \mathbf{1}_2$. In an electromagnetic field we must take account of the noncommutativity of kinetic momentum components.

2) Minimal coupling and the kinetic momentum

For a particle of mass m and electric charge q in the electromagnetic potentials $\phi(\mathbf{r}, t)$ and $\mathbf{A}(\mathbf{r}, t)$, define

$$\hat{\mathbf{p}} = -i\hbar \nabla, \quad \hat{\boldsymbol{\pi}} = \hat{\mathbf{p}} - \frac{q}{c} \mathbf{A}(\hat{\mathbf{r}}, t), \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (7.1.3)$$

Here $\hat{\mathbf{p}}$ is the canonical momentum, whereas $\hat{\boldsymbol{\pi}}$ is the mechanical (kinetic) momentum. The Pauli Hamiltonian is

$$\boxed{\hat{H}_P = \frac{1}{2m} (\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2 + q\phi(\hat{\mathbf{r}}, t)}. \quad (7.1.4)$$

To expand the square, use the Pauli identity in operator form:

$$(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2 = \sum_{ij} (\delta_{ij} \mathbf{1}_2 + i \epsilon_{ijk} \sigma_k) \hat{\pi}_i \hat{\pi}_j = \hat{\boldsymbol{\pi}}^2 + \frac{i}{2} \sum_{ijk} \epsilon_{ijk} \sigma_k [\hat{\pi}_i, \hat{\pi}_j]. \quad (7.1.5)$$

Direct differentiation gives

$$[\hat{\pi}_i, \hat{\pi}_j] = \frac{iq\hbar}{c} (\partial_i A_j - \partial_j A_i) = \frac{iq\hbar}{c} \epsilon_{ijk} B_k. \quad (7.1.6)$$

Consequently,

$$(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2 = \hat{\boldsymbol{\pi}}^2 - \frac{q\hbar}{c} \boldsymbol{\sigma} \cdot \mathbf{B}, \quad (7.1.7)$$

and hence

$$\hat{H}_P = \frac{\hat{\boldsymbol{\pi}}^2}{2m} - \frac{q\hbar}{2mc} \boldsymbol{\sigma} \cdot \mathbf{B} + q\phi. \quad (7.1.8)$$

The spin-field term is $-\hat{\boldsymbol{\mu}}_s \cdot \mathbf{B}$ with magnetic moment $\hat{\boldsymbol{\mu}}_s = q\hat{\mathbf{s}}/(mc)$. In this theory the spin gyromagnetic factor is $g = 2$. For an electron $q = -e$ with $e > 0$, its magnetic moment points opposite to its spin.

3) Orbital and spin magnetic moments in a weak uniform field

To identify the orbital contribution, consider a uniform field and choose the symmetric gauge $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$. For weak B we retain terms through first order in the field:

$$\begin{aligned} \hat{\boldsymbol{\pi}}^2 &= \hat{\mathbf{p}}^2 - \frac{q}{c}(\hat{\mathbf{p}} \cdot \mathbf{A} + \mathbf{A} \cdot \hat{\mathbf{p}}) + O(B^2) \\ &= \hat{\mathbf{p}}^2 - \frac{q}{c}\mathbf{B} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{p}}) + O(B^2). \end{aligned} \quad (7.1.9)$$

Here $\nabla \cdot \mathbf{A} = 0$ and $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ is the orbital angular momentum. Substituting into Eq. (7.1.8) yields

$$\hat{H}_P = \frac{\hat{\mathbf{p}}^2}{2m} + q\phi - \frac{q}{2mc}\hat{\mathbf{L}} \cdot \mathbf{B} - \frac{q}{mc}\hat{\mathbf{s}} \cdot \mathbf{B} + O(B^2). \quad (7.1.10)$$

Define the two magnetic moments by

$$\hat{\boldsymbol{\mu}}_L = \frac{q}{2mc}\hat{\mathbf{L}}, \quad \hat{\boldsymbol{\mu}}_s = \frac{q}{mc}\hat{\mathbf{s}}. \quad (7.1.11)$$

The orbital coefficient is half the spin coefficient. The same ratio will reappear from the relativistic Dirac equation. For atoms, these couplings underlie the Zeeman splitting and, after including the nuclear spin, hyperfine structure.

Interactive investigation IC-07-01. Open From Schrödinger to Pauli. After deriving the Pauli magnetic term, vary field direction, charge, and spin state. Use the matrix products to check the sign and strength of the spin-field coupling and its nonrelativistic interpretation.

7.2 From Relativistic Energy to the Dirac Equation

The relativistic energy-momentum relation is

$$E^2 = c^2\mathbf{p}^2 + m^2c^4. \quad (7.2.1)$$

We seek a wave equation compatible with this relation. It will have to accommodate a spinor and a Hamiltonian linear in the momentum.

1) The Klein-Gordon equation

Using the same operator substitutions as in the Schrödinger equation, $E \mapsto i\hbar\partial_t$ and $\mathbf{p} \mapsto -i\hbar\nabla$, Eq. (7.2.1) gives

$$-\hbar^2\partial_t^2\phi = -\hbar^2c^2\nabla^2\phi + m^2c^4\phi. \quad (7.2.2)$$

Thus a relativistic scalar field obeys the Klein–Gordon equation

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{m^2 c^2}{\hbar^2} \right) \phi = 0. \quad (7.2.3)$$

With $x^\mu = (ct, \mathbf{r})$, metric $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ and $\square = \partial_\mu \partial^\mu$, this is $(\square + m^2 c^2 / \hbar^2) \phi = 0$. The equation is second order in time. It has positive- and negative-frequency plane-wave solutions. For a complex scalar field, the natural conserved current has a temporal component of indefinite sign, motivating a quantum-field interpretation when particle creation becomes relevant.

2) A first-order relativistic Hamiltonian

Dirac's strategy is to factor the expression on the right of Eq. (7.2.1) using matrices:

$$\hat{H}_D = c \boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta m c^2, \quad i\hbar \frac{\partial \psi}{\partial t} = \hat{H}_D \psi. \quad (7.2.4)$$

The matrices α_i and β are independent of position and momentum. Expanding the square and demanding $\hat{H}_D^2 = (c^2 \hat{\mathbf{p}}^2 + m^2 c^4) \mathbf{1}$ gives

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \mathbf{1}, \quad \{\alpha_i, \beta\} = 0, \quad \beta^2 = \mathbf{1}. \quad (7.2.5)$$

The anticommutation relations cannot be satisfied by ordinary numbers or by 2×2 complex Hermitian matrices simultaneously for three independent spatial directions and one mass matrix. A four-component representation works:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix}. \quad (7.2.6)$$

Consequently the Dirac wavefunction is a four-component spinor,

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}, \quad \varphi, \chi \in \mathbb{C}^2. \quad (7.2.7)$$

The two Pauli spinors are coupled by the momentum operator. Positive- and negative-energy solutions occur; in the quantum-field interpretation, the latter are related to antiparticles.

3) Covariant form

Define $\gamma^0 = \beta$, $\gamma^i = \beta \alpha_i$ and the four-gradient $\partial_\mu = (c^{-1} \partial_t, \nabla)$. Multiplying the Hamiltonian equation by β gives

$$(i\hbar \gamma^\mu \partial_\mu - mc) \psi = 0, \quad \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \mathbf{1}_4. \quad (7.2.8)$$

Indeed, the anticommutators from Eq. (7.2.6) imply $(\gamma^0)^2 = \mathbf{1}_4$, $(\gamma^i)^2 = -\mathbf{1}_4$, and $\{\gamma^0, \gamma^i\} = 0$. The covariance of the equation follows when a Lorentz transformation of x^μ is accompanied by the corresponding spinor transformation. For the present discussion its essential feature is that the Hamiltonian is linear in the time and spatial derivatives.

Interactive investigation IC-07-02. Open Klein–Gordon versus Dirac. After deriving the Klein–Gordon and Dirac equations, compare their dispersion relations and matrix requirements. Check the Clifford anticommutators and why a first-order relativistic equation needs a multi-component spinor.

7.3 The Nonrelativistic Limit of the Dirac Equation

1) Free-particle energy and the rest-energy phase

For a free particle, the eigenvalues of the Dirac Hamiltonian satisfy

$$E_{\pm} = \pm \sqrt{m^2 c^4 + c^2 \mathbf{p}^2}. \quad (7.3.1)$$

Near the positive-energy rest branch, $|\mathbf{p}| \ll mc$,

$$\begin{aligned} E_+ &= mc^2 \sqrt{1 + \frac{\mathbf{p}^2}{m^2 c^2}} \\ &= mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{8m^3 c^2} + O\left(\frac{p^6}{m^5 c^4}\right). \end{aligned} \quad (7.3.2)$$

The constant rest energy creates a rapid overall phase. Remove it by writing

$$\psi(\mathbf{r}, t) = e^{-imc^2 t/\hbar} \begin{pmatrix} \varphi(\mathbf{r}, t) \\ \chi(\mathbf{r}, t) \end{pmatrix}. \quad (7.3.3)$$

In the Pauli–Dirac representation, $\hat{H}_D$ has block form

$$\hat{H}_D = \begin{pmatrix} mc^2 \mathbf{1}_2 & c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \\ c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} & -mc^2 \mathbf{1}_2 \end{pmatrix}. \quad (7.3.4)$$

Substitution into the Dirac equation gives two coupled Pauli-spinor equations:

$$\boxed{\begin{aligned} i\hbar \partial_t \varphi &= c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \chi, \\ i\hbar \partial_t \chi &= c\boldsymbol{\sigma} \cdot \hat{\mathbf{p}} \varphi - 2mc^2 \chi. \end{aligned}} \quad (7.3.5)$$

The large rest-energy separation makes the lower spinor small on the positive-energy branch. When kinetic energies and the time variation of χ are small compared with $2mc^2$, the second equation gives

$$\chi \simeq \frac{\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}}{2mc} \varphi. \quad (7.3.6)$$

Inserting this into the first equation produces

$$i\hbar \partial_t \varphi \simeq \frac{1}{2m} (\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2 \varphi = \frac{\hat{\mathbf{p}}^2}{2m} \varphi. \quad (7.3.7)$$

Thus the positive-energy Dirac equation reduces to the free two-component Pauli equation. The lower component has a controlled suppression of order $|\mathbf{p}|/(mc)$. The approximation selects positive-energy, low-momentum states; the complete relativistic theory still contains antiparticle solutions.

2) Recovering spin coupling for a charged particle

Under minimal electromagnetic coupling, replace $\hat{\mathbf{p}}$ by $\hat{\boldsymbol{\pi}} = \hat{\mathbf{p}} - (q/c)\mathbf{A}$ and add $q\phi$. After extracting the rest-energy phase, the equations become

$$\begin{aligned} i\hbar \partial_t \varphi &= q\phi \varphi + c\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}} \chi, \\ i\hbar \partial_t \chi &= c\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}} \varphi + (q\phi - 2mc^2) \chi. \end{aligned} \quad (7.3.8)$$

At leading nonrelativistic order,

$$\chi \simeq \frac{\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}}}{2mc} \varphi, \quad (7.3.9)$$

so the upper equation reads

$$i\hbar\partial_t\varphi = \left[\frac{1}{2m}(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\pi}})^2 + q\phi \right] \varphi + \text{higher-order relativistic corrections.} \quad (7.3.10)$$

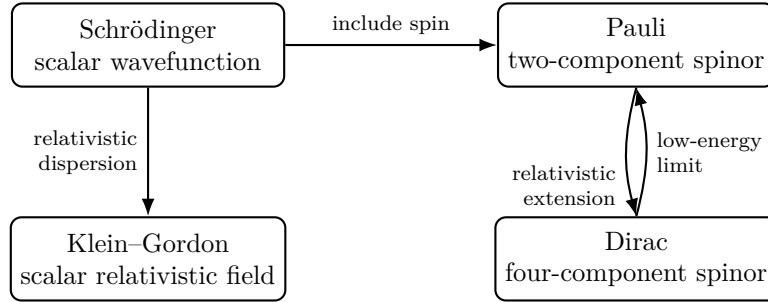
Using the commutator of the kinetic momentum components obtained in the first section yields

$$i\hbar\partial_t\varphi = \left[\frac{\hat{\boldsymbol{\pi}}^2}{2m} - \frac{q\hbar}{2mc}\boldsymbol{\sigma} \cdot \mathbf{B} + q\phi \right] \varphi + \cdots. \quad (7.3.11)$$

This is precisely the Pauli Hamiltonian in Eq. (7.1.8). In particular, the spin magnetic interaction follows from the matrix factorization of relativistic energy. Higher-order terms bring additional effects such as relativistic kinetic corrections and, for inhomogeneous fields, spin-orbit and Darwin terms.

7.4 The Relation among the Schrödinger, Pauli, and Dirac Equations

The preceding derivations connect three familiar descriptions. The scalar Schrödinger equation is the nonrelativistic theory of orbital motion. Enlarging its state space to a Pauli spinor incorporates spin 1/2 and magnetic coupling. The relativistic Dirac equation describes a four-component spinor; at small velocities and after factoring out the rest energy, its positive-energy sector returns the Pauli equation. The Klein–Gordon equation results directly from the relativistic energy relation, and for scalar particles it admits a quantum-field interpretation.



In particular, the two-component Pauli equation can be viewed both as the spinful extension of nonrelativistic Schrödinger theory and as a systematically controlled limit of the Dirac equation.

7.5 Free Dirac Spinors

For a plane wave $\psi = u(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{r}-Et)/\hbar}$ with $E > 0$, the upper and lower components obey

$$(E - mc^2)\varphi = c(\boldsymbol{\sigma} \cdot \mathbf{p})\chi, \quad (E + mc^2)\chi = c(\boldsymbol{\sigma} \cdot \mathbf{p})\varphi. \quad (7.5.1)$$

Consequently $\chi = c(\boldsymbol{\sigma} \cdot \mathbf{p})\varphi/(E + mc^2)$. For a normalized two-component spin basis ξ_s with $s = \uparrow, \downarrow$, choose

$$u_s(\mathbf{p}) = \sqrt{\frac{E + mc^2}{2E}} \begin{pmatrix} \xi_s \\ \frac{c\boldsymbol{\sigma} \cdot \mathbf{p}}{E + mc^2} \xi_s \end{pmatrix}, \quad u_s^\dagger u_{s'} = \delta_{ss'}. \quad (7.5.2)$$

This particular normalization is to unit probability density, $u_s^\dagger u_s = 1$; alternative Lorentz-covariant normalizations are common in quantum field theory. The lower two components are suppressed as $|\mathbf{p}|/(2mc)$ in the nonrelativistic limit. Negative-energy solutions are obtained from the $E < 0$ branch of the same 4×4 Hamiltonian and are interpreted consistently in a quantized Dirac field.

Worked calculation: normalize a Dirac spinor and take its low-energy limit

The four-component spinor just introduced can be understood by calculating its norm directly. Let $\xi_s^\dagger \xi_{s'} = \delta_{ss'}$ and define

$$\chi_s = \frac{c \boldsymbol{\sigma} \cdot \mathbf{p}}{E + mc^2} \xi_s, \quad E = \sqrt{m^2 c^4 + c^2 \mathbf{p}^2} > 0. \quad (7.5.3)$$

Because $(\boldsymbol{\sigma} \cdot \mathbf{p})^\dagger = \boldsymbol{\sigma} \cdot \mathbf{p}$ and $(\boldsymbol{\sigma} \cdot \mathbf{p})^2 = \mathbf{p}^2 \mathbf{1}$,

$$\begin{aligned} \chi_s^\dagger \chi_{s'} &= \frac{c^2}{(E + mc^2)^2} \xi_s^\dagger (\boldsymbol{\sigma} \cdot \mathbf{p})^2 \xi_{s'} \\ &= \frac{c^2 \mathbf{p}^2}{(E + mc^2)^2} \delta_{ss'}. \end{aligned} \quad (7.5.4)$$

The positive-energy dispersion relation implies $c^2 \mathbf{p}^2 = E^2 - m^2 c^4 = (E - mc^2)(E + mc^2)$. Thus

$$\begin{aligned} \delta_{ss'} + \chi_s^\dagger \chi_{s'} &= \left[1 + \frac{E - mc^2}{E + mc^2} \right] \delta_{ss'} \\ &= \frac{2E}{E + mc^2} \delta_{ss'}. \end{aligned} \quad (7.5.5)$$

Multiply the stacked spinor $(\xi_s, \chi_s)^\top$ by $\sqrt{(E + mc^2)/(2E)}$. The normalization factor cancels the previous fraction exactly: $u_s^\dagger(\mathbf{p}) u_{s'}(\mathbf{p}) = \delta_{ss'}$.

Explicit expansion for small momentum. Introduce $\epsilon = |\mathbf{p}|/(mc)$ and assume $\epsilon \ll 1$. Expand the positive energy:

$$\begin{aligned} E &= mc^2 \sqrt{1 + \epsilon^2} \\ &= mc^2 \left(1 + \frac{\epsilon^2}{2} - \frac{\epsilon^4}{8} + O(\epsilon^6) \right) \\ &= mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{8m^3 c^2} + O\left(\frac{|\mathbf{p}|^6}{m^5 c^4}\right). \end{aligned} \quad (7.5.6)$$

Consequently

$$\frac{c \boldsymbol{\sigma} \cdot \mathbf{p}}{E + mc^2} = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{2mc} + O(\epsilon^3), \quad (7.5.7)$$

where the big- O term is understood relative to the dimensionless spinor components. The lower two components are suppressed by one power of $|\mathbf{p}|/(mc)$. The difference $E - mc^2$ becomes the nonrelativistic kinetic energy at leading order. These two limits explain why the two-component Pauli spinor is a reliable low-energy description, together with the magnetic coupling derived earlier in this chapter.

Checkpoint. For momentum along $+z$, $\mathbf{p} = (0, 0, p)$ and $\xi_\uparrow = (1, 0)^\top$, write the complete four-component $u_\uparrow$ and check all four squared moduli sum to one. Then repeat for $\xi_\downarrow$ and verify orthogonality.

Chapter 8

Topological Condensed Matter Physics

8.1 From a Periodic Chain to Energy Bands

A crystal contains a very large number of repeating microscopic units. Its atomic spacing provides a natural length scale, and the periodicity allows us to describe the one-particle states using a wavevector k . Before introducing topology, we construct this description for the simplest possible chain.

1) A tight-binding chain

Consider one localized orbital $|n\rangle$ in each unit cell n of a one-dimensional lattice, with lattice spacing a . The states are orthonormal, $\langle m|n\rangle = \delta_{mn}$. An electron can move between neighboring sites by the *hopping amplitude* t :

$$\hat{H} = -t \sum_n (|n+1\rangle \langle n| + |n\rangle \langle n+1|). \quad (8.1.1)$$

With periodic boundary conditions, $|n+N\rangle = |n\rangle$. Define the normalized Bloch states

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{ikna} |n\rangle, \quad k = \frac{2\pi m}{Na}, \quad m = 0, \dots, N-1. \quad (8.1.2)$$

Translation by one cell multiplies $|k\rangle$ by e^{-ika} , and the Hamiltonian acts diagonally:

$$\hat{H} |k\rangle = -2t \cos(ka) |k\rangle. \quad (8.1.3)$$

The dispersion $E(k) = -2t \cos(ka)$ repeats under $k \mapsto k + 2\pi/a$. The interval $[-\pi/a, \pi/a]$ is one *Brillouin zone*. A band is the energy of an eigenstate of the periodic Hamiltonian as a function of crystal momentum.

2) More than one orbital per unit cell

Suppose a cell has two orbitals, A and B . The Bloch wavefunction has a two-component coefficient $|u(k)\rangle = (u_A(k), u_B(k))^T$, and its energies follow from a 2×2 Hermitian *Bloch Hamiltonian*:

$$\hat{h}(k) |u_n(k)\rangle = E_n(k) |u_n(k)\rangle, \quad n = 1, 2. \quad (8.1.4)$$

Any 2×2 Hermitian matrix can be written

$$\hat{h}(k) = d_0(k) \mathbf{1}_2 + \mathbf{d}(k) \cdot \boldsymbol{\sigma}, \quad E_{\pm}(k) = d_0(k) \pm |\mathbf{d}(k)|. \quad (8.1.5)$$

The bands are separated by a gap whenever $\mathbf{d}(k) \neq 0$. At fixed particle filling, a band insulator results if all occupied bands are separated from the empty bands by a finite gap. These are single-particle bands: interactions may require a more general many-body description.

3) The role of geometry

Besides its energy, a Bloch eigenstate has a direction in the internal space of orbital amplitudes. A smoothly evolving normalized eigenvector $|u_n(k)\rangle$ is defined only up to a k -dependent phase: $|u_n(k)\rangle \mapsto e^{i\chi(k)} |u_n(k)\rangle$. Exactly this freedom appeared in the Berry-phase discussion. As k traverses the Brillouin zone, the eigenvector can accumulate a geometric phase. Certain quantized properties of the eigenvectors remain unchanged when the Hamiltonian is smoothly deformed without closing the gap and without breaking the symmetry that protects the invariant. Such properties are the starting point of topological band theory [3, 4].

8.2 The Su–Schrieffer–Heeger Chain

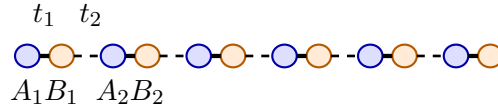
The Su–Schrieffer–Heeger (SSH) model is a chain with two orbitals per unit cell and alternating hopping amplitudes. It provides a solvable example of a topological phase and of boundary states [3].

1) A chain with two hopping amplitudes

Label the sites of cell n by A_n and B_n . Write the one-particle Hamiltonian in second-quantized form,

$$\hat{H} = \sum_n \left(t_1 \hat{a}_n^\dagger \hat{b}_n + t_2 \hat{a}_{n+1}^\dagger \hat{b}_n + \text{h.c.} \right), \quad (8.2.1)$$

where $t_1, t_2 > 0$ are the intra-cell and inter-cell hopping amplitudes. Here $\hat{a}_n$ and $\hat{b}_n$ annihilate a particle on the two sites; the same one-particle Hamiltonian can be studied without specifying Bose or Fermi occupation statistics. For convenience we set the lattice constant $a = 1$ in this chapter.



Choose the Fourier convention $\hat{a}_n = N^{-1/2} \sum_k e^{ikn} \hat{a}_k$ and similarly for b_n . In the basis $(\hat{a}_k, \hat{b}_k)$ the Bloch Hamiltonian is

$$\hat{h}(k) = \begin{pmatrix} 0 & q(k) \\ q(k)^* & 0 \end{pmatrix}, \quad q(k) = t_1 + t_2 e^{-ik}. \quad (8.2.2)$$

Equivalently,

$$\hat{h}(k) = (t_1 + t_2 \cos k) \sigma_x + t_2 \sin k \sigma_y. \quad (8.2.3)$$

Its energies are

$$E_{\pm}(k) = \pm \sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos k}. \quad (8.2.4)$$

The direct band gap is $2|t_1 - t_2|$, reached at $k = \pi$. Only at $t_1 = t_2$ does it close. A gap closing therefore separates two parameter regions.

2) Chiral symmetry

The operator $\Gamma = \sigma_z$ distinguishes the two sublattices. Because the SSH Hamiltonian couples A only to B ,

$$\boxed{\Gamma \hat{h}(k) \Gamma = -\hat{h}(k), \quad \text{or} \quad \{\Gamma, \hat{h}(k)\} = 0.} \quad (8.2.5)$$

This is called *chiral* or *sublattice* symmetry. It pairs an eigenstate of energy E with one of energy $-E$. In the Bloch-vector language it forbids a σ_z contribution and constrains $\mathbf{d}(k)$ to the xy plane. Adding a staggered on-site potential $\Delta\sigma_z$ would break this symmetry; the plane winding introduced below would then lose its quantized protection.

3) A winding number

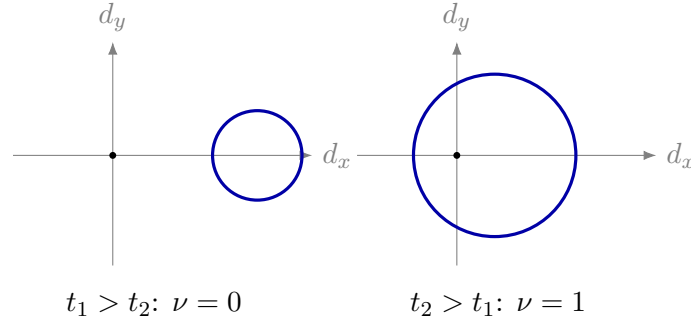
As k traverses one Brillouin zone, the vector

$$(d_x(k), d_y(k)) = (t_1 + t_2 \cos k, t_2 \sin k) \quad (8.2.6)$$

traces a circle of radius t_2 centered at $(t_1, 0)$. The origin is enclosed if $t_2 > t_1$. Define the phase $\varphi(k) = \arg[d_x(k) + id_y(k)]$ and

$$\boxed{\nu = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{d\varphi(k)}{dk} dk = \begin{cases} 0, & t_1 > t_2, \\ 1, & t_2 > t_1. \end{cases}} \quad (8.2.7)$$

In the gapped phases $\varphi(k)$ is continuous locally, although it may differ by an integer multiple of 2π when the Brillouin zone closes. Therefore ν is an integer and cannot change under a smooth, gap-preserving deformation respecting chiral symmetry. Its sign depends on the orientation and choice of unit cell; here the cell and Fourier convention were specified in Eq. (8.2.1).



This simple geometrical picture is the essential topological content of the SSH bulk Hamiltonian.

Interactive investigation IC-08-02. Open SSH bands and winding. After deriving the Bloch matrix and winding number, vary the two hopping amplitudes. Watch the band gap close and reopen, trace the complex winding curve, and revisit the same experiment after studying open-boundary edge states.

4) The Berry (Zak) phase

Write $d_x + id_y = |q|e^{i\varphi(k)}$. A normalized lower-band eigenvector of Eq. (8.2.2) is

$$|u_-(k)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\varphi(k)} \\ 1 \end{pmatrix}. \quad (8.2.8)$$

The Berry connection defined in the preceding dynamics chapter is

$$\mathcal{A}(k) = i \langle u_-(k) | \partial_k u_-(k) \rangle = \frac{1}{2} \partial_k \varphi(k). \quad (8.2.9)$$

Integrating over the Brillouin zone gives the *Zak phase* [4]

$$\gamma_{\text{Zak}} = \oint_{\text{BZ}} \mathcal{A}(k) dk = \pi\nu \pmod{2\pi}. \quad (8.2.10)$$

An eigenstate phase change shifts the connection by a derivative, while the physically meaningful closed-loop phase is unchanged modulo 2π under single-valued gauge choices. The Zak phase also depends on the orbital embedding and the unit-cell convention; its quantization to 0 or π in this SSH convention follows from chiral symmetry. The winding number contains the oriented integer information.

Worked derivation: from SSH hopping to a winding number

We now unpack the short step from the real-space hopping Hamiltonian to its two-band Bloch matrix. Choose a ring of N unit cells, momentum values $k = 2\pi m/N$, and the Fourier conventions of Eq. (8.2.1):

$$\hat{a}_n = \frac{1}{\sqrt{N}} \sum_k e^{ikn} \hat{a}_k, \quad \hat{b}_n = \frac{1}{\sqrt{N}} \sum_q e^{iqn} \hat{b}_q. \quad (8.2.11)$$

To transform the intracell term, substitute both expansions:

$$\begin{aligned} \sum_n \hat{a}_n^\dagger \hat{b}_n &= \frac{1}{N} \sum_{n,k,q} e^{-ikn} e^{iqn} \hat{a}_k^\dagger \hat{b}_q \\ &= \sum_{k,q} \left[\frac{1}{N} \sum_n e^{i(q-k)n} \right] \hat{a}_k^\dagger \hat{b}_q = \sum_k \hat{a}_k^\dagger \hat{b}_k, \end{aligned} \quad (8.2.12)$$

where the term in square brackets is δ_{kq} . For the intercell term a shift of the position index matters:

$$\begin{aligned} \sum_n \hat{a}_{n+1}^\dagger \hat{b}_n &= \frac{1}{N} \sum_{n,k,q} e^{-ik(n+1)} e^{iqn} \hat{a}_k^\dagger \hat{b}_q \\ &= \sum_k e^{-ik} \hat{a}_k^\dagger \hat{b}_k. \end{aligned} \quad (8.2.13)$$

Together with the Hermitian conjugate,

$$\hat{H} = \sum_k \begin{pmatrix} \hat{a}_k^\dagger & \hat{b}_k^\dagger \end{pmatrix} \underbrace{\begin{pmatrix} 0 & t_1 + t_2 e^{-ik} \\ t_1 + t_2 e^{ik} & 0 \end{pmatrix}}_{h(k)} \begin{pmatrix} \hat{a}_k \\ \hat{b}_k \end{pmatrix}. \quad (8.2.14)$$

There is now a precise reason for the exponent sign in the upper-right matrix entry.

Derive the two energies and locate the gap. Let $q = t_1 + t_2 e^{-ik}$. The characteristic polynomial is

$$\det[h(k) - EI] = \begin{vmatrix} -E & q \\ q^* & -E \end{vmatrix} = E^2 - |q|^2. \quad (8.2.15)$$

Compute the modulus explicitly:

$$\begin{aligned} |q|^2 &= (t_1 + t_2 e^{-ik})(t_1 + t_2 e^{ik}) \\ &= t_1^2 + t_2^2 + t_1 t_2 (e^{ik} + e^{-ik}) \\ &= t_1^2 + t_2^2 + 2t_1 t_2 \cos k. \end{aligned} \quad (8.2.16)$$

Thus $E_\pm = \pm|q|$. For $t_1, t_2 > 0$, the minimum $|q|$ occurs at $k = \pi$, because $\cos k \geq -1$. Hence $|q(\pi)| = |t_1 - t_2|$ and the direct spectral gap is $2|t_1 - t_2|$.

One explicit lower-band eigenvector. Write the *upper-right* off-diagonal element as $q = |q|e^{-i\varphi}$, so that $q^* = |q|e^{i\varphi}$. With $\varphi = \arg(t_1 + t_2e^{ik})$, choose

$$|u_-(k)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\varphi(k)} \\ 1 \end{pmatrix}. \quad (8.2.17)$$

Multiplication verifies its lower-band eigenvalue:

$$h(k)|u_-\rangle = \frac{|q|}{\sqrt{2}} \begin{pmatrix} e^{-i\varphi} \\ -1 \end{pmatrix} = -|q||u_-\rangle. \quad (8.2.18)$$

Now differentiate the state:

$$\partial_k |u_-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} ie^{-i\varphi} \partial_k \varphi \\ 0 \end{pmatrix}. \quad (8.2.19)$$

Its Berry connection is

$$\begin{aligned} \mathcal{A}(k) &= i \langle u_-(k) | \partial_k u_-(k) \rangle \\ &= \frac{i}{2} (-e^{i\varphi}) (ie^{-i\varphi} \partial_k \varphi) = \frac{1}{2} \partial_k \varphi. \end{aligned} \quad (8.2.20)$$

The map $(t_1 + t_2 \cos k, t_2 \sin k)$ is a circle of radius t_2 centered at $(t_1, 0)$ in the complex plane. The argument φ winds by 2π if the circle encloses the origin (when $t_2 > t_1$), and by zero if it does not (when $t_1 > t_2$). Therefore

$$\gamma_{\text{Zak}} = \oint \mathcal{A}(k) dk = \frac{1}{2} [\varphi(\pi) - \varphi(-\pi)] = \pi\nu \pmod{2\pi}. \quad (8.2.21)$$

Physical interpretation and limits. A nonzero winding requires a spectral gap and a fixed choice of sublattice/unit-cell convention. At $t_1 = t_2$, $q(\pi) = 0$, so the phase becomes undefined exactly where the gap closes. Adding a staggered onsite σ_z term changes the symmetry class, allowing the planar loop to move away from the origin in three dimensions. The interactive investigations introduced alongside this derivation allow the bulk winding, Berry phase, and finite-chain spectrum to be checked independently.

Interactive investigation IC-08-04. Open Zak phase and Berry gauge. After computing the lower-band eigenvector and Berry connection explicitly, change the unit-cell or phase convention. Compare discrete overlap phases with the analytic Zak phase and identify which quantities remain gauge invariant.

8.3 Boundary States and Bulk–Boundary Correspondence

We now open the chain, retaining a leftmost A_1 and a rightmost B_N . On a finite chain the real-space Hamiltonian is Eq. (8.2.1) without the last bond between the two ends.

1) The limit of separated dimers

When $t_2 = 0$, each A_n bonds to B_n : the chain consists of N isolated dimers, with energies $\pm t_1$ and no leftover sites. When $t_1 = 0$, the inter-cell bonds connect B_n to A_{n+1} . The end orbitals A_1 and B_N are unpaired and have exactly zero energy. They are localized at the boundaries.

2) Constructing an edge-state wavefunction

For finite t_1/t_2 , let a possible left-edge state have amplitudes only on the A sublattice, $|\psi_L\rangle = \sum_{n \geq 1} u_n |A_n\rangle$. At zero energy, the amplitude of each B_n component of $\hat{H}|\psi_L\rangle$ must vanish:

$$t_1 u_n + t_2 u_{n+1} = 0 \implies u_{n+1} = -\frac{t_1}{t_2} u_n. \quad (8.3.1)$$

Thus on a semi-infinite chain,

$$|\psi_L\rangle = \sqrt{1 - \left|\frac{t_1}{t_2}\right|^2} \sum_{n=1}^{\infty} \left(-\frac{t_1}{t_2}\right)^{n-1} |A_n\rangle \quad (8.3.2)$$

for $|t_1/t_2| < 1$. The sum of squared amplitudes equals unity because it is a geometric series. The localization length (in units of the cell spacing) is

$$\xi^{-1} = \ln \left| \frac{t_2}{t_1} \right|. \quad (8.3.3)$$

There is an analogous state on the right end, supported on the B sublattice. On a finite chain the two edge states hybridize weakly; their energies split by an amount exponentially small in the chain length. Exact zero energy follows directly only for an idealized isolated edge or a suitable sublattice imbalance.

3) What is protected?

The winding number $\nu = 1$ for the periodic chain and the localized end modes for the open chain describe the same phase. This is a simple case of *bulk–boundary correspondence*: a topological invariant of the gapped bulk constrains boundary phenomena. Boundary termination matters, and perturbations preserving chiral symmetry and the bulk gap preserve the appropriate topological distinction. If chiral symmetry is broken, boundary energies can shift and the SSH winding need not remain quantized.

The main result of the SSH model can be summarized in a single connected sequence:

$$\text{alternating hopping} \longrightarrow \hat{h}(k) \longrightarrow \nu \text{ and } \gamma_{\text{Zak}} \longrightarrow \text{localized edge states.} \quad (8.3.4)$$

It provides a starting point for the higher-dimensional invariants studied next [3].

Interactive investigation IC-08-03. Open SSH boundaries and symmetry. After constructing the edge-state wavefunction, vary chain termination, system size and chiral-symmetry-breaking terms. Compare edge localization and finite-size energy splitting with the predicted exponential decay.

8.4 Berry Curvature and the Chern Number

The SSH invariant is an integer obtained by winding around a circle. In two dimensions, the Brillouin zone is a torus. An isolated band can carry a different integer invariant, the *first Chern number*, whose physical manifestation includes the quantized Hall response [5, 6].

1) Berry connection in two dimensions

Let $\hat{h}(\mathbf{k})$ be a periodic Bloch Hamiltonian with $\mathbf{k} = (k_x, k_y)$, and let a normalized isolated band satisfy $\hat{h}(\mathbf{k}) |u_n(\mathbf{k})\rangle = E_n(\mathbf{k}) |u_n(\mathbf{k})\rangle$. For a nondegenerate band define

$$\mathcal{A}_{n,i}(\mathbf{k}) = i \langle u_n(\mathbf{k}) | \partial_{k_i} u_n(\mathbf{k}) \rangle. \quad (8.4.1)$$

This is a connection in momentum space; it is the same geometric object used in the Berry-phase derivation but evaluated as the Bloch momentum varies. The corresponding *Berry curvature* is

$$\Omega_n(\mathbf{k}) = \partial_{k_x} \mathcal{A}_{n,y} - \partial_{k_y} \mathcal{A}_{n,x} = i \left(\langle \partial_{k_x} u_n | \partial_{k_y} u_n \rangle - \langle \partial_{k_y} u_n | \partial_{k_x} u_n \rangle \right). \quad (8.4.2)$$

Under a gauge transformation $|u_n\rangle \mapsto e^{i\chi(\mathbf{k})} |u_n\rangle$, the connection changes by $-\nabla_{\mathbf{k}}\chi$, whereas Ω_n is gauge invariant. Stokes' theorem relates the Berry phase around a contour to the curvature passing through an enclosed surface.

2) The Chern number

Integrating the curvature over the full Brillouin zone gives

$$C_n = \frac{1}{2\pi} \int_{\text{BZ}} \Omega_n(\mathbf{k}) d^2k \in \mathbb{Z}. \quad (8.4.3)$$

The integral is an integer for a smooth, gapped, isolated band over the periodic Brillouin torus. Intuitively, the curvature is locally analogous to a magnetic field in momentum space, while the integral measures the global twisting of the band eigenstates. A nonzero C_n prevents one from choosing a single smooth, periodic eigenvector gauge everywhere on the torus. Changing an integer requires an occupied band to touch another band or a defining assumption to fail.

For a two-band Hamiltonian $\hat{h}(\mathbf{k}) = \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}$ with $\mathbf{d} \neq 0$, introduce the unit vector $\mathbf{n}(\mathbf{k}) = \mathbf{d}(\mathbf{k})/|\mathbf{d}(\mathbf{k})|$. The lower-band curvature can be computed directly from the eigenprojector $P_- = (\mathbf{1}_2 - \mathbf{n} \cdot \boldsymbol{\sigma})/2$:

$$\Omega_-(\mathbf{k}) = \frac{1}{2} \mathbf{n} \cdot (\partial_{k_x} \mathbf{n} \times \partial_{k_y} \mathbf{n}). \quad (8.4.4)$$

This gives a geometric picture: as $\mathbf{k}$ sweeps the Brillouin torus, $\mathbf{n}(\mathbf{k})$ traces the unit sphere, and the Chern number counts the oriented number of sphere coverings. For an upper band the sign is reversed.

3) A minimal Chern insulator

Consider the square-lattice two-band model

$$\hat{h}(\mathbf{k}) = \sin k_x \sigma_x + \sin k_y \sigma_y + (M + \cos k_x + \cos k_y) \sigma_z. \quad (8.4.5)$$

Energy bands are $E_{\pm} = \pm|\mathbf{d}|$. They touch only when $d_x = d_y = d_z = 0$, which requires $k_x, k_y \in \{0, \pi\}$ and $M = -2, 0, 2$. These critical values divide the parameter axis into four gapped regimes.

Near any crossing, a change $\mathbf{k} = \mathbf{K} + \mathbf{q}$ gives a two-dimensional Dirac form $\hat{h} \simeq v_x q_x \sigma_x + v_y q_y \sigma_y + m_{\mathbf{K}} \sigma_z$. The four points have masses and orientation factors

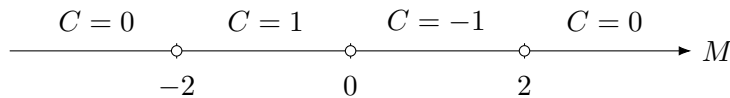
$\mathbf{K}$	$(0, 0)$	$(\pi, 0)$	$(0, \pi)$	(π, π)
$m_{\mathbf{K}}$	$M + 2$	M	M	$M - 2$
$\text{sgn}(v_x v_y)$	$+1$	-1	-1	$+1$

(8.4.6)

A regularized massive Dirac cone changes the lower-band Chern number when its mass crosses zero. Summing the four contributions using Eq. (8.4.4) gives

$$C_- = \frac{1}{2} [\text{sgn}(M + 2) - 2 \text{sgn}(M) + \text{sgn}(M - 2)] = \begin{cases} 0, & M < -2, \\ 1, & -2 < M < 0, \\ -1, & 0 < M < 2, \\ 0, & M > 2. \end{cases} \quad (8.4.7)$$

The integer is undefined at the gap-closing points themselves.



This model is a convenient algebraic route to a Chern insulator. The honeycomb-lattice Haldane model established that a quantum Hall phase can arise from an appropriate pattern of complex hopping amplitudes without a net external magnetic field [7].

4) Hall conductance and boundary transport

For completely filled, isolated noninteracting electron bands and under the usual clean-gap assumptions, the transverse conductivity is quantized in units of the conductance quantum:

$$\sigma_{xy} = \frac{e^2}{h} \sum_{n \in \text{occupied}} C_n, \quad (8.4.8)$$

with sign fixed by the chosen electric-charge and Hall-orientation conventions. This is the TKNN result [5]. At an interface where the Chern number changes, the spectral gap must be bridged by conducting boundary modes. For $|C| = 1$ there is one net chiral edge channel, whose direction is tied to the sign of the invariant. A disorder perturbation that respects the bulk mobility gap cannot remove the net chirality of an isolated edge.

Interactive investigation IC-08-06. Open Chern insulator phase diagram. After working through the two-band Chern number and Qi–Wu–Zhang model, scan the mass parameter. Locate bulk gap closings and compare the computed integer Chern number across the topological transitions.

8.5 Other Topological Phases and Present Perspective

1) Symmetry-protected phases

The SSH winding relies on chiral symmetry. Two-dimensional time-reversal-invariant electronic systems can instead support a $\mathbb{Z}_2$ classification, exemplified by the quantum spin Hall effect: opposite propagation directions can carry opposite Kramers-related spins. The resulting *helical* edges are protected against elastic backscattering by nonmagnetic, time-reversal-symmetric perturbations under the standard single-particle assumptions [6]. Three-dimensional topological insulators similarly exhibit characteristic surface states. The protecting symmetry and the nature of the gap must always be stated when asserting robustness.

2) How to recognize a topological problem

A useful sequence of questions is: What is the Hilbert-space basis? Does the model have translation symmetry, giving a Bloch Hamiltonian? Is the occupied spectrum gapped? Which symmetries are present? Can we calculate a Berry phase, winding number, or Chern number? What do the corresponding finite-boundary eigenstates look like? The SSH chain answers all these questions analytically, while higher-dimensional examples require increasingly sophisticated band geometry.

Beyond these introductory models, current research investigates interacting topological phases, topological superconductivity, higher-order boundary states, driven and dissipative systems, and non-Hermitian spectra. The derivations above provide the basic language for approaching these topics. A deeper introduction to band topology is given in *A Short Course on Topological Insulators* [3]; the broad review of topological insulators and superconductors by Qi and Zhang [6] explains the connection to real materials and symmetry classifications.

8.6 Geometric Phases in a Periodic Lattice

The Bloch-band topology can be computed numerically without differentiating an eigenvector whose overall phase is arbitrary. This is also useful for understanding which quantities depend on the choice of gauge.

1) A discrete Berry phase

Let $k_j = -\pi + 2\pi j/N$ for $j = 0, \dots, N-1$ and identify k_N with k_0 on the Brillouin circle. Consider a normalized, nondegenerate Bloch eigenstate at each point. Define the overlap product

$$W_N = \langle u(k_0)|u(k_1) \rangle \langle u(k_1)|u(k_2) \rangle \cdots \langle u(k_{N-1})|u(k_0) \rangle. \quad (8.6.1)$$

Under the independent phase choices $|u(k_j)\rangle \mapsto e^{i\chi_j} |u(k_j)\rangle$, each overlap acquires $e^{i(\chi_{j+1}-\chi_j)}$. Their product cancels every phase factor. Thus the quantity

$$\gamma_N = -\text{Im} \log W_N \pmod{2\pi} \quad (8.6.2)$$

is gauge invariant. For closely spaced momenta, the overlap is

$$\langle u(k)|u(k+\Delta k)\rangle = 1 + \langle u(k)|\partial_k u(k)\rangle \Delta k + O(\Delta k^2), \quad (8.6.3)$$

so, taking the principal value consistently as the mesh is refined,

$$\gamma_N \longrightarrow i \oint \langle u(k)|\partial_k u(k)\rangle dk. \quad (8.6.4)$$

This yields the Zak phase of the occupied band. The limit requires a nonvanishing gap and a sufficiently fine mesh; accidental orthogonality between adjacent numerical eigenvectors makes the discrete formula ill-conditioned.

2) A simple SSH check

In the SSH model set $t_1 = 0$, $t_2 > 0$. Then $d_x + id_y = t_2 e^{ik}$, so $\varphi(k) = k$ and Eq. (8.2.7) gives $\nu = 1$. The occupied eigenstate is

$$|u_-(k)\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-ik} \\ 1 \end{pmatrix}, \quad \mathcal{A}(k) = \frac{1}{2}. \quad (8.6.5)$$

Consequently the Zak phase is π . For $t_2 = 0$ and $t_1 > 0$, the Bloch Hamiltonian is independent of momentum, and the Zak phase is zero. This difference persists while the gap remains open and chiral symmetry is respected. By contrast, the energy spectra alone contain no information about the sign of the winding: both models can be flattened to bands at ± 1 without losing the topological distinction.

8.7 Charge Transport from a Slowly Varying Hamiltonian

Topology also appears in a dynamical process. Suppose a one-dimensional insulating Bloch Hamiltonian depends periodically on a control parameter $\lambda(t)$, with $\lambda(0) = \lambda(T)$ and an energy gap that stays open during the evolution. The pair (k, t) spans a two-dimensional torus, even though the physical lattice is one-dimensional.

For one completely occupied isolated band, the geometric contribution to the adiabatic transported charge over one cycle is proportional to the Berry-curvature flux in this (k, t) parameter space:

$$Q = eC_{kt}, \quad C_{kt} = \frac{1}{2\pi} \int_0^T dt \int_{\text{BZ}} dk \Omega_{kt}(k, t). \quad (8.7.1)$$

The sign of Q depends on the electron-charge and transport-direction convention. Its magnitude is an integer multiple of the elementary charge when the band is filled, the evolution is adiabatic and the gap stays open. This phenomenon is called a *Thouless pump*: a one-dimensional adiabatic cycle realizes the same two-dimensional Chern-number geometry as a quantum Hall

band. It connects the Berry phase derived after the AB effect to the transport observables of a lattice system.

A useful example is a *Rice–Mele* chain, obtained from the SSH model by adding staggered site energies:

$$\hat{h}(k, t) = [t_1(t) + t_2(t) \cos k] \sigma_x + t_2(t) \sin k \sigma_y + \Delta(t) \sigma_z. \quad (8.7.2)$$

A closed loop in parameter space $(t_1 - t_2, \Delta)$ can encircle the gap closing at the origin while keeping the instantaneous spectrum gapped. The integer pumped charge then distinguishes loops that enclose the degeneracy from loops that can be shrunk continuously to a point without crossing it. This interpretation is a direct physical use of Berry curvature.

Interactive investigation IC-08-07. Open Thouless pumping. After relating pumped charge to Berry curvature over a full cycle, vary the adiabatic pump path. Compare accumulated polarization and transported charge, and check how gap closure disrupts quantization.

8.8 Further Worked Checks

1) A symmetry-breaking perturbation

Add $\Delta \sigma_z$ to the SSH Hamiltonian. Its spectrum is

$$E_{\pm}(k) = \pm \sqrt{t_1^2 + t_2^2 + 2t_1 t_2 \cos k + \Delta^2}. \quad (8.8.1)$$

For $\Delta \neq 0$, the gap stays open even when $t_1 = t_2$. Hence the two SSH limits can now be connected continuously by varying $t_1 - t_2$ while keeping $\Delta \neq 0$. Their chiral winding distinction no longer applies: the perturbation has broken the protecting symmetry. This calculation illustrates why every statement of topological protection requires a statement of the relevant symmetry.

2) Scaling of an edge state

From Eq. (8.3.2), the probability on cell n of a normalized semi-infinite left-edge state is

$$P_n = \left(1 - \frac{t_1^2}{t_2^2}\right) \left(\frac{t_1^2}{t_2^2}\right)^{n-1}. \quad (8.8.2)$$

The mean distance of the state from the end is

$$\langle n \rangle = \sum_{n=1}^{\infty} n P_n = \frac{1}{1 - (t_1/t_2)^2}. \quad (8.8.3)$$

As $t_1 \rightarrow t_2^-$, the edge state spreads across an increasing number of cells and ultimately ceases to be normalizable on the semi-infinite system. The diverging localization length and the closing bulk gap are two complementary signatures of the same transition.

Chapter 9

Quantum Neural Networks and Machine Learning

9.1 What Is Machine Learning?

Machine learning studies how a model uses examples to learn a rule that can make predictions on new inputs. We first introduce the classical problem, because the learning objective is the same whether the model is evaluated by an ordinary computer or by a quantum circuit.

1) Inputs, labels, and a model

In supervised learning, a *dataset* consists of input–target pairs

$$\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \dots, (x_M, y_M)\}. \quad (9.1.1)$$

An input x_i can be a real number, an image represented by a vector, a measurement signal, or any other record. The target y_i is a desired answer, for example 0 or 1 in binary classification. A *model* $f_\theta(x)$ maps an input to a prediction. Its parameters $\theta = (\theta_1, \dots, \theta_p)$ are adjusted to fit the examples.

For regression with real targets, a convenient objective is the mean squared error

$$\mathcal{L}(\theta) = \frac{1}{M} \sum_{i=1}^M [f_\theta(x_i) - y_i]^2. \quad (9.1.2)$$

A step of gradient descent has the form

$$\theta_j \longleftarrow \theta_j - \eta \frac{\partial \mathcal{L}}{\partial \theta_j}, \quad (9.1.3)$$

where $\eta > 0$ is the *learning rate*. For class probabilities $p_\theta(y|x)$, a common alternative is the negative log-likelihood (cross-entropy) $-M^{-1} \sum_i \log p_\theta(y_i|x_i)$.

A model that describes the training examples perfectly may still perform poorly on new data. We therefore separate the examples into training and held-out test sets. *Generalization* refers to prediction quality on inputs not used to adjust the parameters. Model complexity, data volume, statistical noise, and the choice of training objective all affect generalization; the fact that a circuit is quantum does not remove these requirements.

2) Quantum data and classical data

Two types of learning problems must be distinguished. *Classical-data QML* starts from ordinary numbers x and uses a quantum device to prepare a state and evaluate a prediction. This

may require substantial input-encoding effort. *Quantum-data learning* starts from quantum states ρ_x produced by an experiment or a physical system, and aims to classify, estimate, or characterize their properties. Quantum-data problems can have a natural physical reason for using a quantum processor: the input itself is quantum. Most introductory quantum neural networks use the classical-data setting, which we study first.

9.2 Qubits, Quantum Gates, and Readout

1) One qubit

A qubit has an orthonormal basis $|0\rangle, |1\rangle$. A pure qubit state is

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (9.2.1)$$

A projective measurement of $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ gives $+1$ with probability $|\alpha|^2$ and -1 with probability $|\beta|^2$, hence

$$\langle\sigma_z\rangle_\psi = |\alpha|^2 - |\beta|^2. \quad (9.2.2)$$

A general density operator gives the same mean as $\text{tr}(\rho\sigma_z)$; the formalism was developed in the earlier measurement chapter. Finite experimental sampling returns an estimate of this expectation, with statistical uncertainty.

2) Gates as adjustable rotations

A gate is a unitary operator. Three common single-qubit rotations are

$$R_\mu(\theta) = e^{-i\theta\sigma_\mu/2}, \quad \mu = x, y, z. \quad (9.2.3)$$

Since $\sigma_\mu^2 = \mathbf{1}_2$, their exponential can be evaluated exactly:

$$R_\mu(\theta) = \cos\frac{\theta}{2} \mathbf{1}_2 - i \sin\frac{\theta}{2} \sigma_\mu. \quad (9.2.4)$$

For the y axis,

$$R_y(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}. \quad (9.2.5)$$

A *controlled-NOT* gate (CNOT) acts on two qubits. Its first qubit is the control; it flips the second, called the target, if and only if the control is $|1\rangle$:

$$|00\rangle \mapsto |00\rangle, \quad |01\rangle \mapsto |01\rangle, \quad |10\rangle \mapsto |11\rangle, \quad |11\rangle \mapsto |10\rangle. \quad (9.2.6)$$

It can create entanglement. For example,

$$\text{CNOT}(|+\rangle \otimes |0\rangle) = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}. \quad (9.2.7)$$

These gates suffice for the examples below.

9.3 A Quantum Neural Network as a Parameterized Circuit

The term *quantum neural network* (QNN) often means a *parameterized quantum circuit* used as a prediction model [8, 9]. The device repeatedly performs four operations: prepare an initial state, encode input data with gates, apply adjustable gates, and measure an observable. A classical procedure changes the adjustable parameters using the measurement results.

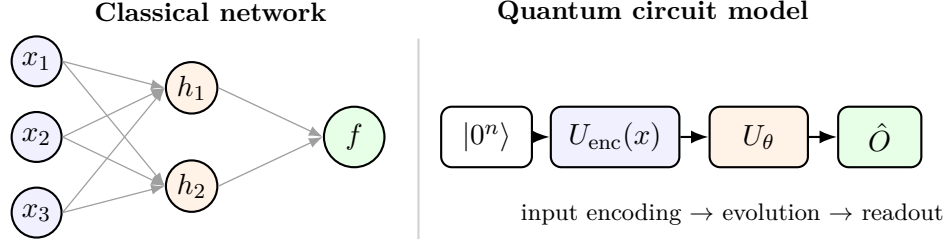


Figure 9.1: Two predictive models. Classical networks combine weighted activations; a quantum circuit encodes inputs, applies trainable unitaries and reads out an observable. Both require a loss function and an optimization procedure.

1) A precise mathematical definition

For n qubits, let $|0^n\rangle = |0\rangle^{\otimes n}$. An input-dependent unitary $U_{\text{enc}}(x)$ prepares the encoded state, and a parameterized circuit U_θ acts on it. The output is

$$f_\theta(x) = \langle 0^n | U_{\text{enc}}^\dagger(x) U_\theta^\dagger \hat{O} U_\theta U_{\text{enc}}(x) | 0^n \rangle, \quad (9.3.1)$$

where $\hat{O}$ is a Hermitian readout observable, such as σ_z on one qubit. Because $\hat{O}$ is Hermitian, $f_\theta(x)$ is real. If the spectrum of $\hat{O}$ lies in $[-1, 1]$, the prediction is also in that interval. A probability can then be constructed as $p_\theta(y = 1|x) = [1 - f_\theta(x)]/2$ when $y = 1$ corresponds to the -1 measurement outcome.

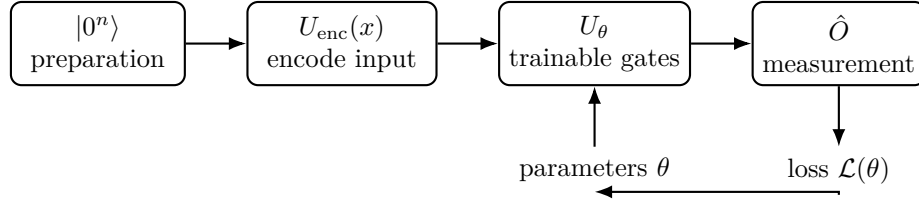


Figure 9.2: Hybrid quantum-classical optimization loop. Circuit measurements produce predictions and a loss, and the classical optimizer updates the gate parameters θ .

The quantum circuit evaluates the function; the classical optimization loop adjusts θ . An implementation must repeat measurements to estimate expected values and gradients.

2) An exactly soluble one-qubit model

Encode a real input x as

$$|\phi(x)\rangle = R_y(x) |0\rangle = \cos \frac{x}{2} |0\rangle + \sin \frac{x}{2} |1\rangle. \quad (9.3.2)$$

Apply a second, trainable rotation $R_y(\theta)$ and measure σ_z . Rotations around the same axis add, so

$$R_y(\theta)R_y(x) = R_y(x + \theta). \quad (9.3.3)$$

The predicted value can be computed without approximation:

$$f_\theta(x) = \langle 0 | R_y^\dagger(x + \theta) \sigma_z R_y(x + \theta) | 0 \rangle = \cos(x + \theta). \quad (9.3.4)$$

The learnable angle shifts the phase of the cosine. This is a nonlinear function of the classical input even though the quantum dynamics is linear in state vectors. The nonlinearity is introduced through the input encoding and the measurement probability, followed by the classical loss function.

For one training example (x, y) , the squared loss is $[\cos(x + \theta) - y]^2$, with gradient

$$\frac{\partial \mathcal{L}}{\partial \theta} = -2[\cos(x + \theta) - y] \sin(x + \theta). \quad (9.3.5)$$

This simple example makes clear that a quantum circuit does not automatically produce a better predictor than a classical cosine model. Its purpose is to show the precise form of the quantum learning procedure.

Interactive investigation IC-09-01. Open Parameterized quantum circuits. After computing the exact one-qubit expectation value, vary gate angles and input encodings. Compare the measured circuit output with its analytic Bloch-vector and Pauli-expectation predictions.

3) Two qubits and entangling gates

Start in $|00\rangle$, encode two real inputs by $R_y(x_1) \otimes R_y(x_2)$, then apply a CNOT with the first qubit as the control. Measure σ_z on the second qubit. The CNOT transforms the measured operator in the Heisenberg picture as

$$\text{CNOT}^\dagger(\mathbf{1} \otimes \sigma_z)\text{CNOT} = \sigma_z \otimes \sigma_z. \quad (9.3.6)$$

Therefore the measured expectation is exactly

$$\boxed{f(x_1, x_2) = \cos x_1 \cos x_2.} \quad (9.3.7)$$

The gate may entangle the intermediate state, yet this particular output has a simple classical expression. With additional noncommuting trainable rotations and repeated input embeddings, circuits can represent more complex coupled dependencies. The attainable functions depend strongly on the encoding, circuit architecture and measured observable.

9.4 How Quantum Circuits Are Trained

1) Expectation values from repeated measurements

Suppose a Pauli readout produces outcomes $z_s = \pm 1$ for $s = 1, \dots, N_{\text{shots}}$. The sample mean

$$\hat{f}_\theta(x) = \frac{1}{N_{\text{shots}}} \sum_{s=1}^{N_{\text{shots}}} z_s \quad (9.4.1)$$

is an unbiased estimate of $f_\theta(x) = \langle \hat{O} \rangle$. For statistically independent shots,

$$\text{Var}(\hat{f}_\theta) = \frac{1 - f_\theta(x)^2}{N_{\text{shots}}} \leq \frac{1}{N_{\text{shots}}}. \quad (9.4.2)$$

Thus the typical sampling error shrinks as $N_{\text{shots}}^{-1/2}$. Larger circuits and complicated observables may require many repeated measurements to estimate the loss and its gradients accurately.

2) The parameter-shift rule

Consider one gate $R_\mu(\theta) = e^{-i\theta\sigma_\mu/2}$ inside a circuit, with all other gates fixed. Its expectation value has the form

$$f(\theta) = A + B \cos \theta + C \sin \theta, \quad (9.4.3)$$

where the constants depend on the other gates, initial state, and observable. It follows that

$$\boxed{\frac{df(\theta)}{d\theta} = \frac{1}{2} \left[f\left(\theta + \frac{\pi}{2}\right) - f\left(\theta - \frac{\pi}{2}\right) \right].} \quad (9.4.4)$$

This is the *parameter-shift rule*. It expresses an exact derivative as a difference of two expectation values measurable on the device. It applies under the two-eigenvalue generator assumption just used; gates with more general generators may require other shift formulas. Combining the chain rule with Eq. (9.1.2) gives

$$\frac{\partial \mathcal{L}}{\partial \theta_j} = \frac{2}{M} \sum_{i=1}^M [f_\theta(x_i) - y_i] \frac{\partial f_\theta(x_i)}{\partial \theta_j}. \quad (9.4.5)$$

Each parameter derivative may therefore require two or more circuit settings. With many parameters and repeated shots, measurement cost becomes part of the optimization problem.

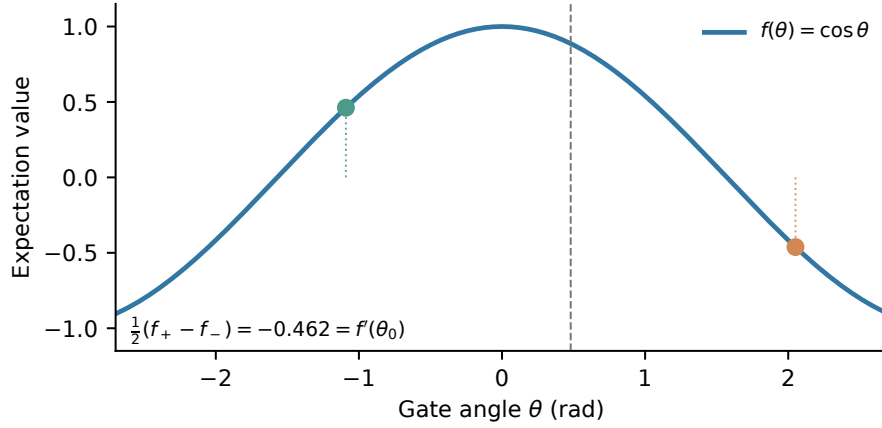


Figure 9.3: The parameter-shift rule for the exactly soluble $f(\theta) = \cos \theta$. The two dots mark the evaluations at $\theta_0 \pm \pi/2$ (with $\theta_0 = 0.48$); their difference, divided by two, gives $f'(\theta_0)$. The plot is generated analytically.

Interactive investigation IC-09-03. Open Parameter-shift training. After deriving the two-shift gradient formula, run a short circuit-training experiment. Compare shifted expectation values against the analytic derivative, then vary shot noise and step size to inspect optimization behavior.

3) More expressive data encodings

A single use of $R_y(x)$ produced the simple trigonometric function in Eq. (9.3.4). A generic data-encoding layer may be written $e^{-ix\hat{G}}$ for some Hermitian generator $\hat{G}$. Inserting its spectral decomposition reveals factors $e^{-ix(g_a - g_b)}$. Consequently a circuit expectation can be written as a finite Fourier-type series in the input:

$$f_\theta(x) = \sum_{\omega \in \Omega} c_\omega(\theta) e^{i\omega x}, \quad c_{-\omega} = c_\omega^*. \quad (9.4.6)$$

The same input can be encoded more than once, as in the following elementary one-qubit sequence. The available frequency set Ω is constrained by the eigenvalue differences of the

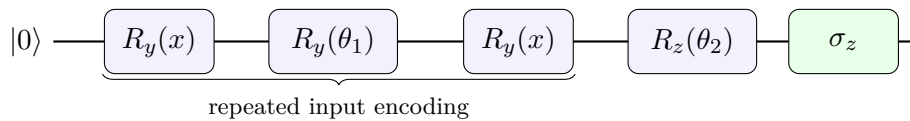


Figure 9.4: Data re-uploading: the same input x appears in multiple encoding gates separated by adjustable rotations. The measured output depends jointly on x and θ_1, θ_2 .

encoding generators. Repeating data encodings between trainable gates (*data re-uploading*) can enrich the possible frequency spectrum. This precise relation between data encoding and expressivity is developed in the work of Schuld, Sweke and Meyer [10]. Richer possible functions still need a useful loss, sufficient data, and a trainable parameterization.

9.5 Quantum Kernel Methods

Parameterized circuits are one way to use quantum states in learning. A different approach is to encode each input as a quantum feature state

$$|\phi(x)\rangle = U_{\text{enc}}(x) |0^n\rangle \quad (9.5.1)$$

and compare two inputs using their overlap. Define

$$K(x, x') = |\langle \phi(x) | \phi(x') \rangle|^2. \quad (9.5.2)$$

This is a *quantum kernel*: it measures similarity of the corresponding density operators $\rho_x = |\phi(x)\rangle \langle \phi(x)|$ because $K(x, x') = \text{tr}(\rho_x \rho_{x'})$.

1) Why it is a valid kernel

For any complex coefficients c_i ,

$$\sum_{ij} c_i^* c_j K(x_i, x_j) = \text{tr} \left[\left(\sum_i c_i \rho_{x_i} \right)^\dagger \left(\sum_j c_j \rho_{x_j} \right) \right] \geq 0. \quad (9.5.3)$$

Hence the matrix $K_{ij} = K(x_i, x_j)$ is positive semidefinite. This mathematical property allows its use in kernel regression or support-vector classification. A quantum device may estimate kernel entries by preparing $U_{\text{enc}}(x')^\dagger U_{\text{enc}}(x) |0^n\rangle$ and measuring the probability of obtaining 0^n .

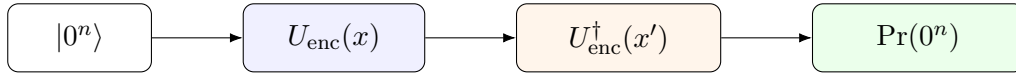


Figure 9.5: An overlap-based kernel evaluation: prepare $U_{\text{enc}}^\dagger(x') U_{\text{enc}}(x) |0^n\rangle$ and measure the all-zero probability, equal to $K(x, x')$.

2) A one-qubit example

For $|\phi(x)\rangle = R_y(x) |0\rangle$,

$$\langle \phi(x) | \phi(x') \rangle = \cos \frac{x - x'}{2}, \quad K(x, x') = \cos^2 \frac{x - x'}{2}. \quad (9.5.4)$$

The result is simple enough to evaluate classically. The research question for larger quantum feature maps is whether useful kernel values can be estimated efficiently on suitable quantum hardware while the corresponding classical computation remains costly for the relevant data distribution. The existence of an exponentially large Hilbert space alone does not establish such a benefit [11].

Interactive investigation IC-09-05. Open Quantum kernel geometry. After proving that the overlap matrix is positive semidefinite, generate a small quantum kernel Gram matrix. Compare overlap probabilities, matrix eigenvalues and the one-qubit analytic result.

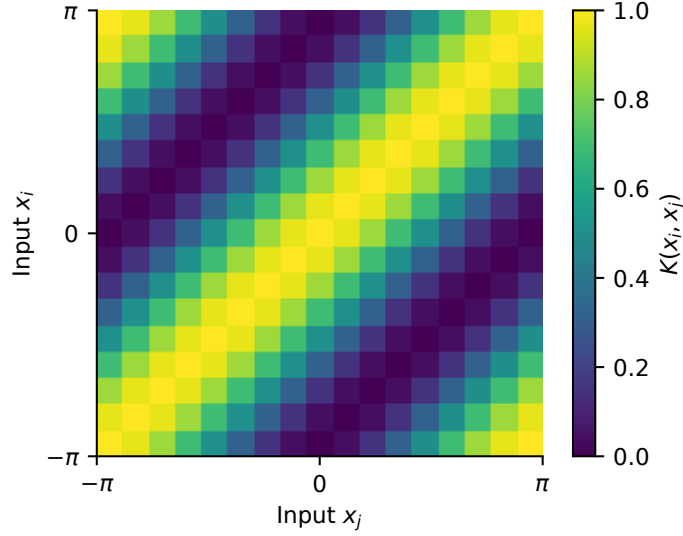


Figure 9.6: Exact one-qubit quantum kernel $K(x_i, x_j) = \cos^2[(x_i - x_j)/2]$ on 17 equally spaced inputs in $[-\pi, \pi]$. The diagonal entries equal one; off-diagonal similarities depend on the relative encoded angle.

Worked derivation: parameter shift and the measurement cost

The parameter-shift formula above is simple to quote. To see why it is exact, start from a single rotation $\hat{R}(\theta) = \exp(-i\theta\sigma_\mu/2)$. Since $\sigma_\mu^2 = \mathbf{1}$, even and odd powers in the exponential sum separately to

$$\hat{R}(\theta) = \mathbf{1} \cos \frac{\theta}{2} - i\sigma_\mu \sin \frac{\theta}{2}. \quad (9.5.5)$$

For any fixed input density matrix ρ and Hermitian measurement operator $\hat{O}$, write $f(\theta) = \text{tr}[\hat{O}\hat{R}(\theta)\rho\hat{R}^\dagger(\theta)]$. Substitute the two-term rotation decomposition and expand the four products. Terms containing two factors $\cos(\theta/2)$ or two factors $\sin(\theta/2)$ give $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$; the two cross terms give $\sin(\theta/2)\cos(\theta/2)$. Using

$$\cos^2 \frac{\theta}{2} = \frac{1 + \cos \theta}{2}, \quad \sin^2 \frac{\theta}{2} = \frac{1 - \cos \theta}{2}, \quad 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} = \sin \theta, \quad (9.5.6)$$

we obtain

$$f(\theta) = A + B \cos \theta + C \sin \theta, \quad (9.5.7)$$

for real coefficients A, B, C independent of θ . This holds with arbitrary fixed unitaries before and after the displayed rotation; they simply modify the effective ρ and $\hat{O}$. Differentiate analytically:

$$\frac{df}{d\theta} = -B \sin \theta + C \cos \theta. \quad (9.5.8)$$

Now evaluate the function at shifted arguments:

$$f(\theta + \pi/2) = A - B \sin \theta + C \cos \theta, \quad (9.5.9)$$

$$f(\theta - \pi/2) = A + B \sin \theta - C \cos \theta. \quad (9.5.10)$$

Subtracting and dividing by two gives

$$f'(\theta) = \frac{1}{2} [f(\theta + \pi/2) - f(\theta - \pi/2)], \quad (9.5.11)$$

which is exactly Eq. (9.4.4).

The gradient is exact; measurements introduce noise. For each of the two shifted circuits, suppose we measure a Pauli observable with independent outputs $z = \pm 1$. If the true expectation is $f_{\pm}$, then $E[z] = f_{\pm}$, $E[z^2] = 1$ and $\text{Var}(z) = 1 - f_{\pm}^2$. With M independent shots per shift, the sample mean $\hat{f}_{\pm}$ has variance

$$\text{Var}(\hat{f}_{\pm}) = \frac{1 - f_{\pm}^2}{M}. \quad (9.5.12)$$

The two batches are independent, so for $\hat{g} = (\hat{f}_+ - \hat{f}_-)/2$,

$$E[\hat{g}] = \frac{1}{2}(f_+ - f_-) = f'(\theta), \quad (9.5.13)$$

$$\begin{aligned} \text{Var}(\hat{g}) &= \frac{1}{4}[\text{Var}(\hat{f}_+) + \text{Var}(\hat{f}_-)] \\ &= \boxed{\frac{2 - f_+^2 - f_-^2}{4M} \leq \frac{1}{2M}}. \end{aligned} \quad (9.5.14)$$

The estimator is unbiased, and its uncertainty depends on measurement outcomes as well as the shot budget.

A useful counterexample to a universal claim. Consider n independently rotated qubits $|\psi(\theta)\rangle = \bigotimes_{j=1}^n R_y(\theta_j) |0\rangle$. Measuring the global observable $Z^{\otimes n}$ gives

$$f_{\text{global}}(\theta) = \prod_{j=1}^n \cos \theta_j, \quad \partial_{\theta_1} f_{\text{global}} = -\sin \theta_1 \prod_{j=2}^n \cos \theta_j. \quad (9.5.15)$$

When each θ_j is independent and uniform on $[-\pi, \pi]$, the averages $E[\sin^2 \theta_j] = E[\cos^2 \theta_j] = 1/2$ give

$$E[(\partial_{\theta_1} f_{\text{global}})^2] = \left(\frac{1}{2}\right)^n. \quad (9.5.16)$$

For the local observable Z_1 , however, $f_{\text{local}} = \cos \theta_1$ and

$$E[(\partial_{\theta_1} f_{\text{local}})^2] = \frac{1}{2} \quad (9.5.17)$$

for *every* n . The same product circuit can therefore have exponentially small global gradients and a dimension-independent local gradient. The precise cost observable and parameter ensemble matter. The [gradient concentration lab](#) tests these exact moments against reproducible random ensembles.

Checkpoint. At $n = 10$, calculate the expected squared global gradient and compare it with the local value. For $M = 512$, compute the conservative standard-deviation bound $\sqrt{1/(2M)}$ for a two-shift estimator. Which gradients are easier to distinguish from shot noise?

9.6 Optimization, Trainability, and Practical Limits

A quantum prediction model must satisfy three different requirements: it should represent a useful relationship between inputs and outputs, its parameters should be learnable from available data, and it should run at a reasonable cost relative to alternative methods. The first is a question of *expressivity*; the second concerns *trainability*; the third concerns *resources and quantum advantage*.

1) Barren plateaus

A common difficulty for parameterized quantum circuits is the *barren plateau* phenomenon. In some families of random-like circuits, the typical gradient becomes extremely small as the number n of qubits increases. Schematically,

$$\text{Var}_\theta \left(\frac{\partial \mathcal{L}}{\partial \theta_j} \right) \sim O(b^{-n}) \quad \text{for some } b > 1, \quad (9.6.1)$$

under the assumptions of the particular theorem. Then exponentially many measurements may be needed to resolve a typical tiny gradient from sampling noise. This behavior was identified in the foundational analysis of McClean *et al.* [12] and is reviewed, including the influence of architecture, noise, initialization and measurement, in Larocca *et al.* [13].

A barren plateau is a property of particular circuit families and training objectives, not an inevitable characteristic of every quantum circuit. Shallow local circuits with carefully chosen initializations or local objectives can have different gradient behavior. At the same time, a model that is easy to train may be representationally restricted or classically simulable. Both questions must be considered together.

2) Noise and finite shots

Physical devices have imperfect gates, readout errors, and limited coherence times. With noise, the ideal pure circuit state is replaced by a density operator $\rho_\theta(x)$, and the prediction becomes

$$f_\theta(x) = \text{tr}[\rho_\theta(x)\hat{O}]. \quad (9.6.2)$$

Noise modifies the loss landscape and can suppress the magnitude of observable-dependent information. A theoretical parameter-shift identity for an ideal gate must be applied with attention to which parts of the noisy implementation depend on the shifted parameter. Finite-shot estimates add statistical uncertainty to every measured output.

3) Data input and fair comparisons

Classical input data must first be encoded into the quantum system. If preparing the encoded state is expensive, the preparation cost must be included in any end-to-end complexity estimate. A model must also be evaluated on genuinely held-out data. Useful baselines include simple trigonometric regression for the one-qubit example, classical kernel methods for the quantum-kernel example, and competitive conventional machine-learning models on the same inputs and splits.

The term *quantum advantage* has a specific meaning: a demonstrable improvement in a defined task and resource measure relative to the best relevant classical alternatives. It is stronger than showing that a quantum circuit can solve a toy task or that its state space has dimension 2^n . Establishing such advantages for useful classical-data machine-learning tasks remains an active research problem. The review of variational quantum algorithms [8], the expressivity analysis [10], and the 2025 overview of trainability obstacles [13] provide a basis for assessing new claims.

Interactive investigation IC-09-06. Open Gradient concentration and trainability. After discussing barren plateaus, compare global and local costs as circuit size increases. Estimate gradient variances over parameter samples and connect the results with the measurement-shot requirements.

9.7 Beyond the Introductory Models

Quantum learning also includes *quantum convolutional* and recurrent-like architectures, generative modeling, quantum-data classification, learning properties of quantum phases, and hybrid quantum–classical methods. Their physical motivations differ. For example, classifying states of a many-body system can start directly from a measured or prepared quantum state, while training a classical-label classifier usually requires a chosen encoding of ordinary inputs. The preceding derivations establish a minimal mathematical language for these extensions:

$$\begin{aligned} \text{input} &\longrightarrow \text{quantum state} \longrightarrow \text{parameterized evolution} \\ &\longrightarrow \text{measurement} \longrightarrow \text{loss and optimization.} \end{aligned} \quad (9.7.1)$$

A physical model of learning is useful when each arrow has an explicit definition and a measurable cost. The broader connection between circuits and learning is analyzed in the review by Cerezo *et al.* [8]; Abbas *et al.* [9] discuss the representational capacities of certain quantum neural networks, while recent work continues to clarify limitations due to statistical concentration and efficient classical representations [13].

9.8 Worked Training and Noise Examples

1) Fitting the simplest model

Suppose a dataset is generated by $y_i = \cos(x_i + \theta_*)$ for some unknown parameter θ_* . Using the one-qubit prediction in Eq. (9.3.4), the squared-error objective is

$$\mathcal{L}(\theta) = \frac{1}{M} \sum_i [\cos(x_i + \theta) - \cos(x_i + \theta_*)]^2. \quad (9.8.1)$$

The true parameter θ_* , modulo 2π , gives zero training loss. To understand local training, write $\theta = \theta_* + \delta\theta$ and expand:

$$\cos(x_i + \theta) - \cos(x_i + \theta_*) = -\sin(x_i + \theta_*)\delta\theta + O(\delta\theta^2). \quad (9.8.2)$$

Therefore

$$\mathcal{L}(\theta_* + \delta\theta) = \left[\frac{1}{M} \sum_i \sin^2(x_i + \theta_*) \right] (\delta\theta)^2 + O(\delta\theta^3). \quad (9.8.3)$$

The coefficient in brackets tells us how informative the training inputs are about this parameter. If all inputs happen to sit close to a cosine maximum or minimum, the derivative is tiny and estimating the parameter becomes difficult, even for one qubit. The relation between data distribution and trainability is already visible in this elementary example.

2) Depolarizing noise

Consider a noise channel on a qubit, after the ideal circuit, that replaces its density operator with a mixture of the original state and the maximally mixed state:

$$\mathcal{E}_p(\rho) = (1 - p)\rho + p\frac{\mathbf{1}_2}{2}, \quad 0 \leq p \leq 1. \quad (9.8.4)$$

Since $\text{tr}(\sigma_z) = 0$, the noisy readout of the previous one-qubit model is

$$f_{\theta,p}(x) = \text{tr}[\mathcal{E}_p(\rho_\theta(x))\sigma_z] = (1 - p) \cos(x + \theta). \quad (9.8.5)$$

Its parameter derivative is reduced by the same factor $1 - p$. Strong depolarization reduces the information about the trainable parameter available through this measurement. The effect is specific to the stated channel and observable; other error processes change predictions in different ways.

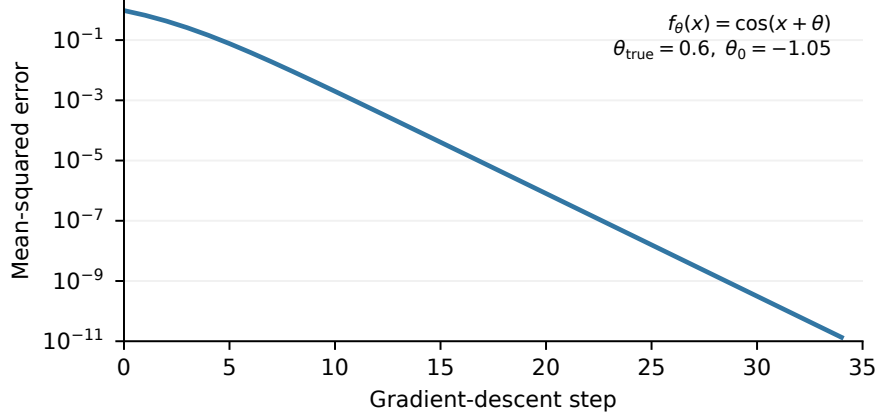


Figure 9.7: Full-batch gradient descent for $f_{\theta}(x) = \cos(x + \theta)$ using 31 equally spaced inputs in $[-1.35, 1.35]$, exact targets $\cos(x + 0.6)$, initial $\theta_0 = -1.05$, and learning rate 0.34. The graph plots the exact mean-squared error over 35 steps, without sampling noise.

3) How many measurements are needed?

Suppose we want standard error at most ϵ in a Pauli expectation estimate. The variance inequality in the preceding section implies the sufficient condition

$$N_{\text{shots}} \geq \epsilon^{-2}. \quad (9.8.6)$$

For example, $\epsilon = 0.01$ requires up to 10^4 shots for this worst-case variance estimate. To estimate a parameter-shift derivative with two independent batches of N_{shots} shots, the variance is bounded by $1/(2N_{\text{shots}})$, because each term carries a factor of $1/2$. Training many parameters over many data points may therefore use substantially more circuit executions than one prediction.

9.9 Learning from Quantum States

The use of quantum inputs can be closer to experimental quantum physics than classical-data classification. Suppose an experiment prepares an unknown two-qubit state ρ . We would like to determine whether it contains entanglement without reconstructing every matrix entry. One possible test uses an *entanglement witness* associated with the Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$:

$$\hat{W} = \frac{1}{2}\mathbf{1}_4 - |\Phi^+\rangle\langle\Phi^+|. \quad (9.9.1)$$

For every pure product state $|a\rangle \otimes |b\rangle$, the squared overlap with $|\Phi^+\rangle$ is at most $1/2$. Indeed,

$$|\langle\Phi^+|a, b\rangle|^2 = \frac{1}{2}|a_0b_0 + a_1b_1|^2 \leq \frac{1}{2}, \quad (9.9.2)$$

by the Cauchy–Schwarz inequality and the normalization of each qubit. A separable mixed state is a convex combination of product states, so it also satisfies $\text{tr}(\rho\hat{W}) \geq 0$. Therefore

$$\boxed{\text{tr}(\rho\hat{W}) < 0 \implies \rho \text{ is entangled.}} \quad (9.9.3)$$

This is a useful classification statement about genuinely quantum data: the expectation of a Hermitian operator witnesses a physically meaningful property. Some entangled states will not be detected by this particular witness, since the criterion is sufficient but not necessary.

The operator $\hat{W}$ can be decomposed into Pauli products,

$$|\Phi^+\rangle\langle\Phi^+| = \frac{1}{4}(\mathbf{1} \otimes \mathbf{1} + \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z), \quad (9.9.4)$$

so its expectation can be estimated from local correlation measurements. This directly connects quantum learning tasks with the density-matrix, Bell-state and measurement material introduced earlier in the textbook.

9.10 Statistical Concentration and a Trainability Check

Why do randomly chosen circuits sometimes provide little training information? Let $|\psi\rangle$ be a Haar-random pure state in a d -dimensional Hilbert space, and let $\hat{O}$ be a traceless Hermitian observable. Then the second moment of the random state gives

$$\mathbb{E}[\langle \hat{O} \rangle_\psi] = 0, \quad \text{Var}[\langle \hat{O} \rangle_\psi] = \frac{\text{tr}(\hat{O}^2)}{d(d+1)}. \quad (9.10.1)$$

For an n -qubit Pauli string with eigenvalues ± 1 , $d = 2^n$ and $\text{tr}(\hat{O}^2) = d$, so

$$\text{Var}[\langle \hat{O} \rangle_\psi] = \frac{1}{2^n + 1}. \quad (9.10.2)$$

Thus expectation values of a fixed traceless observable concentrate near zero under uniformly random pure states in high dimension. We can also examine a particular derivative rather than the expectation of a Pauli observable. Draw $|\psi\rangle$ from the Haar distribution in dimension $d = 2^n$, apply $e^{-i\theta X_1/2}$, and measure the probability $f(\theta) = |\langle 0^n | e^{-i\theta X_1/2} |\psi\rangle|^2$. Here X_1 acts on the first qubit. If $a = \langle 0^n | \psi \rangle$ and $b = \langle 10^{n-1} | \psi \rangle$, then $f'(0) = \text{Im}(a^*b)$. The Haar fourth moment yields

$$\text{Var}[f'(0)] = \frac{1}{2d(d+1)}, \quad d = 2^n. \quad (9.10.3)$$

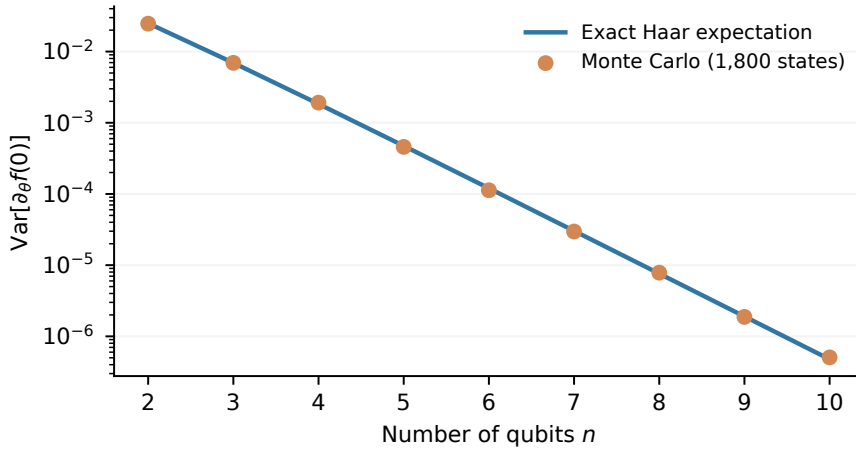


Figure 9.8: Gradient concentration for this specific global-projector objective: exact Haar variance $[2d(d+1)]^{-1}$ versus 1,800 independent numerically sampled Haar states at each n . This illustrative ensemble of random states is a precisely defined toy model, not a universal statement about practical QNN architectures.

This calculation by itself is *not* a barren-plateau theorem for derivatives, but it illustrates the concentration mechanism underlying many rigorous gradient-suppression results [12, 13]. The correct gradient scaling depends on the circuit generators, initial state, measurement locality, and objective function.

Bibliography

- [1] Yuto Ashida, Zongping Gong, and Masahito Ueda. Non-hermitian physics. *Advances in Physics*, 69(3):249–435, 2020. <https://doi.org/10.1080/00018732.2021.1876991>.
- [2] Ramy El-Ganainy, Konstantinos G. Makris, Mercedeh Khajavikhan, Ziad H. Musslimani, Stefan Rotter, and Demetrios N. Christodoulides. Non-hermitian physics and PT symmetry. *Nature Physics*, 14(1):11–19, 2018. <https://doi.org/10.1038/nphys4323>.
- [3] János K. Asbóth, László Oroszlány, and András Pályi. *A Short Course on Topological Insulators: Band Structure and Edge States in One and Two Dimensions*, volume 919 of *Lecture Notes in Physics*. Springer International Publishing, 2016. <https://doi.org/10.1007/978-3-319-25607-8>.
- [4] J. Zak. Berry’s phase for energy bands in solids. *Physical Review Letters*, 62:2747–2750, 1989. <https://doi.org/10.1103/PhysRevLett.62.2747>.
- [5] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs. Quantized hall conductance in a two-dimensional periodic potential. *Physical Review Letters*, 49:405–408, 1982. <https://doi.org/10.1103/PhysRevLett.49.405>.
- [6] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Reviews of Modern Physics*, 83:1057–1110, 2011. <https://doi.org/10.1103/RevModPhys.83.1057>.
- [7] F. D. M. Haldane. Model for a quantum hall effect without landau levels: Condensed-matter realization of the “parity anomaly”. *Physical Review Letters*, 61:2015–2018, 1988. <https://doi.org/10.1103/PhysRevLett.61.2015>.
- [8] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3:625–644, 2021. <https://doi.org/10.1038/s42254-021-00348-9>.
- [9] Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1:403–409, 2021. <https://doi.org/10.1038/s43588-021-00084-1>.
- [10] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103:032430, 2021. <https://doi.org/10.1103/PhysRevA.103.032430>.
- [11] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, 2019. <https://doi.org/10.1038/s41586-019-0980-2>.

- [12] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9:4812, 2018. <https://doi.org/10.1038/s41467-018-07090-4>.
- [13] Martín Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J. Coles, Lukasz Cincio, Jarrod R. McClean, Zoë Holmes, and M. Cerezo. Barren plateaus in variational quantum computing. *Nature Reviews Physics*, 7:174–189, 2025. <https://doi.org/10.1038/s42254-025-00813-9>.